ASSESSMENT AND IMPROVEMENT OF METHODS FOR ANALYTICAL HILLSHADING

Karl Lukas and Robert Weibel

Department of Geography
University of Zurich
Winterthurerstrasse 190
8057 Zurich (Switzerland)
Fax: +41-1-362 52 27
{kalle, weibel}@gis.geogr.unizh.ch

Abstract

The study reported in this paper had two major objectives. In a first step, methods published in the literature were implemented to form the basis for a comparative assessment. Secondly, based on the previous evaluation, existing methods were improved further. Four basic classes of methods of increasing algorithmic complexity were implemented: 1) analytical hillshading with a single constant light source; 2) multiple constant light sources; 3) multiple light sources, weighted according to a locally varying aspect map; and 4) local variation of a single light source according to a map of ridges and valleys. The assessment consisted of systematic experiments, varying the relevant control parameters. Based on a quantitative and visual assessment, it was concluded that the first and third classes of techniques — extended by the enhancements described here — were superior to the other two. The best potential, however, is expected to lie in the combination of different methods in order to add the strengths and alleviate the deficiencies of the individual methods. Composites of classes 1 and 3 seem to be particularly useful (class 1 emphasizes large forms, while class 3 highlights details).

1 Introduction

Hillshading is widely used as a method of terrain visualization in topographic as well as thematic cartography. The automation of the manual hillshading technique, commonly known as analytical hillshading, has been one of the first tasks addressed by research in digital cartography, with first experiments dating back to the 1960s [10]. Today, most digital cartography systems and geographic information systems (GIS) offer functions for analytical hillshading. The overwhelming majority of these systems, however, use a very simple shading technique based on the Lambertian illumination model of diffuse reflection [4], with a single light source at a constant direction. The problem with using such a simple model is that ridges and valleys extending in the direction of the constant light source remain 'expressionless'. That is, because the shading values obtained from the model are the same on both sides of the landforms, even prominent forms are not emphasized well enough and disappear. Conversely, landforms whose general direction is perpendicular to the lighting direction are overemphasized.

In contrast to analytical hillshading, in manual work the cartographer usually locally adjusts the horizontal as well as vertical angle of the light source to better 'shape' the topographic relief [5]. Some researchers have therefore attempted to develop better techniques for analytical hillshading working with multiple light sources and/or a local adjustment of the angle of the incident light [2, 7], but require further improvement before they can be used operationally. Manual hillshading, therefore, is still the preferred method for most maps produced today, which is mainly due to the insufficient cartographic quality of present analytical hillshading. This is true despite of the fact that digital terrain models (DTMs), which form a prerequisite of analytical hillshading, are now available in sufficient quality in many countries. Once DTMs exist, analytical hillshading becomes an inexpensive method of producing an expressive relief display, for purposes ranging from topographic and thematic mapping to geological or geomorphological interpretation. Analytical techniques should therefore be improved further and brought closer to a practical use, so that they become ready to be implemented in commercial systems.

The study reported here had two major objectives [6]. Firstly, methods published in the literature should be implemented to form the basis for a comparative assessment. Secondly, based on the previous evaluation, existing methods should be further improved, and possibly combined into model composites. The major target application was the production of cartographic relief shadings for the purpose of printed topographic or thematic maps. Some of the observations made below may thus not apply to other applications of hillshading (e.g., geomorphological interpretation). The following sections highlight selected issues addressed in the study (for more details, see [6]): the basic hillshading methods implemented for the tests; some of the extensions made; and conclusions which can be drawn.

2 Review of Methods for Analytical Hillshading

2.1 *Overview, implementation, illumination models*

Overview of hillshading methods: Starting off from the early theoretical work of [9] and the first experiments in automated hillshading by Yoeli [10], a variety of methods have been developed and published in the literature. These techniques can be assigned to four basic classes of increasing algorithmic complexity: 1) analytical hillshading with a single constant light source; 2) multiple constant light sources; 3) multiple light sources, weighted according to a locally varying aspect map; and 4) local variation of a single light source according to a map of ridges and valleys.

Implementation: Representatives of each of these classes of hillshading methods as well as variants thereof have been implemented and integrated into the Geographical Resources Analysis Support System (GRASS) developed by the US Army Construction Engineering Research Laboratory. GRASS is a public domain GIS which is (mainly) raster-based, a property which was useful given the fact that this study used grid DTMs. For user-friendly operation, the individual programs developed here have been integrated into the graphical user interface of GRASS — XGRASS. Further implementation details are given in [6].

