ACTIVE CONTOUR SEGMENTATION GUIDED BY AM-FM DOMINANT COMPONENT ANALYSIS

Nilanjan Ray[1], Joebob Havlicek[2], Scott T. Acton[1] and Marios Pattichis[3]

[1]Department of Electrical and Computer Engineering, University of Virginia
[2]Department of Electrical and Computer Engineering, University of Oklahoma
[3]Department of Electrical and Computer Engineering, University of New Mexico

ABSTRACT

For the first time, we explore the application of active contours in the modulation domain by computing snakes on image modulations. As we demonstrate in the examples, such snakes are able to utilize information inherent in the dominant image modulations to acquire and track visually and semantically meaningful structures within the image. We use nonlinear AM-FM image representations to capture regions that are homogeneous in intensity and in texture. A geometric snake approach utilizing a fuzzy classifier is then applied to the image modulations. The combination of AM-FM analysis and the active contour evolution produces an efficacious image partition. As a preliminary demonstration of this novel approach, we apply the modulation domain snakes to the classical texture segmentation problem.

1. INTRODUCTION

In this paper, we propose a segmentation approach based on active contours. The segmentation is achieved by evolving a set of geometric active contours on the classified results of an AM-FM dominant component analysis. The AM-FM representation facilitates both intensity-based and texture-based segmentation. A set of synthetic and real image results demonstrate the potential of the proposed approach for use in a variety of practical applications.

2. AM-FM IMAGE MODELS

AM-FM functions are nonstationary quasi-sinusoidal oscillations that admit simultaneous amplitude and frequency modulations [1], [2]. A general 2-D AM-FM function takes the form

$$t(x,y) = a(x,y)\exp[j\varphi(x,y)], \qquad (1)$$

where $a(x,y) \geq 0$ is the *amplitude modulation function*, or *AM function* of $t(x,y)$ and $\nabla\varphi(x,y)$ is the *frequency modulation function* or *FM function* of $t(x,y)$. The AM function $a(x,y)$ captures the local contrast of the complex-valued image $t(x,y)$, while the FM function $\nabla\varphi(x,y)$ captures the local texture orientation and granularity. The vector-valued FM function may be further decomposed into an instantaneous horizontal frequency function $U(x,y) = [1\ 0]\ \nabla\varphi(x,y)$ and instantaneous vertical frequency function $V(x,y) = [0\ 1]\ \nabla\varphi(x,y)$.

Note that the model (1) is complex-valued, whereas almost all practical image processing applications are concerned exclusively with real-valued images. For a real-valued image, agreement with the complex model (1) is achieved by adding a unique imaginary part equal to the multidimensional directional Hilbert transform of the image [1], [2], [3]. The resulting complex image, which admits many of the most attractive properties of the well-known 1-D analytic signal, is known as the *analytic image*. Given the analytic image $t(x,y)$, the AM and FM functions of the real-valued image may be obtained using the demodulation algorithm [2]

$$\nabla\varphi(x,y) = \mathrm{Re}\left[\frac{\nabla t(x,y)}{jt(x,y)}\right], \qquad (2)$$

$$a(x,y) = |t(x,y)|. \qquad (3)$$

Discretization of the demodulation algorithm (2), (3) was detailed in [2]. This algorithm is based on a *quasi-eigenfuntion approximation* [2], [4], which assumes that, over *sufficiently small* spatial neighborhoods, the image $t(x,y)$ is approximately sinusoidal. Such images are termed *locally coherent*.

The local coherence assumption is violated for many images of interest in important practical applications. Therefore, it is generally necessary to consider that any given image is not just one, but rather a sum of multiple image *components* of the form (1). This gives rise to the multicomponent AM-FM image model

$$t(x,y) = \sum_{k=1}^{K} a_k(x,y)\exp[j\varphi_k(x,y)]. \qquad (4)$$

One popular strategy for isolating the multiple image components in (4) is to apply a multiband Gabor filterbank as described in [5]. Appropriate modifications to the demodulation algorithm (2), (3) and its discrete counterpart are discussed in [2] such that the modified algorithm can be applied directly to the filterbank channel responses to simultaneously estimate the multiple AM and FM functions $a_k(x,y)$ and $\nabla\varphi_k(x,y)$ in (4).

