TOWARDS AUTOMATIC ROAD EXTRACTION FROM DIGITAL IMAGES

Youssef Ibrahim Abd-Elaziz
Cairo University , Faculty of Engineering
Department of Public Works

Hazem Ahmed Fathy Barakat
Cairo University , Faculty of Engineering
Department of Public Works

Amr Hanafi Ahmed Ali
Benha University , Faculty of Engineering-Shoubra
Surveying Engineering Department

Heba m. M. Ahmed El-Qadi
Benha University , Faculty of Engineering-Shoubra
Surveying Engineering Department

ABSTRACT

Automatic extraction of road networks from digital images has been an active research subject for over a decade. This is due to its significant role in various application areas, like Intelligent Transportation (IT) Systems as well as GIS applications. While the manual extraction of road data is time consuming, the urgent necessity is automatic extraction of road data.

In this paper, an original framework for automatic extraction of road networks from digital images is developed. The Proposed algorithm is based on the theories of image processing that is oriented towards the mathematical morphological operation that is based on the theories set for spatial structure analysis. Implementing the road extraction process passed through two main stages: the first was extracting street network from a raster image and defining the road centerline and the second was identifying the coordinates of intersection points of the streets network.

In order to guarantee and validate the methodology of the proposal a software package was developed. Moreover, a case study is conducted to analyze the efficiency and performance of the proposed algorithms.

KEY WORDS:

Automatic road extraction, Morphological operations, Feature extraction, Image processing, Intelligent Transportation.

1. INTRODUCTION

Today's vehicle navigation systems have reached a high level of maturity, using huge up-to-dated data with a high coverage. Substantial work on road extraction has been accomplished since the 70's in computer vision and digital image processing directions [1].

Extracting features, "Digital Images, Scanned Maps" is one of the standard data sources for the acquisition of topographic objects, like roads for geographic information systems (GIS).

Road data in GIS are of major importance for applications such as car navigation or guidance systems for police, fire services, urban emergency services, and disaster management systems. And since the manual extraction of road data is time consuming, there is a need for automation [2]. For such applications, the road data needed to be represented in a separate layer in vector format. In order to get the desired data in the proper format, we need to extract the road features individually and to convert them from raster to vector by defining the start and end points for each road. And at the same time, identifying the junction and points..

2. FEATURE EXTRACTION

The main purpose of extracting features is reducing data by measuring certain properties that distinguish objects or their parts. Features can be classified into two classes, the natural features which are natural in the scene that are defined by the visual appearance of an image. These include the luminance of a region of pixels and grayscale texture regions. While, the second are artificial features resulting from specific manipulations of an image, where, image amplitude histograms and spatial frequency spectra are examples of artificial features [3].

Feature extraction concerns on finding shapes in computer images, as feature extraction process can be viewed as similar to the way we perceive the surrounded world. In features extraction, we generally seek invariance properties so that the extraction process does not vary according to chosen or specified conditions [4].

Feature extraction, usually, is associated with another method called feature selection. During this process, only the salient features necessary for recognition are retained [5].

One of the most widely used step in the process of reducing images to information is segmentation; dividing the image into regions that hopefully correspond to structural units in the scene or distinguish objects of interest.

Segmentation is often described by analogy to visual processes as a foreground / background separation, implying that the selection procedure concentrates on a single type of feature and discards the rest. Selecting features within a scene or image is an important prerequisite for most types of measurement or understanding of the scene [6].

A simple shape extraction technique is the thresholding, where the images can be viewed after separating the feature from the background. If it is assumed that the shape to be extracted is defined by its brightness, then thresholding an image at that brightness level will define the shape. Thresholding is clearly sensitive to any change in illumination. It does not require much computational effort. If the illumination level changes in a linear fashion, then using histogram equalization will result in an image that does not vary. The major drawback comes out where the result of histogram equalization is sensitive to noise, shadows and variant illumination.

An alternative approach is to subtract an image from a known background before thresholding. This assumes that the background is known precisely, otherwise many more details than just the target feature will appear in the resulting image, and clearly the subtraction will be unfeasible if there is noise on image.

Even though thresholding and subtraction are attractive because of their simplicity and speed. The performance of both techniques is sensitive to partial shape data, noise, and variation in illumination and to occlusion of the target shape by other objects. Accordingly, many approaches to image interpretation use higher level information in shape extraction, namely how the pixels are connected within the shape. This can resolve these factors [4].

