COUPLED GEOMETRIC AND TEXTURE PDE-BASED SEGMENTATION

Anastasia Sofou, Georgios Evangelopoulos, and Petros Maragos

National Technical University of Athens, School of E.C.E, Athens 15773, Greece.
Email:[sofou,gevag,maragos]@cs.ntua.gr

ABSTRACT

In this paper, along with recent trends in segmentation using multiple image cues, we examine the integration of modulation texture features, image contrast and region size for decomposition of an image in homogenous regions. First, we propose the use of a morphological PDE-based segmentation scheme of the watershed type, based on seeded region-growing and level curve evolution with speed depending on contrast and size. Second, we analyze object surface texture by modelling image variations as local spatial modulation components estimated via multi-frequency filtering and instantaneous energy-tracking operators. By separately exploiting contrast and texture information, through multiscale image decomposition, we propose a PDE-based coupled segmentation method. Experimental results on various classes of images such as soilsections, aerial and natural scenes indicate that the combined effect of image decomposition and multi-cue segmentation improves the overall segmentation process.

1. INTRODUCTION

Although image segmentation forms the basis of almost any image analysis task, it still remains a hard to solve problem since it appears to be application dependent with usually no a priori information available regarding the image structure. Moreover, the increasing demands of image analysis tasks in terms of segmentation results' quality introduce the necessity of employing multiple cues for improving image segmentation results. In this paper we attempt to incorporate cues such as intensity contrast, region size and texture in the segmentation procedure and derive improved results compared to using individual cues separately.

Based on the well known morphological paradigm of watershed transform segmentation which exploits intensity contrast and region size criteria, we introduce a watershed-like segmentation scheme which couples contrast and texture information. The modelling of the proposed scheme is done via Partial Differential Equations (PDEs) since they ensure better and more intuitive mathematical formulation, direct connections with physics, and better approximation to the continuous geometry of the problem. By well-motivated texture modelling, we propose efficient extraction of image features capable of quantifying important characteristics like geometrical complexity, rate of change in local contrast variations and orientation. Finally, by selecting the $U+V$ image decomposition as an efficient way to incorporate the available contrast and texture information, we propose a coupled segmentation scheme driven by two separate image components: Cartoon U (for contrast information) and Texture component V.

This work was supported by the Greek GSRT research program PENED-2001 , the European NoE MUSCLE, the Greek Educ. Ministry program 'Pythagoras' and the NTUA research program 'Protagoras'.

2. GENERALIZED WATERSHED AND PDES

Image segmentation is one of the most important but yet difficult tasks in computer vision, as it requires to some extent a semantic understanding of the image. Amongst segmentation methods, the morphological watershed transform has proved to be robust, powerful and effective, especially when coupled with nonlinear multiscale morphological operators [1, 2]. Apart from the well-known morphological flooding approach implemented either via immersion simulations [2] or hierarchical queues [1], the watershed transform has also been modelled in a continuous way via the eikonal Partial Differential Equation (PDE) [3, 4], and implemented in [5] using curve evolution and level sets. Using PDE modelling in the flooding process of watershed transform, each emanating wave's boundary is viewed as a curve, which evolves with predefined speed. In the case of uniform height watershed flooding, let us consider a moving smooth closed curve, which is the boundary of the marker, $\vec{C}(p,t)$ where $p \in [0,1]$ parameterizes the curve and t is an artificial marching parameter. The PDE that implements the watershed flooding is of the form:

$$\frac{\partial \vec{C}}{\partial t} = \frac{c}{A(t)\|\nabla I\|} \cdot \vec{N} \tag{1}$$

where c is a constant, $\|\nabla I\|$ is the gradient magnitude of the image function I, $\vec{N}$ is the unit outward vector normal to the curve, and $A(t)$ is either 1 if we perform only contrast-based segmentation (height flooding) or $A(t) = \mathrm{Area}(\vec{C})$ in case of contrast and size segmentation (volume flooding) [6]. The above propagation PDE implies that the evolution speed is inversely proportional to the intensity (volume) variation at each image point, in the direction of the outward normal vector. For implementation we use the level set approach [7] where at each time the evolving curve is embedded as the zero level set $\Gamma(t) = \{(x,y) : \Phi(x,y,t) = 0\}$ of a higher dimension space-time function $\Phi(x,y,t)$. Then this embedding function Φ evolves in space-time according to the following PDE:

$$\frac{\partial \Phi}{\partial t} = \frac{c}{A(t)\|\nabla I(x,y)\|}\|\nabla \Phi\| \tag{2}$$

