A Methodology to Adjust Digital Cadastral Areas in GIS

Tong Xiaohua[1,2] Liu Dajie[1] Gong Jianya[2]

1 Department of Surveying and Geo-informatics, Tongji University, 1239 Siping Road, Shanghai, China, 200092

E-mail: xhtong@celiang.tongji.edu.cn

2 National Lab. for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan, China, 430079

Abstract

In this paper, the error processing in area of the digitized cadastral parcel is discussed. The principles for adjusting the digitized parcel areas are first presented. The adjustment models are then derived, including the condition equations for areas, areas with arcs, rectangular angles and circular arcs. The methodologies to process multi-areas are further presented, including adjustment models for a single and independent parcel area, the parcel areas with “holes”, the multi-areas that are correlated with each other, the multi-areas with fixed vertexes and fixed parcels and the graded adjustment method to solve multi-areas with the larger data volume. Based on these, an area adjustment system for digitial cadastral parcels is developed. The implementation of the models and methodologies are illustrated through case studies. The results are further discussed and analyzed, leading to conclusions that adjustment processing for digital cadastral areas is helpful to ensure the quality of the data in GIS data capturing and database establishment.

Keywords: Error processing; Adjustment model; Digital cadastral area; Data quality

1. Introduction

Error analysis and processing for spatial data is one of the key issues in GIS research. In the land and housing fundamental geographic information system, the cadastral parcel is one of the most important objects. According to the feature classifications in GIS, a cadastral parcel belongs to one kind of closed polygon objects composed of series of digitized vertexes. The area of a parcel is the key attribute with legal authorization. However, in cadastral parcel digitization for capturing data, it is unavoidable to have errors (including systematic error and random error). As a result, with the propagation of errors in the vertexes of a parcel, the digitized cadastral area is not equal to the authorized area (true area) usually calculated by higher accuracy surveying method. Therefore, it is one of the focus problems to minimize the effects of the digitized errors to upgrade the accuracy of digitized vertexes to ensure the precision of the area attribute in GIS database.

Chrisman and Yandell (1988) derived a statistical model for the variance in area based on the assumption that the nodes of the polygon have errors and discussed the correlation of the nodes on area error calculation. However, how to process the area error was not further investigated. Najeh and Burkhard (1995) presented a methodology to create a digital cadastral overlay through upgrading digitized cadastral data. Merrit and Masters (1999) explored the the parametric adjustment approach implemented in the Spatial Adjustment Engine (SAE) to improve the accuracy of the digital cadastral databases. Tong and Liu (1998) further derived the conditional adjustment models for digitized data and the iterative calculation method. However, more efforts should be investigated into the detailed discussion of the error processing on the area.

The digitized data of a parcel can be treated as observations that are the coordinates in the ground system obtained from the digitized coordinates in a digitizer or scanner by orthogonal or affine transformation. In a parcel, the known authorized area, the rectangular angles and circular arcs constitute the constraints of the digitized vertexes. For more correlated parcels, the constraints also become more. As a result, the redundant observations and adjustment problems are put forward.

In this paper, the error processing in area of the digitized cadastral parcel will be discussed, including the principles and mathematical models for adjusting the digitized parcel areas. The methodologies to process the error in parcel area will be then presented according to the different kinds of circumstances. The implementation of the models and methodologies will be illustrated through case studies. The results will be further discussed and analyzed, leading to conclusions.

