U.P.B. Sci. Bull., Series C, Vol. 76, Iss. 1, 2014 ISSN 2286 – 3540

A GRANULOMETRIC COLOR STRUCTURE DESCRIPTOR FOR CONTENT-BASED IMAGE RETRIEVAL

Constantin VERTAN[1], Răzvan George CONDOROVICI[2], Laura FLOREA[3], Bogdan IONESCU[4]

We show that the MPEG-7 color structure descriptor can be particularized down to the classical color histogram. The indexing of the structuring element used in the computation of the CSD is inducing a pseudo-granulometry over the uniform colored parts of the image. We propose the extension of the CSD via the symmetrization of the morphological granulometry, yielding a two-dimensional structure that characterizes both the colors and the color layout. This approach is useful in content-based image retrieval and in color texture recognition.

Keywords: Color Image Analysis, Image description, Content-based Image Retrieval, MPEG-7.

1. Introduction

Content-based image retrieval (CBIR) became in the last two decades a very popular research direction, supporting real-life applications such as digital libraries, image archives, video-on-demand, and specific image database management.

Image classification and image retrieval are tightly related fields, by requiring a relevant image description at a higher level of interpretation. In the case of classification, the image descriptor is used to label the image as belonging to a given class within a pre-defined class set. In the context of retrieval, the descriptor is used to search in a database for images that are similar to the given image. Image description involves extracting relevant features from images within the database, such that a mathematically computed distance between feature vectors corresponds to the visual distance (or visual similarity) between the images. The main goal of an image description scheme is to provide a compact feature vector, embedding most of the visual cues characterizing the image: color,

[1] The Image Processing and Analysis Laboratory, LAPI, University POLITEHNICA of Bucharesti, Romania, e-mail: cvertan@alpha.imag.pub.ro

[2] The Image Processing and Analysis Laboratory, LAPI, University POLITEHNICA of Bucharesti, Romania

[3] The Image Processing and Analysis Laboratory, LAPI, University POLITEHNICA of Bucharesti, Romania

[4] The Image Processing and Analysis Laboratory, LAPI, University POLITEHNICA of Bucharesti, Romania

texture, structure. The most successful descriptors take into account both the color balance and the spatial color distribution, spanning, if possible, several spatial resolution levels.

2. The MPEG-7 image description framework

The MPEG-7 standard [3], [5] is an international standard since September 2001 and specifies metadata for describing multimedia content. The visual content is described by several groups of descriptors, dealing with color, shape, motion, texture and localization. Within the group of color descriptors, which provide the simples approach for dealing with the overall visual appearance of a color image, the standard proposes several descriptors, namely the Color space, Color Quantization, Dominant Colors, Scalable Color, Color Layout/Structure, and Group of Frames/Pictures Color [2]. Color is one of the most important and studied features in describing images. There are three main types of color descriptors: histogram descriptors, layout descriptors and dominant color descriptors. The Color Histogram Descriptor defines a uniform quantization of a color space, selected from RGB, HSV, YCbCr or the newly-defined HMMD. The number of bins used for quantization can be selected from a prescribed set, such that the descriptor can be adapted to a wide range of applications.

The spatial distribution of color constituted an effective descriptor for sketch-based image retrieval. The Color Layout Descriptor is designed to capture the spatial distribution of colors within an image. The CLD is a compact descriptor that uses representative 8 × 8 colors (YCbCr color space) on a grid followed by a DCT and encoding of the resulting coefficients.

The MPEG-7 color structure descriptor (CSD) [2] expresses local color structure in an image using an 8 × 8 structuring element. CSD counts the number of times a particular pixel value (color or gray level) is contained within the structuring element as the structuring element scans the image. If the image has M values (colors, gray levels, etc.) denoted by $g_1, g_2, ..., g_M$, then a color structure histogram is defined as $h(m)$, $m = 1,2,...M$, where the value in each bin represents the number of structuring elements in the image containing at least one pixel with value g_m.

