ROADS EXTRACTION AND THEIR CHANGES DETECTION IN AIRBORNE IMAGES WITH THE HELP OF GIS DATA

Pengfei Liu, Jianqing Zhang, Li Pan and Liang Zhong

School of Remote Sensing Information Engineering,Wuhan University
Luoyu Street 129, Wuhan city, Hubei Province, P.R.China
aiyuer2835@gmail.com

Commission ThS-18

KEY WORDS: Linear features, GIS, error correction, gradient, edge tracking

ABSTRACT:

This paper presents a novel routine to detect the *changes of main roads* with airborne images of medium-scale of 1:10,000. It mainly focuses on the *changes detection* of primary streets not the extraction of roads though it sounds an essential step. Here comes the brief work flows of this method: first, auto dodging the original image with Wallis Filter (Deren Li, 2006) as the pretreatment, second, get the variance and gradient images for the seeking of roads seeds, and third, a special searching processing is employed to locate the precise positions of the *roads seeds* in the above result and the original image in a special extension defined by the *poor GIS data*, and last, with several assumptions based on the gray properties of the roads in the image, preliminary changes of the primary roads can be detected. Results of all the steps mentioned above should have been refined by the *GIS (Geographical Information System)* data—the so-called *DLG (Digital Linear Graphic)* data, unfortunately, the *DLG* data at hand is too poor to identify the roads' locations but to provide only initial positions of the main roads while the secondary access ones have to be ignored. Experiments prove that this work flow effective for roads with homogeneous gray distribution.

1. INTRODUCTION

As the development of the Earth Observation technologies, the 3S techniques (GPS (Global Positioning System), GIS (Geographical Information System) and RS (Photogrammetry and Remote Sensing)) come more and more closer to our daily life have made great changes to the development of Surveying. More and more these data especially the RS images are available easily even they may drown the earth. However, how to make the full utility of these data to update the *GIS* database timely is still a hot spot in Surveying, GIS and RS realms
.
Changes detection, especially the changes of the main natural or man-made features such as *residental*, *water* and *roads* is of great importance to updating timely of the *GIS* database, and with the availability of huge amount of these RS images, more and more researches on the updating of GIS using images had been devoted to, and is still an active researching interest. As far as the roads changes detection are concerned, just as described in the reference (Wenzhong Shi, 2001), there had been a lot efforts put in and many methods are introduced by their work.

Concerning here is the extraction of the road's center points—called as the road seeds in this paper, under the supervising of the *poor GIS data*, and so that they will be used directly to determine the changes of the roads, and due to the *poor GIS data*, only the main roads and their changes are investigated in this paper.

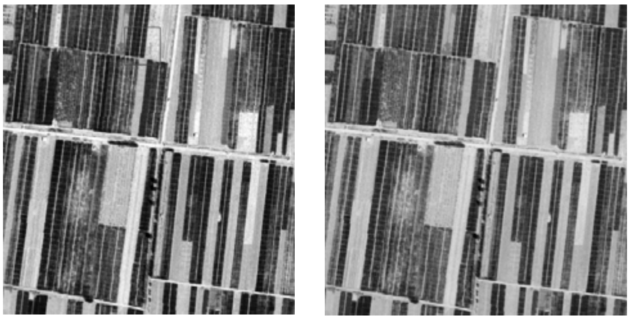
Usually, road center points extraction is the very first step for roads extraction, here comes some brief introductions to this step proposed by the papers mentioned above. With regard to the semi-automatic roads extraction this information is always offered by the operator that is why called semi-automatic. For details go to the papers (Fangju Wang, 1998, Yassin M.Y. Hasan, 2000, Boggess, 1993).

Doctor Li provides two semi-automatic extraction techniques: a model driven linear features extraction scheme and a new type of Snakes: LSB-Snakes—Least Squares B-spline Snakes (Haihong Li, 1997, M. Kass and Terzopoulos, 1988). While concerning the automatic roads extraction (or the roads seeds extraction), many algorithms are proposed: based on the image classification, edge detection and model matching, etc.

Doctor Hu offeres a least-square model matching algorithm to extract the road center-points by matching the image with a series of local models with mutable width changing dynamically according to the road's width (Xiangyun Hu, 2001).