2. Mathematical models for area adjustment

2.1 The conditional equation for a polygon area

In cadastral map digitization, a parcel is treated as a closed polygon whose vertexes are digitized and the coordinates of the vertexes are obtained. Therefore, the coordinates of these vertexes are regarded as digitized observations. Due to the error of these coordinates in digitization, the parcel area calculated by these vertexes is not equal to the true value of the area. So, how to process the error in area to maintain the calculated area the same value as the true area is a key problem. Since the observations are the vertex coordinates of a parcel, and the true area, rectangular angles and arcs in the parcel constitute the constraints of the vertexes of the parcel, it is pertinent to adopt condition adjustment model. Assuming $(x_i, y_i)\ (i = 1,2,\cdots,n)$ are the digitized coordinate observations of the vertexes in a parcel, and corresponding adjustment values and corrective values are $(\hat{x}_i, \hat{y}_i)$ and (v_{x_i}, v_{y_i}), and

$$\begin{aligned} \hat{x}_i &= x_i + v_{x_i} \\ \hat{y}_i &= y_i + v_{y_i} \qquad (i = 1,2\cdots,n) \end{aligned} \tag{1}$$

Taking the parcel polygon in Figure 1 as example, the vertexes are P_i $(i = 1,2,\ldots n)$. Assuming the true area of the parcel is known as S_0, therefore, we have

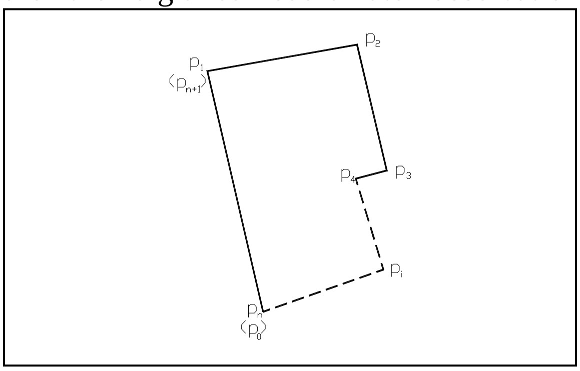

Fig.1 An Arbitrary Polygon Parcel

$$\hat{S} = \frac{1}{2}\sum_{i=1}^{n} \hat{x}_i(\hat{y}_{i+1} - \hat{y}_{i-1}) = S_0 \tag{2}$$

where $P_0 = P_n$; $P_{n+1} = P_1$, i.e., $x_0 = x_n$, $y_0 = y_n$; $x_{n+1} = x_1$, $y_{n+1} = y_1$

Therefore, the linearized form of the above conditional equation is

$$\sum_{i=1}^{n} a_i v_{x_i} + \sum_{i=1}^{n} b_i v_{y_i} + \omega = 0 \tag{3}$$

Where

$$\boldsymbol{a_i} = \frac{1}{2}(\boldsymbol{y_{i+1}} - \boldsymbol{y_{i-1}}); \quad \boldsymbol{b_i} = \frac{1}{2}(\boldsymbol{x_{i+1}} - \boldsymbol{x_{i-1}}); \quad \omega = S - S_0; \; S = \frac{1}{2}\sum_{i=1}^{n} x_i (y_{i+1} - y_{i-1}) \tag{4}$$

2.2 The rectangular conditional equation

From Figure 1 and 2, we can see that the parcel polygons maybe have rectangular angles, for example, vertexes p_2, p_3 and p_4、p_3, p_4 and p_5 constitute 90 angles. Therefore, these vertexes should obey the the rectangular condition to ensure the vertical constraints. Assume that an angle β is composed of three points *i, j*, and *k*. Thus, a conditional equation between these three points is

$$\hat{\alpha}_{ik} - \hat{\alpha}_{ij} = \beta \tag{5}$$

where $\hat{\alpha}_{ik}, \hat{\alpha}_{ij}$ are the estimators of the azimuths in directions *ik* and *ij*.

The linearization form of the above conditional equation is

$$a_{ik} v_{x_k} + b_{ik} v_{y_k} - (a_{ik} - a_{ij}) v_{x_i} - (b_{ik} - b_{ij}) v_{y_i} - a_{ij} v_{x_j} - b_{ij} v_{y_j} + \omega_i = 0 \tag{6}$$

where a_{ij}, b_{ij}, a_{ik}, b_{ik}, s_{ij}, s_{ik}, ω_i, α_{ik}, α_{ij} are calculated in [5]. It is easy to see that when $\beta = \pi/2$ (or $3\pi/2$), equation (6) is the vertical condition.

