Designing advanced GIS visualisations using cognitive ergonomics theories, models and procedures
A. Turk (Melbourne, AUS)

1. Introduction

Much useful and interesting information about the world contains aspects which draw their significance from the spatial distribution of phenomena. Since the earliest days of civilisation such information has been presented in the form of maps and diagrams. The manner of representation has changed exceedingly slowly until recent times. The ability to encode maps digitally has evolved into geographic information systems (GIS) with powerful analysis functionality and display capabilities. Traditional cartographic educational, decision-support and entertainment objectives can now be achieved via a multitude of interactive multimedia products. While this (technology led) revolution has greatly enhanced the potential to fulfil these objectives, it has not eliminated the basic question: what is a good map? Indeed the problem of design has become considerably more complex.

The map designer could be said to be faced with a choice between:

- the easy solution - giving users what you have;
- the popular solution - giving users what they want;
- the complex solution - giving users what they need.

Traditionally, cartographers have often opted for the easy solution, especially in the case of topographic maps. Map design concentrated on the objectives of some well defined specific initial user, usually the one who paid for the survey and map production costs. Subsequent users had to take what was offered. While the cost of manually produced special purpose maps was prohibitive and cartographic products were scarce, users were usually satisfied with this arrangement, if not uniformly content.

As users became less desperate for any map at all, they started voicing increasingly loud demands for products which matched their individual desires. Some cartographers welcomed this unrest as an expression of legitimate user involvement in the cartographic communication process. However, others highlighted the fact that user demands were sometimes made in ignorance - patients don't tell their doctor how to treat them so why shouldn't map users trust cartographers? There was also the real problem of conflicting user demands. Prior to the development of effective digital mapping and GIS technology, one cartographic product frequently had to (inadequately) cover the expectations of a wide range of potential users. This often lead to the informal application of the concept of the 'optimal map' defined, for instance, as that producing the 'least sum of (weighted) user dissatisfaction'. As well as the extreme technical difficulty of achieving such an optimisation in any formal sense, there was the problem of obtaining the required user feedback. Questionnaire surveys, map users forums, user advisory committees, and like mechanisms were often ineffective and short-lived.

The developments in cartographic technology of the late 1970's and the 1980's provided greatly increased potential to satisfy the needs of users. Computer-assisted cartography could provide products designed to meet specific user-requirement specifications. Various approaches to the optimisation of cartographic design could be adopted, including the following:
Expert Opinion:
People who are recognised as expert cartographers could be used to design the map;

Conventions:
The cartographic design could use accepted conventions or principles of map 'goodness';

Specifications:
Some maps (especially those which are part of a series) may be designed by adherence to an explicit set of specifications;

User Feedback:
Users may be requested to subjectively rate alternative map designs;

Utility:
The map design may be objectively assessed in terms of how well it carries out specific predetermined functions, which may include aesthetic as well as more utilitarian objectives.

Clearly the last alternative is the most scientific, rational and objective. This approach led cartographers to undertake psycho-physical experiments to test specific hypotheses concerning map design parameters. It was at this point in history that the initial development of GIS started to significantly expand the analysis and output capabilities of computerised mapping systems. Increasingly, the expanded display options and use scenarios rendered the psycho-physical experiment approach inadequate. It was simply impossible to adequately define the optimisation parameters, let alone conduct the exponentially increasing number of experiments which would be required. Clearly a new approach to designing GIS visualisations is needed. It must take advantage of the understanding of human-computer interaction developed from other application domains and be soundly based on psychology theory.

2. Developments in GIS Visualisations

GIS are becoming more flexible and powerful and more closely integrated with modelling software. Hence, the communications between users and GIS are increasingly interactive and complex, especially when GIS are configured as spatial decision support systems (SDSS). Developments in GIS visualisation techniques have been driven by these factors, as well as by improvements in analysis and display functionality. Thus GIS visualisations are becoming increasingly sophisticated. They may display relationships between conceptual as well as physical entities and may involve image draping and stereoscopic viewing [1; 2].

As well as the trend towards more realistic visualisations, there is increasing use of abstract graphics to represent the distribution of non-visual phenomena and to describe processes [3]. Applications of this type often utilise sequences of graphics/images to provide an animated message, especially to communicate changes in phenomena over time, for example, in global change research [4; 5].

3. Approaches to the Design of Advanced GIS Visualisations

The role of visualisations in advanced, interactive information systems is receiving increasingly scientific assessment. This is essential because of the rapidly increasing representational potential of new systems, especially those employing multimedia or hypermedia approaches. In the GIS application domain the potential for systems to produce enhanced graphics has outstripped the ability of system designers and users to understand which visualisations are most appropriate to particular communication objectives. Visualisations should not be selected merely on the basis of tradition or expediency, but rather, so as to provide the greatest utility for the user in terms of their cognitive work requirements.

If GIS visualisations are to be optimised it is necessary to address rationally the relationship between communication objectives and the nature of the display, within a user-centred, cognitive ergonomics framework [6; 7]. This may be achieved by the use of cognitive task analysis procedures. The means-ends structure of any GIS-based decision process defines the cognitive task requirements and the sets of potential mental strategies which may be used. A cognitive ergonomics analysis enables the identification of representation and interpretation requirements. The interaction of these requirements with the viewers' roles and characteristics may be analysed to infer the visualisation design parameters.
Visualisation design may be undertaken as a formal procedure which implements a user-GIS interaction model through a cognitive task analysis procedure. However, a vast array of interaction models and task analysis procedures exist [8; 9; 10; 11; 12; 13; 14]. A Cognitive Ergonomics in GIS Reference Model may be used as a means of integrating alternative approaches and tailoring procedures to suit specific aspects of GIS design and evaluation, including the design of visualisations [7]. The initial structure of such a reference model is shown in Figure 1.

Figure 1 – Initial ‘Cognitive Ergonomics in GIS’ Reference Model Structure and Components.
4. Taxonomies for Visualisation Design/Selection Procedures

A formalism (model) linking visualisation objectives and products needs to utilise multi-dimensional, generic task and form taxonomies. Hence, a necessary aspect of the development of effective visualisation design/selection procedures is the development of appropriate taxonomy dimensions and categories. Whether any particular taxonomy dimension is useful will depend upon how it reflects the causal factors which dictate the degree of success of a visualisation sequence. This question partly turns on the nature of the design optimisation procedure being adopted. Hence, any development of visualisation design procedures needs to be embedded within a broader approach to system design methodology and visualisation tasks must form a subset of an overall GIS task taxonomy used to design and evaluate user-GIS interaction [7].

The purpose of visualisations may be formalised in terms of taxonometric dimensions such as the one suggested by Ganter [15]. This dimension classifies visualisation graphics in terms of the following broad categories of use:

1. Exploratory graphics:
   Graphics which portray the information generated from numerical simulations or other modelling, especially where there is a need to simplify the presentation or render it less ambiguous or more convincing. "These graphics usually mimic the appearance of the object or process being studied, and are often dynamic, showing behavior over time." [15, p. 234]

2. Design graphics:
   These graphics are "... an externalization of non-verbal creative thought which permit preliminary testing and comparison of solutions to technological problems" [15, p. 234-235]. The most common example is CAD graphics.

3. Reference graphics:
   Graphics of this type "... such as maps, diagrams, and curves are archives of displayed data which can be extracted and put to new uses" [15, p. 235]. They are frequently prepared for a variety of purposes, possibly some considerable time prior to their use, and their accuracy and completeness may be subject to constraints beyond the control (or even knowledge) of the user.

4. Presentation graphics:
   These are (usually simplified) graphics designed to communicate specific concepts in a particular context. Their general form may be similar to that of reference graphics.

A visualisation designer must determine what phenomena need to be displayed, and the form of the representation, so that the defined communication objectives (cognitive tasks) will be achieved. A taxonomy dimension which may facilitate this process is illustrated by the following list:

A. Phenomena visualisation:
   Depiction of natural or man-made phenomena, recorded in terms of either point, local or global variables;

B. Meta-phenomena visualisation:
   Display of the content/coverage, quality, accuracy, etc., of information sets representing particular phenomena;

C. Phenomena change visualisation:
   Depiction of phenomena change (over some specific time period) or the rate of change of the phenomena or one of its attributes;

D. Visualisations of relationships between phenomena:
   Display of specific, spatially based, relationships (e.g. correlation) between phenomena of interest;

E. Causal visualisation:
   Depiction of cause-effect relationships, known or inferred, involving the phenomena;
F. Meta-causal visualisation:
   Displays of the reliability, etc., of inferred causal relationships;

G. Information system (GIS) structure visualisation:
   Depiction of the information system's analysis and display functionality;

H. Analysis process visualisation:
   Graphic depiction of the processes of analysis used to generate a particular visualisation;

I. Motivational visualisation:
   Graphic displays designed to catch and hold the viewer's attention.

The visualisation intent must be implemented through a specific set of graphics. This requires consideration of another taxonomy dimension covering the form of presentation. An example of the sort of classification of form which may be used is the following list:

1. Direct display:
   Use of a 'realistic' visual display to depict a phenomenon which is intrinsically visual, or at least a key aspect under study is visual;

2. Indirect display:
   Graphics/images used to depict a non-visual phenomenon, where the viewer is consciously or unconsciously aware that the visual display is acting as a surrogate for something real but invisible;

3. Abstract graphics:
   Cases where information is rendered in abstract terms;

4. Metaphorical displays:
   Where the graphic display is in terms of some (explicit or implicit) metaphor;

5. Aesthetic graphics:
   Visualisations designed to produce some emotional response in the viewer.

It is important to note that the visualisation design taxonomy dimensions which are appropriate, in any particular instance, will depend (to some extent) on the nature of the information to be depicted, and on the task analysis procedures adopted. In practice, generic visualisation task and form taxonomy dimensions may need to be supplemented by dimensions which support the design process in terms of the theory of interaction on which it is based. For instance, a mode of engagement dimension may be appropriate for a task analysis procedure which is based on the 'levels of cognitive control' theoretic model [11].

For such a dimension, the viewer's mental engagement with the visualisation may be considered to be at one of the following levels:

Level I. Theory/knowledge based:
   Decision-making by the application of theories and mental models relevant to the visualisation sequence;

Level II. Principles/rules based:
   Decision-making through the use of sets of principles and rules, triggered by appropriate codes or visual cues;

Level III. Automated/skill based:
   Decision-making through automated (skilled) responses to familiar tasks represented in the visualisation.
5. Visualisation Design Within an Integrated Methodology

A useful step in the refining of design procedures for GIS visualisations is the development of taxonomies, such as that detailed above, so that the study of the cognitive aspects of visualisation may be approached in a logical manner. Whether any particular taxonomy dimension is useful will depend upon how well it reflects the causal factors which dictate the degree of success of a visualisation sequence. This question partly turns on the nature of the design optimisation procedure being adopted. Hence, any development of visualisation design procedures needs to be embedded within a broader approach to system design methodology.

Where such a design methodology implements models and procedures to ensure the cognitive effectiveness of the GIS and its outputs, it may be termed a 'Cognitive Ergonomics Analysis Methodology' (CEAM) [7]. The cognitive aspects of such analysis procedures must be based on a set of theoretical constructs operationalised through human-computer interaction models. It is also important that the integrated system development methodology incorporates organisational design procedures.

A generic CEAM could include models and procedures which address the following stages of GIS design and evaluation, in an integrated (and possibly iterative) manner:

1. Problem definition and decomposition (including defining causal relationships);
2. Analysis of decision environment and design of decision support processes;
3. Determination of required outputs and outcomes;
4. Preparation of quality management specifications;
5. Definition of goals, tasks and required information transformations;
6. Design of information processing procedures;
7. Assessment of data requirements and database design;
8. Definition of software and hardware functionality requirements;
9. Cognitive task allocation between users and software;
10. Assessment of organisational structure and personnel requirements and user characteristics;
11. User interface design or customisation;
12. Design of decision support visualisations;
13. Design of user instruction and help facilities;
14. Usability evaluation of the system and its products;
15. Audit of efficiency and effectiveness of overall system design process.

The selection of an appropriate sequence of system design and evaluation procedures from the vast array of available models and techniques is a daunting task. However, the definition of a suitable CEAM will be easier once the initial version of the proposed reference model for cognitive ergonomics in GIS is available. In the meantime, GIS design and evaluation studies utilising specific models and task analysis procedures can be used to identify the virtues and problems of different approaches.

6. Limited Case Study - Dam Surveillance SDSS Graphics

The Melbourne Water Corporation (MWC) is a corporatised government agency responsible for the building and management of water storage structures (dams) in the Australian State of Victoria. The principal responsibilities of the Corporation’s Structural Surveillance Section are the analysis of stress and deformation measurement information to determine any requirements for urgent action, the development of structure management action proposals with engineers from other sections, and the provision of risk management advice.

There could be said to exist a 'corporate' model of the stress and deformation behaviour of any of these dams: 'corporate' in that it is both a shared model and because it constitutes the official position of the organisation. This model is constituted of various explicit documents and data sets summarising the current best understanding of the situation, as well as implicit shared theoretical and practical knowledge embodied in the expertise of the relevant personnel. It is important that this 'corporate' model be as explicit as possible, since it must be used by personnel other than structural and civil engineers (e.g. for risk management), and so that MWC management are able to understand these matters and balance structure management considerations with financial and other matters.
A knowledge base system based on GIS for thematic mapping
B. Su, W. Zhang, H. Li, X. Zhang, Y. Zhang, X. Zhu, J. Li (Wuhan, RC)

ABSTRACT
Along with the development of science and technology, the actuality of various geographical information, of map products in particular, is more required. But it is a creative procedure to design and produce a map (or an atlas). This procedure requires both professional expertise and artistic accomplishment. So it is never an easy task even though we are now in the digital information era, because existing CAC systems cannot satisfy all the requirements of a 'real' design. To solve this problem, it is necessary to make CAC systems 'intelligent' and to make as much profit from GIS system as possible. MAPKEY is such an intelligent system developed by the authors, which is based upon GIS systems and integrates Database, Knowledge base system and computer graphics system. MAPKEY can cover almost all procedures of thematic map making, such as map type selection, symbol design, data grading, color design, four color separation and film output, and provide a color environment, 'What you see is what you get' in its real sense. In this paper, the 'inside story' of MAPKEY is revealed and some concepts of map design are also discussed.

KEYWORDS
Knowledge base, Data base, Computer graphics, Thematic Mapping, Automatic color assignment, Four color separation, MAPKEY

1. INTRODUCTION
Thematic map design is a decision—making process requiring expertise

* The Project Supported by National Natural Science Fundation of China
for map content select and representation, map material analysis, map
data processes, the selection and design of symbols, data processes, the
selection and design of symbols, legends, colors and typefaces. Thematic
mapping or thematic map production is a productive process including
planning, design, drawing, color — separating, plates — making, pre —
press proofs and printing. So it has a very long production period for
producing an atlas. 【1】

Map production becomes easier, faster and more efficient due to the de­
velopment of mechanics and electronics, especially, the application of
computer science and technology to map design and production has
brought about great changes of cartography. However, to make use of
computer techniques accomplishes automation of map design and produc­
tion, there still exists a lot of problems.

1. 1 THE EXISTING PROBLEMS IN THEMATIC MAPPING

(1) More and more poor — quality maps are being produced by non­
cartographers who use some widespread thematic mapping package.  
【2】 【3】

(2) Map design becomes the “bottle neck” problem in the thematic map­
ing software since it is a creative procedure that general programming
 technique can barely describe such procedure.

(3) Mapping procedure is very complicated one even if in GIS age, for
example, “US ATLAS: 1990 Road Atlas” was completed by 4 groups
using 2 kinds of computer and 6 software packages. 【4】

1. 2 NEW DEFINITION OF CARTOGRAPHY AND MAP (Board, 1989)

(1) CARTOGRAPHY: The organization and communication of geo­
graphically related information in either graphic or digital form. It can
include all stages from data acquisition to presentation and use.

(2) MAP: A holistic representation and intellectual abstraction of geo­
graphical reality, intended to be communicated for a purpose or purpos­
es, transferring relevant geographical data into an end—product which is
1.3 FEATURES OF MODERN THEMATIC MAPPING

1. GIS and Database maintain the accuracy of original data of map.
2. New kinds of maps with Multi-form or Multimedia emerge continuously (Dynamic map, CD-ROM map, Electronic map, Sound map, Video map, etc).
3. It becomes possible for computers to simulate design procedures of experts because of the development of a technique. The concept of map design has changed a lot. Map design is a process transforming map data stored in GIS into "Visual, digital or tactile products" instead of transforming the artworks.
4. MAP user can enter into design procedure under the guidance of knowledge base.

1.4 THE WORK DONE IN THIS PAPER

In this paper, a structure of unified intelligent thematic mapping production software system is developed (Fig. 1)

<table>
<thead>
<tr>
<th>Unified software system</th>
<th>Unified User Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to use, easy to understand</td>
<td>Text, Desktop, Color, Publish, 4 color, Separating, &amp; Correcting, Film, Output</td>
</tr>
<tr>
<td>Various facilities needed for map production (Improving transformation efficiency and speed)</td>
<td>Expert System (Knowledge base, Inference)</td>
</tr>
<tr>
<td>The kernel of system control (Emphasizing map design, embodying art style and simulating creative procedure)</td>
<td>GIS &amp; CAC, Database Management System</td>
</tr>
<tr>
<td>Basic graphics and data processing (Maintaining accuracy and improving mapping efficiency)</td>
<td></td>
</tr>
</tbody>
</table>

*Fig. 1 Structure of Intelligent Thematic Mapping Software*
1. 5 FEATURES OF THE SYSTEM STRUCTURE

(1) The system is based on GIS (computer graphics and database).
(2) The knowledge of experts and creative thinking simulation are presented by techniques of expert system.
(3) Text procession, color electronic typesetting and 4 color separation printing are introduced.
(4) A unified user friendly interface which is easy to be used and understood is emphasized.

The 5 existing views in cartography could tend to unification based on the system in this structure and the problem mentioned above could be solved effectively.

2. MAPKEY — A KNOWLEDGE—BASED SYSTEM BASED ON GIS FOR THEMATIC MAPPING

Based on the structure described in Figure. 1, a practical Expert System —MAPKEY is implemented. [5] [6] [7]

Figure 2 shows the main elements of MAPKEY.

(1) Reasoning (inference engine) is the control part of MAPKEY. It decides the inference strategy of the system.
(2) **Knowledge base management** is the core of MAPKEY. It manages all map design knowledge represented by Context, Frame and Rule, etc.

(3) **Unified interface** is an intelligent interface between users and the system. Users can conveniently use and control the system with it.

(4) **Database management** part includes geo information management and attribute information management.

(5) **Text process** provides an environment editing text file.

(6) 4 color separation finishes to transform RGB color system into CMY color system and adjust colors. It is the fundament for the system to accomplish "What you see is what you get".

(7) **Desktop color publishing** performs characters publishing (including cartographic lettering) and map publishing. It can output 4 color separated films.

(8) **GIS interface** is a kind of data transforming mechanism between the system and other GIS. Through it, the system can receive geo and attribute information from popular GIS (e.g. Microstation, Mapinfo, AutoCAD, etc) and other database systems.

(9) **Computer graphics** performs functions of graphics operation environment.

### 3. MAIN FUNCTIONS OF MAPKEY

MAPKEY can implement almost all procedures of thematic map making. The following is its main functions:

- Data preparation
- Map type selection
- Data classification
- Symbol design and adjustment
- Legend design
- Color design
- Map lettering
- Four color separating
- Color separated films output
4. THE BASIC WORK PRINCIPLES OF MAPKEY

In MAPKEY, many new techniques of computer are used. Expert system technique in AI is one of the most important techniques.

4.1 KNOWLEDGE REPRESENTATION METHODS

The methods of representing and managing knowledge are CONTEXT, FRAME and RULE, etc.

**CONTEXT** — to describe map content

![Diagram of Thematic Map Context]

**FRAMES** — to present the hierarchy of maps and knowledge

- To represent the hierarchy of map details
• To express the structured knowledge
• To describe the relationship of map elements

**Fig. 4 Simplified Frames in MAPKEY**

**Fig. 5 Working Flow Chart of MAPKEY**

**RULES**—to express the process of design

**FORM:**

IF condition is true
   THEN take actions
ELSE take actions

• Build the reasoning chain
• Connect frames, database, procedures, graphics and knowledge base

**EXAMPLES:**
Rule 097:
IF mapdata. property = "total" and 
frame—get(mapframe = mapslot) = "symbol"
    THEN symbolsel = TRUE
    DO symbolsel (total)
        symbnum = 1

Rule 082:
IF symbnum = 1 and accuracy = TRUE
    THEN symbolname = "bar"

4.2 ARCHITECTURE OF THE SYSTEM

MAPKEY is an intelligent system integrated with techniques of expert 
system, database system, computer graphics, typesetting and color pro­cessing, etc. The Figure 6 shows its architecture in detail. Inference En­gine controls the action of knowledge bases. It consists of Blackboard sys­tem, Strategy of "generate and test", Forward — chain, Backward —chain and Mixed—chain, Fuzzy reasoning, etc. The Blackboard system 
is the core of Inference Engine. With Generalized Frame, it manages 
all reasoning procedures of the system. In MAPKEY, through General­ized Frame, it is possible for frames, rules, procedures, database, text 
processing and graphics processing to reference one another. This is the 
fundament for us to develop MAPKEY.【9】
According to the architecture described in Figure '6, the working flow of 
MAPKEY is shown as Figure 5.

5. SYSTEM CONFIGURATION OF MAPKEY

Aim of designing MAPKEY is low price and high performance. The sys­tem configuration is shown as Figure 7.

6. APPLICATION

"The Atlas of Economic Data of Agriculture in Hubei Province" has 
been designed and produced by MAPKEY.
7. CONCLUSION

(1) MAPKEY is a practical intelligent thematic mapping system which represents the main tendency of modern thematic mapping.

(2) Experience based on MAPKEY has proved that AI combining with mechanic and electronic techniques is a powerful techniques for modern thematic mapping.

(3) The method of Expert system has solved the problem of automatic color assignment (a long—unsolved problem).
(4) Text processing, electronic typesetting and 4 color separation printing can run in a same software system so that it embodies real "what you see is what you get" for color graphics text.

The most important point is that "Data is only data, knowledge is really power".

REFERENCE

User differences in a GIS environment: a protocol study
C. McGuinness, A. van Wersch, P. Stringer (Belfast, GB)

INTRODUCTION

We are engaged in a research programme on how users interact with map displays within a GIS environment. Our primary focus is on the cognitive processes of different types of user and we are developing different methodologies for examining user-GIS interactions. Previously, we have argued that the interactive nature of GIS places high demands on the spatial-information-processing capabilities of users. GIS use requires a cognitive analysis, not only of map displays, but also of the prior knowledge and spatial abilities of the users. This paper presents the first in a series of experiments on user differences in a GIS environment, using a protocol methodology.

If it does not seem too obvious a point to make, what is special about GIS, compared to other information systems, is its "spatialness". GIS permits users not only to produce computerised maps but to interact with databases to learn facts about spatial data, to hypothesise, analyse and synthesise spatial relationships. For this reason we concentrate on the users' interactions with the map display. From the system point of view GIS display maps: from the cognitive point of view GIS demand that users become map-makers which implies some knowledge of cartographic design principles. From the system point of view GIS supports spatial analysis - database queries, spatial queries, attribute queries, overlaying, networking, and spatial statistics. But from the cognitive point of view, GIS demand pattern recognition about the incidence and the distribution of features as a result of a query, inferences about spatial relationships, hypothesising about the next query, spatial memory in a sequences of queries, complex pattern recovery as a consequence of overlaying, and so on. As yet, we know very little about the information-processing demands of even the simplest GIS tasks.

Our research programme adopts a firmly cognitive perspective of the user - a model of the "map user as thinker". Our general view is that understanding is concept-driven as well as data-driven. This perspective contrasts sharply with the model of the map user in traditional cartographic evaluation research: there, the map user was often portrayed as a passive recipient of cartographic information - a blank screen on to which map information was registered. When the map user is construed as a thinker who interacts with the map display in terms of his or her own cognitive processes, then considerable variability can be expected between users who are doing the same task.

The origins of these individual differences are many - the general ability of the user, prior knowledge and training, purposes and goals, and so on. Indeed, one of the most striking and consistent findings in cognitive psychology over the past few years is the importance of expertise in complex information processing, particularly in the interpretation of visual and graphical displays. For example, studies have examined the impact of prior knowledge and expertise on memory for chess positions, for electronic circuit diagrams, for architectural plans, for x-ray diagnosis, and for interpreting hardcopy maps. The studies all conclude that the pattern recognition which is typical of the expert has a major impact on how a graphical
display is interpreted and searched, and ultimately, on the problem solutions generated from that display. Pattern recognition ability, they argue, is deeply embedded in the knowledge structures of experts.

The influence of prior knowledge and expertise, as a source of differences between users, is one of the primary variables of interest in our research programme.

While the application of GIS is at present restricted, it is argued that in a short time they will make available map-making and spatial analysis facilities to a much wider community of users - a community who may not be as adept at spatial reasoning (or computer use) as environmental scientists, surveyors, cartographers, etc. GIS may increasingly be used to educate and communicate with planners, politicians and the public. A number of writers have already pointed to user variability within a GIS context - referring exclusively to variations in computer skill. For example, in their paper on lessons learned from a post-mortem on a failed GIS, Openshaw et al. refer to a number of user issues "GIS is quite different from many other types of computer activity in that it is essentially end-user technology and end-users are increasingly not computer specialists". Raper and Bundock's purpose in designing a "layered user interface" is to allow access to the functionality of the software to multiple groups of users - the infrequent or job-limited user, the spatially aware professional, the database administrator, the system developer. It could well be argued that the same set of users vary substantially in their map-making and map interpretation skills. For these reasons the research aims to study a wide range of users with different levels of expertise.

Hardcopy map evaluation research has been criticised for using artificial map stimuli in experiments and for studying only limited and rather simple map reading tasks. In developing experimental tasks for our studies we have sought to increase the ecological validity of the work by designing "problem scenarios" which mimic as closely as possible what a GIS user might be required to do in a real GIS environment. To achieve this we have used existing spatial databases which were part of on-going GIS projects in the Northern Ireland Regional Research Laboratory.

Because a cognitive perspective is adopted in the research, we are particularly interested in the knowledge structures of the expert and novice users as they interact with the maps in the environment. In making expert/novice comparisons, a number of experimental techniques are commonly used. Subjects are asked to exercise their skill in the normal way. They are given a mini-problem to solve which is very similar to the type of problem which they encounter as part of their everyday practice. Their responses are timed and a record is made of their solution attempts. But because much of what is interesting about their cognitions is going on inside their heads, the method of "thinking aloud" or verbal protocol analysis can be used to externalise their thought patterns. Using this method the subjects are asked to give a running commentary on their thoughts and decisions; these are recorded verbatim and then analysed. Verbal protocol analysis is clearly very time-consuming but it does provide in-depth qualitative information on cognitive processes. It can be very helpful in finding out how experts and novices spend their time on a task and why they generate certain solutions. In our first study a variation on the thinking aloud method was used (see method section).

The study reported in this paper is the first in a series of experiments which examines user interactions with a map display within a GIS environment. The study aims to observe and compare how expert and novice users of GIS create and manipulate map displays when presented with a relatively open-ended spatial data exploration task. The emphasis of the study is on the cognitive processes of users as they interact with the map, using a variation on verbal protocol methodology we want to study, not just what they do, but how they do it, and why.

**METHOD**

**Subjects:** Eighteen users (10 men and 8 women) were studied. Nine users had some GIS experience and constituted the 'expert' group, although their GIS experience was limited (7 were geography undergraduates who had recently completed a 10 week hands-on course on GIS, 2 were research fellows who had self-taught GIS skills). The other nine users - the novices - were psychology undergraduates with no knowledge of GIS. All users were
computer literate. Additional user information was collected on the level of geography education (course participation, e.g. O-level, A-level, 1st year undergraduate, degree level, or higher), a self-rated questionnaire (ratings from 1-5) on frequency of map use, degree of map enjoyment and map competence; two psychometric tests of spatial ability were completed by each user, Map Memory and Hidden Figures from the Kit of Factor-Referenced Tests. These tests were designed to measure individual differences in general visual memory for map-related materials and the ability to find simple forms in complex figures.

Tasks: As the focus of the research programme was on user interactions with the map display, tasks had to be designed (and software chosen) which gave subjects access to the map display with the greatest ease and, at the same time, generated map displays of sufficient variability to allow user differences to be examined. Additionally, the difficulty level of the task had to be such that the experts would not find it trivial and that the novices would be able to make a reasonable attempt. Two problem scenarios were developed to meet these criteria. Both tasks involved data exploration and plotting coverages/variables on maps and the chosen spatial databases were part of on-going projects in the Northern Ireland Regional Research Laboratory. The tasks were run through ARC/INFO (the display module, ARCPLOT) on an IBM PC compatible: using the ARC/INFO macro facility we have succeeded in making the map plotting readily accessible, even for novice users.

1. The Peat Scenario: This task drew on a database to do with machine peat cutting in the Sperrin Mountains in Co. Tyrone in Northern Ireland. Nine variables were available for plotting - incidence of machine cutting, type of peat, contours, roads, rivers, rainline, rainstations, areas of natural beauty/scientific interest, density of wading birds. ARCPLOT permits a number of coverages to be generated, e.g. point data (incidence of machine cutting), line/arc data (rivers, road, contour lines), polygon/area data (type of peat, area of natural beauty) or any combination. The purpose for which the data was originally collected was to examine how environmental and social variables were related to the incidence of machine cutting, with particular reference to any conflicts which might exist between economic and conservation pressures. This general storyline formed the background to the scenario and the experimental task mimicked as closely as possible what a "real" GIS user might do when s/he first viewed the data - examine the distribution of variables, explore the relationships and begin to test for patterns.

Subjects were invited to pretend that they had to prepare a brief for a talk to the local conservation society on the impact of machine peat cutting in the area. They were presented with a blank screen on which to draw their series of maps, and the list of variables which could be plotted. Essentially the subjects had to choose what variables to plot, how many variables (out of 9) to overlay on a single map, what combinations to view, and the order in which to view the displays.

2. The Health Scenario: The second task drew on a database which consisted of health and deprivation variables which had been collected at the level of District Council (N=26) in Northern Ireland. Nine variables could be plotted: rate of premature mortality, % permanently sick and disabled, % low birth-weight (these were the health indices); % unemployed, % households without a car, % rented accommodation, and % overcrowding (these were the deprivation indices); additional variables were social class and religion. These are quantitative variables and the amount of a variable for each District Council (polygon) can be indicated through shading, colour, or graduated point symbol. To simplify the data, each variable was dichotomised into High/Low and a large or small triangle signified the status of variables for each District Council. The purpose for which the data was originally collected was to examine the spatial distribution of health and deprivation indices, to see the relationship between the variables, and to look for spatial patterning. This general storyline formed the background to the problem.

Subjects were asked to pretend that they were acting as research assistants for their local Member of the European Parliament; they had to prepare a brief about the spatial relationships between health and deprivation indices in Northern Ireland. On the screen they
were presented with a map of the District Council regions of Northern Ireland and a list of variables to be plotted. Only two of the statistical variables could be plotted at any one time. Subjects had to choose what variable to plot, how many variables (1 or 2) to plot simultaneously, what combinations to view and what order to view the displays.

For the two scenarios, how the subjects proceeded, what variables they plotted, the reasons they gave for their choices, and the conclusions they reached, constituted "the interaction" with the map display which is the focus for this study.

Procedure: Subjects were tested individually and completed the two interactive problem scenarios in a single 2-3 hour testing session. All subjects completed the Peat scenario first, followed by the spatial ability tests, the Health Scenario, and finally, the self-rated map questionnaire.

Subjects were given written instructions about the nature of the task requirements and were encouraged to read them and to ask for clarification before they started. They were also given practice on how to use the mouse to select variables. During the task they were allowed to keep notes; at the end of each scenario they had to prepare summary notes to explain and justify the conclusions they had reached.

Initially, it was planned to collect concurrent thinking-aloud verbal protocols as the users completed the tasks. On extensive piloting however it was found that concurrent verbal protocols disrupted performance; subjects often fell into silences and found prompts distracting. We decided to adopt a slightly different approach and to videotape the subjects as they did the tasks. The videos were immediately replayed after each scenario and subjects provided a running commentary on their own performance which was dubbed onto the video. This method allowed subjects to give their full attention to the task and, during the replay, the experimenter could collect additional information about reasons and choices.

Measures: For each subject the following data were available for analysis.

1. Performance: Time spent on each scenario, including time to read the instructions (in minutes); number of trials per scenario (a new trial began when previously plotted variables were zeroed); absolute number of variables explored per scenario (maximum of 9); number of variables plotted per trial; number of repetitions (variables replotted over a series of trials).

2. Summary notes: Essentially these notes revealed how well the subjects had performed the tasks. There were no absolute correct answers because of the open-ended nature of the tasks, so criteria to assess the quality, goodness or appropriateness of the answer were designed: the criteria used were fulfillment of task requirements, reference to spatial distribution, level of detail, level of interpretation, level of planned structure. When the summary notes were rated according to these criteria (by two independent raters) they successfully distributed the notes along a dimension - some of the "answers" were judged as very good while others were judged as much poorer (when the notes were scored, the raters were blind to their status as expert or novice).

3. Verbal protocols on video: A number of techniques are available to analyse verbal protocols depending on the level of information which is required: for example, each sentence of assertion can be coded into a category; short sequences of assertions can be coded as evidence of a strategy or stage; or more global characteristics of the protocol can be identified based on larger chunks of information. For this study thirty-eight dubbed verbal protocols (two per subject, one for each scenario) were viewed and transcribed. A global analysis was adopted as no detailed hypotheses were being tested and some 40-60 minutes of protocol was analysed for each subject. Analyses revealed that whole protocols, or segments of protocols, could be categorised in terms of the search strategies which the subjects were following: for example, random or exploratory search, focussed search around a limited number of variables, systematic search through most of the variables, or some combination.

Thus, three sources of evidence - time and variables plotted (from the screen interaction), quality of answers (from the summary notes), and reasons and strategies (from the verbal protocols on video) were analysed.

481
RESULTS

Before reporting how the subjects performed on the tasks additional information about the characteristics of the groups will be described.

Characteristics of the subject groups: The subjects were recruited as expert or novice depending on the amount of previous GIS experience and training. The experts' GIS experience was not extensive; however, 8 of the 9 subjects were either final year geography undergraduates or recent geography graduates. In contrast, five of the novices had finished their formal geography education during secondary schooling (after O-levels, at 15 years) and only two of this group had studied 1st year geography at university. Statistical comparisons between the group showed that the experts had significantly more formal education in geography than the novices, t(16)=4.5, p<.001. With regard to their self-rated frequency of using maps, self-rated map competence, and self-rated enjoyment of maps, only frequency of use yielded a significant difference between the groups; experts reported that they used maps more frequently than the novices, t(16)=2.2, p<.05, although self-rated competence was marginally significant (p=.08). No significant differences between the groups were evident on the psychometric tests of spatial ability. All subjects scored high on the Map Memory test (indicating that perhaps the test was too easy for the group overall). The experts scored higher (mean=5.8) than the novices(mean=3.8) on the Hidden Figures test, but the differences were not significant. Overall the scores on this test were relatively low.

On-Screen Performance: For both scenarios, the experts spent longer exploring the database than the novices, but these differences were not statistically significant. For the Peat task, the experts' mean time on the task was 29 minutes compared to the novices’ mean time of 24 minutes, t(16)=1.3, p=.24; for the Health task, the respective means were 22 and 18 minutes, t(16)=1.45, p=.17. The correlations between subjects' performances on the two tasks was reasonably high (r=.5), indicating that those who explored for longer on the Peat task also spent longer on the Health task.

During that time subjects explored the data by plotting and reploting variables as map displays. There was no significant difference between the groups in the number of trials they attempted (a new trial began when previously plotted variables were zeroed and a new display was generated). In general, experts had only marginally more mean trials than the novices; 10 trials compared to 8 trials on the Peat task, t(16)=1.5, p=.15, and 17 trials compared to 16 trials on the Health task, t(16)=.3, p=.76.

For each scenario, nine variables were available for plotting. With regard to how exhaustively the data was explored, all the experts plotted 8 or nine variables at least once for the Peat scenario; there was more variability among the novices. The same pattern was evident for the Health scenario; 8/9 experts explored all the variables compared to 5/9 novices.

In the Peat scenario where it was possible to overlay from one to nine variables on the same map display, there was a slight tendency for novices to overlay more variables than the experts. The mean number of overlays per map for the experts was 2.2 variables and they rarely went beyond 3 variables. While the mean number of overlays for the novice group was only marginally higher (mean=2.6), there was much greater variability in the group - three novices overlaid very few variables, while three others overlaid many variables - one novice plotted all nine variables simultaneously!

Experts had a higher number of repetitions than novices; they replotted combinations of variables which they had previously viewed more frequently than the novices. For the Peat scenarios, experts had 30 repetitions compared to 9 for the novices; for the Health scenario, they had 20 repetitions compared to 11 for the novices.

Summary notes: After each scenario was completed the subjects wrote summary notes to fulfill the requirements of "preparing a brief". These notes were rated according to the method described above, to assess how well the subjects had completed the tasks. The summary notes were then ranked in terms of their quality from 1 to 18 (rank 1 indicating the best answer). Experts' answers were judged to be significantly higher in quality than novices' answers. The mean ranking for the expert group on the Peat task was 6.2 compared to the novices' ranking
of 12.8 (Mann-Whitney U=10.5, p<.01). On the Health task the experts' mean ranking was 5.5 compared to 13.5 for the novices (Mann Whitney U=4.5, p<.001). The correlation between the subjects' rankings on both scenarios was highly significant (r=.69) showing that there was consistency within subjects' performance across tasks.

As well as discriminating between expert and novice groups, quality of summary notes were negatively correlated (indicating a positive relationship) with other characteristics of the subjects as well as with several on-screen performance measures. Ranked quality was positively related to scores on the Hidden Figures test (r=-.27 for Peat, r=-.32 for Health), to self-rated frequency of map use (r=-.33 for Peat; r=-.33 for Health), and to level of geographical education (r=-.59 for Peat; r=-.44 for Health). The quality of the summary notes for the Peat scenario was positively related to other Peat scenario variables - time spent on task (r=-.6) and absolute number of variables explored (r=-.46). An identical pattern emerged for the Health scenario where the corresponding correlations were r=-.51 for time spent on task and r=-.31 for number of variables used.

**Verbal protocols on video**: When the videos were replayed and used to prompt subjects to give running commentaries on their performances, experts spent on average 53 minutes (28 minutes on Peat, 25 minutes on Health) talking about what they had done compared to novices who spoke 45 minutes on average (22 minutes on Peat, 23 minutes on Health).

The protocols consisted of descriptions of what they saw on the maps (e.g. "cut peat is at lower altitudes"), inferences about the relationships (e.g. "what would be expected......lower altitudes are more accessible for machines") and hypothesis testing (e.g. "now I will look at the road system to see where the roads are in relation to the cut peat"). There was also evidence in the protocols of more global hypothesis testing (e.g. after viewing the spatial distribution of low-birth weight babies, an expert commented "I thought it was more a core and periphery thing.....I'm surprised"). Comments were also made about the reasons for overlaying (e.g. an expert commented on the Health scenario "it was good to do overlays but not multiple overlays....it was too hard to see the relationships", while a novice said about the Peat task "the peat was more helpful and detailed .....could put more on, as much as you wanted"). Reasons were also given for repetitions (e.g. "I'm doing machine cutting again, for a quick check" and, referring to the density of wading birds "I'm plotting it on it's own, the other one was too cluttered").

These comments are examples of the types of individual statements which were made by the subjects as they described and justified their own performance. A full analysis at the level of individual statements will not be reported here.

From watching the videos and reading (and re-reading) the transcripts, it was clear that subjects adopted (or developed) an overall approach or strategy to the data exploration problem which determined their choice of variables to plot on a single display and the overall sequence of their queries. We called these overall approaches "search strategies" and three different types were evident in the protocols.

1. **Exploratory or random search**: subjects who followed this strategy appeared to have no overall plan or purpose in selecting variables to display other than just to "see" the relationship. They did not test the relationships between variables in a systematic way; hypothesis testing was superficial and not pursued over a sequence of plots.

2. **Focussed and selective search**: subjects showed evidence of this strategy by planning their exploration systematically around a small number of variables and relentlessly pursuing these relationships, forming and testing hypotheses.

3. **Systematic search**: subjects who used this strategy adopted a planned approach to the exploration of the variables and systematically tested the relationship between variables, continually hypothesis testing and theory building. Systematic searchers differed from focussed searchers in their ability to build theories around a larger number of variables and to shift hypotheses in a more flexible way. They also checked and rechecked their interpretations.

Type of search was sensitive both to level of expertise and problem scenario. In general the Health scenario afforded systematic searches more readily than the Peat scenario which was more likely to be searched in a focussed way (appropriately around machine cutting). Experts tended to be more systematic in their searches than novices.

For the Health task 6/9 experts showed evidence of systematic search either by itself (4
Experts) or in combination with focussed search (2 experts). Two experts were exclusively focussed searchers and the remaining expert combined a focussed with an exploratory style. In contrast, 6/9 novices were exploratory searchers on the Health scenario; three of them showed this style exclusively and the other three combined it with a more focussed search. One novice had a consistently focussed style and the other two were systematic searchers. For the Peat task 5/9 experts were systematic or combined searchers, two others were focussed, and the remaining two were combined exploratory/focussed; 6/9 novices explored the peat database in an exploratory fashion, one was a focussed searcher, and the remaining two were classified as systematic.

CONCLUSIONS

The first point to be made is that all users - experts and novices - made a reasonable attempt at meeting the requirements of the tasks set by each scenario. With very little initiation, they sat at the computer, read the instructions, explored the databases for 20-30 minutes per scenario, summarised their conclusions, viewed their performance on video, and explained what they were doing and why. The expert group did not think that the task was trivial while the novices were not overwhelmed. Both groups were able to examine the distribution of variables, hypothesise about the relationships and patterns, and bring their findings together in a comprehensive and sensible manner. Only one subject - a novice - arrived at very inaccurate conclusions.

The expert group did not have extensive GIS experience so the range of GIS-specific expertise represented in this study is not very wide. The experts did differ significantly from the novices on other relevant variables - level of geographical education and self-reported frequency of map use. Their general levels of spatial ability were very similar.

The number of subjects studied in both groups is relatively small so inferences about differences between groups and/or relationships between measures (whether statistically significant or not) are suggestive only. Notwithstanding the overlaps between the two groups, a consistent profile of expert and novice interactions with the map displays did emerge. Experts spent longer on the tasks than novices. Although they generated a similar number of map displays (trials) as novices, they explored more variables (absolute number of variables plotted); they plotted and replotted the same combinations of variables more frequently then novices (repetitions). In the Peat Scenario where it was possible to overlay large number of variables, the experts were conservative about overlaying. In contrast to the novices, they were intent on keeping the map display simple and on plotting and reploting variables in sequential displays rather than cluttering a single map display with too many variables.

From this profile of interaction with the database the experts seemed to gain more knowledge and were able to give better quality answers to the general problem posed by both tasks. Compared to the novices, expert answers (from summary notes) met the task requirements more adequately, made more frequent references to spatial relationships, and were more planned and interpreted.

From the verbal protocols, experts were more likely to make systematic searches through the databases and to explore them in a planned and careful way (plot more variables); they consistently tested hypotheses, checked their interpretations (high number of repetitions) and reached comprehensive conclusions (quality of answers). In contrast, the novices often randomly searched, missed out on variables, and engaged only in local and superficial hypothesis testing. With regard to viewing variables some novices believed that it was possible to successfully "read" many overlays and were frustrated by the restriction to two variables in the Health problem. In contrast, some experts commented that two variables were adequate for the type of information being presented: they appeared to know the limitations on their own information-processing capacities. They strove to maximise their understanding through minimising the complexity of the map display.

Although these profiles are merely suggestive they do give some indication of an emergent model of expert and novice map user as thinker which will be further tested in the next stage of the research programme.
Acknowledgements

This research programme is supported by the United Kingdom's Economic and Social Research Council Grant No. R-000-23-2337. We would like to thank Dr Roy Tomlinson and Dr Margaret Cruikshank, School of Geosciences, The Queen's University, Belfast, for permission to use their database on machine peat cutting for the Peat Scenario; and Ms Paula Devine, Northern Ireland Regional Research Laboratory, for technical assistance in developing the problem scenarios on ARC/INFO.

REFERENCES


4. EGAN, D. & SCHWARTZ, B. 1979: Chunking in recall of symbolic drawings. Memory and Cognition 7, 149-158


10. POTASH, L.M. 1977: Design of maps and map-related research. Human Factors 19, 139-150


Session 10

New Tasks, New Techniques, New Terms II

Chairman:
E. Bos, ITC (Enschede, NL)
Dynamic maps as a new type of cartographic production
A.M. Berlyant, L.A. Ushakova (Moskva, CIS)

Introduction

Recent trends in thematic cartography are characterized by a tendency to create synthetic imagination of the mapped object and to show nature and society in dynamics.

Dynamics in geographic sense is a motion peculiar to any geographic category. Dynamics forms structured links and relationships between different elements of the geosystem. Mapping the dynamics of phenomenon is depicting its appearance, development, spatial and temporal changes.

Two main methods are widely used in cartography: representation of the dynamics on the same map using special ways of depicting, classification or indices. The other method is based on the compilation of the series of multitemporal maps. In practice these methods are often mixed: dynamic map design in some cases is accompanied by the comparison of different maps while the compilation of series of multitemporal maps does not dismiss representation of dynamic features in each map.

Efficiency of current spatial and temporal research on the objects and processes is connected with the geoinformation mapping (GI) or computer assisted cartography (CAC) on GIS base. Geoinformation mapping is forming now on the base of integration of CAC, remote sensing, GIS technology and system map design. Dynamic cartography is one of its main trends. Dynamic cartography represents dynamic map series of thematic and spatially related frames, which display the sequential stages (dynamic phases, trajectories) of events and processes thus producing simulation of motion and changes. Dynamic map series may consist of multitemporal images or maps both 2D and 3D and may represent real objects or abstract models.

Review of dynamic mapping techniques in geosciences

Dynamic mapping is often used in geosciences, such as geography, geology, oceanology, meteorology, etc. It forms a special class of geoinimages - spatially and temporally scaled and generalized models of terrestrial objects and processes presented in graphic mode. Dynamic map series are most suitable for monitoring and dynamic cartometry. One can suppose that in the nearest future the dynamic mapping will be a routine tool for the analysis of the environment like hardcopy maps, space images or digital maps.

Dynamic mapping is based on animation. Frequently it is called animation cartography and dynamic map series are called animation sequences. Early research on application of animation principles for animation cartography was conducted by N. Thower (15). His ideas were realized 20 years later when the results of the first computer animated map sequence were published. N. Thower noted some interesting peculiarities in the filmed sequences made between 1936 and 1957. Scales were rarely used,
although a time scale such as 'one second equals ten years' was employed. Systems of orientation, such as lines of longitude and latitude, also were used infrequently. Legends were virtually absent in view of the fact that information was probably expressed verbally in a soundtrack much quicker than it could be read by the viewer. These are still important elements of animation sequences design (6, 14).

The development of computer processing in the 1960s increased interest in data manipulation and analysis particularly in the area of animation cartography. The computer allowed the researcher to analyze huge amounts of data, to detect and correct errors and to display whole animation sequence in a short period of time.

One of the first computer animation sequence was a film of Bell Laboratories which displayed the orbit orientation of a satellite. The first article addressing the potential of computer animated films in cartography was published in 1966 by Cornell and Robins. They proposed the use of a computer vector screen display and an open-shutter camera to capture successive images on film, creating an animated sequence (8).

In 1970 in USA there was a small peak of interest in animated cartography. W.Tobler designed a display film of 3-D presentation of a population growth model for the city of Detroit. D.Douglas produced an animated demonstration of relief shading, showing what happens when a light source rotates around a topographic feature. Historical geographer S.Hiller made an animated movie of Indian land cessions in the United States (6, 16).

In mid 1970s there were published several review papers devoted to animated cartography, some programming languages were designed oriented towards the animation. The researchers from the Department of experimental cartography developed a method for dynamic mapping on the example of Ireland Sea using displays and TV information system of VIEWDATA type (4). Their display films represented the motion of streams, currents and sea fronts. These films allowed not only to model and visualize rapid changes of the environment, but as well to make forecasts, to make solutions on deep sea and sea shelf resources management, on the protection of sea environment, etc. The researchers suggested to design two types of films: inexpensive and primitive "data films" for specialists to evaluate the collected data, and "demonstration films" to distribute the processed data among the specialists in related fields of research.

However, the animated cartography did not developed rapidly in 1970s. The real conditions for its advance came with the progress of the automation and microcomputer technique. The development of increased power, speed, and efficiency in computing, from PC to mainframe, monitor display improvements and progress in the area of videocassette recording have made for a more favorable environment for the increased distribution and usage of animated cartographic techniques.

There are few methods for the design of animated sequence using PC.
1) The computer under control of the operator drafts a sequence of individual frames thus allowing for easy changes and updates. 2) Separate animation frames are designed using special computer drafting programs, stored on a hard disk, and then recalled to produce animated sequence with any speed controlled by the operator. A large number of the software for such type of animated sequences design is developed now including commercial products.
3) The direct transference of graphic images from computer to videotape with the later viewing on a normal television screen.

Many researchers were engaged in the design of dynamic geoimages (9, 10, 13, 14, 16). The fields of current research concern the design of 2D and 3D animated sequences in color and monochrome mode, the usage of cardiograms, graduated symbols, multi-variable and isoline maps, holograms in real-time animation.
There were conducted some successive research works on animated cartography by the geographers of Georgia. GIS of Georgia was used as a data base for the design of animated sequences. This GIS functions in the Laboratory of the Environmental Research in Tbilisi University. The design of animated sequences was fulfilled using automated mapping system on the base of "Atari" personal computer with the color display and printer (3).

One display film created by the Georgian researchers depicted the disastrous situation in Georgia in winter 1987 where one could detect the changes in landscape and ecology day after day beginning from the 1 of January till the 8 of February. The other sequence showed the dynamics of the environment of Georgia from the 1 till the 10 of September 1987. These films showed a good presentation of spatial and temporal dynamics of the environmental processes. They revealed the advances of the cyclones moving over the territory of the republic, the regions of heavy precipitation and rapid changes of the landscapes. Sparkling cartographic symbols attracted attention of the viewer to the regions of the dangerous and hazardous natural phenomena. Avalanche areas were outlined by the symbols of red pulsating areal while the paths of possible avalanches were depicted by the sparkling arrows. The researchers used moving borders and changing background colors to show the areas of heavy precipitations and disastrous flooding of plains.

Basing on the experience of the design of display films for the territory of Georgia these researchers created an animated sequence of the seasonal dynamics of the landscapes of the Earth. Theoretical base for such film has been the methods of spatial and temporal analysis and synthesis of the landscapes. Much interest aroused the display film showing the dynamics of the daily changes of the environment at the beginning of the monsoon period in the Eastern Asia.

Summarizing the experience in the design of display films N.L.Beruchashvili outlined some methods for spatial and temporal mapping: 1) creation of spatial and temporal models of the environment basing on the landscape maps; 2) compilation of the series of maps, atlases and slide-films; 3) design of cartographic movies and video films with discrete maps-frames; 4) compilation of the display films on PC using data bases of GIS; 5) usage of holograms (3).

Some experiments on cartometric processing of the display films were conducted as well. As an example we can introduce the design of animated sequence to analyze the changes of glacier surface altitude (12). Multitemporal digital terrain models of the glacier surface relief calculated from the results of processing of air and space images for the period of 30 years has been used as a data base. On the next step the belonging of every primitive of the image to the glacier surface was derived for the fixed time intervals and the velocity of changes of glacier surface altitude was calculated. The velocity was encoded by some color of blue-red scale, while the surrounding glacier areas were depicted in permanent neutral background color. Such animated sequence produced the imagination of continuous changes of surface altitude and borders of the glacier. Concurrently there has been conducted the cartometric analysis: the altitude, square, volume and depth of ice for the whole glacier and for specific altitude zones could be calculated for any time period.

Display film of north-western Pacific

The authors of this paper created animated sequences of temperature and temperature gradient changes for the surface layer of the water in north-western part of Pacific Ocean. They were created on automated system of image processing PERICOLOR-3000. We used as a data base the decade maps of the surface temperature for the interested part of the Pacific Ocean produced by the Japanese Meteorological Agency for February-July 1983. We processed this data using special software for linear interpolation and
calculated 15 multitemporal digital models of surface temperature distribution. As a result we obtained map-frames of temperature distribution for the period from 21 February to 20 July 1983. Each digital model allows to create map-frames of the surface temperature and gradients distribution and to obtain different parameters of statistics and probability.

North-western part of the Pacific Ocean near the shores of Japan and Kurily Islands is characterized by the presence of a strong hydrologic front. This front is produced as a result of the interaction of subarctic water of the cold current Oiyasio, warm subtropical water of Kurosio current, and deep water of the Pacific Ocean. This zone is often referred as "Subpolar Front" or "Kurosio Front".

Produced animated sequences allow to detect mesoscale peculiarities of surface temperature distribution and their spatial and temporal changes. Continuous animated sequence produced on the display at any speed much better characterizes spatial distribution and seasonal changes in surface temperature of the ocean comparing with static maps. The frontal zone can be deciphered most successfully while looking at the display film of temperature gradient, since its mean value equals $0.035^\circ C/km$ here exceeding by more than the order the mean longitude climatic gradient, which equals to $0.003^\circ C/km$ for the northern part of the Pacific Ocean. Frontal zone in characterized by the seasonal complex horizontal structure of the vortex type.

The designed display films of seasonal distribution and changes of the temperature show only part of the complex hydrodynamic process. Nevertheless the smoothed temperature distribution produces the background for a short-period hazardous processes.