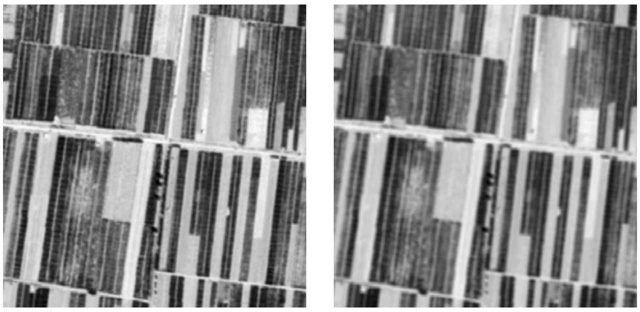
This paper provides a new idea to get the changes information of the main roads directly by extracting the road seed points in the normal direction initially provided by the *poor GIS data*, then compare the extracted key road nodes with the original nodes toobtain the changes.

The structure of this paper is as following: section 2 introduces the pretreatment of the images to enhance the road as much as possible, and also describes detailedly the algorithm and the work flow for roads center points extraction, section 3 and 4 give some examples using the work flow in section 2 and section 5 gives some discussion and a conclusion.



(a) Part of The Original Image (b) Wallis result of Original Image

Figure 1: Part of The Original Image and The Result of Wallis Filter



(a) Median Filter with window kernel: 3 × 3 (a) Median Filter with window kernel: 5 × 5

Figure 2: Median Filter results in ENVI

2. WORK FLOWFOR ROAD CHANGES DETECTION

2.1 Roads Image Pre-processing—Road Sharpening

Like the preprocessing in Remote Sensing realm, some operations have to be done before the next step based on the following two facts: first, due to the natural effects like lights and clouds and the man-made factors like heterogeneous light distribution caused by images jointing, images are unavoidably contaminated by these noises and cannot be used directly, second and moreover, in order to obtain roads as many and good as possible, enhancing the contrast between the roads and non-roads pixels is also of essence even if without concerning the affects of the noises.

Image pretreatment for roads (or roads seeds) extraction is usually removing the noises of the image. In this paper, two convolution filters is exploited: Median Filter: a global filter (Guihua Zhang, 2007) and Wallis Filter (Li Zhang, 1999) the local filter.

Images in Figure 1 shows both part of the original image in Figure 1(a) and the result of Wallis Filter in Figure 1(b). And in Figure 2 depicts the result of Median Filter with window kernel: 3×3 in Figure 2(a) and window kernel: 5×5 in Figure 2(b), respectively.

2.2 Road Seeds Extraction

In this part, algorithms for road center points extraction of each single road and rules for *changes detection* are introduced at length.

Features extraction from image is always the very first essential step of processing. They are existed in image in the forms of crossing points, edges and some others like magnitude, gradient, variance and so on. Concerning the roads extraction, features utilized are the center points with some clear gray distribution characteristics used to identify them as listed in section 2.3.

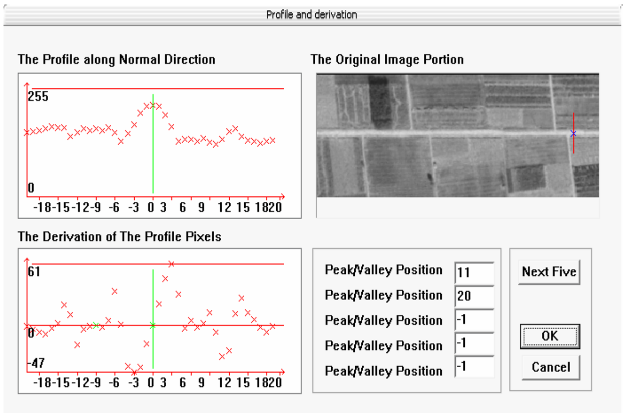


Figure 3: The Portion of Image and Profile and the Seed Point

Road center points, which can be called here the road seeds, can be located in the region near the *poor GIS data* data, the special characteristics of the road seeds are defined below.