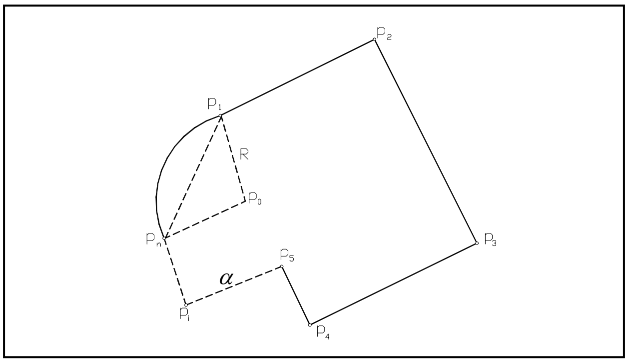

Fig.2 A Parcel Polygon with Arcs

2.3 The area conditional equation with arcs in a parcel

When a parcel polygon has arcs as shown in Figure 2, the area of the parcel is then the sum of the area of polygon P_0, P_1, …, P_n and the arc $P_nP_0P_1$. Assuming the radius of the arc is known as *R*, taking the center point $P_0 (x_0, y_0)$ of the arc as the unknown parameters, we have following area conditional equation

$$\frac{1}{2}\sum_{i=0}^{n} \hat{x}_i (\hat{y}_{i+1} - \hat{y}_{i-1}) + \frac{1}{2} R^2 \hat{\alpha} - S_0 = 0 \tag{7}$$

where $P_{-1} = P_n$; $P_{n+1} = P_0$, i.e., $x_{-1} = x_n$, $y_{-1} = y_n$; $x_{n+1} = x_0$, $y_{n+1} = y_0$; and α is the center angle of the arc and calculated by

$$\sin\frac{\hat{\alpha}}{2}=\frac{1}{2R}\sqrt{(\hat{x}_2-\hat{x}_1)^2+(\hat{y}_2-\hat{y}_1)^2} \tag{8}$$

Linearing equation (8) and Substituting into equation (7), we have

$$\sum_{\substack{i=0\\ i\neq 1,2}}^{n} a_i v_{x_i}+\sum_{\substack{i=0\\ i\neq 1,2}}^{n} b_i v_{y_i}+(a_1-c_1)v_{x_1}+(b_1+d_1)v_{y_1}+(a_2+c_1)v_{x_2}+(b_1-d_1)v_{y_2}+w=0 \tag{9}$$

where

$c_1=\dfrac{R}{2\cos\dfrac{\alpha^0}{2}}\cos\alpha_{12}$, $d_1=-\dfrac{R}{2\cos\dfrac{\alpha^0}{2}}\sin\alpha_{12}$, $w=\dfrac{1}{2}\sum_{i=0}^{n}x_i(y_{i+1}-y_{i-1})+\dfrac{1}{2}R^2\alpha^0-S_0$, α^0 is the approximate value of α .

In the meantime, there should be have at least three digitized points in an arc including its two endpoints. Therefore, we have following equations for the digitized points in an arc

$$(\hat{x}_{ri}-\hat{x}_0)^2+(\hat{y}_{ri}-\hat{y}_0)^2-R^2=0 \tag{10}$$

By linearization equation (10), we have the following conditional equation with unknown parameters

$$\Delta x_{0i}v_{x_{ri}}+\Delta y_{0i}v_{y_{ri}}-\Delta x_{0i}\delta\hat{x}_0-\Delta y_{0i}\delta\hat{y}_0+w=0 \tag{11}$$

where $w=\dfrac{1}{2}[(x_{ri}-x_0^0)^2+(y_{ri}-y_0^0)^2-R^2]$

If there are total *m* arcs in a parcel polygon, and the radius are R_i *(i = 1,2,…m)*, we have following conditional equation

$$\sum_{i=0}^{n}a_i v_{x_i}+\sum_{i=0}^{n}b_i v_{y_i}+\sum_{i=1}^{m}\pm\frac{1}{2}R_i^2\alpha_i^0\delta\hat{\alpha}_i+\omega=0 \tag{12}$$