**Types of dynamic geoimages**

Thus we see a variety of dynamic geoimages now. We mean images and map series, animation (cartographic, video and stereoscopic), movies (cartographic, computer graphic, etc.). Dynamic images can be 3D and 4D. The first ones are cartographic movies, animated films, slide films, computer animated films, multitemporal images, metachrone block diagrams, dynamic atlases. 4D images consist of stereo films, animated stereo films, hologram movies, dynamic block diagrams, dynamic holograms (7, 2). We can classify dynamic images according to the speed of frame change: slow, normal, middle-fast, fast, and high fast. Geoimages can be classified according to the methods of depicting: dynamic cardiograms, dynamic block diagrams, dynamic holograms, etc. We can find some cartographic products which one can hardly classify, since they are combining different approaches. Such products are the computer maps used for ship and plane navigation or for monitoring of spacecrafts. The application of the dynamic geoimages signifies a new level of spatial and temporal cartography. These new possibilities pose new problems in cartography and computer graphics: the development of dynamic symbols (special language of map films), temporal generalization, correlation of discreteness and continuity in processing of geoinformation, etc.

Special attention should be paid to perception and understanding of geoimages. American researchers C.Campbell and S.Egbert consider that it is necessary to conduct vast experiments to find the best solutions for the display of cartographic films (6). They consider to be successful the computer-animated sequence presented during the 1988 Winter Olympic Games in Calgary, Alberta, where the television viewer was aerially guided from the city of Calgary along a "flight" path into the mountains where athletic events were taking place. The scene moved smoothly and was easily interpreted. The same authors consider to be unsuccessful the sequence "L.A.: The Movie", a 3D animated terrain representation of the Los Angeles basin generated by the NASA's Jet Propulsion Laboratory. The sequence starts with logical bearing, azimuth, and verbal description, but soon deteriorates into a rapidly twisting
and turning fly-over, which results in almost complete disorientation for the
person unfamiliar with that landscape.

Application of the new methods of dynamic mapping require special
training for the students of geoscience faculties. Dynamic films distributed on
videotapes can be used in training classes in addition to traditional graphics
(maps, slides, etc.). The advanced training courses should involve the rapidly
developing trend in computing usually referred to as multimedia: the use of
video, audio, music, animation and other media to enhance the computer's
ability to communicate (6).

Dynamic geoimages are a very powerful tool for visualization of the
information. They are not replacing traditional static cartography, but adding
new capabilities to analyze the spatial and temporal relationships between the
objects and the processes. With the help of the computer the researcher can
manipulate and retrieve the dynamic sequence, slow or stop it, run it forward
or backward, change its parameters in real time. The researcher has the
opportunity to set the initial parameters and to extrapolate the dynamic data
in the future or to reconstruct situation in the past, i.e. to design forecasting
map films.

Conclusion

Wide application of geoimages in science and technology should be
based on GIS technology, which allows to generate a variety of dynamic,
overviewing and forecasting geoimages, to make cartometric and statistic
calculations, extrapolations and other transforms. Dynamic data bases and
complex dynamic GIS should be designed, which could take into account the
temporal changes of spatial information, the history and trends of changes.
Dynamic GIS should incorporate different variants of the system
development and should allow the transfer and intersection of data. There has
been done some experiments on the development of such systems of dynamic
map production for ship navigation (5, 11). Much attention should be paid to
the problems of spatial and temporal resolution of animated sequences, to the
design of dynamic thematic maps of different types, to the right solutions in
the employment of color and other tools for map compilation in the design of
animated sequences. Special field of interest concerns the problems of
dynamic generalization. Extensive experiments should be done to find the
optimum speeds of computer films display. A very important task should be
the right selection of the media for animated sequences distribution. In most
cases they can be distributed on diskettes or videotapes. Maybe in future the
leading journals in geography and cartography will distribute supplementary
diskettes and videocassettes with the animated sequences.

We consider, that the modern cartographic science and technology
should contain a special branch for design and application of the operator
controlled dynamic cartographic images, i.e. spatial and temporal models of
the surrounding world.

REFERENCES

1. Berlyant, A.M. Geoinformation cartography (in Russian). In:
14, Moscow, VINITI AN SSSR, 1991, 80-117.
2. Berlyant, A.M. Geoimages and Geoiconics (in Russian). Moscow,
3. Beruchashvili, N.L. Mapping the dynamics of the environment using
4. Bicmore, D. & Linton, P. Ocean and sea mapping. Ten year
experience of the Department of experimental cartography (in Russian).
The use of flow diagrams for describing new techniques in map production
S. van der Steen, A. Brown (Enschede, NL)
1. INTRODUCTION

In traditional map making, flow diagrams have been used for many years. The need for these flow diagrams is not under discussion. The use of such a description of production has proved its advantages over a long period time. It is common for many organisations to overview their map production by means of flow diagrams.

For digital map production, where computer facilities have partly or completely replaced manual map production tasks, a system like in conventional map production showing every detailed procedures has not been introduced so far. In manufacturers' brochures or in reports from mapping organisations an overview of their production is sometimes shown in a very general way with much emphasis on the hardware which is used. This kind of description is not detailed enough, however, for production planning, and has no use for discussions with clients and no use for personnel responsible for revision of maps and archiving of data. The existing descriptions often do not contain details like the creation of a new file, the manipulation of data or the conversion of map data formats to other formats.

This paper describes a very detailed system, developed at the ITC, in conjunction with the Commission on Map production Technology of the ICA and tested in some institutions in Europe. Newly created symbols and good compatibility with the symbolisation used in traditional map making flow diagram make the new flow diagrams a very useful instrument for staff at all levels dealing with map making aspects:

* Management staff responsible for:
  - production planning (staff and equipment)
  - time planning
  - control of continuation of the work flow
  - quality control
  - purchase/order of equipment/software facilities
  - education of staff
  - cost calculation/control

* Staff responsible for the actual production in order:
  - to know which steps to make
  - to take over certain steps in an other part of the production flow in case of absence
  - to communicate about actions and details
  - to carry out map revision

* Teaching staff responsible for the education of staff dealing with mapping and managerial tasks at all levels of the entire mapping process.

2. THE ESSENTIAL NEED TO PRODUCE FLOW DIAGRAMS FOR DIGITAL MAP PRODUCTION

The flow diagrams have proved their use already for many years. An example of such a system is provided by Ketael (Cartographic design and production) The International Institute for Aerospace Survey ITC in The Netherlands and Horsens College in Denmark, use flow diagrams in their map production lectures. Most of the systems based on symbols to represent materials (represented as boxes), equipment (specially created symbols) and actions.

In the production of maps or mapping materials using computer technology there is no real established procedure to show the production flow by means of symbols as used in conventional techniques so far. There are only very general descriptions showing mainly hardware with some information about software or parts
of software programs. Important matters are partly or completely lacking. A lot of detailed information has to be added in order to make a flow diagram of real practical use. That this shortcoming is very serious has been proved at the ITC, where staff and students sometimes had great difficulties to explain to each other the actual situations in the digital production of mapping materials. Details in the entire process were left out and there was no way to control the production other than by writing down on text description way. A flow diagram system as used in traditional map making was lacking. Other organisations have also reported problems which the use of detailed flow diagram would have alleviated. Examples include:
- High inefficiency of expensive equipment and labour leading to extremely high production costs
- Discussable results relating to quality of output resulting to bad quality products
- No control of continuation of production steps (bad management)
- Improperly instructed/educated staff for the demanding tasks
These needs have led a group of ITC staff to set up a system of flow diagrams for digital map production usable for staff engaged in production, management and education.

3. PRELIMINARY DISCUSSIONS

The first step was to make a detailed inventory of what should be contained in a flow diagram for digital map production. The next was to find if such a system already existed. A possible contender appeared in the form of a system by the ICA Commission on Map Production Technology. This system, which is based on equipment used, was adapted at ITC, to meet the perceived needs. The adapted system is shown in figure 1.

After several months of use serious defects were revealed in this proposal. More details were visible than the manufacturers' brochures, but not enough elements were available by far for accurate planning and for knowledge of exact production necessary for teachers and planners. Therefore the need for a new approach brought ITC staff together again for further discussion. The result was an approach similar to the already used flow diagram symbolisation for traditional map making.

With the newest flow diagram system a list of contents and or styles can be added to provide the user with more information on separate elements, because they cannot be specified in a flow diagram and still retain a legible overview.
4. THE ACTUAL PREPARATION FOR SYMBOLISATION IN FLOW DIAGRAMS USING NEW TECHNIQUES

The discussions and conclusions of groups of users within the ITC, some institutes in Europe and a working group of the Map Production Technology Commission of the ICA resulted to the introduction of flow diagrams consisting of the following main groups of symbols:

A. The files creation/file editing and template group
This group of symbols include all the important aspects of computerised mapping in which file preparation are involved. These are:

- creating new files
How a box should be filled with logical symbols, text and numbers (if necessary)

- actions within files such as editing or geometric transformation

- preparation of templates, sometimes called ‘styles’ or ‘windows’ for pre-determined standard symbols, layouts, etc. (Somewhat smaller than the file symbols)

The above listed group is represented by fairly large symbols. These symbols can be compared to material symbols for films, paper and printing plates as used in conventional map making flow diagrams.

B. The equipment and data external storage group

- The only equipment separately symbolised are output devices

- External storage of digital files is represented by circle, made small if there is no need for it to contain text

C. The group of actions such as digitizing, scanning, manipulation, conversion and keyboard input

These symbols are always used in combination with a process line. The symbols only vary by letter and are all of the same size. They can be compared to the reproduction symbols used in flow diagrams for conventional map production.
- Digitising. An addition of a name indicates the kind of digitiser which is used (automatic or manual).

- Keyboard input

- Scanning. The type of scanner can be mentioned by name.

- File manipulation. Any action resulting in a new file including the creation of a data compressed file. (Within one application program)

- Conversion is the creation of a new file created from an existing file in another software program.

D. The group of process lines, cable connections and reference lines

Single or double lines with optional arrows link files, output devices and external storage.

- Process lines give the sequence of activities. They may split, or cross without joining

- A cable connection permits direct data flow

- Reference lines show references to existing files or hard copies. An example is text placement with references to lines and symbols. Reference lines go from what is made first to what is made later

E. The group of software/hardware environment and batch file operations.

- Quite often more than one application program must be used. It frequently happens that a file has to be converted to achieve the required results. A particular combination of computer hardware and application software is called environment. In the flowdiagram this is represented by a heavy dashed line which encloses all files used and activities done within that environment.
If the hardware does not change, but different software has to be used, environments envelopes can be nested or connected. The hardware and software names are included within the environment envelopes. This symbol can also be used to enclose GIS and digital image processing operations.

- It can happen that a particular set of operations which is used repetitively is produced as a batch file. A batch file like this can be named and shown as an envelope nested in the general environment. In this case, for clarity, a thinner dashed line can be used for envelope.

F. The addition of lists of contents and styles

The flow diagram above described is an extremely helpful instrument for managers and users. It describes any digital mapping project and should be constructed in advance, so that an overview of all the mapping process is available. However, no specifications and/or attributes can be defined in the symbolisation of an entire production process. If this were to be done the flow diagram would become an unclear lump of information without a proper overview and overwhelmed by details. It would then not be the tool required for users, managers, planners and teachers.

To meet this problem an additional list of contents, attributes, styles etc. is produced as a separate element. This can be compared to the flow diagram for traditional map making. A list with information about masks, peeled areas, colours and other information is also added to this kind of flow diagram.

<table>
<thead>
<tr>
<th>Features</th>
<th>Levels</th>
<th>Line type</th>
<th>Line gauge</th>
<th>Colour</th>
<th>Patterns</th>
<th>Other information/remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point - churches</td>
<td>10</td>
<td>single</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line - Roads</td>
<td>20</td>
<td>double</td>
<td>4</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Built up - Town c.</td>
<td>35</td>
<td></td>
<td>3</td>
<td></td>
<td>20%</td>
<td></td>
</tr>
</tbody>
</table>

A table of contents as shown is a kind of matrix which shows the following information:

- Symbolised point, line and area features such as churches, roads, built up areas.
- Levels (also called layers)
- Coded line type like dashed lines or double lines
- Gauges, sometimes called weights
- Colours in combination with percentages values, coded by name, numbers or figures
- Patterns like a line or forest pattern
- Possible remarks
Of course the list can be extended to any detail which is useful for the organisation producing maps or mapping products.

In some applications the user must preconstruct all the map details as a kind of library, also called a template. Map making management should have a list of all existing map contents available for single maps or map series in order to have the exact and accurate styles. If properly used the library of styles contains all the information as described in the list of contents.

5. THE CONSTRUCTION OF FLOW DIAGRAMS

Flow diagrams for conventional and digital map production can be produced manually, aided by drawing templates. It is also possible to construct them using almost any graphic application software. It is advisable to create a separate symbol library in order to avoid repetition of work any time a flow diagram has to be produced. This library must be an independently produced file which can be recalled for any newly created or existing file.

The library available for the proposed system is depicted in the figure underhere.
Production of statistical maps from GIS data files, using Carto/graphix software.

(Note: same basic method for each map)

This flow diagram shows a production process of a statistical map using GIS data and applying Carto/graphix software.
A cartographic map production process showing a combination of traditional and digital techniques.
Cartographic map production with the use of ILWIS and Aldus Freehand software.
6. CONCLUSION

The flow diagram system for digital map production described above has been in use at ITC for less than a year. It is not possible to decide whether it is the ideal solution. More experience. However, the existing basic idea has already proved to be a very strong basis for map production in several very different environments. In particular, ITC staff and students have been able to make a smooth transition from the conventional map production flow diagram to the new system for digital production. The system has been used for a variety of graphic, cartographic, GIS and digital image processing application software on Apple Macintosh, PC-AT and Intergraph InterPro platform. We are already convinced that this system, perhaps with modifications suggested by users, will prove very useful for managers, production staff, manufacturers, teachers and students.

The ICA Commission on Map Production Technology has undertaken to coordinate the responses and suggestions of users of this system. Remarks can be submitted to any member of this commission, or directly to the chairman of the Flow Diagram working group of this Commission.

(Sjef van der Steen, P.O. Box 6, 7500 AA Enschede, The Netherlands)

With this paper a strong support has been stated to create flow diagrams for any production task in digital environment as well as with conventional techniques. Of course this only makes sense prior to the actual execution of map production.
Hybrid cartographic data processing up-to-date techniques to accomplish official tasks

E. Jäger (Hannover, D)

Summary:
The topographic cartography is approaching to replace most of the manual processes by digital means. This change is involved in the strong turn to the development and usage of geographic information systems (GIS). The paper demonstrates different examples of how an official agency like the State Survey Department of the Land Lower Saxony applies new methods of hybrid (vector and raster) data processing to meet all upcoming requirements and to take advantage of the new technologies.

Especially, the cartographic pattern recognition of scanned cadastral maps 1:1000, the hybrid updating of the Topographic Map 1:25000 (TK 25), and the establishment and use of raster data archives of topographic map sheets will be discussed. Finally, some future developments in the field of cartographic data processing and its consequences for official cartographic institutions are pointed out.

1 Introduction

In the Federal Republic of Germany the official surveying and mapping activities are tasks of the Länder. The Länder are bounded by law to provide cartographic data (in paper and digital form) for economic, administrative, and legal purposes. The demand for cartographic data is more and more changing from paper to digital form. But, changes in customer's behaviour cause fundamental changes in the way of production and handling of cartographic data within official services.

Digital Data are to be used in two forms:
   a) as primary models of the landscape in ground related information systems for the purposes of analysis and forecast and
   b) as secondary models of the landscape in graphic related form for the purposes of visualization and planning.

To avoid uneconomical multiple data capturing it is one of the most important tasks of the official surveying and mapping agencies to satisfy this immense demand for digital data within a short time. Therefore, in Germany the Digital Cadastral Map (ALK) is established until the year 2005 to meet all requirements in the range of large scales (1:1000). The medium and small scales will be covered by the ATKIS-data (Authoritative Topographic-Cartographic Information System). ATKIS-DLM-(Digital Landscape Model)-data 1:25000 will be available for the whole country in a first completion step in 1995 (DLM25/1).

The derivation of Digital Cartographic Models (DKM) out of DLM-data - as planned in the ATKIS project - is not yet realized, as algorithms for the computer-assisted data preparation and cartographic generalization do not work proper until
now. But, great efforts are invested to automate this process too. By substitution, to get digital data in graphical form rapidly, it is possible to scan the required data from analogous map sheets. The result of the scanning process are raster data which may be used in different ways.

The most important advantages of the raster format are the
- fast data capture process by scanners and
- fast data output on raster devices.

The combination of these advantages with well working vector based digital processes and algorithms increases the economy of computer-assisted manufacturing methods to an optimum. So, the combined use of vector and raster data - the so called hybrid data processing - is one of the opportunities to be used to accomplish official tasks.

2 Automatic digitization and pattern recognition of cadastral maps 1:1000

2.1 General remarks

The first example relates to the process of data capturing. It is investigated to use the automatic digitization of cadastral maps 1:1000 - beside the manual vector digitization - to build up the ALK (Digital Cadastral Map) more rapidly. The automatic digitization process includes the scanning of maps, the raster-to-vector-conversion, the automatic pattern recognition, and the interactive correction and completion of the data.

The standard procedure to digitize cadastral maps is the manual table digitizing. The time estimation to complete all conversion steps until 2005 is based on this standard procedure. To realize the ALK project within a shorter time period the applicability of the automatic digitization was investigated. This investigation in the State Survey of the Land Lower Saxony should answer the following questions:

- What is the recognition rate of the automatic process regarding the two main features of the cadastral map - buildings and boundaries of the land register?
- Does interactive correction and completion of the automatically derived data save time compared to the manually digitization?
- What graphic quality must cadastral maps have to make the automatic digitization an economic alternative?

To answer these questions the State Survey has tested the program CAROLA (Computer-Assisted Recognition of Linear and Areal Features in Maps) of the Institute of Cartography, Hannover University /Illert 1991a/.

2.2 Pre-processing

Basis of the automatic digitization and cartographic pattern recognition is the scanned data of the cadastral map sheets 1:1000. The pixel resolution is 200 lines/cm (= 500 dpi). First of all the grayscale raster data has to be binarized in black/white pixels by a simple threshold operation. After the erosion of isolated black or white pixels ("scan noise") the raster-to-vector-conversion is performed as the recognition of map features is a vector based process.

The raster-to-vector-conversion is carried out in the following steps /Jäger 1992a/:

- calculation of the distance-matrix (for each line (black) pixel the shortest distance to the background (white pixel) is calculated); the distance-matrix contains information about the thickness of map elements.
- execution of the topological skeletonizing; each map element is thinned
from the outline to the center by using the distance-matrix (at first all distances "1" are examined to their topological importance, then all "2"-distances ...). The topological skeletonizing produces new pixel values:

- pixel value "1" indicates → beginning of a line (black pixel with only one black neighbour pixel).
- pixel value "2" indicates → element of a line (black pixel with two black neighbour pixels).
- pixel value "3" indicates → node of three or more lines (black pixel with three or more black neighbours).

- vectorizing: extraction of nodes and line following process; based on the line skeleton the extraction of nodes and beginnings of lines is performed. The raster coordinates of nodes and beginnings of lines are transformed into the geodetic X,Y-coordinate system. Each pixel between nodes or nodes and beginning of lines is followed automatical. For each change in direction inside the followed line skeleton a X,Y-coordinate calculation is realized.
- line smoothing; the raster-to-vector-conversion causes zigzag lines which have to be smoothed. The Douglas-Peucker-Algorithm is a suitable method to smooth these lines.
- improvement of geometry; at last small gaps in lines have to be closed, typical stubbles have to be erased and the positions of close nodes have to be centered.

The result of the pre-processing steps is an unstructured vector data set with topological connected files of nodes and arcs.

2.3 Cartographic pattern recognition

Based on the files of nodes and arcs the automatic pattern recognition algorithm is started. Therefore, it is extremely important to build segments within the data set. Segments are: a) very long networks of lines which are candidates for boundaries and b) short networks of lines with short dimensions in X and Y which may be candidates for text and symbols (Fig. 1).

Now, these candidates are examined with numerical and syntactical algorithms to assign them to a feature class. Statistical insecure candidates are not assigned; they have to be structured by interactive

![Diagram](image)

**Fig. 1:** Strategy for pattern recognition in large scale maps /Illeit 1991b/
means. In detail, the automatic pattern recognition process for buildings and boundaries works as following:

- segmentation and erasure of text, numbers and symbols,
- recognition of buildings (of their contours) by detecting their hachures (three or more small parallel lines within a contour line),
- recognition of small circles which represent the boundary stones in the German cadastral maps,
- line following process between recognized boundary stones,
- calculation of the center of the boundary stones (X,Y-coordinate),
- improvement of geometry of all recognized buildings (right angles),
- storage of all recognized features in different files (building-file and boundary-file) and storage of the non-recognized features.

2.4 Post-processing

The resulting files of the automatic cartographic pattern recognition process are transferred to a hybrid workstation where they are corrected and completed interactively. On the background of the original scanned raster data the recognized buildings and boundaries are presented on the graphic screen in different colors. So, it's easy for the operator to identify incorrect or missing vector elements and to digitize all necessary objects to complete the contents of the cadastral map.

2.5 Results of the investigations

The automatic cartographic pattern recognition is not able to structure a cadastral map completely and without faulty classifications. But, nevertheless this method is able to save time compared to the manually digitization. Based on four cadastral maps the following recognition rates have been noticed (Table 1):

<table>
<thead>
<tr>
<th>recognized correctly</th>
<th>mistakenly</th>
<th>not recognized</th>
</tr>
</thead>
<tbody>
<tr>
<td>boundary stones (circles)</td>
<td>82 %</td>
<td>1 %</td>
</tr>
<tr>
<td>boundaries</td>
<td>85 %</td>
<td>5 %</td>
</tr>
<tr>
<td>buildings</td>
<td>99 %</td>
<td>25 % **</td>
</tr>
</tbody>
</table>

*) Many circles did have gaps because of the small line width.
**) 25 % of the recognized buildings have false contours (see Fig. 2).

Table 1: Recognition rates of the automatic pattern recognition process

---

Fig. 2: Typical result of the automatic pattern recognition process
Compared to the time for the manually table digitizing this presented method saves 30% up to 60% of operators work - depending on the contents of a certain map and its graphic quality. Even if the graphic quality of a map is bad, this method is favourable because of the opportunity of the on-screen-digitizing which is comfortable in accuracy and ergonomics. The hybrid on-screen-digitizing process takes advantage from the accuracy of scanning and avoids the inaccuracy of a table digitizer. Through an online-plotting of the captured vectors on top of original raster data background, incorrect vector points or lines are to be detected and corrected immediately. Summing up, it may be said that the pattern recognition is a suitable method that will be used with economic benefits to establish the ALK-project in Lower Saxony.

3 Hybrid updating of the Topographic Map 1:25000 (TK 25)

In 1989, the State Survey of Lower Saxony and the Siemens Nixdorf Ltd. (SNI) started a cooperation to produce a computer-assisted hybrid system to update the German topographic map series. This system - the SICAD-MAP-REVISOR - has been developed in 1989, tested in 1990, and is used in production to revise the TK 25 since the beginning of 1991.

3.1 Basic conception

The hybrid method is based on scanning the 7 original out-of-date foils of the TK 25 (planimetry, woodland mask, water mask, water lines, contours, lettering "black", and lettering "blue") and on scanning the manually produced compilation sheet (CS 25) on a Hell scanner CTX 330 (see Fig. 3). The scan resolution is 320 lines/cm (812 dpi). The CS 25 contains new features and features to be erased in different colors; both colors are stored in separated files.

All scanned foils are binarized, run-length-encoded, and geometrically rectified. Each foil contents is assigned to one specific bit of a 16-bit computer word. So, the whole pixel information may be presented on the hybrid color-workstation (Siemens WS 2000), and by using the related look-up-table the digital image looks like the analogous TK 25 - extended by two further colors (red: new features; yellow: features to be erased).

On this cartographic workstation the cartographer revises the TK 25 interactively by erasing parts of the raster information and by symbolizing new map elements in vector format. A vector-to-raster conversion combines the new map data with the old raster background. When a whole map is worked out, four TK 25 plot-files (black, green, blue, brown) are

Fig. 3: Computer-assisted process of updating the TK 25
created for the output on the Hell laserplotter CTX 502. These films are used to produce the printing plates for offset printing of the new TK 25. In an additional step the revised TK 25 foils are stored on optical disks (WORM - Write Once Read Many) to build up a raster data archive 1:25000 (see chapter 4).

3.2 Details about the hybrid updating process

To update a whole map sheet parts of maximally 10 cm * 15 cm (3200 * 4800 pixels) are building the actually working set in an interactive session. The red and yellow elements of the scanned CS 25 are graphically non-perfect as they are drawn manually without care. The cartographically highly skilled operator sets one bit level of the 16-bit computer word active (e.g. the "black" planimetric foil). Then, he works out all revising tasks accurately by using the raster editor functions to model the raster background and the vector symbolization software to create new map elements. Some important functions are:

- interaction by using procedures, table- and screen-menues,
- rubberbanding,
- erasio and creation of single pixels and pixel areas,
- paste and copy,
- digitising of axis lines to place linear symbols,
- undo-function in case of incorrectly performed processes,
- zooming and panning.

All new symbols are created according to the TK 25 map symbology. For this reason, about 200 different line symbols and about 80 different point symbols have been generated and are stored in the symbol library. The vector symbols are presented by overlaying the raster background. In this way, one foil of the working set after the other is revised - with the exception of the two lettering foils which are still worked out in the conventional way. This procedure continues until the last working set of a TK 25 has been completed.

Fig. 4 Hybridly updated TK25-cutout

A digitally updated TK25-cutout is demonstrated in Fig. 4; Fig 5. contains those map features which are generated by the described method and which have been added to the old and revised raster background.
3.3 Experiences and benefits

At the end of 1991 the first TK 25 that has been updated by use of the hybrid SICAD-MAP-REVISOR was completed, printed, and published. Compared to conventionally updated maps no differences are recognizable. Up to now (January 1993), more than 30 TK 25 map sheets have been updated on four hybrid cartographic workstations WS 2000.

Although some important improvements of the hard- and software are desirable the goal - to update 25 % of the TK 25 digitally - has been reached. All the involved cartographers are working on "their" digital map four hours daily. In the remaining time they work conventionally. The time period from the scanning up to the printing takes about 9 months on average. This is about 30 % shorter than the conventional process requires.

Some other essential advantages of the hybrid method are:
- the "analogeous" cartographers get used to digital techniques (this is more and more important by looking to the rising GIS era),
- economizing of environmental critical chemical reproduction processes,
- optimizing of the digital layer (foil) method,
- standardization of the working results as shapes and graphical covering of the symbols are always constant,
- dimensional stability of the digital map and
- release of physical exhausting manual work (bended carriage).

In a future phase, the SICAD-MAP-REVISOR should be modified to derive Digital Cartographic Models (DKM) in vector or raster format out of vector-based Digital Landscape Models (DLM), which are established momentary. This modification will contain rules for an automatic establishment of cartographic objects, rules for graphical priority of map features, automatic interpretation of symbol catalogues, and the computer-assisted consideration of the cartographic generalization /JÄGER 1992b/. So, the revising task of the system will more and more change to a map design task within this hybrid system.

4 Raster data archives of topographic maps

4.1 General conditions

Each TK 25 map sheet that has been updated with the SICAD-MAP-REVISOR is not only prepared for the offset printing, but also to be entered into the raster data archive of the TK 25 (see Fig. 3). The following conditions have to be noted to establish a raster data archive:
- homogeneous data capture (uniform maps and adequate uniform pixel resolution),
- uniform coordinate system,
- secure and long-term storage and
- short-term retrieval of the data.

The adequate coordinate system in Germany is the 3°-Gauß-Krüger-System for the topographic scales 1:25000, 1:50000, and 1:100000 (TK25, TK50, TK100). Although this system is not free from distortions the single map sheet fits to another without gaps.

As optical disks are very secure storage mediums they are used for the raster data archive in Lower Saxony. The WORM (Write Once Read Many) technology makes it possible to store large data amounts (2 GByte) on one 12"-disk with a security of 30 years. Another suitable solution would be the insertion of erasable magneto optical disks. For a short-time retrieval of the raster data a good organization of
the data in sub-files is important. At the State Survey of Lower Saxony one sub-file is made of one Topographic Map (about 30 MByte). So, 50 up to 60 TK's may build a data block on one WORM-disk.

Basically, a map may be scanned and inserted into the archive in one of the following manners:

- all original foils of a map may be scanned,
- all printing foils are scanned or
- the printed color-maps are scanned (including color separation).

To have a homogeneous raster data set all map sheets of a topographic map series should be scanned in the same way.

### 4.2 Establishment of raster data archives in Lower Saxony

The establishment of raster data archives of topographic map series is an additional new task within an official surveying and mapping agency. From the economical point of view this new task has to be cost-covering and useful for own purposes.

Because of the utilization for the hybrid map updating process (chapter 3) all original foils of the TK25 are scanned with a pixel resolution of 320 lines/cm. The scanning process of all seven foils takes about 5 hours on the drum-scanner Hell CTX 330. All digital foils of a map sheet are rectified to their specified size, merged together (one computer-word for each pixel), visually checked on a graphic color screen, and added into the archive.

At the beginning of 1993, about 20% of more than 430 map sheets of the TK25 are ready. A total completion of the TK25 raster data archive is planned for the end of 1994. If any external user needs the whole raster information within a shorter duration there are two possibilities:

- scanning of the printed color-maps or
- scanning of all foils has to be done with a sheet-feed scanner, which is much faster but not as accurate as drum-scanning.

To satisfy the demand of many users for raster data of the TK100 within a short time all about 35 printed color-maps of Lower Saxony have been scanned with the Hell scanner CTX 330 with a 200 lines/cm resolution. The CTX 330 is able to recognize up to 15 colors per scan-process by splitting the remitted light signal in red, green, and blue portions and by relating these signals to specified colors. So, the TK100 is separated in 7 colors (digital foils):

- black: planimetry and lettering (black),
- blue: water mask, water lines, lettering (blue),
- brown: contours,
- light-green: woodland mask,
- dark-green: woodland outlines,
- yellow: street filler and
- red: street filler.

Although some disadvantages appear - the lettering is not on separated files, the woodland mask file contains other elements in negative form - this TK100 archive data has a high acceptance because of the good quality when all colors are displayed.

Some Länder have completed establishing a TK50 raster data archive and have sold these data to a lot of customers. So, the State Survey of Lower Saxony will follow these examples. Within half a year, starting at the beginning of 1993, all about 120 printed TK50 maps will be color-scanned.

The minimum pixel resolution of scanned topographic maps should be 400 dpi.
(about 160 lines/cm). Beyond this level map details would melt into one another. On the other hand a lot of color output devices (e.g. ink jet and electrostatic plotters) are able to reproduce this resolution.

4.3 Utilization of topographic raster data

Topographic raster data may be used in manifold ways. Within the agency the TK25 raster data will be used above all for the hybrid updating (see chap. 3). Another application is the combination of TK50 raster data with vector information to build up a digital archive of centers of aerial photographs for the purposes of inquiries.

For external users raster data of the related scales from 1:25000 up to 1:100000 are the suitable topographic background for their thematic maps. By designing such maps digitally these raster data are put in the background of the screen of a hybrid workstation. On top of this background graphical vector symbolization processes may be done. The raster archive data can be delivered to customers in different forms:

- single bit levels (digital foils) in binary form (1 bit/pixel),
- logically merged digital foils to one binary data set (1 bit/pixel) or
- merged digital foils (1 or 2 byte/pixel) for colored representation.

Moreover the delivery of raster data is not limited to the original scan resolution. In the case of the 320 lines/cm of the TK25 raster data users who work on PC's or middle range workstations may get this data transferred into a lower resolution without too much waste of information. An example in Fig. 6 illustrates this:

![Fig 6: TK25 raster data in different pixel resolutions (8-times enlarged)](image)

On the left side a TK25 cutout scanned with 320 lines/cm is shown, on the right side the same cutout with 200 lines/cm scan resolution. In the middle the result of the digital reduction from 320 to 200 lines/cm is demonstrated. Certainly the resolution of the left cutout is evidently better. But it can be realized that there are almost no differences between the calculated and the scanned 200 lines/cm cutouts. So, no special scan process is necessary, if a customer wants to have data with a lower resolution than is available from the raster data archive.
5 Future developments

As it has been suggested in chapter 3.3 the derivation of Digital Cartographic Models (DKM) out of Digital Landscape Models (DLM) is one of the challenges for the next few years. The DKM is the basis to derive digital topographic maps by using symbolization processes in vector or raster mode. The DLM 25/1 which is established momentary does not contain the whole topographic information of the wanted DKM 1:25000 to generate a digital TK25 completely. E.g. buildings, a lot of lettering, and drains are not captured in the DLM 25/1. But these lacking data might be added from the existing raster data archive 1:25000. For a transitional period this method which still has to be investigated could be useful to produce computer-assisted up-to-date topographic maps. The precondition for this are tools for an automated transition of DLM-objects to DKM-objects including a computer-assisted solution of the cartographic generalization process. Hybrid workstations like the SICAD-MAP-REVISOR are necessary to master this problem.

Once DLM’s contain all objects to derive a complete DKM and digital map no scanned (old) raster data is necessary any longer. This moment – approximately in 1996 – will be used to change the map graphics of the German Topographic Map Series. Main aspects of the new map design are:

- increased use of colors – as well for the lettering,
- use of one single lettering type (univers),
- insertion of street names (for main streets only) and
- enhancement of graphical dimensions of map features.

The goal of the recent developments is to derive maps fully digital (automated and computer-assisted) from any data base with any map symbology in any scale you want. The consequences for official cartographic institutions are to be armed against these development by:

- using digital (hybrid) techniques today,
- observing and influencing recent trends and
- early planning of staff training and the necessary equipment.

6 References


Scan-vectorisation as a means for the acquisition of digital data for purposes of a computer based planning cadastre

H.F. Kern (Karlsruhe, D)

Geographic Information Systems (GIS) are already widely available and the Query Languages for retrieving information by highly sophisticated questions are steadily improving. The acceptance of such systems is growing and of no problem any longer. The real bottleneck is to fill the GIS with information, geometric as well as alphanumeric. In traditional cartographic offices and enterprises there is a vast amount of colour separated, quasi layer oriented, analogue "sheets".

It is a challenging problem to convert these analogue information into a digital form. Whereas the raster oriented approach can be solved quite easily, the vectorisation of analogue maps is not yet an accepted and broadly used procedure.

For different cartographic application areas the paper will document the acquired experience:

- Character recognition of scanned name and place layers,
- Vectorisation of hydrographic sheets,
- Vectorisation of contour lines,
- Vectorisation of planning related layers.

Typical problems and their possible solutions will be dealt with:

- Compound characters, relation of number of erroneous characters to number of words,
- Closing of gaps in contour lines,
- Computational avoidance of artificial nodes, segments and meshes,
- Process of area generation,
- Attribution of nodes, segments and meshes,
- Transformation from table coordinates into Gauss-Krüger,
- Exact clipping at sheet borders and continuation of graphic elements over sheet borders.

The results of scan-vectorisation will be compared with manual digitization with respect to geometric precision, ease of attribution and necessary resources.
1. Introduction

In the near future two main tasks will receive attention in countries with extensive cartographic coverage: the current process of up-dating the contents ("continuous revision") of existing maps and the smooth transition to cartographic information systems. A successful and efficient solution for both tasks requires the integration of highly advanced computer techniques.

The field of cartography can contribute substantially to the efforts being undertaken in establishing national GIS data bases. Maps contain not only an immense variety of information on our environment (contents and geometry generalized), but often also show extensive coverage of an area. Usually they are available immediately and at a low price, thus being a valuable and economic source for spatial information systems.

Whereas the digitization of maps has been solved long ago through the application of scanner technology, difficulties are encountered in interpreting the raster data as map contents. That means, structuring of map details into logical features still poses a lot of problems. Mass digitization of topographic maps of average and small scales fails largely because of the lack of automatic acquisition methods.

Even though, the complete automatic structuring of the entire map contents seems to be rather unrealistic at the moment, it would be definitely worth finding partial solutions to particular problems, however only if theoretical concepts can be integrated into cartographic processes.

2. Pattern recognition in the process of map revision

Despite of all developments in the field of digital cartography, traditional cartographic products will be important for a long time yet. Up-dated maps complying with the needs of users do not always have to be immediately and completely transformed into digital maps. Existing map images can be up-dated and completed by using raster techniques, or they may be used for other purposes (e.g. visualization).

The quality of these map images depends on the number of photographic copies that an original goes through in the up-dating process. Scanning a map can actually only freeze its quality. The aging process of graphic elements such as lettering, lines and symbols cannot be reversed without tedious manual correction. In Switzerland there are several 1:50 000 maps with poor graphic elements which should be renewed.

Pattern recognition is an economically interesting method of solving this problem. The original image is scanned, the elements which are to be renewed are identified and automatically replaced with new symbols from a special font. This is adequate to a digital copy, whereas the copy is of a better quality than the original. One advantage of pattern recognition is the possibility of using a special font that can be renewed.
recognition is its success even if only a part of the searched symbols (lettering or a particular type of line) is recognized. The new image of the symbol is better than the original and, with sufficient resolution, the rest of the image remains practically unchanged.

An indirect result of this method is the possibility of using the identified symbols for other applications. Identified objects can be stored, for example in a data base of geographic names, or identified elements can be separated from the rest of the image to form a new data layer.

Fig. 1: Pattern Recognition and Map Revision

These additional applications all lead to the automatic acquisition for GIS, an application which is very much in demand today. The presented results are limited to only a few map elements and are therefore to be considered as examples. A fully automatic process from a paper-image to a GIS (Leica-Infocam) has been successfully tested for geographic names and symbols.

The paper describes a joint-study-project of the Swiss Federal Institute of Technology in Zurich and the Swiss Federal Office of Topography in Wabern which, after its first stage of completion, leads to some remarkable results.
3. Review of cartographic pattern recognition

The digitization of conventional map images comprises two parts: the geometry of the map contents has to be scanned (keeping the loss of accuracy as low as possible) followed by structuring map details and contents into logical features. Automation of the second part requires the efficient integration of pattern recognition methods.

Following specifications must be considered when using pattern recognition for cartographic applications:

• The goal is to recognize and classify letters, numbers, symbols, lines and areas in view of the special fact, that these elements are often not completely isolated but are frequently juxtaposed, they cross each other or overlap themselves. Without context, their logical interpretation is sometimes rather difficult.

• The information is graphically coded according to a strict cartographic system of rules. The standardization of the map symbols was due to good visual identification. The variations in the individual shape of particular symbols are limited.

• According to this deterministic character, classical pattern recognition methods seem to be well-suited (e.g. template matching, geometric-statistical classification etc.).

The software products developed up to now correspond to these specifications and can be appraised as follows:

• Almost all known and available systems are limited to the use of line-oriented images such as cadastral maps, street-maps, the Deutsche Grundkarte (DGK5) etc. with the exception of the work being done on small scale topographic maps at the IfAG, Frankfurt a. Main.

• A recognition rate of 80% or a time reduction to 60% can already be considered as economic.

• The spectrum of the applied methods and procedures is very wide. Each kind of recognition problem requires the programming of specific strategies and algorithms. However, all systems are based on a procedural method allowing specific combinations of modules, depending on the respective problems.

• As soon as particular sequences of steps (or parts thereof) have proved worthwhile, these individual procedures can run as batch jobs, thus increasing the degree of automation with a certain degree of flexibility.

4. Identification in the raster image

It is usually assumed that a vectorization process is performed prior to the actual recognition procedure when applying pattern recognition methods for technical and cartographic purposes (exception IfAG). This method definitely seems to be effective - considering an efficient data-reduction - for line-oriented plans and maps.

There are nevertheless several reasons to question the compulsory raster-vector conversion for structuring (topographic) maps:

• Maps are explanatory, two-dimensional representations of reality. The line feature is neither in reality nor in its cartographic representation the dominant element. Also the visual perception of human beings is without doubt more area-oriented than line-oriented.

• The process of pattern recognition is not carried out with the original, scanned data. According to communication-theories a loss of information and distortions must be taken into account each time something is communicated.

• A raster-vector conversion is often problematic, especially where there is a high density of information (juxtaposition, overlaps etc.) or in case of images with poor graphic quality (interruptions, faulty areas etc.). The automatic choice of fiducial points leads to distortions of the geometric structure (unsmooth straight lines, rounded corners etc.), thus making post-processing inevitable. The choice of adequate parameters for vectorization requires a lot of experiences in digital cartography.

• Availability of powerful hardware resources

• Availability of software modules (raster editors, algorithms for digital image processing)
Raster data are particularly well-suited for visualization. A high-quality cartographic reproduction is possible only by use of raster data.

The work in raster-oriented pattern recognition carried out at the IfAG can be considered as quite successful. The methods are based on simple operations (e.g. thinning, spreading, averaging areas and nodal points, etc.) which can be performed with the basic software on their raster data processing unit Scitex R280. Transferring these non-portable solutions onto another hardware platform is out of question. The capacity of modern workstations with RISC processors, however, makes cartographic pattern recognition in the raster mode a worthwhile and realistic objective.

5. Possibilities to optimize template matching algorithms

The most traditional form of raster-oriented pattern recognition is the reliable method of template matching, where the candidates are matched with the template pixel by pixel. As soon as the defined degree of similarity is reached, an element is recognized.

The basic aspects of this method can be described as follows:
- the algorithm is actually very simple
- the number of elements to be searched should be not too large
- variations within a classified group of objects should be held relatively low
- the method is robust and yields a high rate of recognition
- the large number of operations required by comparing corresponding pixels usually leads to enormous, almost intolerable processing times.

Therefore, the potential for optimizing this method is to be found primarily in accelerating the process. The following principles are founded on two basic ideas: first of all, the number of possible candidates should be rigorously reduced in the sense of progressive data reduction, and secondly, the individual operations should be arranged according to a strict hierarchy, which is typical of many algorithms in digital image processing.

Step 1
The problem of matching for cartographic applications should be reduced to binary raster images (background or line elements) because the individual colors can either be scanned separately or be easily computed with binary transformations. Matching-operations can then be implemented bit by bit very quickly with logical operations such as AND, OR, etc.

Step 2
Fig. 2 and Fig. 3 show the mathematical definition of similarity: A standard method for matching in digital image processing is based on the Euclidean Distance which assumes the value of zero in the case of an exact match. The minimum condition of the Euclidean Distance is adequate to the maximum value of the Cross-Correlation function between the raster image and the template. In comparison to Cross-Correlation a more less complex, translation-variable measuring of similarity can be the Laplace-Distance.

Note: The computation of the Laplace-Distance as a measure of similarity always starts under the assumption that there is an exact match (initial value for the distance = 0!). The value of the distance increases when corresponding pixels do not match. Since a particular feature usually does not occur very often per map image, the tolerable distance is reached rapidly and the computation can be terminated. The fact that a "non-match" is the actual value being measured speeds up the whole process.

Step 3
A massive acceleration is reached by the trivial fact, that an object never can be found in areas consisting exclusively of background pixels. Therefore, computing of similarity only makes sense for feature-colored pixels. Computing the Laplace-Distance, however, is somewhat more time-consuming (see Fig. 3 and L' in Fig. 7).
A standard method to measure similarity at \((X,Y)\) of Template and Image:

Template \( \{ T(z, s) \mid 1 \leq z \leq m, 1 \leq s \leq n \} \)

Image \( \{ B(x, y) \mid 1 \leq x \leq M, 1 \leq y \leq N \} \)

\[
d^2(X,Y) = \sum_{z=1}^{m} \sum_{s=1}^{n} \left( B(X+z, Y+s) - T(z, s) \right)^2 \quad \text{("Euclidean Distance")}
\]

\[
d(X,Y) = \sum_{z=1}^{m} \sum_{s=1}^{n} \left( B(X+z, Y+s) - 2 \cdot B(X+z, Y+s) \cdot T(z, s) + T(z, s)^2 \right)
\]

- \(d=0\) for an exact match, otherwise \(d>0\)
- \(d \rightarrow \text{min.}\) searching a minimum value
- A simpler measure of match instead of the square sum

\[
\text{"Cross - Correlation"} \quad \sum_{z=1}^{m} \sum_{s=1}^{n} \{ B(X+z, Y+s) \cdot T(z, s) \} \rightarrow \text{Max.}
\]

\[
\text{"Laplace - Distance"} \quad L(X,Y) = \sum_{z=1}^{m} \sum_{s=1}^{n} \left| B(X+z, Y+s) - T(z, s) \right|
\]

Fig. 2: Pattern Classification by Template Matching

Fig. 3: Template-Definition

Fig. 4: "Hierarchical" Template

**Step 4**

Analogous to the hierarchical structure of the modules, the matching itself should also be carried out step by step. Of great importance is the fact that the probability to find a feature-colored pixel in the nucleus of an object is surely larger than in the margins (see "Hierarchical" Template in Fig. 4). Logically, the comparison for an object begins in these significant areas and is only continued if the match in these areas of deep significance is sufficient. The comparison of an entire template can be graduated and arranged in any number of hierarchically structured steps, whereby the amount of detail increases with each step. Often the comparison can be terminated after comparing only a few pixels.

**Step 5**

A strictly rectangular image is often inappropriate for finding a feature with a particular shape. Near-by pixels - belonging to other objects - may have a negative influence on the similarity-measurement, resulting in erroneous rejections. Therefore, certain areas are filtered out (see Fig. 5), which can be done very easily by a special code in the template-definition. The computation of the Laplace-Distance must be modified respectively (see \(L'\) in Fig. 7). In
this manner, not only the maximum number of pixels to be compared is reduced, but at the same time the recognition-rate will increase, especially in areas of high information contents.

**Step 6**
Furthermore also the margins of objects (the edges of objects) have to be filtered out (see Fig. 6). The result is, that variations in the width of lines, small differences in size, breaks and contacts to other objects do not have a negative influence on the result of the similarity-measurement. In addition, there must be a difference in coding the filtered area between the feature-colored pixels and the background. This is essential, for example, in the case of updating maps, where the recognized candidates in the raster image are to be replaced with the standardized templates from a pattern library.

![Fig. 5: Insignificant Background](image)

![Fig. 6: Modified Template](image)

**Step 7**
Before starting the comparison of a template with the raster image, a preparation of the template data is necessary for minimizing the number of arithmetic operations (see Fig. 7). First of all, the template data have to be transformed appropriate to size and resolution of the objects in the raster image. In addition the individual pixels of the template are stored in a one-dimensional array with direct access, whereby the order corresponds exactly to the sequence of the pixel-comparison (see step 4, 5 and 6).

\[
L' = \sum_{z=1}^{m} \sum_{s=1}^{n} | B(X+z', Y+s') - T(z', s') | \quad \text{with} \quad z' = z - \Delta x \quad s' = s - \Delta y
\]

\[
L'' = 1/\text{num} \sum_{z=1}^{m} \sum_{s=1}^{n} | B(X+z', Y+s') - T(z', s') | \cdot \begin{cases} 1 \text{ if } T(z', s') = \square \text{ or } \square \text{ or } \square \\ 0 \text{ else} \end{cases}
\]

with \( \text{num} = \sum \) of all elements in \( T \), where \( T = \square \text{ or } \square \text{ or } \square \)

**Fig. 7: Modified Laplace-Distances \( L' \) and \( L'' \)**

**Step 8**
Principally, template matching can be done in the original image. Prior manipulations such as segmenting candidates are not absolutely necessary but may nevertheless be useful, depending on the circumstances. An interesting possibility is positive or negative masking,
which is also accomplished in matching suitable templates. For example, large areas of feature-colored pixels (e.g., great houses in urban zones) can be eliminated by comparing sufficiently large rectangular templates containing only feature-colored pixels.

6. An exemplary application

The efficiency of the steps mentioned above will be illustrated by an exemplary application:

**Image**

  - Scale: 1:25000
  - Size: 70 x 48 cm
- Scanner-resolution: 512 dpi = 20 lines/mm
- Raster-matrix: 14000 x 9600 pixel
- Black layer: ca. 10.4%

**Template**

- 10 different objects to be identified numerals 0 to 9
  - in the black layer, used for coding heights size 1.2 mm
  - of single points and triangulation points
- Size of searching templates 24 pixel
- Size of original templates in the library 120 pixel

**Results**

- Processing-time (10 templates, total map sheet) ca. 63 min.
- Number of
  - objects in the map sheet 2558
  - classified objects 2566
  - not correct classified objects 42
  - not recognized objects 34
- Rate of errors (= 40/2566) $e = 1.56 \%$
- Rate of refusal (= 34/2566) $r = 1.33 \%$
- Rate of Recognition (= 1-e-r) 97.1 \%

**Hard-/Software**

- Workstation IBM RISC System/6000 Model 530 (ca. 32 Mips)
- Program "KAMU", Institute of Geodesy and Photogrammetry, ETH Zuerich, Jan. 93

**Note**

It's worth mentioning, that the application is not restricted to an identification of characters. Quite different objects are possible to be searched!

If requested the recognized numerals of the original layer (black color) can now be replaced by the standardized templates. They can also be deleted and/or stored in a separate layer with a higher resolution as desired. Furthermore, the individual numerals can be used for creating z-coordinates, which can be used themselves for a lot of applications. Many other kinds of map-informations can be extracted by exact the same procedure. Using the methods of pattern recognition with raster data, the automatic data flow from analogous maps to GIS is definitely a realistic goal. Even if only the lettering of a map is extracted, it can be used for a wide variety of applications (i.e. data base of geographic place names, simple digital terrain model etc.).
7. Evaluation

The above example of course can not be sufficient for a realistic evaluation. Extensive tests with the most varied templates lead to the conclusion that some particular problems of automatic data acquisition can be solved efficiently by the use of template matching methods. The main criteria can be summarized as follows:

- For several reasons it is preferable to do pattern recognition directly in the raster image, thus avoiding any pre-processing, especially raster-vector-conversion.
- Template matching algorithms are robust, that means high recognition rate and low error rate, also for:
  - areas with high information contents
  - objects that touch or cross each other
  - objects with petty variations in size, orientation and line width
  - topologically distorted objects (due to graphic shortcomings)
- The great flexibility in defining the template allows the user:
  - an almost unlimited number of objects to be searched for
  - to adapt the tolerance-parameters to marginal conditions (quality of the map etc.)
  - to realize positive or negative masking by using suitable templates
  - to manipulate and graduate the hierarchy
  - to incorporate a priori knowledge by different coding of significant and detailed objects respectively. Instead of defining rules, the "knowledge base" can be applied graphically by defining the template.

- Processing-time

  The algorithms used for template matching are of course unsuitable for on-line applications. However, it has been shown that if a lot of single steps are optimized and if the parameters are chosen intelligently, an acceptable processing-time may be attained. Additional time can be saved by combining appropriate methods for recognizing areas of common interest such as pre-classification in a raster image with a lower resolution ("image pyramids"), pre-classification with templates of a lower resolution or searching for geometric features with sub-templates.

It is not yet evident if template matching can also be used to solve difficult identification problems such as complex structures or linear objects. Some examinations being done show, that this method is not strictly limited to recognize isolated objects with a fix orientation. The lettering of contours, individual symbols and certain linear objects such as ski-lifts are contents of topographic maps which can very well be automatically classified using template matching methods. Not the theoretic feasibility but the needed processing-time will remain as main criterion.

REFERENCES

Brennecke J.
Versuche zur automatischen Mustererkennung von Gebäudeflächen der TK50
Nachrichten aus dem Karten- und Vermessungswesen 1, No. 99, 1987

Crosetta F., Piccin L.
A new approach for pattern recognition in cadastral cartography
GIS, No. 4, 1991

Giebels M.
Automatische Erfassung von Höhenpunkten durch Mustererkennung im Kartenwerk TÜK200
Nachrichten aus dem Karten- und Vermessungswesen 1, No. 103, 1989

Grüneich D.
Welche Rolle spielt die Kartographie beim Aufbau und Einsatz von Geo-Informationssystemen?
Kartographische Nachrichten, No. 1, 1992
Illert A.
Automatische Erfassung von Kartenschrift, Symbolen und Grundrissobjekten aus der DGK5
Wissenschaftliche Arbeiten Universität Hannover, No. 166, 1990

Illert A.
Automatisierte Digitalisierung von Karten durch Mustererkennung
Kartographische Nachrichten, No. 1, 1992

Lehmann H.-J.
Automatische Objekterkennung spezieller Punkte in Karten
Vermessungstechnik, No. 6, 1990

Spieß E.
Bemerkungen zu wissensbasierten Systemen für die Kartographie
Vermessung - Photogrammetrie - Kulturtechnik, No. 2, 1990

Weber W.
Kartographische Mustererkennung
Kartographische Nachrichten, No. 3, 1988
Automatic digitization of geographical names on 1/25,000 scale maps
K. Inaba, K. Kamada (Tsukuba-shi, J)

ABSTRACT
Geographical Names Database from 1/200,000 scale maps of Japan developed in Geographical Survey Institute is described in detail. Then the approach for digitization of geographical names from 1/25,000 scale maps is shown and some elementary experiment results are reported.

1 Introduction
The names on a map are vital because they identify where things are and that is one of the most important purpose of maps. Recently computers have been used for geographical information processing (GIS etc.), and the demand for preparation of digital geographical names has increased. In order to reply these request, Geographical Survey Institute (GSI) has developed Geographical Names Database (GNDB) in 1991. The database includes about one hundred and twenty thousand (120,000) geographical names placed on the 1/200,000 scale regional maps. The digitization was achieved manually.

After completion of GNDB, GSI started the research on automatic digitization of geographical names placed on the 1/25,000 scale topographic maps, which are the largest scale national base maps covering all of Japan. Automatic digitization of geographical names printed in Japanese requires technical achievement in the field of artificial intelligence especially character recognition. Our digitization procedure can be divided into three steps:

(1) Segmentation
Many types of characters are printed on a map: omni-font, variable size, multi form exist and string direction is not fixed. Every character must be classified in accordance with each type, then some characters constituting a name are selected.
(2) Recognition of character

Each character is recognized and corresponding code is stored with relevant x,y coordinates.

(3) Interactive digitization

Geographical names not recognized automatically are digitized by interactive digitization system. Although completely automatic digitization of geographical names is desirable, it is rather unrealistic up to now. Our main approach is not developing completely automatic system but developing over all efficient digitizing system which also includes man-machine interactive operation.