3. METHODOLOGY

Recalling that the objective of this research acts upon automation of feature extraction process, which is useful in isolating a specific feature as roads from surrounding features. The applied algorithm is based on the theories of image processing oriented on mathematical morphological operations. Automatic Road Extraction are mainly divided into two procedures:

- Road finding.
- Seed extraction.

3.1. The Road Finding Procedure

The basic assumption used for this procedure is that roads are generally lighter than the background, "i.e. contrast rule" where roads usually contrast sharply with background. The Mathematical Basic Concepts are explained through the next section.

3.1.1 Structuring element

An essential part of any morphological operations is the structuring element used to probe the input image. A structuring element is a matrix consisting of only 0's and 1's that can have any arbitrary shape and size. The pixels with values of 1 define the neighborhood.

Two-dimensional, or flat, structuring elements are typically much smaller than the image being processed. The center pixel of the structuring element, called the origin, identifies the pixel of interest -- the pixel being processed. The pixels in the structuring element containing 1's define the neighborhood of the structuring element. These pixels are also considered in dilation or erosion processing.

Three-dimensional, or non-flat, structuring elements use 0's and 1's to define the extent of the structuring element in the x- and y-planes and add height values to define the third dimension [7].

3.1.2 Morphological dilation and erosion

Dilation and erosion operations are fundamental to morphological processing. In fact, many of the morphological algorithms are based on these two primitive operations.

These operations are limited to object modification by a single ring of boundary pixels during each iteration of the process [3].

In the dilation process, first, define A as the reference image and B is the structure image used to process A. The dilation of A by B, denoted A $\oplus$ B, is defined as

$$A \oplus B = \left\{ z \mid \left[\left(\hat{B} \right)_z \cap A \right] \subseteq A \right\}. \tag{1}$$

Where $\hat{B}$ is B rotated about the origin. Dilation has many uses but a major one is bridging gaps in an image due to the fact that B is expanding the features of A. Dilation on the other hand can be considered a narrowing of features on an image.

Again defining A as the reference image and B as the structure image:

$$A \ominus B = \left\{ z \mid \left(B \right)_Z \subseteq A \right\}. \tag{2}$$

Many times dilation can be used for removing irrelevant data from an image [8].

3.1.3. Morphological opening and closing

The combination of erosion followed by dilation is called an opening, referring to the ability of this combination to open up spaces between just touching features. It is one of the most commonly used sequences for removing pixel noise from binary images. There are several parameters that can be used to adjust erosion and dilation operations, particularly the neighbor pattern and the number of iterations. In most opening operations, these are kept the same for both erosion and dilation [6].

The process of "opening" an image will likely smooth the edges, break narrow block connectors and remove small protrusions from a reference image.

The opening of set A by structuring element B, denoting A ° B, is defined as

$$A \circ B = \left(A \ominus B \right) \oplus B. \tag{3}$$

Thus, the opening A by B is the erosion of A by B, followed by a dilation of the result by B.

Also, The "Closing" process of an image will smooth edges but will fuse narrow blocks and fill in holes.

The closing of set A by structuring element B. denoting A·B, is defined as

$$A \bullet B = \left(A \oplus B\right) \theta B \tag{4}$$

This, in words, says that the closing of A by B is simply the dilation of A by B followed by the erosion of the result by B [9].

3.1.4. Road finding algorithm

The strategy started with some morphological operations where it is completely pixel based without using any edge information.

This approach is summarized through the following specific steps [10]:

- In short, the image is represented as a matrix. Each element of this matrix is called an image element or pixel. When applying a function to the image, it means applying it to each pixel. Here in, a grayscale intensity image from color image created by eliminating the hue and saturation information while retaining the luminance and the threshold value can be predicted.
- Subsequently, we can convert gray level image to binary image by acquiring binary image which is a logical array of 0's and 1's. The binary image can be obtained by applying thresholding to the gray level image as a test for each pixel. If its value is greater than or equal to threshold value, then the resulted binary image will be of value 1; otherwise, the resulting binary image will be equal to 0.
- After that, we started removing errors "noise" by applying dilation process using structuring element of specific origin, shape and size. At the same time, we started filling the holes in the input binary image. A hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image. These elements varied from case to another. These problems vary according to the feature themselves "polygons, polyline…"
- The last step is thinning road networks by reducing binary image shapes in an image to stroke that are a single pixel wide.