Illumination models: The Lambertian illumination model is used for all hillshading methods in cartographic applications. This model assumes diffuse reflection of incident light from the surface of a physical object. The intensity of the diffusely reflected light is then proportional to the cosine of the angle between the surface normal N and the light vector S (for a review of illumination models, see [4]). Other illumination models, such as Phong shading which adds specular reflection, or Gouraud shading which allows to interpolate shading values are frequently used in domains such as computer aided design (CAD), yet, they do not meet the expectations of cartographers who usually view terrain relief as an opaque, diffuse reflector. Thus, the simple Lambertian illumination model serves the purpose of hillshading well, while more sophisticated illumination models seem only useful for special effects.

2.2 *Single constant light source*

Most cartographic systems and GIS which offer hillshading capabilities use this simple model which represents a direct implementation of the Lambertian illumination model. A single light source is available which is positioned at a constant direction; it can normally be specified to be located at an arbitrary horizontal and vertical angle. For cartographic purposes, however, the light source is usually at azimuths ranging between 270° (W) and 360° (N), with a standard azimuth of 315° (NW). As is well known, azimuths at a southerly position will cause an effect called 'relief inversion' with most individuals [5]. The vertical light angle is usually assumed at around 40-45°.

The intensity values I obtained from the Lambertian model will range between 1 for DTM facets directly facing the light source, and −1 for DTM facets which are facing away from the light. Therefore, assuming a digitizing resolution of the intensity values of 8 bits [0, 255], the initial intensities must be transformed so that the full contrast range can be exploited:

```
intensity_8bits = (intensity + 1) * 127.5;
```

Figure 1 shows an example of this simple hillshading technique. Because of its widespread use in commercial systems, it serves as a reference model.

2.3 *Multiple constant light sources*

As was explained in section 1, hillshading models with a single light source at a constant angle commonly cause problems, as they do not allow to 'shape' the relief as a cartographer would do. A possible extension that immediately comes to mind is the introduction of additional light sources in order to achieve a more differentiated picture. Intensities are computed for multiple (but constant) light sources, and the final intensity values obtained by weighted averaging. Figure 2 shows an example in which three light sources were used: the main source is at 315° (NW), with a weight of 2; the other sources are 75° apart from the main light source (30° and 240°, respectively), and are weighted with 1. The vertical angle of all light sources is 45°.

While this method improves the illumination of parts that lie in the 'shadow' of the DTM, it does not really solve the problem of 'expressionless' relief forms.

Moreover, as a consequence of weighted averaging, the contrast is reduced quite seriously.

2.4 *Multiple light sources, weighted according to an aspect map (Mark shading)*

For the example of the island of Hawaii, Mark [7] has presented a method which extends the previous model of multiple light sources by the use of a more sophisticated weighting technique. In Mark's method, four individual hillshadings are first computed using the Lambertian model, at azimuths 225°, 270°, 315°, and 360°, each at a vertical angle of 30°. The individual shading intensities are weighted according to a grid of weighting factors obtained from an aspect grid. In the original method described in [7] this aspect map is simplified by coarsening the resolution of the DTM to 1000 m prior to computing the aspect values; additionally, the aspect map is smoothed by three passes of a 3x3 moving average. The grids of weighting factors for the four lighting directions are then obtained as follows:

$$W(\theta_i) = \sin^2(\mathit{aspect} - \theta_i); \qquad \text{where } \theta_i = 225°, 270°, 315°, 360° \qquad (1)$$

Figure 3 shows an example of the Mark shading method applied to alpine topography. This technique appears to have a good potential for portraying the fine structure of the relief, while larger forms are less prominently depicted. Yet, some artifacts due to problems of the computation of the aspect map are visible (but see 3.3 for an enhancement).

2.5 *Local variation of a single light source according to a map of topographic structure lines*

The model proposed by Brassel [2] is probably the most complex one developed to date. Instead of introducing multiple weighted light sources at constant angles, a single light source is locally varied according to a map of topographic structures (particularly ridges and drainage channels). Brassel's method represents an attempt at translating directly the principles of manual hillshading into an automated procedure. In a first step, a rasterized map of the major topographic structures is produced. In the original article, this map was obtained by digitizing structures from a map, subsequently rasterizing them. Conversely, if procedures are available for automated extraction of ridges and drainage channels (e.g., [8]), this step can be automated. A grid giving the azimuth of the light source is then computed. While cells lying on a topographic structure line are assigned the azimuth of the structure line in that cell, directional values are interpolated for the remaining cells. During the application of the Lambertian illumination model, the azimuth of the light source for a particular DTM point is then taken from the corresponding point in the azimuth grid. Thus, the light source is locally varied depending on relief structure.

