

DEVELOPMENT OF HYBRID VECTORIZING SOFTWARE FOR DIGITIZATION OF CADASTRAL MAPS

Byoungjun SEO, Jaejoon JEONG, Jaebin LEE and Prof. Yongil KIM, Korea

ABSTRACT

The Cadastral map is a basic data that prescribes lot numbers, the classification of land category, and the boundaries and ownerships of land parcels. In Korea, the government has tried to digitalize cadastral maps and inputted the attribute informations of cadastral maps. But, the efficiency problems have been raised, because the cadastral services mainly depend on the manual works. Also, the figure informations of cadastral maps were not digitalized that there have been many problems in establishing the efficient land information systems.

In this study, we developed the hybrid vectorizing software to enhance the deficiency of the analogue methods and to ensure the accuracy of the cadastral maps. The hybrid vectorizing adopts a screen-digitizing method as a prototype and automates the procedure of searching the intersection of lines with efficiency. Consequently, in the aspect of the accuracy, there is no difference with the screen-digitizing method because the hybrid method is based on the screen-digitizing. In the aspect of the efficiency, we can input the neatline layers at regular intervals and the deformed neatline layers, and shorten the data input time.

1. INTRODUCTION

1.1 Background

The cadastral map is the base data that defines a lot of number, the classification of land category, and the boundaries and ownership of land parcels. Cadastral services occupy about 80% of the administrative affairs, especially at the regional administration. This service includes the location information, the utility limitation, and the resource management of parcels.

Because the cadastral services had been achieved by manual works, problems of efficiency have been raised. Also, the problems of accuracy and management have been raised, because the cadastral maps which were generated under the rule of Japanese imperialism for the first time have been in use until now. Therefore, the government has attempted to digitize cadastral maps. By the policy, the attribute information of the land and the forest land parcels has been digitized since 1982. But, there were many difficulties in constructing the efficient land information systems, because the digitization of cadastral maps which is the graphic information, is not yet completed.

Since 1978, the Administration of Seoul Metropolis and the Korea Cadastral Survey corporation introduced CGP/1000 systems of Computer Vision co. and carried out the reconstruction of cadastral maps. This procedure has the limitation of working hours for digitization by manual inputs while it guarantees the high accuracy. It takes about one

day to input a cadastral map of 1/600 scale. Therefore, we need 57 years for digitization of 720,000 sheets of cadastral maps, assuming 270 working days per one year and that the number of CGP/1000 systems are 50. Thus, we can say that the digitization of all the cadastral maps is impossible without introducing new input methodology.

To overcome the weakpoints of manual input method, screen digitizing and automated input method can be introduced. Merits of these methods are that the original data is not changed, because all the processes are applied to the scanned cadastral maps. But the screen digitizing method has no advantage of shortening the input time compared to the manual input method, and the automated vectorizing method has the different result in the number and position of vectors as in the original data. Besides, it can be influenced by the quality condition of the map. Therefore, these two methods cannot be the efficient tools in digitizing the cadastral maps. Though there were numerous attempts to induce the automated vectorizing for the digitization of the cadastral maps, the automated vectorizing methods could not be adopted because it could not satisfy the accuracy needed in the cadastral maps and these works are closely related with the property rights of the people.

1.2 The scope and methodology

In this study, considering the characteristics that the cadastral maps can't allow for the error, we developed the precise vectorizing tools, and the methodology of shortening the input time to increase the operation efficiency. Therefore, the vectorizing tools are based on the screen digitizing to insure the resultant accuracy, and we developed the hybrid type vectorizing method which automated the look-up procedure of intersection points with the line searching algorithm to increase the operation efficiency. We developed and realized the line searching algorithm which best fit for the characteristics that all lines on the cadastral maps are straight lines, which link the points determined by the actual field survey. And many functions that can provide convenience in the screen digitizing procedure with the operator were attached to the vectorizing tool.

The constitution of this paper is as followings. We raised problems of existing vectorizing systems in section 2, and described the methodology for developing the hybrid vectorizing tools in section 3. In section 4, we tested the accuracy and efficiency of this approach. Finally, the conclusion and further researches were presented.