3. Linking the CSD to the histogram

As it was defined previously, the CSD accounts the number of structuring element positions within the image that contain at least one pixel of the given color (or, in general, value). This means that for any specific pixel location within the image (traversed at some moment by the structuring element used in the CSD) one adds a 1 in the corresponding, current CSD bin if and only if there is at least one pixel with the corresponding color within the structuring element. This understanding of the underlying processing is identical to the one associated with the basic, binary dilation operation from the mathematical morphology.

The morphological dilation [4] of a set *A* via the structuring element *B* is defined as the set of points to which one can translate the structuring element such that it still has a non-empty intersection with the processed set *A*:

$$A \oplus B = \{x \mid B_x \cap A \neq \varnothing\} \quad (1)$$

In order to obtain a compact definition of the CSD, we introduce some notations. Let *I* be the color image to be described; after the quantization of the colors we will refer the binary image containing only color *k* by I_k. If *B* is the structuring element used for the computation of the CSD (it should be noted that the term structuring element may have been especially chosen in MPEG-7) and *Hist* is the counting operator, then we can define the CSD as follows:

$$CSD(k) = Hist(I_k \oplus B)(k) \quad (2)$$

Since this processing is performed within a binary framework, we can also rewrite (2) as:

$$CSD(k) = Area(I_k \oplus B)(k) \quad (3)$$

From equation (2) above is now clear that the usual image histogram is a particular case of the CSD, when the structuring element is reduced to a single pixel. This observation entitles us to claim that we may construct an entire set of CSD's of a given image, by indexing the size of the structuring element used in the computation. This behavior (emphasized by the (3) form) resembles closely to the granulometric curve of a shape.

4. The morphological granulometry

The granulometric curves associated to a plane shape are real functions, based on the measuring (via the area/number of pixels operator) or the sets resulted from applying some morphological operators to the shape to be characterized. The real-valued variable indexing the granulometric curve is linked to the size of the structuring element used by the morphological operation. The obtained descriptor is invariant with respect to the translation and the rotation of the measured set.

The granulometry associated to any plane shape *A* is a sequence of sets $(\Phi_A(\lambda))_{\lambda \in \Re^+}$, where $\Phi_A(\lambda)$ is the result of processing the shape *A* by a morphological operation with structuring element of size λ [4]. The morphological transform must be chosen such that the following conditions hold: the transform is increasing with respect to the processed set (which means that, for a fixed $\lambda \in \Re^+$ and two planar shapes A_1, A_2 if $A_1 \subseteq A_2$ then $\Phi_{A_1}(\lambda) \subseteq \Phi_{A_2}(\lambda)$), anti-extensive (which means that for a fixed parameter $\lambda \in \Re^{+}$, $\Phi_A(\lambda) \subseteq A$) and verifies $\forall \lambda, \mu \in \Re^{+}, \Phi_{\Phi_A(\lambda)}(\mu) = \Phi_{\Phi_A(\mu)}(\lambda) = \Phi_A(\max(\lambda, \mu))$. This last condition implies the idempotency of the morphological transform (putting $\lambda = \mu$ in the condition above). Usually, the structuring element used is a centered disk.

The simplest morphological transform respecting the conditions above is the opening (O); the granulometry will be then defined as:

$$\Phi_A(\lambda) = A \mathrm{O} \lambda B = (A \Theta \lambda B) \oplus \lambda B^S) . \tag{4}$$

It should be noted that in many practical approaches the mathematical constraints defining the granulometry are relaxed, accepting thus simpler morphological operations as $\Phi_A(\lambda)$ transforms, the morphological erosion (5) being a preferred choice.

$$A \Theta B = \{x \mid B_x \in A\} . \tag{5}$$

The areas of the set resulted by the morphological processing are usually normalized to the area of the convex hull of the set *A* to be processed. In the following of this contribution we will not use this normalization, such that the simplest granulometry of the binary set *A* is

$$\Omega_A(\lambda) = Area(\Phi_A(\lambda)) = Area(A \Theta \lambda B) . \tag{6}$$