Fig (1) represents the proposed stages of feature extraction step.

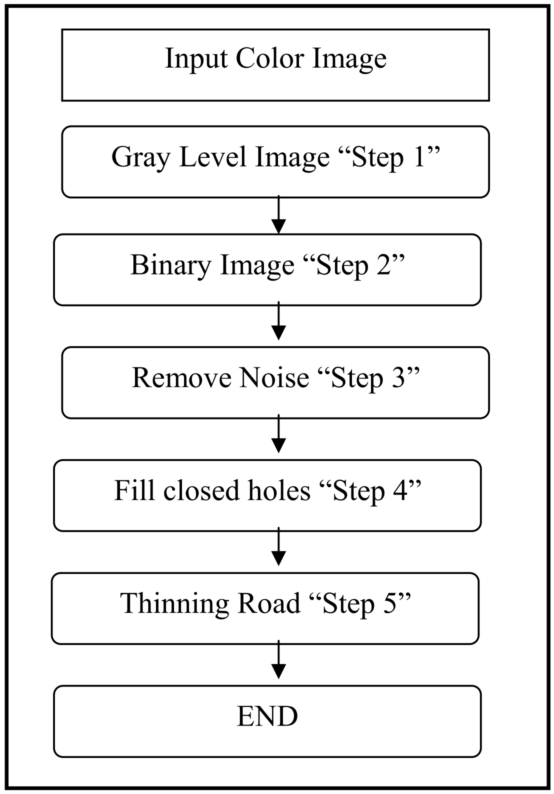

Fig. (1): Block Diagram Represents Road Finding Steps.

After these Steps, the image is converted into binary image and the street network is extracted as a separate raster layer. Next stage is done by extracting the street seeds, then vectorizing the road layer. For further information see [7].

3.2. Seed Extraction Procedure

The proposed algorithm for solving this problem is based upon pixel neighbors, where, a pixel p at coordinates (x, y) has four horizontal and vertical neighbors whose coordinates are given by

$$(x + 1, y), (x - 1, y), (x, y+ l), (x, y - 1)$$

This set of pixels, called the 4-neighbors of p, is denoted by N4 (p). Each pixel is a unit distance from (x, y), and some of the neighbors of p lie outside the digital image If (x, y) is on the border of the image. The four diagonal neighbors of p are denoted by ND(p) with coordinates

$$(x + 1, y + 1), (x + 1, y – 1), (x - 1, y + 1), (x - 1, y - 1)$$

These points, together with the 4-neighbors, are called the 8-neighbors of p, denoted by N8(p). As mentioned before, some of the points in ND(p) and N8(p) fall outside the image if (x, y) is on the border of the image[9].

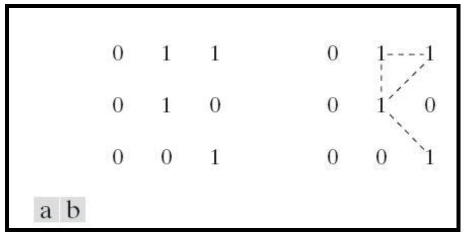

Fig. (2): Adjacency concept.
(a) Arrangement of pixels;
(b) Pixels that are 8-adjacent (shown dashed) to the center pixel.

As shown in Fig. (2), road pixels are the pixels of value 1. For that pixel, if two successive neighbor pixels in the N8(p) set are different in value so the condition satisfied and more than four condition satisfied for a complete circle, so the center pixel is a seed.

3.2.1. Seed Extraction Algorithm

Based on the following scenario, the seed extraction algorithm is performed.

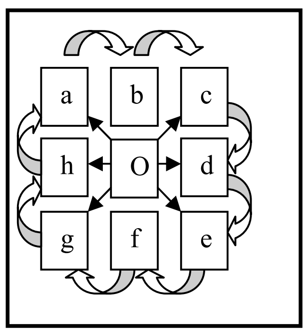

Fig. (3): Pixels Neighbors.

Step 1: Origin or center pixel must be a road pixel of value 1, which is represented by "O" in Fig. (3).

Step 2: As shown in Fig. (3) for neighbor pixels a, b, c, d,….., h. If at least four of theses conditions are fulfilled, then the tested origin considered as a seed.

B (b) < > B (a), B(c) < > B (b), B (d) < > B(c), B (e) < > B (d),

B(f) < > B(e), B(g) < > B(f), B(g) < > B(h) , B(h) < > B(a) .