Modelling generalized watersheds via the eikonal has the advantage of a more isotropic flooding but it also introduces some challenges in the implementation. In general, efficient algorithms [8] to solve time-dependent eikonal PDEs are the narrow-band level sets methods, and more specifically, the fast marching method, an algorithm for stationary formulations of eikonal PDEs.

3. TEXTURE MODULATION ANALYSIS

Elementary natural texture components can be interpreted as locally smooth modulations and hence assumed nonstationary signals, well localized within a narrow band in the spatial frequency plane. As so, textured surfaces can be modeled by a sum of 2D spatial *Amplitude and Frequency Modulation* (AM–FM) signals

$$I(x,y) = \sum_{k=1}^{K} a_k(x,y)\cos[\phi_k(x,y)] \tag{3}$$

where each of the K components is a 2D nonstationary sine with a spatially-varying amplitude $a_k(x,y)$ and a spatially-varying instantaneous frequency vector $\vec{\omega}_k(x,y) = \nabla\phi_k(x,y)$. The amplitude is used to model local image contrast variations and the frequency vector contains rich information about the locally emergent spatial frequencies [9, 10].

Estimation of the 2D modulation signals can be done simply, efficiently and with small estimation error [11] based on an energy operator $\Psi(f) \triangleq ||\nabla f||^2 - f\nabla^2 f$, which is a multidimensional extension of the 1D Teager energy operator. Applying Ψ to a 2D AM–FM signal $f(x,y) = a_k(x,y)\cos[\phi_k(x,y)]$ modelling a texture component yields

$$\Psi[a_k \cos(\phi_k)] \approx a_k^2 ||\vec{\omega}_k||^2 \tag{4}$$

The product in (4), which couples the squares of the instantaneous amplitude and frequency magnitude may be called the *texture modulation energy*. By the assumption that the instantaneous amplitude and frequency do not vary rapidly in space or too greatly in value compared with the carriers, the above approximation error becomes negligible.

Modulation models are applied on bandpass filtered versions of an image [9] through filtering mechanisms with sufficient spatial and spectral localization, usually done by banks-of-filters spanning various radial frequencies and orientations. To obtain representations indicative of the dominant texture components we recently proposed an energy tracking mechanism in the multidimensional feature space consisting of the filter responses [12]. Specifically the narrowband texture components in the outputs of a Gabor filterbank are subjected to energy measurements via the 2D energy operator Ψ, post-averaged by a local averaging filter h_a and compared pixelwise. The filter with the *Maximum Average Teager Energy* per pixel, given by

$$\Psi_{\mathrm{mat}}(I(x,y)) = \max_k \Psi[((I * h_k) * h_a)(x,y)], \tag{5}$$

where $*$ denotes 2D signal convolution and h_k the response of filter k, indicates the most prominent texture component. The derived Ψ_{mat} is a slowly-varying indication of texture modulation energy, which can classify various energy levels and thus different textures with respect to their Teager energy signatures.

By demodulating the most-active in this energy sense texture components we can also derive a low-dimensional feature set which was found to be useful for unsupervised curve-evolution based textured image segmentation [13]. In this work however we examine and propose the modulation energy Ψ_{mat} as a compact, efficient and well localized texture detection cue.

4. COUPLED SEGMENTATION SCHEME

A recently proposed method for image decomposition is the $I = U + V$ model [14, 15], where U is the "cartoon component" and consists of relatively flat plateaus for the object regions surrounded by abrupt edges, whereas V is the "texture oscillation" and contains texture plus noise information. By treating and processing the two components separately a powerful joint segmentation scheme is proposed. Contrast variations are taken into account from the U part and texture oscillations are approached through modulation analysis on the V component.