As defined in equation (6) above, the granulometric curves measures the behavior of the binary set *A* under successive reductions via erosion, that is, the curves measure how the set looses its points. In order to measure the symmetric effect of adding points to the set until its “saturation”, the granulometry can be extended by the use of definitions involving the dual morphological transform of $\Phi_A(\lambda)$, that is $(\Phi_{A^C}(\lambda))^C$. Conventionally, this is introduced for negative indexes $\lambda < 0$, such that the resulting granulometry (without convex hull normalization) is a function defined as $\Omega_A : \Re \to \Re^{+}$, with:

$$\Omega_A(\lambda) = \begin{cases} Area(A \oplus (-\lambda) B, \lambda < 0; \\ Area(A), \lambda = 0; \\ Area(A \Theta \lambda B), \lambda > 0. \end{cases} . \tag{7}$$

We can now see that the CSD value computed for one color from the image, as defined in section **2**, is a single value taken from one half of the granulometric curve, namely from the symmetrized curve obtained by dilation with $\lambda < 0$ in (7). The histogram value corresponding to the same color is a particular case of the CSD (that can be obtained for a structuring element that contains only its origin) and is the value of the granulometric curve at $\lambda = 0$. Finally, the classical granulometric curve (that is, the values obtained via erosion with $\lambda > 0$) is not exploited under the current definition of the CSD. As such, we propose a new, dual CSD, defined as:

$$CSD(k) = Area(I_k \Theta B)(k) \tag{8}$$

Fig. 1 shows a schematic representation of the structure of the proposed dual granulometric extended color structure descriptor, depicting its particularization for various parameters that modulate the size of the structuring element used for the processing.

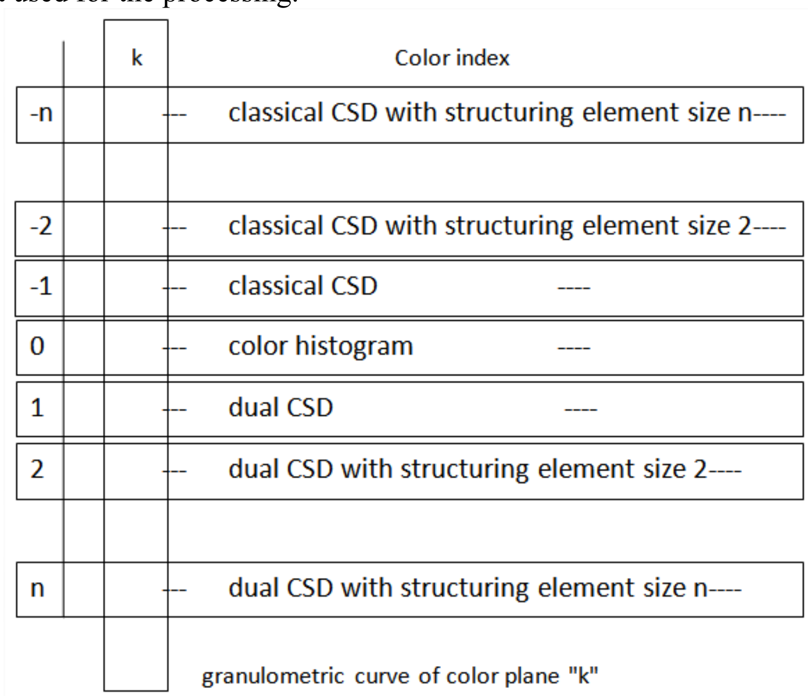

Fig. 1 Scheme of the dual granulometric CSD structure for various sizes of the used structuring element.

5. Applications

Although the main scope of this contribution is theoretical, showing the links between the MPEG-7 color structure descriptor, the color image histogram and the morphological granulometric curves, we can still show that the use of the granulometric CSD (that is a sequence of CSD computed for both various structuring elements and dual morphological operators, as depicted in the scheme in Fig. 1) can outperform by various margins the classical approaches.