While this method was implemented for this project, we have not yet been able to run extensive experiments with it. In order to be capable of fully exploiting the potential of this model, it is necessary to carefully classify the individual structure lines according to their relative importance in shaping the relief (cf. [2]). As this step is rather time-consuming, we have not yet accomplished it. Although some minor extensions to the original model were made, it is not yet possible to reach a definitive judgment. It should still be noted, however, that some problems were observed. For instance, because the azimuth of the light runs in the direction of structure lines in cells which are coincident with such linear features, the problem of 'expressionless' ridges is still not solved satisfactorily. In general, however, the model does appear to carry potential for further enhancement.

2.6 *Additional effects*

In addition to the four basic hillshading models described above, further effects can be introduced. In the 'Swiss manner' of manual hillshading, an technique called the *aerial perspective* is commonly applied [5]. This procedure simulates the effect of haze lying over lower regions. Brassel [2] has proposed a model to automate this effect which is simple to implement and works well. The model consists of a non-linear contrast modification as a function of altitude: in lower regions, the contrast is reduced, while it is enhanced in higher areas (for details, see [2]). Since the technique only modifies the intensity values, it can be applied as a postprocessing step to any of the described hillshading models.

Color shading is a further option which can be generated. If three hillshadings are computed for different light azimuths, they can be thought to represent the intensities of the three additive colors red, green, and blue (RGB), and combined into a color composite. This technique, just as similar methods such as slope-aspect maps [3], portrays the fine structure of the relief well and seems thus suited for structural interpretation. On the other hand, because of the significant variations in color, such methods do not allow to overlay further information and are thus of limited use in cartography.

3 Extensions and Enhancements

Various extensions and enhancements of the basic methods outlined above are possible. For lack of space, we are restricting the discussion to selected modifications, dropping, for instance, the description of alterations implemented for the method by Brassel [2] (cf. 2.5).

3.1 *Global modifications: Optimization of contrast and resolution*

Some possible enhancements relate to *all* hillshading techniques. First of all, contrast of analytical hillshadings is usually too low to be output to paper or film as calculated (this is different for CRTs, since the possible contrast range is usually greater due to the use of additive colors). *Contrast enhancement* is therefore needed following the computation of the analytical hillshading. This postprocessing operation is best performed interactively in an image processing system; many current cartographic systems and GIS do not offer this function. We found that a sinusoidal (s-shaped) contrast curve worked best because it reduces value in darker areas while it increases it in lighter regions.

Additionally, the resolution of the hillshading has to be optimized in two ways. Firstly, the *resolution of the DTM* should be such that individual DTM pixels are invisible. For instance, given a grid spacing of 25 m, the pixel size on a scale of 1:25,000 would be 1 mm, which is still easily visible. Thus, the DTM needs to be resampled (e.g., by bilinear interpolation) to a resolution which is finer than can be resolved by the human eye. Additionally, the *resolution of the halftone screen* used for plotting needs to be optimized depending on the output medium and the reproduction technique used.

3.2 *Improvement of Mark shading*

As was mentioned in section 2.4, the original version of the algorithm by Mark [7] exhibits problems in alpine relief which typically consists of diverse (e.g., large and small) relief forms extending in different directions (Figure 3). The weaknesses observed do not stem from the hillshading method *per se*, but from the approach used to filter the aspect map. A different technique for obtaining smoothed aspect values was thus developed. Following the resampling of the DTM to a user-specified resolution, aspect values are calculated and the aspect map (not the DTM) is then smoothed. Smoothing aspect values presents particular problems, which are taken care of in this method (e.g., the average of 10° and 350° is not 180°, but 0°). The method smooths aspects values in a 3x3 moving window and proceeds as follows: 1) sort the nine values in the window in ascending order and add 360 to all values which have a difference to the greatest value of ≥ 180 (this takes care of the gap 0° to 360°); 2) find lowest and highest value in the new data range; 3) if the difference is smaller than or equal to a parameter *threshold*, compute average of nine cells (otherwise, take old value of center point of window); 4) translate values back to the valid data range [0, 360]. For *threshold*, a value of 120 showed best results.

