MODULATION-FEATURE BASED TEXTURED IMAGE SEGMENTATION USING CURVE EVOLUTION

Iasonas Kokkinos, Georgios Evangelopoulos and Petros Maragos

National Technical University of Athens,
Computer Vision, Speech Communication and Signal Processing Group
School of Electrical and Computer Engineering, Athens, 15773, Greece.
{jkokkin,gevag,maragos}@cs.ntua.gr, http://cvsp.cs.ntua.gr/

ABSTRACT

In this paper we incorporate recent results from AM-FM models for texture analysis into the variational model of image segmentation and examine the potential benefits of using the combination of these two approaches for texture segmentation. Using the Dominant Components Analysis (DCA) technique we obtain a low-dimensional, yet rich texture feature vector that proves to be useful for texture segmentation. We use an unsupervised scheme for texture segmentation, where only the number of regions is known a-priori. Experimental results on both synthetic and challenging real-world images demonstrate the potential of the proposed combination.

1. INTRODUCTION

The segmentation of textured images is a long standing problem in Computer Vision, which has been addressed from various perspectives, with variational models, e.g.[11, 21, 16, 20] and MRFs, e.g. [5, 13] being the most common approaches. As image intensity is a poor cue for texture segmentation, filtering the image with a Gabor filterbank is commonly used as a feature-extraction preprocessing step which results in detecting texture information that resides at different frequency channels. Region based segmentation algorithms which subsequently group pixels into regions according to the proximity of the filter responses at these points are more global and therefore more efficient, contrary to edge based algorithms [12] which result in spurious edges, due to the inherently random nature of textures.

In this paper we examine the potential benefits of using a more compact model for texture analysis that has been recently presented in [8], namely Dominant Component Analysis (DCA). This representation summarizes all the texture information in a 3 dimensional feature vector, which in a loose sense best models the texture at each point using the AM-FM modulations model of images. This technique provides us with a low-dimensional, yet rich feature set, that proves to be useful for segmentation. As a segmentation algorithm we use curve-evolution implemented with the level-sets technique, where the forces that drive the evolution of the contours are determined by a region-based probabilistic criterion, as in [21, 15].

In the following section we present the necessary background for our segmentation algorithm, while briefly mentioning previous work, subsequently we mention the details of our approach and finally we present experimental results that testify the power of the combination of the modulation features with the region-based curve evolution approach.

This work has been supported partially by the GME research program Hrakleitos, the GSRT research program PENED-2001, the NTUA research program Protagoras and the EC NoE MUSCLE.

2. BACKGROUND

2.1. Curve Evolution for Textured Image Segmentation

In the variational framework for segmentation, which we adopt in our approach, a labelling of the image is searched for that minimizes a certain energy functional which encodes the desired features of a segmentation. This labelling is initialized by assigning the same label to pixels inside a closed contour while the energy criterion is expressed in terms of these contours; The segmentation of the image is derived by numerically solving curve evolution PDEs, where the forces that drive the evolution are determined by Euler-Lagrange equations which give the direction of steepest descent of the energy functional.

In [21] the functional that was proposed to be minimized was the likelihood of the data inside each region R_i contained in the interior of curve Γ_i:

$$E[\Gamma, P_i] = \sum_{i=1}^{N} \iint_{R_i} -\log(P_i(I))\mathrm{d}x + \nu/2|\Gamma_i| \quad (1)$$

where P_i is the PDF of the feature vectors inside each region, N is the number of regions and ν is a weighting factor, punishing nonsmooth curves. The probability distribution inside each region R_i is considered Gaussian; specifically, for texture segmentation, where high-dimensional feature vectors are commonly used, multivariate Gaussians are used to model their distributions inside each region:

$$P_i(I; \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}} e^{-\frac{1}{2}(I-\mu_i)^T \Sigma_i^{-1}(I-\mu_i)} \quad (2)$$

Minimizing (1) w.r.t Γ results in the curve evolution equations:

$$\frac{\mathrm{d}\Gamma_i}{\mathrm{d}t} = (\nu\kappa_i + \log(P_i(I)/P_c(I))\vec{N}_i \quad \forall i \tag{3}$$

where P_c is the likelihood of the feature I under the competing neighboring hypothesis, c, κ_i is the curvature of Γ_i and $\vec{N}_i$ is the normal to Γ_i. Minimizing (1) w.r.t. μ_i, Σ_i results in setting the parameters of the Gaussian distributions to their ML estimates. Iterating the minimization w.r.t. μ_i, Σ_i and Γ_i results in a greedy algorithm for the minimization of (1) conceptually similar to the EM algorithm. The parametric distributions employed in [21, 15] have been recently replaced with non-parametric distributions e.g. in [9, 17], which result in similar evolution equations.

