# Rural road extraction from SPOT images based on a Hermite transform pansharpening fusion algorithm

Boris Escalante-Ramírez[a,b*] and Alejandra A. López-Caloca[b]

[a]Universidad Nacional Autónoma de México, Facultad de Ingeniería
Cd. Universitaria, México, D.F., 04510.
[b]Centro de Investigación en Geografía y Geomática "Ing. Jorge L. Tamayo", A.C.
Contoy 137, Lomas de Padierna, México, D.F., 14240.

## ABSTRACT

Roads are a necessary condition for the social and economical development of regions. We present a methodology for rural road extraction from SPOT images. Our approach is centered in a fusion algorithm based on the Hermite transform that allows increasing the spatial resolution to 2.5 m. The Hermite transform is an image representation model that mimics some of the more important properties of human vision such as multiresolution and the Gaussian derivative model of early vision. Analyzing the directional energy of the expansion coefficients allows classifying the image according to the local pattern dimensionality; roads are associated to 1D patterns.

**Keywords:** Road extraction, Hermite transform, Image fusion, Steerable filters, Pattern recognition

## 1. INTRODUCTION

Rural road extraction is considered one of the essential studies that help consolidate the strategic vision of regional development. It is particularly relevant as a fundamental element for the design and elaboration of an inventory of rural roads. The use of satellite images facilitates mapping and updating this task[1]. Among the existing methods for road detection in satellite images, we find pixel-oriented approaches, which associate each pixel in the image to a determined class according to its grey level. Implementations of these methods allow for relatively simple and quick solutions, but neglect the notions of proximity and connectivity in pixels. Region oriented methods explore the notion of connectivity in order to group image zones with similar statistical properties. This is the case of algorithms based on Markov Random Field theory (MRF)[2] whose main disadvantage is the long computational time required to obtain the solution.

Other methods use mathematical morphology schemes to study the shapes and contours of roads. The two main morphological operators are dilation and erosion. Dilation amplifies objects by filling up holes and joining disjoint regions, where erosion reduces the objects. These methods however do not normally behave well in bifurcations (split roads) and interrupted roads. Examples applied to radar images were described by Chanussot and Lambert in 1998[3].

On the other hand, methods based on multiscale and multiresolution analysis represent a good alternative for road extraction. Renaud Péteri & Thierry Ranchin (2004)[4] propose a method using multiresolution analysis (MRA) with an "á trous" wavelet transform (WT)[5]. Their extraction algorithm combines the use of specific active contours (snakes) with multiresolution analysis (MRA).

The Hermite transform is an alternative image representation model that mimics some of the most important properties of human visual perception, namely local orientation analysis and the Gaussian derivative model of early vision. This work introduces the Hermite Transform (HT)[6,7,8,9] as an efficient representation model for road extraction. Shift invariance assures that no artifacts are introduced during image analysis. Furthermore, the isotropic property (rotation) of the HT allows for the detection of oriented patterns such as edges and lines, which are important components of road

* e-mail: boris@servidor.unam.mx; web: verona.fi-p.unam.mx; phone: +52-55-56161719

Signal Processing, Sensor Fusion, and Target Recognition XVIII, edited by Ivan Kadar,
Proc. of SPIE Vol. 7336, 733619 · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.820330

structures. These patterns may be discriminated from other structures according to their local energy and orientation. The use of the locally-rotated HT provides a coefficient expansion set with high energy compaction, in such a way that few coefficients are needed to represent one-dimensional (1D) image patterns.

One of the advantages of the HT over the WT is the use of a free subsampling parameter limited only by the support of the analysis window, i.e., the only constraint for this parameter is that local analysis windows overlap with each other. This allows for the existence of decimated as well as undecimated HT decompositions, the latter being known for the property of shift invariance.

