Fully Automatic Road Network Extraction from Satellite Images

Onur Tuncer, *Member, IEEE*

***Abstract*— In this paper a fully automatic road detection algorithm is introduced. It comprises of pre-processing the image via a series of wavelet based filter banks and reducing the yielding data into a single image which is of the same size as the original optical grayscale satellite image, then utilizing a fuzzy inference algorithm to carry out the road detection which can then be used as an input to a geographical information system for cartographic or for other purposes that are in need. We use a trous algorithm twice with two different wavelet bases in order to filter and de-noise the satellite image. Each wavelet function resolves features at a different resolution level associated with the frequency response of the corresponding FIR filter. Resulting two images are fused together using Karhounen-Louve transform (KLT) which is based on principal component analysis (PCA). This process underlines the prominent features of the original image as well as de-noising it, since the prominent features appear in both of the wavelet transformed images while noise does not strongly correlate between scales. Next a fuzzy logic inference algorithm which is based on statistical information and on geometry is used to extract the road pixels.**

***Index Terms*— road extraction, satellite imagery, fuzzy logic**

I. Introduction

THIS paper deals with a fully automatic road detection algorithm. Road detection algorithms can be classified into two major groups; semi-automatic and automatic. The first approach necessitates that the user specify some initial conditions usually in the form of seed points entered manually by a human operator thorough some graphical user interface (GUI). Examples of these include and are not limited to differential snakes, graph connection etc. The other approach which is fully automatic on the other hand does not require input from an operator and works on its own. In this paper a fully automatic approach is presented. Suggested approach comprises of pre-processing the image via a series of wavelet based filter banks and reducing the yielding data into a single image which is of the same size as the original satellite image, then utilizing a fuzzy inference algorithm to carry out the road detection whose output can then be used as an input to a geographical information system (GIS) for cartographic or other purposes for that matter.

We use "a trous" algorithm twice with two different wavelet bases in order to filter and de-noise the satellite image. Each wavelet function resolves features at a different resolution level associated with the frequency response of the corresponding FIR (finite impulse response) filter. Resulting two images are fused together using Karhounen-Louve transform (KLT) which is based on principal component analysis (PCA). This process underlines the prominent features of the original image as well as de-noising it, since the prominent features appear in both of the wavelet transformed images while noise does not correlate well between high and low resolution scales as it lacks coherence. Principal component analysis is a powerful statistical tool commonly used for pattern recognition or data compression. This approach takes contextual spatial information into account and exploits the complementary characteristics of the passive optical and/or SAR data. Furthermore it is also possible to use more than two images to begin with and fuse them into a single one in this step without any loss of generality if doing so were justified to show better performance at the expense of computational time. On this image obtained thorough wavelet filtering and KLT, road detection is carried out using a fuzzy-logic inference algorithm. Linguistic variables used for this task are mean, standard deviation which are computed within a 5x5 pixel size image window and also another linguistic variable based on geometry. The inference algorithm then classifies each pixel as road or non-road with regard to the fuzzy inference rules yielding in a binary image.

Satellite images used in this paper are passive optical images which are converted to gray-scale representation prior to processing. Nevertheless, same process can be used for SAR (synthetic aperture radar) microwave images as the first step in the algorithm shall take care of the speckle noise (a multiplicative type of noise that often corrupts SAR images [1]) and shall highlight the prominent contextual features in the image at the same time. Results are presented for both urban and non-urban areas. It can be observed that the road network can be extracted with a high accuracy especially in the suburban areas.

Besides road detection fuzzy logic is a powerful and intuitive (i.e. in the sense of human friendliness) tool for identification of other features like runways, moving and/or stationary targets, other man-made objects on earth or any combination of these thereof, etc. For these other possibilities which are not in the scope of this paper another set of linguistic variables and inference rules would ultimately become necessary. Especially for military applications real-time image processing speed is of utmost importance [2]. For example an airborne platform equipped with SAR (and other necessary re-connesance tools) needs to process data as it gathers (i.e. at the sensor rate) for a fast and reliable decision making. Therefore these air platforms carry their own parallel

Dr. Tuncer is with the Louisiana State University Turbine Innovation and Research Center (TIER), Baton Rouge, LA 70808 USA (e-mail: otunce1@lsu.edu).