Where the sign (±) in above equation is determined by the shapes of arcs. α_i is calculated by endpoints and radius of the arc similar to equation (8), and

$$\omega=\frac{1}{2}\sum_{i=0}^{n}x_i(y_{i+1}-y_{i-1})+\sum_{i=1}^{m}\pm\frac{R_i^2}{2}\alpha_i-S_0$$

2.4 The model considering the error in area

Since the true parcel area is also obtained by measurement, therefore, areas are a kind of observations with errors. So, it is rigorous to derive the adjustment model considering the error in area. Regarding areas and radius as observations, and assuming that the their variance and co-variance matrixes are known as $\boldsymbol{Q_S}$、$\boldsymbol{Q_R}$., then equation (4) becomes

$$\sum_{i=1}^{n}a_i v_{x_i}+\sum_{i=1}^{n}b_i v_{y_i}-v_{S_0}+\omega=0 \tag{13}$$

Equation (9) becomes

$$\sum_{\substack{i=0\\i\neq 1,2}}^{n} a_i v_{x_i} + \sum_{\substack{i=0\\i\neq 1,2}}^{n} b_i v_{y_i} + (a_1 - c_1) v_{x_1} + (b_1 + d_1) v_{y_1} + (a_2 + c_1) v_{x_2} + (b_2 - d_1) v_{y_2} + (R \cdot \alpha - e_1) v_R - v_{S_0} + w = 0 \quad (14)$$

where $e_1 = -Rtg\dfrac{\alpha}{2}$

Equation (11) becomes

$$\Delta x_{0i} v_{x_{ri}} + \Delta y_{0i} v_{y_{ri}} - R v_R - \Delta x_{0i} \delta\hat{x}_0 - \Delta y_{0i} \delta\hat{y}_0 + w = 0 \quad (15)$$

And equation (12) becomes

$$\sum_{i=0}^{n} a_i v_{x_i} + \sum_{i=0}^{n} b_i v_{y_i} + \sum_{i=1}^{m} \pm \frac{1}{2} R_i^2 \alpha_i^0 \delta\hat{\alpha}_i + \sum_{i=1}^{m} \pm R_i \alpha_i v_{R_i} - v_{S_0} + \omega = 0 \quad (16)$$

Usually the areas are adjusted in a region composed of one block or several blocks. Assuming the digitized coordinate observation is *L*, corresponding co-variance matrix is *Q*, and the corrective value is *V*, therefore, equations (4), (6), (9), (11) and (12) are represented by following generalized form

$$\boldsymbol{AV} + \boldsymbol{A_x}\delta\hat{\boldsymbol{X}} + \boldsymbol{w} = 0 \quad (17)$$

Where coefficients in $\boldsymbol{A}$, $\boldsymbol{A_x}$ and $\boldsymbol{w}$ are calculated based on equations (4), (6), (9), (11) and (12).

Thus the normal equation is

$$\begin{bmatrix} \boldsymbol{AQA^T} & \boldsymbol{A_x} \\ \boldsymbol{A_x^T} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{K} \\ \delta\hat{\boldsymbol{X}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{w} \\ 0 \end{bmatrix} = 0 \quad (18)$$

Where $\delta\hat{\boldsymbol{X}}$ is the unknown parameter matrix, $\boldsymbol{K}$ is Lagrange multiplier. And $\delta\hat{\boldsymbol{X}}$ and $\boldsymbol{K}$ are calculated by

$$\begin{bmatrix} \boldsymbol{K} \\ \delta\hat{\boldsymbol{X}} \end{bmatrix} = -\begin{bmatrix} \boldsymbol{AQA^T} & \boldsymbol{A_x} \\ \boldsymbol{A_x^T} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{w} \\ 0 \end{bmatrix} \quad (19)$$

The adjustment values of the coordinates is

$$\hat{\boldsymbol{L}} = \boldsymbol{L} + \boldsymbol{V} = \boldsymbol{L} - \boldsymbol{QA^T K} \quad (20)$$