## 2. THE HERMITE TRANSFORM

The HT is a special case of polynomial transforms whose basis functions are derivatives of Gaussian functions. It was first introduced as a model for image representation by Martens[6]. The extension of this model to the multiresolution case was later formulated[9]. The HT uses overlapping analysis windows and projects the localized image onto a basis of orthogonal polynomials. The Gaussian window is the best option from a perceptual point of view and is in agreement with the scale space theory. The Gaussian window is separable into Cartesian coordinates and is isotropic, thus it is rotationally invariant. For discrete implementations a discrete polynomial transform may be considered. In this case a binomial analysis window is a good choice since it approximates the Gaussian function. Another important parameter is the window spread. Fine local changes are better detected with small windows, but on the contrary, representation of low resolution objects need large windows. We have chosen binomial windows of orders $N$=2 however, it is known that the binomial function of order $N$ approximates a Gaussian function with spread $\sigma = \sqrt{N/2}$ . $N$ also represents the maximum order of the transform coefficients since the binomial function has compact support. As mentioned before, the subsampling factor, this is, the distance between adjacent window functions is a free parameter. Shift invariant decompositions can be achieved with subsampling period equal to one, i.e., no subsampling. Larger subsampling periods account for decimated decompositions that are computationally more efficient.

The polynomial coefficients $L_{m,n-m}(p,q)$ are calculated by convolution of the original image $L(x,y)$ with the function filter $D_{m,n-m}(i,j) = G_{m,n-m}(-i,-j)V^2(-i,-j)$ followed by subsampling at positions *(p,q)*, i.e.,

$$L_{m,n-m}(p,q) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} L(i,j) D_{m,n-m}(i-p, j-q)\, di dj \tag{1}$$

where $G_{m,n-m}(i,j)$ are the set of orthogonal polynomials of order $m$ in $i$ and order $n-m$ in $j$ . For the case of the Gaussian window, the corresponding orthogonal polynomials are the Hermite polynomials:

$$G_{n-m,m}(i,j) = \frac{1}{\sqrt{2^n (n-m)!m!}} H_{n-m}\left(\frac{i}{\sigma}\right) H_m\left(\frac{j}{\sigma}\right) \tag{2}$$

where $H_n(i)$ denotes *nth* Hermite polynomial. $V(i,j)$ is the Gaussian analysis window.

Resulting Hermite coefficients are arranged as a set of $N$x$N$ equal-sized subbands, one coarse subband $L_{0,0}$ representing a Gaussian-weighted image average and detail subbands $L_{n,m}$ corresponding to higher-order Hermite coefficients, as shown in Fig. 1.

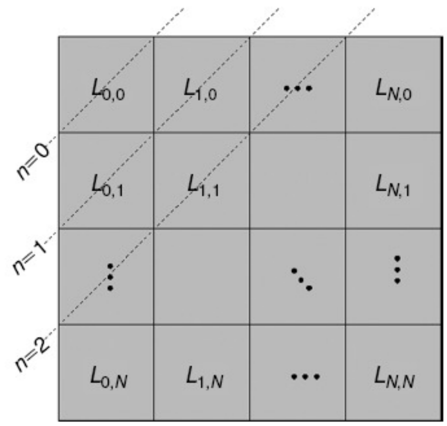


Fig. 1. A spatial representation of Hermite transform coefficients as a set of *NxN* subbands $L_{n,m}$. The subband size depends on the subsampling factor used in the expansion. Diagonals depict zero-order coefficients (n = 0), first-order coefficients (n = 1), etc.

### 2.1 Rotated Hermite transform.

The result of a Hermite transform is an overcomplete signal description, however, high energy compaction can be obtained through adaptively steering the transform[9,10]. The term steerable filters describe a set of filters that are rotated copies of each other. They can be constructed as a linear combination of a set of basis filters. Based on the steering property, the Hermite filters at each position in the image adapt to the local orientation content. This adaptability results in significant information compaction. The local rotation into the domain transform can be seen like a mapping of the expansion coefficients into a local coordinate system whose main axis corresponds to the direction of maximal signal energy. A new set of oriented Hermite coefficients $L^{\theta}_{m,n}(p,q)$ is then generated.

Furthermore, the steered HT offers a way to estimate the dimensionality of the underlying local image pattern, either 0D, 1D, or 2D. By analyzing relations between the oriented energies, it is possible to classify image patterns. In practice, the local gradient angle, calculated from the expansion coefficients as $\theta = \arctan L_{0,1} / L_{1,0}$, where $L_{0,1}$ and $L_{1,0}$ are the first order coefficients of the HT, can be an alternative estimator of the rotation angle $\theta$. This choice would imply $L^{\theta}_{0,1} = 0$.