1-4244-1057-6/07/$25.00 ©2007 IEEE.

processing capabilities on board. Parallelizability of any detection algorithm as a direct consequence is of utmost importance should it be intended for military use as well. For each sequence in the proposed fully automatic algorithm very efficient parallel algorithms already exist in the literature. Wavelet transforms and PCA can be carried out very efficiently in a parallel manner. Fuzzyfication and de-fuzzyfication can also be carried out in a parallel manner by breaking the image into a suitable number of pieces (i.e. depending on the computing power available) and then ultimately processing each one of them individually.

a. Sample Road Network in an Urban Environment

b. Sample Road Network in a Non-Urban Environment

Figure 1 Unprocessed Optical Satellite Images

Figure 1 shows the unprocessed optical satellite images. Figure 1a shows an urban area whereas Figure 1b shows a non-urban area. Both images have 512 x 512 pixel resolutions. The spatial resolution on earth surface is about 1.5 m/pixel for each image with only slight variation from one another. This assueres that a fair comparison is given with regard to the results. Both images were obtained from Google Earth™ software. Both of the images are recorded within the resort town of Cesme which lies within the provincial boundaries of Izmir, Turkiye (see Figure 2 towards the lower left side).

Figure 2 Satellite View of Izmir Province

In Figure 1a one observes a road network that comprises of a large pixel width main highway, slightly narrower than that main roads and finally narrow side roads. Besides this varying in hierarchy (and thus width) road network, we observe blocks of housing, some vegetation cover, also a shoreline at the sea and land interface. Besides all these there are empty yards that have similar grayscale values as the roads themselves. Note that the city is not structured on an orthogonal grid and some of the inner city roads are not straight and have substantial curvature at some locations. On the other hand in Figure 1b one observes a superhighway with service roads and an "8" shaped turnpike all built on top of a suburban landscape. Landscape itself contains features like vegetation cover, some irrigated land, bare earth etc.

II. STEPS OF THE ROAD IDENTIFICATION ALGORITHM

In this section the processing algorithm is introduced and explained step by step. All computations in each step are carried out in MATLAB environment and thus coding is done in M language.

A. Wavelet Filtering

Since satellite images are finite energy (i.e. square integrable) functions a wavelet transform exists. Numerical algorithm utilized to deconvolve the normalized temperature images is referred as "algorithme à trous". This particular algorithm is due to the efforts of Holschneider *et al.* [3]. It is a translation invariant form of DWT (discrete wavelet transform) since it does not involve any decimation

(introduction of zeros) during down-sampling. Steps of the à trous wavelet algorithm (AWT) are briefly explained below.

The smoothed data c_j, which is of the same size as the original NxN image, at a resolution level j where $0 < j \leq \log_2 N$ can be calculated by the 2D wavelet kernel $\Psi(x_1)\Psi(x_2)$ (a tensor product of 1D kernels) translated at $k = (k_1, k_2)$ using the below convolution (1). Here j denotes the scale. In order to implement the algorithm N (pixel number) needs to be dyadic. Inner product operations in Hilbert spaces as in the following equation (1) are carried out in the computer by using FIR (finite impulse response) filters h, g associated with Ψ and Φ respectively. Also the 2D convolution can be achieved by separating the convolution process into two one-dimensional steps along the principal orthogonal directions x_1 and x_2. First a series of convolutions are applied along the rows of the image matrix and then the same process is repeated along the column-wise direction.

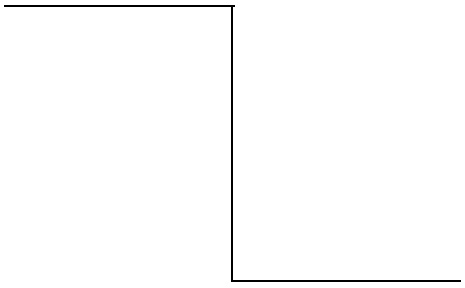
$$c_j(x_1, x_2) = \frac{1}{2^j}\langle I(x_1, x_2), \Phi\left(\frac{x_{1-k_1}}{2^j}\right)\Phi\left(\frac{x_{2-2}}{2^j}\right)\rangle \quad (1)$$

If the smoothing operation is stopped at resolution p, reconstruction of the original image I is achieved by adding the detail coefficients w_j of each level $1 \leq j \leq p$ to the smoothed image c_p (2). This linear simple additive reconstruction formula is the unique convenience of AWT along with its translation invariance and both account for its ease of interpretation. As a consequence of these two properties this algorithm is often employed in object detection. Only drawback of this algorithm is its redundancy, which requires (p+1) times larger storage space than the original image unlike the non-redundant counterparts such as the famous pyramidal Mallat algorithm [4]. Pyramidal algorithms are often better suited to image compression (e.g. JPEG 2000 format employs such compression schemes) and noise removal applications, rather than deterministic image processing, which better suits to the AWT algorithm. Here p is the resolution level at which the algorithm is stopped.