The variance-covariance matrixes of the adjustment coordinates are

$$\begin{aligned} \boldsymbol{Q_{\hat{L}}} &= \boldsymbol{Q} - \boldsymbol{QA^T}(\boldsymbol{AQA^T})^{-1}(\boldsymbol{I} - \boldsymbol{A_x}(\boldsymbol{A_x^T}(\boldsymbol{AQA^T})^{-1}\boldsymbol{A_x})^{-1}\boldsymbol{A_x^T}(\boldsymbol{AQA^T})^{-1})\boldsymbol{AQ} \\ \boldsymbol{D_{\hat{L}}} &= \hat{\sigma}_0^2 \boldsymbol{Q_{\hat{L}}} \end{aligned} \quad (21)$$

And the estimator of the unit root mean square error is represented by

$$\hat{\sigma}_0^2 = \frac{\boldsymbol{V^T PV}}{\boldsymbol{r}} \quad (22)$$

Where $\boldsymbol{r}$ is the number of redundancy. For a block constituted by *M* parcels, The number of $\boldsymbol{r}$ is $\sum_{j=1}^{M}(1+m_j+\sum_{i=1}^{n_j}(k_i-3))$.

3. Methodologies of parcel area adjustment

In the cadastral map digitization, a parcel belongs to a block composed of many parcels that have no sleeks among them. In a block, due to the many parcels, there are many conditions, and they are interrelated with each other. So, if just adjusting the parcel area one by one not considering the correlation among the parcels, then the boundaries among the parcels maybe overlapped. As a result, the area can not be adjusted properly, either can the parcel not be sent into the GIS spatial database. Considering all kinds of area adjustment circumstances, it is necessary to use different methods to adjust areas in different cases.

1) Single parcel area adjustment. This is a simple case that a parcel is single and independent with others. For example, the conditional adjustment models are derived for the parcels shown in Figure 1 and 2 according to equation (6), (11) and (12).
2) Parcel with “holes” area adjustment. A parcel with a “hole” is the one including another parcel in it. Therefore, it has two area conditions, i.e., the area of the hole and the area containing the hole. These two parcel areas should be adjusted simultaneously.
3) Multi-parcel areas adjustment. This is the most usual case. As shown in Figure 3, the block is composed of several parcels. Under the circumstances that the data volume of a block is not very large, for example, the total numbers of the parcels are less than 50, and the total numbers of the points in the block are less than 1000, multi-parcel areas should be adjusted integrally. The conditions for each parcel in the block are firstly derived, then all parcels in the block are adjusted simultaneously. In this case, the key problem is to ensure that the shared vertexes and boundaries among interrelated parcels are moved simultaneously, therefore, the topologies between parcel polygons remain undamaged.

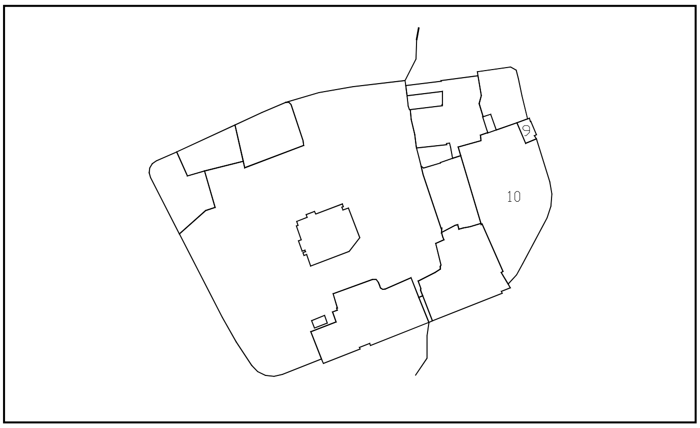

Fig.3 The All Parcels Included in One Block

4) Parcel area adjustment with fixed points and parcels. This case is suit for the condition that there are some points and parcels whose vertexes are measured with higher accuracy, thus these points and parcels remain unchanged in area adjustment. These points are called fixed points, and the parcels are fixed parcels.
5) Graded adjustment for multi-parcel areas. When the data volume of the processing parcels and blocks are very large, a graded adjustment approach for multi-parcel