The results of several elementary experiments are reported.

2 Construction of Geographical Names Database (GNDB)

A geographical names database of Japan was constructed by the GSI in 1991. This database includes one hundred and twenty thousand (120,000) geographical names placed on the 1:200,000 scale regional maps. The database is expected to be an efficient tool to make the list of the geographical names which is to be compared with the corresponding list prepared independently for charts. Background, purposes, contents, construction procedure and use of the database are described below:

(1) Background

Geographical names, which are important means to designate geographical entities, are widely used in social activities from ancient ages. For example, it is in the eight century that the old Government of Japan ordered the local governments to compile the "Hudoki", which is an old gazetteer and topographic description.

In addition, the Japanese language is written in three character systems such as ideogram "Kanji" system, phonogram "Hirakana" system and another phonogram "Katakana" system. Official Japanese documents are written with a combination of "Kanji" and "Hirakana" characters, while "Katakana" characters are used only for imported words and proper nouns such as some personal commodity and geographical names. Since "Kanji" characters usually have more than two pronunciations and most of geographical names include "Kanji" characters in it's official spelling, it is difficult even for native speakers to pronounce or to read written geographical names correctly and to write down pronounced geographical names correctly.

By these reasons, many geographical names databases including gazetteers have been constructed in Japan. In these existing data bases, the location of the entity indicated by the geographical
name is shown simply by the administrative division name or the rough geographical coordinates. Furthermore, official databases, which have usually authoritative information about spellings and pronunciations, contain names of administrative divisions only, while in private company’s databases, which include other feature names such as the names of natural features and transport facilities, information about spelling and pronunciation of geographical names is not necessarily authoritative. In other words, there is no sufficient database for the standardization of geographical names.

(2) Purposes of the database

It became clear that the standardization of the geographical names placed on the 1/25,000 scale topographic maps takes unfeasibly long time by manual means. The most troublesome part of the work is to make two respective lists of geographical names on maps and charts and to compare them. The progress of computer technology makes it possible with less cost to exempt the above mentioned job by means of database.

Under this situation, GSI decided to set up a geographical names database in order to use it as a tool for standardization of geographical names. In addition there is another purpose of the database to eliminate duplication in time and money spent by government agencies, private companies and institutions to organize similar databases for specific needs, in particular, to automate the processes of lettering the maps and to provide an indexing system of regions, areas and places for establishing geographical information systems.

To achieve these purposes, the new database should cover entire Japan and all kinds of features, and it should contain precise information on such items as official spelling, pronunciation, location, etc. of the geographical names.

Hence the following characteristics are indispensable for primary sources of the geographical names in the new database.

1) The sources cover all over Japan in the same accuracy;
2) The sources contain all classes of features;
3) The sources contain the geographical names which were adopted according to a consistent criterion;
4) Correct pronunciations and spellings can be obtained from the sources; and
5) Precise positions can be measured from the sources.

To meet these requirements, the most suitable sources are the 1/25,000 scale topographic maps, which cover all over Japan with approximately four thousand (4,000) sheets and contain approximately four hundred thousand (400,000) geographical names. However the 1/200,000 scale regional maps, which cover all over Japan with one hundred and thirty (130) sheets and contain one hundred and twenty thousand (120,000) geographical names, was
selected as sources of the new database, because of budgetary limitations and of the urgent need for the database.

(3) Contents of the database

The information items of the database are as follows:

1) Map sheet name
The sheet names of both a 1/200,000 scale regional map and a 1/50,000 scale topographic map are recorded. The sheet name of the 1/50,000 scale topographic map indicates where the geographical name is found in the 1/200,000 scale regional map.

2) Identity number
The sequential number of the geographical name is recorded.

3) Feature classification
Features are categorized into ten (10) classes such as administrative division, summit, lake, river, sea area, beach, island, broad natural feature, transportation facility and others. Some classes are divided into subclasses. The feature class of the geographical name is recorded. The size and spacing of characters can be known by the feature class because this classification is consistent to the lettering rules of the 1/200,000 scale regional maps.

4) Administrative division code
The code number of the administrative division which includes the area delineated or pointed by the geographical name is recorded. The code number is decided by the Ministry of Home Affairs.

5) Spelling
The official spelling, which usually includes ideogram "Kanji" characters, is recorded.

6) Pronunciation
The pronunciation derived from the "Geographical Names Description" reported by the local government is recorded.

7) Number of the characters
The total number of characters in the official spelling of the geographical name is recorded.

8) Position of the geographical name on the 1/200,000 scale regional map
The lower left corner point coordinates of both the first and the last characters of the spell string are recorded. The spacing can be calculated from this data and the number of characters. The spacing of characters of the geographical names of subdivided administrative division or railway station is fixed according to the lettering rules. Therefore only the corner point of the first character is recorded for this kind of name.

9) Relation between character orientation and spell string direction
The character orientation with respect to the spell string direction, parallel or cross, is distinguished and recorded.

10) Mispronunciations
Possible mispronunciations of the geographical name is
recorded. Therefore the geographical name can be found even if correct pronunciation is unknown.

11) Data source
   A principal source material of the geographical name in the database, namely the sheet name of the 1:200,000 scale regional map which contain the geographical name, is recorded.

(4) Procedure of construction of the database

There are four main stages in preparing the database, namely, collection, listing, digitization, and compilation.

1) Collection
   Primary sources such as the 1/200,000 scale regional maps and the Geographical Names Descriptions are collected. Then all geographical names placed on the source maps are marked for digitization and numbered for listing.

2) Listing
   The geographical names are listed with the attribute information such as official spelling, pronunciation, feature class and so on obtained from the sources.

3) Digitization
   Necessary points are digitized with accuracy of one tenth second (0.1") in longitude and latitude.

4) Compilation
   The listed attribute information and digitized coordinates are combined into a database, which is stored on thirty (30) megabyte hard disk.

3 The problem of digitization of letters on Japanese maps

(1) Characteristic of Japanese characters

Automatic recognition of Japanese characters is more difficult than alphanumeric characters, on the grounds that Japanese characters have following characteristics.

1) Japanese language is written in three character systems such as ideogram "Kanji" system, phonogram "Hirakana" system and another phonogram "Katakana" system.
2) Japanese language has a lot of characters: "Kanji" has about 3,000 characters, "Hirakana" and "Katakana" has about 70 characters respectively.
3) The form of characters is complicated.
4) Some characters are similar in form.
5) There are various letter forms and letter styles.

(2) Characteristic of letters on a map

Letters on a map are regulated by lettering rules and have following characteristics. Therefore automatic recognition is difficult.
1) There are several letter sizes depending on feature.
2) There are two letter styles depending on feature: "Gothic style" and "Ming letter style".
3) There are four letter forms depending on feature: "Upright", "Slant right", "Slant left", "Syoken".
4) The location of characters of a text is parallel, cross or curved, so the segmentation of characters is difficult.

4 The basic strategy of geographical name digitization and some idea

(1) Basic strategy

There proposed many techniques for the automatic recognition of Japanese character ("Kanji") and some efficient techniques have been reported. However these reports says that such techniques are applied to several sample texts. So it is not clear how much time is needed for automatic digitization of geographical names on a map which has more than 100 names. Although completely automatic digitization is attractive as a scientific problem, from the point of practical use, applying both automatic digitization technique and interactive digitization technique jointly is realistic.

(2) Some idea for automatic digitization

1) Application of GNDB

   GNDB above described includes 120,000 records including coordinates of letter place. This database is a powerful means for digitization of geographical names of 1/25,000 scale maps. Total number of geographical names on 1/25,000 scale maps covering all Japan is estimated about 400,000, and retrieving the corresponding name by coordinates expectedly helps digitization of 30% of all geographical names on 1/25,000 scale maps.

Table 1  The appearance rate of most frequently appearing characters to all characters

<table>
<thead>
<tr>
<th>Feature</th>
<th>5 Characters</th>
<th>10 Characters %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative division</td>
<td>12.1</td>
<td>20.5</td>
</tr>
<tr>
<td>Summit</td>
<td>33.1</td>
<td>39.2</td>
</tr>
<tr>
<td>Lake</td>
<td>33.2</td>
<td>39.1</td>
</tr>
<tr>
<td>River</td>
<td>32.1</td>
<td>37.5</td>
</tr>
<tr>
<td>Sea area</td>
<td>32.0</td>
<td>41.5</td>
</tr>
<tr>
<td>Beach</td>
<td>33.6</td>
<td>40.0</td>
</tr>
<tr>
<td>Island</td>
<td>38.2</td>
<td>44.3</td>
</tr>
<tr>
<td>Broad natural feature</td>
<td>30.6</td>
<td>47.3</td>
</tr>
<tr>
<td>Transportation facility</td>
<td>10.7</td>
<td>18.0</td>
</tr>
<tr>
<td>Others</td>
<td>14.9</td>
<td>24.4</td>
</tr>
</tbody>
</table>
2) Utilization of knowledge about geographical names

Some characters are frequently appear in geographical names. Frequency of appearance of some characters in geographical names is estimated by above GNDB for each features and the result is shown in Table 1. From the table, only 10 characters occupy about 20 to 40% of total characters.

5 Some elementary experiments for automatic digitization

(1) Purpose

In order to cutout characters constituting a text, the recognition of character size and letter form is essential. The item useful for segmentation is as follows.

1) Character size
   There are five (1.5mm, 2.0mm, 2.5mm, 3.0mm, 3.7mm) sizes.
2) Letter style
   There are two fonts (Gothic, Ming)
3) Letter form
   There are four forms (Upright, Slant right, Slant left, Syoken)

This time we performed elementary experiments to classify characters.

(2) Procedure

The procedure of experiments is shown in Fig. 1. An annotation symbol plate is scanned. The scanning pixel size is 25 and 50 micron meter.

1) Data input by raster scanning
   A linear scanner is used for the data collection, so the digitized data is a raster format. In the case of the raster format data, the precision of the data is controlled by the scanning resolution. Therefore, in order to improve the precision, high scanning resolution is needed, but it makes larger amount of digital data. So, after some investigation on both the precision of the raster data and the volume of the data, 25 and 50 micron meter was decided as the scanning resolution.

2) Vectorization
   The data has to be converted into a vector format (the conversion is called vectorization) in the next process. Vectorization method detects the boundary lines of the area, trace the lines and form a polygon. The polygon is so-called "Outline font".

3) Estimation of character size
   The estimation is as follows.
   a) To calculate a circumscribed rectangle for each polygon of outline font.
   b) To group several polygons of the rectangles, which are
located close enough to be estimated that these polygons constitute one character.
c) To calculate the size of a character from the result of grouping.
4) Discrimination of letter style
   The algorithm is as follows.
   a) To trace the outline font
   b) To analyze piecewise lines by the length and the angle of two adjacent lines
   c) To classify a character into two classes (Gothic, Ming)
5) Discrimination of letter form
   This time we tried to classify into two classes, that is upright and slant. The algorithm is as follows.
   a) To calculate the azimuth of every line and make a histogram.
   b) To classify a character into two category by analyzing the histogram.

![Fig.1 Procedure of Experiments](image)

(3) The results of experiments
1) Estimation of character size
   The rate of characters classified to correct size is 87% and 13% are not classified to any groups. There are some example that originally one character is recognized as smaller two characters.
2) Discrimination of letter style
   Characters belonging to Gothic are completely correctly classified. 75% character of Ming letter form are correctly classified and remaining 25% are not classified.
3) Discrimination of letter form
   "Upright" form characters are correctly classified upright group. 73% of "Slant" form are recognized as slant but 27% are remaind unclassified.
7 Development of interactive digitization system

At present, completely automatic digitization of geographical names is not yet practical, so we developed interactive digitization system for geographical names. This system shows a scanned annotation symbol plate raster data on the display and the coordinates of character can be measured on the display. On the same display, the system can show another window in which the geographical names are listed from a different file, and we can easily connect the coordinates of geographical name in raster form with the code of that name. The display is shown in Fig 2.

Fig 2. Display of an interactive digitization system
8 Conclusions and future work

GSI has completed development of the Geographical Names Database from 1/200,000 scale maps. This database is useful itself.

GSI has started the research for digitization of geographical names from 1/25,000 scale maps. The task is being performed through following method.

(1) Utilization of pre-developed Database

(2) Automatic digitization (under experiment)

(3) Interactive digitization system

In order to increase overall efficiency of digitization, we intend to combine above three method mutually. Although only few elementary experiments were done for automatic digitization, we need to develop practically powerful automatic digitization technique.

REFERENCES


Yamamoto, K. and Mori, S., Recognition of Hand Printed Characters by an Outermost Point Method, Pattern Recognition, 12, pp. 229-236, 1980

Lipton, R. J. and Tarajan, R. E., A, Separator Theorem for Planar Graphs, SIAM Journal on Applied Mathematics, Vol. 36
Session 11

Map Based Information Systems II

Chairman:
M. Wood, Centre for Remote Sensing (Aberdeen, GB)
The coastal zone marks the triple-boundary between the sea, the land and the overlying atmosphere. One of these three media - the land - is relatively stable, the sea and the air are each highly fluid. While we can intuitively recognise the existence of the shoreline, its precise definition in time or space is extremely problematic. It has long been known that as you increase the resolution of measurement, the length of a given stretch of coastline also increases: the coast is a classic example of an entity with a fractal dimensionality.

However, just as we can pick up ever more detail when we focus in on finer spatial resolutions on the coast, so too may we recognise multiple scales of temporal activity. Coastal processes operate over spatial scale measured from hundreds of kilometers or more down to the sub-atomic, and over timescales ranging from millenia down to micro-seconds. Many of these different movements and changes may have an apparently cyclical component, although closer examination reveals that, while we may possibly recognise upper and lower bounds of any given change, any given state of coastal configuration will never be precisely returned to. Thus, the behaviour of the coast appears to have affinity with aspects of the newly-emerging scientific paradigms based on Chaos theory.

This paper explores some of these concepts and issues, and discusses some possible implications that they may have for the design and implementation of Coastal GIS.

1. INTRODUCTION

Geographical information systems (GIS) first emerged as a tool for spatial analysis some thirty years ago (Tomlinson, 1988; Coppock and Rhind, 1992). Since then, the technology has found successful application in a wide range of disciplines and professions, including many branches of natural and human resource management. However, for all of their utility, it should also be acknowledged that current GIS technology still contain significant shortcomings and weaknesses with regard to certain types of application, and certain forms of data manipulation and analysis (see, e.g. Burrough, 1986; Aanengenrug, 1991). In particular, Chrisman (1987) has cautioned that "some of the current success [in GIS] is achieved by exploiting the easy parts of the problems. The tough issues, temporarily swept under the rug, will reemerge, perhaps to discredit the whole process."

Chrisman's assertion is particularly apposite when coastal applications of GIS are concerned. Analysis of the literature (e.g. Bartlett, 1990, 1993) reveals that coastal scientists have been aware of GIS and related technologies for over twenty years: in 1972, for example, Ellis wrote of the need for a "coastal zone management information system [which] relates data,
information, predictive techniques, environmental interactions, methods of analysis and applications into one system of procedures, tools and instructions, for use by planners" (Ellis, 1972, p95). In the subsequent two decades, several attempts have been made to harness spatial data-handling tools to aid understanding of coastal systems, to assist in devising integrated management strategies, and for use as decision support systems (e.g. Clark et al, 1990; Law, 1990, 1991).

The success of these efforts is debatable. Clark et al are enthusiastic, and suggest that "...although [the coastal] zone shares many of the problems and solutions common to the broad fields of land, landscape and environmental management, it reflects some unique characteristics which render it particularly well-suited to representation within a geographic information system". This enthusiasm may however be misplaced. While the application of GIS to coastal problems have undoubtedly led to improved knowledge and shoreline management practices, it nevertheless has to be recognised that they have mostly achieved their success by "exploiting the easy parts of the problem". It may even be suggested that many coastal implementations have achieved their success almost in spite of, rather than because of, the fundamental character of current GIS. Closer examination of the problem suggests that the use of GIS for coastal management represents the convergence of several distinct "tough issues", each prominent in its own right on the GIS research agenda, and the unique characteristics of the coastal zone render it particularly badly suited to representation within a GIS.

2. CREATING A COASTAL DATABASE: FUNDAMENTAL PRINCIPLES

Coastal management may be seen as a special category of resource management and allocation (Bartlett et al, 1992) and therefore requires decisions to be made. Sound decision-making in any field is ultimately dependent on the quality of information available, and provision of information in turn requires data to be processed. However, simply collecting data at random and without any particular framework is not enough. In order to be used efficiently, they have to be organised into an appropriately structured database. This requires careful and methodological abstraction of data and observations from the real world, and the translation of these into forms that are meaningful to the computer.

Unfortunately, although the coast is acknowledged to be a distinct feature of the earth’s surface, and one which most people recognise on a subjective level, it is an inherently difficult part of the real world to model in formal terms. For the cartographer, the world's coasts always have presented distinct problems: this was succinctly expressed by Guth (1972), who wrote that "part of the problem of mapping coastal boundaries is that it hasn't been clear what to map or how to map it". Byrnes et al (1991) suggest that the interaction of five primary factors at the shore, namely coastal processes; relative sea level; sediment budget; climate; and human activities "makes shoreline mapping especially difficult" (Byrnes et al, 1991). Even arriving at a formal definition of what constitutes "the coast" as a unit is not easy: Bartlett (1990) has described the coast as a "highly dynamic interface between three key components of the global ecosystem: land, sea and the overlying atmosphere", while Carter defines the coastal zone as "that space in which terrestrial environments influence marine (or lacustrine) environments and vice versa" (1988). Additionally, if the coast as a unit of territory is itself an inherently difficult feature to define, the terminology used to describe components and entities within the coastal system is also generally unstandardised. Opinions therefore differ considerably as regards the identification and delineation of discrete
morphodynamic and other coastal features. A comprehensive account of traditional coastal zone mapping in the United States is given by M.Y. Ellis (1978); while, more recently, discussions of local-scale coastal mapping are provided by Leatherman (1983), Anders and Byrne (1991) and Byrne et al (1991) among others.

3. SPACE AND THE COAST

In a classic study of the geometry of the coast, Kapparott (1986) explains that most people's concept of a coastline derives from two sources, observations of coastlines on a map and a vacation visit to the seashore. A coastline shown in an atlas, he suggests, "gives the impression of the coast as a comparatively smooth and 'ordered' curve of finite length easily determined from the scale of the map. However, a visit to the coast generally reveals rugged and chaotic terrain with rocks jutting out to the sea and an undulating shore dynamically varying over time and space". This dichotomy between the coastline as portrayed on a map, and that as seen by the observer on the ground, is echoed closely by the problems encountered in trying to derive an adequate representation of the coast within a GIS.

The coast as a one-dimensional entity.

The essentially linear nature of the coast is, at least at first sight, its most obvious feature. Even our use of the terms coastline and shoreline reflect our tendency to regard the coast in one-dimensional terms. Accordingly, it is not surprising that many attempted coastal zone GIS have adopted a data model based on a simple one-dimensional line (e.g. Weyl, 1984; Bartlett, 1988; Debusschere et al, 1991; Townend, 1990; etc). However, despite its seemingly-attractive conceptual simplicity, this linear model is in practice both extremely limiting and extremely limited when it comes to handling coastal data.

Two generic approaches may be identified in treating the coast as a line. In the first, the coast is considered as a one-dimensional entity in complete isolation from other phenomena or processes. This means that its only geometric property is length. Two examples of this type of approach may be cited: firstly, in a study of coastal erosion in Northern Ireland (Bartlett, 1988), the coastline was represented as an ordered string of coordinates (vertices) stored as a single table in a relational database. Attributes and features of interest were referenced to, and retrieved by, their position with regard to the nearest vertex as identified by its unique sequence number. Selective retrieval of data, and some basic statistical manipulation, could be undertaken by the database management system. Where graphical output was required, the coordinates were passed a mapping package for plotting as a generated line. A broadly similar, though more recent and more sophisticated example is the study by Debusschere et al (1992) of Louisiana's shoreline systems. The technique adopted here also forsakes any attempt at geometric fidelity, and considers the coastline purely in diagrammatic form as a drawn-out ribbon. In their discussion of this approach (Debusschere et al, 1992), the authors suggest that the linear nature of the technique "increases its analytical potential" when used in this way.

The second generic approach to linear shoreline representation is to again consider the coast as a line, but this time as one located in multi-dimensional feature space. In this approach, the convolutions and meanderings of the line are considered significant. It is becoming the much more commonly encountered situation since, while early attempts at applying GIS to the coast often used purpose-built software (e.g. the RAMS database: see Eberhart and
Dolon, 1980), the trend since the mid-1980s has been firmly towards applications based on "off-the-shelf" proprietary GIS software packages. This method of depicting the coast conforms more closely to the manner in which line objects are handled in most conventional vector GIS (e.g. ESRI’s ARClINFO or Intergraph’s MGE).

Whichever of the two approaches is adopted, a number of specific problems arise out of use of a line metaphor for the coast. These include:

(a) Which line is to be selected to represent the shore? There are several lines that might be candidates: for example, one might fix the line with respect to the position of the land-water divide. But, particularly in a meso- or macro-tidal environment, the position of this line is constantly changing. Furthermore, in such a situation the rate at which the tides wax and wane can vary considerably over even moderate lengths of coast; and the state of the tide at any one point in time will vary considerably over even short distances along the shore. Therefore establishing the absolute position of this line at any moment in space and time is virtually impossible, and normally an approximation to this line has to be used instead. But this still poses the question of which line is to be selected? Mean sea level? Mean high or low water mark? The highest extent of ordinary spring tides? All of these, and more, have been used as delimitators of the shore in conventional mapping of coastal areas.

Alternatively, one might select the line corresponding to the lowest point of land-based vegetation, as is sometimes done for ecological surveys on the shore. But this, too, poses difficulties. Often the limit sought will not be sharply-delineated, but will take the form of a gradual zone of transition with passage seaward. In extreme cases, particularly in enclosed, estuarine or transgressing shorelines, this position might even occur to the seaward to the upper reach of the tides. Finally, we can define a purely arbitrary line, approximating to the actual line of the shore, and attempt to project actual occurrences of coastal phenomena onto this line. This was the approach used in early developments of the Canadian Coastal Information System (CIS) of the Geological Survey of Canada (Fricker and Forbes, 1988), but again this results in a database built around a charactature of the coast rather than the true configuration of the shore.

(b) Determining the appropriate level of detail for the line. Komar has suggested that there are some 440,000 km of shoreline girdling the World’s oceans (Komar, 1983). This, however, is a gross generalisation, and an over-simplification of what is in practice a very much more complex measure. The pioneering studies of fractals by Mandelbrot (1967), based on earlier empirical work by Lewis Richardson, established that if the scale of measurement of the coast is progressively increased (whether by measuring from increasing scales of maps, or by actual measurement of the coast on the ground), the length obtained for the coastline concerned increases progressively towards infinity as ever-increasing levels of detail are considered (Kappraff, 1986). (When one considers an island, this raises the intriguing paradox that the boundaries of the land have theoretically infinite length, while the area that they enclose remains a finite quantity!). Furthermore, coastlines exhibit the properties of statistical self-similarity. This means that the mathematical properties of any segment of the coast are similar to those of the coastline as a whole: that is, the statistical distribution of their projections and embayments are the same, irrespective of the scale at which they are measured (Kappraff, 1986). These two attributes of the coast - seemingly infinite length as the detail of measurement increases, and properties of statistical self-similarity - mean that the coastline fulfills the classic requirements of a fractal curve. Goodchild and Mark (1987)
explain the term 'fractal' as a set, whether of points, lines, areas or volumes, whose measure increases by more than the ratio of any two scales at which it is measured. While Goodchild and Mark concern themselves with fractals in space, it will be suggested later in this paper that the temporal aspects of the coast may also have some fractal affinities.

In terms of mapping the coast, this fractal dimensionality of the line means that it is inherently impossible to include every detail of the shore; the line adopted has to be simplified and greatly generalised. But many, if not most, coastal management problems require data and concepts at the synoptic or local level to be integrated and compared with those at a much more local, site-specific scale of resolution (Bartlett et al., 1992). Thus, a typical study may easily cover several tens, or even hundreds, of kilometers of shore, and yet at the same time may easily have to take full account of even comparatively small perturbations of the line of the coast, either natural or anthropogenic, which can frequently have a disproportional impact on the behaviour of ocean waves, sediment transport, shore erosion patterns, biological activity, etc., and may thus have important implications for human use of the shore (Carter, 1988; Carter et al., 1983). If selective inclusion or omission of detail has to be undertaken, the essential question then arises as to which projections and embayments are significant and which may safely be ignored? Unfortunately, despite the importance of answering this question to the success of the intended information system, there are no clear guidelines.

These are, of course, problems that has taxed coastal scientists and cartographers since the dawn of map-making, but they gain increased significance where GIS is concerned, since the analytical capabilities of these systems allow not only the retrieval of stored information, but the manipulation of data and the creation of new spatial objects from the combination or reorganisation of old ones. This means that data held within the GIS database may suffer much degradation of quality, through loss of precision, fidelity and information content.

(c) Difficulties of recording variations in attributes along the line. Most topologically-structured vector GIS represent line objects as an ordered set of vertices marked at each end by a node. Normally, attributes and properties attached to these line objects apply to the entire extent of the line between the nodes. However, there are few natural nodal points along the coast and yet it may frequently be desirable to assign different properties to different sections of the line describing a stretch of coast.

One solution to this might be to attempt to redefine all possible combinations of all possible attributes, and divide the line up according to the resulting segments. An approximation of this approach was adopted at Cork (Scalise and Bartlett, 1992; Bartlett et al., 1992), in an analysis of shoreline vulnerability to sea-level rise. In this study, the shoreline was conceptualised as a line divided into segments, each having more-or-less homogenous characteristics. This process was repeated in a series of layers, each layer of data dividing the same line object differently according to the distribution of the attribute in question along the shore (i.e. one layer divided the line of the shore according to the dominant sediment type encountered; a second according to presence and type of coastal defences; etc.). However, as Ducker and Vrana (1992) note, this type of approach is unwieldy. One problem encountered and not fully resolved at Cork was the difficulty of integrating and correlating the different layers of data: while conventional vector GIS are now routinely capable of overlaying polygon data, they are still surprisingly deficient in supplying equivalent facilities for overlaying single (though complex) line objects. An alternative approach is to represent
the coast as an ordered string of arbitrarily-determined segments or arcs, either of variable length (e.g. Fricker and Forbes, 1988), or else all of the same constant length, and storing all attributes for each unit (Ducker and Vrana, 1992). However, this would inevitably result in a high degree of duplication and redundancy of data (Ducker and Vrana, 1992), and also would be expected to present difficulties when multiple data layers or coverages need to be integrated.

In the past two years or so, a new approach to the division of line objects in vector databases has emerged. Known as the technique of dynamic segmentation (Ducker and Vrana, 1992), it allows a single line object to have different attributes or properties attached "on the fly" to different sections, without having to actually divide the line into smaller logical segments each separated by its own start- and end-node. At least four proprietary GIS packages now incorporates dynamic segmentation of the line within their range of data-handling facilities (Ducker and Vrana, 1992), but so far the practical application of this concept has tended to focus on transportation studies, with minor application in streams and utility-management sectors. In theory, dynamic segmentation of the line should offer distinct possibilities for more effective representation of coastal variation within a GIS, but the extent of and limitations to this potential have yet to be tested empirically.

(d) The difficulty of integrating data recorded at positions some distance removed from the nominal line of the coast. In many instances, phenomena of interest to the coastal scientist are located a short distance away from the nominal line of the shore selected for data representation. In the case of a reasonably straight and uncomplicated extent of coastline, this might be quite easily resolved by extrapolating the position to the line of the shore. However, where the coast is heavily indented, there may be a wide range of equally-valid points along the shore to which the observation might be nominally attached.

The coast in two dimensions

Despite the conceptual simplicity of the coast as a linear feature, in practice the coastal system also has width: maritime influences may extend considerable distances inland, while terrestrial influences may be felt far off-shore. Coastal variables having (at least) two dimensions include the distribution of differing sediment types offshore and along the coast; the extent of coastal littoral cells (units of basic morphodynamic functioning of the shore); the area and modes of use of recreational beaches; the geography of inter-tidal mudflats exposed at low tide; the area of harbours and ports; the location of marine outfalls; etc.

Both the tesserl and the vector models have their strengths and weaknesses for coastal zone GIS. In the tesserl model, the focus is on areas and the boundaries between areas are implied. By contrast, in the vector model, the primary emphasis is on boundaries, and the interior space bounded by these lines is implied (Berry, 1987). Since the coastal zone represents an area of contiguous space, this would suggest that tesserl modes of data representation are more appropriate for coastal systems.

The nature of the coastal boundaries adds further support to such a conclusion. Attention has already been drawn to the inherent difficulties frequently encountered in defining coastal entities. The coastal system is one thus dominated by poorly-determined, inherently fuzzy boundaries and relatively imprecise demarcations: between sediment domains on a beach, between vegetation assemblages in a salt marsh, or in salinity distributions in an estuarine
environment, etc. At the present time, almost all GIS depend on hard, boolean logic and sharply-delineated entities and relationships. Investigation and development of spatial data processing techniques incorporating alternative methods based on statistical probability and related techniques, is prominent on the GIS research agenda (e.g. Burrough, 1989; Kollias and Vouliotis, 1991), and there is no doubt that incorporation of fuzzy reasoning methods will be of great value to coastal GIS. Most of these emerging new techniques appear based on tesseral data models, and at this juncture it is difficult to visualise how comparable concepts and methods might be developed for handling vector data.

However, a few critical weaknesses of the model also need to be considered. Firstly, particularly within the off-shore portion of the coastal zone, the majority of the data to be incorporated within the GIS are likely to be based on point sampling techniques (e.g. data obtained from wave or current meters, sea-bed sediment analyses, etc), or else will be captured along transects (e.g. geophysical soundings, survey ship cruise lines, etc.). Thus, the use of raster data within a coastal GIS may therefore require considerable interpolation between these discrete and non-contiguous data points. (One major exception to this is the increasing quantity of remotely-sensed data that are becoming available for coastal zone studies, which may be considered as representing a continuum of sampling across an area, at least to a first-order approximation. The fact that these remotely sensed data are themselves captured and provided in a raster format is a further strong argument in favour of adoption of tesseral data models).

A second, and somewhat more serious criticism of the tesseral model for coastal data handling concerns issues of computer data storage efficiency. Given the large areal extent of many coastal studies, and the high levels of cell resolution that are needed for meaningful study of the shore, such issues become particularly critical. Allied to this, in most coastal GIS applications, the greatest level of interest, and thus the greatest requirements for precision and high-resolution data, will be in those areas closest to the meeting place of land and water, while interest and requirements for detail diminish progressively with distance from the shore. As the complexity of the shore increases, it becomes progressively more difficult to control the amount of redundancy. When a long expanse of convoluted shoreline requires consideration, as might be the case on the fiorded coasts of Norway or Scotland, for example, or in the study of a deeply-indented enclosed embayment such as Cork Harbour in Ireland, a substantial amount of storage space will be occupied by pixels that do not contribute in any significance to the actual study.

Thus, for a variety of reasons, a fixed-resolution tessellation will have distinct disadvantages over one where the resolution of the cells may be varied in response to differing requirements of detail. One possible solution to this problem is the use of the variable-resolution tessellation known as the quadtree. By dividing the raster recursively, regions of great homogeneity (e.g. large expanses of sea) can be stored and represented as larger units, while areas where there are greater levels of detail and more rapid change may be represented by progressively smaller cells. A number of coastal applications of quadtree GIS have been described (e.g. Clark et al., 1990; Law et al., 1991), based on use of the SPANS product (Tydac Technologies, 1990) but, as far as can be determined, no empirical study has yet been carried out to assess objectively the advantages, strengths and weaknesses of the quadtree model compared to the conventional raster, with regard to coastal GIS.
**Depth: the third spatial dimension of the coast**

The coastal zone has depth, as well as length and breadth, and coastal process analysis and modelling requires access to all three of these dimensions. However, representing the third spatial dimension in coastal GIS is barely possible at the present day, partly because of shortcomings in the technical abilities of GIS software, but also because accurate, sufficiently-detailed elevation data relating to the coast is virtually non-existent except, perhaps, for a few areas that have been subject to particular study. This latter question could be anticipated as being of particular importance in studies of relative sea-level change, since changes in vertical position of the sea by even a metre may have major effect on both the neighbouring on-shore and off-shore coastal processes.

In technical terms, representation of three-dimensional objects is routinely undertaken in computer-aided design (CAD) systems. However, true handling of 3-d data within GIS still lies very much at the cutting edge of current research (Raper, 1989) and so, although the need for 3-d data handling capabilities for coastal GIS may be acknowledged, its resolution awaits further development of concepts and technological capabilities.

**4. TIME AND THE COAST**

Goodchild (1988, p 317) has asserted that "much spatial data is relatively stable through time". However, while this may be true for many applications, it does not apply to the coast. However, as with the extension of GIS into three spatial dimensions, so too incorporation of true temporal databases and modelling capabilities within GIS has not yet occurred, although it is high on the present research agenda (Langran, 1989; Langran and Chrisman, 1988). In the mean time, coastal applications of GIS are severely handicapped and restricted by the lack of appropriate methods of representing shoreline dynamics.

The coastal zone is in constant motion. It is characterised by cross- and long-shore movements of matter, energy and information at all spatial scales from millimetres to kilometres or more, and at temporal scales measured from fractions of a second to millennia. But separating the temporal from the spatial aspects of coastal processes is itself not always straightforward, as will be apparent in the discussion that follows. Even the tidal and related dynamics of the shore, as discussed earlier, mean that the absolute position of the line between land and water can never be determined: the position of the land-water interface is constantly changing and, furthermore, each small section changes its position independently of its neighbours, so that one wave may be advancing up the shore at one point, while a neighbouring one (at greater or lesser distance away) is in the process of receding temporarily. And, when the next wave comes, the extent of advance and retreat of each successive wave will differ in detail from the preceding ones, and the swept surface of shore will likewise be different.

Superimposed upon these small, local-scale changes are larger-scale dynamics. The tidal cycle, for example, causes the periodic, twice-daily rise and fall of the level of the sea and, with it, associated changes in the submerged extent of the littoral zone. Within certain limitations, the upper and lower extent of these tidal changes may be determined but, as with the movement of small waves upon the shore, the precise levels of rise and fall of the tide can vary significantly both longitudinally (that is over time at the same point in space) and in cross-section (where variation along the shore is examined at a single defined moment in
time). The monthly lunar cycle, in particular, causes significant variation between the neap and the spring tides, while the action of storm surges and other climatic events may perturb the normal pattern of the tides in less-easily predicted ways.

Coastal changes also take place at annual time scales: for example, winter storms on open coasts may frequently cause sediment to be transferred from shore to deep-water storage, as wave energy is expended on the shore (Carter, 1988). This phenomenon may cause radical differences in the character of the shore; for example, a few kilometers along the open coast from the mouth of Cork Harbour, southern Ireland, is a small embayment which, in summer, contains a fine, sandy beach. In winter, all of the sand is routinely stripped away and deposited in a sand bar some distance to seaward. At these times, the "beach" consists of coarse boulders and gravel. Clearly, such changes can have severe implications for maintaining an up-to-date and representative spatial database of shoreline morphology within a GIS. Finally, one may consider the changes that take place on the shore at yet longer timescales. Global sea-level rise, for example, is a currently topical issue of concern. While early estimates of world sea-level rise of the order of several metres (e.g. Hoffman et al., 1983) have since been revised (e.g. Devoy, 1990, 1992), it is certain nonetheless that many coastal communities will be threatened with flooding and marine encroachment over the coming few decades. GIS is already being harnessed as a tool to help assess the magnitude and impacts of such changes (e.g. Gornitz et al., 1991; Devoy and Bartlett, 1992). Geological and recent oceanographic studies have also shown clearly that world sea levels periodically rise and fall over timescales measured in millennia, as ice sheets advance and retreat, changes in the geoid occur and so on (for fuller discussion of such changes, the reader is referred to the volume of papers on the topic of global sea levels, edited by Devoy, 1987).

Although these changes operate on very different scales, both spatial and temporal, one is nevertheless tempted to suggest tentatively that the vertical and horizontal changes in shoreline position, as identified above, exhibit statistical self-similarity and fractal dimensionality in all three spatial dimensions. Furthermore, though even more tentatively since no attempt has yet been made to assess this mathematically, it may also be suggested that the temporal aspects of these changes also exhibit fractal characteristics, since changes taking place over small time-periods (minutes, hours) have distinct similarities in overall behaviour to those that are measured in millennia.

A second point of note is that, while all of the changes that have been described have a basic cyclical nature, in that the advance and retreat of the shore and the rise and fall of the sea each operate between upper and lower bounds (though, with progressively longer time-scales, these upper and lower bounds diverge spatially in all three dimensions), in all cases and at all scales of examination, the precise configuration of the shore, once left is never ever returned to. One wonders, therefore, whether the concept of sensitive dependence on initial conditions, which is central to the newly-emerging ideas of Chaos Theory, may not also have relevance and application to coastal studies? Again, this suggestion is offered extremely speculatively, to invite research into this question, but it has not been tested mathematically.

5. CONCLUSION: SPACE, TIME, CHAOS AND COASTAL GIS

Critical comparison of the nature of the coastal zone to the data-handling capacities of presently-available GIS leads to the overwhelming conclusion that GIS in its present state is
not an entirely appropriate technology for coastal management operations. GIS may be used in support of, or additional to, other tools of investigation, but it cannot take the lead role in managing the coastal scientist’s information to the degree that GIS have successfully been incorporated into other endeavours. Applications of GIS to the coast to date have been obliged to their data to conform to awkward and inappropriate data structures which do not inherently reflect the true nature of the real-world coastal system.

Steve Smyth (pers comm.) has drawn attention recently to the fact that existing GIS are based firmly on the conventional map as the underlying conceptual metaphor. This is partly due to the parentage of the technology, and is also in part because all currently-existing proprietary GIS have been developed with essentially terrestrial applications in mind. Terrestrial data, in both the socio-economic domain and also in the natural sciences, uses a system of basic spatial objects and modes of representation which lend themselves naturally to map-derived metaphors. However, it is arguable whether such simplistic models can truly serve the information requirements of integrated coastal management. More accurate and meaningful representation of the coast within a GIS requires elaboration of new approaches, and new data structures and, perhaps, completely new paradigms and metaphors for representing the real world in computer storage. This will require a return to first principles, and close re-examination of the actual data- and information needs of the coastal manager, and also continuing detailed study of the geography of the coastal zone.

ACKNOWLEDGEMENTS

Many people have contributed substantially to the development of ideas contained in this paper, although in many cases they are almost certainly unaware of the fact. My grateful thanks must go, in particular, to Steve Smyth for pointing out to me that alternatives to the map-based metaphor for GIS might exist and need to be looked for; Mike Goodchild may remember a discussion that we had in Pisa, Italy, where the idea of the coast as a four dimensional fractal was raised: many of the ideas contained in this paper stem directly from that conversation. Gail Langran must be thanked for cajoling and bullying me (in the nicest of ways) into writing this paper at all. And finally, I must thank my colleagues in Cork, especially Robert Devoy and Armando Scalise for their many stimulating discussions on the use of GIS for coastal zone studies.

REFERENCES


Bartlett, D.J. (1993). The Design and Application of Geographical Information Systems for
the Coastal Zone: an Annotated Bibliography. Compiled under sponsorship of the International Geographical Union Commission on Coastal Systems. (Publication forthcoming)


INTRODUCTION

As cartographers have accepted the use of computer technology in their field, they have been forced to reexamine many of the processes that they have traditionally performed on spatial information. It is becoming increasingly apparent that this reexamination is pronounced with respect to data used to make cartographic products. The structuring of digital data bases, which is required by computerized cartography, is a new and vital process. A necessary part of the digital data base is metadata, information required for the efficient access and use of the data.

Metadata is information about the digital spatial data in the data base. This paper outlines the need and suggests potential categories for the content of a comprehensive statement of data quality that must be imbedded in the metadata that accompanies a digital spatial data file. Ideally, an international metadata standard and an international data quality standard could be agreed upon by cartographers for digital spatial data.

Using analog technology, a map-producing organization makes an overall decision that the quality of the available data to be portrayed is sufficient to produce the multi-purpose map product. Often the producing agency includes in the map marginalia a statement about the quality of the map. In the case of the United States Geological Survey’s topographic map series, the statement is “This map complies with National Map Accuracy Standards.” This statement, drafted over fifty years ago, announces that a given map, as part of a single mapping project, meets a stated set of parameters defining the accuracy of the location of a set of positions (x,y coordinates) on the earth (see Appendix 1). If elevations, in the form of spot heights, are also portrayed, the statement of quality also refers to the accuracy of the elevations (z coordinate).

Quality information available to the user of an analog produced map includes, at best, a quality statement furnished by the producing organization such as the one referenced in the preceding paragraph, the reputation of that producer organization, and any experiences resulting from using the product, which are gained by performing tasks based on information from the cartographic product. In some instances the user’s need for additional quality is discovered only after a potentially costly mistake. Alternatively, the product may have a higher quality level than the user needs, and the expense of that excessive quality is therefore wasted.

The digital cartographic environment is different. Using digital technology the producing agency collects the data. These data are then structured into a digital
database. A user accesses that database and selects the data sets that he determines are necessary to solve the problem for the production of a desired cartographic product. The producer makes the basic decisions concerning data quality and provides metadata to the user. Both then decide whether or not the data are useful for their specific cartographic purpose.

In the digital cartographic world the situation has changed so that producers are collecting data and structuring the spatial data into databases that may contain topology and enhanced attribute information. The fact that these digital databases may also be distributed allows other data producers to add attributes to features already in the database or to even add new features. In some databases, relationships among features are also specified. Therefore, any given data file may be the product of a number of producers.

The presence of these digital data bases allows today's user of analog maps to become the user of digital products. Given a digital spatial data base, the potential user first selects the data files that are to be portrayed. On what basis is this selection made? In analog cartography, the potential user of a map has never been forced to assume this selection responsibility before. Are most users adequately trained to accept that responsibility? How does or should a user learn of the quality of the digital data? It is certainly not acceptable to finish the entire process of creating a map and then attempt to use that product in the real world in order to be able to run a quality check. Quality concerns must be dealt with earlier in the digital cartographic process.

THE PRESENT SITUATION

Today it is truly a "let the buyer beware" market place in digital spatial data. It is perhaps truer than ever before that to produce a map visualization in the digital world it is not necessary for the user to have any experience in cartography or data assessment, or any knowledge of data quality. A hardware platform with a user-friendly mapping package can enable any person to create a map. In the analog cartography environment, by comparison, through repeated experiences, cartographers develop a sense of appropriateness for data use. They have the responsibility for the data collection and the selection of data to include on a map. Relying on that sense the cartographer could give his "stamp of approval" of the quality of the created map, after testing to assess the quality of a given data set, knowing those errors that were carried forward from the collection phase through the analytical phase, and knowing the errors introduced by a change of map projection and/or a change of scale. Digital data users are not always so well trained. They may lack even an introductory knowledge of map projections, cartographic generalization, or the errors propagated through a series of manipulations of the spatial data.

A closer examination of the consequences of this change in technologies indicates a secondary change as well. Cartographers have rarely dealt with the data quality issue as it relates to nonpositional data. By the time positional accuracy has been obtained to the satisfaction of the cartographer, the other attributes of the features that are to be mapped have been collected with accuracies that also satisfy the cartographer. In other words the cartographer feels comfortable with the overall quality of all the data on the map if it is possible to attain sufficient accuracy in the positioning of that data in the map context.
Any quality statement is ultimately tested by the user in the real world. However, due to the presence of the newly empowered user in the digital world, the statement of quality in digital cartography should be of a different nature from that applied to analog maps. Users must be made knowledgeable enough about the cartographic processes to be able to assess the true quality of the available data and their created product. How then, in the digital environment should quality be described? The answer is by including information about the data quality in the required metadata of a digital spatial data file. Such metadata must be more inclusive than mere statements that a data set exceeds a given statistical value (most often related to positional accuracy). The metadata must contain descriptive statistics of all the features on the map, and digital data users must learn how to use or assess these metadata statements.

TOWARDS A SPATIAL DATA QUALITY STANDARD

In developing a standard for the exchange of digital spatial data, a committee of professional cartographers in the United States during the 1980’s working solely on the issue of digital spatial data quality, identified five facets of data quality that should be included with the metadata for a digital data file to enable the most effective use of the digital data file. In the digital environment it is recognized that in addition to positional accuracy, feature-attribute accuracy was important. Thus, a rather broadened consideration of data quality is proposed for use in digital cartography. The categories identified by this committee, and now formally part of the U.S. Federal Information Processing Standard 173, (FIPS 173), were positional accuracy, attribute accuracy, completeness, logical consistency, and lineage.

Condensed definitions of the five categories of quality are listed below. It is expected that each category shall contain references to temporal information as well.

LINEAGE, is a description of the source material from which the data were derived and the methods of derivation, including all transformations involved in producing the final digital files. The description shall include the dates of the source material and the dates of ancillary information used for update. The lineage portion shall also include reference to the specific control information used, and describe the mathematical transformations of coordinates used in each stage from the source material to the final product.

POSITIONAL ACCURACY, includes the degree of compliance with the spatial registration standard (latitude and longitude being the preferred standard). Descriptions of positional accuracy shall consider the quality of the final product after all transformations mentioned in the lineage portion. Reports on any tests of positional accuracy along with the date of such a test are to be included.

ATTRIBUTE ACCURACY, accuracy assessment for measures on a continuous scale shall be performed using procedures similar to those used for positional accuracy (i.e. providing a numerical estimate of expected discrepancies). Reports on any tests of attribute accuracy shall include the date of the test and the dates of the materials used.

COMPLETENESS, includes information about selection criteria, definitions used for the features, and other relevant mapping rules employed (e.g., thresholds such as minimum area or minimum width). The report describes the relationship between the objects represented and the abstract universe of all such objects. In particular, the report shall
describe the exhaustiveness of a set of features, and use standard definitions and interpretations if possible.

LOGICAL CONSISTENCY describes the fidelity of relationships encoded in the data structure of the digital spatial data. The report details the tests performed and the dates and results of the tests. Tests of valid attribute values, graphic data representation, and topology may be included.

In the digital cartographic environment it is important to include the values of the tests that were conducted on the individual digitized map features. This "truth in labelling" approach allows the user to make the necessary determination that a given data set is of sufficient quality to use in creating a cartographic product or decision document. The quality of data needed does vary with use. By expanding the statement of quality to include all features on the cartographic product, it is possible to have a better sense of the overall reliability and the variance of reliability within the product.

In the analog environment the cartographer would assess the product and determine qualitatively (and for some data quantitatively) if the product met a given level of quality. Today such an assessment is impossible because the user may be obtaining spatial data files from a variety of different producers and may have more than one digital source for a given feature. We are approaching an era where multiple sets of spatial and attribute data for the same feature will be available from different sources. It will be necessary to select the one data set that is of a quality most consistent with the other data to be portrayed on the map. We will neither want to use higher quality data (which may obscure the message of the map or slow down the production time, which translates into additional costs) nor lower quality data (which may make the map less useful and the decisions drawn from the map suspect). The most appropriate data set for the problem addressed should be selected, and the quality of that selection will depend on the quality of the metadata available to the knowledgeable user.

OTHER STATEMENTS ON SPATIAL DATA QUALITY

It is not only in the United States of America that this need for an expanded statement of spatial data quality for digital spatial data is recognized. One international group and several nations have recognized this need and have produced spatial data quality templates as described below. In addition, recently the Centre Europeen des Normalisation has established Technical Committee 287 to study and propose European standards for spatial data.

United Nations Environmental Programme, UNEP

The United Nations Environmental Program has developed a classification for a statement of spatial data quality. Only the first two levels of the classification will be illustrated in this discussion. The first level division is between data sets verified by relevant international sectoral or thematic scientific bodies. This level only covers the issues of attribute quality and lineage. The second level breakdown is into the rather subjective categories of "limited quality," "medium quality," and "highly accurate." Appendix 2 gives examples of Dataset Qualifier Categories using the first two levels of the UNEP data quality classification system taken directly from the referenced source.
France

French scientists have produced a document that "...clearly and accurately defines the...data, defines the parameters of data quality measurement or assessment in an accurate, complete and consistent manner, and presents the results of the data quality measurement or assessment. The statement may be extremely detailed and complete or reduced to the simplest possible form." Nine categories are defined for the reporting of spatial data quality: génetalogie (lineage); actualité (up-to-dateness); position planimétrique (plane positional accuracy); précision altimétrique (elevation accuracy); précision métrique (metric accuracy); exhaustivité (completeness); précision sémantique (semantic accuracy); cohérence logique (logical consistency); and qualité spécifique (specific quality). It is evident that the five categories in FIPS 173 correspond closely with categories in the EDIGeO system.

The up-to-dateness criteria is a date stamp. This information can be included in each of the five data quality categories of FIPS 173. Additional information is required by EDIGeO, however, which suggests that a date, marking the end of validity of the data, and a value, for the annual rate of change of the data, be specified. The criterion, metric accuracy, is defined as the mean value for a given type of measurement made on the exchanged data set. The type of measurement must be described; this criterion applies to all measurements other than plane positional and elevation measurements. The criterion, semantic accuracy, defines the number of objects, relations, or attributes which have been encoded correctly in accordance with specifications. And the final criterion, specific quality, allows the user to define a specific quality measurement for his data; this criterion is used for qualitative data and therefore a corresponding characteristic property must be specified.

Republic of South Africa

South African cartographers have adopted the five categories for the specification of digital spatial data quality as defined in FIPS 173 and have added a currency category, which is given as a date stamp and in some cases also a time stamp, indicating the date or time of the sampling or measurement of the source data. This is a required criterion.

United Kingdom (AGI)

The Standards Data Working Group of the Association for Geographic Information (AGI) has also endorsed similar categories for the specification of digital spatial data quality: "...across the many international standards initiatives, there is a measure of agreement that data quality criteria fall into five basic categories: lineage, positional accuracy, attribute accuracy, logical consistency, and completeness." It is vitally important for cartographers today to make available and to use quality-related information contained in the metadata that accompanies digital files of spatially-referenced data. This requirement for a statement of data quality at first glance may appear to be long and involved, but experience has already taught many users that a complete data quality statement is perhaps the most important part of the metadata attached to a spatial data file. The data themselves are, of course, important, but without a clear statement of the quality of the data, the data are not completely useful. Today, the importance attended to the data quality statement accompanying digital
spatial data is a good measure of a person's knowledge about the use of digital spatial data. It can be predicted that a user who pays close attention to the data quality statement accompanying the spatial data that are used to create a cartographic product will produce a superior product. This will only increase in the future.

ROLES OF THE ICA

The International Cartographic Association established a Commission on Spatial Data Quality in 1990. This Commission is in the process of developing a draft of a standardized statement of the data quality for spatial data files as called for in the Terms of Reference of this Commission.

The Commission will develop, document, and publish criteria and methodology for assessing the quality of digital spatial data sets.

1. Develop and document a comprehensive set of data quality criteria.
2. Develop and document a standardized rating scheme against those criteria.
3. Develop a methodology for data quality testing.

From the various studies mentioned in this paper, it appears there is common agreement on the elements of data quality criteria. The next step, the development of a rating scheme, is just beginning. The Commission would welcome additional inquiries and ideas into its deliberations.

SUMMARY

One might speculate that eventually all digital data collected will be of sufficient quality to satisfy the needs for more than 90% of its eventual uses. However, the lineage and currency components will still be necessary and specific to each and every cartographic feature that may appear in a visualization. Already there appears to be general agreement on some of the categories of information that a spatial data quality standard should require; indeed, from the brief statements outlined above, several categories appear repeatedly. The research to develop rating and testing methodologies is now underway by the ICA Commission on Spatial Data Quality.
REFERENCES


APPENDIX 1

UNITED STATES NATIONAL MAP ACCURACY STANDARDS

With a view to the utmost economy and expedition in producing maps which fulfill not only the broad needs for standard or principal maps, but also the reasonable particular needs of individual agencies, standards of accuracy for published maps are defined as follows:

1. **Horizontal accuracy.** For maps on publication scales larger than 1:20,000, not more than 10 percent of the points tested shall be in error by more than 1/30 inch, measured on the publication scale; for maps on publication scales of 1:20,000 or smaller, 1/50 inch. These limits of accuracy shall apply in all cases to positions of well-defined points only. Well-defined points are those that are easily visible or recoverable on the ground, such as the following: monuments, or markers, such as bench marks, property boundary monuments; intersections of roads, railroads, etc.; corners of large buildings or structures (or center points of small buildings); etc. In general what is well defined will also be determined by what is plottable on the scale of the map within 1/100 inch. Thus while the intersection of two road or property lines meeting at right angles would come within a sensible interpretation, identification of the intersection of such lines meeting at an acute angle would obviously not be practicable within 1/100 inch. Similarly, features not identifiable upon the ground within close limits are not to be considered as test points within the limits quoted, even though their positions may be scaled closely upon the map. In this class would come timber lines, soil boundaries, etc.

2. **Vertical accuracy,** as applied to contour maps on all publication scales, shall be such that not more than 10 percent of the elevations tested shall be in error more than one-half the contour interval. In checking elevations taken from the map, the apparent vertical error may be decreased by assuming a horizontal displacement within the permissible horizontal error for a map of that scale.

3. **The accuracy of any map may be tested** by comparing the positions of points whose locations or elevations are shown upon it with corresponding positions as determined by surveys of a higher accuracy. Tests shall be made by the producing agency, which shall also determine which of its maps are to be tested, and the extent of such testing.

4. **Published maps meeting these accuracy requirements** shall note this fact on their legends, as follows: "This map complies with National Map Accuracy Standards."

5. **Published maps whose errors exceed those aforesaid** shall omit from their legends all mention of standard accuracy.

