MEASURING SIMILARITY BETWEEN COLOR IMAGE REGIONS

*M. Akcay, A. Baskurt, B. Sankur**
LIGIM (Computer Graphics, Image and Modeling Laboratory)
Bat. 710, 43 bd du 11 novembre 1918, 69622 Villeurbanne cedex France
*BUSIM Laboratory, Boğaziçi University, Bebek, Istanbul, Turkey
Email: makcay@ligim.univ-lyon1.fr , abaskurt@ligim.univ-lyon1.fr, sankur@boun.edu.tr

ABSTRACT

This paper presents a methodology for measuring the similarity between two color image regions. An isolated object or a segmented region is described with the Color Angular Radial Transform (CART) introduced in this study. This new region-based descriptor uses both the shape information in the luminance component and the spatial distribution of the dominant colors in the region. The similarity between two regions results from the weighted sum of the shape and color features. Results obtained on MPEG-7 data set are presented and discussed.

1 INTRODUCTION

With the increased availability of multimedia data over the Internet, the identification and retrieval of the desired audiovisual documents has become an important but challenging task. To this effect the MPEG-7 standard [1] aims to provide standardized core technologies allowing the description of the audiovisual data. The basic visual features that MPEG-7 standardizes are color, texture, shape and motion [1].

The shape descriptors are classified in three categories, contour-based, region-based and 3-D [1,2,3]. The Angular Radial Transform (ART) is one of the region-based shape descriptors. Like other well-known shape descriptors such as Zernike moments [4,5] and recent ones in [6,7], the ART Transform applied on binary objects, is also invariant on translation, rotation and scale changes.

In this paper, we propose the Color ART (CART) as an extension of ART in order to describe and retrieve similar color regions issued from any image segmentation method.

Section 2 presents the basics of ART and the similarity matching of binary objects using ART. In section 3, CART and the similarity measure to retrieve color regions are detailed. Results are presented and discussed in section 4.

2 THE ART TRANSFORM

2.1 Definition

ART (Angular Radial Transform) is a 2-D complex transform defined on a unit disk in polar coordinates [1,2] given by,

$$\begin{aligned} F_{nm} &= \langle V_{nm}(\rho,\theta), f(\rho,\theta)\rangle \\ &= \int_0^{2\pi}\int_0^1 V_{nm}^*(\rho,\theta), f(\rho,\theta)\rho d\rho d\theta \end{aligned} \tag{1}$$

where F_{nm} is the ART coefficient of order n and m, $f(\rho,\theta)$ is an image function in polar coordinates and $V_{nm}(\rho,\theta)$ is an ART basis function that is separable along the angular and radial directions, i.e.,

$$V_{nm}(\rho,\theta) = A_m(\theta)R_n(\rho)$$

The angular and radial functions are defined as follows:

$$A_m(\theta) = \frac{1}{2\pi}\exp(jm\theta)$$

$$R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2\cos(\pi n\rho) & n \neq 0 \end{cases}$$

where ***n*** and ***m*** are respectively radial and angular indices which define the order of the coefficient F_{nm} .

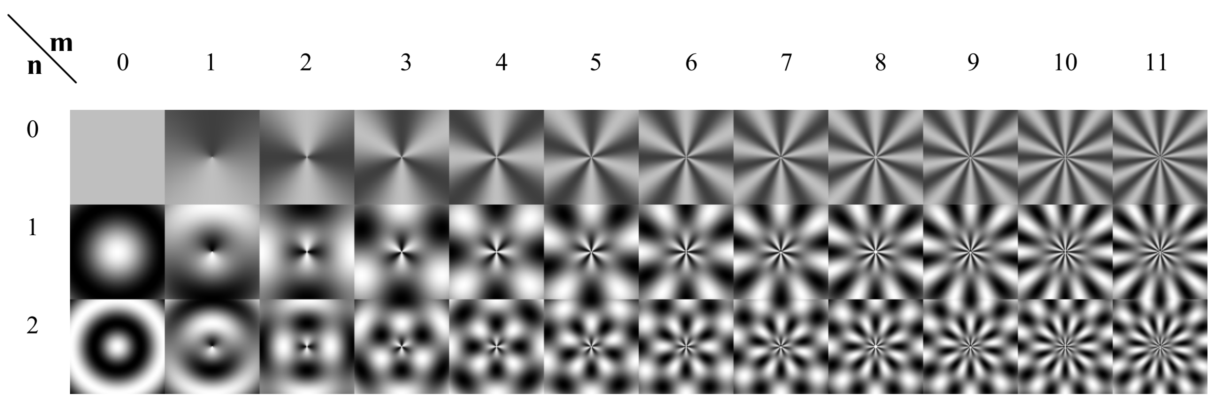


Figure 1. Real parts of ART basis functions (N=3, M=12).