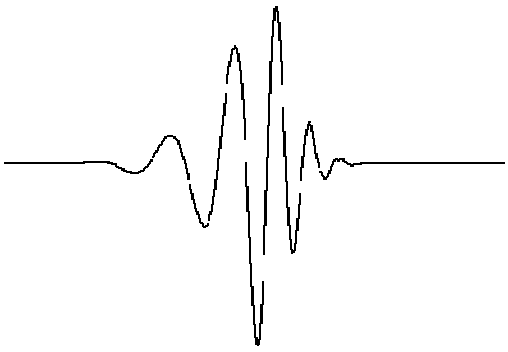
$$I(k_1, k_2) = c_p(k_1, k_2) + \sum_{j=1}^{p} w_j(k_1, k_2) \quad (2)$$

Above convolution (1) can be cast into the following discrete form (3). j denotes scale as emphasized earlier, x_1 ,x_2 are the pixel coordinates and k_1,k_2 are dummy integers which the summations are made onto and g is the corresponding FIR filter associated with the scaling function Φ.

$$c_j(x_1, x_2) = \frac{1}{2^j}\sum_{k_1}\sum_{k_2} I(x_1, x_2) g\left(\frac{x_1 - k_1}{2^j}\right) g\left(\frac{x_2 - k_2}{2^j}\right) \quad (3)$$



a. Haar Wavelet Function

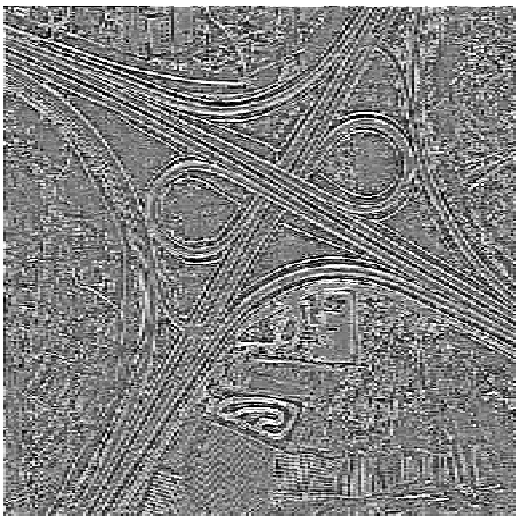


b. db8 Wavelet Function

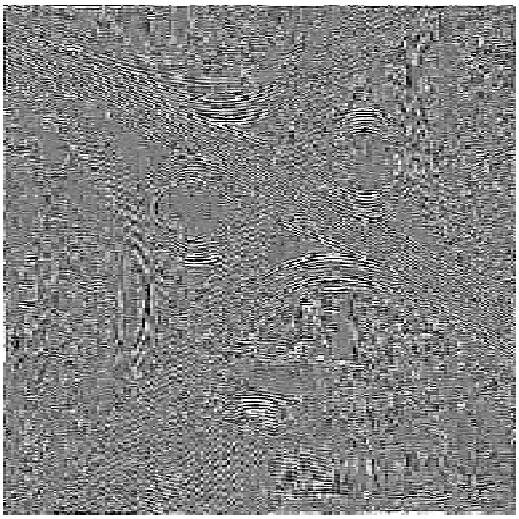
Figure 3 Wavelet Functions Used In the Study

Figure 3 shows the two wavelet functions used for implementing wavelet filtering. The first one is the Haar function which is also known as the db2 wavelet. Haar system is the unique one that satisfies bi-orthogonality, symmetry and compact support properties at the same time. It is also the earliest coined wavelet function. Haar system has a symmetric scaling function, an anti-symmetric wavelet function, a single vanishing moment and has finite support length in the interval [0,1]. The second one shown in Figure 3b is the db8 function. Each have associated FIR filter banks with their own frequency responses. Such effects will be examined while discussing the results of wavelet filtering within the context of the next figure.