In [15] this idea was implemented and extended using the level-set technique, which has many computational advantages, since it handles automatically topological changes and lends itself to efficient implementations.

A problem that arises when filtering the image with a Gabor filterbank is the high dimensionality of the derived feature vector at each point, which results in many potential suboptimal segmentations as grouping the data in high dimensional spaces becomes hard. In the supervised texture segmentation case e.g.[16] this problem can be bypassed by choosing these channels that maximally separate different textures. It is however harder to tackle the unsupervised problem, since choosing the 'best' channels - which is equivalent to projecting the features onto a subspace where some *unknown a-priori* classes become maximally separated- is usually performed using heuristic criteria, as e.g. in [18].

In a recent attempt to alleviate this problem, Rousson et al. [17] used a vector valued diffusion procedure to smooth a low dimensional image feature set, which gave good texture segmentation results, using a 4-dimensional feature vector. In the information theoretic approach of [9] textured image segmentation is accomplished without using a feature extraction stage, using as the sole criterion the maximization of the mutual information between region label and image intensity. In the work presented here we are using features that are based on a different approach to texture analysis, which builds on the AM-FM model [3, 7] of textured signals.

2.2. Dominant Components Analysis

Locally narrowband 2D signals can be modeled as spatial AM–FM structures

$$f(x, y) = a(x, y)\cos[\phi(x, y)], \quad \vec{\Omega}(x, y) = \nabla\phi(x, y) \tag{4}$$

that are 2D nonstationary sines with a spatially varying amplitude $a(x, y)$ and a spatially-varying instantaneous frequency vector $\vec{\Omega}(x, y) = (V(x, y), U(x, y))$. Particularly in image signals, the amplitude is used to model local image contrast and the frequency vector contains information about the locally emergent spatial frequencies. These modulation models have been proposed by Bovik et al. [3] and extended by Havlicek [7] for wideband image signals. Any image can be thought of as a sum of locally smooth, narrowband modulating signals as:

$$I(x, y) = \sum_{k=1}^{K} a_k(x, y)\cos[\phi_k(x, y)], \tag{5}$$

An efficient and computationally simple approach for extracting the 2D amplitude and frequency signals was developed in [14] based on the 2D Teager-Kaiser Energy Operator $\Psi(f) \triangleq \|\nabla f\|^2 - f\nabla^2 f$. Applying Ψ to a 2D AM–FM signal f yields approximately the product of the instantaneous amplitude and frequency magnitude squared. Using as f bandpass filtered versions of an image has a regularizing effect, thereby dealing with stability issues of the Teager operator. If we apply the energy operator on the image derivatives $\partial f/\partial x$ and $\partial f/\partial y$, it is possible to separate the energy into its amplitude and frequency components via a nonlinear algorithm called Energy Separation Algorithm (ESA) [14]. Multiband filtering approaches [2, 1] for texture analysis are applied to decompose an AM–FM modelled image of the form (5) into narrowband, locally varying components. The Dominant Component Analysis (*DCA*) scheme [8, 6], chooses at each pixel the most powerful of these channels and estimates the AM-FM model parameters at that point using the outputs of that channel. This way, the outputs of the filtered images are combined, resulting in a low dimensional texture descriptor.

The only previous work we are aware of where it has been attempted to couple modulation features with curve evolution models is [22] where a geodesic active contour has been used to perform texture segmentation on a modulation based feature space. However the segmentation problem is formulated as a data clustering problem and a purely statistical algorithm performs most of the segmentation task.