ART basis functions $V_{nm}(\rho,\theta)$ are complex functions. In Figure 1, the real parts of the first 36 basis functions are shown, their imaginary parts being similar except for quadrature phase difference. While for image reconstruction one may need infinite number of ART coefficients, for description a finite, usually small, number suffices. In fact we have adopted, in what follows, the values proposed by MPEG-7 committee, that is N=3 and M=12.

2.2 Region description with ART

Figure 2 summarizes how a region descriptor can be formed with ART transform. The region-preprocessing step consists of two steps, the extraction of the object region and interpolation (scaling to standard size). Since the transform is always centered on the centroid of the object it is inherently translation-invariant. Similarly it is scale-invariant since the size of the region of interest (containing the object) is calculated and the object is scaled to a standard diameter size.

The "normalized" region is projected onto the basis functions to compute the corresponding ART coefficients. Instead of converting each image $f(x,y)$ to polar co-ordinates, it is more convenient to consider the ART basis functions $V_{nm}(\rho,\theta)$ in the Cartesian co-ordinates. Thus one has:

$$F_{nm} = \int_y \int_x V^*_{nm}(x,y) f(x,y) dxdy$$

Finally one takes the normalized modulus of the complex ART coefficients, ArtM[n,m]:

$$ArtM[n,m] = \frac{|F_{nm}|}{|F_{00}|}$$

where F_{00} is simply the number of non-zero pixels in a binary image. Obviously the magnitude of the Art coefficients, ArtM, are rotation invariant. The MPEG-7 standard provisions for the 4-bit quantization of the ArtM coefficients, the quantization table being constructed based on an exponential distribution model [1].

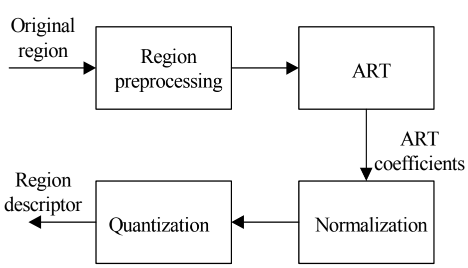


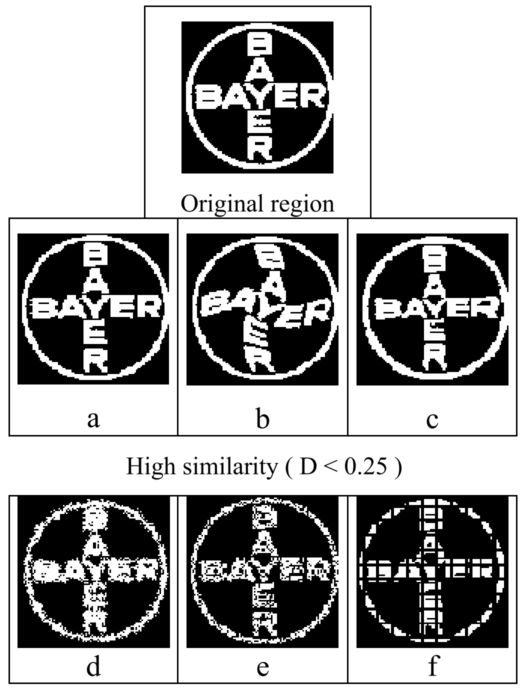
Figure 2. Extraction of region descriptors with ART.

2.3 Similarity matching with ART

The similarity measure between two regions is measured by the L1 distance between the two sets of inverse-quantized ART coefficients. Thus the distance D between two regions A and B is determined by,

$$D(A,B) = \frac{1}{35} \sum_{i=0}^{34} \left| Q^{-1}(ARTM_A[i]) - Q^{-1}(ARTM_B[i]) \right| \quad (2)$$

where D(A,B) is in range of [0;1]. Figure 3 presents an original image from the MPEG-7 database and decreasing sample scores for increasing deformations of the object. The ranges of the scores correspond to our intuitive notion of similarity category between objects, of high and intermediate similarity (first and second rows). It can be noted that the ART transform is quite robust to deformation and noise [1] .



Intermediate similarity (0.25 ≤ D < 0.35)

Figure 3. Similarity measures with ART transform between the original region and the same region under deformation and noise.

3 EXTENSION OF THE ART TO COLOR

In this section, we propose a generalization of this transform for color images.

