

HIERARCHICAL SEGMENTATION VIA A DIFFUSION SCHEME IN COLOR/TEXTURE FEATURE SPACE

I. Vanhamel, A. Katartzis, H. Sahli

ETRO/IRIS, Vrije Universiteit Brussel
Pleinlaan 2
1050 Brussels, Belgium
(iuvanham,akatarzi,hsahli)@etro.vub.ac.be

ABSTRACT

This paper presents a segmentation scheme for images containing both smooth regions and textures. It is based on a vector-valued anisotropic diffusion on a combined color/Gabor feature space, followed by a hierarchical segmentation using dynamics of multiscale 'generalized gradient' watersheds. The proposed method gives good segmentation results and is shown to be more effective than its counterpart, which uses only multiscale color information.

Keywords:

Segmentation, color, texture, multiscale watersheds, anisotropic diffusion.

1. INTRODUCTION

Texture segmentation is a long-termed research field in computer vision [1]. It has been studied using various approaches. Since the early days of computer vision, the feature classification approach has become the essential texture analysis technique for the treatment of textured surfaces [1]. Recently, methods combining region and boundary information have been suggested [2]. These methods generate a texture feature space by filtering the image using Gabor filters. Texture information is then expressed (e.g. statistical measurements) and segmentation is achieved using geodesic contours. These approaches have been extended by using the Beltrami flow [3] for denoising the Gabor features, which is used as an inverse edge detector in the geodesic snakes mechanism [4].

In this paper, an early presented anisotropic diffusion scheme for color images [5] is applied to vector-valued images that describe both color and texture features, obtained using Gabor filtering. In the case of grey level images, a similar approach was proposed by Rubner and Tomasi in [6]. After the generation of the multiscale tower, a region-based segmentation is performed using a multiscale hierarchical scheme [5] [7], which allows the creation of a hierarchy among generalized gradient watersheds. The notion of a generalized gradient for color/texture contrast originates from differential geometry and is often used to estimate gradient of multi-spectral images [8].

The hierarchical segmentation scheme consists of three basic modules. The first module retrieves the texture feature images and combines them with the color information. The second module is dedicated to attribute a saliency measure to each contour arc at the localization scale (the finest scale in the scale-space stack), taking into account the whole multiscale tower. The entire process to

retrieve the saliency measure for the gradient watersheds requires three steps: (i) scale-space stack generation by vector-valued non-linear diffusion filtering. (ii) Linking: At each scale the generalized gradient is estimated using the gradient-tensor method [8]. At the localization scale, the watershed transformation is performed to identify the position of all the contours in the image. At the higher scales, the duality between the regional minima of the gradient and the catchment basins of the watershed is exploited to make a robust region based parent-child linking scheme. (iii) *Contour valuation by downward projection*. A combination of the dynamics of contours [9] over the scale-space stack - the dynamics of contours in scale-space - is used to valueate the contours detected at the localization scale. Finally, the last module of our method identifies the different hierarchical levels through a hypothesis testing criterion.

The paper is organized as follows. An overview of our hierarchical segmentation scheme is presented in Section 2. In particular, Section 2.1 deals with the Gabor transform and the generation of the color/texture feature space, while Sections 2.2 and 2.3 present the saliency measure and hierarchical level retrieval modules, respectively. Finally in Section 3 we illustrate segmentation results on a set of artificial and natural scene images and quote some conclusions.

2. HIERARCHICAL MULTISCALE SEGMENTATION

A large class of hierarchical schemes based on superficial or deep image structure have been adapted to deal with the segmentation problem. Among them we can cite [10][11][12][13][14].

Our main goal is to create a hierarchy among the generalized gradient watersheds that preserves the topology of the watershed lines at the localization scale and extracts objects of a larger scale. Let us first define what we mean by a hierarchy. Let $P^0 = \{S_1, S_2, \dots, S_n\}$ be the initial partitioning of the image at the localization scale after the application of the gradient watershed transformation. A **hierarchical level k** (HL_k) is defined as the partitioning $P^k = \{S_1^k, S_2^k, \dots, S_{m_k}^k\}$ which preserves the inclusion relationship $P^k \supseteq P^{k-1}$, implying that each segment of the set P^k is a disjoint union of segments from the set P^{k-1} . A **hierarchy of partitions** is defined as a family which consists of all the hierarchical levels HL_k , where $k \in [0, K]$. It corresponds to a **hierarchy of Region-Adjacency Graphs**, $\mathcal{G}_k(P^k, A^k)$, that are generated by applying successive mergings.

