An automatic system for urban road extraction from satellite and aerial images

S. IDBRAIM [(*)(**)] , D. MAMMASS[**] , D. ABOUTAJDINE* & D. DUCROT***
* GSCM-LRIT laboratory, Faculty of Science - University Mohammed V Agdal- Rabat- Morocco
** IRF –SIC Laboratory, Faculty of Sciences, Ibn Zohr University – Agadir - Morocco
*** CESBIO, University Paul Sabatier-Toulouse -France
soufianeidbraim@yahoo.fr

Abstract: - we present in this paper an automatic system of urban road extraction from satellite and aerial imagery. Our approach is based on an adaptive directional filtering and a watershed segmentation. The first stage consists of an automatic procedure which adapts filtering of each block band to the dominant direction(s) of roads. The choice of the dominant direction(s) is made from a criterion based on the calculation of a factor of direction of detection. The second stage is based on watershed algorithm applied to a Shen-Castan gradient image. This process provides a decision map allowing correcting the errors of the first stage. A ratio of surface on perimeter is used to distinguish among all segments of the image those representing probably roads. Finally, in order to avoid gaps between pieces of roads, the resulting image follows a treatment, based on proximity and colinearity, for linking segments.

The proposed approach is tested on common scenes of Landsat ETM+ and aerial imagery of the city of Agadir in Morocco. The experimental results show satisfactory values of completeness and correctness and are very prominsing.

Key-Words: - Road extraction; Satellite and aerial imagery; Urban areas; Adaptive directional filtering; Segmentation; Grouping; Evaluation.

1 Introduction

Automatic road extraction from remotely sensed imagery has been an active area of research in computer vision and digital photogrammetry for over two decades. A number of semiautomatic and automatic methods and algorithms for road extraction have been developed. Conventional methods of road extraction usually consist of three main steps: road finding, road tracking and road linking. In road finding, local properties of the image are tested and road candidates are found using certain criteria. The detected road candidates are then traced to form road segments. The separated road segments are finally linked to generate a road network using geometric constraints. In semi-automatic extraction. Many researches have been conducted for this purpose based on geometrical constrained template matching [3], active contours or snakes [6,8], fuzzy set and morphological operators [13,14,22]. In automatic extraction, numerous approaches have been developed [1,2,3,4,5,10]. Most of them concern the extraction of roads in rural areas. Compared to the relatively high number of research groups focussing their activities on high-resolution imagery, there are many approaches designed to process satellite or low-resolution aerial images which generally describe roads as curvilinear structures [16,19,21], and only few groups work on the automatic extraction of roads in urban environments [7,11]. Since many tasks related to automatic scene interpretation are involved.

In this paper, we present an automatic system for urban road extraction from low-resolution multi-spectral satellite image. The system runs automatically without the need of preselected points, although some algorithm parameters must be set. In our approach, the road extraction system is based on an adaptive directional filtering and segmentation based on watershed algorithm applied to a Shen-Castan gradient image. It is tested on LANDSAT ETM+ imagery of the City of Agadir in Morocco. In this study, the PAN data are fused with the MS data. Quantitative evaluations of the extracted road from representative test sites are reported. The quality and accuracy of the road extraction is greatly improved.

This paper is organized as follows: in section 2, we present an overview of the integrated system and

the preprocessing techniques. In Section 3, the road extraction methods used in each block of the system are described and illustrated by intermediate results. In Section 4, a numerical evaluation of the results currently achievable with our system is given. Finally, conclusions are given in Section 5.

2 The proposed system overview

2.1 The flow chart

Because of the complex nature of road extraction in urban environments, our conceptual workflow incorporates three blocks. The first block is devoted to an adaptive directional filtering, the second is related to a segmentation process and in the third block, the results from the two parallel blocks are combined together and completed to generate the extracted roads image.

Before starting these procedures, a preprocessing step is needed to improve image quality and to mask specific image pixels.The workflow outlining steps of our system is shown in Fig. 1