Figure 4 shows a result of this improved version of Mark shading. It should be noted that the output of this procedure depends largely on the degree of smoothing of the aspect map. In general, it holds that the more the aspect map has been smoothed, the more contrast the hillshading exhibits. Therefore, it does even make sense to smooth an aspect map up to 20 times or more in some cases.

3.3 *Contrast enhancement by adaptive vertical exaggeration*

Besides the global contrast enhancements described in section 3.1, further modifications to contrast variations are necessary on the local level, particularly in areas of gentle to moderate slope. Figure 5 shows such a region which is part of a DTM which includes higher areas with steep slopes as well. The contrast range of the hillshading is so low that smaller forms are hardly visible. This deficiency can be overcome to some extent by global (i.e., constant) vertical exaggeration of the DTM, as depicted in Figure 6. With such a rigid global scheme, however, noise is added to the hillshading due to the fact that even minor forms appear excessively enlarged in the vertical dimension. A better approach would thus have to be more locally adaptive to variations in slope. The method developed in this project modifies the vertical exaggeration factor for each DTM point using the following formula:

$$zfactor_{temp} = zfactor - magnitude \cdot \left(zfactor - \frac{slope \cdot zfactor}{midslope} \right) \qquad (2)$$

where *zfactor* is the user-defined vertical exaggeration factor; *magnitude* represents a user-defined scaling factor which influences the intensity of the vertical exaggeration effect; *slope* is the value for the slope (gradient) in the DTM point; and *midslope* is the largest value of slope present for the DTM, divided by 2. Thus, if *slope* is lower than *midslope*, vertical exaggeration is reduced; conversely, if *slope* is greater than *midslope*, it is increased. In our experiments, the values used for *zfactor* were up to 5, while *magnitude* was usually set to 1 (i.e., neutral). In order to suppress noise and resolution-dependent effects, it is recommended to obtain the values for *slope* from a smoothed slope map.

3.4. Combination of individual hillshadings into a composite

As is quite obvious, no single hillshading method is superior to all others for all purposes that may be conceived. Therefore, it makes sense to explore the possibility of combining individual shadings into a composite. In our experiments, we used a simple and global technique: for each pixel, shading values of two or more hillshadings were combined by computing a weighted average of the input values. The results look promising, with best results obtained by a combination of single source hillshading (which was adaptively vertically exaggerated) and Mark shading. Figures 8 and 9 show two sample results using this technique, with weighting factors of 3 (for single source shading) and 1 (Mark shading), respectively. The single source shading emphasizes larger relief forms, while Mark shading brings out the fine structure.

5 Conclusions

The main purpose of the study reported here was to implement the known techniques for analytical hillshading in order to assess their quality and characteristics, and propose further enhancements. These extensions are rather practical ones which could easily be implemented in commercial systems. In [6], a quantitative analysis — by means of descriptive statistics (moments, histograms) — as well as a visual assessment of the various hillshading methods was performed. It is thus possible to draw some conclusions, and make some recommendations about the usability of different techniques. First of all, easy-to-use facilities for optimization of hillshading parameters, the resolution of the DTM, screening, as well as for contrast enhancement should be made available in cartographic systems and GIS. The improved Mark shading method is useful for depicting the fine structure of the relief. Adaptive vertical exaggeration significantly improves the contrast variation in the hillshading, leading to enhanced portrayal of areas of gentle to moderate slope. Finally, the combination of different hillshadings into composites has great promise. In our study, the best results were obtained by a combination of adaptively exaggerated Lambert shading with a single light source and Mark shading, with aerial perspective added subsequently.

Future research should address various issues. Firstly, more adaptive methods for the combination of hillshadings should be developed (in contrast to the simple global approach used here). Secondly, we still feel that Brassel's method [2] has considerable potential. In this study we did not have the time to more fully exploit it, however. Thirdly, interactive imbedding should be improved. For instance, a user should be able to interactively designate regions which should be shaded using a different method and/or parameter settings, with blending functions applied at region boundaries. Likewise, interactive retouching of the hillshading result should be possible. An initial set of facilities of this kind has recently been developed [1]. Finally, hillshading is intimately related to generalization [8, 1], a problem (i.e., the lack of generalization) which becomes all too obvious in the illustrations shown here. A comprehensive solution of analytical hillshading also requires solving aspects of relief generalization.

References

[1] Bär, H.R. (1995): *Interaktive Bearbeitung von Geländeoberflächen — Konzepte, Methoden, Versuche.* Ph.D. Dissertation, Department of Geography, University of Zurich, 113 pgs.