Figure 4 shows the wavelet planes of the suburban image. Here a trous algorithm is truncated at p=4 level which gives adequate information at that point. The wavelet planes shown in Figure 4 are accumulation of $w_1 \ldots w_4$, that is $\sum_{j=1}^{4} w_j$. Note that due to the additive reconstruction property of the a trous algorithm this is equivalent to $I - c_4$. Therefore in implementation we continue with the computation of the coarse scale image using (3) until c_4 then subtract it from the original image. Note that in Figure 4a high frequency details are more vivid than in Figure 4b. This is due to the frequency response of associated FIR filters. Principal component analysis is used to fuse these two images together back into a single image.



a. Decomposed Image Using the Haar Basis Function



b. Decomposed Image Using the db8 Basis Function

Figure 4 Wavelet Planes of Suburban Image

B. Karhounen-Louve Transfrom (KLT)

After the wavelet transforms we have two images, let us call them $\widetilde{I_1}$and $\widetilde{I_2}$. Then we can proceed to define matrix X which has the information about these two images in its two rows (4). Operator "vec" denotes vectorization of matrices.

$$X = \begin{bmatrix} vec(\widetilde{I_1}) \\ vec(\widetilde{I_2}) \end{bmatrix}_{2xN^2} \tag{4}$$

Given X next step is to compute the Karhounen-Louve transform of X and collapse the information into half along its principal components [5].

$$Y = \mathbb{KLT}\{X\} \tag{5}$$

Next calculating the mean column vector as in (6),

$$u_j = \frac{1}{2}\sum_{i=1}^{2} X_{ij} \tag{6}$$

Then computing the mean subtracted matrix as follows (7),

$$B = X - u.1_{1x2} \tag{7}$$

Covariance matrix can then be calculated as follows,

$$C = \mathbb{E}\{B \otimes B\} \tag{8}$$

Next the matrix of eigenvectors V that diagonalizes the covariance matrix C. W is the first column of V which corresponds to the largest eigenvector. With a final projection step which reduces the dimension of data, the algorithm is complete. Y vector can be re-shaped into an NxN matrix afterwards.

$$Y = W^T.B \tag{9}$$

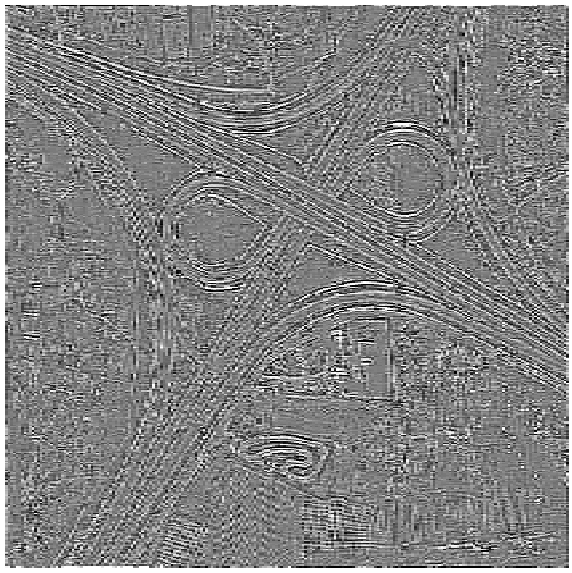


Figure 5 Fused Image by KLT

C. Fuzzy Logic Inference Algorithm

Mathematically a fuzzy set is defined as asset whose members have a degree of membership. $\mu(x)$represents the membership function which determines the level at which its argument x belongs to that particular fuzzy set (10). This fuction can take values within the [0 1] probability interval. Unlike classical set theory where x can either be an element of a set or not, fuzzy sets allow an ambiguous classification through a membership function [6]. Hence they are called fuzzy sets. By virtue of this fuzziness it can mimic well the human logical reasoning.

$$A = \{x, \mu_A(x)\}; x\epsilon X \tag{10}$$

First two linguistic variables which are used as inputs for the fuzzy logic classification are the statistical quantities of the processed image after the fusion step (see Figure 8). For each pixel a 5x5 pixel size window is centered around the pixel of interest and within this window mean and standard deviation, are calculated. This statistical approach is quite similar to the one described in [7].