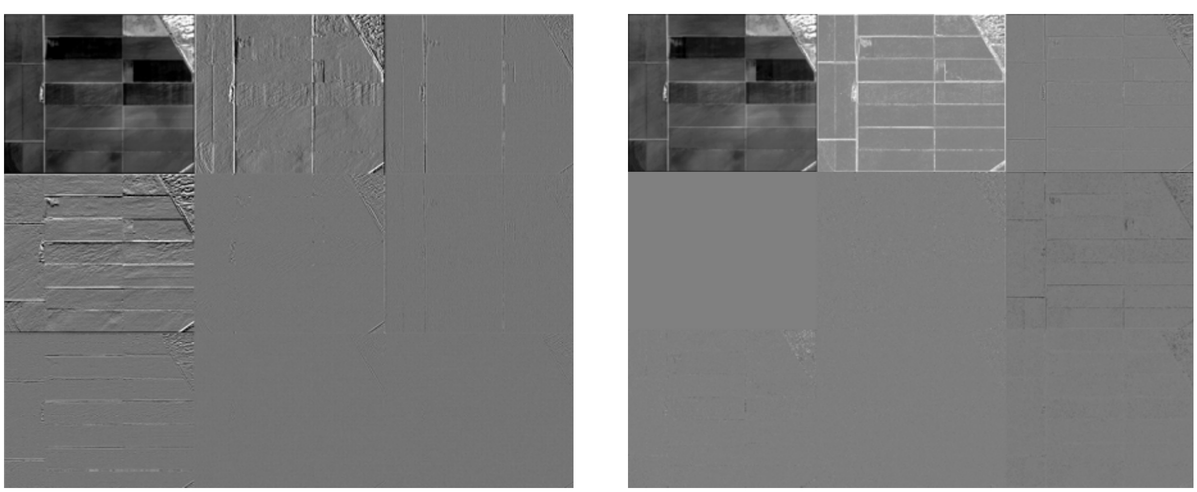

Fig2. Left, original Hermite transform, N=2. Right, rotated Hermite transform. Note that in this case $L^{\theta}_{0,1} = 0$

# 3. METHODOLOGY

## 3.1 Image fusion

In the case of SPOT images, we have observed that contrast between roads and background is often larger in band 1 than in all other bands, including the panchromatic band, however the low spatial resolution of spectral bands limits the task of fine road detection. We applied a pansharpening algorithm based on the HT in order to improve the spatial resolution of band 1. The fusion algorithm has been previously formulated and proved to preserve the spectral properties of the original band while improving considerably the spatial resolution[7,8]. The general framework for fusion through Hermite transform includes six steps:

(1) Perform the Hermite transform over the images to be fused.

(2) Detect local image orientations by maximizing the energy measure $E_N^{1D}(\theta) = \sum_{n=1}^{N} \left[ L_{n,0}^{\theta} \right]^2$ for each window position. In practice, one estimator of the optimal orientation $\theta(p,q)$ can be obtained through $\theta(p,q) = \tan^{-1}\left[ L_{0,1}(p,q) / L_{1,0}(p,q) \right]$ where $L_{0,1}(p,q)$ and $L_{1,0}(p,q)$ are the first-order Hermite transform coefficients at window positions $(p,q)$.

(3) Steer the transform coefficients of each image towards directions $\theta(p,q)$.

(4) Select coefficients from each source image at every window position. The coefficient variance at each window position is computed as a measurement of the activity associated with the central pixel of the window.
A strong value indicates the presence of a dominant pattern in the local area. A binary decision map is constructed indicating which image has dominant patterns at each window position. This binary map is subject to consistency verification[11].

(5) Construct a new set of transform coefficients from the coefficients of both source images according to the binary map.

(6) The final step is the inverse transformation from the new coefficient set. This results in the fused image.