6) The limit value of parcel area adjustment. In order to adjust the parcel areas properly, it is necessary to give a limit error for parcel area adjustment. If the error is less than the limit error, then the area adjustment could be carried out, otherwise, further efforts should be paid to check the problem. Assuming that the prior root mean square error of digitized coordinate is 7cm[2], based on to equation (2), we have the variance in area

$$\sigma_S^2 = \boldsymbol{A}\boldsymbol{D}_\zeta \boldsymbol{A}^T \tag{23}$$

Where $\boldsymbol{A} = [\boldsymbol{a}_1 \quad \boldsymbol{b}_1 \quad \boldsymbol{a}_2 \quad \boldsymbol{b}_2 \quad \dots \quad \boldsymbol{a}_n \quad \boldsymbol{b}_n]$; $\zeta = [\boldsymbol{x}_1 \quad \boldsymbol{y}_1 \quad \boldsymbol{x}_2 \quad \boldsymbol{y}_2 \quad \dots \quad \boldsymbol{x}_n \quad \boldsymbol{y}_n]^T$, $\boldsymbol{D}_\zeta = \boldsymbol{diag}\{0.49, 0.49, \dots, 0.49\}_{2n\times 2n}$, $2\sigma_S$ or $3\sigma_S$ is chosen as the limit error of the area.

4. Case studies and analysis

Based on the previous theoretical discussion, a case study will be presented to illustrate the implementation of the models and the methodologies, and the results will be thorough discussed and analyzed. Taking the parcels included in the block shown in Figure 3 as example, we will discuss the area adjustment for these parcels. In this block, there are total 17 parcels among which the parcel numbered 9 and 10 are fixed parcels. Here, we choose the prior root mean square error of digitized coordinate as 5cm, and $2\sigma_S$ as the limit error of the area. Table 1 shows the data for the true areas of the parcels, the calculated areas, the differences between these two areas, the relative errors of the areas and the limit errors of the areas.

Tab.1 The Area Calculation of the Parcels

True Areas (m^2)	Calculated Areas (m^2)	Differences (m^2)	Relative Error (‰)	Limit Errors (m^2)
2521.000	2516.165	-4.835	2	23
5516.000	5513.318	-2.682	0.5	38
44938.000	44945.702	7.702	0.2	115
110.000	108.962	-1.038	9	4
2857.000	2857.780	0.780	0.3	24
2748.000	2750.091	2.091	0.8	26
1873.000	1873.226	0.226	0.1	18
4144.000	4141.147	-2.853	0.7	39
2704.000	2709.042	5.042	2	26
5183.000	5182.701	-0.299	0.1	36
107.000	106.667	-0.333	3	7
579.000	582.477	3.477	6	11
483.000	483.818	0.818	2	10
2197.000	2199.293	2.293	1	21

142.000	144.048	2.048	1.5	6

The graded adjustment approach is adopted to adjust the multi-parcel areas in this case. The outer points of the block are firstly processed, and these points become fixed points, then the parcel areas in the block are adjusted. In the processing, the total numbers of the adjusted vertexes in the parcels are 115. After iterative calculations (usually 2 times), the parcel areas calculated by the digitized coordinates are equal to the corresponding true areas of the parcels, and the unit weight root mean square error is 0.081m. Table 2 shows parts of the corrective values and adjustment values of the digitized coordinates.