6. When a published map is a considerable enlargement of a map drawing (manuscript) or of a published map, that fact shall be stated in the legend. For example, "This map is an enlargement of a 1:20,000-scale map drawing," or "This map is an enlargement of a 1:24,000-scale published map."
7. To facilitate ready interchange and use of basic information for map
construction among all Federal mapmaking agencies, manuscript maps and published
maps, wherever economically feasible and consistent with the uses to which the map is to
be put, shall conform to latitude and longitude boundaries, being 15 minutes of latitude
and longitude, or 7.5 minutes, or 3-3/4 minutes in size.

Last revised June 17, 1947 by U.S. Bureau of the Budget

APPENDIX 2

EXAMPLES OF DATASET QUALIFIER CATEGORIES
USING TWO LEVELS OF UNEP CLASSIFICATION SYSTEM

Class A-1 Data sets (or outputs derived from such data sets) validated and verified by
relevant scientific groups; considered to be highly accurate, precise and (if
relevant) timely.

Class A-2 Data sets (or outputs derived from such data sets) validated and verified by
relevant scientific groups; considered to be of medium quality, restricted, for
example, by limitations of geographic range, heterogeneity in quality, or
being out of date.

Class A-3 Data sets (or outputs derived from such data sets) which have been
examined by relevant scientific groups; considered to be of limited quality
but important enough to be used as 'best available' with qualifications and
the understanding that effort must be made to improve them.

Class B-1 Data sets (or outputs derived from such data sets), unverified or under
question by the relevant scientific group; believed by UNEP or consultants
to be acceptable with qualification, i.e. 'the best we have for the time
being'.

Class B-2 Preliminary data sets; may or may not have accurate geo-referencing;
unverified or under question by the relevant scientific group; in majority of
cases should not be used for analytical purposes.
Exploratory analysis of error in geographic information systems

H. Veregin (Kent, USA)

Introduction

A frequent but often overlooked problem in map-based spatial analysis is the use of data for which the magnitude and spatial distribution of error is unknown. Although some maps conform to known accuracy standards (e.g., the US Geological Survey's National Map Accuracy Standard), many maps contain undocumented errors and many map producers do not consistently apply procedures for assessing map accuracy. Reliance on data of low quality can have serious repercussions in decision-making, especially if these data are used in policy enforcement and litigation, where data quality standards are high. The replacement of traditional cartographic techniques with GIS has exacerbated rather than diminished this problem.

One of the advantages of GIS is that it provides a convenient environment in which source map layers can be transformed to derive new map layers. GIS data transformation functions induce modifications in the error characteristics of source layers and pass these modified characteristics on to derived layers. This process is referred to as error propagation. As a result of error propagation, derived layers may contain significantly higher levels of error than the sources from which they were derived, making them unfit for certain applications.

The importance of data quality in GIS is reflected in the recent adoption by the National Institute of Standards and Technology (formerly the National Bureau of Standards) of a set of spatial data quality guidelines as a Federal Information Processing Standard (FIPS) to serve all segments of the US federal spatial data processing community. Other federal agencies in the US are also in the process of developing systems for storing and tracking data quality information, including the US Geological Survey and the US Environmental Protection Agency. Accuracy of spatial databases also formed the first of twelve research initiatives by the National Center for Geographic Information and Analysis (NCGIA). This initiative produced several key reports detailing current and knowledge and future research directions. Recent books and conferences also point to growing awareness of the importance of data quality issues.

Various authors have described methods of accuracy assessment in the context of spatial databases. Some of these methods have been employed in the design and development of automated systems for documenting errors in databases and propagating error through sequences of GIS transformation functions. Our previous research in this area has focused on the design and implementation of an automated system for modeling the propagation of one component of database error through several commonly-employed data transformation functions. This system is designed to track source layer errors as data transformation functions are applied, based on lineage relationships established between derived layers and the sources from which they are derived. As layers are passed through transformation functions, a graphical representation is constructed showing source layers, derived layers and the transformation functions that connect them. The error indices associated with each layer are also displayed showing the modification of index values (Fig. 1).
Figure 1: Screen dump from automated forest proportionation system showing example application to identify oak woodlands at risk from cattle grazing.
This system can be used to provide information about the accuracy of products derived through the course of GIS-based spatial analysis. This paper discusses an enhancement to the system to facilitate exploratory analysis of error propagation. This capability is based on inverse error propagation, whereby the error propagation model associated with a particular sequence of data transformation functions is inverted to express source layer accuracy in terms of the accuracy of the derived layer. The system can be used to assess the relative importance of different sources in terms of derived layer accuracy, explore the impacts of alternate scenarios for improving the accuracy of one or more sources, and devise optimal strategies for improving derived layer accuracy.

Error Propagation Modeling in GIS

In layer-based GIS, geographical data are organized as registered thematic or temporal overlays. Error propagation modeling in layer-based GIS is based on the ability to model the lineage relationships between source and derived layers. These lineage relationships provide information on input-output links between the layers in the database through the course of GIS processing. This information includes parent-child associations between layers and the data transformation functions used in the derivation of new layers.

Error propagation modeling depends on the availability of several types of information:

- error indices for each source layer;
- error propagation functions for each data transformation function; and
- ancillary information for each layer.

Error indices are numerical values describing the magnitude of a selected error component in a database. Error indices may be scalar quantities, matrices, mathematical functions, or even co-registered layers containing reliability information for different portions of the database. An error propagation function describes the ways in which error indices for a layer are transformed by a given data transformation function. The error propagation function modifies the error index attributed to a source layer and attributes this modified index to the derived layer. The way in which error propagation is modeled depends on ancillary information describing such factors as the spatial distribution and co-occurrence of error on different layers.

In Figure 1, error propagation modeling is applied in the context of a simple GIS application. This application is designed to identify areas in which Oak tree regeneration is at risk from cattle grazing activity. Areas of cattle grazing are defined as areas that are ranches or areas for which grazing permits have been obtained. In the context of this application, each derived layer can be defined as a function of the layer from which it was derived. For example, the reselect function is applied to the source layer LANDUSE to derive a layer called RANCHES, which depicts those locations for which the landuse is "ranch". The data flow is given by

\[
\text{RANCHES} = \text{reselect}(\text{LANDUSE})
\]

In this case the derived layer is based on one input layer. It is also possible to represent functions with two input layers. For example, the layer GRAZING is derived by applying a union function to two layers called RANCHES and PERMITS. This data flow can be written as

\[
\text{GRAZING} = \text{union}(\text{RANCHES}, \text{PERMITS})
\]

The flow of data through the entire application can likewise be modeled as a nested set of data transformation functions operating on source layers, i.e.,

\[
\text{AT_RISK} = \text{intersect}(\text{union}(\text{reselect}(\text{LANDUSE}), \text{PERMITS}), \text{reselect}(\text{VEGETATION}))
\]
Note that this approach allows for the derived layer to be described as a function of the source layers, such that intermediate derived layers can be ignored.

Modeling the propagation of error from sources to derived layers is achieved in a similar manner by nesting a set of error propagation functions. In this case, it is the error indices attributed the sources, rather than the sources themselves, that are propagated forward through the sequence of data transformation functions. Numerous error indices have been developed to describe spatial error (also known as positional or locational error), thematic error (also known as attribute error) and temporal error. To model thematic error for categorical data, it is common to construct a classification error matrix showing a cross-tabulation of the actual and estimated thematic classes for a sample of locations. Element $c_{ij}$ in the matrix is the number of sample locations assigned to class $i$ that actually belong to class $j$. An error index called $PCC$ (proportion correctly classified) can then be defined as the trace of this matrix (i.e., the sum of all $c_{ij}$ where $i = j$) divided by the number of sampled locations. If the sample has been obtained randomly then $PCC$ may be interpreted as the probability that a location selected at random is correctly classified.

Propagation of the $PCC$ index through the application described above mirrors the flow of data through the application. The $PCC$ of the derived layer $AT\_RISK$ is determined using the $PCC$ of each of the three sources and error propagation functions ($epf$) that modify the $PCC$ index appropriately for each of the GIS functions employed in the application. Thus,

$$PCC_{AT\_RISK} = epf\_intersect\_pcc(epf\_union\_pcc(epf\_reselect\_pcc(PCC\_LANDUSE), PCC\_PERMITS), epf\_reselect\_pcc (PCC\_VEGETATION))$$

Note that each error propagation function is defined in terms of a given error index (in this case, $PCC$) and a given data transformation function (e.g., reselect, union, intersect). In fact, a variety of error propagation functions might be applied to each combination of error index and data transformation function, depending on the assumptions about error propagation mechanisms and the spatial distribution and co-occurrence of error. Consider error propagation for the $intersect$ function. Given

$$C = intersect\ (A, B)$$

then the $PCC$ of derived layer $C$ can be expressed as

$$PCC_C = epf\_intersect\_pcc(PCC_A, PCC_B)$$

Under the assumption of uncorrelated errors, $epf\_intersect\_pcc$ states that

$$PCC_C = PCC_A \times PCC_B$$

A more general form of the error propagation function is

$$PCC_C = PCC_A \times PCC[B\mid A]$$

where $PCC[B\mid A]$ is the conditional probability of observing a correct classification at a location on layer $B$ given that the location is correctly classified in layer $A$. The maximum value of this conditional probability is 1 (in which case $PCC_C = PCC_A$) and the minimum value is 0 (in which case $PCC_C = 0$). If $PCC[B\mid A] = PCC_B$ the error propagation function is identical to the uncorrelated case defined above. If $1 \leq PCC[B\mid A] < PCC_B$ then correctly classified locations tend to co-occur spatially and the accuracy of layer $C$ will be higher than for the uncorrelated case. If $PCC_B < PCC[B\mid A] \leq 0$ then the correctly classified locations tend not to co-occur spatially and the accuracy of layer $C$ will be lower than for the uncorrelated case.
Inverse Error Propagation

Propagated error indices are useful for assessing the reliability of data derived through the course of GIS-based spatial analysis. In many cases, it is also desirable to be able to improve the quality of these derived data. For example, one might wish to determine the increase in derived layer accuracy that would result from a unit increase in the accuracy of any source layer. This capability would make it possible to prioritize source layers in terms of their role in improving the accuracy of a derived layer.

This capability is based on the ability to write the application in inverse functional form. Given a data flow in which layer B is derived from layer A using the reselect function, i.e.,

\[ B = \text{reselect}(A) \]

then the inverse form is

\[ A = \text{reselect}^{-1}(B) \]

The error propagation function for reselect for the PCC index is defined as

\[ \text{PCC}_B = \text{epf_reselect_pcc} (\text{PCC}_A) \]

where \( \text{epf_reselect_pcc} \) states that

\[ \text{PCC}_B = \text{PCC}_A + K (1 - \text{PCC}_A) \]

where

\[ K = \frac{(r (r - 1) + (k - r) (k - r - 1))}{(k (k - 1))} \]

where \( k \) is the number of classes in the source layer and \( r \) is the number of classes collapsed to form the reselected class.

The inverse error propagation function is defined as

\[ \text{PCC}_A = \text{epf_reselect_pcc}^{-1} (\text{PCC}_B) \]

where \( \text{epf_reselect_pcc}^{-1} \) states that

\[ \text{PCC}_A = \frac{(\text{PCC}_B - K)}{(1 - K)} \]

A more complicated case exists when layer C is derived from two source layers A and B, e.g.,

\[ C = \text{intersect} (A, B) \]

In this case the inverse form is defined as

\[ A = \text{intersect}^{-1} (C, B) \]

Note that there is also a second inverse form, i.e.,

\[ B = \text{intersect}^{-1} (C, A) \]

The PCC of derived layer C is expressed as
The inverse error propagation function (for the first inverse form described above) is defined as

\[ PCC_A = \text{epf}_\text{intersect}_\text{pcc}^{-1}(PCC_C, PCC_B) \]

In the uncorrelated case, \( \text{epf}_\text{intersect}_\text{pcc} \) states that

\[ PCC_C = PCC_A \times PCC_B \]

and thus \( \text{epf}_\text{intersect}_\text{pcc}^{-1} \) states that

\[ PCC_A = \frac{PCC_C}{PCC_B} \]

This function can be used to compute the required PCC for source layer A in order to achieve a given PCC for derived layer C. However, as layer B is also a source layer, the PCC of this layer must be fixed at some arbitrary value. Without fixing the accuracy of B, an infinite number of solutions exists.

One can likewise define the inverse error propagation function for the \text{union} function, as follows. Given

\[ C = \text{union}(A, B) \]

and

\[ PCC_C = \text{epf}_\text{union}_\text{pcc}(PCC_A, PCC_B) \]

where \( \text{epf}_\text{union}_\text{pcc} \) states that

\[ PCC_C = 1 - (1 - PCC_A)(1 - PCC_B) \]

then the inverse spatial data transformation function is

\[ A = \text{union}^{-1}(C, B) \]

and the inverse error propagation function is

\[ PCC_A = \text{epf}_\text{union}_\text{pcc}^{-1}(PCC_C, PCC_B) \]

where \( \text{epf}_\text{union}_\text{pcc}^{-1} \) states that

\[ PCC_A = \frac{1 - (1 - PCC_C)}{1 - PCC_B} \]

These inverse error propagation functions can be nested to define the inverse propagation of error through an entire GIS application. Consider the example application shown in Figure 1. Assume that one is interested in computing how accurate input layer LANDUSE would need to be in order for derived layer AT_RISK to achieve a given level of accuracy. In this case, the accuracies of input layers PERMITS and VEGETATION will be held constant.

First, the inverse set of data transformation functions is written as

\[ \text{LANDUSE} = \text{reselect}^{-1}(\text{union}^{-1}(\text{intersect}^{-1}(\text{AT_RISK}, \text{reselect}(\text{VEGETATION})), \text{PERMITS})) \]
Next the inverse error propagation functions can be written as

$$\text{PCCLANDUSE} = \text{epf_reselect}_{\text{pcc}}^{-1}(\text{epf}_{\text{union}}_{\text{pcc}}^{-1}(\text{epf}_{\text{intersect}}_{\text{pcc}}^{-1}(\text{PCC}_{\text{AT_RISK}}, \text{epf}_{\text{reselect}}_{\text{pcc}}(\text{PCC}_{\text{VEGETATION}}), \text{PCC}_{\text{PERMITS}})))$$

**Inverse Error Propagation as an Optimization Tool**

In the example application, the PCC of LANDUSE is 0.75 ($k = 5$), the PCC of PERMITS is 0.9, and the PCC of VEGETATION is 0.7 ($k = 4$). The computed PCC of AT_RISK is approximately 0.84. Consider an example in which the desired PCC of AT_RISK is 0.94 (an increase of 0.1). Application of the inverse error propagation functions as described above indicates that in this case LANDUSE must have a PCC of approximately 0.87, assuming that the PCCs of the other sources (PERMITS and VEGETATION) are held constant. The same approach can be used to define the required accuracy of PERMITS (holding LANDUSE and VEGETATION constant) and VEGETATION (holding LANDUSE and PERMITS constant).

The required PCC of each source layer can also be computed for a range of PCC values for AT_RISK. This is illustrated in Figure 2, which shows the relationship between changes in the PCC of each input layer and the resulting PCC of AT_RISK. The lines for the three sources intersect at the computed PCC value for AT_RISK in the example application (i.e., 0.84). The layer with the steepest slope (i.e., VEGETATION) yields the highest rate of increase in the PCC of AT_RISK. Assuming that the cost of improving accuracy is the same for all three data sources, the accuracy of AT_RISK can most economically be improved by increasing the accuracy of VEGETATION.

![Figure 2. Relationship between change in PCC of source layers and change in PCC of AT_RISK.](image)

It is also possible to identify the solution space for all source layers yielding a specified accuracy level for the derived layer AT_RISK. Given three source layers the solution space can be represented as a surface for which the x, y and z coordinates are defined by the PCC values for the three source layers. Any point on the surface gives a combination of PCC values for the source layers that yields the required PCC of the derived layer. Figure 3 shows such a surface for various levels of required accuracy. Note that the most critical layer in deriving the desired accuracy is VEGETATION. Either LANDUSE or PERMITS can be relatively inaccurate without significantly impacting AT_RISK, because these two layers are passed through the union function (which tends to inflate accuracy).
Figure 3. Solution space for LANDUSE, PERMITS and VEGETATION for PCC for AT_RISK of 0.9 (top) and 0.95 (bottom).
Inverse error propagation also allows the user to define optimal strategies for increasing derived data accuracy by improving or replacing one or more sources. If the cost of improving or replacing all sources is equal, then the source yielding the greatest increase in derived layer accuracy per unit increase in source layer accuracy is the logical choice (e.g., the VEGETATION layer, as shown in Fig. 2). In practice, the cost of improving accuracy will vary as a function of the nature of the data contained in the source, the availability and cost of data, the costs of performing ground-based accuracy assessments, and other factors. Various strategies might be pursued. One strategy is to improve the accuracy of the source that yields the required accuracy at the lowest cost. An alternate is to improve the accuracy of the source that yields the greatest increase in derived layer accuracy per unit of cost. The latter strategy can be represented in the form of a graph similar to Figure 2, in which the vertical axis represents the change in accuracy of the derived layer divided by the cost of a unit change in accuracy of each source. An example is shown in Figure 4, in which the relative costs of improving the accuracy of the three sources LANDUSE, PERMITS and VEGETATION are expressed by the ratio 0.1:0.5:1 (i.e., the cost of a unit improvement in the accuracy of VEGETATION is ten times that for LANDUSE and two times that for PERMITS). As Figure 4 shows, LANDUSE yields the greatest return in derived layer accuracy per unit investment.

![Figure 4](image-url)

**Figure 4.** Relationship between change in PCC of source layers and change in PCC of AT_RISK, weighted by the cost of improving source accuracy.

**Conclusion**

By providing a mechanism for users to examine the implications of variations in source layer accuracy, inverse error propagation allows users to explore the error characteristics of their databases and devise optimal strategies for improving derived data products. Inverse error propagation functionality is currently being incorporated within the lineage-based system for error propagation we have described elsewhere.\(^\text{10,11,12}\) The inverse error propagation module is designed to interface with the lineage component of the system. This will require information transfer of several forms, including input of lineage information to the error propagation module (error indices and ancillary data attributed to sources, and parent-child relationships between layers encoded as semantic links), and output of transformed error
indices from the error propagation module to the lineage component. In addition, transformed error indices must be output to the graphical user interface component of the system for display and manipulation. Exploratory analysis of error will be most effective if users can manipulate iconic representations of layers and error indices directly through the graphical user interface. As icons are manipulated, the display is then dynamically updated to illustrate the implications for all sources and derived layers.

References


ABSTRACT
The shift from an almost totally manual production of maps and other geographic products to an almost totally automated production brought about a number of new products, many with new and unanticipated capabilities. The Map Image Metafile (MIM) representation of maps is one of the developments resulting from the use of automation to produce maps. MIM's are the means now used for exchanging digital map files between the various hardware and software environments in the United States Census Bureau's publication map production system, and MIM's are the digital form in which many of the the 1990 publication maps will be released to the public. The extremely simple, yet extensible structure of MIM's makes them valuable for most cartographic products. The ability of a MIM to fully describe a map image in a human readable form provides a way to faithfully re-produce the map image in the future, regardless of changes in hardware and software.

MIM's are not a spatial data transfer means. Indeed, spatial relationships are not part of the structure such as they are in the Spatial Data Transfer Format (SDTS). However, MIM's are not simply binary representations of map images or image production commands such as in the Computer Graphics Metafile (CGM). MIM's do have a level of intelligence that allows the user to either faithfully reproduce the original map image, or selectively display or modify the image to suit their own needs. Except for the way a country encodes its language into the ASCII character set, and the terms used by a country to describe the various parts of a map, the MIM structure is machine and software independent. Because of the temporal and technological independence provided by MIM's, they promise to have a significant impact on the future of cartography.

INTRODUCTION
Historically the U. S. Bureau of the Census has made the maps to be printed and bound with the data volumes or issued separately. The excellent quality of the printed maps provided the public with inexpensive, readily available, and flexible research tools. They were easy to duplicate either photographically or xerographically. The durability of the paper made them useful both as working bases and as archived documents.
The automation of map production at the Census Bureau has changed things dramatically. The number of map sheets produced has increased ten-fold or more. For example, over 70,000 different sheets were plotted electrostatically for just one map type! And there were several different map types. No longer could the Census Bureau provide the public with maps in the that they had done in the past. The shear volume presented significant problems to individuals, companies, other government agencies, academic institutions and libraries. They are neither prepared to receive, store, and manage such a large number of map sheets, nor to pay for them from limited budgets. Also, the somewhat lower quality of paper of the plotted maps verses printed maps affects durability during use and archiving. Further, the contrast and sharpness of the plotted maps are much more difficult to reproduce.

The initial plan was to duplicate the plot tapes and sell them to the public who could then make as many map copies as desired. However, the plot files were produced in a format for a specific set of hardware. Thus, anyone purchasing the plot files would have to also have access to the same type of plotter.

Producing publication quality maps is difficult because the programs that generate the map images (plot files) cannot account for all the special conditions that occur in a map. Text placement, with or without overlap, position of a north-arrow within image area for different shapes of geography, editing features that were erroneously coded in the TIGER/File are the examples of tasks that cannot presently be economically and efficiently performed by batch (non-interactive) mapping programs. This type of task is most efficiently done interactively. Therefore, some way had to be found to edit and manipulate the image.

At the Census Bureau, the map generation process is made more difficult by a lack of stability in the production environment. The main computer used for TIGER processing has changed three times in the last seven years. Also the plot file generation software has had to be developed in three different environments (UniSys Exec 8, VAX VMS, and UNIX on a variety of workstations).

Given that the map production environment in terms of hardware, software and operating procedures will continue to change, there is yet another major problem: How should map images be stored for future reproduction and use? People will need to be able to reproduce the census maps for years to come. For the near future, the plot files can be used to replot maps as needed. However, plot files may become unreadable or the plotters themselves will need replacing. Storing the map images as hardware specific plot files and forcing ever burdensome backwards compatibility on new plotters is neither desirable nor realistically possible in the term of decades of time.

Capturing the map image as a bit-image at the scale of map creation is well within the technology of today, but is unsatisfactory for several reasons. First, displaying the bit-image on the screen of a graphics terminal severely limits the amount of the map visible since most screens today cannot display more than about 1,000 by 1,000 dots. Thus, at 200 dots-per-inch (dpi), only about five inches of map can be seen at one time, though the map may be 24 by 30 inches or larger. Further, a bit-image lacks intelligence to allow for
selected display of say, boundaries without text, or specific boundaries in red, all others in green. Finally, an uncompressed bit-image file of a 24 by 30, 200 dpi map would take 3,600,000 bytes. Compressing the file means that if the information on how to uncompress becomes lost, the map could only be retrieved with great difficulty.

Other storage forms were investigated, including, but not limited to the American National Standard computer graphics metafile (CGM). This is a powerful and clear standard in which many valuable ideas were found. However, the CGM standard, as well as several other of the standards investigated, suffer from the same difficulty of use. They all produce a file that is very close to hardware levels of image data representation and require knowledge of some external facts to read the file. Admittedly, they could have been forced to perform for census map image operations, but the shortcoming are serious when viewed from a future position.

It is difficult, if not impossible to predict the future uses of the 1990 census maps. If the past if any indication, the users will need ready access to the maps and ease of manipulation to not only show census data on them, but user data as well. Past maps were fully comprehensible, free standing documents of great use; digital files of map images should be no less.

Clearly map production at the Census Bureau needed a simple and easy-to-implement solution to the many problems. A solution that allows a map image to be generated in whatever computer environment available, and edited, manipulated, and plotted in any other hardware and software environment, and in a format that is readable in years to come. The answer was a simple image transfer format that would not suffer under hardware limitations such as how bits are stored by different machines and so forth. The solution developed is a form of meta4 map image file, known in the U. S. Census Bureau as the map image metafile (MIM).

**WHAT IS MIM?**

A Map Image Metafile is a flat file of ASCII characters that is a self-documenting, full-image description of a census map. This means several things: The contents of a MIM file are readable without special software. In fully expanded form, no external documentation is needed to describe the data fields. Differences in the bit-structure of the real, integer, and character data types between computers and programming languages are of little consequence. And, most importantly, the information in the file can be used to faithfully reproduce a copy of the original census map.

There are two basic characteristics of a MIM, it's object nature and flat ASCII format. The map is considered to be an object that can be described in terms of other cartographic objects such as legends, north arrows, etc. These in turn can be described in terms of other cartographic objects or cartographic primitives such as text, lines, and areas; each with it's own attributes. The flat file format removes the need to interpret hardware codes while the object structure allows for showing relationships between graphic elements. This is very powerful, because whole portions of a map image may be (1) readily identified within the file and (2) treated as a single unit when interactively manipulating the image. Neither of these are generally true of ordinary plot files.
The production of a map is generally a multi-phased operation. This is particularly true of a publication map. MIM’s serve as the exchange format between the various production steps.

Finally, MIM’s provide a simple means for archiving census map images in a digital form. Their form and content provides admirably for future reproduction of the original census map image. A MIM can either contain the full description of a map image or when many images are grouped on a single storage medium, MIM’s allow for sharing of common file information.

**WHAT MIM IS NOT**

A MIM is not a compact binary transfer format. Part of the value of a MIM comes from its almost universally readable flat, ASCII exchange format, and ASCII is not a compact storage form.

A MIM is not intended to be used directly by mapping operations programs. MIMs almost always will be read into an internal structure that is efficient for a given hardware and software installation. The relative ease of reading a MIM is a strength that permits the use of many different types of hardware and software for display and processing.

MIM’s are not designed to be randomly accessed. MIM information is communicated by command and data sequence, much as one reads a letter or novel. Since MIM information is sequential, when an attribute such as line width is set, it stays in force for all succeeding lines plotted.

MIM information is not stored in fixed field position or fixed record size. In some cases a data field is generally only wide enough to contain the value in it. Similarly, many arguments for the commands are optional, so the record size will vary according to the number of arguments actually used.

Finally, MIM commands are not designed to account for every possible special condition in every census map. Even if all the possible exceptions or special conditions could be enumerated and described, to do so would weight the system down with specialized commands. Different map symbols also are not pre-established in the MIM definitions. Rather, the philosophy is to treat special conditions through the existing commands.

**KEEP IT SIMPLE AND SUPPLE**

The guiding principle during development of the MIM file structure and content has been for simplicity in form and content and supple in structure. A problem with defining strict structures and standards is that they tend to evolve into highly complex encoding schemes as technological and/or social changes occur and attempts are made to accommodate them. There is always the tendency, in the name of management ease, to enforce formalized structure in situations when it is at best inappropriate, and at worse, totally repressive to creativity and advancement.

Simplicity is maintained by defining the absolute minimum number of commands and providing for extensibility. This minimal number can be determined by only allowing commands if they (1) cannot be constructed efficiently and effectively out of existing meta-commands, (2) apply to most map images, and (3) are neither hardware, software, nor operationally...
Specific. Some degree of common sense must apply when deciding on a new command, but in general, the fewer the better.

Suppleness is maintained by providing a way to accommodate special situations and changing conditions. The MIM achieves this by providing for extensibility to the commands. Suppleness implies a flexibility and moldability but not a change in the basic nature of the underlying structure. In mathematical terms, a torus shaped piece of clay can be molded into a coffee cup without losing the underlying nature of the torus.

**COMMAND PRIMITIVES**

MIMs are constructed from a series of primitive commands. A detailed description of the commands is given in the U.S. Census Bureau's Map Image Metafile documentation. Meta-commands are divided into three categories, (1) required commands, (2) basic class and attribute commands, and (3) grouping or extending commands.

**Required commands** These commands must appear in all metafiles. There are three commands in this category: *int, *cls, and *msz. The *int and the *cls command bracket a map image. Any command found outside these two are not considered as part of the image. The *msz command defines the fundamental design parameters of the meta image. All map images have been designed to give a certain "look" to the map product. This means that they are designed for a certain size and to be produced on a certain output device at a given resolution. The *msz command sets the visual size and resolution parameters.

**Basic class, attribute, and helper commands** These commands make up the bulk of the metafile commands. They define the look-and-feel of the map image. There are three class commands, *str, *pgX, and *vtx. The *str command defines a line in terms of its shape points. The *pgX command defines a polygon in terms of its shape points. The *vtx command defines a line of text to be plotted. While text can be considered to consists of lines and areas, for MIM purposes it is considered as a class itself. Therefore, these three commands are the fundamental cartographic entities of all maps images.

Class commands optionally can have alphanumeric codes that identify the geographic entity being represented. This provides a way to be able to work with the entities as say, state boundaries, water fill, city names, and so forth without having to work with all strings or all text.

The way an entity appears in the final map depends upon the attributes of the entity. A line has a width, a type such as solid, dashed, and so forth, and other attributes. Text have fonts. Polygons have pattern fill. There are commands that are used to set attribute conditions. Each attribute setting commands must be proceeded by a defining command. For example, the *dpa command defines the characteristics of a pattern that is then referred to by the line or fill pattern commands. Table I shows the relationships between cartographic entities, attribute setting, and attribute defining commands.

The helper commands, *cmt and *inc, are there to make metafiles easier to build and understand. The *cmt command allows comments to be included in the file. They can be used to elaborate on the map image, "non-standard"
things in the file such as a new scheme for geocoding, map projection, map scale, and so forth.

<table>
<thead>
<tr>
<th>Strings</th>
<th>Polygons</th>
<th>Text</th>
<th>Defined</th>
</tr>
</thead>
<tbody>
<tr>
<td>*str</td>
<td>*pgX</td>
<td>*txt</td>
<td>by</td>
</tr>
<tr>
<td>Fill color</td>
<td>fcs</td>
<td>fcp</td>
<td>fcv</td>
</tr>
<tr>
<td>Fill pattern</td>
<td>n/a</td>
<td>fpp</td>
<td>fpv</td>
</tr>
<tr>
<td>Line color</td>
<td>lcs</td>
<td>lcp</td>
<td>lcv</td>
</tr>
<tr>
<td>Line pattern</td>
<td>lps</td>
<td>lpp</td>
<td>lpv</td>
</tr>
<tr>
<td>Line type</td>
<td>lts</td>
<td>ltp</td>
<td>ltv</td>
</tr>
<tr>
<td>Line weight</td>
<td>lws</td>
<td>lwp</td>
<td>lwv</td>
</tr>
<tr>
<td>Text font</td>
<td>n/a</td>
<td>n/a</td>
<td>sft</td>
</tr>
</tbody>
</table>

The *inc command is used for meta commands common to many maps so they do not have to reside in every metafile. A separate file of common commands can be produced and included in the map image file when appropriate. For example, a set of pattern descriptions that are common to a map series can be put in a file, then linked to each map image file by the *inc command.

Combining and extending commands These commands provide a means for grouping cartographic objects and for extending the number and type of commands. They define the beginning and end of a group of commands that produce either a given map element directly or thru a process.

The *bef and *enf commands denote the beginning and end of a collection of meta-commands that should be treated as an object. These can be used to define things like a legend, which is a group of text, lines, polygons, and even other *bef - *enf objects. Now the object components can be treated as a single object and can be manipulated as needed. To ease creation of "sets" of entities encapsulated within a *bef and relative to each other, the *rel is used to indicate that all coordinates within the *bef are relative to the map inches position given in the *rel command. The *rel command is only active within the *bef in which it appears.

The commands that provide extensibility to the basic command set are the *dep and *enp commands. They denote the beginning and end of a process to be performed. For example, the process may be to display a triangle in a specific location and at a given angle. Using a *dep, a triangle symbol need only be defined once. Then adding arguments for x,y-position and angle creates a new command, say a *triangle command, which will produce a triangle each time it is used. The symbol can be produced as many times as needed on the image, but without having to describe all the bounding coordinates, the fill pattern, and so forth at each instance of the triangle.

In addition to all other MIM commands, the following are examples of operators that can be used within a *dep defined procedure:

<table>
<thead>
<tr>
<th>Operation type</th>
<th>Action denoted</th>
</tr>
</thead>
<tbody>
<tr>
<td>+, -, *, /, sqrt</td>
<td>Arithmetic</td>
</tr>
<tr>
<td>cos, sin, tan, atn</td>
<td>Trigonometric</td>
</tr>
<tr>
<td>=, push, pop</td>
<td>Assignment, stack</td>
</tr>
<tr>
<td>&lt;, &gt;, ==, !=</td>
<td>Logical</td>
</tr>
<tr>
<td>integer, float, char</td>
<td>Variable typing</td>
</tr>
<tr>
<td>if...[else]...endif, while...</td>
<td>Flow control</td>
</tr>
</tbody>
</table>
A *bef does not expect arguments to modify or control instantiation while the *dep does. Put another way, *bef’s are singular expressions of the contents of the *bef while *dep is instantiated as many times as needed, i.e., a legend would be a *bef while a school symbol would be a *dep. One occurs once; the other occurs as many times as needed. The *bef-object is a description of a specific map element. It has no ability to reproduce itself elsewhere on the map, or change its characteristics. These must be done to it externally. A *dep-extension is an abstract map element that will act on itself or other map components to reproduce the image. Only the instantiation of the extension can be manipulated, not the extension itself.

RULES FOR MIM CONSTRUCTION
A MIM is constructed much as one would write a verbal description of a map image. The rules for MIM construction are predicated upon the use of information in the sequence presented.
1. A map image starts with an *int and must close with a *cls.
2. All map images must contain one and only one *msz command. When placing the *msz command, it must occur before of any commands that include coordinates or patterns.
3. Defaults cannot be assumed. The MIM should always start by setting the attributes and similar values. Note, there are two exceptions to the "no defaults" rule. As generally found throughout the industry, patterns 0 and 1 are already defined as all bits off and all bits on respectively. Also, line type 0 means a solid line.
4. *inc’s must occur before their information is used. Includes can be placed anywhere in the MIM, but before any reference to it’s contents.
5. Attribute values continue to apply. Once an attribute value has been set, it is applied to all subsequent commands as appropriate.
6. Use of optional entity geoID argument. The creator of a given map design is free to establish any geoID coding scheme desired. When a coding scheme is used, the codes and their meanings should be placed at the top of the file and within comment commands.

NOTE: geoID’s are limited to a character string of not more than 31 characters.

A geoID must be one and only one argument. For example:

"Green Briar Springs" Good Multiple words enclosed by quotes.
Green Briar Springs Bad No quotes, therefore three arguments

Comments help provide information to users of the MIM about what codes are to be encountered and how they should be interpreted. It is the responsibility of the program developer of the map specific software to produce programs that read this map file and handle the geocodes appropriate to that map. There is no general procedure for handling geocode information, it is all map specific. An example may be,

*cmt "geoID PONDWATER denotes a small water body"

NOTE: The manner in which a feature is displayed depends upon the display program and hardware. In no way is the display constrained by the color, line type, and so forth that are stated in the MIM. The
MIM commands for color, line type, and so forth are followed, if and only if one wishes to faithfully reproduce the original map image.

7. **Use of optional entity state argument.** There are five states in which an entity may exist in a MIM.

<table>
<thead>
<tr>
<th>State</th>
<th>Denoted by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>Absence of state argument</td>
</tr>
<tr>
<td>Deleted</td>
<td>Xdel argument</td>
</tr>
<tr>
<td>Changed</td>
<td>Xchg argument</td>
</tr>
<tr>
<td>Added</td>
<td>Xadd argument</td>
</tr>
<tr>
<td>Reference</td>
<td>Xref argument</td>
</tr>
</tbody>
</table>

These arguments are provided so that entities modified by an interactive program may be flagged. For example, suppose one were investigating how a human operator designs a map. After a computer program has produced its version of the map image, a human operator may use an interactive program and modify the image. The program flags each modified entity with one of the state flags and retains the copy of the computer-generated entity. After a series of map images have been processed, the resulting MIMs can be examined to compare the human actions with the original image file to derive an understanding of what rules were followed. This could lead to these rules being incorporated into future mapping programs.

The Xref argument is used to denote a feature or entity that is included in the MIM as an aid to processing. It denotes a feature or entity that is not to be displayed as part of the image, but is to be used to aid in creating the image. This is valuable in such operations as splitting text for stacking when the original text string needs to be retained. Another example is when providing a coordinate string for automated positioning of text or a polygon within which text is to be constrained. These entities are not part of the image; they are for reference only.

8. Portions of this documentation are in a MIM.doc file. The file is not necessary to use the MIM for reproducing a map, but it helps. It can be included with a *inc-command.

**COOL MAP COMPONENT DEFINITIONS**

Every map image consists of certain components such as legends, image area, graphics scale, and so forth. For purposes of describing these to a computer program, certain COOL5 conventions were adopted to use in MIMs. The following are selected examples of some COOL definitions used.

**Frame line** - this line defines size and shape of the area in which components of the map itself are displayed. Folio information such as page numbers, map title, and so forth may appear outside this frame -- or may not, depending upon the map design.

**Outer Image Line** - this defines the maximum envelop of a map image. Generally no parts of a map image are expected to go beyond this.

**Neatline (Map Image Area)** - this defines the limits of the area within a map in which an image will be displayed. There may be multiple map image areas per map page.

**Entity Latitude/Longitude Envelope** - this is the maximum and minimum latitude/longitude of the entity being displayed. These values relate to the
entity maximum/minimum map units envelope. This relationship allows the computation grid ticks, particularly when moving the entity's position within the image area.

**Entity Map Units Envelope** - this is the maximum and minimum map units of the entity being displayed. It is related to the entity latitude/longitude envelope and used during computation of grids.

The diagram in figure 1 graphically describes some COOL definitions.

![Diagram](image)

**Figure 1. Selected COOL cartographic definitions**

**PRODUCTION EXAMPLE OF MIM USE**

MIMs have served as the foundation for five projects at the Census Bureau and are planned for many more. Two recent examples are the maps for a Congressional District atlas and maps of urbanized areas to accompany published data. Over 1,000 publication quality maps were produced for the atlas and about 800 for the urbanized areas project. Figure 2 shows, in reduced form, an example of an urbanized area map.

It is important to note that each part of the map image can be manipulated separately. Not only can the text be manipulated as text, i.e., turned on or off, but it is even coded as to the type of geographic entity the text represents and, therefore, manipulated accordingly. Similarly, the shoreline "knows" that it is a shoreline and the neatline "knows" that it is a neatline. For example, the following code was taken from the MIM used to produce the map in Figure 2.

```plaintext
*cmt "geoID Census place codes are a 'p' followed by 7-digit code"
*cmt "geoID County codes are a 'C' followed by a 5-digit code"

*lwv 0.005
*sft RBHev.Mft B
*vtx 12.634 12.966 0.15 24.0 C12021
"COLLIER"
*sft RFTrip.Hft
*vtx 10.576 15.410 0.01 90.0 p1219325
"East Naples"
```
This example shows that each string of text "knows" not only its display attributes such as font or line weight, but what geographic entity it represents. Further, the definition of each code is included within the MIM.

**SUMMARY**

Map Image Metafiles (MIMs) developed from the efforts to use automation to produce maps at the U. S. Bureau of the Census. They are the answer to questions of how to lessen the impact of a constantly changing software, hardware, and operating systems environment. MIMs fully describe a map image in a human readable form; one that if faithfully, followed will
reproduce the map image in years to come. But MIMs are more than just plot files. They have a level of intelligence in the coding scheme so that the various cartographic entities "know" what kind of entity they are and, in some cases, how to reproduce themselves. These are new and powerful capabilities for cartographic files; capabilities that promise to have a significant impact on the future of cartography.

REFERENCES


1 The use of vendor names in association with hardware and/or software, is for describing the operating environment during map production and is in no way intended as an endorsement by the government.

2 Almost all census maps are created as black and white raster plots. Equating a raster to a bit, if the bit is on, a dot is plotted there, or if off, nothing is plotted.

3 The Federal Information Processing Spatial Data Transfer Standard is a data exchange standard that satisfies geographic data transfer needs very well. Indeed, the definitions of cartographic entities used in MIM were extracted from the SDTS.

4 The term meta was selected in the sense of beyond the mere image, transcending a given computer or display environment, and more comprehensive than a purely machine readable hardware image exchange format.

5 Several cartographic texts were referenced when developing these definitions. Among them, though not all of them, are Elements of Cartography by Robinson et al and Cartographic design and production by J. S. Keates.
Session 12

Atlas Cartography I

Chairman:
D. Pearce, International Cartographic (Mount Pleasant, AUS)
Regional, national and international atlases – towards an integrated system
S. Jaatinen (Helsinki, SF)

Introduction

Present time accomplishments in modern thematic cartography, here focused on the atlas cartography, are predominantly committed to cybernetic processing and technical problem areas. However, the ultimate aims of cartographic information distribution like the quality of the data treatment, the validity and perception of results, and the integration problems are neglected. There is thus a great danger, that the prevalent orientation in the field of thematic cartography has a too narrow range.

It seems as if neither the essentially scholarly character of the thematic cartography, nor the necessity of cooperation and integration is enough observed. One explanation might be the fact, that too little attention is paid to the fundamental features (except semantics) of cartographic information distribution. One has to admit, however, that even the scholarly basis of cartography is not a self-evident matter. On the contrary, behind the facade there is actually a discrete structure of both apparent and hidden objectives directing the scholarly selection, treatment and handling of the basic themes. These subjective preferences and predispositions are affecting also the processing and visual expression of thematic data cartographically 1.

Another problem area is the rational integration of atlas cartography in relation to the regional frames applied (area covered, hierarchical status of the area and utilized scales). The integration involves on one hand a predominantly technical question: how to adapt the methods (like choropleths) and symbols (like rasters, dots, conventional signs, etc.) to different but integrated settings of regional frames. This problem area is especially acute when the cartographic expression of data on maps of different scales will be conveyed by electronic devices like PC's as for instance the options for simultaneous comparison are here very limited.

On the other hand, and may be of even more essential importance, is the fact that the integrated expression of thematic data on maps of various scales (and regional frames) and different contexts (single maps, national atlases, thematic atlases and so on) ought also to be considered in terms of truthful content, logic, comparable structures and scholarly accuracy. This problem area of consistency has been dealt with in connection with the exceptionally large and thematically comprehensive fifth edition of Atlas of Finland, published since 1977 and finished this year.

585
Thus there was initiated in the early 1980's a specific project, the computer-assisted KALSKE-system, in order to create a framework for the semantic structure of the spatial data field covered in this and earlier editions of Atlas of Finland and to make it possible to comprehensively catalogue logically and to retrieve swiftly thematic maps published in Finland since the later part of the 20th century (both separate maps, and maps in books and atlases). Information about the KALSKE-system, now applied in most of the map archives in Finland, can be obtained through the University Library, (P.O. Box 15), 00014 University of Helsinki, Finland).

When different types of atlas works are integrated the ultimate aim must be to maintain a holistic structure of the thematic expression and an adequate standard of interpretational distinctness. The aspects of integration must also be discussed in an international dimension. This is obvious as the development of standards for both scholarly terms and cartographic concepts proceeds swiftly in all sectors of information gathering, processing and communication. The intention of this paper is to survey a few of these questions in order to stimulate further discussions and practical implementations.

Background

Thematic cartography with the intention to produce comprehensive atlases has had a variety of starting points. In the light of improved integration of atlas work on different levels, a number of diverse aspects can be mentioned. One is the purpose of the atlas work: education, thematic surveys, prognostic or planning intentions, and so on. Further, the content and form of atlases vary considerable and the same can be said of the typographic, cartographic and scholarly standards.

Finally there is the crucial question of representation media: on paper or in electronic form. Of ultimate importance are the selection of user orientation, the degree of generalization and other aspects of cognitive interpretation. The administrative organization of atlas work both on national and international levels should also be observed. A suggestive and preliminary analysis of the different aspects related to the frames, problem areas and prospects will be given below.

Frames and organisations

The editorial setting of thematic atlas work has been given different solutions both in time and place. Sometimes the responsibility has been carried by governmental institutions of great permanency, but occasionally by ad hoc committees with rather ephemeral mandates. National survey organisations, geographical research institutions and different planning bodies are examples of the first mentioned type of frames. On the other hand there has also been numerous examples where the thematic atlas work has adhered closely, but rather temporarily to the scholarly and academic realm of a country. A good example of this last-mentioned type are the 4 editions (1899, 1910, 1925 and 1960) of Atlas of Finland, whereas the fifth one has been realized as a joint venture between the Geographical Society of Finland and the National Board of Survey, thus representing a mixed type, also applied in connection with the new National Atlas of Sweden, which started to appear in 1990.
The organization of thematic cartography on different levels in a country is also a point to be considered. Very seldom atlas works on the regional level are rationally integrated with the corresponding tasks on the national level. Contrary to the obvious assumption, that governmental bodies, like ministries, central and regional offices, ought to be able to cooperate, one can often observe jealous competition of the right to publish thematic spatial information. Sometimes this results in conflicting or disparate expressions of the same themes or data. These types of detrimental situations must be replaced by more open attitudes, easier access to the primary data and positive criticism. A diverse representation of information is of course as such desirable as long as different approaches are allowed and taken as complementary for a holistic understanding of the themes.

Functional structures between primary data resources (like statistical offices), scholarly analysis of the themes, editorial processing and cartographic realisation have been organized in a variety of ways. Both the data processing and the cartographic accomplishment demand nowadays sophisticated and expensive investments. It is crucial to scrutinize the atlas work in a comprehensive fashion. Integration demands interaction of a number of different data producing and processing units. In many countries this has already been given a definite and cohesive form in national GIS programmes. From the point of thematic cartography this is not always enough, as the cartographic potential of most GIS programmes is not flexible or adequate enough to serve directly as a base for thematic cartography.

The integration must include more "open ends" or contact possibilities varying according to the themes in question. Finally it is worth to observe, that the verbal structure (symbols and their explanations, textual analysis of the maps and cartograms, and so on) of thematic cartography in relation to the languages used in the atlas context is of profound importance, especially for small countries, where the primary purpose is to serve the domestic population in an intelligible way and secondarily to present the information for the international public. Accurate translation is not only a question of linguistics, but in addition a problem of scholarly accuracy, international standards and categories. The work on the fifth edition of Atlas of Finland (1977-1993) has shown what a formidable task this really is, as this atlas has been published completely in 3 languages (Finnish, Swedish and English). For more information about the national atlases of Finland see & 3.

National integration

Thematic cartography including atlas cartography has long standing traditions in Finland. Thus, on these experiences it might be fruitful to analyse some aspects of scale related integration. To start with, one can mention the integrated use of information on maps in a large scale and primarily produced for other purposes than thematic surveys. Maps of this type can be considered as regional according to the relative scale (showing only a part of the country). An example of this are the GT-road maps of Finland published in the scale of 1:200 000, the whole of Finland covered by some 19 separate sheets (size 15" x 30" or 82 cm x 100 cm). Information from these maps with quite a varied content of spatial features was derived for several different thematic maps in the 5th edition of Atlas of Finland (later abbreviated AoF:5).
For instance, a map on the relief energy was constructed on the basis of the contour lines on the GT-road map\(^4\). Furthermore, these same road maps have served as base material for several different land use maps in AoF:5, for example a map of the arable area\(^5\) and a map on the settlement pattern\(^6\). All these thematic maps are published in the scale of 1:1 million in AoF:5 and have thus involved special methods of scale reduction and processing. These cases are not in a strict sense illustrations of primary integration of two different types of maps (road map and thematic maps), but show how - by scholarly interpretation - spatial information from very different sources can be utilized. These pioneering efforts indicate future ways of more effective use of spatial information across conventional cartographic boundaries.

Another approach to integrate source material on a national level is represented by the use of proper regional thematic maps and atlases. Local planning is increasingly supported by cartographically presented information. Thematic maps and atlases of this large scale type, already quite common in Finland and other countries, present an important interface between the public concerned and the different planning authorities with duties to analyse, anticipate and explain comprehensively the geographical and regional features of the actual plans and thus to create a basis for the articulation of public opinions.

Examples of source material which could be more widely used for national thematic maps can be interpreted from a map on the general planning situation in Finland\(^7\), map 7d, and a map on the state of the building culture inventory\(^7\), map 7f, also a survey of the general topographic maps in scales 1:500 - 1:5000\(^8\), map 18a could be used as a basis for retrieval of information to be presented on small scale thematic maps.

Regional thematic atlases have also been produced for educational, nationalistic and scholarly purposes. There are further numerous maps of various scales and themes included in different geographical either scientific or popular treatises of small areas, villages, tracts, and so on. This kind of material has rather seldom been used for thematic maps on a national scale. An attempt was made to survey types of regional divisions of Finland proposed in relevant scholarly literature (covering both physical and human geography) with the aim to derive for AoF:5 a basis for a description of the spatial pattern of Finnish landscapes. However, this attempt could not be finished, but gave nevertheless some experiences of the use of this valuable source material, which now through the KALSKE-system (see above, p.2) will be much easier available.

Regional atlases have also been produced for educational, nationalistic and scholarly purposes. From Finland an example of this kind of regional atlases is the "Atlas of Archipelago Finland\(^9\), describing nature, settlement and economy of the southwestern archipelago region, one of the "faces" of Finland. Here the application of some common - if not intentionally integrated - features are to be discerned in relation to the earlier and later editions of AoF. For instance the purposeful aim to describe diverse phenomena in minute geographical details, the proposition of geographical subregions and the use of sample areas in order to illustrate in an intelligible fashion the environmental context in place and time.

At least on a national level it ought to be easy to adhere to some common principles as regards the scale structure, thematic colour schemes and symbols both in the regional atlases and maps and those which can be considered as national maps.
and atlases. On a more profound level one ought to achieve a good correspondence of the scholarly data interpretation in all these cases. Thus fruitful ideas could be exchanged and a more efficient parallel use of thematic maps in different scales would result. In some cases again, thematic analysis on a national scale, may be originally made for the use in a national atlas, could be directly applied on a regional scale. For instance, information on thematic maps showing distribution of generalized topographic features, population, land use, industrial plants and traffic flow can be adapted to either planning maps on a regional scale or used for regional syntheses on a national or even international scale. Thus integration ought to developed into a two-way interaction of both methodology and concrete information.

The understanding and interpretation of thematic maps and cartography at large, must also be viewed in an educational light. Thematic maps are seldom easy to interpret, thus at least in a national context some common standards for thematic maps in different scales (regional and national) and also with varied thematic orientations (for instance environmental, historical, social, communication atlases) would be most desirable. One way would be to prepare a guidebook on thematic cartography, in Finland such one is presently planned.

Thematic maps are an increasingly important medium for information distribution at all levels of education. But more generally, in a modern society, where the environmental, social and economic matters are interwoven in increasingly complex systems, the public enlightenment is of crucial importance. Thematic maps and atlases are - if designed in a rational and integrated manner - one of the most efficient ways to attain this goal and to strengthen the democratic processes.

International integration

Today thematic cartography on an international scale, dealing with groups of countries, continents or the whole Earth, has grown to a very considerable sector of map production. Nations are of course still the main actors, although internationally organized operators both in welfare (for instance the bodies of UN), business (international companies), politics (EEC) and education (the great publishing companies) produce internationally composed thematic maps and atlases. When scrutinizing these one can not avoid to observe the occurrence of considerable discrepancies both in methods and - even worse - in the quality and accuracy of the data matters and their interpretation. This was apparent, when at the National Survey of Finland an educational atlas was prepared with the title "Finland in Europe"10, comparable with Pierre Vallaud's "Atlas Hachette - la puissance économique"11. It would be of ultimate importance if international cartography would be supported efficiently by well and carefully developed sets of some common standards and methods. This must of course not hinder innovative developments. In the field of land use mapping ambitious international efforts has been made already more than ten years ago, but has not so far resulted universally accepted standards.

An example of an international thematic atlas produced by close cooperation between a number of countries is the Norden-atlas12. This publication (116 pages inclusive texts) gives a comprehensive geographical image of the five Nordic countries (Denmark, Finland, Iceland, Norway and Sweden). This Norden-atlas
adhered to traditional cartographic methods and typography of national atlases. Typical for this atlas is the use of a variety of scales and regional frames.

Comparable to this Norden-atlas, but of much larger scope are the atlases of Danubian countries\(^{13}\) and the eastern and southeastern Europe\(^{14}\). However, the reality described by these two atlases has to a great extent been swept away and new attempts are needed. Similar in principle, although of much more modest design (printed in black-white only) are the Eurogeo publications\(^{15}\). These serve effectively the need of actual spatial information through thematic maps and figures including condensed texts. Finally one could also refer to the international atlas works with a specific theme, like Atlas Florae Europaeae\(^{16}\) and Atlas Linguarum Europaeae\(^{17}\). These specific thematic atlases have of course demanded close international integration of the data handling, processing, regional frameworks and visualization of the content matter.

New challenges

Both regional, national and international atlas work has to deal to an ever greater extent with themes of similar character. The still prevalent systematic structure of contents reflecting the traditional disposition of geographical matters starting with the themes of physical geography, then continuing with population, economy, social affairs and finally with cultural features, will have to be changed into a problem oriented direction, or at least to be complemented by this last-mentioned trend.

Questions like population development, urbanization, resource management, environmental deterioration and cultural interaction are actually extremely complex and interwoven. Thus the cartographic visualization must take in account this constitution and try to show the different contacts and implications inherent in each theme or subtheme. Besides this, it is necessary to remember, that the spatial features (and the associated actual problems) represent often unbroken chains of causes and effects from the local to the global and conversely. This is of course not a very original insight, and thus there has already been published quite a few thematic atlases approaching these aims, the above mentioned examples (10 & 11) may illustrate this.

The technique of "multimedia", still in its infancy, can be considered as an advancement in this problem oriented and holistic direction. However, a serious warning is necessary: carefully prepared, designed and printed thematic maps (the same concerns also topographic maps) and atlases will never loose their importance and advantages both from the point of everyday (and everyone) use and permanent sources of great and (often earlier unexpected) potential information. Electronic media have all too little analysed shortcomings in relation to the neurobiological and psychological limits of the human mind!

The integrative and compatible use of different media forms for thematic cartography is apparently an urgent question. Only by applying complementary holistic methods and techniques more adequate solutions will be found in this field. The present emphasis on national atlases (often with an exhibitional bias) has to give way for more attention on local and international projects. Finally, also the individual thematic maps and cartograms have to be designed with the view to integrate and harmonize the visualization of different aspects or elements inherent in
the theme in focus. A preliminary sketch of the field of integration is presented in the accompanying figure.