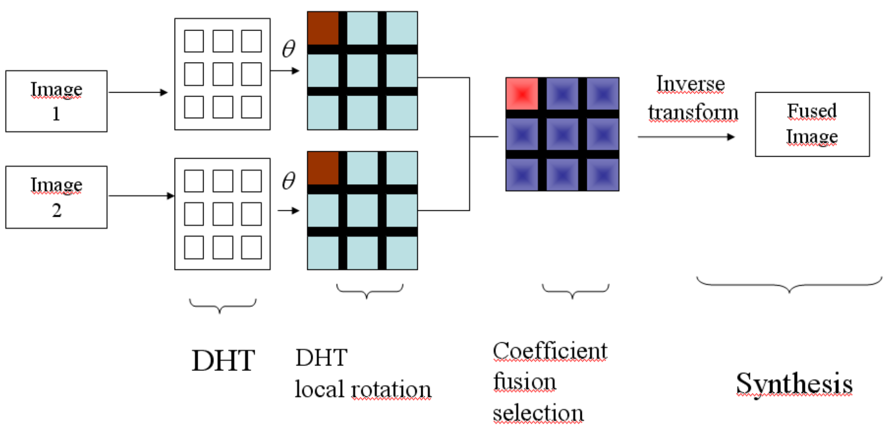


Fig. 3. Hermite transform fusion for multispectral and panchromatic images.

## 3.2 Pattern classification

The steered HT offers a way to estimate the dimensionality of the underlying local image pattern, either 0D, 1D, or 2D. By analyzing relations between the oriented energies, it is possible to classify image patterns. Locally seen, roads are mainly one dimensional structures; therefore our approach consists of classifying the image according to the dimensionality of the underlying local pattern. The dimensionality of a pattern is directly related to the number of coefficients needed to code the block. For blocks over a constant-luminance background all the AC-coefficients vanish. Therefore, only a DC-coefficient ($L_{00}$) is needed to code the block. For blocks containing a strongly oriented structure, only the coefficients $L_{i,0}^{\theta}$ for $i = 0,\ldots, N$, are required for representing that pattern (for simplicity we omit the superscript hereafter). The third general class comprises all the non-oriented patterns like corners, junctions, random dots, etc.

To produce a partition of the block space we assume that each of these three classes can be well described by the following variables:

$$L = L_{0,0}, \qquad C = \left[ \sum_{i=0}^{N} \sum_{j=0}^{N} L_{i,j}^2 - L^2 \right]^{1/2}, \qquad \Delta C = \left[ C^2 - \sum_{i=1}^{N} L_{i,0}^2 \right]^{1/2} \tag{3}$$

which are measures of the mean luminance, contrast, and 1D residual contrast respectively.

Let the three classes labeled as 0D, 1D and 2D. The classification is done in two steps. First, the 0D class is separated by means of the comparison

$$C \underset{>_{1D \cup 2D}}{\overset{<^{0D}}{}} k_0 C_{thr}(L) \tag{4}$$

where the curve $C_{thr}(L)$ represents the light adaptation threshold(ref) This threshold can be inferred from the image of Fig. 4(a). The image presents white Gaussian noise with mean varying linearly along horizontal direction, while the variance varies linearly along vertical direction. The curve plotted over the image represents a typical detection threshold. This threshold determines the limit from which local contrast becomes visually relevant. The light adaptation threshold was fitted following the approach presented in a previous work[12]. Fig. 4-(a) shows the theoretical curve of this threshold and the measured curve along the image profile on the image.

In the second step, we separate the 1D blocks from the set of remaining blocks by mean of the following comparison

$$\Delta C \underset{>_{2D}}{\overset{<^{1D}}{}} k_1 \Delta C_{thr}(L, C) \tag{5}$$

where $\Delta C_{thr}$ is the contrast masking threshold. Ideally, for 1D patterns all the energy is fully contained along the gradient orientation and the residual contrast becomes zero. In practice, however, there is a residual contrast different from zero even in the cases when oriented structures are well perceived. Contrast masking refers precisely to the reduction in the visibility of a component in the image due to the presence of another. Again, the model used for contrast masking has been previously formulated[12]. In Fig. 4(b) we have plotted a stimulus image built with a sinusoid wave grating plus uncorrelated Gaussian noise. The sinusoidal amplitude was varied linearly long the vertical axis, while the standard deviation of noise was varied linearly long the horizontal axis. In the figure it is also shown the theoretical curve and the measured values for the image profile shown to the left. Figure 5 shows a pattern classification over a natural scene.