It then becomes possible to extract the dominant modulations on a pointwise basis. This approach is known as *Dominant Component Analysis,* or DCA [2], [5]. At

0-7803-6725-1/01/$10.00 ©2001 IEEE

each pixel, the dominant modulations are the AM and FM functions that correspond to the component in (4) that dominates the local image spectrum. They provide a powerful characterization of the local texture structure and have been used with great success in solutions to a wide variety of image processing and computer vision problems.

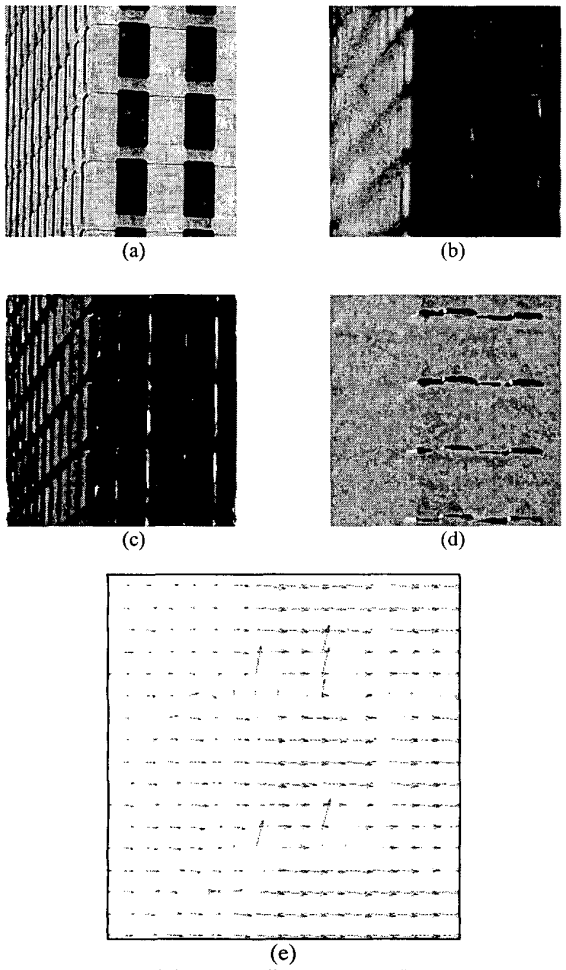

Figure 1: (a) Original "Building" image"; (b) Dominant AM function; (c) Dominant FM U component; (d) Dominant FM V component; (e) Dominant FM function shown as a needle diagram.

The natural building image shown in Figure 1(a) is from the MIT VisTex database. The dominant AM and FM functions computed from Figure 1(a) are shown as images in Figures 1(b)-(d). The dominant FM function is also shown as a needle diagram in Figure 1(e), where the arrow lengths are proportional to the reciprocal of the instantaneous frequency magnitude. With this convention, long arrows correspond to low frequencies which are typically dominant in the neighborhood of large objects in the image. From Figure 1(e), it is clear that the dominant FM function provides powerful cues for discriminating between the two faces of the building that appear in the left and right halves of the image.

Figure 2(a) shows the image "Grass-flowers". This synthetic image is a juxtaposition of two Brodatz-like textures. The computed dominant AM function is shown in Figure 2(b) and is clearly useful for discriminating between the two textures. In this case, the dominant FM function (Figures 2(c)-(d)) admits a variety of orientations and frequency magnitudes in both textured regions.