2. PROBLEMS OF THE EXISTING VECTORIZING METHOD

The methods for the digitization of cadastral maps could be classified as three categories. These are the manual input method using the mapping table, the screen digitizing method using scanned cadastral maps on the CRT screen, and the automated input method by the use of line detection algorithms. Following sections describes the problems of each method.

2.1 The manual input method

Compared with other methods, the vector products from the manual input method using mapping digitizers guarantee the high accuracy. When reading coordinates from a digitizing table, the resultant accuracy is influenced by the personality, the degree of

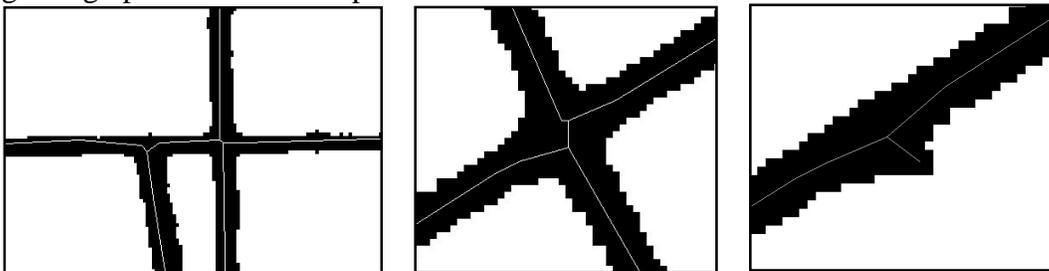
skillfulness, and the mental state of the operator. Therefore, the resultant accuracy can be varied. And there is a fault that the temperature and humidity of a operation room can affect map sheets and the reference coordinates can be changed. Besides, it need so much time for digitizing that this methods are inapplicable to the digitization of cadastral maps.

2.2 The screen digitizing method

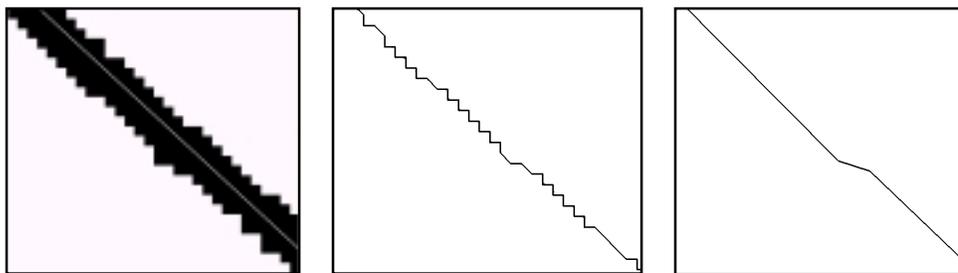
The screen digitizing method had advantages that the original data is not changed during the operations and that there are no limitation of devices. Lastly, the accuracy can be guaranteed as the manual input method. But in this method the operator has to locate the parcel boundaries with the moving pointer on a computer screen manually as in the former method. If there are extensive parcels in the scanned cadastral map or the scanned resolution is very high, the amount of works for an operator increases. Consequently, this method cannot be appropriate for the vectorization of cadastral maps.

2.3 The automated input method

The automated vectorizing with the line searching algorithm are mainly utilized in the area of the map management and developed for the efficient digitization of already existing map sheets. Digitization of map sheets which are in good state can be fully automated, and incorrect locations and inputs of vectors can be revised with the digitizing operation on a computer screen.



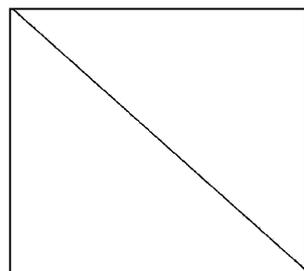
<Figure 1> The vectorizing result using the fully automatic method



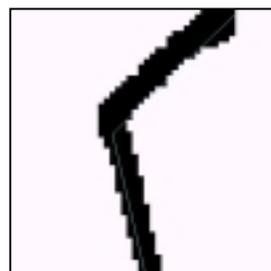
(a)

(b)

(c)



(d)