## 4. RESULTS

We used SPOT images from a crop field with a multispectral image resolution of 10m and a panchromatic image resolution of 2.5. Figure 6 shows the results of detecting 0D, 1D, and 2D patterns. Figure 7 shows 1D patterns superimposed to original image. Roads are associated to 1D patterns. Since the detected 1D patterns are often discontinuous, it was necessary to implement a semi-automatic connecting process based on active contours (snakes). At the right, detected 1D patterns are superimposed to the original image.

## 5. CONCLUSIONS

Results show that the proposed methodology indeed classifies the image according to its local dominant dimensionality. There are some spots where the underlying noise or the complexity of patterns produces confusion in the classification

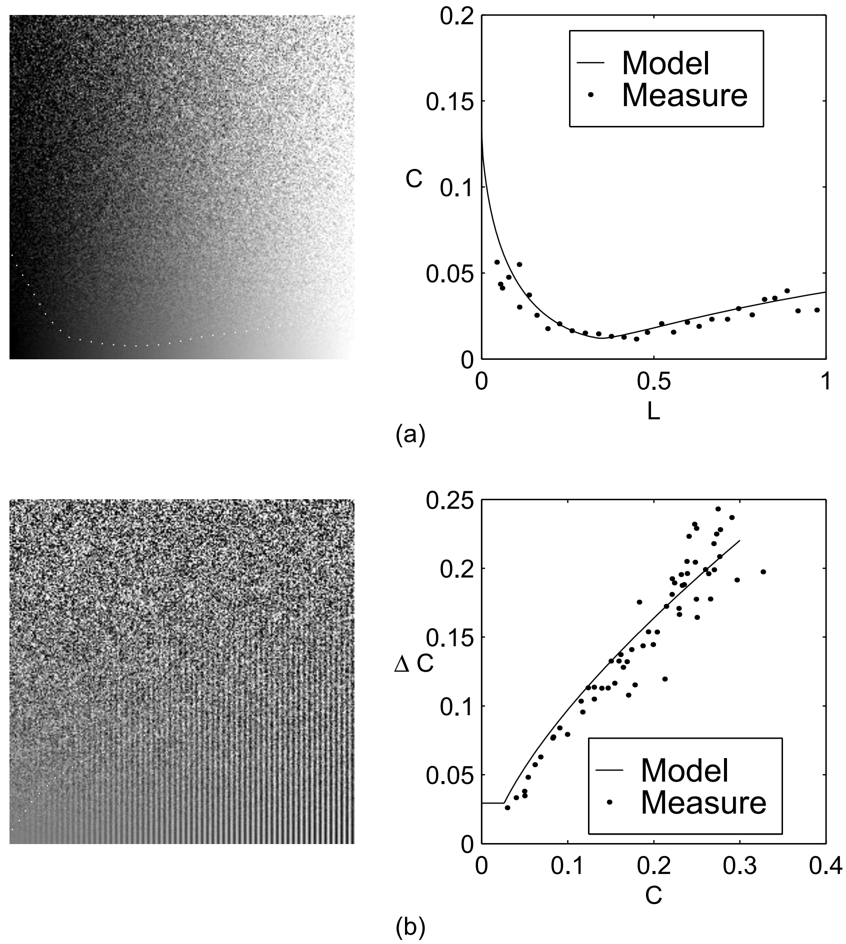


Fig. 4.(a) light adaptation and (b) contrast masking thresholds.