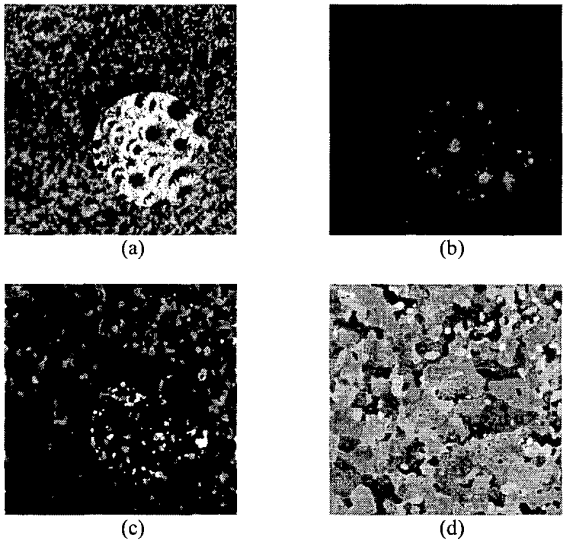

Figure 2: (a) Original "Grass-flower" image"; (b) Dominant AM function; (c) Dominant FM U component; (d) Dominant FM V component.

3. CLASSIFICATION

Given the AM-FM decomposition, we essentially have three inputs to our segmentation process: the AM image **A**, the horizontal component of frequency **U** and the vertical component of frequency **V**. We use these three images to formulate a multicomponent segmentation problem. The first step is to combine these three into a single vector-valued image

$$\xi = \{\mathbf{A}, \mathbf{U}, \mathbf{V}\}, \qquad (5)$$

where $\xi(x, y)$ represents the AM and the two FM components at location (x, y) in the image.

This three-component image can be converted into a class map via fuzzy c-means classification [6]. Given that the number of classes N is known *a priori*, we minimize the following energy functional:

$$E(\Xi, \mathrm{X}) = \sum_{x,y} \sum_{i=1}^{N} (u_i(x,y))^2 \left| d_i(x,y) \right|^2 \qquad (6)$$

where Ξ dictates class membership, and X is the set of

cluster centers. The distance of the AM-FM components at position (x, y) from the center μ_i of cluster i is given by

$$d_i(x,y) = \left|\xi(x,y) - \mu_i\right|.$$

At each point (x, y), the image has a fuzzy class membership value $u_i(x, y)$ for the i^{th} class that is recomputed at each iteration of the algorithm:

$$u_i(x,y) = 1/[\sum_{x,y}\sum_{j=1}^{N}(d_i(x,y)/d_j(x,y))^2],$$

Concomitantly, the cluster centers are iteratively updated according to

$$\mu_i = \sum_{x,y}(u_i(x,y))^2 \xi(x,y) \Big/ \sum_{x,y}(u_i(x,y))^2.$$

The results of the fuzzy c-means classification do not provide a segmentation, however. To find closed regions, we utilize active contours that are often referred to as *snakes*.

4. ACTIVE CONTOURS AND CURVE EVOLUTION

An active contour may be considered as the intersection of a plane and a conical surface [7], [8]. Let $\Phi(x,y,t)$ be an evolving 3-D cone, and let $\Phi(x,y,t) = 0$ be the *zero level set* at any instant t. Assume that our evolving curve exists within the zero level set. In this framework, we achieve a distinct advantage over parametric snake evolution: we can enact contour splitting and merging effortlessly. So, the level set model facilitates topological changes [7]. The zero level set governs the curve evolution and the corresponding differential equation is given by [8]

$$\partial\Phi/\partial t + F\left|\nabla\Phi\right| = 0, \qquad (7)$$

where $F(x,y)$ is the speed of the curve evolution in the direction to the outward normal of the curve at (x,y). In implementing the level set approach to curve evolution, design of the speed variable $F(x,y)$ is the key to success. We formulate a speed term to accommodate multiple snakes which will be associated with multiple image segments.