Fig.1. Tentative framework of integration in the field of thematic cartography.
Future prospects

In order to be more efficient both from the point of truthful information distribution and of economical resource management, two different criteria have to be fulfilled. The one is closer cooperation between all engaged in thematic cartography. This is an area where the ICA Commission on Regional and National Atlases could start to act purposefully. Different aspects of thematic methodology in cartography ought to be discussed with the aim to harmonize and develop in an innovative fashion thematic orientation, scale structures, symbol designs, and so on. Exchange and common use of primary data should be improved and cooperation between different scholarly units, institutions and individuals both in the own country and abroad should be stimulated in order to improve the quality, accuracy and swiftness of thematic information distribution by different types of thematic maps and atlases.

Secondly, it is of crucial importance to bring thematic cartography closer to the local level by more frequent and cartographically more efficient use of thematic information associated with the everyday issues most frequent and of the most essential importance to the inhabitants concerning the region which they conceive as their own. A structured strategy, both national and international is needed in order to achieve these goals of information service and corresponding production of maps and atlases.

It is essential to realize, that the most potential directions of thematic cartography lie in two extreme ends: the local and the international. The technical visualization through different media will improve with great steps. Because of this the simultaneous enhancement of scholarly interpretation is most urgent. Correspondingly it is imperative to tie together the local and the international dimensions in a way which will help understanding the geographical elements, their structures and interactions on the Earth.

The essence of the opinions in this paper can be illustrated by a quotation from the meeting of the IGU Commission on the History of Geographical Thought, August 4-7 1992 in Fredericksburg (USA) on the general theme "Geography is Discovery". Here the Finnish geographer Anssi Paasi claimed: "While voyages of discovery have helped to fill in the blank spaces on the globe, there still lurks in the human consciousness and mindscape an area whose mapping present a constant challenge for geographers. It is when installed in this national and international mindscape that meanings, ideologies and power elements confront each other in the same way as they did in the national and colonial objectives set for traditional geography." Mapping and cartographic conveyance of truthful enlightened ideas and images present a most effective way of attaining positive reactions to the problems facing mankind today.
References

1 Jaatinen, Stig (1992), The scholarly structure of the national atlases of Finland. Paper, given at the meeting of the ICA Commission on Regional and National Atlases in Madrid, in May 1992 (yet unpublished).


10 Suomi Eurooppa (Finland in Europe), an atlas published by the National Survey (so far only in Finnish), 1993 Helsinki.


17 Atlas Linguarum Europaeae, started in 1990, yet uncompleted.

Product analysis and market research to explore regional atlas content and design
C.P. Keller, D. Hocking (Victoria, CDN)

0. ABSTRACT

This presentation points out the need for product analysis and market research in regional atlas design, production and distribution. Argued is that some contemporary atlas initiatives may be driven by technological innovations in cartography and geographic information systems, failing explicitly to consider the needs of atlas users. Three preliminary product analysis and market research surveys have been conducted to study the atlas market, and to solicit user input in British Columbia, Canada. This presentation introduces and justifies the three surveys, gives a general summary of survey methodologies, and reports selected highlights of survey findings.

Key Words: Regional Atlases, Product Analysis, Market Research, Content and Design, User Survey, Atlas Reviews.

1.0 INTRODUCTION

This presentation reports on product analysis and market research conducted to investigate aspects of atlas usage and design. The research was undertaken for two reasons. Firstly, it was discovered that there exists little published literature on atlas market research. Such enquiry, however, was felt to be of potential benefit to the overall atlas research agenda. Secondly, it was discovered that few published guidelines are available to those interested in the design and production of a new atlas, and those placed in charge of evaluating atlas proposals. Comparable product analysis and market research, therefore, were thought to have the potential to yield findings that might assist those interested in planning or reviewing a new atlas initiative.

Research conducted focused specifically on usage and production of state or provincial atlases. However, it is thought that findings are equally applicable to the concept of any general thematic atlas of regional scale. This paper, therefore, will keep referring to regional atlases.

In specific, research was conducted to solicit answers to a number of questions. They were:
What can be learned from existing regional atlases?
Who uses regional atlases?
For what purpose?
How often, when and where?
What are users' opinions concerning regional atlas design and content?
Are users satisfied with regional atlases marketed?
What awareness, interest and opinions do atlas users have concerning non-traditional methods of atlas presentation?
What are experts' opinions concerning atlas design and content?
What should a new and innovative regional atlas look like and contain?

Three surveys were designed and conducted to seek answers to the above questions. Firstly, a preliminary survey of existing state and provincial atlases was carried out to search for commonality and differences in regional atlas content and design. Secondly, a set of atlas user surveys was carried to solicit user input into thematic content and design for a regional atlas, and in an effort to identify general atlas usage patterns, to enhance our knowledge of atlas users' awareness and understanding of atlases in general, and to identify what the public's expectations are of a regional atlas, especially given contemporary changes in publication and distribution media. Thirdly, a survey of atlas reviews was conducted in an attempt to seek indirectly expert's judgement of existing regional atlases.

This presentation points out the need for product analysis and market research in atlas design, production and distribution. It introduces and justifies why we opted for the three surveys conducted. Detailed description of the surveys and research results have been published in Hocking (1991), Hocking, Keller and Peterson (1991) and Hocking and Keller (1992, 1993a, 1993b). This presentation, therefore, will limit itself to a general summary of the survey methodologies, reporting only highlights of survey findings. The presentation concludes by commenting on the worthiness of the market research conducted, arguing the need for additional work.

2.0 THE NEED FOR ATLAS MARKET RESEARCH

Examination of contemporary atlases on the market reveal great diversity in products. Atlases marketed extend from strict collections of bound or loose leaf maps drawn to consistent specifications to thematic atlases where maps are complemented by textual information, photographs, drawings, graphs and other visual components, with maps often taking on a secondary role. Atlases available vary considerably in size, price and scope. They range from cheap school atlases to expensive historical and regional atlas tomes suitable for coffee table display. Distribution media vary from the traditional bound paper format to digital products.

Despite the contemporary diversity of products on the market, there appears to be a general trend for the concept of an atlas to have broadened progressively from strictly a collection of paper maps towards a multi-media spatial information system, utilizing digital technology for production, and more recently as well for occasional publication. The transition of an atlas towards a contemporary digital multi-media system has been recognized by Laurini and
Thompson (1992), who refer to the contemporary printed atlas as a forerunner of the modern form of hyperdocument.

Given the above trend in atlas design and production, one should expect a considerable scientific debate in the cartographic literature critically reviewing atlas editorial and production innovations, and critically evaluating the success of innovations in atlas content and design. It came as a surprise to us, therefore, to discover that there exists relatively little such published critical atlas research. We conclude that atlases represent a relatively neglected area of research in cartography, geography and information sciences.

Given this paucity in recent published atlas research, the questions beckon as to what guidelines are available to those interested in the design and production of a new atlas, and those placed in charge of evaluating atlas proposals, especially given the rapid advances in cartographic technologies. What was the rationale for progressively moving atlases from a collection of maps towards multi-media documents? Was this move merely technology driven? Have recent innovations in atlas design and production received praise from atlas users? What are the most effective and appropriate methods of packaging atlas information given contemporary technology and societal values? For what markets are the different atlases produced, and do they meet the needs of the readership? What ought the atlas of the future to look like? How can we educate atlas users to accept innovative atlas products and how do we market innovative atlases?

We posed some of these questions in relation to regional atlases when setting out to conceptualize the content, format and design for a new atlas for the Canadian province of British Columbia. Having failed to find a satisfactory debate and solutions in the literature, we were faced with two options. On the one hand, we could opt to follow the advice expressed in Robinson (1971) that cartographers, as experts, should decide what to show in atlases and how to show it; and thereafter to go ahead and do it. This 'specialist approach', of course, makes the assumption that we know what is best for the atlas user, and that we know what the atlas user wants or should get. A second option open to us was to accept a general trend in production and society in the latter part of the twentieth century to commence by conducting product analysis and market research. This would involve actively setting out to study products on the market, and to solicit input and opinion from the atlas users, including the public and special interest groups.

No doubt, there are intuitively obvious privileges to be derived from pursuing the specialist approach. However, we decided that the potential added benefits to be gained from a strategy of product analysis and market research would justify the extra cost, effort and invariable delay. We decided, therefore, to conduct the three surveys outlined below.

3.0 THE SURVEYS

Comparable product analysis and market research initiatives generally are constrained by availability of funds, facilities and labour. The principal team to conduct the research outlined below consisted of a member of faculty and a graduate student in the Department of
Geography at the University of Victoria in Victoria, British Columbia. Neither received payment for work conducted. A small operating budget was obtained from the University's pool of Social Sciences and Humanities Research Council (SSHRC) funds (approx. Can. $2,000). Other funds were found to hire a research assistant over a three month period. Access to computing and related research facilities were made available by the University's Spatial Sciences Laboratories. The above constraints should be kept in mind when evaluating the scope and merit of the following three surveys.

3.1 Survey 1: Content and Design of Regional Atlases

It was felt that considerable knowledge and insight could be gained from a study of past atlas initiatives. A decision was made, therefore, to compare aspects of format, organization and thematic content of fifteen Canadian provincial atlases. Findings were contrasted to a 1969 study of provincial and state atlases by Stephenson and Galneder, and to a more recent study comparing thematic content of state atlases by Kent and Tobias (1990).

3.2 Survey 2: Set of Atlas User Surveys

In order to solicit input and opinion from atlas users, we thought it most appropriate and effective to conduct a direct user survey. Three issues had to be addressed, namely:

- How to identify the population of atlas users?
- What is the most appropriate method of interview?
- What questions to ask?

Identifying the population of atlas users:

Atlas users are an elusive and varied population, not readily identifiable via some existing list. Kirby (1970) had identified atlas users by targeting special interest groups, but wanted also to identify the level of interest in atlases among the general public. A decision was made, therefore, to obtain input from both the general public and from specific populations more likely to use atlases. In the latter category we included geography and history university graduates, undergraduate geography students, and school pupils. Constraints of finance and time did not allow us to target other specific populations, for example natural resource managers or urban and regional planners.

Method of interview:

A questionnaire survey was felt to present the best means of soliciting the information sought. Given that random distribution of questionnaires to the general public was anticipated to result in a very high percentage of non-responses, a decision was made to prescreen the general public by telephone-contacting a systematic random sample. In order to keep survey costs manageable, the population of the general public from which to draw the sample had to be confined to a population of approximately 300,000 living in the Greater Victoria region in British Columbia. We had to accept the fact that this would give the survey results a potential bias towards an urban population living in a westcoast city.

Individuals contacted by telephone were asked a number of preliminary questions concerning their atlas usage, demographics, and whether
they would be willing to participate in a more lengthy mailed questionnaire. Out of 131 individuals contacted, 62% agreed to answer preliminary questions, and 54 agreed to participate in the follow-up survey. In the end, 37 full questionnaires were completed by the general public, representing 28% of the initial contacts, and 69% of those who agreed to answer the longer survey.

A list of names of geography and history alumni was supplied by the University of Victoria’s Office of the Registrar. A full survey was mailed to a stratified systematic random sample of this list. 29% of the 268 alumni contacted completed the full survey. A class of third year students taking a course in the regional geography of Canada were selected as the sample of geography students. This yielded an additional 47 responses. Pupils in grade ten (average age approx. 16) in the Canadian high school system from both, an urban private and a rural public school, were selected as an appropriate sample for the last group. This yielded an additional 60 and 68 responses respectively. In total, 574 individuals were contacted and 289 completed surveys were received, an overall response success of 50%.

What questions to ask:

A set of questions were draw up to solicit answers to some of the queries posed in Section 1.0. These questions were put into questionnaire format following design recommendations made by Dillman (1987). Related questions were grouped by cognitive ties, and demographic questions were kept until last. Respondents were given an incentive to complete the survey by informing them that all completed responses would be entered for a draw for the prestigious Historical Atlas of Canada Volume I (1987). A pilot questionnaire was tested on a sample of pupils and teachers. A copy of the questionnaire and additional design and detail can be found in Hocking (1991).

3.3 Survey 3: Expert Atlas Reviews

Once published, atlases are subjected to peer review. It was felt that a detailed study of peer reviews of regional atlases would yield an insight into what experts look for, praise and condemn in these atlas initiatives. Therefore, ninety-eight reviews of fifty-three state and provincial atlases of the United States, Canada and Australia were examined using the methodology of content analysis. An attempt was made to identify what judgement criteria experts use when evaluating state or provincial atlases, and what aspects of the different criteria the reviewers praise and condemn. As well, a search for expert consensus was undertaken in an attempt to identify a regional atlas review recipe.

4.0 MAJOR FINDINGS

4.1 Preliminary Survey: Content and Design of Regional Atlases

This survey revealed few surprises. It would appear that the vast majority of state and provincial atlases continue to preserve an atlas tradition observed over 40 years ago by Nicholson (1952) of dividing thematic content into physical, economic and socio-cultural topics. We found ourselves in agreement with Stephenson and Galneder (1969) and Kent and Tobias (1990) when observing a sharp increase in socio-cultural themes through time at the expense of both physical
and economic coverage. We found that thematic coverage tends to attempt to emphasise a region's unique character, with maps complemented by some textual explanation. Some valiant efforts have been made at breaking with tradition (for example the Economic Atlas of Ontario, 1969), but these efforts appear to have failed to translate into a widely adopted new regional atlas paradigm.

Comparison of the above findings were made to the only regional digital atlas at our disposal at the time, namely the digital Atlas of Arkansas (see Smith, 1987). Content and organization of this atlas clearly was modelled on the traditional approach, despite its innovative publication format. More detail concerning the above findings can be found in Hocking, Keller and Peterson (1991).

4.2 Survey 2: Set of Atlas User Surveys

This survey proved to be of considerable interest. Following are a number of highlights of survey findings. More detailed results can be found in Hocking and Keller (1993a and 1993b).

Nearly half the members of the public randomly contacted by telephone indicated no use for atlases. Of all those found to use atlases, a surprising number claimed to use them on a frequent basis, with the modal class being "more than monthly". Reasons for use vary, but are dominated by the need to locate places. Few commented on the value or usage of an atlas to study thematic data and their relationships. A surprising number of respondents did note, however, that they use atlases for daydreaming and imaginary travel. In general, therefore, atlases at present appear to be used predominantly for locational fact finding and imagination, rather than for analysis.

Asked to comment on the value of non-map components, users stressed the need for an index and gazetteer, which verifies their primary atlas use for location of topographic features. Reactions to other non-map components appear less positive, with photographs and interpretive text judged least useful. We found weak preliminary evidence to support the hypothesis that respondents fall into one of two groups: those who think that an atlas ought to consist exclusively of maps, and those who see the atlas as a multi-media spatial information system.

Most respondents recognizes that information should be up-to-date, stressing the need for clearly stated date and source of data. The fact that statistical tables and diagrams are the first part of an atlas to become outdated was commented upon by a number of respondents, and text received considerable criticism for being too general, irrelevant or unnecessary. Some respondents criticised atlases for not showing alternative or former spelling, and for failing to give adequate precision when referencing location in indexes or gazetteers.

Excessive complexity of design elicited concern. It would appear that the majority of users want atlases to be easy to read and understand, and in which it is easy to find locations. The physical size of an atlas is a point of contention, but there is no consensus, with some wanting atlases bigger, others smaller.

Respondents were asked to comment on alternative methods of atlas production and distribution, that is methods other than the bound collection of maps and accompanying information in book format.
Loose leaf distribution of maps did not receive favourable comment. Over sixty percent of respondents had seen maps displayed on a computer screen. Of interest here may be that chi square analysis ($\alpha = 0.05$) of cross-tabulation with gender revealed that men are significantly more likely to have seen them than women (2 in 3 males versus 1 in 2 females). Reactions to computer maps generally were favourable. Respondents recognized the potential benefit of map customization in a digital environment, and the ability to more efficiently update information. Pupils expressed considerable enthusiasm for computer maps, frequently commenting that they would be easier to access and use than the traditional atlas. However, computer maps did come under criticism for their poor relative quality and lack of clarity, and for assuming access to computers.

An attempt was made to identify whether users perceive electronic atlases as a substitute for the traditional atlas. The majority was found still to prefer the book format, and the vast majority of those expressing an interest in electronic atlases still noted the wish for a traditional book atlas to complement any digital initiative. Gender analysis revealed that twenty five per cent of men compared to ten per cent of women would be satisfied with merely an electronic atlas.

Users were asked how they would like to receive an electronic format atlas. Given the option of floppy diskettes, CD-ROM or video tape, users stated a clear preference for diskettes, but did not rule out the other media. Asked if users were willing to receive an atlas via cable television sometime in the future, two thirds expressed some interest, but less were willing to pay extra for this service.

4.3 Survey 3: Expert Atlas Reviews

This survey, too, proved to be of considerable interest. We found that, specific topics discussed in any one review can vary enormously, but that there exists a general review pattern where roughly equal proportions of text are devoted to issues of content and design.

It appears to be the norm to list an atlas's topical coverage, and to pass overall judgement on topical suitability. Comments generally tend to be positive although basic disagreement exists amongst experts over what thematic content is acceptable in a regional atlas. Some argue in favour of encyclopedic coverage, that is comprehensive coverage including content not particularly unique to a region, while others suggest focusing efforts by concentrating on the unique topics which capture the regional flavour. As one reviewer aptly notes, "regional atlases, especially, should have personality or 'soul'" (Norwine, 1976). Some reviewers praise placing a state or province in a broader regional context, a practise condemned by others.

Basic disagreement appears also to exist concerning the merit of non-map material. Some purists feel that an atlas ought to focus attention exclusively on maps, while others praise elaborate addition of non-map material.

Design and presentation attract a higher proportion of critical comments, notably when discussing organization of contents. Common criticism appears to focus on layout and appropriateness of scale, on suitability and quality of colour, on symbolization and on
We found the vast majority of reviews analyzed to be positive in nature and conservative in style, suggesting that experts generally appear to be happy with the status quo of regional atlas production. Beyond factual description, reviews tend to point to failure to meet expected standards of content and design, or to point out where an atlas deviates from the expected norm. However, a number of reviews were found to be highly disparaging of the traditional approach to regional atlas production, critically questioning the traditional approach, and calling for innovations in content and design. More detail concerning the above results can be found in Hocking and Keller (1993a).

5.0 DISCUSSION

Did the research prove of value to us when setting out to conceptualise a new regional atlas for British Columbia, and is additional research of this nature desirable?

The three surveys certainly have been a learning experience. Our investigations have revealed that there prevails an established status quo in regional atlas production, although there do exist variations in format, design and contents. This became evident when studying regional atlases on the market, and when examining experts' reviews of published works. With respect to our new atlas initiative, we could opt, therefore, to conform to what we learned about the traditional approach, fine tuning the established regional atlas recipe in an attempt to communicate the geography of British Columbia as well as possible.

However, our preliminary user survey yielded insights that suggest that the market may be ready to experiment with an innovative approach to regional atlas design. We learned that regional atlas users are not opposed to innovative methods of regional atlas presentation and delivery, as long as simplicity is maintained; a regional atlas should be ease to use and understand.

We learned that, at present, the traditional regional atlas is used predominantly for browsing and to locate topographic features; it appears to have proven too cumbersome and too tedious to find widespread use for analysis. The question arises, therefore, whether there does exist an innovative regional atlas approach using new technologies that will support not only daydreaming, browsing and topographic feature location, but that will also be simple enough to use for analysis.

We must acknowledge that our survey findings are based on a relatively small and possibly bias sample. Constraints of finance and other resources did not allow us to conduct market research of a scale suitable to properly ask and answer atlas user concerns and needs. Our findings, therefore, are preliminary, and should be corroborated by additional work.

We live in the "Information Age". We are in a period of history where tremendous advances are being made in the management, processing and dissemination of information, and where access to digital information technologies is becoming increasingly widespread.
We also live in an age of increasing instrumentation of the universe to foster real-time data capture and display. We now have satellite technology to monitor the earth's surface, we have global positioning systems to monitor movements on earth, and we have access to a multitude of solar powered field stations capable of automatically measuring and transmitting field data to central computers. Never before have we had more access to up-to-date geographic data, some of which are ideally suited for display in atlas format. We will need to develop innovative ways of integrating these data into our atlases.

We also live in an era where visualization of scientific patterns and results has become vogue, where spatial analysis has moved from static to dynamic process analysis, and where an increasing emphasis is placed on the study and visualization of interrelationships of environmental and social phenomena. These features, too, should find their way into the atlas of the future.

The above information trends already have impacted cartography and geography. Major efforts are ongoing to build digital national, regional and local cartographic information inventories, to develop digital geographic information analysis systems, and to capitalise on digital technologies to derive innovative cartographic products. Systems are now in place or under development to instantaneously digitally display mapped weather patterns, including animated temporal sequencing. We now have access to digital maps showing instantaneous probabilities of fire strikes during lightning storms, and we already can digitally access and query multi-media map based systems showing up-to-date real estate markets. Other examples of innovative mapping exist, too numerous to mention.

A number of ongoing atlas initiatives are attempting to make us of these rapidly advancing technological innovations in cartography and geographic information systems. However, many of these efforts are based on the 'specialist approach', failing explicitly to consult the needs and wishes of the atlas end-users.

Given the industrial competitiveness we are encountering in the latter stages of the twentieth century, few entrepreneurial initiatives will risk design and production of an innovative products without the conduct of market research. Why should attempts to design innovative atlases differ? Surely the atlas of the future should not be driven by technological advances, adopting them blindly. Should individuals in charge of new atlas initiatives not aim, first and foremost, to better meet the needs of atlas users. Would it not make more sense, therefore, to critically identify areas where we ought to improve on or add to traditional atlas capabilities, and to make selective use of technological advances in cartography and geographic information systems where appropriate?

Our preliminary market research revealed that the atlas user is ready for experimentation, but does have certain expectations. Limited resources constrained the scope of our market research effort, and did not allow us to answer all the questions posed at the outset. We feel that the work should carry on. Small targeted research projects are needed to conceptualize, design and build innovative digital regional atlas modules. Atlas user input ought to be solicited in the conceptualization and design phases. Atlas users should, thereafter, be given a chance to test and comment on innovations. The above strategy would ensure design of an innovative regional
atlas product, developed to suit first and foremost the needs of the market, not the needs of cartographic researchers and atlas editors.

Our children are getting to be known as the "video generation" (Karl, 1992). They will grow up taking contemporary innovations in information packaging and dissemination for granted, including innovations in mapping and cartographic communication. There exists no doubt, therefore, that the atlas of the future will look very different from the atlas as we know it today. It thus is not a question if atlas design, format and distribution will change, but how, when, by whom, and in whose interests.

6.0 BIBLIOGRAPHY


Abstract

The paper describes in detail and concretely the conception and design of the 'Medicinal Herbs Atlas of China' theoretically and practically. Its contents include:
1. introduction;
2. map projection design;
3. introductory map design;
4. the subject selection of thematic comprehensive maps;
5. the selection of representative method for medicinal herbs output maps;
6. the selection of representative method for medicinal herbs distribution maps;
7. the design of thematic symbol system;
8. the design of geographic base maps;
9. the design of front cover;
10. the selection of printing color.

1. Introduction

The Chinese medicinal herbs resources can go back to ancient times, and has long history in medicine, it is a base for developing the Chinese medicine undertaking. In order to find out the present situation and potential in herbs production, in recent years a comprehensive herbs resources general investigation has been carried out in the country, its data will be compiled and edited into writings and maps in volumes. The 'Medicinal Herbs Atlas' is a map form of the herbs general investigation results.

The tasks for the atlas are to sort out and compile the herbs investigation data into an atlas that reflects the present situation of Chinese herbs resources distribution, as an important reference data of plan and design for developing herbs resources. It is required to have highly scientific summerrization and generalization and to represent the investigation results in detail.

This is an octavo of full edition atlas without frame. The largest scale of national maps in the atlas is 1:12 million, which represent county boundary of the whole country, so that it is suitable to present the herbs
output by county unit. If 16 mo atlas is selected, the scale of provincial map is about 1:2 million. Such scale is too small for county range, and quite a number of county boundary could not be represented on the national map, and affect the clarity and readability on map. If quarto is used, although the scale can be enlarged, the volume, weight, cost of the atlas will be increased, and it is not convenient for use and store.

In the case of the same format and mapping region, the scale of full edition map without frame is larger than the map with frame. In the format scale differing one times, about 50% maps of the atlas are the same with quarto scale, another 50% map scale is slightly small. When the map format and mapping region range have been decided, the bigger the map scale, the more contents can be held. Map is a carrier of transmitting space information, the large the carrier capacity, the more richer information can be transmitted.

There are four parts in the atlas design, i.e. introductory map, thematic generalized map (thematic introductory map), herbs output map, provincial map of herbs distribution, 81 maps in total. A statistical graph related to the main map is designed on the back page of the map front; the related writing description is arranged on the back page of the map back, to further add remarks of the main maps, e.g. herbs growth condition herbs property etc.

2. Map Projection Selection

As a result of the maps in the atlas are quite lot, the difference of each characteristics, use, geographic location and scale are fairly big. The selection of projection types should be varied appropriately, but the integrity, comparability and consistency should be considered.

The maps in the atlas can be divided into 3 types according to the contents, introductory maps, herbs output maps, herbs distribution maps. The scale of provincial maps is 1:1.5 million to 1: 4.75 million, the administrative maps of China is 1:14 million, introductory maps and herbs output maps is 1:12 million, climate maps is 1:18 million, 1: 24 million. According to mapping region, it can be divided into national maps and provincial maps.

Three kind of projections are selected in the atlas: oblique equal-area azimuthal projection, normal equal- area conical projection and normal autogonal conical projection.

The Administrative Map of China includes all the waters in the south China Sea, which represents the full view of Chinese territory. As a result of the large map range, small scale, it is available to select the projection of small deformation and well-distributed. The oblique equal-area azimuthal projection can meet these requirements. Its basic characteristics are as the following: the central longitude is a straightline, the other longitude and latitude are all symrical to the
curve of the central longitude, without any area deformation, any
deformation in the projection center, but the far away from the projection
center, the bigger the deformation does, and the shape of distortion
isobaths is more close to the outline of Chinese region. The length
deformation in wide region of China is $\pm 2\%$, the largest angle deformation
is $2.5^\circ$, the length deformation in part border district is $\pm 3.25\%$, the
largest angle deformation is $4^\circ$.

The introductory maps and herbs output maps are the maps with contents
representing relief, cultural features, soil, vegetation, climate, natural
division and herbs output in China, the islands in the South China Sea as
an insert map. The normal equal-area conical projection is used, the
standard longitude is $\phi = 25^\circ$, $\phi = 47^\circ$. The length deformation along the
longitude and latitude direction is $\pm 3.8\%$, the south border is $\pm 3\%$, the
central is $\pm 1.9\%$. The length deformation in wide area between the North
Latitude 20° and 50° is 2%, the angle deformation in most area is all
within $2^\circ$. The largest angle deformation: the north border is $4^\circ$, the
south border is $3^\circ$, the central is $2^\circ$.

The normal autogonal conical projection is used in provincial maps, the
provincial maps are herbs distribution maps which requires geographic
base map either in detail or accurate, no error in azimuth, not too large
deformation in length and area.

China is situated in the central latitude region, most provincial maps are
extended along the latitude, so most are horizontal maps. Therefore, the
conical projection of equal absolute value of length deformation of
mapping region edge latitude and central latitude is selected. So that
the length deformation of most maps is all within $\pm 0.1\%$, only the length
deformation in individual maps is up to $\pm 0.2\%$ to $\pm 0.5\%$. The area
deformation is one times higher than the length deformation.

3. Design Conception of Introductory Maps

The introductory maps are composed of 9 unfold pages, 14 maps, which is
closely related with the atlas contents, it briefly expounds for the people
to the natural conditions providing with herbs production and development
on the land of China, e.g. soil, climate and cultural features and so on.

The introductory maps are consisted of administrative map of China, city
and county maps of China, relief map of China, soil map of China, vegetation type, climate maps of China (including 5 small maps: total amount
of annual solar radiation, hours of annual sunshine, annual average air
temperature, annual precipitation, climate division in China), natural
division in China etc.

(1) The Administrative Map of China represents the administrative
structure in China, and shows the full view and general situation in China,
it mainly presents provincial boundary, provincial government seats, and
important cities, and have the function of both general map and index maps in provincial map group. The provincial maps represent the administrative structure in each province in detail, e.g. county boundary and settlements in county level. The administrative map of China and the provincial maps constitute a series maps of organic connection, it works in concert with each other from its contents to mapping region.

(2) The city, county maps of China take the boundary and name of city and county of administrative division in China as the main contents of the map, it is also a base map of population map and herbs output maps.

(3) The population map of China. Population distribution has close relation with the development, use and cultivation and expanding of wild herbs resources. In the area of frequent mankind activities, the use rate of wild herbs is higher, it easily brings about excessive collecting and damage wild herbs; The demand quantity for wide herbs is great in density population region, in certain extent that the production development of herbs will be promoted. Such an intension will be shown on the population maps.

(4) The relief map of China. The soil map of China are all representing the object condition of herbs production, the complicated terrain and soil type available for various herbs production and reproduction, the vertical zone of close relations with the growth of plant herbs and animal herbs, the soil property having great influence on plant herbs growth and medicinal property are represented on the maps.

(5) The vegetation map of China. Animal and plant growth is not often growing by itself in isolation, they are interdependent and influence on each other in the big family of natural world. So the vegetation environment that herbs growth depended on will be briefly represented on the vegetation map.

(6) The climate maps of China including 4 maps, i.e. total amount of annual solar rediation, the hours of annual sunshine, the annual average air temperature and the annual precipitation, which mainly represents the climate factor closely related to herbs production. It is evident that the production of plant herbs is affected by climate factor, e.g. longan is only growing in the south region of the tropic of Cancer with high temperature and rich precipitation.

(7) The climate division map of China and the comprehensive natural division map of China. On the basis of herbs resources investigation, the requirements of strength the management of herbs production, expanding herbs production scale, working out herbs production plan etc., all should be combined with climate division and comprehensive natural division. These maps provide natural condition of climate, soil, precipitation for herbs production getting high and stable yields, keeping pure medicine property, provide scientific basis for natural condition with local
4. Subject Selection of Thematic Comprehensive Map

The generality and synthetic of the thematic comprehensive map make it has the introductory map characteristics. The contents fully and briefly represent the theme of the whole atlas, the scale is 12 million, mapping region is nation wide. Five maps are selected: the distribution map of wild plant herbs of China, the distribution map of cultivated plant herbs of China (1), the distribution map of cultivated plant herbs of China (2), the distribution map of animal of China, mineral herbs of China and the medicinal herbs division map of China. The purpose of selecting this map group is to briefly present the present situation of the national herbs distribution and the total output. For example, a plan and scientific arrangement can be drawn out for the national herbs production and developing prospect through this medicinal herbs division map.

This map group affect the judging and prompt of herbs output map and provincial herbs distribution map arranged at the back.

5. Selection of Representative Method for Medicinal Herbs Output Maps

The medicinal herbs output maps represent the situation of medicinal herbs output. 126 kind of medicinal herbs most in use are selected within the scope of the whole nation in the atlas, present the statistics numbers in each county. 2-3 kind of herbs output are represented on each map, 43 herbs output maps are in total. The herbs represented on the same map should select their main produce area that are not in the same area, however a few overlay is hard to avoid. So the selection of different color points stand for different herbs type output, in this way 2-3 kind of different herbs output in the same county can be presented on a same map.

The output of wild herbs and cultivated (or breaded) herbs is also distinguished in the atlas. Therefore, the small points of same color with different shape points and small circles are used to differentiate the output of wild herbs and cultivated (or breaded) herbs.

Some herbs photos are put up on the edge of output maps, by which either enrich the map contents, or increase the knowledge of distinguish herbs.

6. Selection of Representative Method for Medicinal Herbs Distribution Maps

In the investigation of herbs distribution in field, the range line of herbs distribution has not been drawn out in detail and the investigation data is summarized. So selecting the representative method of herbs distribution have to choose summarized method. The herbs distribution is areal distribution feature, it appears interruptedly within mapping region, and is available to be represented by generalized range line. The generalized range line of herbs distribution is represented by single
areal symbol, and the symbol pattern itself has no accurate location meaning.

For the herbs connected distribution, 3-5 kind herbs are selected on each map and the distribution range line should be drawn out and represented by coloring.

The herb distribution maps are separately based on provinces, including provincial boundary, prefecture boundary and county boundary. The form of statistics chart can be used within the range of prefecture boundary to represent th percentage of several main herbs output.

Graphes are used on some maps to present the rule of herbs vertical distribution.

7. Design of Thematic Symbol System

Map symbol is map language, and is an important tool and basic form to objectively represent various things and phenomenon. Over 320 medicinal herbs most in use are collected in the atlas. A scientific symbol system design method is required to represent many herbs on map in symbol form with both independent and unity, either containing individual information or collective information.

1) The classification of medicinal herbs

The map symbol classification of the herbs atlas is consistent with the herbs resources. Herbs are classified from classification to sub-classification, there are more than 10 herbs at least in each classification, more than 100 root medicinal herbs. This is the obvious characteristics and difficulty in symbol design, so that a principle of variations and unity should be followed.

2) The requirement of thematic symbol design

Thematic symbol design should correctly deal with the relations between symbol independence and systematization. Each classification of symbols and each symbol should have its evident characteristics and independence. Within a group of symbol, the internal relations should be kept in form, and systematization should be reflected. Symbol design should be concise and in pattern. Composition should be simple and implication clear and characteristics distinct. The key of pictorial symbol should stress on image feature, not rest on detail of symbol. If pattern too concentrate "appearance", so that it will be difficult for symbols plotting and remember, and not good for the effectiveness of reading.

3) The factor of influence on symbol design

Human vision is able to distinguish distance c of two points close, to depending on good vision distance d and visual angle α. Their relations are: $a = c/d$ or $c = d \cdot a$.

Generally the reading distance of table map is 30cm, the minimum distance between two points is 0.75mm.
Registration of printing and plotting techniques should be taken into account in symbol design, any frame of geometric pattern should use the natural color of printing color or interval color containing yellow component, so as to avoid to bring about registration print error.

4) The process of symbol design

Herbs symbols are various, pictorial symbols should be used as more as possible for the consideration of the characteristics of herbs resources. The materials of pictorial symbols come from the insert pictures or photos of atlas, annals, dictionary and handbooks of Chinese medicine, etc. To highly generalize its shape, exaggerate its basic character of herbs, abandon secondary parts, so that a symbol embryonic form can be formed. The embryonic symbol will be classified and induced, the appearance close to is a group, a symbol of generalized and abstracted should be designed in a group, and different color stands for several different herbs. For the herbs inconvenient to use pictorial symbol, e.g. mineral herbs, should all use geometric pattern. Its appearance outline is used to represent its general character, and different internal structure produce various patterns. Many colors for each pattern will produce several decade herbs symbols.

5) The principle of symbol coloring

Map printing color, map surface effectiveness and chromatograph printing should be considered in symbol coloring, and the natural color of herbs should be taken into account properly.

Four colors printing will be used, i.e. yellow, magenta, blue and black.

1) Symbol frame coloring. Yellow is not individually available for frame, because yellow is a weak color, the result is not good to be used for complete symbol or frame, but magenta, blue and black are good. Black should not be used more in order to decrease the depression sense on map surface and to reflect even more alively and distinct atmosphere. Some green and orange can be correspondingly used, they are combined separately by mesh points of yellow with blue, yellow with magenta. The visual sensitive degree of yellow mesh points on map surface is not good, the reflection is not obvious if there is slightly chromatograph printing error.

2) The mesh point of frame coloring should be used for color process within symbol frame, so as to avoid to bring about chromatograph printing error. If the other coloring is required, then the combined color of symbol frame color with yellow should be used.

8. Design of Geographic Base Map

Geographic base maps are geographic base of compiling various thematic maps. The relations between thematic contents and base map features, like 'flower' and 'leave', leave plays a role in setting off and supplement. Geographic base maps have an influence directly on the representation and
The unity, coordinatity of base maps reflect the unity, intersupplementary, coordinatity and comparity of atlas. The new compiled national general atlas can meet these requirements, only the loading quantity for each feature is too great. It can be used as the basic data of compiling geographic base map of the atlas, and is required to be highly generalized and abandoned in large amount. It will be explained in the following taken provincial maps as an example.

Water system: The rivers above three level should all be selected, under the three level should be abandoned accordingly. The distance between rivers is 10-20mm. The area of lakes and reservoirs large than 10mm² can only be selected on map.

Settlements: The settlements above county level should all be selected, under the county level should be selected accordingly. About 60 settlements should be selected in every 100mm² in most density area, its density is only 1/4 of the basic data.

Roadway Network: All the main line of railroad are selected, roadway of first level is mainly selected, the second level roadway is selected a few in minority region. Roadway network on the map should be controlled between 20-30cm² in roadway density region.

The representation of administrative division boundary is to county level and autonomy county.

Vegetation feature will be not represented on the atlas base map, the vegetation condition is briefly represented mainly through the clarity and readability of the base map, so its contents could not be too miscellaneous.

9. Design of Front Cover

The front cover of atlas is a highly generalization for atlas contents, and also an improtant part of atlas decoration design. The herbs atlas takes the herbs output and distribution a theme. The conception and design of front cover should be characterized by typicalness, representativeness and artistry of herbs.

1) Atlas name

Atlas name is the first information to readers. The selection, arrangement form and color of typeface should be coordinated and identical with the subject contents. Atlas name has two ways of writing, i. e. write words horizontal and vertical. The vertical is a traditional way of the Chinese nation writing. The Chinese medicine and the Chinese herbs are gems of the Chinese traditional culture, in order to enable the decoration of the atlas front cover fully present the Chinese nation style, a vertical form of writing map name is going to be used, and is deviated to the right side of front cover, the typeface is running hand. Vertical writing of map name
is not frequent in the published atlas in China, the atlas would be to make a test on that.

2) The pattern of front cover

The atlas published internal and external recently pay more attention to using symbolized pattern to decorate front cover. Although the pattern of front cover occupied secondary place on the main cover, it presents the theme of the atlas and the contents on the main cover will be added and enriched. The atlas selects the well-known rare medicinal herbs ginseng as the front cover pattern and is put on the bellow left side of the front cover, that will have the effect of work in concert with and set off with the theme by contrast.

3) Cover on the front cover

Little dark green is selected in the atlas, green is a color of grass and woods in the nature, there is some meaning of nature and growth, that the people can be called attention to the fresh sense of the nature.

The name of the atlas uses gilding color, so that it has the characteristics of bold, resplendent and serious. The pattern uses relief without color printing, so as to highlight the theme.

10. Printing Color

At present, four color printing is the most advanced and lowest cost coloring program in printing technology. The geographic base maps of the atlas use steel grey, the other colors are all combined by standard four color registered, it includes base color in large area and area color inside small symbols. Only the frame of symbols and the points of point-number method adopt the natural color of magenta, blue and black. The result is ideal through the proof test on color.

REFERENCES

Abstract. The design principles of a user interface to a computer atlas are studied. A computer atlas can be seen to consist of a database and a user interface. The database contains an abstract model of the real world, and the user interface is a tool for viewing this model. The design of the user interface can be divided into conceptual, functional and appearance design. The conceptual design is often based on a metaphor. Suitable metaphors for a computer atlas are examined. The functional design is based on the conceptual design and the appearance follows from the functionality. Some practical experiences of designing display maps are reviewed. Hypermedia is examined as a way to realize the user interface. The main advantage of hypermedia is the use of contexts to organize information. Two types of fundamental spatial contexts are presented: vertical which are based on hierarchy and horizontal which are based on connectivity.

I Introduction

In this paper, we are considering the role and uses of the user interface in a microcomputer based atlas system. We concentrate to the information search, i.e. database query situation, because atlas is a communication medium, a tool for querying and presenting information.

As the CD-ROM technology and high resolution colour displays are now becoming commonplace, a wave of microcomputer atlases has emerged. But our minds are still attached to the old. Most of these new “electronic atlases” look and behave like the old atlas books. This is natural, but it does not justify the use of computer as a medium. Real books are still more practical to carry around and nicer to look at than computers.

Interactivity is the main feature which distinguishes computers from other kinds of media, such as books or television. The user of a computer application can communicate with the system. The working session is not completely predefined. The author gives only the rules of the game, but each user may adopt his or her own strategy in the information search and manipulation. Applications which fail to meet this goal are likely to be dull and useless.
Atlas is a product for a very diverse audience. Some of its users may have very little or no geographical or cartographical knowledge, some may be geographers or cartographers themselves. Therefore, its conceptual clarity and ease of use are very important factors in determining its usefulness. The success and usability of a computer atlas is judged on the basis of how well it performs in relation to an atlas book. People will choose the book if the computer version is too difficult to use. The real challenge in the user interface design is to make the computer atlas as easy - or even easier - to use as an atlas book and at the same time, more powerful in the information search, manipulation and presentation capabilities.

2 What is a Computer Atlas?

First we have to ask what is an atlas. Atlas could be described as a collection of related maps. It is much like an encyclopedia, only the things are not arranged alphabetically, but according to geographic space. Thus also the presentation does not use so much text but space, i.e. maps.

An atlas gives an areally organized view of its subject. It is a summary of the knowledge and beliefs of its time. This idea of an atlas as a presentation of the outlook of the world is seen for instance in how much the contemporary world atlases contain general geographic theories, such as plate tectonics and atmospheric processes.

Monmonier (1981) emphasizes the importance of atlas as a coordinating tool in presenting data. The receiver gets a much more comprehensive understanding of the multitude and unity of the things in the subject area than he or she would get from studying separate map sheets. In an atlas the information is presented in uniform symbology, scales and projections.

The task of a computer atlas is the same: to give a coordinated presentation of the current understanding of its object region. Its structure, however, is different. The
information and its presentation are separated. The former is coded in the database operated by data management system and the latter takes the form of the user interface. Queries to the database form the intersection of these two (Fig. 1).

‘Maps’ in a computer atlas differ from the maps in atlas books. Because the computer is an interactive medium, there is no need to restrict to any predefined (sub)region or theme in a display. All the data from all of the object region is available, so the map display can change according to user's wishes or over time. The region in view, scale, projection, thematic contents, symbology – everything can be made user-definable, if so wished.

2.1 Data Management

The amount of data which is handled even in a moderate geographical application is usually so large, that it needs a database in one form or another. Data management consists of the creation, use and maintenance of a database. A database is a set of related data. Elmasri and Navathe (1989: 3 - 4) list some further properties of a database:

- A database is logically coherent collection of data with some inherent meaning. A random assortment of data cannot be referenced to as a database.
- A database is designed, built and populated with data for specific purpose. It has an intended group of users and some preconceived applications in which these users are interested.
- A database represents some aspect of the real world, sometimes called the miniworld. Changes to the miniworld are reflected in the database.

One essential property of a database is that it contains a description of the structure and organization of the data and that queries can be made on the basis of the description. The individual data items have relationships which are independent of their physical location in files. So the user interacts with the data itself, not with data files.

The description of the structure and organization of the data is called the conceptual schema of the database. This is the very heart of the system. The conceptual schema contains an abstract representation of the properties, contraints and relationships of the data items stored in the database. Thus it is an abstract and formal description of the general structure of the mini-world mentioned above.

The set of procedures used to operate the database is called database management system (DBMS), which can consist of one or more programs. Although we are talking about a separate database here, in practice the data management in a computer atlas may be merged with the user interface into one application. The user may not be able to differentiate between the two. Within the application, however, this distinction has to be made. The data management and user interface should be kept conceptually apart because this promotes modular design. Thus the implementation of one of these can be changed without affecting the other.

2.2 User Interface

User interface is the common language between the user and the application. All information change between the user and the application is part of it. The minimal user interface for a program might be just the starting command:
system> run backup.exe

with no feedback to the user.

Another way of describing the user interface is to compare it to a map. In a query system the user interface is a tool for viewing the miniworld described in the database. In the same way map is a tool for viewing the real world. We could say that map is a query interface for the real world. Thus in a sense, the designer of the user interface is the cartographer, or the other way round. The main difference between paper and display maps is that with the computer the reader has possibility to manipulate the map. The unidirectional communication from the cartographer to the map reader via the map becomes circular interaction with the database, as in Fig. 2.

The task of the user interface is to: 1) inform the user of the state of the system; 2) present choices of action to the user; 3) receive the user's input and translate it to commands to the processing software and DBMS; 4) translate the required information from the database's internal representation or language to a language understood by the user.

What is the language understood by the user? Textual data can be presented as such, statistical data are often presented with tables or diagrams and data with a spatial dimension are usually presented with maps. However, the map is not the only way of presenting spatial data. Cross-sections, perspective views, etc. can be constructed from the data. Digitized photographs, sounds and video can also be used to give further information of the object region. Besides, the map does not have to be static. Animation is an efficient way of presenting change over time.

Whether maps can be considered a language has been a subject of keen debate (e.g. Robinson & Petchenick 1977, Board 1981, Andrews 1990). We can not delve deeper into this subject here, but a more detailed discussion can be found in Lindholm and Sarjakoski (1992). If a language is defined as a set of signs and a set of rules governing the use and interpretation of them, maps are a language.

A communication model for a query interface to a geographic database is depicted in Fig. 2 (Lindholm & Sarjakoski 1992). The ovals are information sets coded in some language, the boxes are translations between these languages and the arrows show the flow of information. The two main parts are the user communicating with the database management system. The user interface is divided into input and output languages. The human user can receive messages through perception, mainly by seeing and hearing, and send messages through physical action, e.g. by pressing the keys on the keyboard or moving the mouse. The DBMS can receive messages as queries and send messages by mapping the result of the queries to the output language. ‘Mapping’ is here understood as any way of making the information accessible to the user. Thus it contains the production of maps, diagrams, text and other kinds of media. Via the input the user can affect the mapping process, i.e. how the data is to be displayed; for instance change the colours in a choropleth map.

3 Levels in the User Interface

Below the user interface is divided into three levels: conceptual, functional and appearance. This outline is a modification of a division by Foley et al. (1991: 391-395) for user interface design. Their model consisted of four design phases: conceptual, functional, syntactical and lexical. The two first phases or levels are about the same as here, but the other two are more concerned with the detailed implementation of the functions.
Here, we have brought up as the third level the visual appearance of the application, which includes the design of the maps. In a cartographic application the visual design can not be considered only as a part of the technical implementation of the user interface but on the same level of importance as the general conception of the application and the operations needed.

These three levels also present the phases of the user interface design, like in the division by Foley et al. Our scheme does not take into account the detailed implementation of the features in the user interface, the actual programming. From the point of view of cartography, design is the most interesting part of the work. The conceptual level is usually outlined first, before any decisions of the functionality and appearance of the application are made. The meaning of the application results in its functions. The functionality in turn, must be known before the outlooks of the application can be created. Form follows function.

The following discussion is based on the assumption that a computer atlases are designed for a group of users with somewhat similar information needs. If several different views to the database for people with different tasks were provided, the database should be fairly independent of individual views. Instead, it should provide a generic abstract model of the reality as in geographic information systems. The way we see it, however, is that computer atlases are likely to be mass market products for personal use at home or in schools in the manner of word processors. Its users will not have very specialized information needs.
3.1 Conceptual level

The conceptual level consists of the general idea of use and operation of the application: what need does the application meet and how is this goal reached? On a very general level, Fig. 2 presents a conceptualization of a database query system. But more detailed knowledge is needed to develop a proper conceptualization of the user interface. The designer has to know the intended audience of the application and what is to be communicated. In an atlas, much of this is known beforehand. The aim is to provide information of some part of the Earth's surface or of the whole of it. This information is communicated mostly in the form of maps. Thus the main design questions are: what information? how is it retrieved?

The conceptual level of the application depends on the conceptual schema of the database. The designer has to know what features and what attributes are stored in the database and what are their relationships. In other words, we have to know what the minidworld is like, into which we are providing a view. The conceptual design is perhaps best combined with the design of the database.

When using a computer application, the users tend to form personal mental models of the structure and functionality of the application, a framework within which the application becomes comprehensible (Norman 1986). This understanding is mostly based on the functionality and appearance of the application. Fig. 3 shows the aim of the conceptual design. On the left, different views of reality are presented, some of them overlapping somewhat. To the right, the ideal situation is depicted. The view given by the user interface of the study region matches the view coded in the conceptual schema of the database. From the user interface the user adopts this same view. Hopefully, this knowledge is not fully contained in the user's earlier conception of the world, so that he or she learns something new. A unifying framework is needed to manage the design of these three different aspects.

Often the user's mental model takes the form of a metaphor (Carroll & Olson 1988). The application is seen in the terms and concepts of a more familiar phenomenon. "Word processor is a typewriter" is a typical metaphor. Metaphors have also been used deliberately in user interfaces to make the applications easier to learn and use (Apple Computer 1987). A well designed metaphor could guide the user to form a 'correct' understanding of the use of the application.

What would be a good metaphor for a computer atlas? Many writers have suggested that the popular desktop metaphor is not very well suited to applications handling spatial data (e.g. Gould & McGranaghan 1990, Raper & Rhind 1990). A straightforward approach is to make the computer atlas to resemble a book. Such an application would contain the normal parts of an atlas: the general maps, the thematic maps, the text part, table of contents and index of place names. Information would be searched by using the table of contents and index or simply by browsing the 'pages' of the atlas.

If well done this kind of application is easy to learn and use. People know how to use an atlas book so they do not need to adjust their existing mental models of an atlas and the product is more easily accepted. On the other hand, though, why should we artificially impose on ourselves the restrictions of another technology when a computer is capable of doing many things a book can not and is inferior to a book in some other ways? Marchioni (1989) studied subjects using an electronic encyclopedia and found out that they tried to use it like an ordinary printed encyclopedia. They searched information using only the article names. They did not use the possibility to look for key words within text and they did not include logical connectives (and, or, not) in their searches. To summarize, atlas book metaphor may be useful especially for new and casual users, but it may severely limit more experienced users.
Another possible metaphor could be the globe, which most people are also familiar with. On the display there could be a virtual globe, which can be rotated in any direction and zoomed in or out. When the globe is 'dark,' the physical features are visible. When the ‘light’ is turned on, the cultural features emerge. The use of the orthographic projection, i.e. perspective view of the Earth, together with the rotation possibility could reduce the false perception of the form and size of the continents caused by distortions in the world map projections.

One step beyond the virtual globe is the virtual reality. Kuhn (1991) has suggested metaphors that are based on the human senses. The application could be thought of as a view into miniature world. This fits nicely together with idea of a miniworld in the database. The task of an atlas is to provide a unified and organized view into the reality, so a quite obvious way to do this is to build a model of (some part of) the reality into the database and then a user interface to view it. This doll's house -approach lends itself well to displaying topographic data, but how about thematic data, much of which can not directly be seen? The bar and circle diagrams, isolines, flow diagrams, etc. do not exist in reality, but we need them to visualize things which are not visible in the landscape. How will the virtual reality support display of thematic data?

If no one metaphor seems to cover all the aspects of the application, Gould and McGranaghan (1990) suggest the use of several nesting metaphors. We have in another paper (Lindholm & Sarjakoski 1992) suggested an ‘extreme’ user interface, where each user could form their own metaphor. Some writers (e.g. Nelson 1990) have questioned the whole use of metaphors. Instead, Nelson promotes iterative design based on some unifying conceptual and functional principles without the burden of having to adhere to any metaphor.

Fig. 3. The aim of the conceptual design is to unify the user interface’s view of the database and the user’s mental model of the application with the conceptual schema of the database.
3.2 Functional level

The functional level of the user interface includes the meaning of all the operations the user can do and the way they are done. In an atlas, the user's tasks consist of retrieving, combining and presenting the data from the database. Additionally, there can be supportive tasks, such as printing a map display or exporting the data to another application. Referring to Fig. 2, the sentences of input and output languages are defined.

To produce a display map the region, scale, projection, theme, the visible features, the map type and symbolization must be specified. Many of these, e.g. a suitable projection, map type and symbolization, can be decided by the designer and implemented as default values. In a series of displays it is natural that all the parameters which are unchanged from the previous map will affect the new one. It might be a good idea to start the working session with a default map so the user has right away something to refer to.

The division of these tasks into subtasks and their actual implementation depends on the decisions made at the conceptual level. If a certain metaphor has been chosen the implementation must fit naturally to it. For instance, if the application was made to resemble an atlas book, the “zoom in” operation could be implemented with a “magnifying glass”-tool which would ‘enlarge’ (i.e. show in a larger scale) a part of the map. It would look like a magnifying glass and it could be moved around on the display. The scale of the enlargement could depend on how “far away” from the map the magnifying glass was held.

Consistency in functionality is very important. Similar tasks should be performed similarly, using consistent terminology. Consistent feedback should also be provided. Errors should be reported always the same way. Confirmation of a successful execution of a task should be the same to every operation. (Shneiderman 1987)

There are still more design aspects to consider. Is the operation done interactively (i.e. instantly, such as selecting an area) or as a batch work (such as printing)? Should the operations be done in a certain sequence or can they run parallel? Is the application modeless (i.e. all operations are all the time available) or does the task need structured working (some operations might be sensible only in a given context)?

3.3 Appearance level

The appearance of the application is the perceptible form of the functionality. It too is based on the conceptualization. If the user interface is built around a metaphor, the audiovisual appearance must support it. The visual appearance consists of the graphical and alphanumerical symbology needed to control of the application and the graphical presentation of data.

Many practical hints and guidelines for general display design and the choice of symbology for different operations in an application are given for instance in Marcus (1992), Helander (1988), Shneiderman (1987) and Durrett (1987). Some important properties of a good user interface appearance are consistency, ease, clear structure and beauty. A high quality user interface makes the user not only to work more effectively, but to enjoy the work as well.
It was not until the last few years that the technology had advanced so much that displays with decent resolution and enough colours came within the reach of most people. Display maps are relative newcomers in the field of cartography and of many of its aspects no experimental data is available (some examples of experimental investigations are Gilmartin (1988) and Gooding & Forrest (1990)). Because we do not know in which ways the perception of display maps differs from that of the printed ones, no definitive guidelines can be given. Some principles of the use of display maps are studied in Makkonen and Sainio (1991). In the following, we make some notes gained from practical experience in designing display maps for an ongoing project, a thematichal computer atlas of Finland.