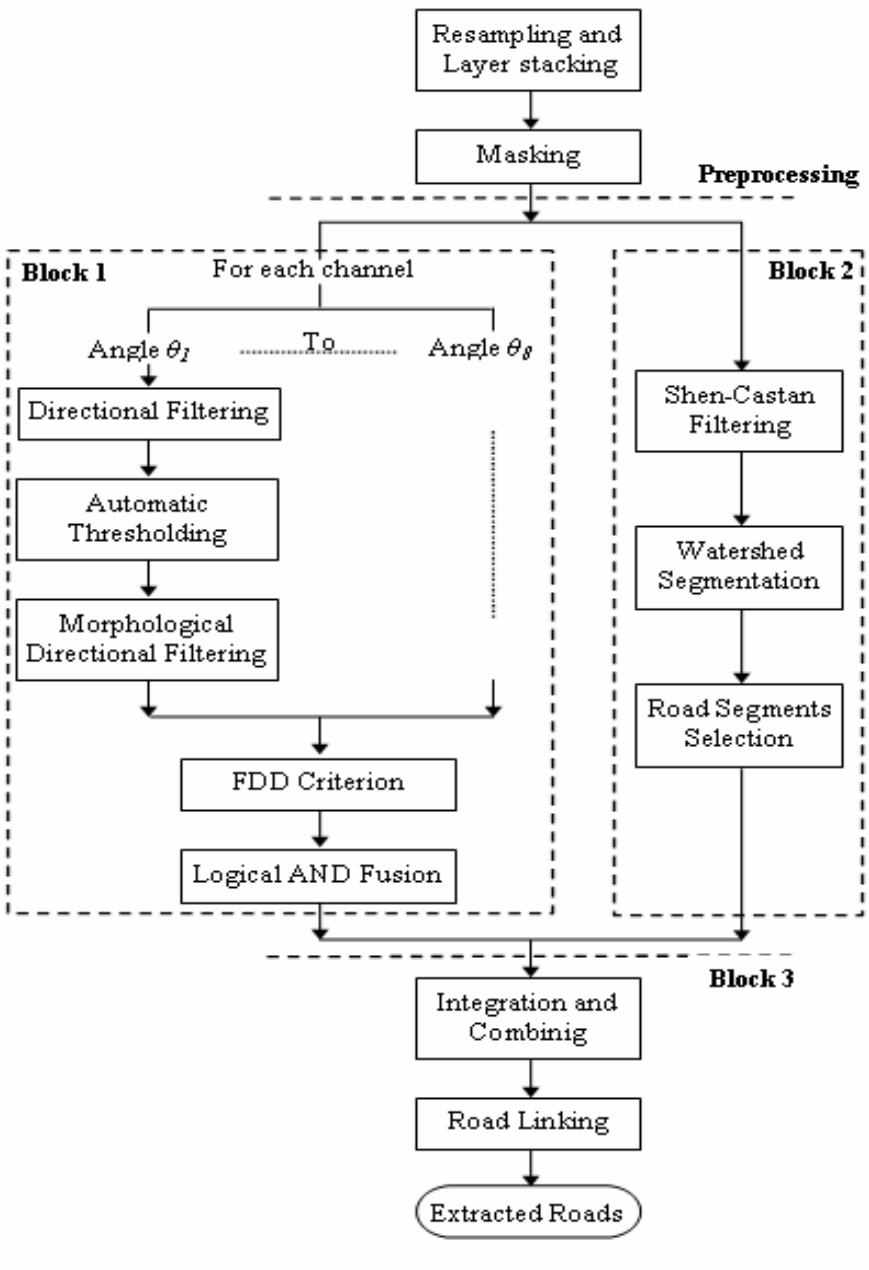

Fig. 1. Processing flow chart of the road extraction system

2.2 Preprocessing

The preprocessing gathers the resampling process and the selection of the scene of interest. It is designed to improve the resolution of the image and to mask the urban area.

In this study, the satellite image used is a multi band LANDSAT ETM+ , representing the zone +29, defined by the coordinates (Path/ Row: 203/39) and acquired in 19 March 2002, date corresponding to a cloud-free scene. The image has a resolution of 28.50 m in MS band 1, 2, 3, 4, 5, and 7, and a resolution of 14,25 m in PAN band. The PAN band is fused with the six MS bands using a color normalization method [18] with the "Nearest Neighbor" resampling method implemented in ENVI 3.5 to generate a six-band pan-sharpened multi-spectral (PS-MS) image with 14, 25 m spatial resolution. Therefore, we select a subset rectangle surrounding the city of Agadir.

It is beneficial to mask out vegetation and water (sea and river) in the PS-MS image in order to reduce computational load and confusion. So, the Normalized Difference Vegetation Index (NDVI) [15] definded by equation (1) is used as a sensitive indicator for the study of global and regional land cover changes.

$$\mathrm{NDVI} = \frac{(\mathrm{TM4} - \mathrm{TM3})}{(\mathrm{TM4} + \mathrm{TM3})} \tag{1}$$

where TM4 is the reflectance value of the near-infrared channel, and TM3 is the reflectance value of the red channel. Vegetation has a high reflectance in the TM4 channel and a low reflectance in the TM3 channel. Consequently, the NDVI equation produces values in range of -1.0 to 1.0, where increasing positive values indicate increasing green vegetation and negative values indicate non-vegetated surface features. After the tests with different thresholds vegetated areas with an NDVI higher than 0.2 were masked out.

Similarly, to hide the sea and the river, we implemented a band math operation which is the opposite of the redness index definded by:

$$\frac{(\mathrm{TM2} - \mathrm{TM3})}{(\mathrm{TM2} + \mathrm{TM3})} \tag{2}$$

Usually, water bodies have a high reflectance in the TM2 channel and a low reflectance in the TM3 channel, thus, areas with values higher than 0.15 were masked out.