Again, we assume that there are N region classes in the image (N classes, but possibly more than N segments). In the AM-FM modeling framework, these classes do not necessarily represent regions of homogeneous graylevel; rather, they correspond to homogeneous textures captured by DCA. Let $\mathbf{C}_i$ be the i^{th} active contour for the i^{th} class, where each contour is capable of splitting to cover multiple segments of the same class. For active contour evolution, we must establish a cost functional for each curve and then define a set of curve update functions that minimizing the cost functionals. The cost functional for the i^{th} curve is given by

$$E(\mathbf{C}_i(x(s),y(s))) = \iint_{\mathbf{C}_i} b_i(x,y)dxdy + \alpha \int_{\partial\mathbf{C}_i} \sqrt{x_s^2 + y_s^2}\, ds. \qquad (8)$$

Here, $s \in [0,1]$ is an index identifying a specific active contour and $\alpha > 0$ is a weight that penalizes the length of the boundary $\partial\mathbf{C}_i$ of $\mathbf{C}_i$. In the first term, we find $b_i(x,y)$, which is a function penalizing each pixel inside the contour that does not belong to the i^{th} class. Likewise, $b_i(x,y)$ encourages encompassing pixels of the i^{th} class [9]:

$$b_i(x,y) = \begin{cases} -1 \text{ if } (x,y) \in i^{th} \text{ class} \\ +1 \text{ if } (x,y) \notin i^{th} \text{ class} \end{cases}.$$

The energy functionals lead to a set of N decoupled descent equations [9], [10] that allow curve evolution. For $\mathbf{C}_i$, we have a speed term of

$$F_i(x,y) = -b_i(x,y) - \alpha\kappa_i,$$

in which

$$\kappa_i = \mathrm{div}\left(\nabla\Phi_i / \left|\nabla\Phi_i\right|\right). \qquad (9)$$

So, with (9), we can evolve the contours on the zero level set using (7).

5. RESULTS AND DISCUSSION

We computed the dominant AM and FM components for the "Building" image (see Figure 1) and the "Grass-flower" image (see Figure 2). The fuzzy c-means clustering gives the two-class maps shown in Figure 3 and Figure 4. The classification results by themselves are noisy and unusable for segmentation.

Figure 5 and Figure 6 show the initial geometric snake positions for the "Building" and "Grass-flower" examples, respectively. After evolution of the geometric snakes based on the energy minimization technique (8), we have obtained the final active contour positions as shown in Figure 7 and Figure 8 respectively. For the "Building" example we have used stiffness parameter $\alpha =$ 20.0, and for the "Grass-flower" example we have taken α = 15.0. Finally we get the segmentation results as shown in Figure 9 and Figure 10 for these two images. We observe that minor regions and rough boundaries have been removed, compared to the results shown in Figure 3 and Figure 4.

In the experiments, we assumed N classes of textured regions. However, the number of textured regions present in the image results may be estimated without *a priori* information by applying the density based clustering approach described in [11] to a scatter plot of the computed dominant modulations. The computed dominant modulations **A**, **U**, and **V** are first normalized by dividing each one by its respective sample standard deviation. The image pixels are then plotted in a feature space with axes corresponding to the normalized modulations. A 3-D Gaussian filter is applied in this feature space to estimate the local density of feature vectors about each point. Gradient ascent is then used to identify local maxima in the filtered result and group the feature vectors into clusters. By thresholding on the number of feature vectors in each cluster, we remove minor clusters and merge them

with the larger clusters via the nearest neighbor rule, where the squared-error or similar validation metrics can be used as a stopping criterion. This approach was demonstrated successfully for unsupervised texture segmentation in [12].

Figure 3. FCM result on "Building" image.

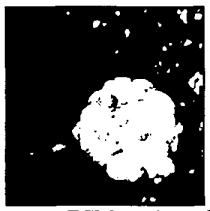

Figure 4. FCM result on the "Grass-flower" image.

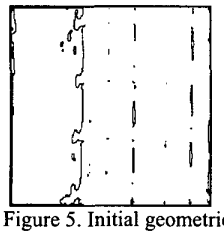

Figure 5. Initial geometric snake position for the "Building" image.