[2] Brassel, K. (1974): A Model for Automatic Hillshading. *The American Cartographer,* 1(1): 15-27.

[3] Brewer, C.A., and Marlow, K.A. (1993): Color Representation of Aspect *and* Slope Simultaneously. Proceedings Auto-Carto 11, Minneapolis, MN, 328-337.

[4] Foley, J.D., van Dam, A., Feiner, S.K. and Hughes, J.F. (1992): *Computer Graphics: Principles and Practice.* Second Edition, Revised Fifth Printing. Reading, MA: Addison-Wesley.

[5] Imhof, E. (1982): *Cartographic Relief Presentation.* Berlin: de Gruyter.

[6] Lukas, K. (1994): *Evaluation und Weiterentwicklung von Methoden der computergestützten Schräglichtschattierung.* M.Sc. Thesis, Dept. of Geography, University of Zurich, 127 pgs.

[7] Mark, R.K. [1992]: A Multidirectional, Oblique-Weighted, Shaded Relief of the Island of Hawaii. US Geological Survey *Open File Report,* **92-422**.

[8] Weibel, R. (1992): Models and Experiments for Adaptive Computer-Assisted Terrain Generalization. *Cartography and Geographic Information Systems,* **19**(3): 133-153.

[9] Wiechel, H. (1878): Theorie und Darstellung der Beleuchtung von nicht gesetzmässig gebildeten Flächen mit Rücksicht auf die Bergzeichnung. *Civilingenieur,* **24**.

[10] Yoeli, P. (1967): The Mechanisation of Analytical Hillshading. *The Cartographic Journal,* 4(2): 82-8.

Figures

All figures show parts of a DTM in the Simplon region (Canton of Valais, Switzerland), taken from the DHM25 model (© Swiss Federal Office of Topography, used with permission (D-2289)). The original DTM has a grid spacing of 25 m which was resampled by bilinear interpolation to a resolution of 8 m for smooth hillshading. The screening resolution is 50 lines/cm for all figures. Three different DTM sections are shown here, each of which represents an area of 5.8 by 4 km.

The original DTM was obtained by vectorization of contour lines of the 1:25,000 scale topographic map of Switzerland. Note that although scale reduction is considerable here (for lack of space), *no generalization* was performed (which is certainly a drawback).

Note that each figure caption also indicates the section which explains the associated technique.

Fig. 1: Single constant light source (cf. 2.2).

Fig. 2: Three constant light sources (cf. 2.3).

Fig. 3: Mark shading, original method (cf. 2.4).

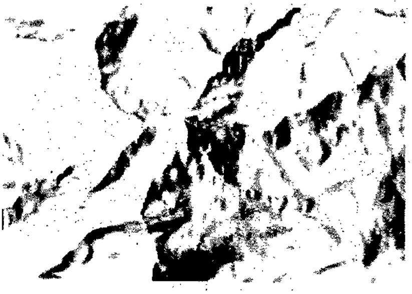

Fig. 4: Mark shading, improved aspect map (cf. 3.2).

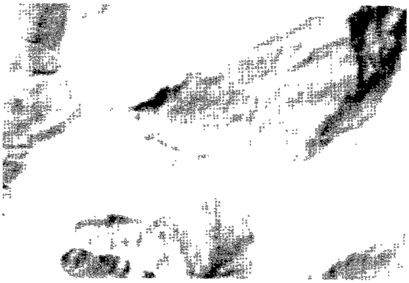

Fig. 5: Single light source, no vertical exaggeration (cf. 3.3).

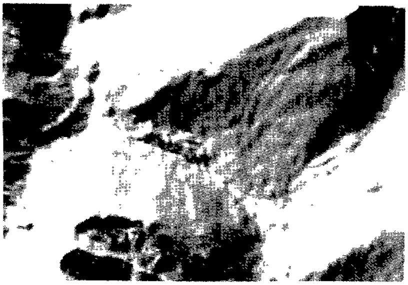

Fig. 6: Single light source, global vertical exaggeration by a factor of 5 (cf. 3.3).

Fig. 7: Single light source, with adaptive vertical exaggeration (cf. 3.3).

Fig. 8: Combination of single source shading (with adaptive vertical exaggeration) and Mark shading. Example 1 (cf. 3.4).

Fig. 9: Combination of single source shading (with adaptive vertical exaggeration) and Mark shading. Example 2 (cf. 3.4).