Although statistical information can give useful insight about where the road pixels are located a fuzzy detector based

solely on those is somewhat hard to tune and even then the false alarm rate can be substantial to the point that where it can render this whole approach useless. However by combining this statistical radiometric information with geometric information, the false alarm rate (i.e. assigning road values to non-road pixels) can greatly be reduced [8]. In order to aid in this one can rely on the a priori information that the road segments are linear features in the image domain. Hough transform [9] is an excellent tool to detect the linear features of an image and is quite commonly used in computer vision applications. The resulting hybrid fuzzy approach can quite simply be formulated as follows (11). Therefore there is one more input variable based on the geometry of the image pixel.

IF (statistical radiometry) AND (geometry) THEN assign ROAD (11)

In order to assess the geometry information Hough transform is utilized as mentioned before. For Hough transform a line segment is defined in its parametric notation as first introduced by Duda and Hart [10],

$$x\cos\theta + y\sin\theta = \rho \tag{12}$$

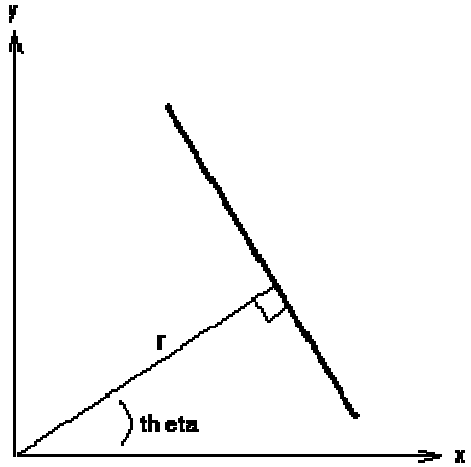


Figure 6 Illustration of a Line in (ρ, θ) Domain

where r denotes the length of a normal drawn from the origin to the line of interest and θ is the orientation of the line measured from the x-axis. Therefore the image from the (x,y) domain is transformed into (ρ,θ) domain. Within this domain one then looks for significant peaks which correspond to strong linear features of the image. Selecting these peaks the linear features can be reconstructed in the (x,y) domain through an inverse transformation. During this transformation pixels where linear features are identified are assigned a value of one and the others are assigned a zero value.

Below is the non-urban image (see Figure 7) in the Hough domain with the corresponding significant peaks marked with a square symbol. For the fuzzy inference geometry input instead of transforming the whole image into Hough space image is broken into windows and Hough transform is computed separately in each window which turns out to be a more efficient way.

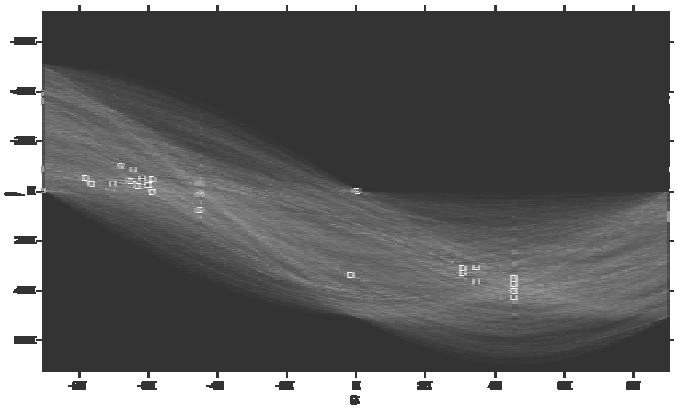


Figure 7 Suburban Image Transformed Into Hough Domain

In the next figure (see Figure 8) the schematic of the fuzzy logic system is shown with three inputs (mean, standard deviation, geometry) and one output (road status). Only 6 rules are used for decision making. Fuzzy network is of Mamdani type [11].Corresponding membership functions μ for some of the variables (13) are shown in the next figure (see Figure 9).

$$\mu_{(x_i)} = \left\{\mu_{x_i}^1, \ldots, \mu_{x_i}^k\right\} \tag{13}$$

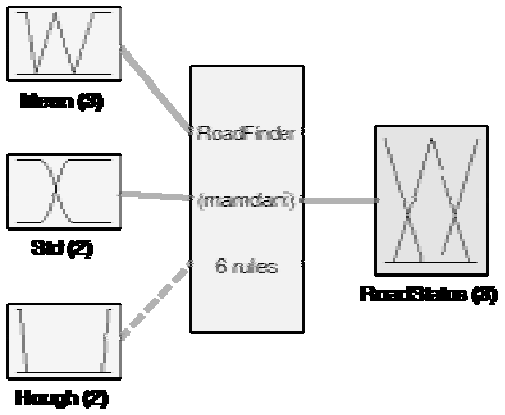


Figure 8 Schematic of the Fuzzy Logic System

Finally a set of rules have been devised in order to make the decision for the road extraction. Fuzzy inference from these rules is a two tier process. First tier is the computation of the "AND" part of the rules and the second tier is the calculation of the "THEN" part of the rules. Rules are structured as in the following equation (14).