3.1 Definition of the Color ART (CART)

The basis functions of the Color ART are the same as those of the ART transform (1). The main difference will be in the context of extraction of the object of interest to apply the transform and the calculation of the similarity distance. The color image is first represented in the perceptually uniform (L*,a*,b*) color space [8]. Each component is independently projected on the basis functions. This allows to adapt the number (N,M) of these functions to the spectral characteristics of each component. Thus the luminance component will be more precisely analyzed (N=3, M=12), while the chrominance components, which possess less spatial detail (in fact about three times lower resolution than the luminance for a mono-CCD color camera), will be projected onto (N=3, M=4) basis functions.

3.1.1 Luminance component

The luminance component is thresholded by Kittler's method [9] in order to obtain a binary image which corresponds to the global aspect of the object (shape). In Figure 4, the original image Fish and its binarized version are illustrated. The similarity distance D_L between two regions A and B binarized with the same threshold, is calculated with (2).

3.1.2 Chrominance component

For the chrominance part of the information, the CART transform measures the similarity between two color regions in terms of the spatial distribution of their dominant colors. The first step determines the first K dominant colors using MPEG-7 dominant color descriptor [1]. In Figure 4, the spatial distributions of three dominant colors of the original image Fish are shown. In each binary image, the white pixels indicate the presence of the associated dominant color. The dominant colors are reduced to binary images, which are projected independently to (N=3, M=4) ART basis functions. Three similarity distances, D_{C_1} D_{C_2}, D_{C_3} are then computed with equation (2).

3.1.3 Similarity matching with CART

The similarity distance between two color image regions A and B is calculated as a weighted sum of the distances between the luminance L and the chrominance descriptors $C_i, i = 1,2,...K$. While the distance D_L measures the similarity of shape, the distance D_C evaluates the matching of the spatial distribution of a specific dominant color in regions A and B. The global distance is given by

$$D_{CART}(A,B) = \alpha D(L(A),L(B)) + \sum_{i=1}^{K} \beta_i D(C_i(A),C_i(B)) \quad (3)$$

where $C_i(A)$ is the binary image associated to the i[th] dominant color of A and $\alpha + \sum_{i=1}^{K} \beta_i = 1$. In the following results, we take $\sum_{i=1}^{K} \beta_i = 0,5$ in order to attach the same weight to the luminance and the chrominance parts, while the K weights β_i are made proportional to the cardinality of the dominant color i of the reference image. Finally if the regions A and B do not have the same dominant color, then $D(C(A),C(B))$ is equal to 1 (the maximum distance). In this study we have considered only the first dominant color, hence K = 1. Note that the coefficients α and β can be changed according to the importance we want to give to the shape and to the dominant color. Obviously the choice of $(\alpha,\beta) = (1,0)$ corresponds to the ART case, while with $(\alpha,\beta) = (0,1)$ one obtains the pure color similarity matching.

4 RESULTS AND DISCUSSION

In Figure 5, an example of similarity matching is presented with decreasing order from left to right for each row, with $D_{CART}(A,B) = 0,5 D_L + 0,5 D_C$, using the first dominant color. These results are plausible in that the scores correspond to our subjective categorization of high, mediocre and low similarity. The image in Figure 5c is classified in the intermediate similarity class because of its heterogeneous color texture and its horizontally deformed shape. Furthermore the images in Figure 5d-e are also in this class only due their dominant color, where both the reference and test images have similar distributions. To demonstrate the contribution of the color component, let's consider in the above example only the luminance component, hence we take $(\alpha,\beta) = (1,0)$. In this case the intuitively more similar objects in the middle row (Figs. 5d and e) become rejected, as their luminance similarity is not adequate. This is an example where CART, which takes into consideration colour information, outperforms ART.

5 CONCLUSION

The Color Angular Radial Transform (CART) is presented. This transform combines the information of shape issued from the luminance component and the information of the dominant colors and their spatial distribution. CART transform enhances the capabilities of the ART transform by considering the shape information inherent in the color. It can be observed that this region-based approach can describe and classify color regions, even in the presence of such disturbances as texture or minor holes.

Original image | Binarized luminance Component L | First dominant color C_1 | Second dominant color C_2 | Third dominant color C_3

Figure 4. Binarized luminance component and binarized dominant color components of the image Fish.