3. Description of the methodology of the system blocks

3.1 The adaptive directional filtering block

The aim of this step is to perform a directional filtering in order to delineate the road contours and to facilitate the extraction step. This task may be done by filtering, both horizontally and vertically, the PS-MS image. The major problem is that roads are oriented in different directions. The innovation in the present implementation consists of an automatic procedure to adapt, for each band, the direction of filtering on the basis of the predominant direction of the roads or the dominants directions if there is more than one. This method is mainly based on linear directional convolution filters with the kernel:

$$\begin{pmatrix} -(\cos\theta+\sin\theta) & -\sin\theta & (\cos\theta-\sin\theta) \\ -\cos\theta & 0 & \cos\theta \\ -(\cos\theta-\sin\theta) & \sin\theta & (\cos\theta+\sin\theta) \end{pmatrix} \tag{3}$$

where θ is the angle of orientation. However, in our study, the road network is not complicated enough, so it is not necessary to know the direction of the road with great precision. So we consider only the following eight directions (Fig. 2) with a step of 22.5 °:

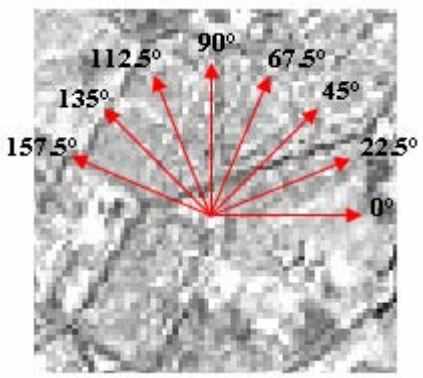

Fig. 2. The used directions in the filtering process

On each direction θ, a directional filtering by blocks on all spectral channels is done and the filtered images are automatically binarized. The choice of threshold for each image channel-direction is based on a method using the grayscale histogram. This method begins with a smoothing with a simply convolution 1D of the histogram since the histogram of the image does not lend itself to a direct search of local minima (as it is not smooth), which introduces a large number of local minima. To make it smooth, a filtering is often necessary. In our study, we successively used a median filter and a mean filter.

Once the histogram is filtered, the local minima can be detected. Formally, it is about finding all points where the derivative of the histogram is null. To distinguish only thresholds, it is sufficient to identify locations where the derivative goes from the negative values to the positive values. In the case of roads extraction, view the grayscale of roads, we look at the dark regions (close to the black), so we take pixels of the image whose gray levels are between [0… S2], where S2 is the second automatically found threshold. We have taken the second threshold in order to maximize the number of the retained roads.

To promote the reconstruction of linear elements in the binarized images, we apply a morphological directional filtering of the type opening with the structuring element line with the size of five pixels, and according to the preceding orientations.

Thereafter, we apply a criterion to determine the predominant or the dominant directions. This criterion is based on the calculation of a factor of direction detection (FDD) which gives us the percentage of pixels that are arranged in a linear way comparing to the image size. The selected images are those with an FDD higher than the average FDDs of all directions. These images will be fused with the logical operator AND to have a binary image of the band concerned. Similarly, an operation of fusion with a logical AND is made on all the six bands to have a result binary image.