We use two texture databases. A first generalist color texture database (*Textures*), consisted, in our experiments of 100 classes of nine 128 by 128 color images of various natural and artificial, regular and irregular textures, for a total of 900 images (part of this database being taken from the well-known MIT Vistex texture database, available at http://www.media.mit.edu/vismod). A second database is the classical Brodatz [1] graylevel texture database, consisting of 112 texture classes, for a total of 1120 sampled texture images.

The performance of the descriptors was tested for both the texture recognition and the texture image retrieval. The image retrieval performance is classically measured by the precision-recall curves [6]. The precision-recall curve plots the precision for all the recall rates that can be obtained according to the current image class population (C=9 for the *Textures* image database) from 1/C to 1, in steps of 1/C, using, in turn, all images in the database as queries. As shown in the plots from Fig. 2, the proposed description scheme provides superior indexing performance. The texture recognition performance is measured by the average recognition ratio (for all images within the database) according to a 1-, 3-, 5- and 7- nearest neighbor (NN) techniques. The retrieval performance is measured by the classical precision-recall curves.

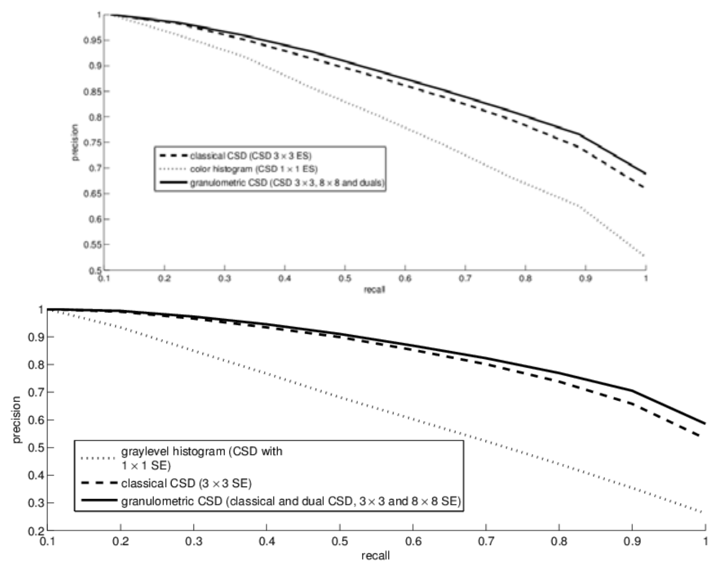

Fig. 2 a) Precision-recall curves for the *Vistex* image database: dotted line – color histogram (CSD 1 × 1 ES) dashed line – classical CSD (CSD 3 × 3 ES), continuous line – granulometric CSD (CSD 3 × 3, 8× 8 and duals). b) Precision-recall curves for the *Brodatz* image database: dotted line – grayscale histogram (CSD 1 × 1 ES) dashed line – classical CSD (CSD 3 × 3 ES), continuous line – granulometric CSD (CSD 3 × 3, 8 × 8 and duals).

Tables 1 and 2 show the average recognition rates and the average precision of recognizing one texture image versus all the remaining textures in the database for its correct class. Typical cases have been selected for the comparison; the results show a slight performance improvement (measured especially by the average precision) in the use of multiple CSDs (obtained via several structuring elements) with or without the use of the proposed dual CSD.

Table 1

Recognition rates and average precision for the generalist Vistex texture image database based on a k-Nearest Neighbor algorithm (with k=1,3,5,7). All the descriptors are computed for a fixed 6 bin/color channel quantization.