3. VARIATIONAL SEGMENTATION ON A MODULATION BASED FEATURE SPACE

We have approached the basic DCA using a modified decision logic based on the Teager-Kaiser Energy Operator. In order to capture image modulation information at various scales and orientations, a bank of Gabor filters h_i has been used since they are compact, smooth and attain the lower bound in a time-frequency uncertainty relationship [4]. Each bandpass image $I_i = I * h_i$ is demodulated via the ESA and at a pixel-wise basis, a value is kept for the amplitude and frequency signals from the bandpass image I_i that maximizes the Energy Operator response:

$$i = \arg\max_j \Psi[(I * h_j)(x, y)] \tag{6}$$

Due to lack of space, details about the methodology used for extracting dominant component features can be found in our companion paper [10].

The feature vector we have used consists of the following components: a) Amplitude b) Phase Gradient Magnitude c) Phase Orientation d) Image Intensity. We include image intensity in our feature set as in [17], since this is still an important feature for non-textured regions.

Curve evolution has been implemented using level-set methods [19], where a very similar architecture with the one proposed in [15] has been used to implement region competition. An explicit scheme has been used, and no boundary based term was used. As in [18, 17, 20] the distribution of the data inside each region is learned in parallel with the evolution process, resulting in an adaptive scheme. For simplicity we model the distribution of the feature vectors inside each region with a multidimensional Gaussian with diagonal covariance matrix, which is reasonable, since the features we use should be uncorrelated. We note that the phase orientation data should be modelled using a von Mises distribution, which is the analog of the normal distribution for angular data. For the examples used in this paper a Gaussian distribution worked well, even though this is not guaranteed in case the orientations inside a region are around 0 and π.

We observed that using more robust estimates of the Gaussian distribution parameters results in greater invariance to initialization: we used the α-trimmed mean of the data and the α-trimmed mean absolute deviation of the points from this value, for $\alpha = 10\%$; the deviation was normalized with the factor $1/.8$ to compensate for the reduction in the variance caused by using a subset of the sample that is closer to this value. This gracefully deals with spuriously high amplitude estimates, caused occasionally by the ESA algorithm, at places where pre-smoothing does not eliminate errors. At a pre-processing level we have experimented with the coupled diffusion of the vector-valued data, like in [17], where an orientation diffusion term is introduced into the evolution equations for the directions feature channel, in order to obtain smoother estimates. No significant changes in performance were observed using the latter enhancements, since the modulation features are sufficiently smooth, contrary to the nonlinear structure tensor [17] data.

4. EXPERIMENTAL RESULTS

In the results presented here our only intervention has been in predefining the number of regions in the image; even though various statistical criteria can be used to find the 'correct' number of regions we believe this is a very hard problem to solve automatically, since even humans may disagree about the correct number of segments in an image.

In Fig. [1] we show how the system performs with some simple synthetic images: one can note from the bottom row that scale information is included in the feature vector, contrary to [17]; the inner texture is a scaled version of the outer texture, so the magnitude of the frequency vector helps discriminate among them. In Fig. [2] results with a real image are demonstrated, which contains both textured and non-textured regions. In Fig. [3] we notice that columns have been detected as a unified textured region, as the amplitude strength is almost constant, while the flowers, the steps and the bushes form separate regions.

5. DISCUSSION & CONCLUSION

Comparing our proposed algorithm for texture segmentation to related work on unsupervised curve-evolution based texture segmentation we believe it is advantageous in the following aspects:

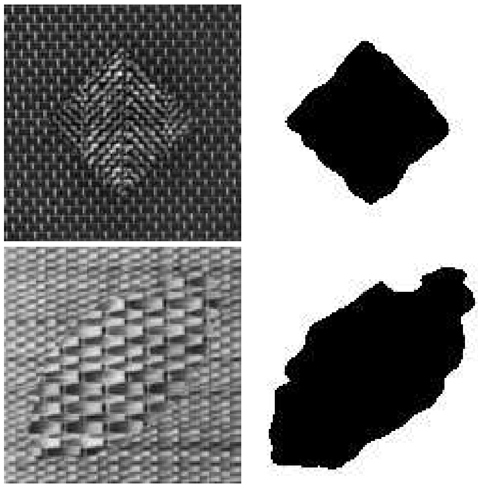

Fig. 1. Results with synthetic textures

- The representation of the texture in terms of its dominant components results in a low-dimensional feature space, where it is no longer necessary to search in a supervised or unsupervised setting for low-dimensional projections of the feature set, as e.g. in [18, 16]. However this low-dimensional feature vector is expressive enough for the discrimination of a wide variety of textures.
- Comparing our work with the most up-to-date publication where low-dimensional feature spaces are used [17], we mention that the feature set we use is richer in its expressive ability, even though it is of the same dimensionality: texture scale is naturally represented by the gradient magnitude, while good feature localization is achieved without anisotropic diffusion being necessary.
- The fact that modulation features can be used to reconstruct a signal, makes them interpretable in terms of *generative* models which model the observed data. As such, they are amenable to a probabilistic treatment, and comparable to other features, that may compete for the 'explanation' of the observed data. In our companion paper [10], we explore the ability of incorporating the extraction of modulation features with the discrimination of textured/non-textured regions, which is very active area of current research, and where we have obtained promising results.
- The Geodesic Active Contour model used in [22] is used at a post-processing stage to eliminate small & fragmented regions. The most important part of the segmentation is accomplished during the previous step, where the image data are clustered using a purely feature-driven criterion, disregarding any geometrical information. This diminishes the benefits of the region competition framework, that allows region, geometry as well as shape based knowledge to be incorporated into the evolution equations, thereby interleaving the use of geometrical and statistical information.

Summing up, we believe that the proposed method is characterized by simplicity and efficiency and combines the best of the DCA and curve evolution methods. Even though a three dimensional feature vector cannot discriminate between *any* set of textures, promising results have been obtained on both synthetic and real images, showing that our

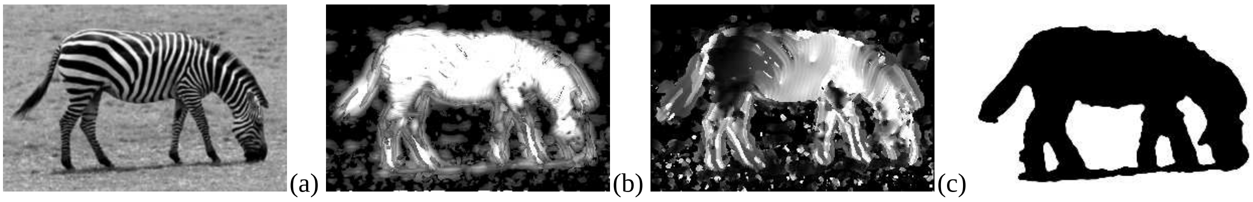

Fig. 2. (a) Input Image, (b,c) Amplitude, ϕ_x estimates using DCA, (d) Segmented image

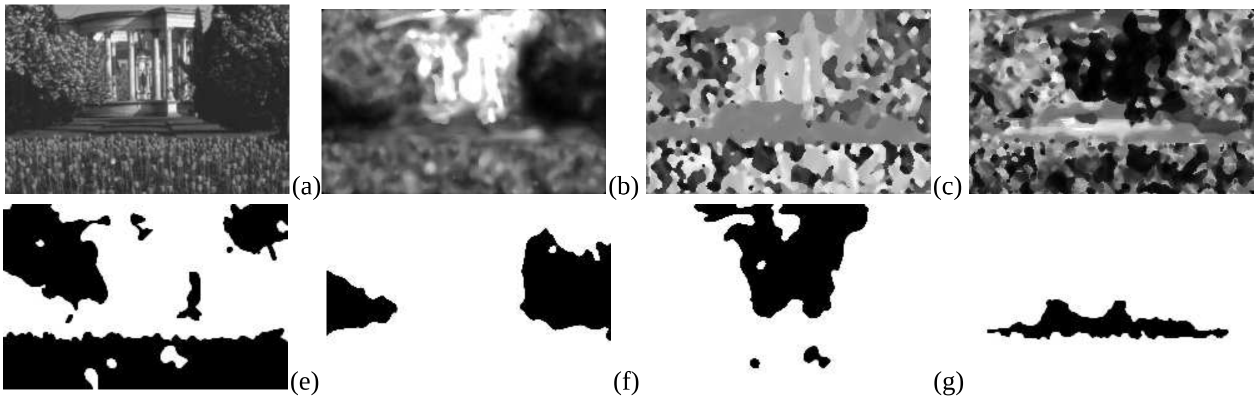

Fig. 3. (a) Input Image, (b-d) Amplitude, ϕ_x, ϕ_y estimates using DCA, (e-h) Detected regions using curve evolution.

method is applicable to a wide variety of textured images.