2.1. Color/Texture Feature Space

Spectral decomposition is a common way to describe texture in image processing. The texture content is usually represented as a vector-valued image, in which each decomposition band describes the energy at a given frequency and orientation. The spectral decomposition using Gabor filtering has often been justified by the fact that it provides a good approximation of the natural processes in the primary visual cortex.

A Gabor function is a harmonic wave modulated by a Gaussian. In [6], the *log*-Gabor filters are used, since natural textures often exhibit a linearly decreasing log power spectrum. In the frequency domain, the *log*-Gabor filter bank according to the implementation of Bigun et al. [15, 6] is defined as:

$$G_{ij}(\omega_r, \omega_\varphi) = G \omega_r - \omega_{r_i^0}, \omega_{\varphi_i^0} \quad (1)$$

where (r, φ) are polar coordinates, $\omega_{r_i^0}$ is the logarithm of the center frequency at scale i , $\omega_{\varphi_i^0}$ is the j^{th} orientation and $G(\omega_r, \omega_\varphi)$ is defined as:

$$G_{\omega_r, \omega_\varphi} = \exp \frac{-\omega_r^2}{2\sigma_{r_i}^2} \exp \frac{-\omega_\varphi^2}{2\sigma_{\varphi_j}^2} \quad (2)$$

where $\sigma_{r_i}^2$ and $\sigma_{\varphi_j}^2$ are the parameters of the Gaussian, $i \in [1, M]$ and $j \in [1, N]$ for M orientations and N scales. The N orientations are taken equidistant (Eq.3) and the scales are obtained by dividing the frequency range $\omega_{max} - \omega_{min}$ into M octaves (Eq.4).

$$\begin{aligned} \sigma_{\varphi_i} &= \frac{\pi}{2N} \\ \omega_{\varphi_i^0} &= 2\sigma_{\varphi_i}(j-1) \end{aligned} \quad (3)$$

$$\begin{aligned} \sigma_{r_i} &= 2^{i-1}\sigma \\ \omega_{r_i^0} &= \omega_{min} + 1 + 3 \cdot 2^{i-1} - 1 \cdot \sigma \end{aligned} \quad (4)$$

where $\sigma = \frac{\omega_{max} - \omega_{min}}{2(2^M - 1)}$ which yields M octaves $2\sigma, 4\sigma, \dots, 2^M\sigma$.

The texture feature images are the magnitudes of the responses of the Gabor filters. The latter encodes the energy content and is independent of the position within the texture. Each vector is further normalized to be a unit vector to emphasize the texture structure information and reduces the dependance of the responses on lighting. This procedure yields a MN hypersphere [6].

For color images spectral decomposition using Gabor filtering is mostly achieved by processing each color band separately. In [16], Gabor opponent features are used to describe texture. The Gabor responses of each color channel interact, according to the opponent process theory, with the responses of another color channel. In [17], the color image is transformed to a complex-scalar image that is based upon the *HSV* colorspace. It is on the complex scalar image that the Gabor filter is applied.

In this paper, the texture features for color images are obtained by filtering each color band separately with the Gabor filter bank for scalar images described above [6]. The latter yields $3 * MN$ hypersphere for the texture features and a $3+3NM$ -channel image for the combined color and texture information. We use $N = 4$ and $M = 4$ and limit the maximum frequency by the Nyquist frequency for all experiments.

2.2. Salient Measure Module

2.2.1. Scale-Space Generation

Anisotropic diffusion for image enhancement and denoising has been extensively studied by a number of authors in recent years [18] [19][20]. The idea is to smooth the image in a direction parallel to object boundaries and prevent diffusion across edges. Traditionally, "edges" are localized in those regions where the gradient of the image luminance is high. However, a measure based solely in luminance or color contrast does not always produce satisfactory results. Looking at other image attributes such as texture provides additional information that can help in localizing image discontinuity. The measure of image features (color, texture, ...) at a single location leads to a multi-valued representation of the image. Several vector diffusion PDE's have been proposed in the literature [21].

In this work, the multiscale tower of the color/texture feature space is generated using an anisotropic diffusion for vector-valued images [7]. It is based upon the regularized Perona and Malik anisotropic diffusion scheme [18] and it uses a system of coupled diffusion equations. The diffusion tensor is a function of the gradient of the color/texture feature space (generalized gradient). The latter is estimated using the gradient tensor method [8]. Furthermore the method adopts the Additive Operator Splitting numerical scheme(AOS) [22], which is computational more efficient than most other schemes.