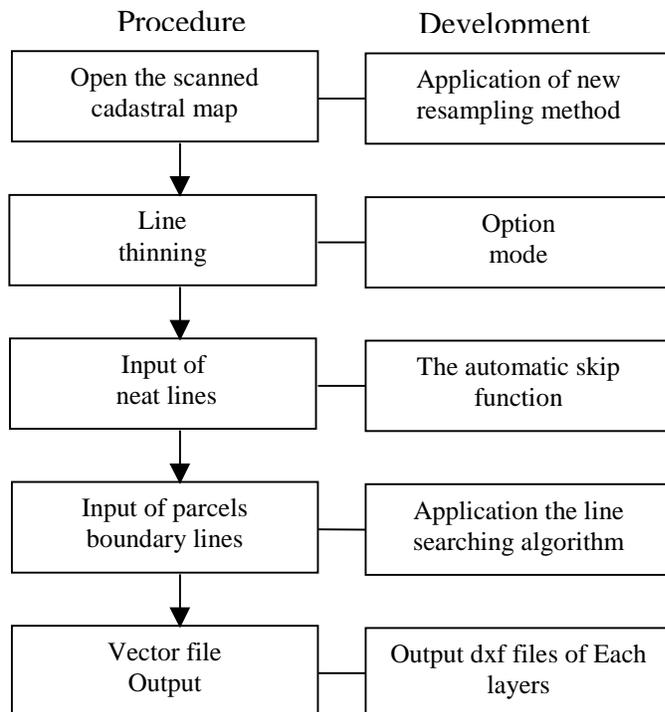
(e)

<Figure 2> The problems of the fully automatic method

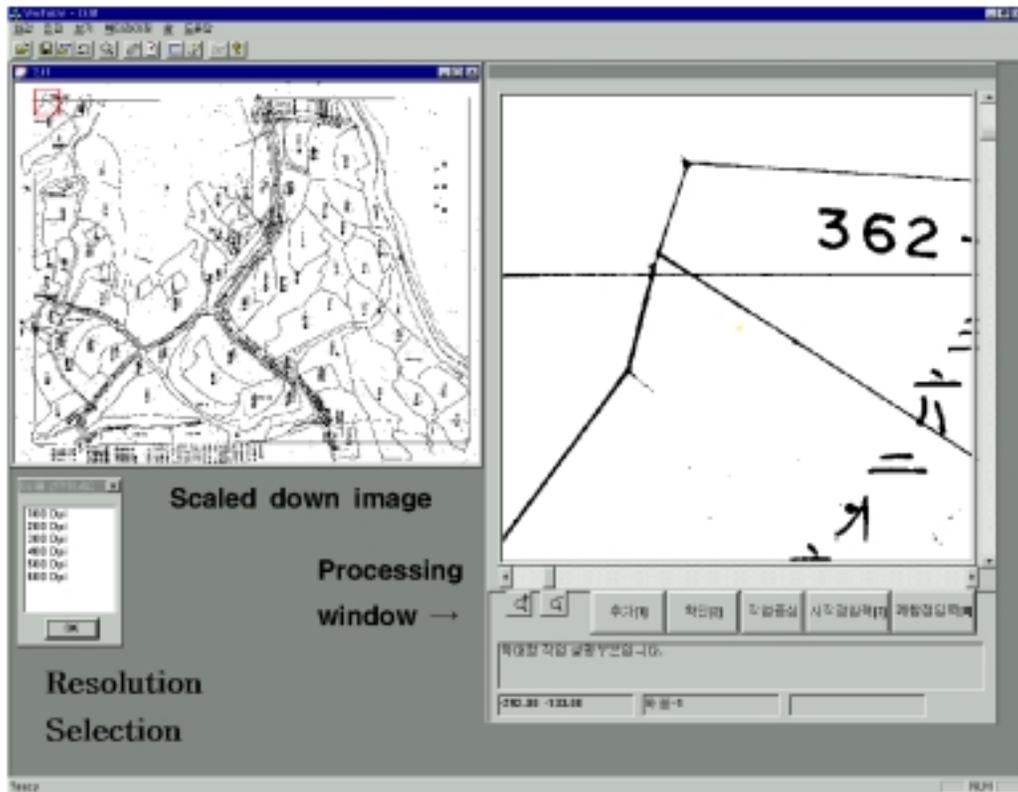
Though it can speed up the overall processes, the accuracy of products cannot be secured and the confidence level of the result from the pattern recognition techniques is low. In the addition, if the maintenance state of a map sheet is not good, automated processes cannot be continued. Generally, a line thinning procedure is performed before the line searching algorithm. These resulting vectors have focuses on a line itself, but lines on the cadastral map are just the linkage of points and have no sense of its own. Therefore, difficulties in the automated vectorizing algorithms are that it cannot locate the significant points on cadastral maps. <Fig. 1&2> shows the result from the commercial vectorizing software. From <Fig.1>, we can see so many redundant vectors, so some softwares settled these problems with the generalization algorithm(see Fig. 2). But this process usually cannot locate the position of lines correctly. In <fig. 2(a)>, the thicker line indicates a scanned line and the thinner line shows the vector that is to be digitized. <Fig. 2(b)> shows the primitive results after the execution of a commercial software and <Fig. 2(c)&2(d)> shows the final result of a generalizing procedure. As shown, all vectors in <Fig. 2(b), 2(c), &2(d)> are different with the vectors shown in <Fig. 2(a)> in their numbers and locations. Their inappropriateness was shown in <Fig. 2(e)>.