Where:

B (a), B (b), B (c),……., B (h): are the brightness of the pixels a, b, c,….,h respectively.

< >: means not equal.

Figure (4) represents the flow chart that explains the Seed Extraction steps.

It begins with performing thinning road pixels, with loop for each pixel, considering it as the origin. After that establishing the neighbors pixels, and examining each two successors neighbor pixels "the equality of their values", if the origin pixels satisfies more than 4 inequality conditions then this origin pixel is considered as seed or junction [10].

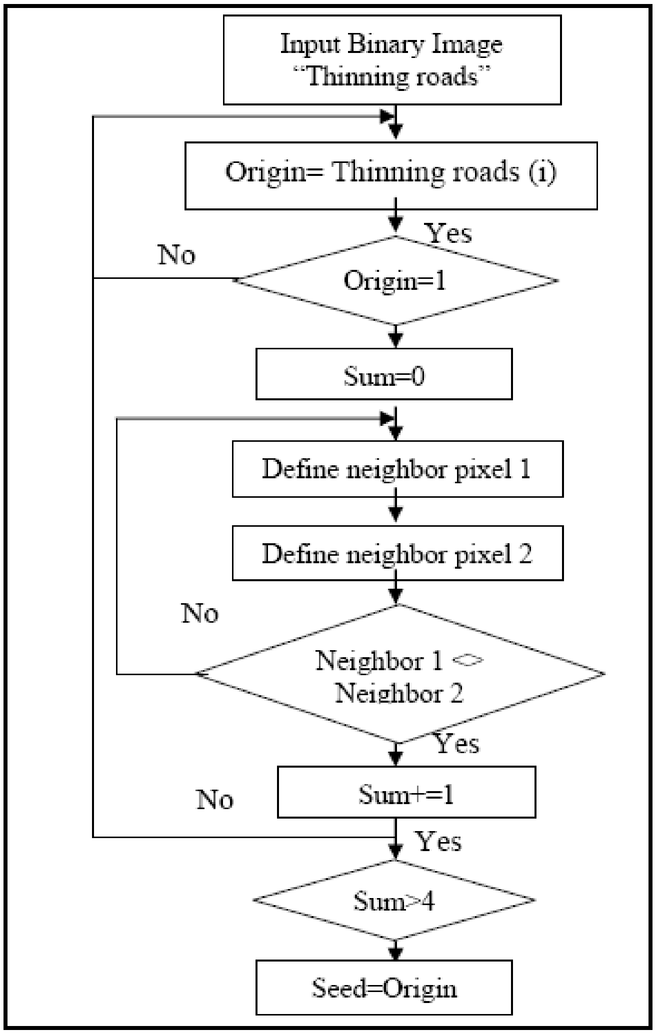

Fig. (4): Seed Extraction Flow Chart

4. RESULTS AND ANALYSIS

The raster images were selected for the benefits of its homogenous colors, the high contrast and well defined shape of roads network which helps in the examination of this sub-module. The study, that includes the above described properties, was selected in Downtown "Kasr El-Nil", Cairo, Egypt. Within this area, and according to the designed package, the study area was examined.

4.1. Case Study

The study area is a part of the road network Downtown, Kasr El-Nil, Cairo, Egypt, which covers about 1km2 with scale 1:5000 and resolution 1 pixel / m. Figure (5) represents the digital road map (Raster image for the study area).

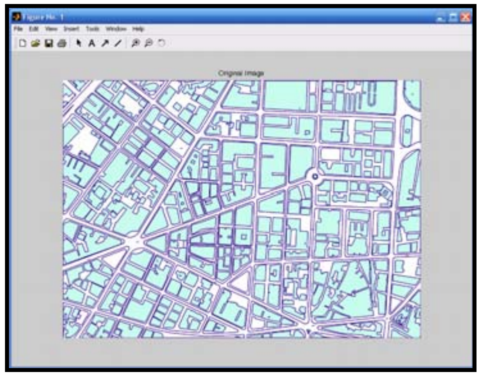

Fig. (5): Input Image.

4.1.1 Road finding procedure

Through this procedure and according to the designed package, the following steps are presented:

Step 1: Converting image into gray level.
This step is done to convert color image into gray level image where the designed algorithm is capable of dealing with either color or gray level image. The converted gray level image is shown in figure (6).

Step 2: Selecting suitable threshold value.
In order to convert the gray level image into binary image, different trials have been done to define the suitable threshold value that completely distinguishes the road feature.