2.2.2. Linking Scheme

The linking scheme aims to track the regional minima in the gradient through the scale-space. The linking process is applied using the approach proposed in [10]. The linking of the minima for successive scales is applied by using the *proximity* criterion [23]. This criterion is limited for projected minima of scale quantization level s_i inside the same **geodesic influence zone** $iz_A^{s_i+1}(B_q)$ of a connected component B_q of B in A at scale quantization level s_{i+1} :

$$iz_A^{s_i+1}(B_q) = \{p \in A, \forall j \in [1, k]/\{q\}, d_A(p, B_q) < d_A(p, B_j)\} \quad (5)$$

Any regional minimum of the set $\{m^{s_i}\}$, that is spatially projected on the geodesic influence zone $iz_A^{s_i+1}(B_q)$ at scale quantization level s_{i+1} , will be linked with the regional minimum $m_q^{s_{i+1}}$. The projected minimum of the set $\{m^{s_i}\} \in iz_A^{s_i+1}(B_q)$, which is the closest to the minimum $m_q^{s_{i+1}}$, is considered as the father. The rest of the projected minima onto the same influence zone are considered annihilated. Closeness is defined with respect to the topographic distance which is a natural distance measure following the steepest gradient path inside the catchment basin. At the end of the linking step, for each couple of neighboring segments (S_i, S_j) that share a common border at the localization scale, a linkage list $\Lambda(S_i, S_j)$ is constructed.

2.2.3. Contour Valuation by Down-projection

The next step is to valuate the gradient watersheds at the localization scale s_0 . For each segment couple $(S_i^{\Lambda(S_i, S_j)_m}, S_j^{\Lambda(S_i, S_j)_m})$, appearing at the branch m (scale quantization level s_m) of the linkage list $\Lambda(S_i, S_j)$, we compute the dynamics of contours $DC_{\Lambda(S_i, S_j)_m}^{s_m}$ [9]. It expresses how much contrasted the adjacent regions (S_i, S_j) are at the scale quantization level s_m . The dynamics of contours at each scale are normalized according to the maximum dynamic at that scale and finally the dynamics of contours in

scale-space (DCS) for the adjacent region couple (S_i, S_j) is defined as the sum for all normalized valuations during the evolution in scale-space:

$$DCS(S_i, S_j) = \sum_{m=0}^{N-1} \frac{DC_{\Lambda(S_i, S_j)_m}^m}{\max_{S_i, S_j \in \mathcal{S}_m} DC_{\Lambda(S_i, S_j)_m}^m} \quad (6)$$

where N denotes the branch of the linkage list $\Lambda(S_i, S_j)$ where the contour formed by the region couple (S_i, S_j) is annihilated. Each arc $a = (S_i, S_j) \in A^0$, of the RAG $\mathcal{G}_0(P^0, A^0)$, is ranked according to their saliency measure $DCS(S_i, S_j)$.

2.3. Hierarchical level retrieval

This module identifies the hierarchical levels. Starting from the watershed segmentation at the localization scale, a successive merging operation is performed until a stopping criterion is satisfied. The merging of adjacent segments is based upon a color similarity measure, while the merging sequence (order) is given by the valuation (dynamics of contours in scale-space) of the contours [24].

3. RESULTS AND DISCUSSION

In this paper, we introduce texture to the multiscale hierarchical segmentation method proposed in [5, 7]. The salient measure (Section.2.2) that creates the multiscale hierarchy, is now based upon color, as well as texture. On the other hand, the extraction of the different hierarchical levels uses solely color information.

In Fig.1, the evolution of both the color gradient and the generalized gradient in the scale-space tower for the image shown in Fig.2.ab, is illustrated. Diffusing the color feature space, enhances the contrast between the textures, while the diffusion the color-texture feature space smooths the textured areas so that the relative contrast of the real segments increases. In Fig.2. we compare the

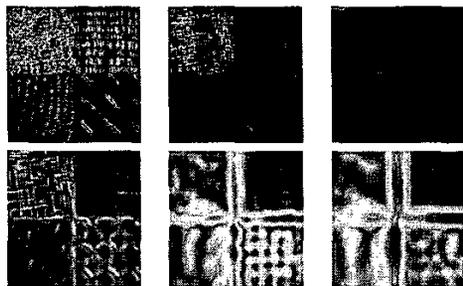


Fig. 1. Top. Color gradient in scale-space. Bottom. Generalized gradient in scale-space. The scale increases from left to right. Blue values correspond to low contrast and the red values denote areas of high contrast.

hierarchical levels that are retrieved from the hierarchy build using only color information (Fig.2a and Fig.2c) and the one generated using both color and texture (Fig.2b and Fig.2d). We demonstrate that for an image that is created using textures from the Vision Texture database¹ (Fig.2ab), the color based salient measure can

¹<http://www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>

not create a hierarchy that enables the separation of the different texture patterns, where as the color-texture based salient measure creates a meaningful hierarchy that allows the extraction of a hierarchical level that separates the four different textures. For natural scene images, both salient measures yield meaningful hierarchies. In the case of the color based salient measure, small but color-contrasting segments persist at the coarser hierarchical levels (Fig.2c). The color-texture based salient measure can identify these segments as part of textured segments and insures that this type of segment is only preserved in the finer hierarchical levels (Fig.2d).