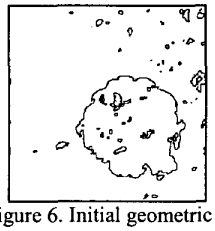

Figure 6. Initial geometric snake position for the "Grass-flower"image.

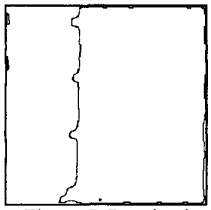

Figure 7. Result of geometric snake evolution on the "Building"image.

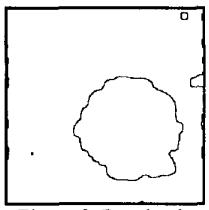

Figure 8. Result of geometric snake evolution on the "Grass-flower" image.

Figure 9. Segmentation result for "Building".

Figure 10. Segmentation result for "Grass-flower".

Following table compares percentages of error (i.e., pixels actually belonging to i^{th} class classified as j^{th} class) of the classification in FCM to the proposed method.

Classification error	FCM	Snake Evolution
Building Image	3.53 %	2.13 %
Grass-flower Image	5.21 %	2.96 %

Table 1. Misclassification in FCM and proposed method.

The segmentation results from the preliminary study of applying snakes in the modulation domain show promise for texture analysis problems. Potential applications include object-based image coding, segmentation of remotely sensed imagery and segmentation for content-based image retrieval.

REFERENCES

[1] J.P. Havlicek, D.S. Harding, and A.C. Bovik, "The multi-component AM-FM image representation," *IEEE Trans. Image Proc.*, vol. 5, no. 6, pp. 1094-1100, Jun. 1996.

[2] J.P. Havlicek, D.S. Harding, and A.C. Bovik, "Multidimensional quasi-eigenfunction approximations and multicomponent AM-FM models," *IEEE Trans. Image Proc.*, vol. 9, no. 2, pp. 227-242, Feb. 2000.

[3] J.P. Havlicek, J.W. Havlicek, and A.C. Bovik, "The analytic image," in *Proc. IEEE Int. Conf. Image Processing*, Santa Barbara, CA, Oct. 26-29, 1997.

[4] A.C. Bovik, J.P. Havlicek, D.S. Harding, and M.D. Desai, "Limits on discrete modulated signals," *IEEE Trans. Signal Processing,"* vol. 45, no. 4, pp. 867-879, Apr. 1997.

[5] J.P. Havlicek, A.C. Bovik, and D. Chen, "AM-FM image modeling and Gabor analysis," in *Visual Information Representation, Communication, and Image Processing*, C.W. Chen and Y. Zhang, eds., Marcel Dekker, New York, 1999.

[6] J.C. Bezdek, J. Keller, R. Krisnapuram, N.R. Pal, *Fuzzy models and algorithms for pattern recognition and image processing*, Boston, Kluwer Academic, 1999.

[7] S. Osher and J.A. Sethian, "Fronts propagating with curvature dependent speed: Algorithm based on Hamilton-Jacobi formulation," *Journal of Comp.Physics.*, vol. 79, pp. 12-49, 1988.

[8] J.A. Sethian, *Level set methods and fast marching methods*, Cambridge University Press, 1999.

[9] N. Ray and S.T. Acton, "Image segmentation by curve evolution with clustering," *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, California, October 29 – November 1, 2000.

[10] A.Yezzi Jr., A.Tsai, and A. Willsky, "Binary and ternary flows for image segmentation," *IEEE International Conference on Image Processing*, vol. 2, pp. 1-5, 1999.

[11] E.J. Pauwels and G. Frederix, "Finding salient regions in images," *Comput. Vision, Image Understand*, vol. 75, no. 1/2, July/August 1999, pp. 73-85.

[12] T.B. Yap, T. Tangsukson, P.C. Tay, N.D. Mamuya, and J.P. Havlicek, "Unsupervised texture segmenation using dominant image modulations," *Proc. 34th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, October 29 - November 1, 2000.