Step 3: Applying thresholding to convert gray level image to binary image.
The selected threshold value was applied to get the binary image representing the extracted road. Figure (7) shows the desired roads network with white color, other features with black color and there were undesirable features with common values as road network.

Step 4: Inverting image.
In this step we inverted the image color to get the road network with black color see figure (8).

Step 5: Removing noise "errors" due to extraction process.
This step is very important when road network colors are shared with other undesirable features; we applied noise removing filters which depend on the Dilation morphological operation. The removed noise image is represented in figure (9).

Step 6: Filling holes in close regions.

After removing noise by using the morphological operation through a defined structure element, other undesirable big features still remains at the image; hence filling the holes in close region is done, which depends also on closing the morphological operation. This is a combination of Dilation and Erosion processes. After implementing this step, the roads networks are completely extracted as shown in figure (10).

Step 7: Thinning road network

The last step of the road finding procedure is done, by converting the extracted road blocks into a linear form of only one pixel width. This is accomplished by applying the skeleton operation. Where, figure (11) shows the final result of this procedure which is the thinning road network.

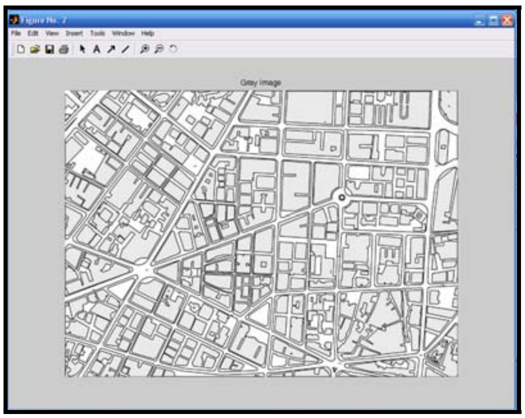

Fig. (6): Converted Gray Level Image

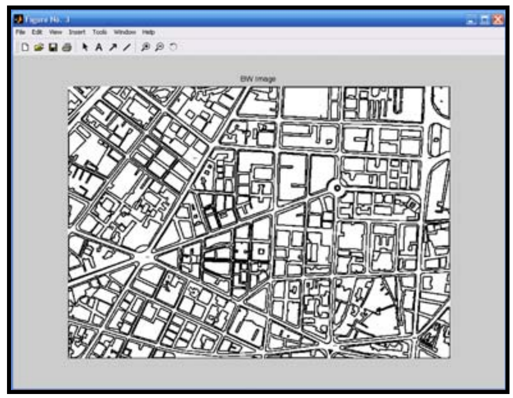

Fig. (7): Binary Image

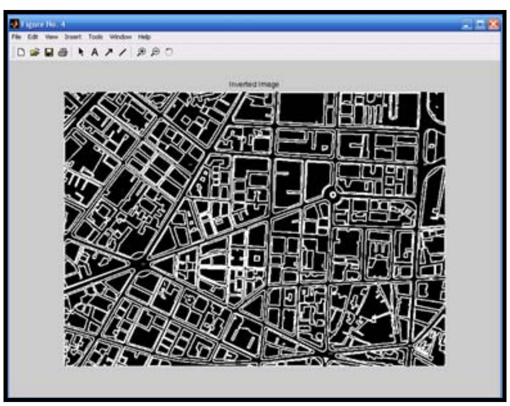

Fig. (8): Inverted Binary Image

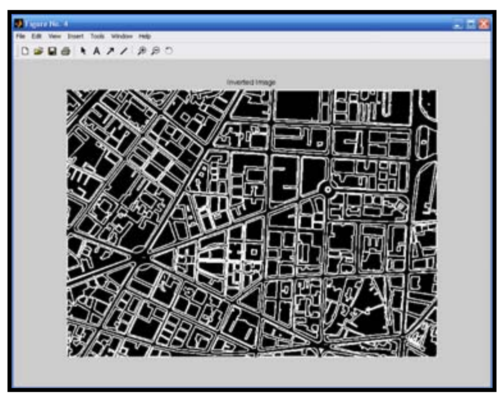

Fig. (9): Removed Noise Image

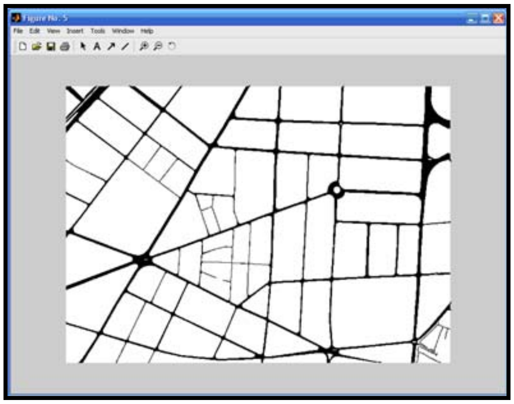

Fig. (10): Removed Big Holes Image