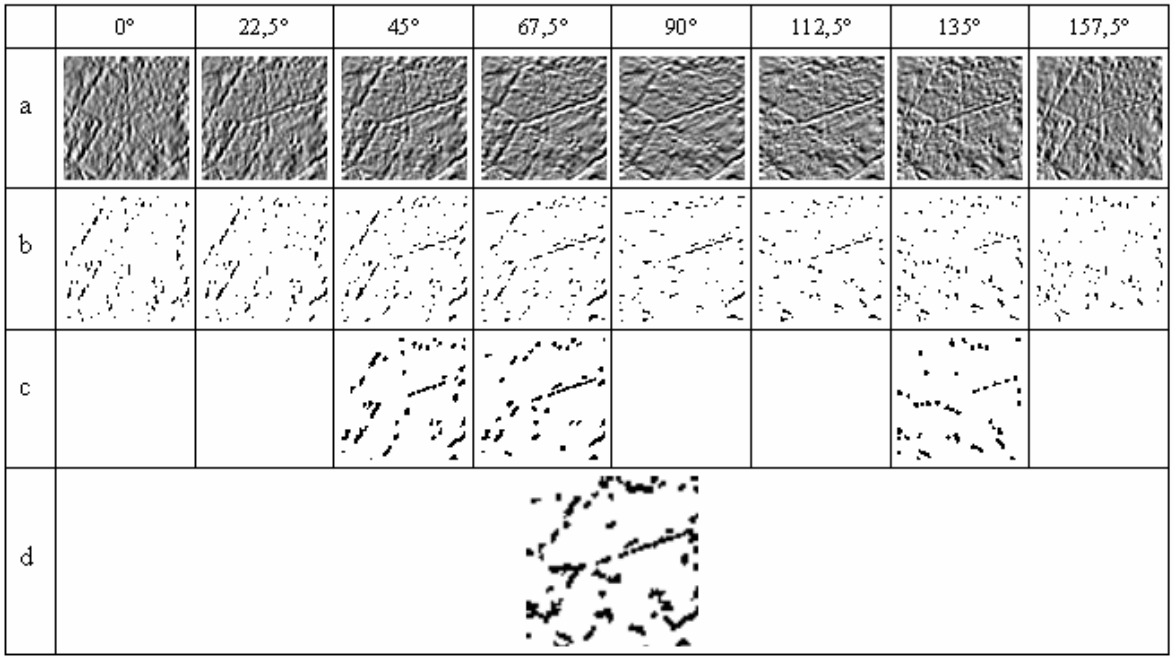

Fig. 3. Directional filtering steps applied to one channel subset: (a) Directional filtering; (b) Automatic thresholding; (c) Dominant directions; (d) Result fusion.

Angle	0°	22,5°	45°	67,5°	90°	112,5°	135°	157,5°	FDDs mean
Band 1	0,042857	0,22857	1,2857	0,25714	0	0,014286	0,28571	0	0,26428288
Band 2	0	0,2	1,3143	0,42857	0	0,11429	0,71429	0,1	0,35893125
Band 3	0	0,2	1,4143	0,4	0	0,042857	0,35714	0	0,30178713
Band 4	0	0,27143	1,5857	0,32857	0	0,085714	0,21429	0,071429	0,31964163
Band 5	0,042857	0,28571	1,3286	0,44286	0	0	0,54286	0,11429	0,34464713
Band 6	0	0,24286	1,5	0,44286	0	0,042857	0,67143	0,071429	0,3714295

Table 1: Calculation of the Factor of Direction Detection (FDD)

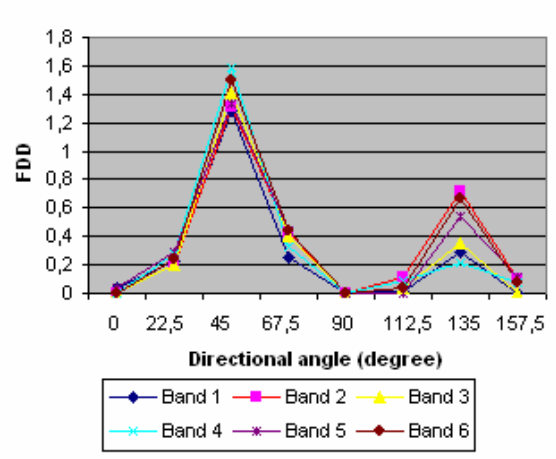

Fig. 4. Illustration showing the FDD of a subset in all channels

3.2 The segmentation process block

3.2.1 The Shen-Castan filtering

For the segmentation, we calculate the image module gradient by the Shen-Castan operator [9], which is based on the difference of weighted averages exponentially. The equation of this filter is:

$$h(x) = ce^{(-\alpha|x|)} \tag{4}$$

c is chosen so as to normalize the filter:

$$c = \frac{1-e^{-\alpha}}{1+e^{-\alpha}} \tag{5}$$

The parameter α determines the width of the filter, more α is small more the smoothing is important. The derivation filter is written:

$$h'(x) = \begin{cases} d\, e^{(-\alpha|x|)} & \text{if } x \geq 0 \\ -d\, e^{(-\alpha|x|)} & \text{else} \end{cases} \tag{6}$$

Likewise, d is chosen for normalising the filter:

$$d = 1-e^{-\alpha} \tag{7}$$

The implemented program calculates the module of horizontal and vertical components. As a multi-channel image, the contributions of different channels are averaged, also an automatic normalization by the factor d such that 0.1% of pixels to be truncated is done.

3.2.2 Watershed segmentation

After, we pass to extract closed and squelettized contour from the image of powers contours known as the gradient image produced by the Shen-Castan filtering. The program will perform watershed algorithm [17] with a simple threshold. It is a method, resulting from the mathematical morphology, systematic and effective to provide directly squelettized and closed contours. If we consider the image of powers contours as a surface in the 3D space, we can use the terminology of geography which defines the watershed as the limit between two catchment basins. The method consists in simulating a filling of the basins of the image of powers contours beginning with the local minima, until the meeting of water of the neighbouring basins which forms the watershed. At this place, we construct a dam to announce the presence of a contour. The threshold of segmentation causes to fill the basins until a height equal to the threshold without constructing the dams (hence the contours), then we repeat the usual process. For a threshold of zero, we find the classical watershed algorithm.

The implemented program allows a maximum of 256 levels of immersion, which means 256 possible thresholds segmentation (0-255).