The hierarchical level retrieval module simply identifies the different levels of abstraction in the created hierarchy. In order to select a final segmentation result, one needs to select the level of abstraction that is appropriate for further processing [7]. In Fig.3, we show the optimal hierarchical level for the hierarchy that is created by the color-texture based salient measure, according to a segmentation quality measure that relies upon (i) intra-segment uniformity, (ii) inter-segment contrast, and (iii) smoothness of the boundaries between segments [25, 24]. Fig.3a and Fig.3c are test images used by Petrou et al. in [26]. Both segmented images render the correct amount of segments. However, there are some deviations with the ideal segment border. In the case of Fig.3e, the segmentation renders 42 segments, which is somewhat over-segmented. Finally, Fig.3b, Fig.3d and Fig.3f are natural scene images that contain textured animals. For all three images, relatively good segmentations are obtained. However, there is a clear undersegmentation present. The latter can be solved by selecting a finer hierarchical level from the hierarchical stack. In the case of natural scene and color-textured images, we can conclude that the incorporation of texture features in our multiscale segmentation scheme leads to better results than the ones obtained using only color information.

4. REFERENCES

- [1] M. Tuceryan and A. K. Jain, "Texture analysis," in *Handbook of Pattern Recognition and Computer Vision*, 1993, pp. 235–276.
- [2] N. Pargios and R. Deriche, "Geodesic active region for supervised texture segmentation," in *Proc. of Int. Conf. on Computer Vision*, 1999, pp. 22–25.
- [3] N. Sochen, R. Kimmel, and R. Malladi, "A geometrical framework for low level vision," *IEEE Trans on IP*, vol. 7(3), pp. 310–318, 1998.
- [4] C. Sagiv, N.A. Sochen, and Y.Y. Zeevi, "Texture segmentation via a diffusion-segmentation scheme in the gabor feature space," in *2nd Int. Workshop on Texture Analysis and Synthesis (Texture 2002)*, 2002, pp. 123–128.
- [5] I. Vanhamel, I. Pratikakis, and H. Sahli, "Hierarchical segmentation using dynamics of multiscale color gradient watersheds," in *Scale-Space and Morphology in Computer Vision*, 2001, pp. 371–379.
- [6] Y. Rubner and C. Tomasi, "Coalescing texture descriptors," in *ARPA Image Understanding Workshop*, 1996.
- [7] I. Vanhamel, I. Pratikakis, and H. Sahli, "Multiscale gradient watersheds of color images.," *accepted for publication in IEEE Trans on IP*, 2003.

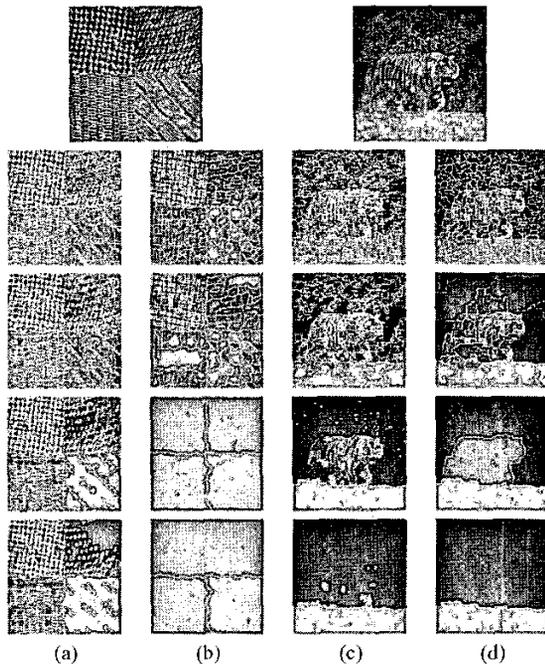


Fig. 2. Hierarchical Levels P^k . Top row: original images, Column (a) and (c): color-based salient measure. Column (b) and (d): color-texture based salient measure. From top to bottom. Column (a): $k = 0, 3, 6, 9$ with 2430, 2043, 1342 and 949 segments. Column (b): $k = 0, 2, 4, 5$ with 647, 215, 4 and 3 segments. Column (c): $k = 0, 5, 6, 9$ with 1879, 988, 166 and 11 segments. Column (d): $k = 0, 4, 8, 9$ with 600, 102, 3 and 2 segments.