Several nonlinear edge-preserving image smoothing schemes can create cartoon approximations of an image such as anisotropic diffusion and image selective smoothing [15, 16]. To obtain the cartoon component U we apply the *leveling* operator [17] on the initial image, motivated by its attractive properties, as well as by the fact that it is also used as a non-linear simplification filter at the pre-processing stage of the segmentation procedure. More precisely, levelings are nonlinear object-oriented filters that simplify a reference image I through a simultaneous use of locally expanding/shrinking an initial seed image, called the marker M, and a global constraining of the marker evolution by the reference image. Specifically, iterations of the image operator $\lambda(F|I) = (\delta(F) \wedge I) \vee \varepsilon(F)$, where $\delta(F)$ (resp. $\varepsilon(F)$) is a dilation (resp. erosion) of F by a small disk, yield in the limit the leveling of I w.r.t. M, denoted as $\Lambda(M|I) = \lim_{k\to\infty} F_k,\ F_k = \lambda(F_{k-1}|I),\ F_0 = M$. The levelings have many interesting multiscale properties, such as reconstruction of whole image objects with exact boundary preservation if the objects are hit by the marker.

As texture component we use the residual $V = I - U = I - \Lambda(M|I)$. We construct multiscale leveling cartoons U_i by using a sequence of multiscale markers M_i, obtained from sampling a Gaussian scale-space. The corresponding residuals $V_i = I - U_i$ constitute a hierarchy multiscale texture components. As an alternative marker selection, we consider the use of anisotropic diffusion [16]. At each sequence step the leveling marker is obtained by a version of the image with blurred regions but adequately preserved boundaries, caused by the constrained diffusion process. It should be noted here that the scope of this paper is not to find the best $U + V$ decomposition but some efficient decomposition that couples with segmentation methods.

The proposed new segmentation scheme is based on the following curve evolution PDE

$$\frac{\partial \vec{C}}{\partial t} = \left(\frac{\lambda_1}{A||\nabla f_1||} + \lambda_2 \Psi_{\mathrm{mat}}(f_2) - \mu\,\kappa \right) \vec{N} \tag{6}$$

where f_1 and f_2 are image transformations related to the original I, but not necessarily the same. Thus, the curve's speed depends on three terms: the first two are eikonal, whereas the third (curvature motion) is diffusive. All terms are linked with some optimality criterion. The first term drives the curve with speed that maximizes the *flooding* of the f_1 image toward its watershed. The second term can be shown to correspond to a flow that maximizes the average texture energy: $\max \iint_{R(C)} \Psi(f) \Longrightarrow \partial\vec{C}/\partial t = \Psi(f)\vec{N}$. This term pushes the curve toward regions with large average texture energy.

Following the level set formulation in [7], we embed this evolving planar curve as the zero-level curve of an evolving space-time function $\Phi(x,y,t)$, and conclude to the level function PDE:

$$\frac{\partial \Phi}{\partial t} = \left(\frac{\lambda_1}{A||\nabla f_1||} + \lambda_2 \Psi_{\mathrm{mat}}(f_2) - \mu\,\mathrm{curv}(\Phi) \right) ||\nabla\Phi|| \tag{7}$$

where $\mathrm{curv}(\Phi)$ is the curvature of the level sets of Φ.

Based on the PDE (6), different scenarios can be obtained by varying the signals f_1 and f_2. The most obvious choice is $f_1 = I$, $f_2 = I$, but we also propose another novel and promising scenario which is $f_1 = U$, $f_2 = V$. The former is a curve evolution

Quality Criteria	Segmentation Method			
	CTS	VTS	WT	KM
YLGC	0.25	0.09	0.6	7.7
MSF	3.25	3.30	3.40	3.69

Table 1. Quantitative evaluation by cost functionals.

with velocity inversely proportional to the intensity contrast (or volume) of the input image and proportional to the Ψ_{mat} energy of the image. The latter is a curve evolution with speed inversely proportional to the intensity contrast (or volume) of the cartoon component and proportional to the Ψ_{mat} energy of the textured component, and is favored and further investigated since it integrates edge and texture information by combining the different signals produced by the $U + V$ decomposition of the image.