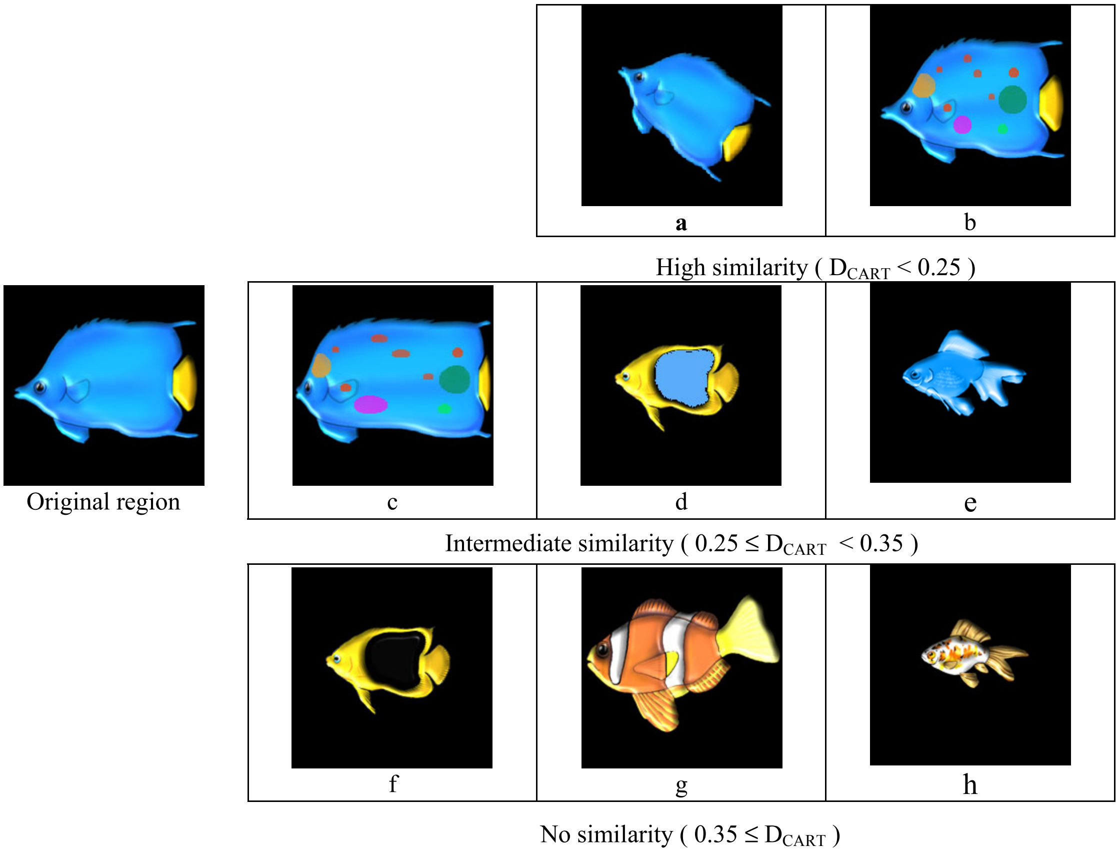


Figure 5. Similarity measures with CART transform.

The use of CART on a large database is under investigation. We work also on how to treat the cases where a region is composed with equal cardinality in their dominant colors. Another aspect to explore is the similarity of these dominant colors in terms of neighborhood in the chrominance histogram in order to compare two regions which have similar (and not only identical) dominant colors by combining CART measure with a color distribution metric.

6 REFERENCES

[1] S. Jeannin , "MPEG-7 Visual part of eXperimentation Model Version 9.0", *ISO/IEC JTC1/SC29/WG11 /N3914*, Mpeg Meeting, Pisa, Italy, Jan. 2001.

[2] S. Jeannin, M. Bober, "Description of core experiments for MPEG-7 motion/shape", *ISO/IEC JTC1/SC29/WG11/N2690*, Seoul, March 1999.

[3] M. Bober, J.-D. Kim, Y.-S. Kim, W.-Y. Kim, K. Muller, "Summary of the results in shape descriptor core experiment", *ISO/IEC JTC1/SC29/WG11/M4869*, Mpeg Meeting, Vancouver, Canada, July 1999.

[4] A. Khotanzad, Y. H. Hong, "Invariant image recognition by Zernike moments", *IEEE Trans. On PAMI,* 12 (5), pp.489-498, May 1990.

[5] R. J. Prokop, A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition", *Graphical Models Image Processing*. 54 (5), pp. 438-513, September 1992.

[6] H.-K. Kim, J.-D. Kim, "Region-based shape descriptor invariant to rotation, scale and translation*", Signal Processing, Image Communication*, 16, pp. 87-93, 2000.

[7] E. Paquet et al., "Description of shape information for 2-D and 3-D objects", *Signal Processing, Image Communication,* 16, pp. 103-122, 2000.

[8] K. N. Plataniotis, A. N. Venetsanopoulos, "Color Image Processing and Applications", *Springer-Verlag Berlin Heidelberg*, Germany, 2000.

[9] J. Kittler, J. Illingworth, "Minimum Error Thresholding", Pattern Recognition, 19, pp. 41-47, 1986.