The main difference between a digital display map and an analog conventional map is in the structure of the image. Paper is a continuous surface, while the display is formed of small discreet dots, pixels. This difference, however, is not qualitative, but quantitative. If the pixels can be made successively smaller at some point there comes a limit where the eye cannot distinguish individual pixels and the surface seems continuous. For the time being, we can think of the display as a grid of colour values.

Because the resolution of the display is much lower than that of a printed product, the display map can deliver much less information per area unit. This combined with the relatively small size of computer screens – usually less than an A4 paper – results in that display maps have to be much more generalized. The symbols and texts on the display map have to be simpler and larger.

The display map has some advantages over the paper map which compensate these shortcomings. Firstly, there is no need to show everything at once. The user’s attention can be directed to things which are considered important at any given moment. Involved phenomena which used to require complex and difficult symbology can be replaced with showing many maps side by side or in sequence, perhaps using animation. Often it is not even possible to show everything at once. A database can store much more information than a map sheet.

Secondly, the display map can react to user’s actions. Features can be highlighted when pointed at, the symbology can be altered, scale and projection can be changed, the application can give spoken or written explanations of each feature over which the cursor travels, etc. If there is a possibility to update the database, the map can update itself immediately after the change. The user can have the ability to make a cross-section profile along any line, and so on.

Thirdly, new ways of communication can be added to maps. Aerial and landscape photographs, video clips, sound environments, spoken guidance, animated maps and drawings can mixed with the map. After successive zoom ins the map display could be changed to an aerial photograph, for instance.

Geometric accuracy of measurement is an important requirement for conventional printed maps. The map reader has to be able to measure locations, directions, distances, areas, slopes and volumes from topographic maps or exact thematic values from the diagrams, dots and isolines of thematic maps. This requirement dates back to the maps earlier role as a storage of data. In modern database based systems the task of the map has changed. The exact data values can always be made available and there is no need to measure them from the map. The display map is like a mental map in this regard (see Gould & White (1968) for mental maps). It can depart from the measured geometric reality arbitrarily. There is an increased opportunity of lying for cartographers, but also an increased opportunity of emphasizing things that are not easily seen in a conventional map.
4 Interaction styles

In modern interactive computer systems the sentences of the input language are usually formed using a keyboard and a pointing device, most commonly a mouse. Also speech-operated applications are gradually emerging. The output language is usually transmitted via the screen and speaker.

Shneiderman (1987) distinguishes five different styles of interaction between human and computer:

**Menu selection.** The user is presented a list of choices from which one or sometimes several alternatives may be selected. Employing menus in the user interface reduces errors, because the user can not give a command the application does not understand. They also help the user to remember all the different commands and options available. Well-defined menus structure decision-making. On the other hand, long menus can be difficult to comprehend and many submenus may slow the work of experienced users. This problem can be alleviated by careful menu layout.

**Form fill-in.** Text is typed into the fields of a form-looking display. This form of interaction is of limited usability. It suits well to data input and to a query-by-example type of database interfaces when dealing with alphanumeric data.

**Command language.** To this class belong also the database query languages, such as SQL. The user types one or more commands to be executed immediately or to be stored in a file for later use. Command languages are the most flexible means of interaction. Very complex operations can be described in a compact form. It is easy to create macros for frequent operations. Command languages are however, difficult to learn and remember. They also invoke many errors, because of wrong parameters or typing errors. In geographic applications, character-based command languages bring an unnecessary abstraction level between the user and the data. A solution to this might be the use of graphical query and command languages, such as described in Ichikawa et al. (1990) and Maiguenaud and Portier (1990).

**Natural language.** The user acts as with a command language, but the commands are given in natural language, e.g. English. Spoken input of natural language would be preferable. Due to the ambiguousness of natural language, there is frequently a need for confirmations and clarifications, which slows down working and annoys users. Furthermore, in an atlas the use of natural language interface may not be useful, because it is quite cumbersome in describing many spatial concepts. That's why we have maps in the first place!

**Direct manipulation.** The application's concepts have a pictorial representation, e.g. icons, which can be manipulated with a pointing device, such as the mouse. If a good visual expression to the application's concepts can be found, then this form of interaction is very efficient. Direct manipulation interfaces are easy to learn and remember. They encourage exploration and experimentation. Quality is an important factor here also. If the visualization is done badly, it is only misleading and will frustrate users.

In practice the designer usually has to blend different types of interaction depending on the task. In an atlas, for instance, the selection of a region to zoom in might be done with directly framing the area from the map display with the mouse. An index of place names in turn, would probably be best implemented as a menu.

Direct manipulation seems to suit well to the interfaces of geographic databases, because the database objects have a spatial/visual representation already. The communication between the user and the database through an interactive map would truly be a step
beyond the conventional atlas books. A graphical query language implemented in a
direct manipulation fashion with the aid of some menus seems to be a powerful means
for controlling a computer atlas.

5 Hypermedia

Above we have stated that the conceptual model of reality which is coded in the
database should become the conceptual model of reality seen in the user interface. One
method of doing this is hypermedia. A hypermedia application is a collection discrete
elements or packets which may contain any kind of information. These nodes are
connected to each other with links which organize the information into semantic constructs.
Information search from a hypermedia application is called browsing.

A node usually has a visual representation on the display, e.g. a window, a text or an
image field. Links from this node to others are expressed in some way on the display:
within or near the node or in some reserved area. They may have the form of a word or
an icon. By activating the link, e.g. by pointing with mouse cursor at the link icon, the
node to which the link points to is brought up. Such direct manipulation as a means of
interaction is a very intuitive way to implement a user interface.

In recent years many studies of the use of hypermedia in GIS and computer-based map
systems has emerged. Wallin (1990) introduced the concept hypermap, Laurini and
Milleret-Rafford (1990) presented a data structure for hypermap applications. Some
further studies are Camara (1991), Lewis (1991) and Raper (1991). However, in most
of these papers maps are seen as single nodes, separate entities connected to each other
and other kinds of documents via the links. In our opinion, the hypermedia concept
should be planted deeper into the structure of the database. The links should be
between features, not map sheets. This is where the greatest improvements in information
search can be made.

A computer atlas contains a lot of clearly distinguishable, individual features which
have a clear visual representation. Between them there are many different relationships
which we use to give order to the space when reading a map. This knowledge can used
to create a hypermedia atlas or hyperatlas. The concept of browsing a hyperatlas
conforms well to the general idea of an atlas as tool for information search and presentation.
In a hyperatlas all relevant information is never further than one step away; we can
“turn pages” (i.e. traverse links) to many different directions, not only forward or
backward. In an atlas book, by turning the page, we get only the next or the previous
map in sequence.

In a hypermedia application the relationships in the database are transmitted as such
into the user interface. Thus the database and the user interface share the same view of
reality. The main advantage of hypermedia is, however, the possibility to give several
different organizations to the same information elements. The same set of nodes can be
given a different structure when looked at from another aspect. A context can be
defined as a set of links which impose a certain structure on a set of nodes.

Geographic objects appear very different in different contexts and there are so many
different relationships between them that it is not possible to show them all at a time.
The use of contexts seems to be a useful way of structuring the user interface and to
reduce the browsing possibilities to a manageable amount. While individual features,
such as lakes, rivers, roads, fields, cities, countries, etc. form the nodes in a hyperatlas,
the contexts represent different spatial structures.
Fig. 4. Spatial contexts. The example region is divided into subregions on three levels of hierarchy. The vertical relationships between the subregions are shown with dashed lines and the horizontal relationships with double-ended arrows. For instance, the vertical context of the subregion a on the level II is that it belongs as a part to the region c on level I. On the other hand, it divides on the level III to subregions a, b and c. The horizontal context of the region a of level II is that it is a neighbour of the regions b and c on the same level.

In Lindholm (1991) two main types of geographical or spatial contexts for a computer atlas are defined: vertical and horizontal (Fig. 4). Vertical contexts are hierarchies within spatial features. They form tree-like branching structures. There are two kinds of hierarchies: functional and classificatory. The functional hierarchies are formed of the ‘real’ functional relationships between features, which are inherently hierarchical. Typical examples are administrative regionalization (e.g. country – county – municipality – village) and the hierarchy of drainage basins of a river system.

Classification hierarchies are created by dividing the face of the Earth into homogenous regions, often using some quantitative criteria. This way the world has been divided into zones on the basis of soil classifications, for instance (an overview of the soil classification system and the corresponding zones is presented in e.g. Strahler & Strahler 1987: 392-415). The main zones divide into subzones which divide into subzones, and so on. A hierarchy of increasingly smaller and homogenous regions is formed.

Typical classification hierarchies are the climatic and vegetation regions. Also cultural phenomena – languages as an example – have been classified this way. But these regions and hierarchies do not exist in reality. They are only conceptual generalizations of geographical reality. The North American prairie as such has nothing to do with the Central Asian plains, but on the basis of their form they are classified to the same category, grasslands.
Horizontal contexts are relationships between features of equal importance. A county for instance has vertical relationships to the country it belongs to and to its municipalities. Besides these, it has also horizontal relationships to other counties. Usually these are neighbourhood or connectivity relationships. Horizontal contexts form network-like structures. The "has common border"-relationship is a typical horizontal context within a set of regions. Communications between cities is another.

The realization of the contexts in a hyperatlas is done so, that the user employs one context at a time to browse the nodes. At any time another set of links can be loaded and the context changes. If the user found a region through hierarchical search, he or she can then change the context and begin to look at its neighbours.

6 Conclusion

In this paper, much emphasis has been given to the conceptual design. This is the most demanding and often neglected part in the creation of a computer application. The designer must develop a comprehensive understanding of what he or she is doing and why. After that, the technical implementation of the application is trivial, despite many practical problems which often arise.

On the implementation side, more scientific studies of display map perception are needed so that guidelines for map design can be developed. Even though most of the conventional cartographic presentation techniques and principles seem to apply well the creation of display maps, the new possibilities of the computer have to be explored. A new cartography is needed for computer atlases.

Some of the ideas presented above are being implemented and tested in an ongoing development project, a thematic atlas of Finland1. The development is done jointly with Finnish Geodetic Institute and National Land Survey. The atlas is aimed to be an end-user product for homes and schools.

References


1 This program will be on view at the map exhibition of the 16th International Cartographic Conference 4.-8.5. 1993, Cologne.


La production d'un atlas environnemental: le cas du Saint-Laurent

J. Raveneau (Québec, CDN)

L'Atlas environnemental du Saint-Laurent est un ouvrage dont le but est la diffusion de connaissances sur l'état de l'environnement du fleuve Saint-Laurent, au bénéfice de publics non spécialisés. C'est une publication du Centre Saint-Laurent d'Environnement Canada (Montréal), préparée dans le cadre du Plan d'action Saint-Laurent. L'atlas est réalisé au Département de géographie de l'Université Laval, à Québec. Il fait partie d'une série d'ouvrages du Centre Saint-Laurent visant à préparer un bilan de l'état du fleuve. Il est également destiné à servir d'instrument pour la recherche de solutions appropriées à la protection, la conservation et la mise en valeur de l'environnement fluvial. Dans cette présentation nous aborderons successivement le rôle de l'atlas comme outil d'éducation environnementale, la structure du plan de l'atlas, sa forme physique, le traitement de l'information, les moyens graphiques utilisés pour communiquer des connaissances environnementales structurées, les méthodes infographiques de production des images et les perspectives offertes par les supports électroniques multimédia.

L'Atlas environnemental du Saint-Laurent : un outil d'éducation à l'environnement

L'Atlas environnemental du Saint-Laurent a pour but de fournir au public une vue synthétique des connaissances existantes sur les composantes de l'environnement fluvial et sur leurs interrelations. Les objectifs de l'atlas sont communs à ceux de l'établissement du bilan de l'état du fleuve; il s'agit d'informer la population sur les ressources et les usages du Saint-Laurent, de prévoir et d'évaluer les tendances et les changements des conditions environnementales et d'influencer les décisions relatives à la gestion des ressources. Par le biais de l'atlas et d'autres documents, le bilan doit également répondre à un objectif d'éducation environnementale. Un changement d'attitude à l'égard de l'environnement dépend avant tout des efforts d'éducation auprès de la population en général et des jeunes en particulier, afin de promouvoir la connaissance de l'environnement et de susciter la prise de décision éclairée dans toutes les couches de la société. Une meilleure prise de conscience environnementale se répercutera sur la qualité des interventions des citoyens, autant dans les processus de décision reliés à des projets particuliers que dans ceux portant sur la planification et l'élaboration de politiques gouvernementales.

En fonction de ces objectifs généraux l'Atlas environnemental du Saint-Laurent s'adresse plus à un public de jeunes (étudiants) et à un public général qu'à des chercheurs spécialisés. Il vise en particulier les clientèles suivantes:

- le public scolaire, depuis la troisième année du secondaire, en passant par les collèges d'enseignement général et professionnel (niveau pré-universitaire), jusqu'au 1er cycle universitaire:
- le public en général, ayant une formation secondaire ou collégiale, intéressé par la problématique du fleuve, sa protection et sa mise en valeur;
- les organismes environnementaux régionaux et nationaux et leurs gestionnaires;
- les municipalités riveraines du fleuve et leurs gestionnaires;
- les groupes communautaires et les groupes d'éducation environnementale;
- les associations touristiques et autres organismes publics ou privés dont la vocation est de diffuser des informations concernant le fleuve.

La structure et le contenu de l'atlas: une approche systémique

La structure générale de l'atlas est fondée sur les préoccupations des citoyens, utilisateurs et protecteurs du fleuve, en fonction des questions qu'ils se posent. Ces préoccupations découlent du fait que le Saint-Laurent constitue un cadre de vie pour plus de quatre millions de personnes qui habitent à moins de 10 km de ses rives. La richesse et la diversité des écosystèmes qui caractérisent le Saint-Laurent représentent un patrimoine d'une valeur unique qui mérite d'être mieux connu. Aussi, les questions que les citoyens se posent face au fleuve et les besoins d'information qui en découlent ont été à la base de l'élaboration du cadre structural de l'atlas:

Les questions considérées

- Qu'est-ce que le Saint-Laurent?
- Que doit-on protéger?
- Que fait-on du Saint-Laurent?
- Quel est son état?
- Comment le protéger?

Les besoins d'information subséquents

- Importance et caractéristiques
- Ressources- Valeurs
- Usages
- Détérioration
- Protection

La structure de l'atlas comprend donc cinq volets qui découlent des questions ci-dessus:

1) Milieu naturel
2) Milieu humain
3) Usages des ressources
4) Dégradations
5) Conservation

Les cinq volets de cette structure sont étroitement liés entre eux et constituent le système Saint-Laurent. Le premier volet est destiné à donner au lecteur les connaissances générales de base sur l'environnement physique et l'écologie du fleuve, connaissances qui sont indispensables pour comprendre les problèmes environnementaux. Le second volet vise à décrire le processus historique de mise en place du peuplement et des activités économiques le long des rives du Saint-Laurent et à expliquer le mode d'occupation actuel du territoire. Dans le troisième volet, les usages du fleuve et l'exploitation des ressources qui y sont associées sont présentés en évoquant les impacts qui en résultent sur la qualité du milieu fluvial et riverain. Dans le quatrième volet, la question de la qualité des eaux sera abordée de même que les phénomènes de pollution, de détérioration des rives et de destruction d'habitats associés aux différentes formes d'exploitation du milieu. Le dernier volet présente les éléments de restauration, de conservation et de mise en valeur de l'environnement du Saint-Laurent, ainsi que les attitudes de la population à l'égard de la protection du fleuve. À l'intérieur du système Saint-Laurent, toute modification d'un élément a des répercussions sur d'autres éléments, c'est pourquoi dans chacune des planches l'accent a été mis, chaque fois que possible, sur les impacts de l'occupation humaine, des usages et des dégradations sur le milieu naturel et sa conservation. À l'inverse, certaines

Tableau 1

**ATLAS ENVIRONNEMENTAL DU SAINT-LAURENT**
Liste des thèmes des planches

**LE MILIEU NATUREL**
- Les grandes divisions hydrographiques*
- Le littoral du Saint-Laurent**
- Les milieux humides*
- Les écosystèmes riverains et insulaires
- Les écosystèmes des eaux douces et saumâtres**
- Les écosystèmes marins*

**LE MILIEU HUMAIN**
- La mise en place du peuplement (17e et 18e siècles)*
- L'urbanisation et l'industrialisation (19e et 20e siècles)*
- La population et l'occupation des rives**

**CONSERVATION**
- La conservation des patrimoines**
- Les attitudes face à la protection du fleuve

**LES USAGES**
- La navigation commerciale*
- La pêche commerciale**
- Les loisirs et la récréation
- L'exploitation des ressources
- Eau potable et eaux usées

**LES DÉGRADATIONS**
- La pollution du Saint-Laurent: polluants et pollueurs
- La qualité des eaux et des sédiments
- La dégradation des écosystèmes

**PLANCHE DE SYNTHÈSE**

* Planches imprimées en 1991-92; ** Planches terminées, à imprimer en 1993

**La forme physique de l'atlas**

Le support physique de l'atlas environnemental a été déterminé en fonction du type de public à rejoindre, tel que défini plus haut. Bien que la technologie de réalisation d'un atlas électronique multimédia soit opérationnelle (voir la dernière partie de cet article), il a été décidé d'opter dans un premier temps pour un atlas sur papier. Celui-ci présente en effet un potentiel de diffusion beaucoup plus grand qu'un atlas électronique dépendant d'un appareillage de lecture encore peu répandu parmi les utilisateurs visés par l'atlas.

Un atlas conventionnel prend généralement l'aspect d'un recueil de cartes rassemblées sur des planches (1 planche = 1 page d'atlas), reliées à l'intérieur d'une couverture plus ou moins rigide. La forme actuelle de l'**Atlas environnemental du Saint-Laurent** consiste en une série de 20 planches autonomes de grand format (64 x 110 cm). Les planches sont imprimées et distribuées séparément au fur et à mesure de leur achèvement; six planches ont été publiées en 1991-92, cinq autres sont terminées et doivent être imprimées en 1993; l'ensemble de l'atlas devrait être complété vers 1995. Les planches sont distribuées en format plié (style carte routière, 18 x 21 cm) ou roulé. La publication se fait simultanément en français et en anglais, en deux éditions distinctes. Cette forme de publication en planches séparées permet de diffuser
l'information environnementale plus rapidement et favorise la mise à jour éventuelle des planches.

Le grand format des planches est justifié par l'ampleur du territoire à représenter et l'obligation de conserver une échelle compatible avec les impératifs de localisation des phénomènes; il est aussi nécessaire pour mettre en œuvre le mode de traitement de l'information et le concept graphique de l'atlas, décrits plus loin, impliquant la visualisation de nombreuses images complémentaires aux cartes sur la même feuille de papier. En effet, chaque planche de l'atlas est organisée selon le concept de «planchette-affiche» ou «d'atlas en une page»1, réunissant plusieurs blocs d'information disposés selon une séquence de lecture matérialisée par des numéros.

**Le traitement de l'information environnementale à caractère spatial: inventaires et SIG vs présentation de connaissances géographiques structurées**

L'une des fonctions conventionnelles des atlas est de servir de support analogique aux inventaires spatiaux sur le territoire qu'ils couvrent. Or cette fonction est de plus en plus prise en charge par les systèmes d'information géographique (SIG). Les SIG permettent de pallier à certains inconvénients des atlas conventionnels dont celui de l'obsolescence rapide des données. Avec un SIG, les données environnementales peuvent (théoriquement) être mises à la disposition des usagers au fur et à mesure de leur collecte et de leur validation; de nombreuses superpositions et comparaisons de variables, ainsi que des analyses spatiales complexes peuvent être effectuées rapidement; enfin, la visualisation cartographique des données sur ordinateur peut être réalisée dans une fraction du temps requis pour la confection d'une carte manuelle, alors que le changement des paramètres graphiques d'une carte est l'affaire de quelques secondes. Dans ce contexte on peut facilement envisager la préparation d'un atlas directement à partir d'un SIG et ce, dans des délais relativement courts.

Cette vision idéale d'une liaison SIG-Atlas n'a toutefois pas pu être mise en œuvre directement dans la phase initiale de la préparation de l'Atlas environnemental du Saint-Laurent. En effet, la principale difficulté en ce qui concerne la préparation de l'atlas est que celle-ci s'est déroulée en parallèle — plutôt qu'après — avec un programme de recherches et d'acquisition de données sur le fleuve, mené par le Centre Saint-Laurent d'Environnement Canada. Des banques de données environnementales à référence spatiale ont été créées au Centre et sont en constant développement à mesure que de nouvelles données deviennent disponibles. Cependant, la mise au point d'un système d'information géographique complet et intégré sur le Saint-Laurent est encore loin d'être achevée. Un tel système n'a donc pu être utilisé directement pour la confection de l'atlas. De plus, des problèmes conceptuels doivent être résolus, en raison notamment de la multiplicité des formes spatiales revêtues par les phénomènes environnementaux, en incluant la troisième dimension et le caractère dynamique (dans le temps et dans l'espace) des phénomènes. De nombreux inventaires sont discontinus dans l'espace et composés de données fragmentaires. En outre, les dimensions du territoire (1 200 kilomètres de Cornwall à Blanc-Sablon), de même que la diversité des écosystèmes présents dans les différentes sections du fleuve, ne permettent pas de traiter de façon uniforme les phénomènes sur toute l'étendue du territoire fluvial. Sur le plan technique, la variété des progiciels et des systèmes d'exploitation utilisés causent des problèmes de transfert et de compatibilité des fichiers de données spatiales environnementales.

Les banques de données à référence spatiale et les SIG permettent de répondre à des besoins d'information précis, pour des individus ayant accès à la technologie appropriée et possédant le niveau de connaissances adéquat pour analyser des
données dont l'interprétation est souvent complexe. Ces instruments, destinés
d'abord aux chercheurs, sont moins adaptés lorsqu'il s'agit de répondre à des
besoins de connaissances plus généraux, à caractère synthétique, sur les
caractéristiques de l'environnement du fleuve Saint-Laurent et sur leurs
interactions, sous une forme compréhensible par un public non spécialisé, comme
celui qui a été identifié plus haut. C'est pourquoi l'Atlas environnemental du
Saint-Laurent a été conçu comme un ouvrage de vulgarisation à petite échelle,
destiné à présenter un ensemble de connaissances structurées, sous une forme
synthétique, plutôt que des inventaires spatiaux exhaustifs à grande échelle (cette
dernière fonction étant accomplie de manière plus efficace par les banques de
données à référence spatiale et les SIG). Un atlas environnemental ne doit pas
servir seulement à communiquer des données spatiales; il doit être aussi un outil
de visualisation apte à soulever des questions concernant les interrelations entre
les données environnementales. Aussi s'est-on attaché à illustrer, dans les
planches de l'atlas, les processus environnementaux et leurs incidences spatiales
autant que la localisation proprement dite des phénomènes. La présentation
visuelle de connaissances environnementales structurées nécessite un traitement
graphique original pour traduire de façon adéquate les interrelations entre les
données.

Le rôle de la visualisation et du traitement graphique dans la
structuration des connaissances environnementales

L'observation des atlas publiés depuis deux décennies révèle que la carte n'est
plus l'unique support de l'information spatiale. La présentation visuelle de cette
information s'est considérablement diversifiée et l'on associe de plus en plus à la
carte des images de formes variées: photographies, photographies aériennes,
images de télé-détection, graphiques, schémas, diagrammes, dessins, matrices
graphiques, tableaux, etc. Au concept d'atlas «cartographique» s'est substitué le
concept d'atlas multimédia, d'abord sur support papier et maintenant sur support
electronique. D'autre part, l'information géographique n'est plus seulement
illustrée par une collection d'images. Elle est de plus en plus présentée sous la
forme d'un ensemble de connaissances structurées à l'aide d'une combinaison
texte-images qui met l'accent sur les interrelations entre les composantes de cette
information. C'est le concept de présentation dit «scripto-visuel», associant le
texte et l'image. On utilise la forme graphique la plus appropriée pour la
communication de chaque catégorie de données ou d'information de manière à
favoriser au maximum la compréhension des relations entre les phénomènes. Les
représentations graphiques autres que la carte permettent aussi de faciliter la
compréhension des processus qui sont à l'origine de la distribution spatiale des
phénomènes. Dans l'Atlas environnemental du Saint-Laurent les méthodes de
traitement graphique de l'information utilisées exploitent pleinement l'approche
multimédia dans le but de faciliter la compréhension et l'apprentissage des
connaissances environnementales.

Chaque planche de l'atlas environnemental constitue en elle-même un ensemble
d'informations autonome, un «atlas en une page», construit en fonction
d'objectifs précis, avec une introduction, une conclusion et un développement
bâti selon un scénario logique d'organisation des connaissances propre au thème
de la planche. Les objectifs généraux et le scénario provisoire (canvases
préalables) de la planche sont établis avant même la collecte des données.
L'organisation graphique des éléments composant chaque planche est facilitée par
l'utilisation d'une grille modulaire qui permet de créer des blocs d'information
bien délimités et ordonnés visuellement (figure 1). Chaque bloc comprend une ou
plusieurs images et un texte destiné à donner des informations ou des explications
complémentaires aux images, à établir les relations entre les phénomènes et à
préciser les liens entre les blocs. Les titres de bloc sont descriptifs et visent à
résumer l'idée principale traitée à l'intérieur de chacun d'eux.
Figure 1

Exemples de planches de l'Atlas environnemental du Saint-Laurent

(les planches originales sont imprimées en couleurs et ont une dimension de 64 x 110 cm)
Figure 1
Exemples de planches de l'Atlas environnemental du Saint-Laurent
(les planches originales sont imprimées en couleurs et ont une dimension de 64 x 110 cm)
La disposition physique des blocs à l'intérieur d'une planche est organisée de manière à refléter la structuration logique de l'information (le scénario) selon une approche systémique visant à favoriser la compréhension du thème et des interrelations entre les phénomènes illustrés, de même que les interrelations avec les thèmes traités sur d'autres planches. C'est le support visuel du message véhiculé à l'intérieur de la planche.

Les recherches sur la perception visuelle des cartes ont montré que la manière dont l'œil humain lit une carte est un processus largement empirique, qui varie d'un individu à l'autre. Dans l'atlas environnemental, deux moyens visuels ont été utilisés pour inciter le lecteur à décroître une planche selon la séquence pédagogique correspondant au scénario d'organisation des connaissances:
- 1) la numérotation des blocs, qui contribue aussi à pallier certaines contraintes d'organisation graphique qui ne permettent pas toujours de respecter l'ordre de lecture conventionnel du haut en bas et de gauche à droite;
- 2) la hiérarchie visuelle des éléments graphiques, de manière à définir plusieurs niveaux de lecture.

Une planche de l'atlas environnemental ne se lit donc pas comme un texte ordinaire, de façon linéaire. Elle se découvre selon plusieurs niveaux de lecture, depuis le coup d'œil général jusqu'à la lecture de détail, en passant par des étapes d'observation intermédiaires (voir les exemples de planches sur la figure 1). Aux niveaux de lecture d'ensemble et intermédiaire, on vise à faciliter l'appréhension globale de chaque thème et de l'organisation de ses composantes par les moyens suivants:

- utilisation d'une hiérarchie appropriée de la typographie (style de caractère et taille) pour identifier les composantes de la planche: titre, sous-titre, titres de blocs, titres d'images;
- formulation du titre de la planche composé d'un titre principal, imagé, évocateur du thème et d'un sous-titre descriptif, dont la formulation est plus technique; ex.: UN FLEUVE, DES ESTUAIRES, UN GOLFE ; les grandes divisions hydrographiques du Saint-Laurent;
- utilisation de macro-blocs graphiques traduisant visuellement l'ordre logique des composantes d'un thème et divisant une planche en plusieurs aires; ces macro-blocs permettent de définir le schéma d'organisation de la planche, lequel correspond au scénario de présentation des connaissances;
- utilisation de la couleur comme moyen de codage et de rappel graphique pour souligner et amplifier visuellement les interrelations entre certains éléments d'information à l'intérieur de chaque planche, tel que décrit plus haut.

Au niveau de lecture de détail, on fait appel aux diverses formes d'images disponibles (cartes, diagrammes, schémas, photographies, etc.) et au texte pour communiquer les connaissances particulières contenues à l'intérieur de chaque bloc. Ces différents paliers de lecture se prêtent à une exploitation pédagogique à plusieurs degrés de détail et de complexité, qui peut être adaptée au niveau des élèves ou du public utilisant les planches. Dans son ensemble, le concept de «planche-affiche» de grand format permet d'établir rapidement des relations visuelles entre les différents types d'images composant la planche. C'est aussi un bon moyen pour exprimer et faire découvrir en quelques coups d'œil les relations thématiques entre les différentes composantes d'un thème. Cette facilité de découverte visuelle est considérablement diminuée dans un atlas conventionnel où l'on doit passer d'une page à l'autre en perdant de vue la continuité graphique et conceptuelle entre les éléments d'un même thème.
Un atlas entièrement produit par infographie

Une fois que toutes les données nécessaires à la confection d'une planche ont été rassemblées, vérifiées, traitées et préparées pour être visualisées, la production technique des textes, cartes, graphiques, schémas, etc., est assurée entièrement par des moyens informatiques, sur des postes de travail Macintosh. L'idéal, pour la production d'un tel atlas, aurait été de pouvoir procéder à la cartographie directement à partir d'un système d'information géographique. Mais l'état de développement des banques de données à référence spatiale sur le Saint-Laurent n'était pas encore suffisamment avancé au moment de la préparation de l'atlas. De plus, les données nécessaires à la préparation de chaque planche proviennent souvent d'une masse de documents non cartographiques très hétéroclites. La cartographie assistée par ordinateur a cependant été utilisée pour construire la version préliminaire de certaines cartes et la plupart des graphiques ont été bâtis par infographie, directement à partir de tableaux entrés dans un chiffrier. Mais un grand nombre de dessins et schémas, de même que la version finale des cartes et graphiques, ont été traités dans des logiciels de dessin assisté par ordinateur, surtout dans FREEHAND. La figure 2 résume la chaîne de production informatique et les principaux logiciels utilisés.

Figure 2

**LE PROCESSUS TECHNIQUE DE PRODUCTION D'UNE PLANCHE**

**CONSTRUCTION DES CARTES**

<table>
<thead>
<tr>
<th>CARTES DE BASE</th>
<th>BASF DE DONNÉES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carte de base numérisée, Atlas national du Canada</td>
<td>Numérisation CARTIER/DESKSCAN</td>
</tr>
<tr>
<td></td>
<td>Fichiers de données EXCEL</td>
</tr>
</tbody>
</table>

**CONSTRUCTION DES GRAPHIQUES**

<table>
<thead>
<tr>
<th></th>
<th>Graphiques DELTAGRAPH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Graphiques, schémas, dessins divers FREEHAND</td>
</tr>
<tr>
<td></td>
<td>Système SCITEX</td>
</tr>
</tbody>
</table>

**LOGICIELS**
Les différents éléments graphiques composant une planche sont mis en place sur la grille de montage à l'aide du logiciel QUARK X PRESS, permettant ainsi de produire une première maquette de l'ensemble de la planche. Une impression préliminaire est effectuée en noir et blanc ou en couleurs, sur une imprimante à faible résolution (300 points par pouce en noir et blanc, 216 en couleurs), en feuilles détachées de format 8,5 x 11 pouces, qui doivent être ensuite assemblées manuellement. Après analyse et critique de la maquette et les corrections appropriées, on peut ensuite passer à la mise en page définitive, en reprenant et en modifiant les dessins originaux si nécessaire. Le dessin final, équivalent du «prêt à photographier» en cartographie manuelle, est remis à l'imprimeur sous forme de fichiers sur disquette en vue de la séparation des couleurs et de la préparation, chez une firme spécialisée, des planches-mères pour l'impression en quadrichromie; la séparation couleur des photos et leur mise en place sur la planche est également effectuée à la même occasion. C'est probablement la première fois au Canada, sinon en Amérique du Nord, que des planches d'atlas d'une telle dimension et d'une telle complexité graphique sont préparées entièrement sur micro-ordinateur jusqu'au stade de l'impression. La qualité graphique obtenue se compare avantageusement avec celle de planches réalisées à l'aide de méthodes manuelles comme le tracé sur couche.

Cette méthode de production, bien que présentant certaines contraintes graphiques, comporte de nombreux avantages en ce qui concerne la possibilité de modifier et de mettre à jour les dessins et les textes à n'importe quel stade de la production. Elle offre aussi une grande souplesse de conception graphique, grâce à laquelle les personnes chargées de la réalisation graphique des planches peuvent donner libre cours à leur créativité et concevoir des images originales, parfois complexes, qui auraient été difficiles à produire manuellement. Les opérations photomécaniques préalables à la confection des planches-mères sont entièrement éliminées. De plus, grâce à cette technologie, les planches de l'Atlas environnemental du Saint-Laurent pourront, le cas échéant, être mises à jour et rééditées au fur et à mesure que de nouvelles données pertinentes à un thème seront disponibles. Les différents fichiers composant une planche sont aussi réutilisables pour l'éventuelle production d'une version «atlas électronique».

De l'atlas papier à l'atlas multimédia sur support électronique

L'essor de la micro-informatique dans les trois dernières années a amené le développement des applications multimédia sur vidéodisque et CD-ROM (Compact Disk - Read Only Memory). Celles-ci peuvent être définies comme étant des applications informatiques comprenant du texte, des graphiques, des cartes, des images (animées ou non) et des sons. L'utilisateur consulte ces données et informations de façon interactive. Cet outil présente un grand potentiel pour faciliter la compréhension de l'état de l'environnement d'un grand fleuve comme le Saint-Laurent et celle des phénomènes écologiques l'influencant.

Le processus de développement des applications multimédia

L'élaboration d'applications multimédia en environnement fait appel à trois principaux types d'intervenants possédant des compétences spécifiques: le spécialiste en environnement, le pédagogue et l'analyste-programmeur. Le schéma de la figure 3 illustre les relations entre les activités et les individus impliqués dans l'élaboration d'une application multimédia de bas ou de haut de gamme.
Processus de développement d'une application multimédia en environnement

Gestionnaire scientifique

Définition de l'application

Décisions d'orientation

Buts et objectifs

Rédaction

Analyste-programmateur en multimédia

Acquisition des données et des informations

Standards de développement

Programmation et tests

Validation

Application multimédia sur support magnétique ou optique

Spécialiste en environnement

Développement de l'application

Conception

Pédagogue

Diffusion de l'application

Utilisateur

Mise à jour de l'application

Appréciation commentée

Modifications
La réalisation d'une application multimédia implique :

- la définition de l'application en fonction du groupe d'utilisateurs visé ;
- l'acquisition, l'analyse et la validation des données et des informations environnementales ;
- le développement de l'application avec des logiciels appropriés et les tests de toutes ses composantes ;
- la reproduction et la diffusion du produit sur des supports magnétiques ou optiques ;
- la mise à jour de l'application en tenant compte des commentaires des utilisateurs, des nouvelles connaissances en environnement et des nouveautés sur le plan informatique .

Les composantes matérielles et logicielles d'une station multimédia

Les applications multimédia nécessitent des processeurs traitant rapidement des instructions informatiques, comme les 80486 d'Intel, les 68040 de Motorola et les RISC (Reduce Instructions Set for Computers). Les capacités de stockage des données doivent être grandes, de l'ordre de 500 megabytes. On utilise donc fréquemment des lecteurs de CD-ROM. Une application de haut de gamme permet l'animation des images et l'obtention de sons de meilleure qualité mais exige, pour la production et la lecture, un matériel informatique spécialisé .

Il existe également une grande variété de progiciels commerciaux et privés permettant la réalisation d'applications en multimédia. Le choix du progiciel dépend de facteurs comme la possibilité de produire des applications qui fonctionnent indépendamment du progiciel, sa convivialité, le support informatique de la compagnie responsable du progiciel, sa capacité d'importer des fichiers créés par différents progiciels graphiques, et son coût. Le Centre Saint-Laurent procède actuellement à des expérimentations pour la production de documents multimédia au moyen de micro-ordinateurs gérés par le système d'exploitation DOS, dans un environnement graphique WINDOWS. Les progiciels TOOL BOOK et MULTIMEDIA RESOURCE KIT sont utilisés pour ces recherches exploratoires qui pourraient déboucher sur la production d'un atlas électronique multimédia.

Des données et des informations environnementales déjà informatisées

L'intérêt et la valeur des applications multimédia reposent non seulement sur le degré de sophistication des technologies utilisées mais aussi sur la qualité des données et surtout des informations contenues dans les applications. En ce qui concerne le Saint-Laurent, les données environnementales acquises sur le terrain et par télédétection aéroportée et satellitaire sont traitées au Centre Saint-Laurent à l'aide de systèmes d'information géographique ou de cartographie assistée par ordinateur. Ces données sont gérées par le Centre d'acquisition et de traitement de données informatiques sur le fleuve Saint-Laurent (Centre ACTIF) et sont utilisables pour le développement des applications multimédia. Par ailleurs on peut aussi exploiter les données de la présente version, imprimée sur papier, de l'Atlas environnemental du Saint-Laurent. Comme celui-ci a été réalisé presque entièrement à l'aide de logiciels d'infographie et d'édition, les informations utilisées pour sa préparation existent déjà sous forme numérique et peuvent donc être modifiées facilement pour l'élaboration d'applications multimédia.

L'organisation de l'information de l'atlas papier permet aussi d'accélérer grandement le design d'un atlas électronique multimédia. Pour l'élaboration d'un tel atlas, la manipulation des fichiers créés par différents progiciels gérés par des systèmes d'exploitation variés cause encore certains problèmes de transfert. Ces derniers deviennent toutefois de plus en plus faciles à solutionner.
L'analyse de marché pour un atlas électronique multimédia

L'analyse de marché pour des applications multimédia doit tenir compte: 1) du matériel informatique que possède l'utilisateur potentiel, 2) des besoins de ce dernier et, 3) des buts et objectifs de l'organisation. Des enquêtes effectuées dans les foyers du Québec ont démontré qu'on y trouve présentement plus de 100 000 micro-ordinateurs (de type IBM surtout), d'une puissance suffisante pour la visualisation de documents multimédia. À l'intérieur de l'échantillon consulté, 56 % des utilisateurs se sont déclarés intéressés par les questions environnementales et 65 % par l'utilisation éventuelle de documents multimédia. Dans les écoles secondaires, 400 000 étudiants et 29 000 professeurs disposent de quelque 21 000 micro-ordinateurs (de type IBM principalement). Le bassin potentiel d'utilisateurs est donc suffisant pour que le Centre Saint-Laurent élabore un atlas environnemental en employant l'approche multimédia sur support électronique afin de contribuer à diffuser les connaissances sur le Saint-Laurent.

Conclusion

La production de l'Atlas environnemental du Saint-Laurent, dans sa version sur papier, a été effectuée en utilisant la technologie infographique la plus récente. Elle n'a pu toutefois profiter du potentiel d'analyse et de visualisation des systèmes d'information géographique puisque le développement et l'intégration des diverses banques de données à référence spatiale sur le Saint-Laurent n'était pas encore suffisamment avancée au moment de la préparation de l'atlas. Les expérimentations en cours au Centre Saint-Laurent permettent cependant d'envisager la possibilité de réaliser un atlas électronique multimédia en utilisant les données stockées et traitées dans les systèmes d'information géographique développés au Centre, de même que les fichiers numériques créés lors de la production de l'atlas sur papier. Le marché et l'équipement requis pour la diffusion d'un tel produit existent. Toutefois, le défi actuel de la préparation d'un atlas environnemental multimédia n'est pas de nature technologique; il réside plutôt dans l'obtention de données et d'informations fiables et à jour et dans leur traitement pédagogique adapté au niveau du public visé. Le design d'un atlas électronique selon le concept d'hypermédia (avec un système de navigation basé sur les principes de l'hypertexte) suppose au départ une structuration de l'information environnementale en fonction d'une organisation conceptuelle des connaissances spécifique à chaque thème traité. La qualité d'un atlas en hypermédia dépend autant de «l'ingénierie de la connaissance» que de la technologie informatique. Enfin, un atlas électronique multimédia devrait exploiter pleinement le potentiel des nouvelles méthodes d'analyse et de visualisation des données spatiales. Citons, entre autres, la liaison interactive entre les données et leur image graphique ou cartographique, l'illustration des variations spatio-temporelles des données environnementales à l'aide de cartes animées en trois dimensions, etc. En utilisant toutes ces possibilités, un atlas électronique multimédia sur le Saint-Laurent pourra contribuer à une meilleure diffusion des connaissances sur les phénomènes et les problèmes environnementaux du fleuve et communiquer au public une image dynamique de son environnement.
Références


2. L’importance du rôle de la visualisation cartographique pour la représentation des relations spatiales est soulignée dans:


5. L’importance du rôle de la visualisation cartographique pour la présentation des relations spatiales est soulignée dans:


Note sur les auteurs et l’Atlas environnemental du Saint-Laurent

Jean RAVENEAU est professeur de cartographie au Département de géographie, Université Laval, Sainte-Foy, Québec, G1K 7P4, Canada.

Michel MÉLANCON est gestionnaire scientifique au Centre Saint-Laurent, Environnement Canada, 1141 Route de l’Eglise, Sainte-Foy, Québec, G1V 4H5, Canada.


On peut se procurer les planches imprimées de l’atlas en s’adressant à :
CENTRE SAINT-LAURENT
Environnement Canada
105, rue McGill, 4e étage
Montréal (Québec)
H2Y 2E7 Canada
Tél.: (514) 283-7000
Télécopieur (fax): (514) 283-9451
Session 13

Cartography Modelling of Geographic Information, Map Revision

Chairman:
S. Rimbert, French CNRS & VLP (Strasbourg, F)
"Navigation" and "narration" strategies in dynamic bivariate mapping
M. Monmonier (Syracuse, USA)

Abstract
Navigation and narration are complementary strategies for the dynamic cartographic analysis of geographic correlation. Although most interactive systems for exploratory data analysis rely largely on the navigation metaphor, software features based on the narration metaphor can informatively complement navigational systems. These features include (1) graphic phrases that generate dynamic maps linked to bivariate scatterplots and (2) graphic scripts embedded in narratively sequenced pull-down menus or command files. Narrative enhancements appropriate for the cartographic examination of bivariate correlation suggest useful strategies for other aspects of cartographic description and analysis.

Introduction
As metaphors for information retrieval and analysis, ‘navigation’ and ‘narration’ describe distinctly different approaches to dynamic cartography. The navigation approach is a highly interactive exploratory strategy in which the analyst uses signposts and pathways provided by the system to navigate from one visual display, or view, to the next. In principle, the user evaluates each new view for its marginal relevance to his or her overall information-seeking goals. In contrast, the narration approach requires the system to address these goals directly by selecting and displaying a coherent sequence of relevant yet largely non-redundant views. Thus, a navigation strategy is comparatively unrestrained and requires the viewer’s active participation, whereas a narration strategy can more efficiently provide an overview or a focused summary of the data.

This paper focuses on the correlation between two variables, or geographic series, measured for the same set of areal units. Its thesis is that effective cartographic analysis in an interactive computing environment requires both navigational and narrative phases. After all, comprehensive examination of a bivariate correlation based on spatial data requires a variety of graphic images. Some of these images are statistical and non-geographic, for example, distribution graphs to describe each variable’s variance, a bivariate scatterplot to describe the covariance, and one or more lines or curves to describe the generalized or smoothed bivariate relationship. Other views are cartographic, for example, maps of the individual variables (either singly or juxtaposed), bivariate cross-correlation maps obtained by partitioning the scatterplot, maps reflecting the fitting of one variable to the other, and residual maps showing variation not accounted for by curve-fitting. Use of the plural form ‘maps’ in this second list reflects the diverse views that might result from applying different symbolization and classification schemes to the same geographic series.

Navigation and narration strategies can differ widely in freedom and thoroughness. A navigation strategy commonly yields the analyst’s idiosyncratic sampling of only some of many relevant cartographic views, whereas a script can ensure a more efficiently and reliably informative selection of maps. Because the latter approach can serve as an intellectual ‘pump-primer’ for the former, these two strategies are inherently complementary.
The first part of this paper describes a prototype graphic narrative designed to explore the correlation of two geographic variables. The second part discusses focus-group testing of both this narrative and a simple navigational variation of one of the graphic narrative’s principal segments. The third part discusses a variety of short graphic narratives that might be especially useful in a largely navigational system for examining spatial correlations.

**A Graphic Script for Bivariate Geographic Correlation**

A carefully ordered series of dynamic maps and graphs served as a working prototype for exploring the role of graphic narratives in geographic analysis. This ordered set of dynamic graphics was one of two prototype graphic scripts developed as part of the Atlas Touring Project. Divided into three acts each consisting of multiple scenes, the “correlation script” moves from introductory scenes that examine the variances of each variable individually, to intermediate scenes that explore the covariation of the two variables, and then to culminating scenes that apply a linear regression model and examine the spatial and regional patterns of the residuals. Individual scenes consist of one or more graphic phrases, each generating a highly focused sequence of maps, graphs, or text. A second prototype, the “historical script,” uses a variety of measurements and views to examine change over a 90-year time span by year and by decade for a single variable.

Titled “Women in Politics,” the correlation script examines the plausible hypothesis that females should hold a relatively larger share of elective offices in local government in states with higher-than-average percentages of adult females in the civilian labor force. An extended stage-and-play metaphor assigned the dependent variable the character name “Female Officials” and the signature hue red. Although the script uses longer, more descriptive titles above its large, full-screen maps, shortened, two-word names are useful for titling smaller, partial-screen maps as well as for labeling the axes of bivariate scatterplots and other statistical graphs. Treating fill patterns as a form of cartographic costume, the script uses red area symbols for all choropleth maps of the dependent variable and red bars for all bar charts. In contrast, the character name “Females Working” and the signature hue blue provide similar identification for the independent variable. Magenta (as a subdued, bluish red) provides an appropriate signature hue for the residuals, introduced and mapped in the script’s final act.

To promote a coherent, informative graphic script, I followed principles of narrative discourse grounded in research on reader expectations and discussed by Joseph M. Williams in *Style: Toward Clarity and Grace.* In observing the effect of narrative structure on intelligibility, Williams and his colleagues noted that readers comprehend a straightforward story more readily and reliably than a narrative with obscure goals and weak transitions. The effective writer establishes a context for ready interpretation by stating “the point” (namely, the goal or purpose) early in the story, linking new information to readily relevant old information, and avoiding needless distractions. In authoring the correlation script, I attempted to minimize narrative complexity by adopting standardized symbols, locations, and labels, by announcing with text the point of each act or scene, and by following a logical sequence from relatively simple to more statistically complex concepts.

The first two scenes of the script’s first act introduce the dependent and independent variables separately, and the third scene compares summary maps developed in the first two scenes. To minimize visual complexity, the two introductory scenes are identical except for the wording of map titles and other descriptive text, the signature hues, and the spatial patterns and statistical distributions. Each scene begins with a text screen relating the variable’s full title and brief title to a comparatively detailed and complete description of the measurements, including any significant adjustments or exceptions. After a suitable pause for reading, the screen shifts to the juxtaposed map and vertical bar graph in Figure 1. The rank-ordered bars are scaled according to a non-zero base so that differences in height represent differences in value. Each bar is linked to its corresponding polygon on the map. At any instant the map highlights all states with a data value less than or equal to the current threshold value represented in the vertical indicator at the left as well as by the extent to which the variable’s signature hue has advanced horizontally across the bar graph. In a sweep-by-rank sequence the fill color advances at a uniform pace across the bar graph from left to right, and then retreats at the same pace back to the left side. The state with the highest value is the last to be highlighted, and the state with the lowest value remains highlighted throughout. During the sweep-by-rank sequence the signature hue rises and falls...
unevenly in the value indicator at the left; movement is slow when several states have similar data values and rapid when adjacent ranks are comparably different. During the sweep-by-value sequence that follows, the indicator level rises and falls uniformly as the signature hue advances and retreats less evenly (in most cases) across the bar graph.

To call attention to extremes in the data, the script highlights first the ten highest states (the highest fifth) and then the ten lowest states (the lowest fifth). The remainder of the introductory scene consists of a series of choropleth maps for which the bar graph serves as the key. With a quintile map of the states, for instance, five progressively darker area symbols fill contiguous groups of ten bars each, the lightest group on the left and the darkest on the right. To point out that different class breaks yield different map patterns, the script follows a quintile (five-category quantile) map with a tritile (three-category quantile) map. For each map, the script simultaneously blinks the polygons and bars for each category several times. It then restores the quintile map and canvasses the country by census division, blinking on and off in rapid succession all states in each of the nation's nine census divisions. Because white is the normal fill pattern for states in the lowest category, these bars and polygons alternate with solid black during a blinking sequence.

The third scene of the first act is an "alternagraphics" sequence in which quintile choropleth maps, in red for the dependent variable and blue for the independent variable, alternate rapidly on the same base, with 13 full red-blue-red cycles in 12 seconds. If the variables were highly correlated, the map would appear merely to change color. For the "Women in Politics" script, where the product-moment correlation is only 0.31, the map appears to throb.

In its second act, script moves gradually into an analysis of the variables' covariance. The first scene juxtaposes maps of both variables with a bivariate scatterplot. To link each variable's map to the corresponding axis on the scatterplot, the script moves a conditioning brush back and forth twice along each axis. Figure 2 illustrates the brushing of the vertical axis, linked to the upper map; in this screen snapshot the upper map highlights the five states with values in the narrow, moderately-above-average range represented by the instantaneous position of the brush during an upward movement along and perpendicular to the vertical axis. After brushing the dependent variable, the script brushes the independent variable and highlights polygons on the lower map. The next graphic phrase invokes quintile classifications on both maps and alternates their keys several times on the scatterplot, toggling every few seconds between (1) five horizontal bands progressively darker from bottom to top and containing ten dots each and (2) five vertical bands progressively darker from left to right and also containing ten dots each. The scene then repeats the sequence for juxtaposed choropleth maps classified first according to equal thirds and then according to equal halves.

To introduce a bivariate map focused more fully on the variables' covariance, the second scene segues from the first scene's two-category choropleth maps with class breaks at the medians to two-category maps with breaks at the two variables' national means. After alternating the keys, the script rapidly moves the two maps toward each other until they collide and dissolve into the single bivariate cross map represented in black-and-white in Figure 3. As before, the scatterplot serves as the map key. On a color screen the script uses four highly contrasting hues: red for states above the mean only for the dependent variable, blue for states above the mean only for the independent variable, yellow for states in the upper category for both variables, and white for states in the lower category for both variables. In a canvass-by-category sequence the script then displays each category individually, with the corresponding category label above the map in the title box and the polygons of states in other three categories temporarily filled with black. Next, to examine the stability of the covariance pattern, the script moves the bivariate breakpoint outward in a spiral and temporarily alters the map pattern as category boundaries change. Then, after spiraling inward toward the original breakpoints representing the national means for both variables, the script displays a full key below the map, as in Figure 3.

The third act uses the same layout, with a map on the left and a scatterplot on the right. After adding a least-squares regression line in magenta and announcing the correlation coefficient, the screen blackens the map and the scatterplot's interior (except for the magenta regression line) and begins a canvass-by-region sequence displaying the name of each of the nine census divisions while highlighting in yellow the division's polygons and dots. An equally brief second scene displays a four-category choropleth
map of the residuals, partitioned according to class breaks at zero and ± one standard error. The script then performs two successive canvass-by-category examinations of the residuals map, each highlighting in turn all positive residuals, all negative residuals, all positive residuals above one standard error, and all negative residuals below one standard error. Its final view is the full map of residuals with a complete legend below.

User Preferences for Navigational Control

A focus-group evaluation of the correlation script confirmed what both self-examination and the informal opinions of colleagues had already made evident: users want more control over both the pace and direction of a graphic script.19 Four group interviews conducted by an experienced focus-group facilitator viewed and discussed two demonstrations, one consisting of the correlation script and the other comprised of two short portions of the historical script and an interactive version of the graphic phrase for automatically sweeping the distribution of data values. The facilitator collaborated with the author in developing a focus-group protocol for probing such key issues as informativeness, coherence, perceived assets and liabilities, and user preferences for customized graphic narratives. The 26 participants were professionals involved in either geographic analysis or the evaluation and design of information systems. Most participants found the graphic scripts engaging, appreciated the underlying graphic logic, and correctly detected salient patterns in the data. But they voted unanimously for the navigational option of being able to slow down the script or to stop, back up, and repeat selected portions. Participants also recognized the need for a voice-over or an oral narrative, such as that provided by the facilitator—although screens with text might appropriately precede a dynamic sequence of graphics, viewers cannot easily watch a moving display and simultaneously read titles, labels, and other accompanying text.

In the second of the two demonstrations, an interactive version of the upward and downward sweeps of the vertical bar graph in Figure 1 generated several important insights. As described in Figure 4, this modified prototype replaced the bar graph with a univariate scatterplot in which small open circles representing data values are linked dynamically to polygons on the map. This design provides a more direct representation of clusters and gaps in the numerical distribution and displays in a small box above the scatterplot the instantaneous numerical threshold represented by the dynamic breakpoint for the two-category map. In its fully manual, navigational mode, the viewer can use a mouse or trackball to move the breakpoint left or right across the scatterplot. Focus-group participants were unanimous in preferring this manual mode to both an automatic mode, which presented two steadily paced horizontal sweeps of the scatterplot, and a semi-automatic "drive" mode, which allowed the user to control only the direction and speed of the horizontal sweep. Although viewers might appreciate a brief demonstration of graphic tools, they also prefer full control over each tool's operation.

Several strategies exist for accommodating viewer preferences in a revised correlation script. A relatively rudimentary solution is to insert "pause points" after all graphic phrases, at appropriate intermediate places within some graphic phrases, and at all text screens. At these points in the graphic narrative, the system would pause and await the viewer's command either to repeat the preceding view or sequence, or to proceed immediately to the next. A second, more fully navigational modification is to insert "interaction points," at which the viewer may manipulate the preceding graphic phrase. Indeed, many if not all of the graphic phrases in the correlation script could be placed under the viewer's direct control. Moreover, a short automatic-demonstration at the beginning of each segment could be mandated, suppressed altogether, or provided as an option.