3. DEVELOPMENT OF THE HYBRID VECTORIZING METHOD

In this study, considering the vectorizing efficiency and the characteristics that the cadastral maps don't allow for the error, we developed the hybrid vectorizing tools. It was based upon the screen digitizing and automated the look-up procedure of intersection points with the line searching algorithm, otherwise the operator should seek them inconveniently(see <Fig. 3&4>). For this purpose, we developed the new line searching algorithm fitted for the vectorizing of cadastral maps and many tools were added to relieve the operator of his discomfort. The Microsoft Visual C++ was used for the software development.



<Figure 3> The procedure of the developed software



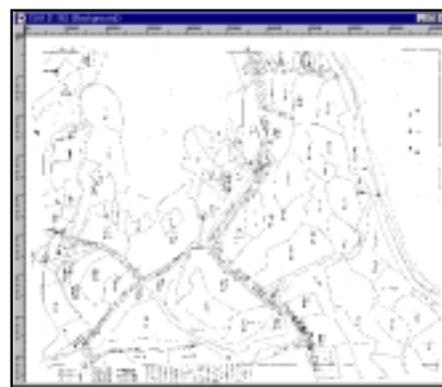
<Figure 4> The processing window

3.1 Application of the new resampling method

Scanning gives the basic data for all of the vectorizing procedures. Scanned cadastral maps can be accessed with the open menu whatever its format is. In this procedure our new resampling method is applied. When a cadastral map is scanned at the resolution of 400 dpi, the raster image is larger than 8000×7000 pixels. So, in the general image processing software, lines in the cadastral map can't be seen in a scaled-down image. In <Fig. 4>, the left is a cadastral map resampled within 600 pixels. Because it plays the role of an index map in the vectorizing procedure, we applied the new resampling method to provide an operator clear-cut lines. <Fig. 5> shows the resampled result of 11000×9000 pixels image. <Fig. 5(a)> is the result using our software, and <Fig. 5(b)> is the scaled-down image of a general image processing software.



(a) The scaled-down image using the new resampling method



(b) The scaled-down image using the general image processing software

<Figure 5> Application the new resampling method

3.2 Line thinning

Line thinning algorithms are widely used in areas such as identifying characters and a fingerprints, a circuit board manufacturing, and etc.. The purpose of it is to minimize the amount of data for the automatic recognition of patterns in the image.

In the development of a vectorizing software, the thinning process is not necessary. But the resultant image created from the thinning algorithm gives more chances for the point which determine boundaries of parcels to be located nearer than the result without it. So, this can offer the convenience to the user.

Algorithms used in this study is the parallel thinning algorithm using weighted-value which is robust to the noise and provides good quality. While other thinning algorithms have the distorted results around the intersection points, this algorithm has less distorted results around the intersection point and the broken points which is the linear features mainly used in the cadastral map.

Followings are the procedures of the thinning algorithm proposed in this study.

- Step 1. Determine the weight of each black pixel.
- Step 2. Each of the exterior pixels has a weight of K. If the eliminating condition is satisfied then, it is eliminated.(Including the case that the weight is 9)
- Step 3. Determine the weight of each black pixel.
- Step 4. The exterior pixels has each weights as K.
- Step 5. Repeat the step 3 and 4, until there are no black pixels which satisfy the eliminating condition.