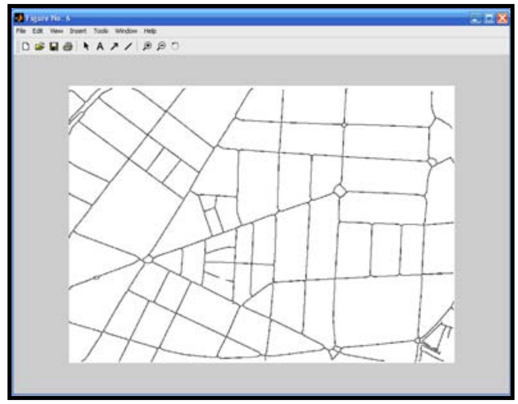

Fig. (11): Thinning Extracted Roads Network

While, the overlapping of extracted roads network and the original image are represented in figure (12).

In order to analyze the resulted shift of the extracted road center line, the digitized center line will be the reference. The analysis is done on the line EF "the longest road on the map". Cross sections was taken each 5 m and the determination of the coordinates of intersections with digitized and extracted roads was done.

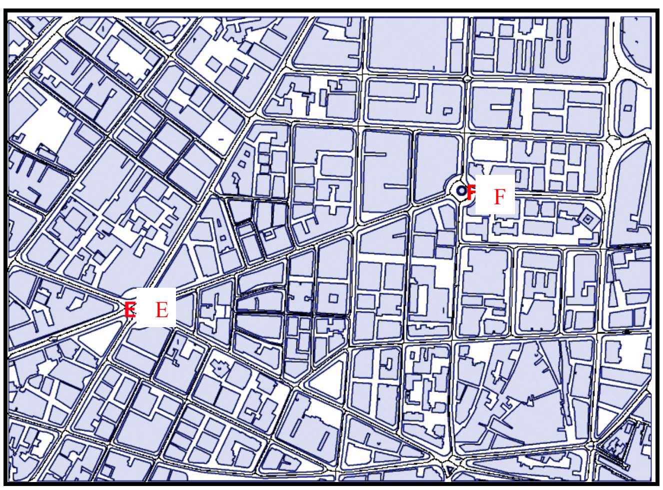

Fig. (12): Extracted Road Overlapped with Original Image.

The shift value between the digitized points and the extracted one is calculated according to the equation (7).

$$\Delta X = X^{\backslash} - X$$
$$\Delta Y = Y^{\backslash} - Y$$
$$\text{Shift} = \sqrt{(\Delta X^2 + \Delta Y^2)} \qquad (7)$$

Where:

$X^{:}$ The East-West coordinates for the digitized points

$Y^{:}$ The North-South coordinates for the digitized points

$X^{\backslash}$: The East-West coordinates for the extracted road points

$Y^{\backslash}$: The North-South coordinates for the extracted road points.

Note that the shift value will be a positive value as it is the output of root squaring to the (squares of the difference between coordinates in East-West direction and North-South direction).

The result of bias components represented as shift value indicated at figure 13. It was expected to get less accuracy and highly bias components as the scale is 1:5000. The maximum bias components were ΔX= 0.36 m, ΔY=0.99 m, and Shift = 1.06 m. The mean value of the shift between digitized and extracted points along EF was about 0.48 m. These values represented a sufficient accuracy due to:

- The resolution of the image which is 1 pixel that represents 1 meter on the ground.
- The high quality output during the digitizing process. This is achieved due to the parallelism of the road edges.