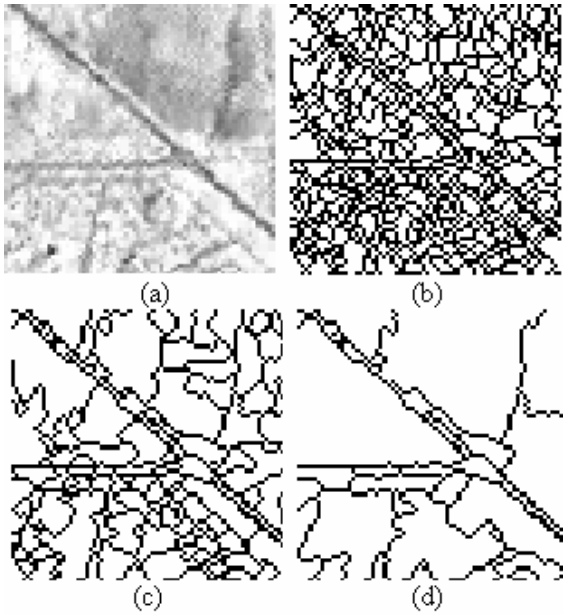

Fig. 5. Watershed segmentation thresholding: (a) test subset; (b) 2; (c) 12; (d) 22.

After a test of several thresholds, we retained a threshold which enabled us to have a better fragmentation of the image in small segments from which we will extract the closed segments representing the pieces of roads according to the next methodology:

Firstly, we labellise all closed segments. Thereafter, we calculate for each segment the surface and the external perimeter. Seen the morphology and the geometric form of the roads, we can say that raods belong to a class with a very close ratio surface on perimeter.

Fig. 6 illustrates principle in a simple way this, there are three segments which have the same surface S=12 but different perimeters. We distinguishes that the segment (a) which has a form similar to that of road, has the greatest surface on perimeter ratio.

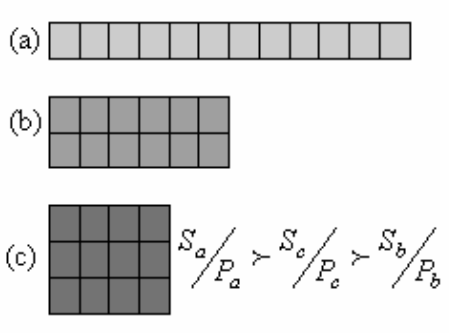

Fig. 6. Examples of surface on perimeter ratio.

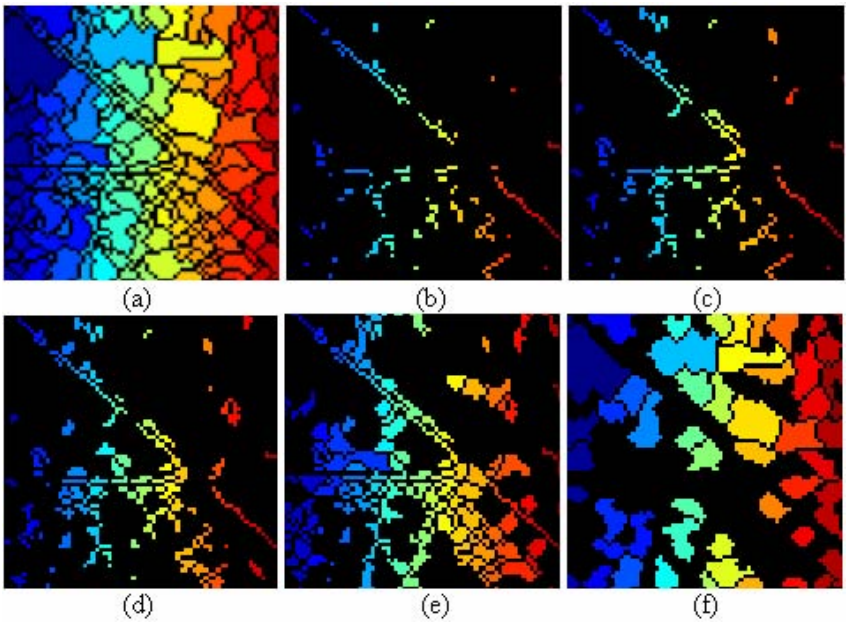

Fig. 7. Different thresholds of the surfaces on perimeter ratio: (a) original labelled image ; (b) <0,5; (c) <0,65; (d) <0,8; (e) <1,3; (f) >1,3.

In our procedure, we have tested different threshold values of the surfaces on perimeter ratio and noticed that more the threshold is high, the extraction of the roads segments is less interesting. So, we have chosen the threshold which gives us a satisfactory results around 0,5. The result image of this stage of segmentation will be used with the one retained in the stage of adaptive directional filtering to improve the extraction.