Other, more flexible navigational strategies might provide access to the script's text blocks and graphics phrases through one or more menus. At the lowest level of menu control, all graphic phrases and text screens might be linked to a set of pull-down menus. In this comparatively "open" graphic narrative, a menu bar at the top of the screen would present the correlation script's three acts as a group of adjacent pull-down menus, with act one to the left and act three to the right. Each pull-down menu could then list its act's text blocks and graphic scripts in a narrative sequence from top to bottom so that a viewer could recreate the script in its intended order as well as jump back freely to a particular segment. The analyst could, of course, view the entire script in its canonical sequence by using the "resume" key to navigate stepwise through all commands in the script's menus.
In a different menu configuration, a command window could list in narrative order a series of short descriptors representing the script’s text blocks and graphics phrases. In its fully automated, “player piano” mode, the system would progress down this list, highlighting individual descriptors in the command window while presenting the corresponding text block or graphic sequence in the system’s view window. In contrast, the viewer might navigate the script in “improvisational mode” by scrolling up or down the command window and viewing individual descriptors or larger portions of the script at random.

Navigational systems can, of course, be used to capture completely “closed” graphic narratives on videotape, as a computer-generated animation, or as command files, which viewers can replay in the manner just described. Closed narratives provided in video format or as a real-time computer movie can be effectively enhanced with a highly informative, carefully edited voice-over. The computer-movie format allows the added flexibility of pause points, replay options, and control over pace and direction. In contrast, fully “open,” navigational systems for data exploration could have narratively organized menus or command windows as options. Most navigational systems for exploring and analyzing geographic data have neither. Nonetheless, developers occasionally provide narratively organized independent tutorials as “stacks” for use with a graphic hypertext system28 such as HyperCard.

Narrative Graphics for Correlation Analysis

Both approaches, narrative as well as navigational, can benefit from a conceptually enriched analysis of geographic correlation that (1) examines the effects of spatial resolution, (2) addresses the possible roles of spatial autocorrelation11 or coherent spatial trends, (3) explores the potential relevance of nonlinear models, such as logarithmic and polynomial regression and lowess smoothing,2 and (4) evaluates the effects on the analysis of extremely high or low data values, or outliers. In a fully narrative presentation the order of these enhancements—and the decision about which ones to invoke—must depend upon their marginal contributions to the viewer’s understanding. For example, if spatial resolution has a substantial influence on the result, with noteworthy differences in either the direction or the magnitude of the relationship for state, county, or minor civil division data, the system should reveal this highly important effect early in the script. Similarly, if the distributions of both the dependent and the independent variable share a common spatial trend, if the salient bivariate relationship is clearly curvilinear rather than linear, or if the suppression of outliers improves the model’s goodness-of-fit, the graphic narrative should highlight these phenomena as close to the beginning of the script as internal coherence and the user’s grasp of concepts permit. An ideal narrative system will first interrogate the data for these and other meaningful relationships and then generate a uniquely customized, “data-driven” visual presentation.

The ideal system will also consider the interests and quantitative ability of the viewer, as represented in a user profile4 containing pattern masks.21 What the viewer already knows, and what he or she can readily comprehend is important information that both a human narrator or an electronic narrative system must weigh seriously. Moreover, a competent navigational system should help the user recapitulate and resume an interrupted analysis by remembering and replaying the salient portions of an exploratory session carried out an hour, day, week, or month earlier. It is essential, therefore, that the navigational system maintain not only a log of the earlier session but also a record of the viewer’s assessment of each segment’s relative informativeness. A recapitulatory graphic narrative thus provides an “analytical bookmark” to help the user answer the question, “Well, now where was I?”

A relevant and informative graphic narrative must also avoid a mechanistic focus on trivially obvious patterns. After all, aberrations or exceptions to a blatantly prominent trend can be far more enlightening than patterns the viewer already knows. Although surprise departures from expectations usually represent defects in the data rather than serendipitous discoveries, awareness of deficiencies in data quality is a prerequisite for competent exploratory data analysis. An intelligent narrative system thus depends not only on the viewer’s user profile but also on an intelligent data base that can apply conventional wisdom and guide the user rapidly and efficiently over familiar, less intellectually promising paths toward interesting and possibly rewarding anomalies.
At a narrower, more focused level, narration can be an effective strategy for assessing the relevance of particular views of the data and the value of further, navigational exploration. Consider, for instance, the dynamically linked bivariate map and scatterplot in Figure 3. As the spiral sequence in the correlation script demonstrates, the bivariate breakpoint represented by the intersecting horizontal and vertical lines on the scatterplot can be moved freely in two dimensions. Before providing the user an opportunity to explore the covariance relationship interactively, the system might usefully execute a brief model-driven narrative based on the regression line and a bivariate density ellipse containing, say, 75 percent of the data points. As described schematically in Figure 5A by the sequence numbers 1-2-3-4-5, the system moves the bivariate breakpoint through an upward sweep along the regression line and then completes a counterclockwise circuit around the perimeter of the ellipse. To de-emphasize places conforming to the linear relationship, the graphic narrative 1 substitutes a gray area symbol in the interior of the ellipse and in all polygons linked to points within the ellipse and 2 then completes downward and upward sweeps along the central part of the regression line described by the sequence numbers 5-6-7 in Figure 5B. Next, a model-driven examination of the residuals map focuses attention on the more extreme departures from the linear relationship. In the first phase of the residuals sequence (Figure 5C), bright magenta represents positive residuals (above the line), a comparatively faint magenta represents negative residuals (below the line), and an expanding white linear band centered on the line focuses attention on progressively more extreme residuals, which retain their hues longer. The second phase (Figure 5D) reintroduces the gray ellipse in order to focus attention on residuals outside the regression line’s 75-percent density ellipse. All polygons linked to points within the ellipse remain gray, so that growth of the white band highlights places that are both more extreme as well as less representative of the linear bivariate relationship. The viewer may repeat this short narrative sequence or begin immediately to explore the correlation interactively.

If a curvilinear model better fits the data, an abbreviated sequence similar to that in Figure 5C might usefully precede a navigational exploration of departures from the overall bivariate relationship.

Another quantitative technique that can drive a graphic narrative is trend surface analysis. A short narrative based on linear trend surfaces might be particularly revealing if two distributions are not only at least moderately correlated but also share a similar linear trend, that is, if 1 the coefficients of determination for both linear surfaces are above, say, 0.8, and 2 the azimuths of these linear trends differ by no more than about 45 degrees. In this case, the scatterplot can be juxtaposed with separate maps for each variable. If the correlation is generally positive, a movable line perpendicular to the regression line partitions the scatterplot into two zones, one generally lower for both variables and the other generally higher for both variables. Filling with magenta the polygons linked to points currently within the lower zone and filling with white those outside the zone provides a treatment similar to the univariate sweeps represented in Figures 1 and 4. As the perpendicular division line moves along the regression line from the lower left to the upper right, the magenta fill will simultaneous flood across both maps in roughly the same direction. Superimposing on each map a single arrow representing the linear trend provides a useful summary of the graphic phrase. A similar graphic narrative might informatively present other common trends, such as a circular diffusion pattern centered on a single major city.

When two highly correlated variables share a readily identifiable geographic trend that might suggest other relevant factors, it can be useful to detect and epitomize the trend early in the analysis. Timely, concise graphic narratives are especially important in a data-rich exploratory environment, where incisive graphic narratives can promote overall effectiveness by leaving additional time for new discoveries.

Concluding Remarks

Narrative graphics can usefully enrich many kinds of navigational analyses, not just explorations of bivariate geographic correlation. Narration can introduce a new user to the system and help an experienced user review and resume yesterday’s analysis. Narrative concepts embedded in the software can also generate quick, revealing overviews; foster thoroughness and vigilance; support the development, use, and evaluation of canonical sequences of tools and views; and promote the critical
discussion of cartographic strategies for analysis and discovery. Alternative graphic narratives can demonstrate dramatically that geographic analysis is both science and art.

Acknowledgments

The author is grateful to the National Science Foundation (grant number SES-90-22845) for supporting the Atlas Touring Project; to Myke Gluck for helping design the focus-group protocol, serving as facilitator, and helping interpret the results; and to Mike Kirchoff at the SU Cartographic Laboratory for assistance with the illustrations.

References

FIGURE 1. Black-and-white screen snapshot from the first scene of the first act shows an upward sweep, by rank, of the dependent variable. In color, the bars and polygons shown here in black would be red, the signature hue of the dependent variable.
FIGURE 2. Black-and-white screen snapshot from the first scene of the second act shows maps for each variable juxtaposed with a bivariate scatterplot. A horizontal conditioning brush is exploring the scatterplot's vertical axis, which represents the dependent variable. In color, the interior of the brush is light red and the highlighted polygons are bright red.

FIGURE 3. Black-and-white screen snapshot from the second scene of the second act shows the bivariate cross map above a full map key. The scatterplot (right) serves as the primary key during most of the scene. See text for a description of colors.
FIGURE 4. Black-and-white screen snapshot showing the modified, interactive version of the graphic phrase used at the beginning to the correlation script's first scene. In color, the black polygons and shaded portion of the univariate scatterplot would be red.

FIGURE 5. Schematic description of a model-driven graphic narrative for examining (A, B) geographic covariance portrayed on a cross map and (C, D) residuals from linear regression.
Generation of high-quality 1:25,000 and 1:50,000 satellite image maps of high-relief terrain

M.F. Buchroithner, R. Fischer (Dresden, D and Graz, A)
Abstract

The objective of the presented study was to exploit the potential of remote sensing data sets in high-mountain cartography for special applications to hiking and climbing tourism. Several problems induced by the high-mountain terrain, especially geometric, radiometric, atmospheric and cartographic ones, had to be tackled. Digital fusion of multiresolution multispectral data (Landsat TM bands 3, 2, 1 and SPOT panchromatic) was a prerequisite for the generation of a high-quality large-scale image map. Both data sets were high-precision geocoded to the common reference system of the Austrian Topographic Map by means of a digital elevation model and resampled to the DEM-corresponding pixelsize of 25 m and 12.5 m respectively. After the actual geocoding, particular correlation techniques were applied to correlate TM and SPOT data to each other with accuracies significantly below half a pixel. Two methods of data fusion, the IHS transformation and the High-Pass Filter method were tested. After the combination of both data sets several methods of postprocessing, like contrast stretching and histogram modifications, were applied. Subsequently, the adjacency effect was taken into account. Through its correction fine details, and in particular linear elements such as small parcels of land or roads, were enhanced.

During the cartographic work trekking routes, rivers, names, heights and other elements important for orientation, were digitally integrated into the image. The results of these tests are presented as high-mountain image maps of the Schladming Tauern Range (Austria) at the scales of 1:50,000 and 1:25,000 respectively.

1. The TADAT Research Program
1.1. Background

The development of remote sensing techniques has reached a level where these methods can now be used on an operational basis for many cartographic applications. However, the large variety of both digital and analogue data handling techniques utilised in remote sensing cartography has led to a profusion of value-added map products that are often difficult to compare or integrate with each other, or to employ for practical purposes.

To achieve these objectives, an interdisciplinary working group for remote sensing, photogrammetry and cartography in Graz (Austria) proposed the establishment of test areas that offer a diversity of terrain types and land uses and for which comprehensive geoinformation is available. A geographical area meeting these requirements and subsequently selected by the working group as a suitable test site is the Tauern-Dachstein Region in the Austrian Alps, which covers an area of approximately 400 km².

1.2. Test Area Description

The TADAT test site consists of (from south to north): the Schladming Tauern Range (Mount Hochgolling, 2863 m) with the planned national park Niedere Tauern (crystalline rocks); a high mountain wilderness area with minimal human impact; the broad valley of the Enns River (the town of Schladming, 750 m) with locally intensive land use; settlements, road and rail routes, agricultural areas, tourism and light industry; the Ramsau Plateau (elevation around 1100 m) with scattered settlements; used mainly for tourist industry and as pasture land; the glaciated Dachstein Massif (Hoher Dachstein Peak, 2996 m) with glacial and karst features (carbonate rocks); mixture of high relief terrain and undulating plateaus, with locally intensive tourist industry (summer skiing centre).
1.3. Objectives

Within the last years the need for high-quality maps in high mountain regions became very important, because of the increasing trekking tourist industry worldwide. This led to the idea, to use the great potential of satellite image data, especially Landsat TM and SPOT panchromatic, to check high-quality satellite image maps for the application to hiking and climbing tourism. The first step is to optimise the geometric and radiometric quality of the satellite images in order to provide an excellent basis for further cartographic processing (cf. Figure 2).

Especially in the high-mountain regions of the Himalaya (Kostka, 1987) and the Andes there is a lack of quality maps at large scales.

![Map generation scheme](image)

Figure 1: Location map of the TADAT test site.

Figure 2: Map generation scheme.
2. High-precision Geocoding

A prerequisite to optimise the absolute geometric location accuracy is the removal of image distortions, caused by topographic relief, through the use of a digital elevation model (DEM), in particular in the case of high-relief terrain. This type of absolute parametric georeferencing is called terrain correction geocoding and provides output data with optimum precision for any type of terrain and sensor geometry.

For the present study, the need to use terrain correction to obtain geocoded multi-sensor image data with high geometric accordance for high-relief terrain is obvious.

- Landsat TM scene no. 191-27:
  Date of acquisition: 31-8-1985
  Centre of scene: \( \lambda = 14^\circ 06' \)
  \( \varphi = 47^\circ 27' \)

- SPOT panchromatic scene no. 64 254:
  Date of acquisition: 14-9-1987
  Centre of scene: \( \lambda = 13^\circ 35' \)
  \( \varphi = 47^\circ 30' \)

It is obvious that the displacement depends on terrain elevation \( H \), the type of the sensor specific mapping geometry. Figure 6 shows a graphic presentation of the interaction between terrain height and relief displacement for SPOT and Landsat 5 imaging geometry, the respective "iso-displacement"-curves being derived by simplified geometric relations. For the present image data the range of absolute displacement errors caused by the actual terrain elevation can be depicted from these graphs with respect to their inherent mapping parameters and the relative location of the test area as follows:

Landsat: \( 0 \text{ m} \leq \Delta d \leq 60 \text{ m} \)

SPOT: \( 0 \text{ m} \leq \Delta d \leq 60 \text{ m} \)

2.1. Approach Used

A prerequisite for high-precision parametric geocoding is the availability of a number of ground control points (GCPs). These are used to determine sensor-specific mapping parameters and, moreover, to refine these parameters in a least squares parameter adjustment procedure in order to optimise the mapping model. Then, high-precision geocoding using a DEM is performed in two basic processing steps:

- Each pixel of the DEM defines the location of an output image pixel (geocoded pixel) in a map projection. For the coordinate triple easting, northing and terrain height \( (E, N, H) \), first the corresponding location in the input image has to be determined by a so-called map-to-image coordinate transformation (Raggam, 1990).

- Subsequently, a grey value can be interpolated for the output pixel from neighbouring input image pixels following certain interpolation criteria.

A detailed description of this approach is given by Raggam (1990).

Figure 3: Axonometric view of part of the the TADAT digital elevation model from north.
2.2. Mapping Accuracy

A rather large number of GCPs has been measured in each of the two input images and in the topographic 1:25,000 map. This is a comparably difficult task in mountaineous areas where man-made features - the most common objects to serve as GCPs - are scarcely available. These points have been input to a parameter set-up and refinement procedure in order to determine appropriate mapping models for the relation of ground points and corresponding image pixels.

Further, this procedure rejects GCPs, which have been identified to be obviously erroneous, thereby retaining 38 GCPs for the SPOT XS, and 95 GCPs for the Landsat TM image, respectively.

For a parametric geocoding procedure generally no need for the availability of so many control points exists, but apart from geocoding, in the present case these should be used also to analyze the fidelity of the individual mapping models. This is determined by GCP residual vectors, defined as the difference between a pixel location measured in the image and the pixel location determined using the mapping model. Overall statistical parameters can be calculated for the residual vectors (Strobl et al., 1990). They are summarized in Table 1 and 2. Usually the RMS values are taken as a quality measure for the overall mapping accuracy.

2.5. Geocoding Procedure of Landsat and SPOT Data

The actual geocoding process was performed with the RSG Software (Remote Sensing Software Graz, Almer et al., 1991). This software package allows a graphic and statistical presentation of the quality control of the geocoded data which is shown in Tables 1 and 2.

Table 1: Quality control (global residual statistics) of geocoded Landsat TM data.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>16.75</td>
<td>20.20</td>
<td>26.24</td>
</tr>
<tr>
<td>MEAN</td>
<td>9.11</td>
<td>2.56</td>
<td>24.03</td>
</tr>
<tr>
<td>STD</td>
<td>14.06</td>
<td>20.03</td>
<td>10.54</td>
</tr>
<tr>
<td>MIN</td>
<td>-20.71</td>
<td>-44.60</td>
<td>0.75</td>
</tr>
<tr>
<td>MAX</td>
<td>34.59</td>
<td>35.10</td>
<td>44.75</td>
</tr>
<tr>
<td>[MIN]</td>
<td>0.75</td>
<td>0.00</td>
<td>0.75</td>
</tr>
<tr>
<td>[MAX]</td>
<td>34.59</td>
<td>44.60</td>
<td>44.75</td>
</tr>
</tbody>
</table>

Table 2: Quality control (global residual statistics) of geocoded SPOT panchrom data.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>8.16</td>
<td>7.09</td>
<td>10.56</td>
</tr>
<tr>
<td>MEAN</td>
<td>-2.46</td>
<td>-1.14</td>
<td>9.74</td>
</tr>
<tr>
<td>STD</td>
<td>7.99</td>
<td>6.99</td>
<td>5.63</td>
</tr>
<tr>
<td>MIN</td>
<td>-16.68</td>
<td>-13.10</td>
<td>1.90</td>
</tr>
<tr>
<td>MAX</td>
<td>9.58</td>
<td>13.00</td>
<td>19.15</td>
</tr>
<tr>
<td>[MIN]</td>
<td>0.00</td>
<td>0.70</td>
<td>1.00</td>
</tr>
<tr>
<td>[MAX]</td>
<td>16.68</td>
<td>13.10</td>
<td>19.15</td>
</tr>
</tbody>
</table>

Check points: 19

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>7.37</td>
<td>5.73</td>
<td>9.24</td>
</tr>
<tr>
<td>MEAN</td>
<td>4.08</td>
<td>-0.18</td>
<td>8.37</td>
</tr>
<tr>
<td>STD</td>
<td>6.14</td>
<td>5.73</td>
<td>4.14</td>
</tr>
<tr>
<td>MIN</td>
<td>-15.10</td>
<td>12.50</td>
<td>1.00</td>
</tr>
<tr>
<td>MAX</td>
<td>7.59</td>
<td>10.00</td>
<td>15.51</td>
</tr>
<tr>
<td>[MIN]</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>[MAX]</td>
<td>15.10</td>
<td>12.50</td>
<td>15.51</td>
</tr>
</tbody>
</table>

3. Correlation of Landsat TM and SPOT Panchrom Data

Geometric congruence has been accomplished by hierarchical feature vector matching, a procedure which is described in the following.

3.1. Feature Vector Matching

A comparatively new approach of stereo matching is based upon the idea of creating a feature vector of each pixel (-surrounding) and comparing these features in the images to be registered (Figure 10). The principles have been described by Paar and Pödlzeiter (1991) and are for the sake of better understanding, briefly referenced in the following.

The first step is to create a set of feature images for both the reference and the search image. These features are derived from local properties in the surrounding of each pixel. The size of the windows depends on the feature to be calculated.
3.2. Feature Selection

To approach the question which features should really be used Paar and Pößleitner (1991) suggest to use feature selection and extraction methods as widely used in statistical pattern recognition. The procedure is the following:

First, for a stereo pair a correct elevation model is derived. This elevation model is used to find the 3d coordinates for the pixels in the 2D images and the respective disparity values. Each pixel pair with its associated feature vector is now used to form two classes: The class of 'matching points', and 'non-matching' points in the neighbourhood of the corresponding pixels. For these two classes statistical methods, e. g. the Mahalanobis distance, can be applied to get the best feature and/or feature combinations.

3.3. Matching Criterion: Minimum Euclidian Vector Distance

Once the feature images have been created, at each pixel location \((i, j)\) a feature vector \(\vec{r}_{i,j}\) exists with the following features: local 5 x 5 average, gradient of Kirsch 5 x 5 edge operator, local 5 x 5 variance, magnitude of a Sobel edge operator, first bandpass filter, second bandpass filter.

A correspondence to a reference vector \(\vec{r}_{i,j}\) is searched in the search space \(\sigma\) with

\[
\sigma = \{\tilde{s}_{k,l}/k \in [i_0 - 2\Delta i, i_0 + 2\Delta i], l \in [j_0 - 2\Delta j, j_0 + 2\Delta j]\},
\]

\((i_0, j_0)\) ... A priori expected coordinates of corresponding pixel

\(\Delta i, \Delta j\) ... Extension of search space in row and column direction.

The best correspondence is found on position \((k_0, l_0)\) in the search image with

\[
\min_{\vec{r}_{k,l}}(||\vec{r}_{i,j} - \vec{s}||) = ||\vec{r}_{i,j} - \vec{s}_{k_0,l_0}||
\]

using the Euclidean distance as metric.

Subpixel accuracy can be achieved if the distances between the reference vector \(\vec{r}_{i,j}\) and the search vectors surrounding the minimum distance vector \(\vec{s}_{k_0,l_0}\) are used as weights for coordinate interpolation.

3.4. Multi-Resolution: Bottom-up Pyramid Generation

Several approaches in stereovision deal with multiresolution for the generation of the disparity map. In our case, a 3 x 3/4 Gauss pyramid is built for each of the two images. The number of levels depends on the maximum expected disparity. For our example (1024 x 1024 pixels, Figure 4, not preregistered), the maximum disparity is about 35 pixels in either direction. So for a search space of 7 x 3 pixels (3 left and right, 1 up and down) at the top level, the selection of 5 levels with top resolution 64 x 64 is sufficient. For each of these levels the above mentioned features are generated. Beginning with the top level, a disparity (parallax) map is created for this resolution.

![Figure 12: Top down disparity refinement scheme (from Paar and Pößleitner 1991).](image-url)
For each of the next higher resolutions, the disparity derived above is used to begin the search in the right image for the best correspondence on each pixel.

3.5.  Multi-Resolution: Top-down Disparity Refinement
Feature vector matching is carried out on all pixels that contain enough information to be reasonable candidates for finding a homologue pixel in the search space. For the currently used features, the only criterion is a threshold in local variance. This threshold is set for naturally textured scenes containing small amount of shadows so that about 90% of all image pixels are matching candidates. Disparities on undefined regions as well as on the image borders are interpolated after matching.

Median filtering is carried out on the disparity map (distinct maps for row and column disparities) to remove "salt and pepper" noise. To provide backmatching (left-to-right and right-to-left), matching is done in both directions.

The result of this process is again a set of undefined pixels on the disparity map which have to be interpolated.

Each pyramid level uses the disparity map derived at the upper one as a priori information.

4.  Data Fusion Landsat TM - SPOT Panchrom

4.1.  Preprocessing
First the Landsat TM data are digitally enlarged by a factor of three in both directions to generate a pixel size similar to the SPOT panchromatic data. The results of this enlargement are smoothed with a 3 by 3 low-pass filter to eliminate the blockiness of the enlarged images (Chavez, 1986).

4.2.  IHS-Transformation
The IHS-Transformation is one of the most often used methods to merge multisensor image data, like Landsat TM and SPOT panchromatic. Successful applications are described by Chavez et al. (1991) and Albertz (1991).

The method uses three bands of the lower spatial resolution data set (Landsat TM) and transforms these data into the IHS space. Then the higher spatial resolution image (SPOT panchrom) replaces the intensity component image before the images are retransformed into the original (RGB) space.

Before replacing the intensity component image there was made a radiometrical adjustment of the SPOT panchrom data to the intensity component image. It is very important that the two images are approximately equal to each other spectrally.

5.  Contrast Enhancement
During a linear transformation the contrast is enhanced through an increase of the grey value range of the respective image. If the histogramme of the original image has an asymmetric shape, the linear transform only partially yields better results. In the present case different transformation curves have been interactively applied to the output bands of the IHS transformation (3,2,1) on the screen. The goal of this procedure was to achieve a colour composition as close as possible to the natural colours and to optimally enhance the image contrast.

For this purpose band specific transformations (partially, linear) whose curve shapes were not identical have been applied to each band (R, G, B).

Besides the above mentioned interactive contrast enhancements a method implemented at the GOP-300 image processing system has been applied (GOP-300 Image Processing Manual, Volume 1). This is an enhancement operation which generates the output image through a parameter-controlled (high-pass enhancement factor NFHP) linear combination of a low-pass filtered image and an image filtered by an isotropic high-pass filter.

6.  Atmospheric Correction - Correction of the Adjacency Effect
According to Richter (1990) the adjacency effect is only taken into account during the second phase of a comprehensive atmospheric correction. As such a procedure would have exceeded the frame of the present study, in modification to the above principle, an approximative correction of the adjacency effect has been performed independent of a preceding atmospheric correction.

For this correction procedure the resulting images of the digital fusion of the Landsat TM and SPOT panchrom data were used as input images. First, the average remission \( p \) in a \( N \times N \) pixel window for a particular pixel of the output image \( p_i \) is calculated.
Then the difference between \( p_1 \) and \( p \) is computed, multiplied by an atmospheric weighting factor \( q \), and added to \( p \), according to the formula: 
\[
p_2 = p_1 + q(p_1 - p) \]

The size \( N \) of the window depends on the pixel size, the atmospheric parameters, the spectral bands, and the partial frequencies of the scene (Kaufmann, 1985).

Commonly this approach is used for unrectified or already geocoded TM raw data. This work demonstrates that the above procedure can also be applied to already processed, in our case digitally fused Landsat TM and SPOT panchrom data. Instead of the so-called "reflectance image" (Richter, 1990, Itten et al., 1992) the output images of the IHS transformation are used as input images.

![Diagram](image)

The correction results in an obviously clearer image in which the colours look more intense in the structure of the forest comes out better. Fine details and linear elements like texture features or roads were strengthened without generating any unintended side effects.

**7. Cartographic Work**

**7.1. Lakes and Rivers**

Due to the relatively low reflection of the water bodies (lakes) they appear very dark in the RGB composite. This is why small water bodies in a comparatively dark environment are not or hardly visible. This also applies to narrow creeks within forests. They are only visible in the lower valleys due to their accompanying bush or tree seems.

For the digital integration the drainage network and the contour lines of the lakes were digitised from the Austrian Topographic Map 1 : 25 000. The drainage system, however, has not been traced completely but a selection of the most important rivers has been made. Subsequently the digitised lake contour lines and rivers have been transferred from vector to raster format. The resulting raster files were used as masks and integrated into the satellite image in the following way:

First, the lakes were treated by a \( 3 \times 3 \) average filter in order to make the sharp border line between water and land look more natural.

\[
R = ((1 - \text{Lake Mask}) \cdot TM3 + \text{Lake Mask} \cdot x/255) \\
G = ((1 - \text{Lake Mask}) \cdot TM2 + \text{Lake Mask} \cdot y/255) \\
B = ((1 - \text{Lake Mask}) \cdot TM1 + \text{Lake Mask} \cdot z/255)
\]

\( x, y, z \) constituting the various colours in the RGB mode. The values \( x = 20, y = 45, \) and \( z = 105 \) have been selected.
In the same way the drainage lines were filtered by a $3 \times 3$ average filter in order to ensure a natural integration into the image. Moreover, the lines are smoothened by this filtering.

The integration into the image was performed in the same way as for the lakes. The following colours were selected: $x = 0$, $y = 50$, and $z = 200$.

Through this combination with the initial enhanced image an optically balanced integration of lines and areas without abrupt demarcation versus the respective surrounding of the mask has been accomplished. Furthermore, the lines and areas fit much better into the (dark or bright) surrounding which results in a more natural appearance.

Figure 6: Hiking trail mask

7.2. Hiking Trails

The hiking trails were also digitised from the Austrian Topographic Map $1:25\,000$ and subsequently transformed from vector into raster format.

In this case the integration into the image was made without previous application of an average filter as the filter would have resulted in a broadening of the lines, so that the trails would have appeared to wide. However, the principle of the integration is the same as for the rivers and lakes, the values being $x = 180$, $y = 0$, and $z = 0$.

7.3. Value-Added Image

The result of the integration of lakes, rivers and hiking trails in the above mentioned way has been shown in a first report by Fischer and Buchroithner (1993). It is actually "only" a "value-added", high-precision geocoded satellite image and not yet a real image map. The complex digital lettering process is not included in this paper and will be described elsewhere.

8. References


Strategies and contents of topographic-cartographic real-world-modelling

G. Vickus (Bonn, D)

Summary: In conjunction with the conversion from analogue to digital production methods in topographic cartography, we are faced with the task of modelling real-world objects and facts. This article will discuss various real-world modelling strategies and describe the contents of this process.

1) Introduction

Together with the conversion from analogue to digital production methods in cartography, we are faced with the task of in future digitally storing and retrieving topographic information in and from information systems and databases which to date has been documented in analogue topographic maps. This data will then serve as spatially related core information for Geo-Information Systems (GIS) used in various disciplines, but especially as a means of deriving topographic maps.

In future topographic maps can be produced as follows in digital processes: The required information is gathered and stored within a Geo-Information System by means of topographic ground surveying and photogrammetric aerial evaluation or by digitising existing maps and adding existing geometric and object data files. By using various processing methods, particularly cartographic generalisation and following a symbolisation process, print copies for analogue maps can be prepared (Fig. 1).

In order to construct such topographic-cartographic Geo-Information Systems, real-world facts and objects must be modelled on the basis of cartographic model and GIS theory.

This modelling process is described as topographic-cartographic-real-world-modelling which comprises the definition of

- real-world facts and objects,
- attributes and relations,
- hierarchisation and structuring,
- data capture, symbolisation and representation criteria,
- data catalogues.
Fig. 1: Map production using GIS

2. Strategies

If one considers the requirements affecting topographic-cartographic real-world modelling, two strategies can be pursued in principle; combined topographic-cartographic modelling (TCM) (i.e. DLG-E as practised by the U.S.G.S.) or separate topographic modelling (TM) and cartographic modelling (CM) (i.e. ATKIS-DLM and ATKIS-DKM as practised by Germany's Federal States).

Fig. 2: Strategies of topographic-cartographic real-world modelling
Apart from qualitative and quantitative attributes the TKM process also provides the features of a landscape with graphic symbols. The content of TKM is orientated towards traditional map content and the feature definition is map feature related.

If TM and KM are separated, then TM is executed without any graphics. The content is then primarily orientated towards the significant real-world features and their appropriate thematic characteristics. The cartographic model then represents a map feature related and generalised extract of the topographic landscape model.

The advantage of TKM lies in the fact that only one closed model needs to be considered, making data administration and map revision easier.

By contrast, separated modelling enables the compilation of more variable and fuller landscape models from which cartographic models of varying scales and content density can be derived. The fuller landscape models are also advantageous for thematic applications, since the addition of specific attributes is made easier.

Regardless of strategy, general and concrete real-world modelling (RWM) needs to be carried out during modelling.

In general RWM (global view) contexts which are independent of the definition of individual special feature classes are described. The geometric model, for example, is classed as general RWM.

Concrete real-world modelling describes the definitions and representation rules of individual feature classes or map feature classes and their structural inclusion into the general model.

These observations lead to two complex diagrams (Fig. 3,6) on whose basis it is possible to carry out real-world modelling in GIS.

3) General real-world modelling

General real-world modelling (Fig. 3) comprises

- global feature definition,
- the core geometric model,
- the topological model,
- the data catalogue,
- and finally the symbolising and generalising model.

The global feature definition stipulates whether the model will be more topographically or more cartographically orientated.

A cartographically orientated model will include map feature classes which are decisive for the respective map. In order to be able to derive a map from such a model, it is necessary to symbolise and generalise the map features.
A topographically orientated model corresponds in content to a greater degree with the manifold requirements of GIS-users. Topographic content is generally structured differently and in greater detail. Here, map production is seen as a special application of the GIS. In order to be able to produce a topographic map, it is initially necessary to derive map features from the topographic features. After this, the created map feature model itself needs to be symbolised and generalised.

The geometric models describes model accuracy, various permitted geometry types, hierarchical structure of the data with subsequent vertical references, integration of vertical information (dimensions) and general qualitative and quantitative data capture and selection rules.

Fig. 3: General topographic-cartographic real-world modelling

<table>
<thead>
<tr>
<th>Geometric Model</th>
<th>General feature definition</th>
<th>Topographic and cartographic modelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model accuracy</td>
<td>Positional accuracy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Approximation of geometry</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Completeness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Raster resolution</td>
<td></td>
</tr>
<tr>
<td>Delineation</td>
<td>Point</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Line</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Area</td>
<td></td>
</tr>
<tr>
<td>General structure of features</td>
<td>Structure of features (vertical)</td>
<td>References</td>
</tr>
<tr>
<td>Dimension</td>
<td>2 D Without 3rd dimension</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 D + 1 D Situation model</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.5 D z-Coordinate as attribute</td>
<td>Elevation model</td>
</tr>
<tr>
<td></td>
<td>3 D 3-dimensional Coordinates</td>
<td></td>
</tr>
<tr>
<td>General conditions for data capture</td>
<td>Quantitative Conditions</td>
<td>Qualitative Conditions</td>
</tr>
<tr>
<td>Topologic model</td>
<td>No representation of topology</td>
<td>Representation of special topologic relations</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Data organization</td>
<td>Vector</td>
<td>Struktur, hierarchy</td>
</tr>
<tr>
<td></td>
<td>Raster</td>
<td></td>
</tr>
<tr>
<td>Model of symbolization and generalization</td>
<td>Symbolization</td>
<td>Point-graphic</td>
</tr>
<tr>
<td></td>
<td>Generalization</td>
<td>Simplification</td>
</tr>
</tbody>
</table>

The topological model defines to what extent topological information, i.e. node, link and polygon structures are to be made available.

Considering the data catalogue, special significance is given to the representation of data hierarchy with corresponding references and the allocated attributes as well as the definition of individual data elements (cf. Fig. 4,5: Extracts from the ATKIS-Data Catalogue).
Fig. 4: Hierarchy of ATKIS-Data (Source: AdV /1/)

Fig. 5: Definition of a DKM-Object (Source: AdV /1/)

Note: * := Iteration
Finally the symbolising and generalisation model illuminates the classical cartographic aspect of real-world modelling. The type and extent of cartographic generalisation correlates to the type and complexity of the symbolising, so that both steps result in one model.

4) Concrete real-world modelling

In concrete real-world modelling a basic differentiation must absolutely be made between the single-step concept of topographic-cartographic modelling (B) and the two-step concept of topographic and cartographic modelling (A).

Each models is identical in structure, but differs in the content of individual feature class definitions.

Fig. 6: Concrete topographic-cartographic real-world modelling

In concrete real-world modelling the real world is initially roughly classified in groups of similar feature classes and is then subjected to increasingly refined classification down to classification of individual feature classes. For each feature class further attributes can be defined to describe particular feature characteristics.

At the same time, this classification is transformed to the appropriate data catalogue, meaning that the feature expression is placed in relation to the individual classification levels with all its upper and lower units.
Both topographic and cartographic modelling of a concrete feature/map feature class comprise:

- the description of the feature/map feature class,
- the verbal definition of the feature/map feature class,
- the geometry type,
- references,
- attributes,
- names,
- feature/map feature creation rules.

Topographic modelling additionally comprises:

- data capture criteria \((A)\),

Cartographic modelling additionally comprises:

- selection criteria \((A)\),
- data capture criteria \((B)\),
- symbol definition,
- symbol usage rules,
- text definition,
- presentation priorities,
- and generalisation information.

Fig. 7: Concrete topographic-cartographic real-world modelling using an ATKIS example

The individual feature/map feature classes must be precisely defined with the assistance of such a diagram (Fig. 6). Figure 7 uses an ATKIS simplified example to demonstrate how concrete real-world modelling can be carried out.
5) Outlook

Apart from the attempt presented here to provide a stronger theoretical background for real-world modelling in Geo-Information Systems by dividing such modelling into general and concrete real-world modelling, the past few years have also seen the development of various topographic-cartographic models, for example, the DLG-E-Model created by the U.S. Geological Survey and the ATKIS-Model created in Germany.

For all models, in order to achieve an acceptable graphical output the problem of semi-automatic cartographic generalisation must be solved. Research into this problem must be given increased significance against the background of GIS development.

References:

(1) AdV (Hrsg.): Amtlich Topographisch-Kartographisches Informationssystem (ATKIS). Das Vorhaben der Landesvermessungsverwaltung zum Aufbau Digi-
taler Landschaftsmodelle, Bonn 1989).


(6) Vickus, G.: Objektgeneralisierung vom DLM zum DRM und ihre Anforderungen an die ATKIS-Modellie-

(7) Vickus, G.: Digitale topographische und kartographische Modelle sowie Entwicklung ihrer Überfüh-
rungsstrukturen am Beispiel von ATKIS. Disserta-
tion. Institut für Kartographie und Topographie der Universität Bonn, publication in preparation.
MKS-aspect™ – a new way of rendering cartographic Z surfaces
H. Moellering (Columbus, USA)

INTRODUCTION

The rendering of continuous surfaces is a crucial part of graphic data display systems in cartography, geography and the spatial sciences. This paper presents the MKS-ASPECT™ Slope-Aspect Color Surface Rendering Process which is a new and more powerful method of rendering continuous cartographic Z surfaces. This method was originally invented by Prof. Harold Moellering and subsequently enhanced by Dr. Jon Kimerling. The early discussion will summarize the important points of the MKS-ASPECT™ and review its significant characteristics.

The Second stage of discussion will provide a number of examples from topography and hydrography which illustrate the use of the MKS-ASPECT™ process on real spatial surfaces. The discussion will be directed towards showing how this process is an improvement over conventional approaches. This innovation is a significant step forward in surface rendering and hence the Ohio State University Research Foundation has been issued U.S. Patent No. 5,067,098 and has another patent application pending.

NOTE: Because the Figures in this paper are in color, it is not practical to include them in the body of this paper.

THE MKS-ASPECT™ SLOPE-ASPECT COLOR SURFACE RENDERING PROCESS

Slope-aspect is the horizontal direction of the gradient (maximum slope) at a particular location on a surface that can be calculated analytically on any digital surface that is parametric and is mathematically continuously differentiable. It turns out that most Digital Elevation Models (DEMs) of topography and hydrography meet these basic assumptions and hence are candidates for such a surface rendering. A traditional approach is to calculate the gray shades of such a surface using an illumination angle from the Northwest and shade the surface with gray shades that follows the cosine illumination law (Horn, 1982). This approach works reasonably well on simple surfaces, but begins to produce ambiguous results on complex surfaces and produces artifacts because it uses one single angle of illumination. A second traditional approach is to calculate the slope-aspect values and try to render the surface in various hues to sidestep the gray level angle of illumination artifact. Earlier attempts to do this have produced ambiguous results with local perceptual inversions in the resulting surface that made it extremely difficult to interpret.
Hence, the MKS-ASPECT™ approach is to improve on these earlier traditional approaches and eliminate their shortcomings by developing a process that can render a spatial surface unambiguously in color, preserve the perceptual relief effect, and eliminate the gray level surface illumination artifact. The process is built on four fundamental hues, yellow, blue, red, green and their intermediate mixture hues that are used to illuminate the surface with the effect that the surface is being illuminated from four directions with the four hues and their mixtures simultaneously. Because the hues with the highest perceptual luminosity, yellows, face to the northwest, and those with the lowest, blues, face to the southeast, the rendering of the surface creates the perceptual relief effect associated with gray shading, but now with hues of color. Left and right illuminations are shaded with mixtures of red/yellow and red/blue, while right illuminations are shaded with mixtures of green/yellow and green/blue. The invention of the MKS-ASPECT™ is the first time that the perceptual relief effect has been correctly achieved using a number of hues and their intermediate mixtures to render a spatial surface for visualization. In this way also, the MKS-ASPECT™ process illuminates the surface with many hues of light simultaneously and hence eliminates the artifacts produced by ordinary gray shading that are due to a single angle of illumination. It turns out that the MKS-ASPECT™ Slope-Aspect Color Surface Rendering Process is consistent with both the standard cosine illumination law and the opponent color theory (Eastman, 1986). A much more extensive discussion of the technical details of this process is presented by Moellering and Kimerling (1990).

Properties of the MKS-ASPECT™ Process

The first fundamental property of the MKS-ASPECT™ process is that it preserves the perceptual relief effect. Figure 1 shows a test area from the La Honda area of California rendered with 16 hues that have been assigned strictly on the basis of the orientation of the slope-aspect value for each representative square cell in the surface. This produces a surface visualization in two dimensions where the highest perceptual brightness (yellow) faces to the Northwest and the lowest perceptual brightness (blue) faces to the Southeast. The intermediate orientations of each surface square cell is rendered with a mixture of either red or green mixed with yellow or blue depending whether the general orientation is to the Southwest (red mixtures) or to the Northeast (green mixtures). This produces a full color rendering of the surface where the major ridges in the image, which are oriented North-South can be clearly distinguished from the valleys between them. On the Northwest sides of these major ridges one can clearly see yellows and mixtures of yellow/red and yellow/green on the bright side of these ridges. Conversely, on the Southeast sides of these ridges one can see blues and mixtures of blue/red and blue/green on the darker sides of these major ridges. In the extreme South of the image one can see a major ridge system running East-West with an accompanying valley just to the North of it. One can immediately clearly distinguish between the major ridges and valleys in the scene. When one examines the image in more detail, one can see a host of small spurs and side ridges that are coming down from the main ridges in the system. These contrasts are very apparent because they are using ranges of hues which are unambiguous. The entire image can be observed and perceived correctly in terms of the perceptual surface.

Because of the way in which the process is designed, surface features are now clearly visible regardless of the axial orientation of that feature, a major advance over conventional approaches. Therefore, a feature with any angle of orientation will be correctly rendered in
this display. Hence, the yellow/blue hue combination renders the major ridges/valleys in the image running in a general NE/SW orientation, while the red/green hue combination renders the ridges/valleys with a SE/NW orientation. All smaller side ridges/valleys will be rendered correctly because of the hue combination that results from their relative aspect angles.

Figure 2 shows the same image as shown in Figure 1 that now has been rotated 180 degrees. This now produces a perceptual inversion of the surface. Because the 180 degree rotation of the image now seems to put the brightly illuminates sides of the ridges and valleys on the "other side" the entire perceptual relief surface inverts and the high points become lows while the low points become highs. In this manner the ridges have become perceptual valleys and the valleys have become perceptual ridges. This simple exercise demonstrates that the MKS-ASPECT™ color surface rendering process preserves the perceptual relief effect in color and is robust enough to invert the perceptual surface when rotated 180 degrees.

The second fundamental property of the MKS-ASPECT™ is that it eliminates the gray level illumination artifact that always occurs on surfaces that have been rendered in shades of gray. This gray level illumination artifact has been known for many years and has been a significant impediment to cartographic surface shading. This problem results from the fact that when implementing gray shading of a surface, the illumination of a particular part of the surface comes from one angle of illumination. Hence, illuminating a spatial surface using different angles of illumination produces very different results.

This was very strikingly demonstrated by Batten and Francica (1986) by two renderings of an area in the Southeastern part of the Black Hills in the U.S. State of North Dakota. This part of the Black Hills is an area with complex geology and is highly dissected by sharp ridges, valleys and other surface features. Figure 3 shows this area

Figure 3. S.E. Black Hills with Gray Level Illumination from 300 Degrees. (From Batten and Francica, 1986)
where Batten & Francica shaded the area with an illumination angle of 300 degrees from North. Since this angle is very close to the standard Northwest angle of 315 degrees, the result looks very much like a standard gray rendering of the surface. However, if one knows this area, one is aware that there are many features that are not clearly visible because they are near to being parallel with the 300 degree gray illumination angle. Batten & Francica were aware of that situation and produced a second shading of the area with a gray shading illumination angle of 180 degrees, directly from the South. Figure 4 shows this gray shaded surface with the 180 degree illumination angle. This surface gray rendering is very different from Figure 3 in that very different features are now highlighted due to the greatly different angle of illumination. Here the perceptual relief effect has been greatly disturbed because of the near reversal of the gray illumination angle. Some features which first appeared as high areas now appear as low areas, and one can detect many areas of local relief inversion. This change of gray illumination angle has greatly changed which features are prominent and recede into the background. Now valleys that were barely visible before and now clearly visible, but unfortunately features that were clearly visible with the gray illumination angle at 300 degrees are now much diminished and some are hardly visible at all. The result is confusion on clearly understanding the surface under study and how it is really organized. Using multiple gray renderings helps, but individual renderings contain local ambiguities.

It turns out that the MKS-ASPECT™ Slope-Aspect Color Surface Rendering Process eliminates these problems and provides a correct rendering of the surface while at the same time correctly preserving the perceptual relief effect in color. Figure 5 shows the same area from the Southeastern Black Hills rendered with the MKS-ASPECT™ process.
where the terrain surface has now been rendered with 16 hues and hue mixtures. This is a very complex terrain surface that shows the power of the MKS-ASPECT™ process for rendering spatial surfaces. Now the main ridges and valleys are clearly identifiable regardless of the directional orientation of the feature. One can also see and identify a host of smaller features, many with subtle detail that would not be recognizable on an ordinary gray shaded rendering of the surface. For example, there is a long arcing ridge and valley that goes from the Northeast corner of the surface to the Southern edge of the surface. This double set of ridges slopes generally to the South and hence has a host of smaller side ridges and valleys sloping in that direction. All of these features can be seen clearly regardless of orientation. Near the middle of the surface there is a very obvious North/South straight line valley that is the product of geological faulting. It shows up clearly in this color rendering and so do the side ridges and valleys that are not so easily seen in the gray renderings. In the Southeastern corner of the surface is the Angostura Reservoir. To its East there is a relatively flat plain that on the gray renderings shows few if any of the subfeatures of that area. In Figure 5 that area can be seen to be composed of many subtle undulations with a number of smaller valley systems running SE to NW with a fewer number running NE to SW and connecting to the other ones. The primary reason all of these complex features along with their subtle subfeatures can be seen clearly is because the MKS-ASPECT™ color surface rendering process preserves the perceptual relief effect in color, and because it shows the feature clearly regardless of its directional orientation, thereby eliminating the gray level illumination artifact.

Additional characteristics of the MKS-ASPECT™ process can be summarized rather straightforwardly. The minimum number colors that can be used is as low as four, for basic kinds of 2-D displays, and range up to more than a hundred for more sophisticated 3-D spatial visualizations. Hence, the decision of how many colors to use is influenced by the capabilities of the generating system, the dimensionality of the display, and the complexity of the surface being analyzed. The mathematical requirements of the underlying spatial surface are that the surface be parametrically scaled Z data which represents a continuous spatial surface. The particular data representation of the spatial surface can be any piecewise cellularization of that continuous spatial surface. This includes square cells, TIN models, parametric patches, and other similar cellularizations.

**EXAMPLE SURFACE RENDERINGS**

In order to more completely show the effectiveness of the MKS-ASPECT™, two examples will be discussed. They are a topographic example which further discusses the Southeastern Black Hills, South Dakota, and a hydrographic example in the Gulf of Mexico named the Mitchell Dome.

**Southeastern Black Hills, South Dakota**

This example from the Black Hills has been discussed above and has shown the MKS-ASPECT™ process to be effective in preserving the perceptual relief effect and in eliminating the illumination artifacts that result from conventional gray shading. Figure 6 shows a part of that same area in a three dimensional visualization that uses more than a hundred colors. Here the third dimension is now made explicit and is more obvious to the viewer. In the foreground one can see the
Angostura Reservoir and behind it some of the complex topographic surface that is typical of this area. Here more colors have been used to enhance the detail of the rendering of the smaller features such as the third and fourth order ridges and valleys in the surface. This also improves the rendering of the subtle detail on the more level area to the Northeast of the reservoir. Such a 3-D visualization of the surface shows an additional effective way in which the MKS-ASPECT™ rendering process can be used.

Mitchell Dome, Gulf of Mexico

The Mitchell Dome is a hydrographic feature located in the Gulf of Mexico offshore from the State of Louisiana in the United States. This area is a set of subtle salt domes which are geologically interesting. Figure 7 is a 2-D rendering of the area using more than a hundred MKS-ASPECT™ colors. The most striking features are the salt domes themselves clustered in the middle part of the surface. In reality they are rather subtle features that do not show themselves very well with gray shading. Here one can clearly see the outlines of each salt dome, its side slopes and the subtle variations on the top of the larger domes. One can clearly see the relationship of the salt domes one to another, and the fact that several have ridges of sediment trailing off to the North and somewhat to the West. When one examines the surface for other features, one can clearly identify large and deep channel crossing the Northeast corner of the surface. On the Southern part of the surface one can also identify a set of very subtle parallel bottom ripples running Southwest to Northeast. Although the origin of these bottom ripples is not known, it was not realized that they were even there until the MKS-ASPECT™ process was applied to this data.

Figure 8 is a 3-D visualization of the same Mitchell Dome area. This 3-D visualization shows the same things as Figure 7 and confirms the effectiveness of the MKS-ASPECT™ rendering process. Everything that is visible on the 3-D visualization is also visible and identifiable on the 2-D rendering. This demonstrates that the MKS-ASPECT™ process works well in both two and three dimensional situations.

SUMMARY AND CONCLUSIONS

This paper has discussed the MKS-ASPECT™ Slope-Aspect Color Surface Rendering Process invented by Moellering and Kimerling. This is the first process to render slope-aspect data in full color that correctly preserves the perceptual relief effect. Because it has the effect of illuminating the surface with many hues simultaneously from different angles, the process eliminates the gray level single angle of illumination artifact that produces ambiguities in a gray shaded surface. The MKS-ASPECT™ process can be implemented with as few as four colors and more than a hundred, depending on the capabilities of the particular system and the goals of the project. The process can be directly used in two and three dimensional representations. It can be used on any kind of spatial surface that is parametric and continuously differentiable, and works correctly with a wide variety of spatial data structures that can be used to cellularize a spatial surface. The process can be implemented in software, firmware and hardware. Because it has all of these desirable properties the Ohio State University Research Foundation has been issued U.S. Patent No. 5,067,098 and has another patent application pending.

With all of these desirable properties it is clear that the MKS-ASPECT™ Slope-Aspect Color Surface Rendering Process has an interesting...
future for the analytical visualization of surfaces in the spatial sciences.

ACKNOWLEDGEMENTS

The Spatial Data Display Project has been supported in part by NASA grant NAGW-973 and an Ohio State University matching grant. Tektronix Inc. is a commercial partner. Prof. Jon Kimerling of Oregon State University made a significant contribution. Mr. Peter Dotzauer implemented the system software. The U.S. Geological Survey and the National Ocean Service provided the data used in the examples. Mr. Alan Edmonds assisted in the creation of the data sets.

REFERENCES


Représentation graphique et analyse de surfaces avec le procédé IRISOS
C. Schneider (Strasbourg, F)

Résumé :

On présente un procédé de représentation graphique de surfaces, où la variation de couleur est redéfinie avec une signification ondulatoire. L'utilisation et l'interprétation spectrale de la couleur aboutit des résultats intéressants sur différents plans : du point de vue méthodologique on réalise la synthèse graphique de l'analyse spectrale ; sur le plan du langage graphique, on obtient une extension des dimensions significatives de la couleur et un accroissement des informations sur le pixel. Le procédé constitue un outil d'analyse et de représentation globale des phénomènes complexes.

IRISOS (pour Imagerie par Reconstruction de l'Information Spectrale des Ondes Spatiales) est un procédé de représentation graphique qui résulte du rapprochement de notions issues de différents domaines, et dont l'intérêt commun est la description et l'analyse de l'espace. Des techniques de cartographie, de traitement d'image, d'analyse des surfaces, et d'infographie y sont combinées ou connectées dans le but de constituer un procédé cohérent, destiné à accroître le potentiel de description et de traitement du langage graphique.

Notre exposé concerne la phase de mise au point du procédé et porte davantage sur la logique et l'enchaînement de la démarche, que sur les problèmes de perception visuelle, ou sur ses applications thématiques. Cela nous permet de mettre en pratique la formule de Gaston Bachelard pour qui "... les instruments ne sont que des théories matérialisées. ", et qui nous rappelle la nécessité d'une confrontation permanente entre les aspects concrets et les aspects théoriques des outils que nous manipulons.

Position du problème :

On dispose de multiples procédés de représentation graphique qui conviennent bien à la description de la forme extérieure d'une surface ou à l'apparence des phénomènes. Celle-ci peut être représentée de façon efficace au moyen de techniques cartographiques éprouvées comme, par exemple, les cartes en isolines, en points proportionnels aux valeurs de la surface etc. Mais les surfaces géographiques, observées ou théoriques, résultent souvent d'un empilement de plans d'informations ou de champs de forces ; lorsque l'on souhaite représenter ces espaces dans leur plénitude, avec l'ensemble de leurs composantes et la multiplicité des liens qui sous-tendent la surface, les symboles
graphiques conventionnels ne permettent de transcrire ces informations que de façon incomplète et artificielle.

En fait la cartographie, qui a pour objet de représenter des milieux physiques ou humains issus d'interactions complexes, ne dispose pas de signes graphiques intuitifs et univoques pour visualiser la complexité à l'échelle du point élémentaire. La question qui nous intéresse alors est de savoir si nous pouvons trouver, dans le domaine graphique, un descriptor susceptible de répondre de façon plus pertinente au problème des phénomènes composés ?

Le concept de composition tel que nous le voyons, se distingue de la simple addition des valeurs, dans la mesure où il inclut l'identification qualitative des différentes composantes représentées, leur hiérarchie, leur degré de singularité, ou leurs discontinuités. Ces notions peuvent-elles être transcrites par une classe de signes graphiques ayant une signification plus large et plus fiable que des conventions sans support théorique explicite ? En définitive, est-il possible d'accroître les possibilités logiques de ces cartes et images qui servent de support à la pensée, en redéfinissant le sens de l'un ou l'autre de leurs signes élémentaires ?