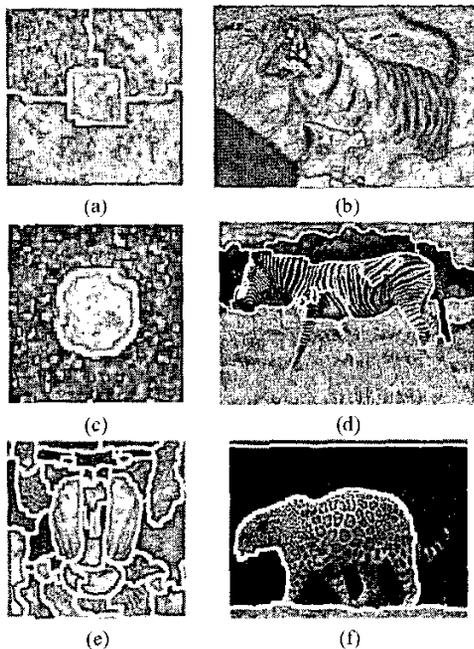


Fig. 3. Optimal hierarchical level selection: Original image with segment boundaries.

- [8] S. Di Zenzo, "A note on the gradient of a multi-image," *Computer Vision, Graphics and Image Processing*, vol. 33, pp. 116–125, 1986.
- [9] L. Najman and M. Schmitt, "Geodesic saliency of watershed contours and hierarchical segmentation," *IEEE Trans on PAMI*, vol. 18(12), pp. 1163–1173, 1996.
- [10] I. Pratikakis, H. Sahli, and J. Cornelis, "Low level image partitioning guided by the gradient watershed hierarchy," *Signal Processing*, vol. 75(2), pp. 173–195, 1998.
- [11] J.M. Gauch, "Image segmentation and analysis via multi-scale gradient watershed hierarchies," *IEEE Trans on IP*, vol. 8(1), pp. 69–79, 1999.
- [12] E. Sharon B. Brandt and R. Basri, "Fast multiscale image segmentation," in *CVPR*, 2000, pp. 1:70–77.
- [13] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans on PAMI*, vol. 22(8), pp. 888–905, 2000.
- [14] F. Meyer, "Hierarchies of partitions and morphological segmentation," in *Scale-Space and Morphology in Computer Vision*, 2001, vol. 3, pp. 161–182.
- [15] J. Bigün and J.M. du Buf, "N-folded symmetries by complex moments in gabor space and their application to unsupervised texture segmentation," *IEEE Trans on PAMI*, vol. 16(1), pp. 80–87, 1994.
- [16] A.K. Jain and G. Healey, "A multiscale representation including opponnent color faetures for texture recognition," *IEEE Trans on IP*, vol. 7(1), pp. 124–128, 1998.
- [17] C. Palm, D. Keysers, T. Lehmann, and K. Spitzer, "Gabor filtering of complex hue/saturation images for color texture classification," in *Joint Conf. on Information Sciences - Int. Conf. on Computer Vision, Pattern Recognition, and Image Processing*, 2000, vol. 2, pp. 45–49.
- [18] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans on PAMI*, vol. 12(7), pp. 629–639, 1990.
- [19] G. Sapiro and D.L. Ringach, "Anisotropic diffusion of multivalued images with applications to color filtering," *IEEE Trans on IP*, vol. 5(11), pp. 1583–1586, 1996.
- [20] J. Weickert, "A review of nonlinear diffusion filtering," in *Scale-Space Theory in Computer Vision*, 1997, pp. 3–28.
- [21] D. Tschumperle and R. Deriche, "Diffusion PDE's on vector-valued images," *IEEE Signal Processing Magazine*, vol. Sept. 2002, pp. 16–25, 2002.
- [22] J. Weickert, B.M. ter Haar Romeny, and M.A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Trans on IP*, vol. 7(3), pp. 398–410, 1998.
- [23] J.J. Koenderink, "The structure of images," *Biol. Cybern.*, vol. 50, pp. 363–370, 1984.
- [24] I. Vanhamel, I. Pratikakis, and H. Sahli, "Automatic watershed segmentation of color images," in *Mathematical morphology and its applications to image processing*, 2000, pp. 207–214.
- [25] J. Liu and Y.-H. Yang, "Multiresolution color image segmentation," *IEEE Trans on PAMI*, vol. 16(7), pp. 689–700, 1994.
- [26] M. Mirmehdi and M. Petrou, "Segmentation of color textures," *IEEE Trans on PAMI*, vol. 22(2), pp. 142–159, 2000.