As demonstrated in Figure 3, they exist in somewhere near the *DLG* corresponding to the peak values if the old *DLG* roads are still there. Obviously, they are almost dwell in the perpendicular direction of the old road linear segment and more important, in the nearby there are also two positions with local maximum gradient values, the distance between the two equals to the road width or so, which can be gained from the edge image through many algorithms, the so called edge detection operator, like Roberts, Sobel, Prewitt and Canny etc. Canny operator with the best result of information is employed here, details please go to (Canny, 1986).

Based on the above conclusion, the work flow for each single road seeds extraction is designed as following:

1. Get the perpendicular direction of an old *DLG* road segment, drawing the profile of this line in that direction, the extension of this depends on the width—w of the old roads, here ±2w is adopted. The horizontal axis represents distance to the seed point, while the vertical the gray value of that point. See Figure 3.

2. Get the peak values in the profile so that the center point can be obtained. According to the mathematics, the local peak or valley values of function are the positions where the first-order derivations come to zero. And in image processing derivation is always approximated by difference. That is: the seed points can be located by the derivation of the profile. The calculation equation is:
$$f'(i,j) \approx D(i,j) = f(i,j) - f(i+1,j)$$
Where the $D(i,j)$ is the difference value of the pixel in (i,j), and $f(i,j)$ is the value of point in the profile. See Figure 3.

3. After getting the local peak or valley values of profile, some further testimonies are needed to pick out the real seed point. This can be done by searching the two road edges points near the center points in the edge image.

4. If the two edge points near the peak or valley values can

be founded, this point can be recognized temporarily as seed points. Maybe in this step not only one of these values are marked as the seed points, they can be assured furthermore in the subsequent processing. Changes information can be got simultaneously from this step, details are in the next section.

5. Find the next seed point along the initial direction of the old *DLG* road segment.

6. Repeat the above routine until all the segments are examined.

The above descriptions are illustrated in the Figure 4.And the main work flow of this paper is shown in the Figure 5.

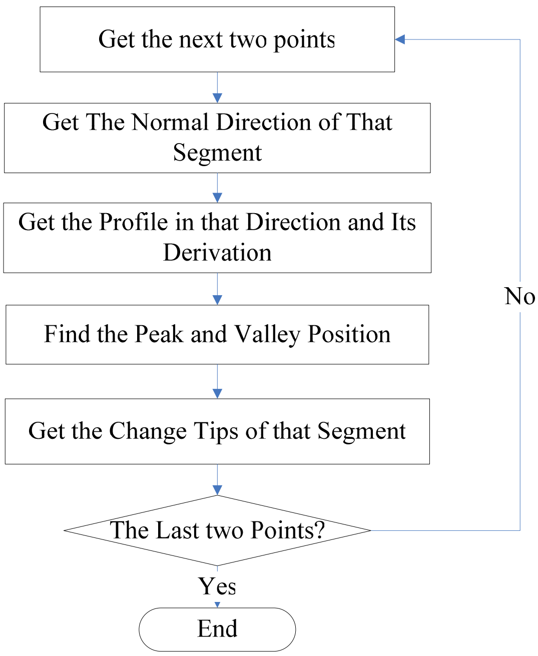


Figure 4: Work Flow for each Single Old Road

If none seed points can be found or the distance the newly found seeds to the old segments are longer beyond a threshold: D_1 in the above direction, two possible results can be got:

Disappeared The special road segment has disappeared. This assumption needs to be assured by the next seed point.
Changed The special road segment may be changed like being extended or being turned to another direction only if the next seed point is still there.

An additional explanation is necessary here: under some circumstances—see Figure 6, the roads in the image are covered by the tall trees along them and roads are shown as a dark linear zone in contrast to the normal roads in the image with higher light value about greater than 220.

This is a special case, two possibilities may be got from this:

1. This is a segment of the road disturbed by the green utilities in the sides.

2. This is a segment of other land-covers like the farm land whatever.

And in this case the roads center points are the positions where the valley values appear in the profile of some point in this specified segment and their first-order derivation in Mathematics are zero too. So that both the peak and valley values in the profile should be considered when extracting the road seed points, and if the above case takes place more attention should be paid in the subsequent procedures for the second possibility may be the situation.

So far, the work flow for each road polyline has been llustrated. Repeat the whole above routine with all the main streets of the old *poor GIS* data, and the changes tips of each road can be gained. This is illustrated in the Figure 5.