3.3 Automatic line searching algorithm.

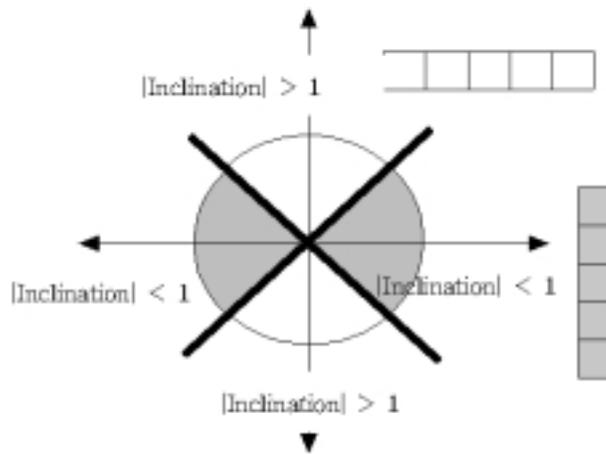
As formerly mentioned, the manual input method has the problem of efficiency and the fully automatic vectorizing method has the problem that it can't guarantee the accuracy. Because the cadastral map is related with the individual right of property, the vectorizing method using allowable errors can't be used in vectorizing. Also, because the lines in the cadastral map just connect the points, the lines has no sense of itself. Only the points i.e. the intersection points of lines is the element composing the parcels. Therefore, to minimize the allowable error we developed the hybrid vectorizing method.

Automatic line searching algorithm is as followings.

- Step 1. Input manually the broken points and the intersection points.
- Step 2. Input the indicative points(the direction of vectorizing) on the lines in cadastral maps.
- Step 3. Calculate the inclination between the intersection points and the indicative points or the broken points. And determine the shape of template to be used.
- Step 4. If the line thinning process isn't used, the center of line is controlled to coincided with the center of a template. And recalculate the inclination of the indicative direction of template.
- Step 5. Move the template along to the line using the direction determined from step

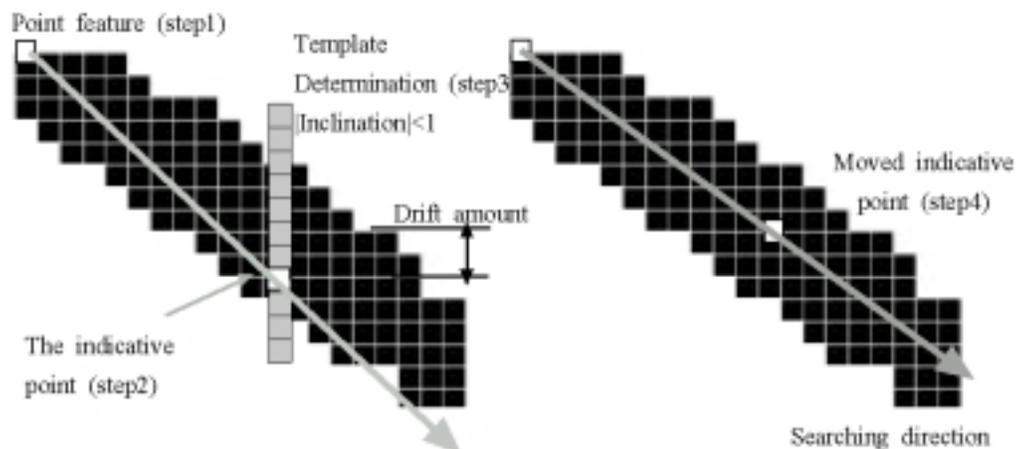
2, the shape of template from step 3, and the inclination from step 4. If the value of 0 exists at both extremities of a template, then it is considered as the broken point or the intersection point. And stop the advance of a template, and wait for the user input.

In step 3, according to the inclination between the broken point or intersection point and the indicative point the rule of determining the shape of a template is as followings. In <Fig. 6>, the center of the coordinates is the point inputted by user. If the inclination has the value between -1 and 1, then we use the vertical shaped template. If the inclination has the value more than 1 or less than -1, then we use the horizontal shaped template. The length of template is determined from the scanned resolution. The higher the resolution is (the thicker the line is), the longer the length of template is.