In the proposed segmentation scheme, there are multiple curves to be propagated, which are initialized as the contours of a set of *markers*, indicative of significant image regions. Since markers provide evidence about the existence of homogeneous regions, a variety of methodologies has been developed for their extraction. Thus, markers can be: (i) contrast-oriented corresponding to peaks or valleys of certain depth (obtained via reconstruction filters); (ii) peaks of $\Psi_{\mathrm{mat}}(I)$, indicating areas with rich texture; (iii) combination of contrast as well as texture criteria; and (iv) manually placed at areas of interest. In general, the methodology of marker extraction is application dependent and beyond the scope of this paper. The implementation of (6) has been done with established techniques from level sets methods. If $\mu = 0$ the PDE is of pure eikonal-type and its implementation is based on the fast marching methodology (FMM) [8], which ensures computational speed. If $\mu \neq 0$ the PDE is implemented using the narrow band method [8], and the segmentation boundaries are smoothed. The decision of whether or not to use the κ-term is application dependent.

5. RESULTS AND APPLICATIONS

The proposed coupled method for texture analysis after decomposition and multi-cue segmentation has been applied to various classes of images. Such experimental results are illustrated in Fig. 1 - 3. In [13], the region based scheme is tuned to texture only, whereas here there is a coupling between texture detection and geometric segmentation.

In Fig.1 the scheme is applied on a soilsection image under polarized light, where grouping of various materials and structures is achieved. In Fig.1(a)-(d) the initial image, the leveling cartoon, texture and energy component, are respectively illustrated. Initial segmentation results produced using the marker set of Fig.1(e) are presented in Fig.1(g). Due to the multi-structured nature of such images, these results can be refined by an area post-merging procedure, as shown in Fig.1(h). Briefly, neighboring regions with minimum squared distance of mean region features are merged, constrained to obey a modified *Fisher criterion* on the variance and size of the combined region. A minimum distance is chosen for convergence by histogram values. Color vector and U, V intensities were used here as merging cues. The proposed scheme (coupling contrast and texture information CTS as well as volume and texture information VTS) was tested against other simpler but established segmentation methods in order to verify its ability to improve segmentation results. Comparisons were performed against traditional watershed (WS), K-means clustering (KM) and recursive shortest spanning tree (RSST) split and merge algorithm [18], as illustrated in Fig.1(i-l)respectively. Quantitative comparison results can be found in Table.1 where as goodness criteria were used the *Liu-Yang Global Cost* (LYGC) [19] and *Mumford Shah Energy Functional* (MSF) [20]. The smaller their values, the better the segmentation results are considered. In Fig. 2 the segmentation scheme is applied on *Lena* image, which is characterized both by edges (strong and medium) and texture, using manual markers as shown in Fig. 2(a). The proposed approach in (e) outperformed the results of the single-cue, contrast-based segmentation and the multi-cue model of Eq.(7) with no decomposition, $f_1 = f_2 = I$, shown in (d). Finally in Fig.3 results on four different image classes are illustrated. In all cases, a sufficient choice of markers led to satisfactory region labeling.

From these and analogous tests we conclude that: 1) Using the combined edge-contrast and texture information, the proposed segmentation scheme ensures better results than using single cues. 2) If the edge-contrast and texture information are separately taken from the U and V components rather than the initial image, segmentation results are further improved, since texture areas are better located, described and extracted.

6. REFERENCES

[1] S. Beucher and F. Meyer, "The morphological approach to segmentation: The watershed transformation," in *Math. Morphology in Im. Proc.*, E.R.Dougherty, Ed. Marcel Dekker, New York, 1993.

[2] L. Vincent and P. Soille, "Watershed in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. PAMI*, 13:583-598, 1991.

[3] L. Najman and M. Schmitt, "Watershed of a continuous function," *Signal Processing*, 38(7):99-112, 1994.

[4] F. Meyer and P. Maragos, "Multiscale morphological segmentations based on watershed, flooding, and eikonal pde," in *Proc. of Scale-Space*:351-362, 1999, Springer.