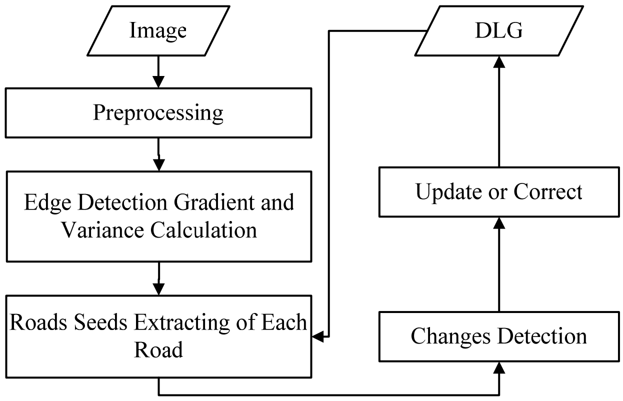


Figure 5: The Whole Work Flow for Roads Changes Detection

As expressed above, to determine the changes of a road involves not only one or two seed points, but also the edge points near the old road and sometimes maybe even more information for the possible confusions should be included.

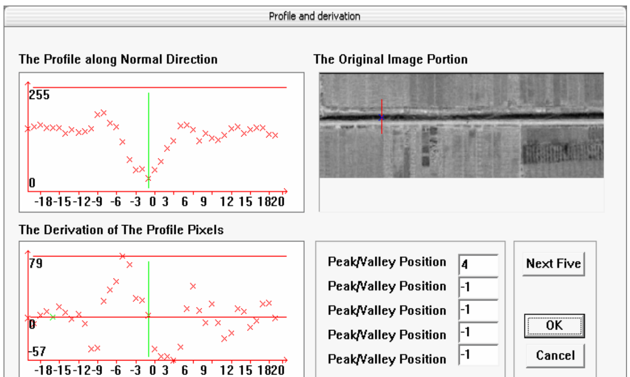


Figure 6: The Dark Portion Original Image and Profile of a Point

2.3 Strategy on Changes Detection

Principles used to determine the changes of a special road are highlighted in this part.

Honestly, it is so difficult to distinct the extraction of the seed points from the determination of the changes of roads. The two steps are actually happen simultaneously for when extracting the seed points some changing tips can be gained immediately, maybe sometimes they are not so clearly, and as the going of the extraction with more and more seed points available some of these tips are eliminated while others are kept and becoming more and more clear. When the extraction of the seed points are finished, the change tips should also been gained and marked to the user.

An imaged object is usually defined and can be identified by its characteristics, which depend largely on both its physical and color attributes and can be classified into five groups: photometric, geometric, topological, functional and contextual.

Here are some most useful properties and they are also used in this paper (Haihong Li, 1997, J. Amini, 2002) , In equations of this part, $f(s)$ is a vector function which maps the arc-lengths to point (x, y) in the image. The preprocessed digital image is represented by $G(x, y)$, $G(f(s))$.

- Road surface is usually homogeneous, at least in a certain portion of the image(photometric); This is equivalent to
$$E_{r1} = \sum_i \int_{\Delta s_i} [G(f(s)) - G_m(\Delta s_i)]^2 \, \mathrm{d}s \Rightarrow \text{minimum}$$
Where Δs_i is a short segment of the road segment, and $G_m(\Delta s_i)$ is the average value of G on s_i , i.e.
$$G_m(\Delta s_i) = \frac{\int_{\Delta s} G(f(s)) \, \mathrm{d}s}{|\Delta s_i|}$$

- Based on this feature, a suitable distance can be jumped while tracking the road seeds from the one has been located and detect the false changes information if an unexpected seed points with obviously different feature comes out—say, point in the valley position of the profile which may be disturbed by the green utilities along the roads, and if this happens just change the judging rule for the road seeds into the valley point in profile so that to be immune to the affects of the trees.