6. REFERENCES

[1] A.C. Bovik, "Analysis of multichannel narrow-band filters for image texture segmentation," *IEEE Trans. on Sig. Proc.*, vol. 39(9), pp. 2025–2043 1991.

[2] A.C. Bovik, M. Clark, and W.S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. on PAMI*, vol. 12(1), pp. 55–73, 1990.

[3] A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, "Localized measurement of emergent image frequencies by gabor wavelets," *IEEE Trans. Info. Theory*, vol. 38(2), pp. 691–712, 1992.

[4] J. E. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *JOSA A*, vol. 2(7), pp. 1160–1169, 1985.

[5] D. Geman, S. Geman, C. Graffigne, and P. Dong, "Boundary detection by constrained optimization," *IEEE Trans. on PAMI*, vol. 12(7) pp. 609–628, 1990.

[6] J.P. Havlicek and A.C. Bovik, "Image modulation models," in *Handbook of Image and Video Processing*, A.C. Bovik, Ed., pp. 305–316. 2000.

[7] J.P. Havlicek, D.S. Harding, and A.C. Bovik, "The mutli-component am-fm image representation," *IEEE Trans. Image Proc*, vol. 5(6), pp. 1094–1100, 1996.

[8] 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(2), pp. 227–242, 2000.

[9] J. Kim, J. Fisher, A. Yezzi, M. Cetin, and A. Willsky, "Nonparametric methods for image segmentation using information theory and curve evolution," in *Proc. ICIP*, 2002.

[10] I. Kokkinos, G. Evangelopoulos and P. Maragos, "Advances in texture analysis: Energy dominant components and multiple hypothesis testing," in *Proc. ICIP*, 2004.

[11] T.S. Lee, D. Mumford, and A. Yuille, "Texture segmentation by minimizing vector-valued energy functionals: the coupled-membrane model," in *Proc. ECCV*, 1992.

[12] J. Malik and P. Perona, "Preatentive texture discrimination with early vision mechanisms," *JOSA A*, vol. 7(5), pp. 923–932, 1990.

[13] B.S. Manjunath and R. Chellappa, "Unsupervised texture segmentation using markov random field models," *IEEE Trans. on PAMI*, vol. 13(5), pp. 478–482, 1991.

[14] P. Maragos and A.C. Bovik, "Image demodulation using multidimensional energy separation," *JOSA A*, vol. 12(9), pp. 1867–1876, 1995.

[15] N. Paragios and R. Deriche, "Geodesic active regions: A new paradigm to deal with frame partition problems in computer vision," *JVCIR*, vol. 13(1-2), pp.249-268, 2002.

[16] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," *IJCV*, vol. 46(3), pp. 223-247, 2002.

[17] M. Rousson, T. Brox, and R. Deriche, "Active unsupervised texture segmentation on a diffusion based feature space," in *Proc. CVPR*, 2003.

[18] B. Sandberg, T. Chan, and L. Vese, "A level-set and gabor-based active contour algorithm for segmenting textured images," *UCLA CAM Report 02-39*, 2002.

[19] A. Sethian, *Level Set Methods and Fast Marching Methods.*, Cambridge University Press, Cambridge University Press, 1999.

[20] A. Yezzi, A. Tsai, and A. Willsky, "A statistical approach to snakes for bimodal and trimodal imagery," in *Proc. ICCV*, 2001.

[21] S.C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation," *IEEE Trans. PAMI*, vol. 18(9), pp. 884–900, 1996.

[22] Z. N. Zray, J. Havlicek, S. Acton, and M. Pattichis, "Active contour segmentation guided by am-fm dominant component analysis," in *Proc. ICIP*, 2001.