IF (Mean is Average) AND (Std is Low) AND (Hough is Line) THEN assign ROAD (14)

Consequently, the output of each rule is a linguistic variable. Therefore it needs to be defuzzified to yield in a

numerical value [12]. This process is referred to as defuzzification. There exist a number of defuzzification schemes. Within the context of this paper center of mean defuzzification method is used (15).

Before deciding upon the rules and carrying out the computations one must map the output region into a fuzzy linguistic variable. Figure 9 demonstrates the output linguistic variable membership functions. There are three triangular membership functions in this case; not road, road unlikely and road.

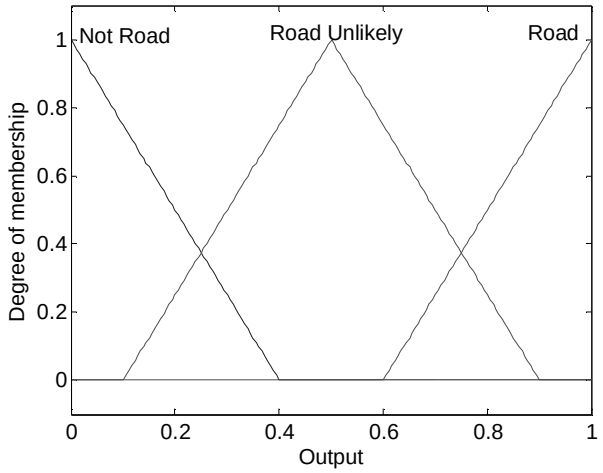


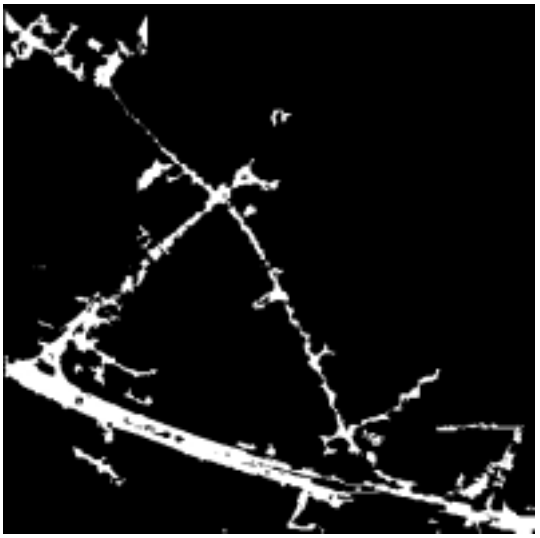
Figure 9 Output Variable Membership Functions

$$y = \frac{\sum_j \mu_y^j(w_j) w_j}{\sum_j \mu_y^j(w_j)} \tag{15}$$

After constructing and tuning the fuzzy logic network parameters, pixels in the original images can be segmented into road and non-pixels in order to extract the road network. These results are presented and an adequately thorough discussion is provided in the next section of article.

III. RESULTS & DISCUSSION

An algorithm for fully automatic road detection is presented. Results of the proposed algorithm have been shown for both urban and non-urban areas in Figure 10. These are the results are the results without any image thinning process being applied. Pixels identified as "road" are shown in white, whereas the "non-road" pixels are shown in black. As it can be seen from these images fully automatic classification process outlined in this paper performs with a high degree of accuracy in both urban and non-urban areas. However, there are some very important concerns particularly regarding the urban area classification which presents a more challenging classification problem than its suburban counterpart. As it can be seen from Figure 10a only roads with rather large pixel widths such as the main highway are recovered well. Main roads and intersections are also recovered with good accuracy but inner city roads which are narrow have not been recovered. In addition, some pixels belonging to rooftops of buildings were falsely identified as roads.