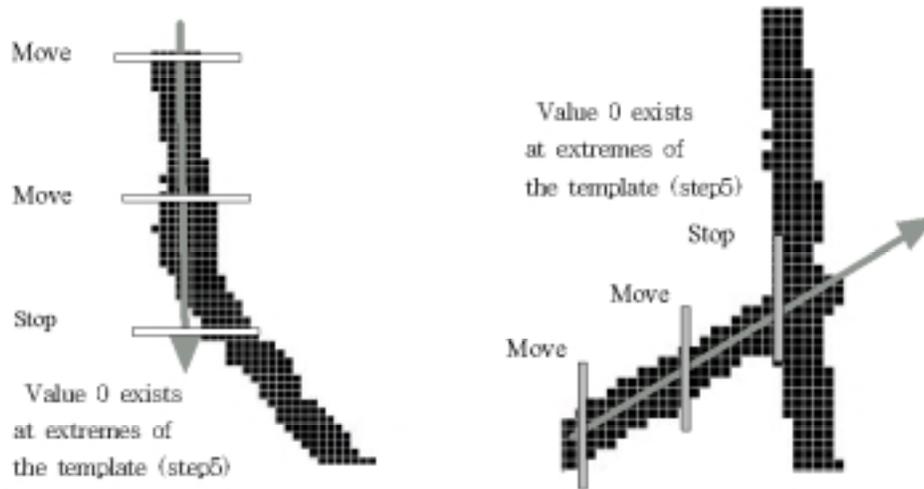
<Figure 6> Determining the template using in the automatic line searching algorithm

Step 4 is the stage that locate the template correctly on the line and recalculate the direction of searching. (see <Fig. 7>) Firstly, the template is located on the line at random. Calculate the drift between the center of the template and the center of line, and move the template by the amount of the drift. Then, the center of the template shall coincide with and the center of line. Also, the inclination between this center and the formerly inputted point is the moving direction of the template. If the line thinning is done, there is no need for step 4.



<Figure 7> The procedure of the automatic line searching

In step 5, the template moves according to the defined rules, and the broken point and the intersection point are automatically found and appear on the screen. When moving the template, it is monitored whether the value of 0 exists at both extremes in the template or not. If the value of 0 exists, then it will stop moving the template and show the message that the broken point and the intersection point are found. <Fig. 8> shows the example that the template is stopped.



<Figure 8> The procedure of determining the intersection point using the automatic line searching

The user can input the accurate point among these points. To find the more accurate point, the thinning procedure need to be performed.

The advantage of this algorithm is that it need less computing time and can detect the broken point in an acute angle, which could not be done in a fully automatic process. Moreover, vectorizing is possible in the case that the boundary of parcels are broken.

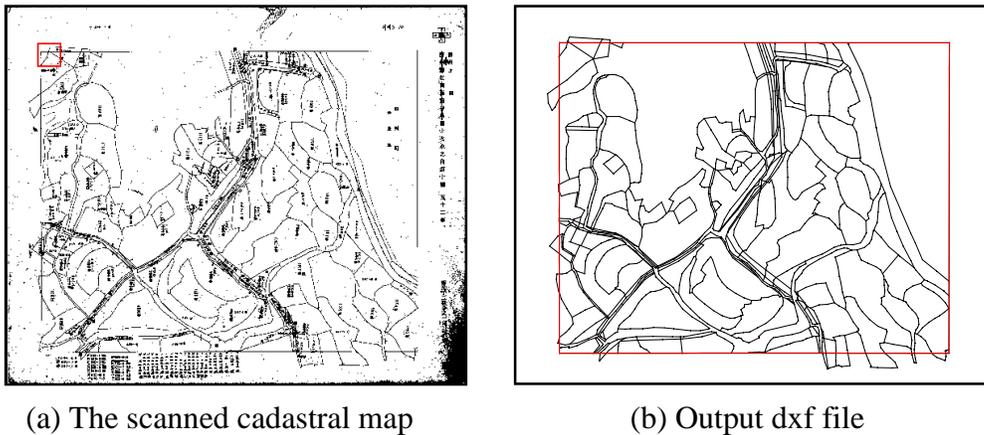
3.4 Input and saving layers

To computerize cadastral maps, 7 layers should be digitalized.(Gr: lattice point, Jm: land category, Jp: lot number, Li: boundary line, Po: matching reference point, Qu: neat line) But, layers to be digitized in the vectorizing program are the boundary line layer and the neat line layer.