3.3 Combinig and linking block

3.3.1 Combinig method

We use the resulting image of segmentation as a decision map for the guard or not of the black pixels of the resulting image from the first phase. Indeed, we parcour the filtered image (Fig. 8. (a)) pixel by pixel and for each considered pixel as an object, we check the 3x3 neighbourhood of the segmented image (Fig. 8. (a)) if there is presence at least of one pixel of a road segment. It is about an intersection by neighborhood operation. As a result, this phase has enabled us to remove enough pixels, thing which will improve our extraction.

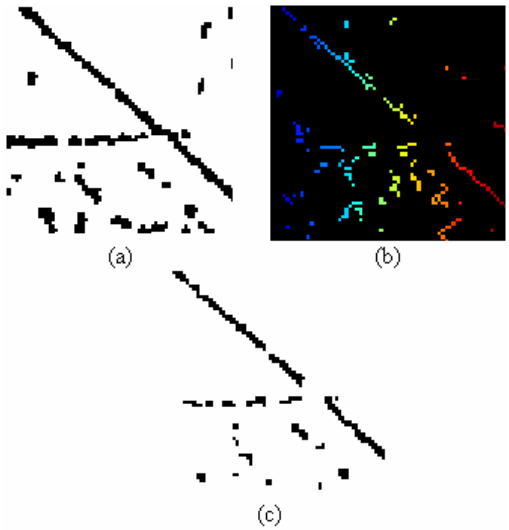

Fig. 8. Use of segmented image in directional filered image

3.3.2 Linking road segments

After identifying segments of roads, it remains to connect them because some roads can be broken into several segments. For humans, perceptual organization is the ability to readily group elements in an image based on various relationships between them such as similarity, proximity, colinearity, parallelism, and connectivity [12]. Here, perceptual grouping operation based on proximity and colinearity is applied on the original segments to group disconnected segments. The perceptual road-seed grouping is done using the following two rules:

1. If the endpoints of two separated segments are within some small distance, then these segments are connected.
2. If the endpoints of two segments are separated by a moderate distance, but the line segments are in the same direction, and there is an other segment in the same line with them, then these two lines are connected.

4. Results and evaluation

Our approach of road extraction is tested using the PS-MS LANDSAT ETM+ image of the City of Agadir in Morocco. A test site of aerial imagery is used after to evaluate the methodology with high-resolution image. Three representative test sites with a size about 1 km^2 were selected. Two test sites were selected from the LANDSAT imagery, which one is less difficult for road extraction than the second. The third test site was selected from high-resolution urban aerial image which include complex situations encountered in urban areas.

4.1 Evaluation criteria

The automatic results were compared with manually plotted data (reference data) according to the method developed in Wiedemann & al. [20]. The adopted measures to evaluate the quality of road extraction are defined as follows:

$$\text{Completeness} = \frac{\text{length of matched reference}}{\text{length of reference}} \quad (8)$$

$$\text{Correctness} = \frac{\text{length of matched extraction}}{\text{length of extraction}} \quad (9)$$

$$\text{Quality} = \frac{\text{length of matched extraction}}{\text{length of extraction} + \text{length of unmatched reference}} \quad (10)$$

The evaluation process developed for our study has a buffer with a width of three pixels along the reference data and extracted roads. If a pixel on the extracted roads is within the reference data buffer, then it is counted as a matched extraction. If a pixel on the reference data is within the buffer of the extracted roads, it is counted as a matched reference.

4.2 Evaluation results

The evaluation results of the road extraction system are summarized in Table 2 for the two LANDSAT test sites and in Table 3 for the aerial imagery test site. The road extraction results are shown in Fig. 9 and Fig. 10 respectively.

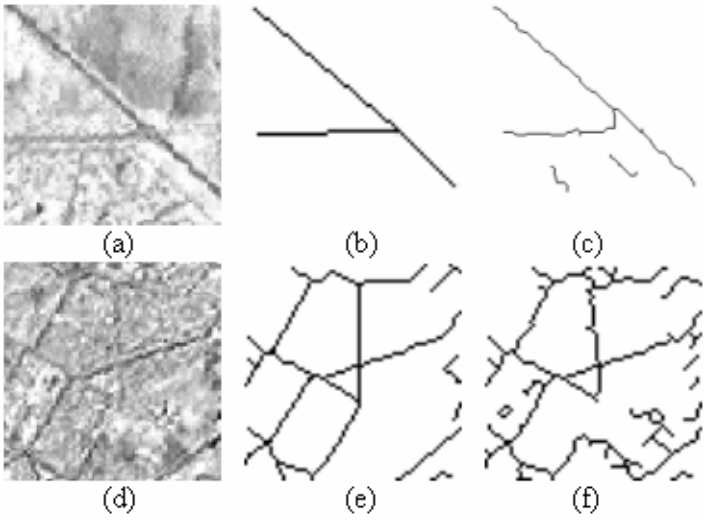

Fig. 9. Road extraction results from the satellite image: (a,d) original bands; (b,e) manually plotted road (reference) and (c,f) road extraction results.