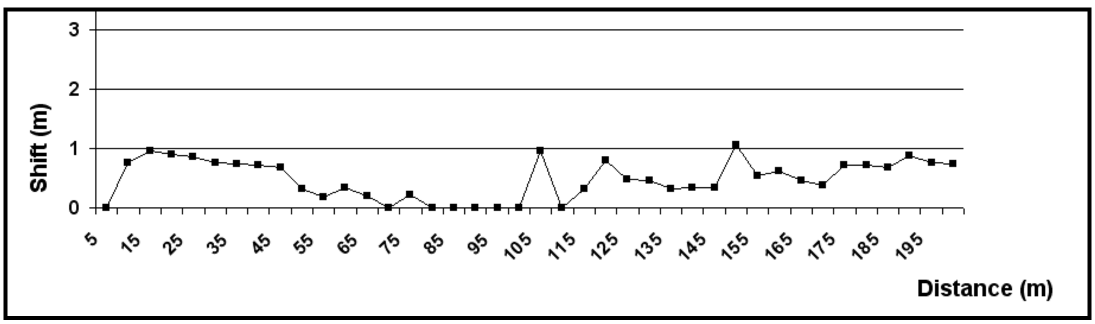

Fig. (13): Shift Variation for Line "EF"

As mentioned before the result of this stage is an isolated road network layer in raster format. The next procedure of this process is prepared for the purpose of vectorize this theme.

4.1.2. Seed Extraction "Vectorization" Procedure:

The binary image of the road theme that represent edge pixels was examined row by row from the top-left to the bottom-right by using the proposed algorithm. The output of this procedure is two dimensional (2D) array representing X and Y coordinates of the intersection points.

The developed package was run through test data. The total number of intersection points was around 80 points, the correct extracted intersection points were 78 points, while the failed points detected were 2 points and there were 5 wrong detected points as shown in figure (14).

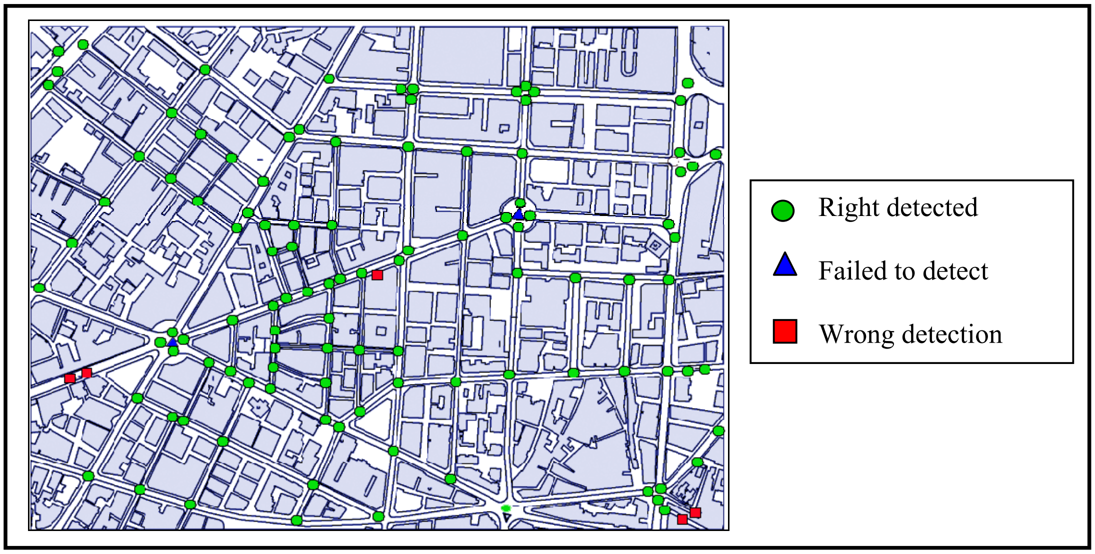

Fig. (14): Extracted Seeds Overlapped with Original Image and Extracted Road.

For the sake of examining the accuracy of the extracted nodes "intersection points", these nodes defined through digitizing process.

Figure (15) shows the shift value between the coordinates of the digitized intersection points and the extracted intersection points which calculated according to equations (7).