The neat line layer is used to correct the distortion of the cadastral map after performing the vectorizing. The neat line of the cadastral map are composed of 4 straight lines. But, because of the distortion of the map sheet program require the user's input at the pixel corresponding to 3-4 cm. Also we used the line detection algorithm to update the vectors in the case that the map boundary line isn't rectilinear. Therefore, the number of input vector at one side of neat line is 10.

The inputted vector data is saved as a dxf file format, and it can hold the information of the layers.

4. TEST OF THE ACCURACY AND THE EFFICIENCY.



<Figure 9> The scanned cadastral map and Output dxf file

To test the accuracy and the efficiency of the developed program in this study, the vectorizing method using the developed software was compared with the screen digitizing method using the autocad program. Comparison with the fully automatic vectorizing method couldn't be performed because the number of vectors and the accuracy of the automatic vectorizing method were variable.

The result of the test is as followings. The computer system equipped the Pentium II - 233MHz CPU and 128 MB RAM was used. The map was scanned at the resolution of 400 dpi and the size of a scanned map was 8000×6700 pixels.

4.1 Comparison between these vectorizing methods

Because these methods are mainly based on the screen digitizing, the accuracy test among these methods is the matter of the operator's sincerity and skill, not the method itself. Therefore, we just performed the comparison among these methods in the aspect of efficiency.

<Table 1> shows the result of the comparison. There were no differences in inputting layers of the map boundary lines. But, using the developed software we could input the map boundary lines at the same interval, and could input the map boundary lines effectively even if the map boundary lines were distorted. And, when inputting the layers of the boundary lines, we could save time(approximately 30 minutes) using the developed software. This means that the operator's charge could be reduced. because the pointer was located within less than 1-2 pixels from the center of the parcel boundary lines.

Consequently, there is no problem in the accuracy. And we could save 30 minutes for the processing.

4.2 The accuracy of the vectorizing.

When computerizing the cadastral map, the most impotent issue is the error between the

raw data(the cadastral map) and the digital data(the result of vectorizing). As the vectorizing method, the expression of the accuracy is different. In the manual input method, the visual interpretation technique has been used. But, when we use the raster file scanned from the cadastral map sheet, the accuracy should be examined in three aspects like followings.

<Table 1> Comparison between these vectorizing methods

Layer	The screen digitizing		The developed software	
	Input time	Remarks	Input time	Remarks
The map neat lines	15 min.	Forming the vectors of map neat lines at non-uniform intervals	10 min.	Forming the vectors of map boundary lines at uniform intervals
The boundary lines	150 min.	Generate the red-tape formalities such as expansion and reduction The operator moves the pointer manually around the parcel boundary	120 min.	The operator can perform the job more convenient because the pointer is located automatically around the parcel boundary
Line thinning			(35 min.)	Excluded in the the time required because it can be processed in a batch mode
Total	165 min.		130min.	

- the accuracy between the cadastral map and the raster file.
- the accuracy between the raster file and the input vector file using this raster file.
- the accuracy between the cadastral map and the input vector file.

Among these, the first is related the accuracy of a scanner. Generally, the drum scanner is better than the flat bed scanner. But, the drum scanner can't be used in digitizing the cadastral map because the cadastral map is made from the aluminum kent paper or the Korean paper. Also, when we use the flat bed scanner, there is the problem that the size of the flat bed scanner does not match with the size of the cadastral map. But, these problems can be solved because the accuracy and the size of a flat bed scanner are increased recently.

Secondary, aspect of the accuracy is dependent on the operator's character, career, and etc.. Therefore, the accuracy can't be tested from this point of view. When we digitize the cadastral map with the developed software, the thickness of lines in the cadastral map is approximately 3 pixels. If we have a mistake in inputting the center point, the error of 1 pixel will be occurred in the x, y direction. It can be calculated as the distance on the map sheet. The value is 0.0898(mm).

It is less than the limit of the allowable drawing error(0.2mm) which is generally used in a map drawing. Also, it is less than the resolution of the automatic drawing instrument(0.1mm) in a digital map. Therefore, there will be no problem if there is no mistake of an operator.

