SAR-Image Classification with a Directional-Oriented Discrete Hermite Transform and Markov Random Fields

Boris Escalante-Ramírez\textsuperscript{a,b}, Penélope López-Quiroz\textsuperscript{a}, José L. Silván-Cárdenas\textsuperscript{b}

\textsuperscript{a}Graduate Division, School of Engineering
National University of Mexico, UNAM

\textsuperscript{b}Center for Geography and Geomatics Research, “Ing. Jorge L. Tamayo”, Centro GEO
Mexico, D.F. MEXICO

boris@servidor.unam.mx, penelope@servidor.unam.mx, jlsilvan@centrogeo.org.mx

Abstract - A novel classification scheme for SAR images based on the perceptual classification of image patterns in the Discrete Hermite Transform domain over a roughly hexagonal sampling lattice has been developed. In order to obtain the DHT referred to a rotated coordinate system the set of coefficients of a given order are mapped through a unitary transformation based on the generalized binomial function. This representation allows a perceptual classification, including constant patterns (0-D), oriented structures (1-D), and non-oriented structures (2-D). Classification is based on light adaptation and contrast masking properties of the human vision. Finally, classification is improved by means of a probabilistic approach based on Markov Random Fields.

Keywords: SAR image classification; Hermite transform; binomial filters; steerable transforms; human perception; Markov random fields.

I. INTRODUCTION

Unsupervised image classification schemes often aim at finding differences among a wide variety of image characteristics such as texture, gray level, pattern structure, etc. Some of the more efficient approaches combine statistical methods with multiresolution analysis methods. Some of these methods are inspired in important properties of human vision such as the Hermite Transform. This transform has been proposed as a robust model for image representation since it carries out similar operations to those present in the primary stages of the human visual system [1]. These operators are modeled as Gaussian derivatives and have been successfully used in several applications, including restoration [2], coding [3] and optic flow estimation [4]. They possess symmetry properties related to translation, rotation and scaling which are ideal for the analysis of the geometric properties of images [5], [6]. The discrete approach of these operators based on binomial filters satisfies a number of recursive relationships that can be efficiently implemented using adders and delays only. Moreover, these discrete filters present interesting orthogonality properties. In this work we use the discrete version of the Hermite transform (DHT) on a roughly hexagonal sampling lattice. For the case of SAR imagery, an algorithm for speckle reduction, also based on the Hermite Transform, is used before classification [2].

II. IMAGE DECOMPOSITION

The DHT maps the image data into a set of spectral coefficients from which it is possible to recover the original image:

\[ W_k = \sqrt{2^{-N}} C_N^k \]

\[ B_{n,k} = 2^{-N} \Delta^n \left\{ C_N^{k-n} \right\}, \quad \text{for } k = 0, \ldots, N, \]

where \( C_n^k = n!/(n-k)! \) are the binomial coefficients for \( k = 0, \ldots, n \) and \( \Delta^n f_k = \sum_{j=0}^{n} (-1)^{n-j} C_n^j f_{k+j} \) is the \( n \)-th order forward difference of the discrete sequence \( f_k \). Then, the DHT is obtained by convolving the input signal with the binomial filters \( D_{n,k} = W_k B_{n,N-k} \) and taking the outputs at points in a given lattice [1]. The binomial functions satisfy a number of properties that can be expressed using the matrix notation \( B = [B_{n,k}]_{n,k=0,\ldots,N} \), where each row represents a mask of analysis. Therefore, the outputs of the filtered signal by the binomial filters at a generic lattice point can be obtained by applying the matrix \( D = WB \) to a vector containing the samples of the signal, where \( W = \text{diag}(W_k)_{k=0,\ldots,N} \) is a diagonal matrix of the window function. The two most appealing properties of the binomial functions are the orthogonality \( BB = I \) and the symmetry \( W^2 B = B^2 W^2 \), which leads to the following identity for the binomial filters: \( D^T D = W^2 \). Thus, if

\[ D = \sum_{n,k} D_{n,k} W_k B_{n,k} \]

This work was partially supported by DGAPA-PAPIIT grant IN107101
F is a matrix containing the data of a block of \((N + 1) \times (N + 1)\) pixels, then \(G = D F D^T\) are the spectral coefficients of the DHT. The DHT is said to be a local expansion of the image. This is expressed by \(W^2 F W^2 = D G D^T\), which is the windowed block of data.