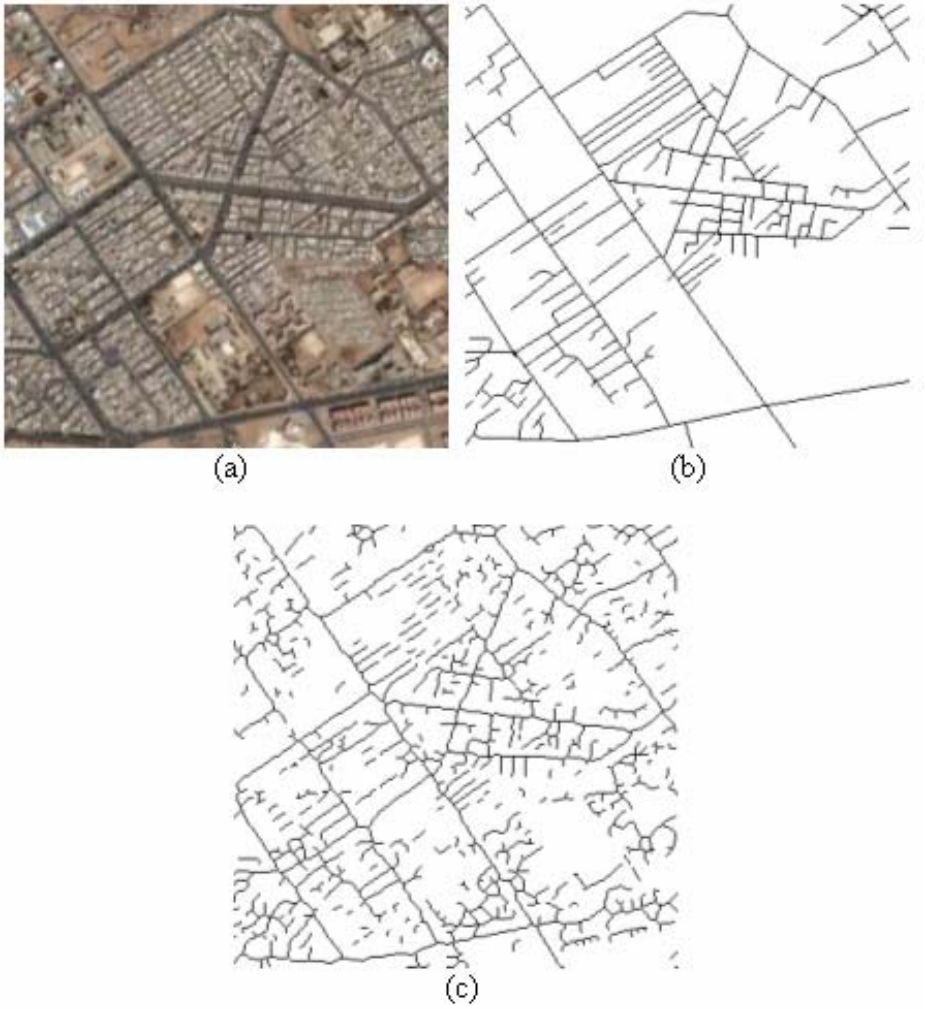

Fig. 10. Road extraction result from the aerial image test site: (a) original image; (b) manually plotted road (reference) and (c) road extraction result.

Table 2
Evaluation results of LANDSAT imagery test sites

	Completeness	Correctness	Quality
Test site 1	0.81	0.67	0.58
Test site 2	0.89	0.69	0.64

Table 3
Evaluation results of aerial image test site

	Completeness	Correctness	Quality
Test site	0.85	0.59	0.53

We can remark that major parts of roads have been extracted in spite of the high complexity of the scenes. We achieve completeness values that range between 81% and 89% for the LANDSAT imagery, and about 86% for the aerial imagery, and correctness values that range between 67% and 69% for the LANDSAT imagery, and about 59% for the aerial imagery because roads with small width in this type of image are confused with building roofs and shadow.

5. Conclusion

We present, in this paper, an automatic system for roads extraction based on an adaptive directional filtering and a watershed segmentation applied to a Shen-Castan gradient image.

The interest of the adaptive directional filtering process is to promote by block the dominant directions according to the calculation of a factor of direction detection (FDD) for eight tested directions. The results are fused to be use later in the system.

In the segmentation process, we need two thresholds, the first for the segmentation, which is based on watershed algorithm applied to a Shen-

Castan gradient image, and the second for the selection of a ratio of surface on perimeter that allows distinguishing among all segments of the image after a step of labilisation those represent probably roads. This process provides a decision map in order to correct the errors of the directional adaptive filtering stage. Thereafter a method of linking pieces of roads is applied.