[5] P. Maragos and M. A. Butt, "Curve evolution, differential morphology, and distance transforms applied to multiscale and eikonal problems," *Fundamenta Informaticae*, 41:91-129, 2000.

[6] A. Sofou and P. Maragos, "PDE-based modelling of image segmentation using volumic flooding", in *Proc. IEEE ICIP*, 2003.

[7] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations," *J. Comp. Physics*, 79:12-49, 1988.

[8] J. A. Sethian, *Level Set Methods and Fast Marching Methods*, Cambridge Univ. Press, 1999.

[9] A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, "Localized measurement of emergent image frequencies by gabor wavelets," *IEEE Trans. Info. Theory*, 38:691-712, 1992.

[10] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "Multidimensional quasi-eigenfunction approximations and multicomponent AM–FM models," *IEEE Trans. Im. Proc.*, 9(2):227-242, 2000.

[11] P. Maragos and A. C. Bovik, "Image demodulation using multidimensional energy separation," *J. Opt. Soc. Amer. A*, 12(9):1867-1876, 1995.

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

[13] I. Kokkinos, G. Evangelopoulos, and P. Maragos, "Modulation-feature based textured image segmentation using curve evolution," in *Proc. IEEE ICIP*, 2004.

[14] Y. Meyer, "Oscillating Patterns in Image Processing and Nonlinear Evolution Equations", University Lecture Series Vol. 22, AMS 2002.

[15] L. Vese and S. Osher, "Modeling textures with total variation minimization and oscillating patterns in image processing," *J. Sci. Comp.*, 19:553-572, 2003.

[16] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. PAMI*, 12(7):629-639, 1990.

[17] F. Meyer and P. Maragos, "Nonlinear scale-space representation with morphological levelings," *J. Vis. Commun. & Im. Repr.*, 11:245-265, 2000.

[18] S. H. Kwok and A. G. Constantinides, "A fast recursive shortest spanning tree for image segmentation and edge detection," *IEEE Trans. Im. Proc.*, vol. 6, no. 2, pp. 328–332, 1997.

[19] J. Liu and Y. Yang, "Multiresolution color image segmentation," *IEEE Trans. PAMI.*, 16(7):689-700, 1994.

[20] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, 42:577685, 1989.