### III. Local Rotation

The rotation in the DHT domain can be seen as a mapping of the spectral coefficients (for each point in the sampling lattice) through a linear transformation [3]. This mapping is done over the set of coefficients of the same total order by means of the generalized binomial functions

\[
B_{n,k}^{(N)} = \tan^{k+n} \frac{\pi}{2} \theta^N \left[ C_{N-n}^{k-n} \cos^{2k} \theta \sin^{2N-2k} \theta \right]
\]

for \(0 \leq \theta \leq \pi / \mu\). Notice that when the slant parameter \(\theta\) is set to \(\pi / 4\) we get the symmetric case, i.e., \(4/\delta\), where \(\delta\) are given by

\[
\delta_n = \frac{n}{2N}, \quad 2N \text{ denote the column vectors formed by the anti-diagonals of matrix } G, \text{ then the coefficients in a coordinate system that is rotated by an angle } \theta \text{ are given by } \begin{bmatrix} g_{\theta}^r \end{bmatrix} = R^{\theta} g_n, \text{ where the rotation matrix } R^{\theta} = \text{WB}^{\theta} W^{-1}. \]

Since \(\text{W is unitary, the inverse rotation is achieved by applying the transposed matrix } R^T\). Hereafter, we refer to the rotated coefficients as the matrix \(G(\theta)^r\) whose anti-diagonals are the vectors \(g_{\theta}^{r,\theta}\). In the remaining work we consider the angle \(\theta = \arctan (G_{0,1}=G_{1,0})\), which is nothing but the gradient orientation.

### IV. Pattern Classification

Since image content is not homogenous, it is convenient to separate the space of blocks into three classes according to the dimensionality of the underlying 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 \((G_{0,0})\) is needed to code the block. For blocks containing a strongly oriented structure, only the coefficients \(G_{i,0}\) and \(G_{0,i}\) are required to represent 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:

\[
L = G_{0,0}, \quad C = \left[ \sum_{i=0}^{N} \sum_{j=0}^{N} G_{i,j}^2 - L^2 \right]^{1/2}, \quad \Delta C = \frac{C^2 - \sum_{i=0}^{N} \sum_{j=0}^{N} G_{i,j}^2}{2}
\]

which stand for mean luminance, contrast, and 1-D residual contrast respectively. Let the three classes be labeled as 0-D, 1-D and 2-D. The classification is done in two steps. First, the 0-D class is separated by means of the comparison. If \(C < k_0 C_{thr}(L)\), then 0-D, otherwise 1-D or 2-D, where \(C_{thr}(L)\) is the light adaptation threshold [7]. In the second step, we separate the 1-D blocks from the remaining blocks by means of the following comparison. If \(\Delta C < k_1 \Delta C_{thr}(L, C)\) then 1-D, otherwise 2-D, where \(\Delta C_{thr}(L, C)\) is the contrast masking threshold [8]. Ideally, for one-dimensional 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. The contrast masking model used here is based on the works of Legge and Foley [8], [9].