- Road surface has often a good contrast to its adjacent areas (photometric); This is the main principle for the designation of work flows in this paper, as described in the Figure 3 and the derivation Figure 3, the positions e_1 and e_2 are there where show the contrast between the roads and non-roads pixels, and the better the contrast is, the better to locate these positions and so that provide higher possibility for the roads center points. This suggests that a squared sum of the gray values (or their second derivatives in the direction normal to the road) along the curve attains a maximum, i.e.
$$E_{r2} = \int [G(f(s))]^2 \, \mathrm{d}s \Rightarrow \text{maximum}$$

- A road is light linear feature. In fact, this is a generalization of the two previous properties. This property can be considered through the appropriate evaluation on the formula:
$$E_{r3} = \int w(d(s))[G(f(s)) + d(s)n(s)]^2 \, \mathrm{d}s \Rightarrow \text{maximum}$$
Where, $d(s)$ is the distance between road line and the linear feature near it, and $w(d(s))$ is a Gaussian weight function that decreases as the distance $d(s)$ increases.

- In terms of basic geometric property, a road is usually smooth and does not have small wiggles. In fact, most roads consist of straight line segments connected with smooth curves, normally circular arcs. This property can be represented mathematically as:
$$E_g \int |f''(s)|^2 \, \mathrm{d}s \Rightarrow \text{minimum}$$
Where, this implies that the road line can be represented as a cubic spline.

- Roads are elongated (geometric); this depends on both the width and the spatial resolution of the image, and the shapes are linear zones or linear polylines.

- Roads have a maximum curvature (geometric)—this attribute is not used in this paper:
$$|f''(s)| < T_1$$
Where, T_1 is a given bound.

- Roads don't ends for no reason (topological); For the extraction of the roads, this rule is used to group the line segments to be a network. However, in this paper, it's explored to make a conclusion whether or not roads existed in the old *DLG* remain in.

- Roads may be indicated by a special distribution of trees or the settlements near the streets (contextual). With the principle of contrast, the road seed points can still be located by this rule even though their contours are actually consist of the outlines of trees or settlements.

Due to the method using road center points to determine directly the changes tips of the main roads, changes information can be gained at latest after extracting all the seed points in a specified road—of course this is the worst which may be caused by exceptions like the trees along the roads or some other noises and so on.

Changes tips got in each step are described detailedly as follows: Here, "In" and "After" are the short of in and after the seed points extraction of each road respectively.

In During the extraction of seed points for each segment of a specified road:

- If none seed point near this segment can be found, this segment may be disappeared and even the whole road.

- In some segment, if only one seed point near the old key points can be extracted, enlarge the searching region in the normal direction of this segment and try to find another road center point according to the features above, and if still no point can be found then the extraction of this segment and the road is terminated, if this is the first segment then the road is marked as disappeared, if not trying hard to find the next seed points in the next segments.

- If the two seed points of this segment are peak and valley values respectively in the profile, shadow happens.

After After searching seed points in each segment, checking the extracted points with the old nodes, if the two numbers are different, this road is marked as changed, and split the original line.

3. EXPRIMENT

In the experiment, an aerial image with its corresponding GIS data in some suburban of Beijing is used. The image has a scale of 1:10,000 and is scanned with a pixel size of 100 μm. The area covered by the image is a rural area in which there are many roads of different widths passing through and intersecting each other, and many farm lands looks like very much short roads.

The steps adopting here are: first, overlapping the image with the corresponding by the affine transformation coefficients provided by the TIFF Word file with the image; second, select a road line to extract the road seed points, of course this can be done automatically; third, change the old poor road polylines according to the extracting lines.

4. RESULTS AND DISCUSSION

Some of the results have been shown in the above figures, and here gives some other detailed images: Figure 7 shows a part of that image with the poor GIS data in red and the extracted line in green.
This graphic show that the idea proposed here can do its job good.

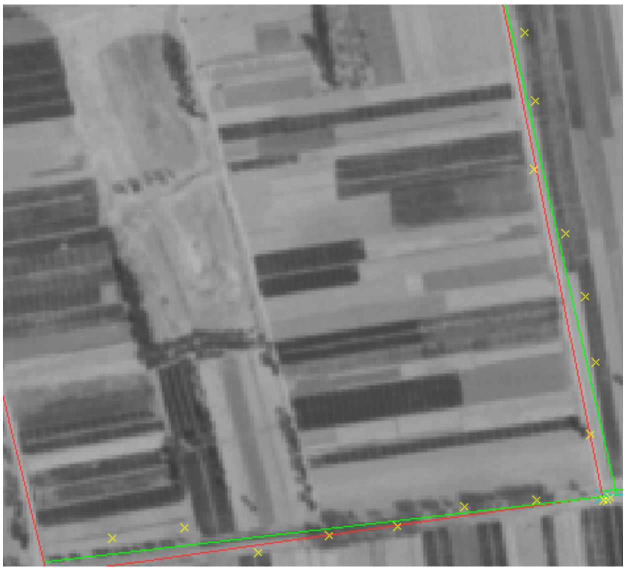


Figure 7: Part of The Original Image, The Poor GIS Data and The Extracted Road Line