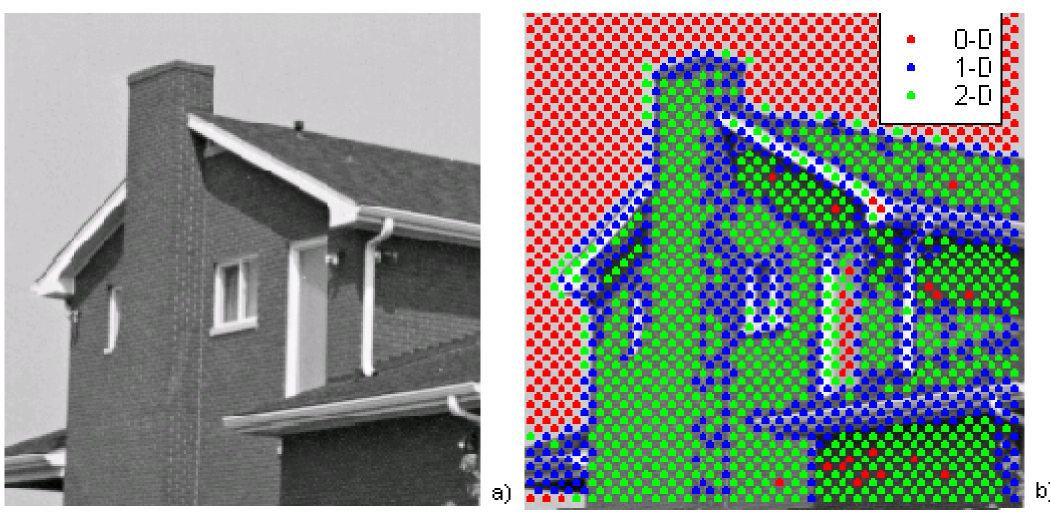


Fig. 5 0D, 1D, and 2D patterns found in a natural scene from dimensional energy analysis with the rotated Hermite Transform.

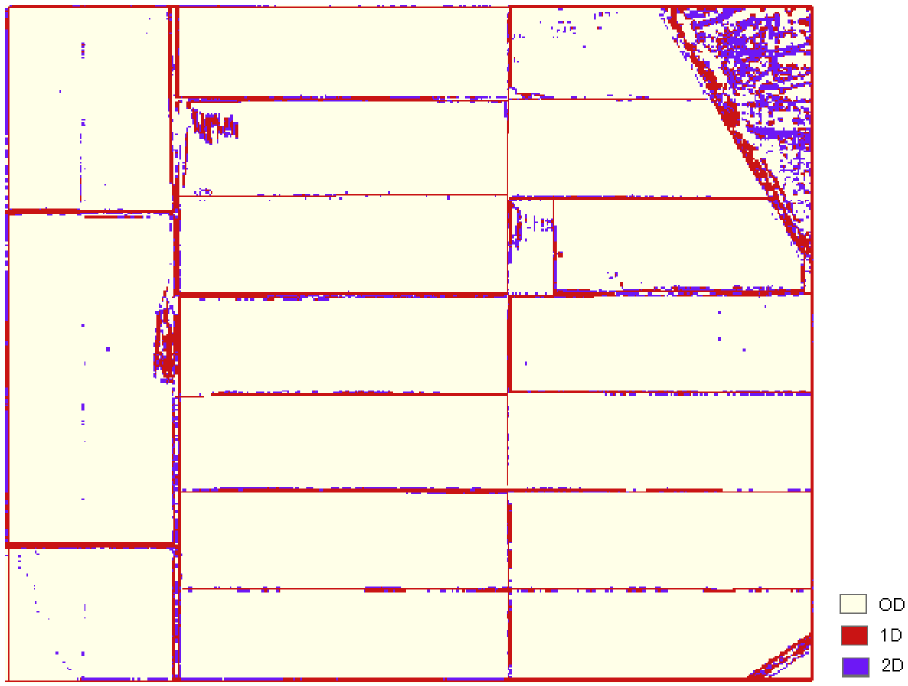


Fig. 6. 0D, 1D, and 2D patterns found in rural area.