### V. Markovian Regularization

In order to estimate the true vector field from the output vector field we follow a similar approach to the problem of image restoration. Let us take a Markov random field (MRF) model for piecewise constant image restoration where the goal is to recover true pixel values \(f\) from the observed image values \(d\). In our case, let \(d = \{d_1, ..., d_N\} \) be an observation, i.e., a rectangular array of vector values, \(f = \{f_1, ..., f_w\} \) the set of sampled vectors, which is assumed to be a realization of a MRF, and \(L = \{l_1, ..., l_M\}\) the set of all possible labelings. The classification of MRFs, an optimal estimate, known as the MAP estimate, is given by:

\[
f^* = \arg\max_{f \in F} P(f | d) = \arg\max_{f \in F} \left\{ p(d | f) P(f) \right\}
\]

If \(P(f | d) \propto e^{-U(f(d))}\), and \(U(f | d) = U(d | f) + U(f)\) then the MAP estimate can be found by minimizing \(f^* = \arg\min_U U(f | d)\), known as the a-posteriori energy. In order to model \(U(f)\), we assume that for piecewise constant surfaces the clique potentials \(V_c\) are defined by using multilevel logistic (MML) considering only the pair site cliques, i.e., \(V_c(f) = V_2(f, f_r) = v_{20}(1 - \delta(f, f_r))\), where \(\delta(\cdot)\) is the Kronecker delta function and \(v_{20}\) is the penalty for “non equal labels on two-site cliques”. The a-priori energy is the sum of all clique potentials, that is, \(U(f) = \sum_{i=1}^{N} \sum_{j=1}^{N} v_{20}[1 - \delta(f_i, f_r)]\).

The likelihood energy \(U(d|f)\) can be determined using a general observation model. Assuming no noise, then \(U(d|f) = U(f)\). In order to minimize \(U(f)\), the easiest algorithm is steepest local energy descent. Given the data \(d \in D\) and \(f_{S[\{i\}]}\), the set of labels at the sites in \(S - \{i\}\), the algorithm ICM (Iterated Conditional Modes) sequentially updates each \(f_{i}^{(k)}\) into \(f_{i}^{(k+1)}\) by minimizing the corresponding a-posteriori potential using the following rule

\[
f_{i}^{(k+1)} = \arg\min_{f_i} V(f_i | d_i, f_{N_i})\]

where

\[
V(f_i | d_i, f_{N_i}) = \sum_{l \in N_i} V(f_i | f_{l}^{(k)}) + V(d_i | f_i)
\]

If we consider
$U(d|f) = 0$, then $V(d|f) = 0$ and the a-posteriori potential $V(f_i | d_i, f_{N}) = v_2 \sum_{i \in N} [1 - \delta(f_i | f_r)]$, where

$\sum_{i \in N} [1 - \delta(f_i | f_r)]$ is the number of neighboring sites satisfying $mag(f_i - f_r) > threshold$. It is important to stress that ICM estimation accuracy depends very much on the initial estimator $f^{(0)}$, which, in our case, is the output vector field.

Fig. 1. Output vector field from pattern classification

Fig. 2. Markovian regularized vector field

VI. RESULTS

Results of the classification scheme were tested on a SAR-ERS2 image, taken in 1995, over Mexico City valley. Fig. 1 shows the output vector field obtained from pattern classification over a detail of the original image. Note that the vector length is proportional to the energy oriented towards the dominant direction. It is therefore expected that longer vectors correspond to 1D structures Figure 2 shows the estimated vector field after Markovian regularization.

VII. Conclusions

We have developed a method for SAR image classification based on the de rotated discrete Hermite transform. This transform is a computational efficient model that possesses interesting mathematical and psychophysical properties. In order for the algorithm to be efficient, speckle reduction should be applied before classification. 0D, 1D and 2D patterns are classified by comparing parameters extracted from DHT coefficients, such as mean luminance, contrast and 1D residual contrast with light adaptation and contrast masking thresholds. A Markov random field restoration approach is used to estimate the true vector field of oriented structures. Results of this classification scheme show that 0D patterns correspond to flat surfaces on the ground, such as water reservoirs and large flat surfaces. 1D patterns correspond to streets, large urban constructions, and boundaries between different regions. 2D patterns include inhabiting zones and irregular surfaces. Analysis of the evolution of classes along different spatial resolutions may help classifying zones accordingly: large objects at low resolutions while fine detail at higher resolutions.

REFERENCES