And what if the results of roads disturbed by the green utilities or the trees along them? For them in contrast with the above figure, their gray values in the image are darker than that of the features along them, Figure 8 shows what happens to them and the bottom portion of this figure shows the detailed information of the region specified by the blue rectangle in the upper.
Roads with dark gray values in the image can still be extracted correctly, though it not always happens.

Things not always go as they are designed or expected, failures are still inevitably as shown in the Figure 9, and it's

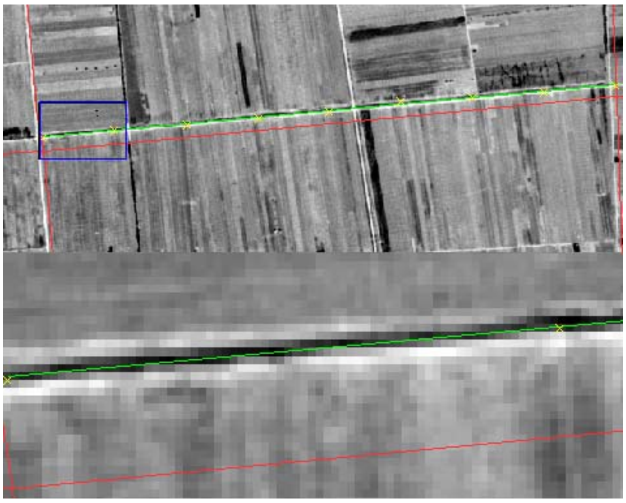


Figure 8: Part of The Original Image, The Poor GIS Data and The Extracted Road Line Disturbed by The Green Utilities

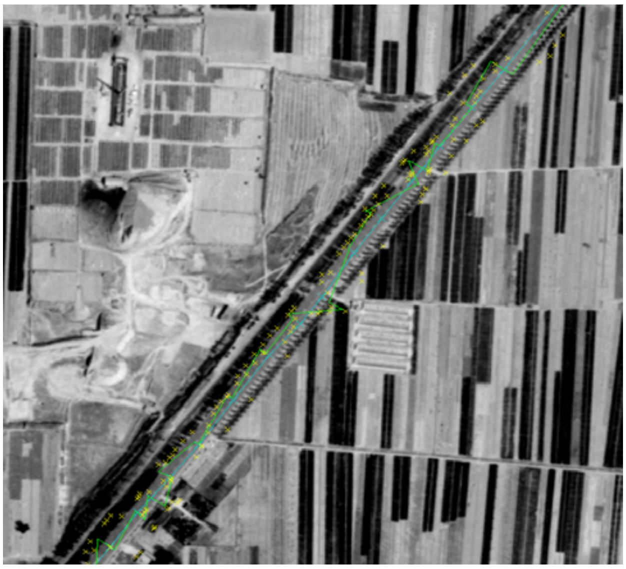


Figure 9: Road Seeds Extracting Failed

more worse when encountering the roads crossing the residential regions as depicted in the Figure 10.

In Figure 9 there is a road of 26m wide still cannot be extracted, due to the green utilities in and along it which bring more noises to the roads and make the situation more complicated.