a. Urban Image

b. Suburban Image

Figure 10 Results of the Fuzzy Classification

On the other hand in Figure 10b one can see that most of the road pixels have been recovered successfully. Two superhighway roads are intersecting at a right angle with of them crossing on top of the other via a bridge and those two are linked with an "8" shaped turnpike, there are also a number of service roads connecting the freeways to other roads. Recovering the highway pixels was not much of a challenge at all as these are wide and straight features appearing in the image. For the turnpike there is substantial curvature in the geometry and therefore some of the pixels are missing. Same is true for one of the service roads. It is seen that the rural landscape and irrigated land is filtered out nicely and they do not falsely appear to be belonging to the road network for most of the image.

IV. CONCLUSION

In this paper an automated method for extracting road networks from satellite images is presented. It is observed that the image classification is carried out to a high degree of accuracy especially for the non-urban areas. In the urban areas

however, only major roads with larger pixel widths have been recovered. Moreover, the presence of buildings and other objects that have linear features made the road extraction somewhat more difficult compared with the non-urban case. In this case some pixels belonging to rooftops of buildings were wrongly identified as road pixels. For urban road extraction with high fidelity this method seems to require higher resolution at which the narrow and small roads shall appear with larger pixel areas. Another concern is since the fuzzy logic system is looking for linear features for road segments areas of high curvature in the roads have been problematic to recover. One good example is the "8" shaped turnpike that is in the suburban image where many of the pixels belonging to the road segment are left unidentified. This error is due to the high curvature of this feature.

Although grayscale optical satellite images are used for identification it is possible to extend this approach to SAR image processing for road or target detection. As mentioned earlier in the paper the approach introduced is structured such that it is quite possible to introduce very efficient parallelization in order to do the processing in real-time for other practical purposes such as target recognition.

Prospective work will be directed towards implementing the inverse transformation method of Kesidis and Papamarkos [13], which is expected to yield in better results especially for the detection of thinner line segments corresponding to narrow urban roads.

ACKNOWLEGEMENT

I would like to gratefully acknowledge financial support received from Louisiana Board of Regents during my post-doctoral studies at Louisiana State University Department of Mechanical Engineering.

REFERENCES

[1] C. Peng and A. Chen, "Speckle Noise Removal in SAR Image Based on SOT Structure in Wavelet Domain", Geoscience and Remote Sensing Symposium, 2001, IGARSS'01, IEEE International **7**, pp. 3039-3041

[2] M. S. Moore, "Model Integrated Program Synthesis for Real Time Image Processing", PhD Dissertation, Vanderbilt University, May 1997, Nashville , TN, USA

[3] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian, "A Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform," Wavelets, time-frequency methods and phase space, Springer-Verlag, Berlin, Germany, 1989

[4] S. Mallat, "A Theory of Multiresolution Signal Decomposition: The Wavelet Representation", IEEE Trans.Pattern Anal. Machine Intell., 1989, **11**, pp. 674-693

[5] Fukunaka, K., "Statictical Pattern Recognition", Elsevier Publishing, San Diego, 1990

[6] L. Zadeh, "Fuzzy Set", Information and Control, 1965, **8**,pp. 338-353

[7] J. Amini, "A Proposed Model for Segmentation of Spot Images", Proceedings of IAPRS **35**, 2004, pp. 1-3

[8] L. L. Yun, K. Uchimura, and S. Wakisaka, "Road Extracted Using Fuzzy Reasoning From Satellite Images", ASPRS Anuual Conference Proceedings, 5-9 May 2003, Anchorage, Alaska, USA

[9] Hough, P.V.C., "Method and Means for Recognizing Complex Patterns", US Patent No: 3069654, 1962

[10] Duda, R. O. and P.E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures", Comm. ACM, **15**, 1972, pp. 11-15

[11] Mamdani, E. H., "Applications of Fuzzy Algorithms for Control of Simple Dynamic Plant", Proceedings of IEEE, 1974, **121**, pp. 1585-1588

[12] "Fuzzy Logic Toolbox User Guide", Mathworks Inc., MA, USA, 2007

[13] A. L. Kesidis, N. Papamarkos, "On the Inverse Hough Transform", IEEE Transactions on Pattern Analysis and Machine Intelligence, **21**, 1999, pp. 1329-1343