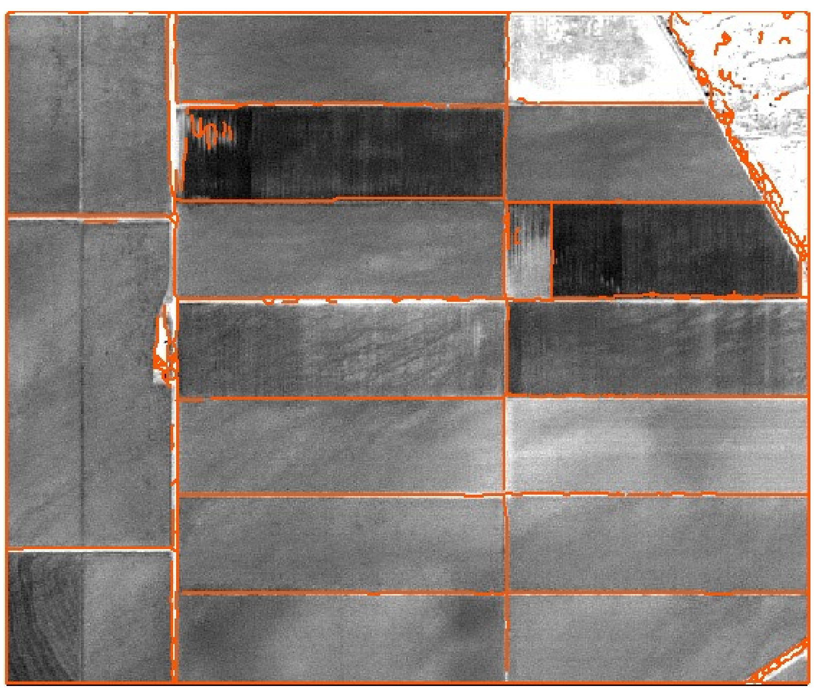

Fig. 7. 1D patterns superimposed on original SPOT image.

process. Furthermore many of the detected 1D patterns presented discontinuities, and a connection algorithm based on active contours had to be used. The proposed methodology is efficient to detect image primitives that can be considered strong candidates for roads, i.e., 1D patterns, can be associated to roads but also to other image structures such as borders, buildings, etc. Further processing is required to fully detect roads, such as contour tracking[13]. Future work must include multiresolution analysis as an interesting alternative to road tracking and road discrimination. The Hermite transform is suited for this kind of analysis[9], since its agreement with the scale space theory provides multiscale image representation that do not introduce artifacts, and it is also computationally efficient in both schemes, decimated and non-decimated.

## ACKNOWLEDGMENT

This work was sponsored by UNAM PAPIIT grant IN106608.

## REFERENCES

[1] Mena, J.B., "State of the art on automatic road extraction for GIS update: a novel classification", Pattern Recognition Letters 24, 3037-3058 (2003).

[2] Mukherjee, J., "MRF clustering for segmentation of color images," Pattern Recognition Letters 23(8), 917-929 (2002).

[3] Chanussot, J., Lambert, P., "An application of mathematical morphology to road network extraction on SAR images", Proc. Internat. Symp. Math. Morphol. Amsterdam, 399-406 (1998).

[4] Péteri, R., Ranchin, T., "Multiresolution snakes for urban road extraction from Ikonos and Quickbird images", 23nd EARSeL Annual Symposium Remote Sensing in Transition, 2-5 (2003).

[5] Holschneider, M., Kronland-Martinet, R., Morlet, J., Tchamitchian, P. "A real-time algorithm for signal analysis with the help of wavelet transform", [Wavelets Time-Frequency Methods and Phase Space], Springer-Verlag, Berlin, 289-297 (1989).

[6] Martens, J.B., "The Hermite Transform –Theory", IEEE Transactions on Acoustics Speech and Signal Processing 38 (9), 1607-1618 (1990).

[7] Escalante-Ramírez, B., López-Caloca, A., "Image fusion with the Hermite Transform," 14th International Conference on Image Processing, (2003).

[8] Escalante-Ramírez, B., López-Caloca, A., "The Hermite transform: an efficient tool for noise reduction and image fusion in remote-sensing", [Signal and Image Processing for Remote Sensing], C.H. Chen, ed., CRC Press (Taylor and Francis Books), Boca Raton, 539-557 (2006).

[9] Silván-Cárdenas, J.L., Escalante-Ramírez, B., "The multiscale Hermite transform for local orientation analysis", IEEE Transactions on Image Processing 15(5), 1236-1253 (2006).

[10] Van Dijk, J.A.M., Martens, J.B. "Image representation and compression with steered Hermite transform", Signal Processing 56, 1-16 (1997).

[11] Li, H., Manjunath, B.S., Mitra, S.K., "Multisensor image fusion using the wavelet transform", Graphical Models and Image Processing 57 (3), 235-245 (1995).

[12] Silván-Cárdenas, J.L., Escalante-Ramírez, B., "Image coding with a directional-oriented discrete Hermite transform on a hexagonal sampling lattice", Proc SPIE 4472, 528-536 (2001).

[13] Péteri, R., Celle, J. Ranchin, T. "Detection and extraction of roads networks from high resolution images", 14th International conference on Image Processing, (2003).