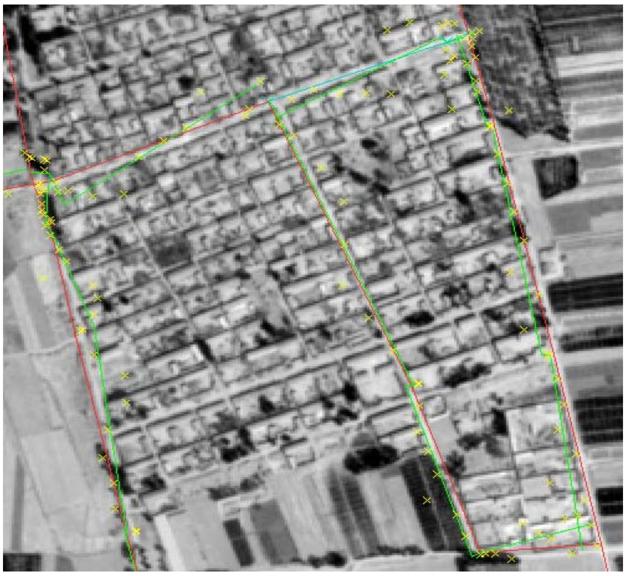


Figure 10: Road Seeds Extracting Failed in The Residential Region

5. CONCLUSIONS AND FUTUREWORK

From the figures in section 4 on page 5, a conclusion can be drawn that this method is an useful means to extract the road seed points and determine their changes information without having to do the complex procession from extracting the roads polylines to then compare them with the old poor DLG data to get the change information which will be time consuming and depend largely on the extracting result and more important this method does not take the old DLG data as an assistant. The idea in this paper exploits the poor DLG data and take it as an important known information for getting the changes information and only extracts the road seeds, so it save the extracting time and get the change information directly.

Of course, it is still not perfect, as shown in the Figure 9 and Figure 10. Due to the complex distributions of the roads, noises are unavoidably especially when facing the roads in the residential regions, it is difficult to tell the roads from the houses. Thus, roads across the residential regions need more attentions.

As for the wide roads like in Figure 9, more details of it are shown in the image, and more noises emerge to confuse the road seed points finding algorithm, therefore another strategy should employed in the future work: extracting the wide roads in the down sampling pyramid image. And of course, more restrains during the road seed points extraction should be introduced in the near future to refine the result.

ACKNOWLEDGEMENTS

The work was supported by the National Key Technology R&D Program of China under grant No. 2006BAB10B01.

REFERENCES

Boggess, J. E., 1993. Identification of roads in satellite imagery using artificial neural networks: A contextual approach. Mississipi: *Mississippi University Press*.

Canny, J., 1986. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 8(6), pp. 679–698.

Deren Li, Mi Wang, J. P., 2006. Auto-dodging processing and its application for optical rs images. *Geomatics and information Science of Wuhan University* 31(9), pp. 783–756.

Fangju Wang, R. N., 1998. A knowledge-based system for highway network extraction. *IEEE Transactions on Geoscience and Remote Sensing* 26(5), pp. 525–532.

Guihua Zhang,W. L., 2007. The realization of median filtering in image processing. *Journal of Jiamusi University (Natural Science Edition)* 25(4), pp. 500–502.

J. Amini, M.R. Saradjian, J. B. e., 2002. Automatic road-side extraction from large scale imagemaps. *International Journal of Applied Earth Observation and Geoinformation* 4(2), pp. 95–107.

Li Zhang, Zuxun Zhang, J. Z., 1999. The image matching based on wallis filtering. *Journal of Wuhan Technical University of Surveying and Mapping* (WTUSM) 24(1), pp. 24–27.

M. Kass, A. W. and Terzopoulos, D., 1988. Snakes: Active contour models. *International Journal of Computer Vision* 1(4), pp. 321–331.

Haihong Li, 1997. Semi-automatic Road Extraction from Satellite and Aerial Images. PhD thesis, Institut für Geodedäsie und Photogrammetrie an der Eidg. *Technischen Hochschule Z¨urish, Swiss Federal Institute of Technology Zurich*.

Xiangyun Hu, 2001. Automatic Extraction of Linear Objects and Houses from Aerial and Remote Sensing Imagery. PhD thesis, *Wuhan University P.R.China*.

Wenzhong Shi, Changeqing Zhu, Y. W., 2001. Road features extraction from remotely sensed image: *Review and prospects. Acta Geodaetica et Cartographica Sinica* 30(3), pp. 257–262.

YassinM.Y. Hasan, L. J. K., 2000. Morphological reversible contour representation. *IEEE Transactions on Pattern Analysis and Machine Intelligency* 22(3), pp. 227–241