Finally, the accuracy between the cadastral map and the vector file is determined by the first and the second accuracy. So, if there is no problem at the first and second accuracy, there will be no problem at the third aspect of the accuracy.

5. CONCLUSION AND FURTHER RESEARCH

The main issue of digitizing the cadastral map is how we can input the cadastral map correctly and efficiently.

The methods to make up for the weak point of the manual input method are the screen digitizing method and the automatic input method. A strong point of these two methods is not transforming the raw data, because these methods are operated after scanning the cadastral map. But the screen digitizing method is not more efficient in reducing the input time than the manual input method. The automatic input method produces different positions and numbers of vectors, and sometimes, it is difficult to be applied according to the condition of maps. So, these two methods can't be adopted for efficient digitizing of the cadastral map. Especially, there was an attempt to use the automatic input method as a tool for digitizing the cadastral map, but the automatic input method couldn't be applied, because the cadastral map is related with the right of property, and cannot allow for the error.

In this study, we developed the hybrid vectorizing software to enhance the deficiency of the analogue methods and to ensure the accuracy of the cadastral maps. The hybrid vectorizing adopted a screen-digitizing as a prototype and automated the procedure of searching the intersection of lines with efficiency. We developed and applied the new line searching algorithm suitable for the cadastral map, and added some functions to give an operator convenience at the screen digitizing task.

The result of developed software was found that there are no vectorized lines which are exceed map boundaries in the visual interpretation. So it is same as the screen digitizing method in the accuracy. Also, we test the efficiency by comparing the software developed in this study with the screen digitizing method using the AutoCad program. As a result, we can save 35 minutes using this software.

As the output of this software is dxf formatted not including topology, the further study about the topological product should be progressed. From this, we can use the cadastral map efficiently and can construct the integrated land information system based on it. Only if this process accomplished, we can use the cadastral map efficiently as the base map on GIS.

ACKNOWLEDGEMENT

This research is performed by the project of "Development of Expert System that contains Expansion-Shrinkage Correction, Map Matching Correction and Quality

Examinations for Computerization of cadastral map", sponsored by Ministry of Science & Technology of Korea.

REFERENCES

1. Ministry of Science & Technology of Korea , 1998, Development of Expert System that contains Expansion-Shrinkage Correction, Map Matching Correction and Quality Examinations for Computerization of cadastral map , pp. 79-81.
2. Korea Cadastral Survey Corporation, 1987, A Study on the Cadastral Map Revise, p. 46.
3. Baek Seung-Chul, 1994, A Study on the Cadastral Map Re-production by Scanning, Master's Thesis, Chongju University.
4. Research Institute of Computer, Information & Communication, Busan National University , 1998, Development of a Vectoring Software for Printed Maps,
5. Seoul Development Institute, 1993, The Study of Seoul City Implementing City-wide GISs, Seoul Development Institute, 93-R-25.
6. Han Nak-hee, Rhee Phil-Kyu, 1996, Parallel Thinning Algorithm using Weighted-Value, Korea Journal of Cognitive Science, Vol.7, No.1, pp. 5-35.

CONTACT

Byoungjun Seo
Ph. D. Course, Researcher
Department of Civil, Urban & Geo-Systems Engineering
Seoul National University
Shillim Dong
Gwanak Gu
Seoul
KOREA
Tel. + 82 2 880 7371
Fax + 82 2 889 0032
E-mail: cctrp@chollian.net

Jaejoon Jeong
Ph. D. Course, Researcher
Department of Civil, Urban & Geo-Systems Engineering
Seoul National University
Shillim Dong
Gwanak Gu
Seoul
KOREA
Fax + 82 2 889 0032
E-mail: hayoon@chollian.net

Jaebin Lee
Graduate Course
Department of Civil, Urban & Geo-Systems Engineering
Seoul National University
Shillim Dong
Gwanak Gu
Seoul
KOREA
Fax + 82 2 889 0032
E-mail: damanegi77@hotmail.com

Assoc. Prof. Yongil Kim
Department of Civil, Urban & Geo-Systems Engineering
Seoul National University
Shillim Dong
Gwanak Gu
Seoul
KOREA
Fax + 82 2 889 0032
E-mail: yik@snu.ac.kr