Tab.2 The Corrective Values and Adjustment Values of the Digitized Coordinates

No.	Observations (m)		Adjustment Values (m)		Corrective Values (m)		Root Mean Square Error (m)		
	X	Y	X	Y	ΔX	ΔY	M_x	M_y	M_p
1	-4031.721	257.794	-4031.694	257.773	0.027	-0.022	0.078	0.079	0.111
7	-4069.340	235.548	-4069.329	235.520	0.011	-0.028	0.080	0.077	0.111
8	-4096.320	246.240	-4096.297	246.207	0.023	-0.033	0.079	0.078	0.111
9	-3972.550	105.050	-3972.529	105.139	0.021	0.089	0.069	0.062	0.093
10	-3949.428	131.123	-3949.451	131.140	-0.023	0.017	0.075	0.077	0.108
26	-3887.501	340.667	-3887.568	340.697	-0.068	0.030	0.070	0.078	0.105
27	-3896.566	343.083	-3896.555	343.126	0.011	0.044	0.080	0.076	0.110
28	-3913.806	347.217	-3913.781	347.291	0.024	0.075	0.079	0.061	0.100
29	-3966.947	365.094	-3966.924	365.159	0.023	0.065	0.079	0.067	0.104
100	-3856.698	406.364	-3856.681	406.426	0.017	0.062	0.080	0.073	0.108
101	-3868.918	410.284	-3868.892	410.362	0.026	0.078	0.079	0.067	0.104
114	-3835.784	445.370	-3835.794	445.325	-0.010	-0.044	0.080	0.071	0.107
115	-3854.112	414.408	-3854.172	414.346	-0.060	-0.061	0.072	0.072	0.101

Analyzing the results of the case studies, we can conclude that

1) The differences between the true areas and the calculated areas of the parcels are always exist, and the range of the area relative errors are 0.1‰ up to 10‰. However, after parcel area adjustment, these two area values are the same. From Table 1, we also see that area limit errors of the parcels is related with the sizes of the parcels and the numbers of the vertexes in the parcels. Therefore, for the parcels with small sizes and numbers of the vertexes, their area relative errors are much larger, as a result the possibilities greater than the limit errors are higher.
2) Based on the known authorized area, the rectangular angles and circular arcs in the parcel, the conditional adjustments are derived to process the parcel area. This approach improves effectively the accuracy of the digitized vertexes of the parcel. Another important facet of the method is to maintain the logical consistency between the parcels. Meanwhile, from Table 2, we see that the range of the corrective values of the digitized coordinates of the vertexes in the parcels is between 0cm and 9cm, the positional error is about 10cm. These mean that during the range of the limit error, it is feasible to move the vertexes to adjust the areas of the parcels.
3) According to the different circumstances, different methods are adopted to adjust the area of the parcels, after this step, the other cadastral features as houses and

5. Conclusions

In this paper, the error processing in area of the digitized cadastral parcel is discussed. The principles and adjustment models are presented. The methodologies to process multi-areas are further developed. The implementation of the models and the methodologies are demonstrated through case studies. The results are further discussed and analyzed, leading to conclusions that adjustment processing for digital cadastral areas is helpful to upgrade the accuracy of the digital cadastral parcel and ensure the quality of the fundamental data in GIS data capturing and database establishment.

References

1. Chrisman, N.R. and Yandell, B., 1988, A model for the variance in area. Surveying and Mapping, 48, 241-246.
2. Liu D.J., Shi, W.Z., Tong, X.H., 1999, Accuracy analysis and quality control of spatial data in GIS. Shanghai Science and Technology Documentation Press.
3. Merrit, R. and Masters, E., 1999, Digitial Cadastral Upgrades – A progress report. International Symposium on Spatial Data Quality, HongKong, 180-188.
4. Najeh, T. and Burkhard, S., 1995, A methodology to create a digital cadastral overlay through upgrading digitized cadastral data. Surveying and Land Information System, 55(1), 3-12.
5. Tong, X.H., Liu, D.J., and Yang, D.Y., 1998, Adjustment models of GIS digitized data and development in software. Journal of Tongji University, 26(6), 682-685.

Acknowledgement

The work described in this paper was substantially supported by the Post-doctoral Funds of China, Visiting Scholar Foundation of Keb Lab. In University of China and the National Natural Science Funds of China.