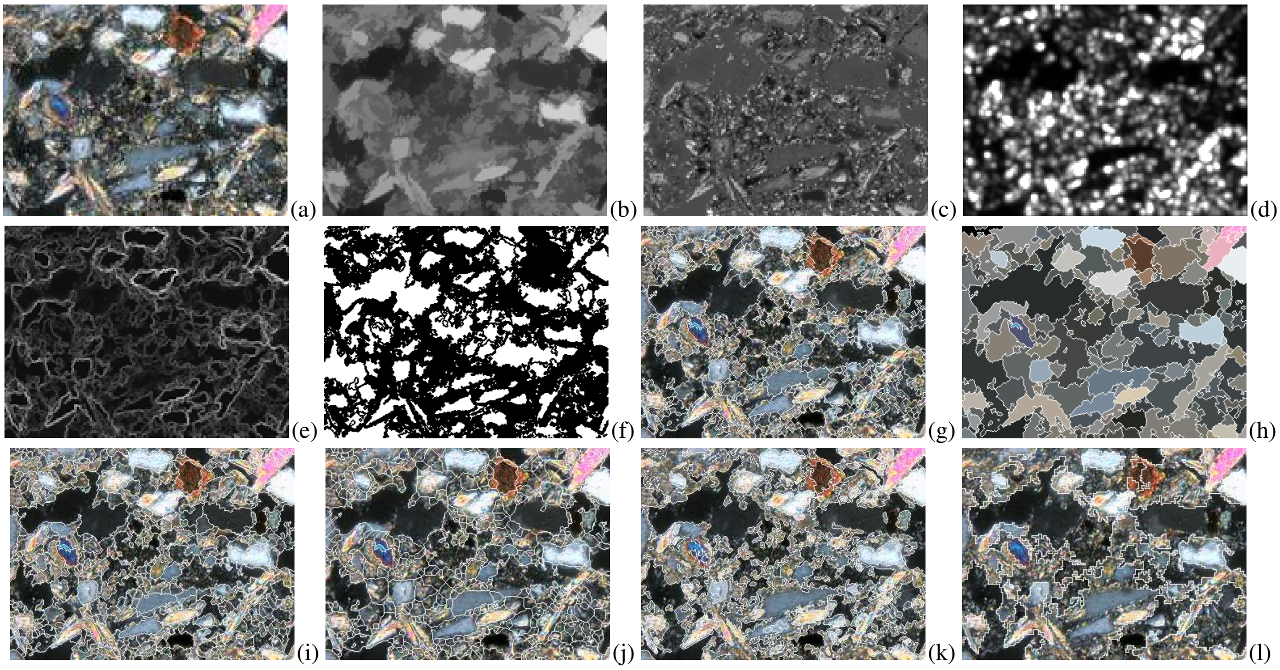

Fig. 1. **Soilsection Image Segmentation**: **(a)** Soilsection Image under polarized light **(b)** Leveling "Cartoon component" U, **(c)** "Texture component" V, **(d)** Texture Energy $\Psi_{\mathrm{mat}}(V)$, **(e)** Gradient of U, **(f)** Markers, **(g)** Coupled Contrast-Texture segmentation, **(h)** Refined Segmentation Regions, **(i)** Coupled Volume-Texture segmentation, **(j)** Watershed segmentation, **(k)** K-means clustering segmentation, **(l)** RSST segmentation.

Fig. 2. **Coupled Segmentation Scheme**: **(a)** Original Image and Imposed Markers **(b)** Leveling "Cartoon component" U **(c)** "Texture Component" V **(d)** Segmentation, $f_1 = f_2 = I$ **(e)** Segmentation, $f_1 = U, f_2 = V$

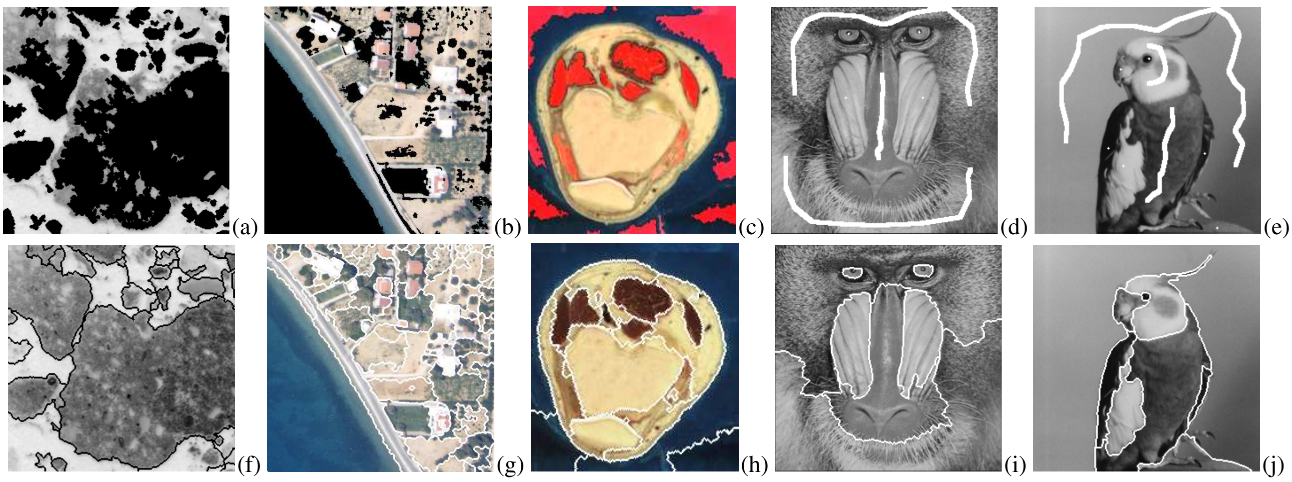

Fig. 3. **Set of Segmentation Results**: **(a)-(e)** *Soil, Aerial, Medical, Mandrill, Bird* images with corresponding set of markers superimposed, **(f)-(j)** Segmentation results from the proposed coupled segmentation scheme