Our system constitutes a proposal of good efficiency for application on satellite images with low-resolution. Similarly, the system can be improved to adapt the treatment to high-resolution image simply by introducing other layers, for example, a layer of detection the road bordes, a layer of skeletonization … etc.

We have tested the system on Landsat ETM + data and on an aerial image. Thus, the evaluation results show a good extraction for the first type of data and satisfactory extraction for the second type due to disturbances from some building roofs and linear shadows.

References:

[1] Agouris, P., Stefanidis, A. (Eds.), *Integrated Spatial Databases- Digital Images and GIS*., Springer-Verlag, 1999.

[2] Baltsavias, E., Gruen, A., Gool, L.V (Eds.), *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*. Balkema Publishers, 2001.

[3] Gruen, A., Kuebler, O., Agouris, P. (Eds.), *Automatic Extraction of Man-Made Objects from Aerial and Space Images*. Birkhauser Verlag, Basel, 1995.

[4] Gruen, A., Li, H., Road extraction from aerial and satellite images by dynamic programming, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol.50, No.4, 1995, pp. 11-20.

[5] Gruen, A., Baltsavias, E., Henricsson, O. (Eds.), *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkhauser Verlag, Basel, 1997.

[6] Gruen, A., Li, H., Semi-automatic linear feature extraction by dynamic programming and LSB-snakes. *Photogrammetry. Eng. Remote Sensing*, Vol.63, No.8, 1997, pp. 985–995.

[7] Hinz, S., Baumgartner, A., Automatic extraction of urban networks from multi-view aerial imagery, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol.58, 2003. pp. 83-98.

[8] Idbraim, S., Mammass, D., Aboutajdine, D., Extraction des routes à partir d'images LANDSAT ETM+ par contour actif, *In Proc. SETIT International Conference: Sciences of Electronic Technologies of Information and Telecommunications*, Tunisia, 2007.

[9] Laligant, O., Reynaud, N., Truchetet, F., Generalization of Shen-Castan and Canny-Deriche filters, *Proc. SPIE, Intelligent Robots and Computer Vision XVII: Algorithms, Techniques, and Active Vision*, Ed, Vol.3522, 1998, pp. 54-65.

[10] Mena, J.B., State of the art on automatic road extraction for GIS update: a novel classification, *Pattern Recognition Letters*, Vol.24, 2003, pp. 3037-3058.

[11] Mena, J.B., Malpica, J.A, An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery, *Pattern Recognition Letters*, Vol.26, 2005, pp. 1201-1220.

[12] Mohan, R., Nevatia, R., Perceptual organization for scene segmentation and description, *IEEE Transactions of Pattern Analysis and Machine Intelligence*, Vol.14, No.6, 1992, pp. 616–635.

[13] Soille, P., Talbot, H., Directional morphological filtering, *IEEE Transactions of Pattern Analysis and Machine Intelligence*,Vol.23, No.11, 2001, pp. 1313-1329.

[14] Soille, P., Pesaresi, M., Advances in mathematical morphology applied to geoscience and remote sensing, *IEEE Transactions Geoscience. Remote Sensing*, Vol.40, No.9, 2002, pp. 2042-2055.

[15] Townshend, J.R., Justice, C.O., Analysis of the dynamics of African vegetation using the normalized difference vegetation index, *International Journal of Remote Sensing*, Vol.7,1986, pp. 1435-1446.

[16] Trinder, J.C., Wang, Y., Automatic road extraction from Aerial images, *Digital Signal Processing*, Vol.8, 1998, pp. 215-224.

[17] Vincent, L., Soille, P., Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations, *IEEE Transactions of Pattern Analysis and Machine Intelligence*, Vol.13, No.6, 1991, pp. 583-598.

[18] Vrabel, J., Multispectral imagery advanced band sharpening study, *Photogrammetry. Eng. Remote Sensing*, Vol.66, No.1, 2000, pp. 73-79.

[19] Wang, Y., Trinder, J., Road network extraction by hierarchical grouping, *International Archives of Photogrammetry and Remote Sensing*, vol.33, Part B3/2, 2000, pp. 943–950.

[20] Wiedemann, C., Heipke, C., Mayer, H., Empirical evaluation of automatically extracted road axes, *In: CVPR Workshop on Empirical Evaluation Methods in Computer Vision*, California, 1998, pp. 172–187.

[21] Wiedemann, C., Ebner, H., Automatic completion and evaluation of road networks, *International Archives of Photogrammetry and Remote Sensing*, vol.33, Part B3/2, 2000, pp. 979–986.

[22] Zhang, C., Murai, S., Baltsavias, E., Road network detection by mathematical morphology, *In: Proc. ISPRS Workshop on 3D Geospatial Data Production: Meeting Applicat. Requirements*, Parıs, 1999, pp. 185–200.