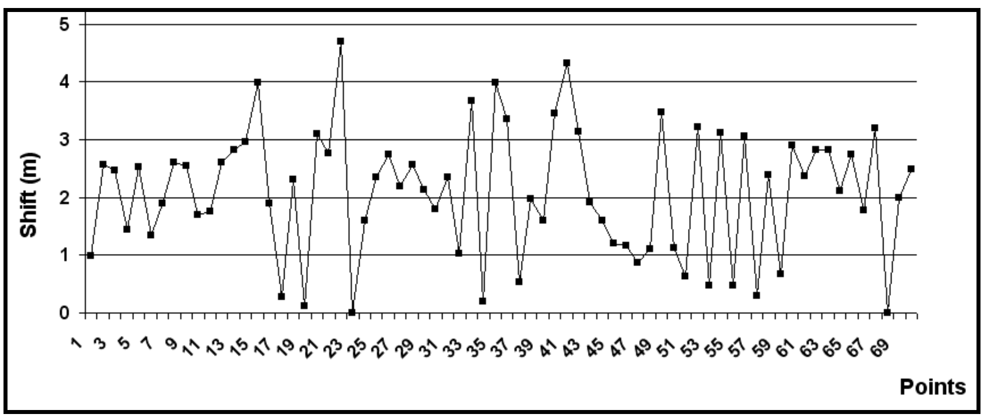

Fig. (15): Shift between digitized intersection points and extracted one

As observed from the graph, the maximum bias components of the shift value between the extracted and digitized points were ΔX=4.37 m , ΔY=3.98 m , and Result = 4.7 m and the mean value of this result was about 2 m.

It can be noticed that the maximum bias components are relatively high. This result can be due to many factors such as personal error during digitizing process, resolution of the image, geometry of the road networks.

5. CONCLUSIONS

The main objective of this research was to define reliable and automatic road extraction process. In other words, getting out the road segments from raster images is performed successfully, which is very useful as a step for preparing spatial data. This process consists of two stages.

The first stage is to extract street network from a raster image. This process is done using developed algorithms based on image processing theories "morphological operations".

The algorithm was tested through a study area that represented as raster image. This image characterized by high resolution image, homogenous color, high contrast and well defined shape of its road networks. The study area covers 1 km2 with scale 1:5000 and resolution 1 pixel /1 m. For the sake of examing the accuracy of this extracted road network , the reference of the test "ground truth" is considered as the digitized center line of the road features. The mean value of the diffference between extracted road and digitized one was about 0.48m.

The second stage is the vectorizing stage which is identifying the seeds "coordinates for the points of intersections of street network". This step was performed using developed algorithms based on image processing theories "pixel neighbors basics".

A novel algorithm was created to define seeds from the extracted road network which is the output of the first procedure. This includes identification of intersections and junctions. The algorithm was tested on the case study to all the map points from upper left corner to bottom right corner. The result showed that 97% of extracted intersection points were correctly extracted in 20 seconds.

In order to examine the accuracy of these extracted seeds ground truth verification was adapted .The mean value of the difference between digitized and extracted seeds was about 2 meters, which is sufficient for tracking purpose. It is observed that the failed seeds points were approximately in L-shape.

6. REFERENCES

[1] Dal-Poz, A. P., Vale, G.D., and Zanin, R. B. (2005): "Automatic Extraction of Road Seeds from High-Resolution Aerial Images", Anais da Academia Brasileira de Ciências, 77(3): 509-520, Brazil, January 5, 2005, www.scielo.br/aabc.

[2] Laptev I., Mayer H., Lindeberg T., Eckstein W., Steger C.,and Baumgartner A. (2000): "Automatic extraction of roads from aerial images based on scale space and snakes ", Machine Vision and Applications © Springer-Verlag 2000.

[3] Pratt, W. K. (2002): "Digital Image Processing", Third Edition, John Wiley and Sons, Inc., New York.

[4] Nixon, M. S., and Aguado, A. S. (2002): "Feature Extraction and Image Processing", First edition, Planet tree, London, England, 2002.

[5] Chanda, B., and Majunder, D. D. (2003): "Digital Image Processing and Analysis", Electronics and Communication Sciences unit, Indian Stastical Institute, Prentice-Hall of India, 2003.

[6] Russ, J. C. (1999): "The image Processing Hand Book", Third Edition, CRC press, IEEE press, 1999.

[7] MathWorks, Inc. (1999): "MATLAB Image Processing Toolbox", The MathWorks, Inc. ,Natick, MA , January 1999.

[8] Smith S. W. (1997): "The Scientist and Engineer's Guide to Digital Signal Processing", copyright © 1997-2007 by California Technical Publishing, 1997.

[9] Gonzalez R. C. and Woods R. E. (2002): Digital Image Processing, Second Edition, Prentice Hall, Upper Saddle River, New Jersey, 2002.

[10] El-Qadi Heba. (2007): "Towards Real Time Vehicle Navigation System", M.Sc. Thesis, Cairo University, Public works department, 2007.