En examinant les différents modes de variation de la tache graphique on constate que l'un d'entre eux, la couleur, possède des caractéristiques physiques remarquables qui n'ont été que partiellement prises en compte sur le plan du langage et de sa logique. D'une part la couleur, qui est issue de la décomposition de la lumière, est décrite en physique par un modèle ondulatoire de propagation de l'énergie. D'autre part l'ensemble des variations de la couleur est représentée en physique par un système à trois dimensions, dont deux pour la seule variation de chromaticité. Or, l'utilisation graphique de la couleur est habituellement monodimensionnelle, et la couleur est généralement réduite à des fonctions de simple différenciation, de signalisation, ou à des effets esthétiques. Peut-on tirer partie des propriétés physiques particulières de la couleur pour en déduire une logique d'utilisation plus élaborée ?

On observe par ailleurs une correspondance intéressante entre le domaine graphique et celui du traitement d'image : c'est le même modèle physique, l'onde, qui intervient dans la décomposition de la tache lumineuse, et qui sert à décomposer et à filtrer les surfaces et les formes complexes par analyse spectrale ; cela veut dire que c'est le même concept qui s'applique à la tache graphique et aux surfaces, avec des propriétés semblables.

On note enfin que la mathématisation des comportements ondulatoires et la mise au point des transformations de Fourier ont fait de l'onde un outil majeur d'analyse de formes et de phénomènes complexes.

Il apparaît ainsi entre la couleur et l'analyse des surfaces une convergence de propriétés et d'analogies, fondées sur la notion d'onde, et encore exploitées sur le plan graphique ; notre objectif est d'utiliser ces possibilités pour donner à la couleur, qui fonctionne sur un mode complexe, une signification et une logique complexes.

Après le rappel de quelques généralités sur l'onde, la couleur, et l'analyse spectrale, nous présenterons un protocole opératoire de procédé graphique fondé sur l'approche ondulatoire de la couleur, que nous désignons par le sigle IRISOS ; enfin nous conclurons par un examen des avantages ou des inconvénients du procédé proposé.

**L'onde et sa faculté de composition :**

L'onde décrit une propagation d'énergie par vibrations périodiques. Elle est caractérisée par sa longueur d'onde qui est la distance entre deux crêtes, et par l'amplitude qui correspond à la hauteur des crêtes. La notion de longueur d'onde correspond à une approche spatiale du phénomène ondulatoire, et intéresse particulièrement la cartographie...
car elle permet de décrire les espacements entre les objets. L'équivalent temporel de l'onde est la période, l'inverse de la période est la fréquence ; l'approche fréquentielle qui est la plus courante, sert à décrire l'évolution des phénomènes dans le temps.

Les ondes se manifestent sous différentes formes : vagues, sons, ondes électromagnétiques, comportements quantiques, qui sont toutes décrites en mathématiques par un modèle unique, une courbe sinusoidale qui peut être considérée comme la projection d'un mouvement circulaire. Tout système, oscillant ou tournant, produit un signal ondulatoire, et l'onde constitue un modèle fondamental de la physique.

Erwin Schrödinger écrivait, en parlant de physique quantique, que ce qui l'intéresse dans la description ondulatoire ce n'est pas tellement qu'elle corresponde avec la réalité observée, mais le fait qu'"elle est censée nous donner des informations au sujet des faits observés et de leurs relations de dépendance mutuelle." Si nous appliquons une description et une signification ondulatoires aux couleurs du domaine graphique, c'est avec un objectif analogue, à savoir une description plus féconde des relations entre les lieux, plutôt que pour la pertinence du modèle.

![Diagramme de longueur d'onde et amplitude](image)

L'onde, qui décrit donc des déplacements d'énergie, se distingue par une propriété spécifique, le phénomène d'interférences qui consiste en une combinaison d'ondes : plusieurs ondes simples s'additionnent pour former une onde complexe, et inversement une forme complexe peut être décomposée en ondes simples. Cette propriété de combinaison se manifeste dans de nombreux effets optiques, notamment dans les phénomènes d'irisation. On la retrouve également sous des formes logiques et opératoires dans les techniques d'analyse spectrale.

**La couleur, une variable graphique de nature ondulatoire**

Si l'on présente la tache graphique comme une tache de lumière, directe sur l'écran, ou réfléchie sur le papier, alors on peut considérer que l'onde occupe une place essentielle dans la représentation graphique. La lumière blanche résulte de la combinaison des ondes de la tranche du rayonnement électromagnétique dite "visible". Newton en a réalisé la décomposition au moyen d'un prisme de verre ; le résultat en est le spectre des couleurs qui consiste en une série ordonnée de tonalités dont les longueurs d'onde s'échelonnent entre 400 nanomètres pour le bleu, et 700 nm pour le rouge.(fig.)

Lors de la recomposition des couleurs, les ondes se combinent par addition de leurs amplitudes, mais gardent leur identité, qui apparaît dans la tonalité de la couleur : le spectre total se synthétise sous la forme de la lumière blanche, achromatique, et les fractions de spectre sous la forme de résultats colorées.

*Chaque couleur peut ainsi être considérée comme l'apparence sensible d'une combinaison d'ondes d'une partie du spectre.*
L'organisation de base de l'espace des couleurs peut être reconstituée à partir du spectre : l'axe spectral linéaire, allant du bleu au rouge, est réorganisé sous la forme d'une boucle fermée dont les extrémités rouges et bleues se rejoignent ; c'est le cercle des couleurs. En élévant un axe perpendiculaire de luminosité au centre de ce cercle des couleurs, on obtient un volume des couleurs. Ce volume constitue un modèle qui décrit toutes les variations de couleur ; sa structure est à la base de tous les autres volumes, sphère, cube, double cône, etc.

Finalement, c'est tout l'espace des couleurs qui peut être considéré comme le résultat d'une décomposition spectrale d'un axe de luminosité Noir/Blanc qui se "déplierait" à partir de son centre. Les axes de cet espace de couleur à 3 dimensions sont alors : la tonalité dominante (T), la valeur de luminosité (L), et la saturation (S) qui est la distance de l'axe achromatique (L) au cercle de chromaticité maximum (T).

L'analyse spectrale et le filtrage des surfaces :

En se fondant sur les propriétés de combinaison et de décomposition des ondes, Joseph Fourier a établi les relations qui permettent de décomposer une forme complexe irrégulière en une série de composantes périodiques simples, liées harmoniquement. Chaque composante est la somme d'une onde sinus et d'une onde cosinus.

La transformée de Fourier (TF) mesure la puissance de chaque onde élémentaire; leur ensemble constitue le spectre des puissances. La transformation de Fourier a été reconnue comme une découverte majeure, et son rôle dans l'analyse des phénomènes complexes a conduit à la mise au point d'un algorithme rapide, appelé FFT pour Fast Fourier Transformation.

L'analyse spectrale d'une surface géographique repose sur l'idée que l'hétérogénéité spatiale peut être considérée comme le résultat de l'addition d'une somme de régularités. La surface à analyser est filtrée par FFT en composantes spatiales périodiques, en deux étapes. Dans l'étape 1, les amplitudes de la surface initiale sont reclassées sous forme de spectre suivant les longueurs d'onde spatiales et l'orientation. Dans l'étape 2, l'image initiale est filtrée en sous-images suivant les longueurs d'onde.
Le filtrage (2) aboutit à un empilement de composantes ordonnées suivant la longueur d'onde fondamentale et ses harmoniques, et il révèle un nouvel axe, "intercomposante" et *spectral*. Cette pile de sous-images a la même géométrie que la vue naturelle initiale, mais ici les amplitudes sont connues pour chaque longueur d'onde, suivant l'ordre du spectre. L'information nouvelle, spectrale, réside dans la description des différentes longueurs d'onde impliquées dans la surface initiale et de leur importance relative. Cette information peut être caractérisée par l'identification des composantes, l'évaluation de leur degré de singularité ou de confusion, et la mise en ordre des dominantes.

Pour accéder à ces nouveaux plans d'information et les rendre utilisables, il faut donc d'abord les définir, puis les reconstruire, et finalement en proposer une visualisation spectrale qui complète l'image des amplitudes. Pour réaliser cela, *nous utiliserons la couleur à la fois comme modèle de définition et de synthèse, et comme moyen de visualisation*.

**La synthèse graphique IRISOS :**

Le problème consiste à trouver une solution autre qu'une simple somme des composantes filtrées. Il s'agit de faire la synthèse des composantes tout en gardant et en mettant en valeur les spécificités du filtrage, c'est-à-dire la hiérarchie des longueurs d'onde et la façon dont elles interviennent dans la surface. On s'écarte donc d'une logique de simple addition consistant à sommer les amplitudes en masquant les singularités, ce qui aboutirait à une image identique à l'image de départ. La solution que l'on recherche est un mode de synthèse qui préserve les informations qualitatives et hiérarchiques contenues dans la pile des composantes.

Or nous avons vu que la synthèse ou composition des couleurs permet de combiner des amplitudes en préservant les différences de qualités - couleurs/longueurs d'onde - qui sont résumées par la tonalité dominante et la saturation.

En raison des constats précédents une *première proposition graphique* peut être faite :

Puisque la couleur peut être considérée comme une apparence sensible d'une combinaison d'ondes, on utilise ses propriétés de composition pour réaliser la synthèse visuelle des filtrages spectraux d'une surface.

De manière concrète cela signifie que, puisque l'on connaît le comportement "intercomposante" de chaque point de la surface, c'est-à-dire sa courbe spectrale spatiale à travers l'espace des filtrages, *on visualise ce comportement spectral au moyen du spectre des couleurs*. La résultante couleur représentera alors des longueurs d'onde spatiales selon la signification des référentiels couleur 3-D.

![Diagramme de synthèse graphique](image)

*Pour matérialiser cette synthèse par la couleur, on passe par les procédés simplificateurs de la reproduction trichrome, qui permettent de restituer toutes les couleurs et leurs mélanges, par synthèse optique à partir de trois couleurs primaires. Dans ces systèmes de*
primaires on utilise trois pigments ou trois lumières : la synthèse y est simplifiée mais la logique des couleurs reste identique, et l'ordre spectral est préservé sur le périmètre du volume. En infographie où l'on utilise des axes de couleur proches de ceux de la physique (TSL pour Tonalité dominante, Saturation, Luminosité), on a mis au point des algorithmes fiables pour passer d'un système de logique TSL à un système de primaires, pour la reproduction sur papier ou sur écran.

Un changement de référentiel dans le domaine de la couleur équivaut à un changement de point de vue en terme d'informations.

![Diagramme de transformation](image)

L'utilisation du principe des trois primaires pour réaliser la synthèse des filtrages spectraux, constitue une phase-clé du procédé IRISOS, et exige l'énoncé d'une deuxième proposition graphique :

Puisque trois couleurs primaires sont suffisantes pour décrire la totalité du spectre et de l'espace des couleurs, alors trois bandes de longueurs d'onde sont suffisantes pour décrire la totalité du comportement spectral d'une surface et réaliser sa synthèse graphique.

Le principe des primaires permet à la fois d'effectuer la synthèse spectrale des informations, et de la concrétiser matériellement dans une image trichrome CMJ ou RVB. Au moment de la visualisation toutes les tonalités du spectre et leur mélanges sont rendues par synthèse optique, à partir des trois primaires. L'œil et le cerveau recomposent l'information suivant l'échelle du spectre comme si celle-ci n'avait jamais été segmentée en trois bandes.

La deuxième proposition aboutit donc à une simplification des filtrages spectraux qui ne sont plus réalisés pour chacune des longueurs d'onde potentielles de la surface, mais seulement pour trois bandes de longueurs d'onde. Le filtrage de la surface initiale en trois bandes de longueurs d'onde, grandes, moyennes et petites, s'appuie sur une partition en trois de la variance totale du spectre. La synthèse graphique permet de restituer l'image suivant l'ordre continu du spectre.

Pour la mise en pratique des deux propositions IRISOS on utilise comme modèle, un algorithme infographique de changements d'espace de couleurs que l'on adapte pour la transformation des données spatiales issues des filtrages spectraux en trois bandes. L'algorithme de transformation - Primaires --> TSL - effectue alors le calcul de la longueur d'onde spatiale dominante, de la saturation de cette dominante, et de l'amplitude globale des ondes spatiales. Ces algorithmes graphiques, d'utilisation très large, permettent d'effectuer un traitement précis de la couleur, et répondent donc aux conditions de fiabilité exigées en imagerie scientifique.

Le principe simplificateur de la décomposition et de la reproduction ternaire est semblable à ceux que l'on utilise dans de nombreux systèmes de reproduction des arts graphiques et du son. Sa fiabilité est suffisante pour pouvoir être appliqué à des données autres qu'analogiques : il a été utilisé en physique pour des mesures par interférométrie, ainsi que dans l'analyse de la turbulence.
Organigramme :

Le procédé IRISOS peut être appliqué à des surfaces numérisées sous forme de tableau de points. A cause du caractère spatial des images, il convient naturellement à tous les phénomènes géographiques qui peuvent être ramenés à des surfaces.

Avec ce procédé on est dans le domaine des techniques de l'imagerie : cela signifie d'une part, que toutes les informations créées sont issues de l'image, exclusivement ; d'autre part, comme le procédé d'analyse est global, le découpage ou le cadrage de l'extrait à étudier influence le résultat.

Schéma d'application :

Intérêt du procédé :

On peut résumer la démarche exposée ci-dessus en présentant le procédé IRISOS comme la *mise en correspondance de deux systèmes d'interférence* : la couleur qui équivaut à des interférences lumineuses, et la surface considérée comme une résultante de l'interférence d'ondes spatiales. *On utilise le premier système pour visualiser le second.*

Si ce principe de l'image est simple, sa mise en œuvre exige néanmoins un nombre d'opérations et des capacités de calcul conséquentes ; leur importance devrait être gommée peu à peu avec la montée en puissance des outils informatiques.

La principale difficulté d'accès au procédé vient de la redéfinition "spectrale" de la variation de couleur, qui rompt avec les habitudes de lecture de l'imagerie, exception faite des spectrogrammes des physiciens où la couleur représente également des longueurs d'onde. Mais ce phénomène de rupture de la signification, qui vient de l'importation d'un modèle de physique dans le langage graphique, donne accès à de nouvelles possibilités méthodologiques, graphiques, et théoriques.
Sur le plan des méthodes d'analyse de données :

Avec le procédé IRISOS, on poursuit la logique ondulatoire de l'analyse spectrale jusqu'à son terme naturel, c'est-à-dire la synthèse colorée fondée sur un référentiel spectral : il n'y a donc pas seulement cohérence entre la description mathématique et la visualisation, mais aussi complémentarité logique entre les phases d'analyse et de synthèse.

Il ne s'agit pas d'imposer l'onde comme modèle explicatif, mais de la proposer à titre exploratoire, et lui emprunter ses solutions pour filtrer et synthétiser les composantes d'un phénomène : l'onde sert à changer de vision.

La synthèse des ondes se manifeste par des effets d'irisations qui visualisent l'organisation spatiale et la structure profonde de la surface ; la couleur sert de marqueur des composantes d'échelles. La nouvelle image est de nature autre que l'image des amplitudes apparentes : c'est son reflet spectral.

Les filtrages et transformations effectuées fournissent leurs informations suivant trois axes analogues aux axes de la couleur ; deux de ces axes sont nouveaux pour l'analyse spatiale : le premier, l'axe de longueur d'onde spatiale dominante, et le second qui lui est lié, l'axe de pureté, ou de validité, de cette dominante ; la notion de dominante indique ici la longueur d'onde qui porte le plus d'informations. Dans l'espace des couleurs ces deux axes correspondent au plan de chromaticité. Ces nouveaux axes peuvent également être utilisées hors visualisation, pour l'étude de la surface, comme variables morphologiques autonomes.

L'information du troisième axe, celui des valeurs d'amplitudes/luminosité, représente la partie non spectrale de l'information contenue dans la surface initiale, celle qui a été isolée par le procédé IRISOS. Cette information est éliminée de la représentation spectrale. Dans la vision spectrale d'un phénomène les amplitudes ont un caractère secondaire ou résiduel.

Du point de vue graphique :

Ce procédé introduit un descripteur graphique, la couleur spectrale, dédié aux champs de propagation de l'énergie, de la même façon que les courbes de niveau sont dédiées à l'intensité de l'énergie. L'interprétation spectrale de la couleur permet de différencier les longueurs d'onde qui interagissent, et de représenter les notions de hiérarchie, de périodicité, de régularité, et les discontinuités de ces champs de propagation : la couleur est un résumé de la signature spectrale.

Les informations de relations et de hiérarchie entre des objets sont représentées ici à l'échelle du point élémentaire, alors qu'en cartographie elles sont habituellement figurées au moyen des deux dimensions de l'espace graphique.

Les trois dimensions de la tache lumineuse deviennent significatives, au lieu d'une ou deux dans l'utilisation graphique habituelle. On a vu que si on "déplie" l'axe graphique blanc-noir, la couleur s'y insère sous forme de volume avec des propriétés logiques étendues aux trois dimensions de ce volume et des possibilités innombrables de variation graphique.

Parallèlement à cette extension de la signification de la couleur, on est amené à préciser le rôle de la composition colorée comme opérateur de traitement graphique : la synthèse colorée permet ici, d'effectuer des fonctions d'addition ou de soustraction graphiques, tout en conservant le contenu qualitatif et l'ordre de l'information. Dans un contexte
d’informations spectrales bien contrôlées, la couleur devient alors un opérateur de synthèse qui peut prendre place dans ce que Sylvie Rimbert appelle "l’algèbre de cartes".

C’est le volume des couleurs à trois dimensions spectrales TSL qui tient lieu de légende pour toutes les images construites d’après ce procédé ; mais l’interprétation des couleurs et de leur situation dans l’espèce TSL peut être faite, soit en termes d’ondes spatiales, soit en termes théoriques :

- luminosité
- amplitude totale
- désordre

-domaine couleur
domaine des longueurs d’onde
domaine conceptuel

domination
saturation
longueur d’onde dominante

distance
ordre

Du point de vue des systèmes de représentation :

Pour analyser des phénomènes complexes, il faut des langages qui supportent une logique complexe.
Le système de représentation graphique proposé permet de transmettre davantage d’informations avec un faible niveau de bruits. Ce résultat est obtenu grâce à la généralité de sa légende, l’espace des couleurs, et aussi grâce à la cohérence entre les différents relais de la communication : l’information émise, la transmission par la couleur, et le système de réception visuel ont une même organisation tridimensionnelle en bandes, en primaires, et en cônes, et reposent sur un principe de combinaison analogue.

La décomposition par analyse spectrale d’une surface et la synthèse graphique des composantes permettent d’étudier les phénomènes spatiaux suivant une approche globale. En modulant les fines strates de longueurs d’onde, la couleur permet d’intégrer plusieurs dimensions de la réalité dans une même image.

Au-delà des longueurs d’onde, le modèle ondulatoire représente un ordre déterministe à l’épreuve duquel on soumet et on teste une surface. Les résultats graphiques et numériques du procédé permettent d’évaluer l’adéquation et les écarts entre la réalité observée et ce modèle ; on interprète l’espace des couleurs selon une logique d’ordre.
L’axe circulaire couleur/longueur d’onde prend alors une signification de "cohérence ou ordre maximum", et l’axe des amplitudes de luminosité noir/blanc qui résulte du mélange des couleurs signifie "désordre ou entropie": les aires où le modèle ondulatoire fonctionne sont colorées, et se séparent de celles où il ne fonctionne pas et qui sont figurées par les tons neutres noir-gris-blanc. La saturation (S) indique la distance entre les deux systèmes, ou l’écart au modèle ondulatoire. Le procédé joue le rôle de séparateur de modèle.

Ce type d’interprétation est directement lisible sur l’image couleur, des tonalités (T) plus leur saturation (S). La saturation peut également visualisée seule, en noir et blanc.
Conclusions :


Pour répondre à une question d'utilisation particulière nous avons exploré et tenté de redéfinir sur un point précis le champ des significations de la représentation graphique. Le procédé proposé sera sans doute confronté à des difficultés, provenant de la rupture des habitudes de lecture de la couleur, ou à des problèmes de mise en œuvre.

En introduisant la logique combinatoire spectrale dans l'utilisation de la couleur, on constate cependant que l'on aboutit à plusieurs perspectives intéressantes. Le procédé devra maintenant subir des tests d'application sur des données géographiques observées.

Remarque: La présentation complète de ce travail exige des images en couleur. En raison des contraintes de cette édition, elles ont été remplacées par des schémas en noir et blanc, ou supprimées.

Références:

BACHELARD, G. 1934 Le nouvel esprit scientifique, PUF, Paris
BOHM, D. 1987 La plénitude de l'univers, Editions du Rocher
FRANCON, M. 1986 L'optique moderne et ses développements, Hachette, Paris
KEATES, J.S. 1989 Cartographic design and production, Second edition, Longman
MAITTE, B. 1981 La lumière, Editions du Seuil
MANDELBROT, B. 1989 Les objets fractals, Troisième édition, Flammarion
MOLES, A. 1972 Théorie de l'information et perception esthétique, Denoël, Paris
PINSON, G. DEMAILL Y A. FAVRE D. 1985 La pensée, approche holographique, Presses universitaires de Lyon
RIMBERT, S. 1990 Carto-graphies, Hermès, Paris
RIMBERT, S. 1992 Données pour une algèbre de cartes, Société Française de Photogrammétrie et Télédétection, bulletin n° 125, pp 14-25
SCHWEIZER, P. 1987 Infographie, t.II, Presses polytechniques Romandes, Lausanne
ZWIMPFER, M. 1992 Couleur-optique et perception, Dessain et Tolra, Paris
42. Deutscher Kartographentag
Köln 1993

Unter dem Leitthema „Deutsche Kartographie im europäischen Umfeld“ wurden die Themenbereiche

- **Topographische Kartographie**
  (J. Schaffer, D. Grothenn)

- **Thematische Kartographie**
  (J. Dodt, F. Ormeling)

- **Kartographie und Geoinformation**
  (W. Gillessen, K. Barwinski)

- **Berufsfeld der Kartographie**
  (S. Schulz, H. Hansen)

in 8 Vorträgen behandelt. Die Kurzfassungen der Vorträge sind auf den folgenden Seiten wiedergegeben.
Kurzfassungen der Vorträge des Deutschen Kartographentages

Topographische Landeskartenwerke im vereinigten Deutschland – die Notwendigkeit zur Angleichung und die Realisierung

von Jörg Schaffer


Mit der im Zeitraum 1955 bis 1970 durchgeführten topographischen Neuaufnahme im Maßstab 1:10.000 und der anschließenden kartographischen Neuherstellung der topographischen Karten Karten 1:100.000, 1:25.000, 1:50.000, 1:100.000 und 1:200.000 wurden in der ehemaligen DDR, den heutigen neuen Bundesländern, eigenständige topographische Kartenwerke bearbeitet und herausgegeben.


Die Angleichung der topographischen Kartenwerke 1:10000, 1:25000, 1:50000, 1:100000 und 1:200000 der ehemaligen DDR an die topographischen Landeskartenwerke der Bundesrepublik Deutschland beinhaltet die Umstellung auf eine einheitliche geodätische Grundlage und einen einheitlichen Blattschnitt (Regelblattschnitt, auch als Normalblattschnitt bezeichnet), die Einführung einer einheitlichen Aufmachung einschließlich Kartenrahmen- und Kartenrandausstattung entsprechend der jeweiligen Redaktionsdokumente (Musterblätter für die topographischen Landeskartenwerke der alten Bundesländer) sowie die Einführung einzelner neuer Kartenzeichnen und Veränderungen in der Anwendung der bisherigen Kartenzeichnen, Schriftzusätze und Charakteristikum. Das topographische Kartenwerk 1:10000 fungiert weiterhin als flächendeckende Grundkarte; die Deutsche Grundkarte 1:50000 wird in den neuen Bundesländern mittelfristig nicht als flächendeckender Grundkartenwerk aufgebaut.

Für das topographische Kartenwerk 1:10000 ergibt sich im Regelblattschnitt ein Blattsiegel in den Ausmaßen A 5' und B 3', die Blattnumerierung wird entsprechend folgendem Beispiel vorgenommen: 7918 – NW (NO, SW, SO).


Im Rahmen des Vortrages werden, ausgehend von einem Vergleich der topographischen Kartenwerke in den alten und neuen Bundesländern, die Inhaltsbestandteile der Angleichung detailliert aufgezeigt und erläutert sowie die eingesetzten kartographischen Technologien und die angewandten Technologien zur Realisierung der Angleichung an ausgewählten Beispielen vorgestellt.


Die Herstellung einheitlicher topographischer Landeskartenwerke für die Bundesrepublik Deutschland bleibt eine langfristige Aufgabe, die aus heutiger Sicht nur über den Aufbau des bundeseinheitlichen Amtlichen Topographisch-Kartographischen Informationssystems (ATKIS) und seiner Präsentation mit einer modernen Kartengrafik zu gewährleisten ist.
Einheitliche Gestaltung der amtlichen topographischen Kartenwerke in Europa?
von Dieter Grothenn


Bei der Suche nach Vereinheitlichungsmöglichkeiten bietet sich vordergründig eine Übernahme militärischer Kartenwerke in den zivilen Kartengebrauch an, da in der Vergangenheit sowohl in Osteuropa wie auch in Westeuropa mehr oder weniger einheitliche militärische Karten geschaffen wurden.

Dieser Weg dürfte aber nicht gangbar sein, da die kartographische Zeichensprache militärischer Karten nicht selbstverständlich als das angemessene Ausdrucksmittel für eine topographische Karte angesehen wird, die für den universellen Kartengebrauch bestimmt ist.


So könnte folgende Stufenlösung zum Ziel führen:

4. Vereinheitlichung der Maßstabsreihe zumindest in den Maßstäben 1:25 000 und kleiner.
5. Vereinheitlichung des Karteninhalts.

In den unter 1. bis 3. genannten Bereichen sind Aktivitäten zu beobachten, die Hoffnung auf eine baldige Vereinheitlichung nähern. Dagegen dürfte eine Verwirklichung der Stufen 4 und 5 noch in weiter Ferne liegen, da sie neben dem Problem der Arbeitsmenge noch nationale Traditionen und psychologische Hürden überwinden müßte.


Trotz dieser Schwierigkeiten muss die amtliche topographische Landeskartographie dafür sorgen, daß der Bedarf an topographischen Basisinformationen nicht nur national, sondern auch europaweit nutzergerecht gedeckt wird. Darüber hinaus wäre einheitlich gestaltete topographische Landeskartenwerke ein sichtbarer Ausdruck europäischer Einheit.
Die Begriffe Altlasten bzw. Altlast-Verdachtsflächen bezeichnen Grundstücke, die – aufgrund ihrer früheren Nutzung als industriegewerbliche Produktionsstätten ("Altstandorte"), Rüstungs- und Militäranlagen ("Rüstungsaltlasten") oder als Flächen der Abfallablagerung, -verfüllung oder -aufhaldung ("Altablagerungen") – in solchem Maße verunreinigt sind, daß sie nachweislich bzw. mit hoher Wahrscheinlichkeit eine Gefahrenquelle für Mensch und Umwelt darstellen. Die Erkundung von Altlasten/Altlastverdachtsflächen, d.h. ihre Lokalisierung, die Ermittlung der Schadstoffeinträge und die Bewertung des Gefährdungspotentials, hat sich zu einem Problem- und Aufgabenfeld entwickelt, das inzwischen von den verschiedensten Fachdisziplinen bearbeitet wird. Hierzu gehört in aller Regel auch die Kartographie, die auf zwei Gebieten zur Altlastenerkundung beiträgt bzw. beizutragen vermag:

1) zur Erfassung der Altstandorte und Altablagerungen in flächendeckenden Erhebungen (Gebietsinventuren) wie in einzelfallbezogenen Standortrecherchen,
2) zur Dokumentation und Visualisierung der im Erfassungsprozess ermittelten räumlichen Daten, Informationen und Erkenntnisse.


Im zweiten Einsatzgebiet der Kartographie in der Altlastenerkundung, bei der graphisch-kartographischen Visualisierung der ermittelten Daten und Informationen, sind mittlerweile die verschiedensten Formen der Darstellung entwickelt worden, wobei durch entsprechende Verordnungen und Zeichenvorschriften in Ansätzen bereits eine Standardisierung erfolgt ist. Entscheidende Bestimmungsfaktoren der Darstellungsformen sind neben dem (problematischen) Aspekt des Datenschutzes – vor allem die Art der Erhebung und – damit i.a. zusammenhängend die Größe des Untersuchungsgebiets sowie das jeweils angestrebte Informationsniveau.

Dementsprechend werden bei Gebietsinventuren größerer Verwaltungseinheiten (Länder, Regierungsbezirke) in der Regel Verbreitungskarten auf der Basis kleinmaßstäbiger Karten der Verwaltungsbezirke mit Gattungs-/ObjektSignaturen für die Grundkategorien der Verdachtsflächen sowie einer bestenfalls groben Kennzeichnung der Bestandszustände erarbeitet. Bei kleineren Untersuchungsräumen (z.B. Stadtgebiet) hat sich die Darstellung auf der Grundlage topographischer Karten in Maßstäben zwischen 1:10000 und 1:25000 mit Flächenkonturierung und Zusatzsignaturen für weitere altlastsignifikante Elementar-Informationen (Stoff- und Gefährdungspotential, Kriegseinwirkungen u.a.) durchgesetzt.


Um die Fülle und Vielfalt der auf verschiedenen Maßstabs- und Informationsebenen ermittelten Daten flexibel zu bewältigen, um ferner die Aufbereitung der Informationen zu beschleunigen und um im weiteren Verfahrensablauf auch Daten aus der Geländesicherung problemlos integrieren zu können, wird die konventionelle Datenaufbereitung in Form analoger Karten zunehmend durch rechnergestützte Verfahren ergänzt bzw. abgelöst. Hierzu werden Beispiele zum Einsatz des GIS Arc/Info und der Software Geo-Map vorgestellt.
Neue Formen, Konzepte und Strukturen von Nationalatlanten
von Ferjan Ormeling


Erzeugung und Vermarktung digitaler topographischer Daten – Arbeitsteilung zwischen öffentlicher Verwaltung und privater Wirtschaft
von Wolfhart Gillessen

Der Vortrag konzentriert sich auf die Vermarktung topographischer Daten und das anzustrebende Rollenverständnis zwischen öffentlicher Verwaltung, privater Wirtschaft und Kunden.


Da selbst andere öffentliche Bedarfsträger sich ihre eigenen Geo-Grunddaten erfassen (lassen werden), ist die Schnelligkeit der Bereitstellung der wichtigsten Grunddaten von existentieller Bedeutung für das öffentliche Vermessungswesen.

Der Kunde – ob öffentlich oder privat – verlangt ein überschaubares Datenangebot mit klaren Ansprechpartnern, einfachen Verträgen, gesicherter Wartung und Preisen die marktgerecht, also nicht das «return on invest» in 3 Jahren kalkulieren, sind. Die föderale Struktur im Vermessungswesen wird vom Autor hierbei als Nachteil angesehen.

Da die GIS-Anbieter von der Datenlieferfähigkeit profitieren, sind sie die besten Vertriebspartner der Vermessungsverwaltung. Der Kunde sieht das gesamte GIS-System und die Kostenblöcke (nachstehend Durchschnittswerte)

<table>
<thead>
<tr>
<th>Grund- und Kundendaten</th>
<th>80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software GIS</td>
<td>12%</td>
</tr>
<tr>
<td>Hardware</td>
<td>8%</td>
</tr>
</tbody>
</table>
und wird die Kalkulier- und Wartbarkeit des Datensegmentes als größten Anteil ansehen! Alle Marktstudien prognostizieren einen Zuwachs des mitteleuropäischen GIS-Marktes um mehr als 30% jährlich, falls (endlich) die amtlichen Geodaten zur Verfügung stehen. Ein Blick über den Atlantik zeigt, daß dort wegen der preis­
werten Nutzungsmöglichkeit der amtlichen (auch für militärische Zwecke erfaßten) Geodaten der GIS-Markt wesentlich früher expandieren konnte.

Den Datenvertrieb über kompetente Partner aus der privaten GIS-Wirtschaft aufzubauen, wird vom Autor als Überbrückungsmaßnahme begrüßt.

Da das Produkt gut ist und der Markt »boomen« wird, könnte auf Dauer bei einem »Dienstleister Selbstver­
ständnis« der öffentlichen Vermessungsverwaltung der Datenvertrieb auch selber gewinnbringend organi­
siert werden.

---

Geoinformation und Kartographie in Europa
von Klaus Barwinski


Die öffentliche Verwaltung sieht sich als Produzent für die Geobasisdaten, auf die bezogen viele Anwendungen aufgebaut werden können. CERCO will sicherstellen, daß von seinem Service-Center Daten bezogen werden können. Diese Daten stammen aus den nationalen Agenturen und werden über ein Informations­netzwerk zusammengeführt. Es ist also nicht daran gedacht, ein europäisches Landes­vermessungs­amt zu schaffen, sondern sicherzustellen, daß zusammenge­arbeitet werden kann, und daß dem Verbraucher die Daten in standardisierter Form zur Verfügung gestellt werden können. Ein solcher Netzverbund über die Länder hinweg kann diese Aufgaben übernehmen. Zum gleichen Zeitpunkt muß man sich auch Gedanken über das »Copyright« und eine gemeinsame europäische Preispolitik machen. Hier gilt es, die unterschied­lichsten Ansätze zu harmonisieren und zu einer europäischen Linie zu finden.

Im Vordergrund der europäischen Aktivitäten steht also das Schaffen einer Datenbasis, die für unterschied­liche Anwendungen direkt und ungeneralisiert zugrunde glegt werden kann oder aus dem für bestimmte Zwecke Karten mit einer schematischen Aussage abgeleitet werden können. Wir müssen uns daran gewöh­nen, daß das »Original« gleich dem »Speicherinhalt« ist. Die Karte wird zum Informations­instrument und dient der Vermittlung und Präsentation von Ideen im privaten, insbesondere aber auch im politischen Raum.


Berufsentwicklung in der Kartographie
von Prof. Dr. Siegmund Schulz

Das Berufsfeld der Kartographie unterliegt derzeit einem tiefgreifenden Wandel. Dieser ist gekennzeichnet
durch den immer rascheren Übergang zur digitalen Kartenherstellung bzw. -fortführung, den Einsatz von
Datenbanken bei der Anwendung von raumbezogenen Informationssystemen, die Verfügbarkeit von Fern-
erkundungsdaten, die Integration elektronischer Verfahren in der Reproduktionstechnik sowie – nicht zuletzt
besonders auch von ökonomischer Bedeutung – die Änderung des Kaufverhaltens bei den Kartennutzern.

Als gemeinsames Charakteristikum für die Veränderungen in den jeweiligen Teilbereichen der Kartographie
ist der ausgeprägte Trend zum Einsatz von Microcomputern zu nennen. Die PCs werden immer leistungsfähi-
ger bei insgesamt sinkenden Investitionskosten, was einerseits zur Verdrängung von Großrechenanlagen
führt und andererseits das Kaufinteresse großer Bevölkerungskreise steigert. Diese Entwicklung hat Auswir-
kungen auf das gesamte Berufsfeld der Kartographie: Abläufe in der Kartenproduktion müssen neu konzi-
piert werden, zusätzlich neue Möglichkeiten der Präsentation von kartographischen Informationen sind zu
berücksichtigen und nicht zuletzt muß im Bereich der Aus- und Weiterbildung eine Bereinigung und Aktua-
lisierung der Lehrinhalte erfolgen.

In Anlehnung an die Entwicklung in der Kommunikations- und Unterhaltungsbranche werden Einzelpersonen
zukünftig neben der gedruckten Karte auch Disketten und CDs mit kartographischen Informationen als
Datenträger für die Visualisierung am heimischen PC-Bildschirm einsetzen. Insbesondere durch audiovisu-
elle Elemente kann dabei die Aussagemöglichkeit einer Karte erweitert werden und zugleich kann eine
gelungene Präsentation zu einem Erlebnis besonderer Art werden.

Das Referat soll außerdem verdeutlichen, daß eine solide Kenntnis und Anwendung traditioneller kartogra-
phischer Grundsätze nach wie vor auch bei der Herstellung mit rechnergestützten Verfahren erforderlich ist.
Andererseits muß die Kartographie gegenüber neuen Technologien im Interesse des Faches stets aufge-
schlossen bleiben, damit originäre kartographische Aufgaben nicht von Fachfremden als Betätigungsfeld
entdeckt werden. Die Berufsentwicklung in der Kartographie wird insgesamt zu einer noch stärkeren inter-
disziplinären Zusammenarbeit zwischen allen relevanten Berufsfeldern führen: Geographie, Geodäsie,
Reproduktionstechnik, Fernerkundung. In einigen Fällen wird das kartographische Know-how fest in neuen
Berufsfeldern integriert sein, z.B. in der Geoinformatik.

Kartographisch gestützte Analyse- und Planungs-Systeme für Marketing und Vertrieb
von Hans Hansen

Die Nutzung der Karte als Medium zur Analyse und zur Visualisierung geographischer Strukturen des Mark-
tes soll unter fünf Aspekten betrachtet werden:

1) Es geht um die Erschließung der geographischen Dimension von Daten und Zusammenhängen für
Marketing und Vertrieb
Mit der Entwicklung der PC-gestützten digitalen Informationsverarbeitung und dem operativen Einsatz
 der DV in den Unternehmen sind für die Modellierung und Visualisierung geographischer Verteilungen und
Interaktionen von Marktdaten und Prozessen neue Voraussetzungen entstanden. Dafür lassen sich 4 Tätig-
keitsfelder in den Unternehmen unterscheiden und die zuzuordnenden Aufgaben beispielhaft nennen:
• Marktforschung; regionale Segmentierung der Märkte nach den geographischen Verteilungsmustern
vom Konsumenten und Marktpotentialen,
• Werbekommunikationsplanung; die ungleiche Verteilung der Zielgruppen verlangt nach geographi-
scher Selektion der Medien,
• Vertriebsplanung, Standortbewertung und Außendienststeuerung; Direktvertrieb senkt Kosten und
fordert die Kundenbindung, der harte Handelswettbewerb steigert die Anforderungen an Standortwahl und
-support, die Kostenentwicklung verlangt nach Optimierungsmöglichkeiten des Außendiensteinsatzes,
• Controlling; erfolgreiche Unternehmensentwicklung basiert auf Kostentransparenz, dabei gewinnt die
Operationalisierung der geographischen Kostendimension zunehmend an Bedeutung.
2) Die mikrogeographische Marktsegmentation gestützt auf ein integriertes digitales geographisches Datensystem liefert die methodische und datentechnische Basis zur Lösung dieser Aufgaben. Das Marketing denkt in produktorientierten Zielgruppen. So, wie sich Konsumenten anhand ihrer demo­ 
graphischen Merkmale oder besser noch über Life-Style-Typologien nach Zielgruppen klassifizieren lassen, können auch geographisch definierte Markt­ 
zellen, also Orte, Ortsteile, Straßenzüge anhand ihrer statistischen Merkmale klassifiziert und dementsprechend nach Zielgruppen-Nähe segmentiert werden. Datentechnische Voraussetzung für eine DV-gestützte Erschließung von Marketinginformationen aus Kundendaten, Erhebungen und anderen Quellen ist ein integriertes digitales geographisches Daten­ 
system bestehend aus: 
  • Referenzdateien mit allen Orts-, Straßennamen, Hausnummern und Gebietskennziffern als «Binde­ 
glied» zur externen Datenwelt, 
  • Lagekoordinaten zur Lokalisierung aller geographischen Elemente des Systems sowie entsprechende Grenzkoordinaten, 
  • Straßennetzdaten zur Modellierung von Interaktionen, 
  • digitale topographische Karten als visuelle Orientierungsgrundlage.


4) PC-gestützte geographische Marktinformationssysteme sind das künftige Arbeitsmittel für die lokale Marktbearbeitung. Die rapide Entwicklung der PC-gestützten kartographischen Technologie manifestiert sich u.a. in den immer leistungsstärkeren geographischen Informationsystemen (GIS). Die Nutzung dieser Tools für die Aufgaben der Marktbearbeitung verlangt jedoch:
  • eine wesentlich stärkere Kopplung mit anderen im Marketing gebräuchlichen PC-Tools, mit einem gemeinsamen Zugriff auf die gleiche Datenbank,
  • eine konsequent benutzerorientierte Menutechnik, die ihn von methodisch komplizierten Aufgaben (z.B. in der thematischen Kartographie) entlastet,
  • eine «laienorientierte» Kartiertechnik unter Verwendung von Hybridsystemen, in denen gerasterte topographische Karten (auf verschiedensten Maßstabsebenen) den «Orientierungshintergrund» für vektorbasierte thematische Nutzerinformationen bilden,
  • die simultane und visuell vernetzte Präsentation der Information auf Karten und in der Form von Tabellen, Datenblättern, Texten etc.,
  • und flexible Interaktionsmöglichkeiten im Sinne eines «digitalen Fingers», der den Benutzer beim Dialog mit der Datenbasis und seinen Rechenmodellen unterstützt, der ihm erlaubt, Kartenebenen oder -ausschnitte zu wählen oder gezeigte Informationen zu markieren u.v.a.m.

Dafür werden beispielhafte Systeme vorgestellt und, sofern die technischen Möglichkeiten es gestatten, soll das Leistungsspektrum an Beispielen demonstriert werden.

| Board Ch. | 19 | Akeno K. | P151 |
| Bos E. | 10 | Anderson J.M. | 91 |
| Brown A. | 2 | Aranaz F. | 92 |
| Dahlberg R. | 6 | Artimo K. | 6, 37, 112, P128 |
| Dörrstock R. | DKT | Asche H. | 126 |
| Guptill S.C. | 9 | Balcerzak J. | P149 |
| Miranda i Canals J. | 5 | Baowen C. | P136 |
| Moellering H. | 1 | Bartlett D. | 62 |
| Morita T. | 15 | Beard K. | 41 |
| Morrison J. | 7 | Boenes M. | 43 |
| Muller J.C. | 17 | Benzecri J.-P. | P159 |
| Olson J.M. | 3 | Berlyant A.M. | 56 |
| Ormeling F. | 16 | Bianchin A. | P138 |
| Papp-Vary A. | 8 | Björke J.T. | 42 |
| Pearce D. | 12 | Blok C. | 105 |
| Peterson M.P. | 21 | Board C. | 2 |
| Rimbert S. | 13 | Brandenberger C. | 81 |
| Siekierska E. | 20 | Bräuninger T. | 14 |
| Spiess E. | 4 | Brooke G. | P130 |
| Tikunov V.S. | 18 | Broome F. | 67 |
| Wirth K. | DKT | Brown A. | 57, 80 |
| Wood M. | 11 | Brüggenmann H. | 50 |
| Zhang Q. | 14 | Buchroithner M. | 75 |
| Buttenfield B.P. | 98 | Capek R. | 97 |
| Carosio A. | 60 | Carter J.R. | 27 |
| Cartwright W.E. | 124 | Chen C. | 36 |
| Cheng B. | 100 | Christian D.G. | P137 |
| Cotter D. | 85 | Coulson M.R.C. | 40 |
| Davies M. | 12 | Dhib M. | 90 |
| Du D. | 109 | Dutton G. | 85, 88 |
| Dymon U.J. | 46 | Elg M. | 93 |
| van Elzakker C.P.J.M. | 80 | Emelyanov A. | 19 |
| Eseka E.N. | 18 | van Essen R.J. | 8 |
| Feinstein A. | 72 | Fischer R. | 75 |
| Fisher P. | 63 | Forer P. | 110 |
| Francula N. | 7 | Freitag U. | 2 |
| Gao J. | 108 | Gao R. | 70 |
| Gartner G. | P155 | Gdowski B. | P149 |
| Goldworm M. | 85 | Green D.R. | 115 |
| Grelot J.P. | 120 | Grosser K. | P131 |
| Grünreich D. | 23 | Gudbergson G.M. | 121 |
| Guthill S.C. | 65 | Hagen D. | 49 |
| Han P. | P141 | Haraguchi K. | 45 |
| Harju E.-S. | P157 | Hehak W. | P129 |
| Helmfred S. | P147 | Herrmann C.M. | 126 |
| Hocking D. | 69 | Hua Y. | 108 |
| Huffman N.H. | 3 | Hurni L. | 5 |
| Hutzler E. | 36 | Ikhuoria I.A. | 18 |
| Inaba K. | 61 | Inazawa Y. | 45 |
| Iwaniak A. | 39 | Jaatinen S. | 68 |
| Jackson I. | 51 | Jäger E. | 58 |
| Jiang B. | 82 | Jiang J. | 52 |
| Jordan P. | 30 | Kadmon N. | 104, P134 |
| Kainz W. | 114 | Kamada K. | 61 |
| Kamiya I. | P152 | Kanakubo T. | 2 |
| Karl D. | 123 | Keling W. | P153 |
| Keller C.P. | 69 | Kerimov G.A. | P158 |
| Kern H.F. | 59, 119 |

*) Articles printed in italics are to be found in Proceeding II

701
<table>
<thead>
<tr>
<th>Name</th>
<th>Programme Nr(s)</th>
<th>Name</th>
<th>Programme Nr(s)</th>
<th>Name</th>
<th>Programme Nr(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kilpeläinen T.</td>
<td>24</td>
<td>Ormeling F.</td>
<td>125</td>
<td>Tuukkanen K.</td>
<td>6</td>
</tr>
<tr>
<td>Kim K.</td>
<td>85</td>
<td>Panasiuk J.</td>
<td>P149</td>
<td>Uçar D.</td>
<td>87</td>
</tr>
<tr>
<td>Klimczak H.</td>
<td>P143</td>
<td>Parker T.</td>
<td>72</td>
<td>Une H.</td>
<td>P151</td>
</tr>
<tr>
<td>Kostka R.</td>
<td>161</td>
<td>Peng W.</td>
<td>20</td>
<td>Unwin D.</td>
<td>63</td>
</tr>
<tr>
<td>Kraak M.J.</td>
<td>111</td>
<td>Peterson M.P.</td>
<td>101</td>
<td>Urabe B.</td>
<td>P152</td>
</tr>
<tr>
<td>Kremers H.</td>
<td>P144</td>
<td>Pillay R.</td>
<td>103</td>
<td>Ushakova L.A.</td>
<td>55</td>
</tr>
<tr>
<td>Krzywicka-Blum E.</td>
<td>106</td>
<td>Pravda J.</td>
<td>2, 88, P142</td>
<td>Uthe A.-D.</td>
<td>29</td>
</tr>
<tr>
<td>Kubik K.</td>
<td>P130</td>
<td>Qian H.</td>
<td>P139</td>
<td>Vasconcellos R.</td>
<td>113</td>
</tr>
<tr>
<td>Kumaki</td>
<td></td>
<td>Raveneau J.</td>
<td>73</td>
<td>Veregin H.</td>
<td>66</td>
</tr>
<tr>
<td>Langran G.</td>
<td>107, P150</td>
<td>Rimbert S.</td>
<td>2</td>
<td>Vickus G.</td>
<td>76</td>
</tr>
<tr>
<td>Lapaine M.</td>
<td>7</td>
<td>Robbins R.G.</td>
<td>9</td>
<td>Vujakovic P.</td>
<td>26</td>
</tr>
<tr>
<td>Laurema M.</td>
<td>112</td>
<td>Ruutiainen T.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laureti L.</td>
<td>96</td>
<td>Samimi E.</td>
<td>95</td>
<td>Wang D.</td>
<td>52</td>
</tr>
<tr>
<td>Lee D.</td>
<td>122</td>
<td>Sarjakoski T.</td>
<td>24, 71</td>
<td>Wang Z.</td>
<td>15, 20</td>
</tr>
<tr>
<td>Li H.</td>
<td>54</td>
<td>Scharle W.</td>
<td>28</td>
<td>Wang J.Y.</td>
<td>10, 25</td>
</tr>
<tr>
<td>Li J.</td>
<td>54, 70</td>
<td>Schewe J.</td>
<td>P156</td>
<td>Wang H.</td>
<td>P164</td>
</tr>
<tr>
<td>Li L.</td>
<td>109</td>
<td>Schneider C.</td>
<td>78</td>
<td>van der Wel F.J.M.</td>
<td>102</td>
</tr>
<tr>
<td>Liang Q.</td>
<td>P141</td>
<td>Sclaf er M.-N.</td>
<td>64</td>
<td>Wen C.</td>
<td>P148</td>
</tr>
<tr>
<td>Lin A.S.</td>
<td>10</td>
<td>Seferi E.</td>
<td>95</td>
<td>van Wersch A.</td>
<td>55</td>
</tr>
<tr>
<td>Lindholm M.</td>
<td>71</td>
<td>Shehu A.</td>
<td>P145</td>
<td>Wilmersdorf E.</td>
<td>83</td>
</tr>
<tr>
<td>Lipej B.</td>
<td>118</td>
<td>Shi Z.</td>
<td>P141</td>
<td>Wolodt schenko A.</td>
<td>P142</td>
</tr>
<tr>
<td>Liu X.</td>
<td>13</td>
<td>Silapacharanan S.</td>
<td>33</td>
<td>Wood M.</td>
<td>11</td>
</tr>
<tr>
<td>Liu Y.</td>
<td>94</td>
<td>Smith R.M.</td>
<td>72</td>
<td>Wu F.</td>
<td>25</td>
</tr>
<tr>
<td>Lu X.Z.</td>
<td>16</td>
<td>Smith J.Y.</td>
<td>72</td>
<td>Wu Z.</td>
<td>25</td>
</tr>
<tr>
<td>Ma Z.</td>
<td>32</td>
<td>Smith B.E.</td>
<td>117</td>
<td>Xie L.</td>
<td>P146, P162</td>
</tr>
<tr>
<td>Mackaness W.</td>
<td>41</td>
<td>Spencer G.F.</td>
<td>P140</td>
<td>Yanfang L.</td>
<td>47, P163</td>
</tr>
<tr>
<td>Malysheva N.V.</td>
<td>48</td>
<td>Spiess E.</td>
<td>38, 81, 121</td>
<td>Yao lin L.</td>
<td>P163</td>
</tr>
<tr>
<td>Marabini F.</td>
<td>P135</td>
<td>Stappenbeck C.</td>
<td>32</td>
<td>Ye T.</td>
<td>34</td>
</tr>
<tr>
<td>Martynenko A.</td>
<td>P132</td>
<td>van der Steen S.</td>
<td>57</td>
<td>Yoeli P.</td>
<td>116</td>
</tr>
<tr>
<td>Maruyama H.</td>
<td>P152</td>
<td>Steinberg J.</td>
<td>44</td>
<td>Yoshinari T.</td>
<td>P152</td>
</tr>
<tr>
<td>Matsuzawa T.</td>
<td>P133</td>
<td>Stengele R.</td>
<td>60</td>
<td>Zetsche J.</td>
<td>84</td>
</tr>
<tr>
<td>McCrorie I.</td>
<td>11</td>
<td>Stringer P.</td>
<td>55</td>
<td>Zhang W.</td>
<td>54, P130</td>
</tr>
<tr>
<td>Mc Guinness C.</td>
<td>55</td>
<td>Su B.</td>
<td>54</td>
<td>Zhang X.</td>
<td>54</td>
</tr>
<tr>
<td>Mél lançon M.</td>
<td>73</td>
<td>Sweet-Escott J.</td>
<td>63</td>
<td>Zhang Y.</td>
<td>54</td>
</tr>
<tr>
<td>Meng L.</td>
<td>21</td>
<td>Syrén M.</td>
<td>93</td>
<td>Zhang Q.</td>
<td>82</td>
</tr>
<tr>
<td>Mezösi G.</td>
<td>P154</td>
<td>Tainz P.</td>
<td>89</td>
<td>Zhaozhong X.</td>
<td>P153</td>
</tr>
<tr>
<td>Midtbo T.</td>
<td>42</td>
<td>Takazawa S.</td>
<td>P151</td>
<td>Zhen T.</td>
<td>10</td>
</tr>
<tr>
<td>Moeller ing H.</td>
<td>77</td>
<td>Tanaka T.</td>
<td>P151</td>
<td>Zhou S.</td>
<td>13</td>
</tr>
<tr>
<td>Monmonier M.</td>
<td>74</td>
<td>Thomassin M.M.</td>
<td>P159</td>
<td>Zhu X.</td>
<td>54</td>
</tr>
<tr>
<td>Morhard H.</td>
<td>119</td>
<td>Thothong W.</td>
<td>35</td>
<td>Zou Y.J.</td>
<td>99</td>
</tr>
<tr>
<td>Morita T.</td>
<td>2, 127</td>
<td>Tijssen T.</td>
<td>111</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morrison J.L.</td>
<td>4</td>
<td>Tiku nov V.S.</td>
<td>86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mue si L.</td>
<td>P154</td>
<td>Török Z.</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muller J.C.</td>
<td>20</td>
<td>Tra inor T.</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Musei bov A.M.</td>
<td>P158</td>
<td>Tsuzawa M.</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nabiyyev A.A.</td>
<td>P160</td>
<td>Turk A.</td>
<td>53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakano Y.</td>
<td>P151</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nickless E.F.P.</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*) Articles printed in italics are to be found in Proceeding II
Generationsfähige Produkte. Fachkompetenter Dialog.

Das SAFIR Programm von RENKER umfasst Materialien und Systemlösungen für Ihren speziellen Bedarf in Technik, Organisation und Verwaltung.

- Spezialmaterialien für die Landkartentechnik
- Reprografie-Materialien vom Zeichnungsentwurf bis zur Vervielfältigung auf Papier und Folie
- Systeme für Lichtpaus- und Kopiertechnik und für Telekommunikation

RENKER - Ihr kompetenter Partner - immer in Ihrer Nähe.
Vertriebsniederlassungen/Auslieferungslager in:

- Augsburg
- Berlin
- Bielefeld
- Chemnitz
- Essen
- Frankfurt
- Freiburg
- Hamburg
- Köln
- Mannheim
- München
- Nürnberg
- Stuttgart

RENKER GmbH & Co. KG
Postfach 101746
Kreuzauer Straße 33
D-5160 Düren
Telefon 02421/597-0
Telefax 02421/5974 52
Telex 833834

RENKER GmbH & Co. KG
Grüßeberstrasse 146
D-1000 Berlin 48
Telefon 030/740 04-0
Telefax 030/740 04 64
Telex 183027