Texture parameters	Recognition rate [%]				Average precision [%]
	1-NN	3-NN	5-NN	7-NN	
RGB histogram (CSD 1×1 SE)	94.0	90.2	83.3	77.8	86.33
CSD 3×3 SE	97.4	94.3	90.2	86.6	92.13
CSD 8×8 SE	97.3	94.9	91.8	87.7	92.91
CSD 3×3 SE and CSD 8×8 SE	97.8	94.8	91.6	87.9	93.00
CSD 3×3 SE and dual CSD 3×3 SE	97.4	94.0	89.8	85.8	91.74
CSD 8×8 SE and dual CSD 8×8 SE	97.3	95.0	91.8	87.4	92.88
CSD and dual CSD 3×3 and 8×8 SE	97.8	94.8	91.7	88.0	93.05

Table 2

Recognition rates and average precision for the grayscale Brodatz texture image database based on a k-Nearest Neighbor algorithm (with k=1,3,5,7). All the descriptors are computed for a fixed 6 bin/color channel quantization.

Texture parameters	Recognition rate [%]				Average precision [%]
	1-NN	3-NN	5-NN	7-NN	
RGB histogram (CSD 1×1 SE)	90.9	81.7	73.7	64.6	77.73
CSD 3×3 SE	98.5	95.5	91.9	88.1	93.49
CSD 8×8 SE	98.4	95.1	91.4	87.4	93.06
CSD 3×3 SE and CSD 8×8 SE	98.8	96.7	93.2	89.1	94.44
CSD 3×3 SE and dual CSD 3×3 SE	98.6	95.8	92.3	88.3	93.72
CSD 8×8 SE and dual CSD 8×8 SE	98.4	95.1	91.4	87.4	93.06
CSD and dual CSD 3×3 and 8×8 SE	98.8	96.7	93.2	89.2	94.48

6. Conclusions

This paper presents a theoretical reformulation of the definition of the classical MPEG-7 Color Structure Descriptor. It has been shown that accounting for the color structure within the image is similar to the construction of the granulometry of the corresponding binary color plane. Several, different CSD values can be obtained for each color by modifying on regular basis the structuring element used for the processing of the CSD. The classical color histogram can be obtained as a particular case, if the structuring element contains only its origin. Further, the CSD can be extended via the dual morphological operation used in its computation.

The new descriptor, that can be named granulometric CSD, is a two-dimensional structure; its size is controlled by the number of different structuring elements used in for the computation of the CSD. It should be noted that there is no need for the computation of the full granulometry.

As simple tests show, the new descriptor can provide better performance without a significant increase in its size. The descriptor can be further developed according to two new directions: first, one can use the morphological operations that allow the construction of a hull, mathematically-correct granulometry in place of the simple, current erosion/dilation. Then, following the idea of "Generalized Structure Descriptor", new, local features can be taken account instead (or in complement) of the pixel color.

Acknowledgement

This work was supported by the Romanian Sectoral Operational Programme Human Resources Development 2007-2013 through the Financial Agreements POSDRU /89/1.5/S/64109, POSDRU /107/1.5/S/76903 and POSDRU /89/1.5.S/62557.

R E F E R E N C E S

[1]. *P. Brodatz*, Textures: A Photographic Album for Artists and Designers, (1966), Dover Publications, New York.
[2]. *B.S. Manjunath, J.R. Ohm, V.V. Vasudevan, A. Yamada,* The MPEG-7 visual standard for content description – an overview, IEEE Trans. on Circuits and Systems for Video Technology, **11**, (2001), No. 6, 703-715
[3]. *ISO/IEC 15938-1,* Multimedia content description interface part I: Systems, (2002).
[4]. *J. Serra,* Image Analysis and Mathematical Morphology, (1982), Academic Press, London.
[5]. *T. Sikora,* The MPEG-7 visual standard for content description – an overview, IEEE Trans. on Circuits and Systems for Video Technology, **11**, (2001), No. 6, 696-702
[6]. *A.W. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain*, Content-based image retrieval at the end of the early years, IEEE Trans. on Pattern Analysis and Machine Intelligence, **22**, (2000), No. 12, 1349-1380.