

# REGION-BASED RETRIEVAL: COARSE SEGMENTATION WITH FINE COLOR SIGNATURE

Julien Fauqueur and Nozha Boujema

{Julien.Fauqueur, Nozha.Boujema}@inria.fr

INRIA, Imedia Research Group, BP 105, F-78153 Le Chesnay, France

## ABSTRACT

The two major problems raised by a region-based image retrieval system are the automatic definition and description of regions.

In this paper we first present a technique of unsupervised coarse detection of regions which improves their visual specificity. The segmentation scheme is based on the classification of Local Distributions of Quantized Colors (LDQC). The Competitive Agglomeration (CA) classification algorithm is used which has the advantage to automatically determine the optimal number of classes.

Then, considering that region description which must be finer for regions than for images, we propose a region descriptor of fine color variability: the Adaptive Distribution of Color Shades (ADCS). Compared to existing color descriptors, the high color resolution of ADCS improves the perceptual similarity of retrieved regions.

## 1. INTRODUCTION

The primary functionality of a Content-Based Image Retrieval system is the global query-by-example approach, in which visual features are extracted from the entire image. But in many cases the user's goal is to retrieve similar regions rather than similar images as a whole. In a generic image database the search for similar regions using global features over images can be highly biased by the surrounding regions and background.

Existing region-based query systems differ by the definition and description of regions. Automatic region determination methods can be performed on-line using features back projection (see [1]). Off-line methods include systematic image subdivision into squares or automatic image segmentation. This latter method is proposed in a couple of systems such as Blobworld [2] and Netra [3]. Concerning description, existing systems represent color as distributions computed over predefined subsamplings of color spaces giving about 200 colors (166 to 256 colors in [1], [2], [3]).

Designing a region-based query system remains a challenging and open problem: automatic detection of regions

of interest is a hard task and existing region descriptors are derived from image descriptors without considering the fact that regions are more homogeneous with less statistics. Our approach differs by how we extract and describe regions. Regions should integrate more intrinsic variability to be visually more characteristic. Besides, region description should not depend on a predefined set of 200 colors, but should rather be adaptively determined. The key idea of coarse region detection and fine signature: the relatively high visual variability inside regions is accurately described by the high resolution of color shades, such that regions are really specific against each other in the database.

In the next section, we'll outline CA algorithm, an essential background technique in our work. Region extraction will be developed in section 3. Region indexing and matching are explained in section 4. Then tests and results will be presented and discussed in section 5. We conclude in section 6.

## 2. BACKGROUND: CA CLASSIFICATION ALGORITHM

Competitive Agglomeration classification, originally presented in [4], has the major advantage to determine the optimal number of clusters.

Using notations from [4], we call  $\{x_j, \forall j \in \{1, \dots, N\}\}$  the set of  $N$  data we want to clusterize and  $C$  the number of clusters.  $\{\beta_i, \forall i \in \{1, \dots, C\}\}$  denote the prototypes to be determined. The distance between data  $x_j$  and prototype  $\beta_i$  is  $d(x_j, \beta_i)$ . Then CA-classification is performed by minimizing the following quantity  $J$ :

$$J = J_1 + \alpha J_2, \text{ where } J_1 = \sum_{i=1}^C \sum_{j=1}^N u_{ij}^2 d^2(x_j, \beta_i) \quad (1)$$
$$\text{and } J_2 = - \sum_{i=1}^C [\sum_{j=1}^N u_{ij}]^2$$

Subject to membership constraint:  $\sum_{i=1}^C u_{ij} = 1, \forall j \in \{1, \dots, N\}$  where  $u_{ij}$  represents the membership degree of feature point  $x_j$  to prototype  $\beta_i$ . Minimizing  $J_1$  separately

is equivalent to perform an FCM classification which determines  $C$  optimal prototypes and the fuzzy partition  $U$  given  $x_j$  and  $C$  using distance  $d$ .

Therefore  $J$  is written as a combination of two opposite effect terms ( $J_1$  and  $J_2$ ). So minimizing  $J$  with an over-specified number of initial clusters classifies data and simultaneously optimizes the number of classes.

$J$  is minimized recursively.  $\alpha$  is the competition weight and should allow a balance between terms  $J_1$  and  $J_2$  in (1). At iteration  $k$ , weight  $\alpha$  is expressed as :

$$\alpha(k) = \eta_0 \exp\left(\frac{-k}{\tau}\right) \frac{\sum_{i=1}^C \sum_{j=1}^N u_{ij}^2 d^2(x_j, \beta_i)}{\sum_{i=1}^C [\sum_{j=1}^N u_{ij}]^2} \quad (2)$$

As iterations go,  $\alpha$  decreases so emphasis is first given to agglomeration process, then to classification optimization.  $\alpha$  is fully determined by parameters  $\eta_0$  and  $\tau$ .

During the algorithm spurious clusters are discarded. Convergence is decided when prototypes are stable. The classification granularity is controlled by factor  $\alpha$  through its magnitude  $\eta_0$  and its decline strength with  $\tau$ . The higher  $\eta_0$  and  $\tau$ , the higher  $\alpha$ , so the more classes are merged. So for a given classification granularity, CA determines the optimal number of classes.

CA will be used at three steps in our work with different levels of granularity: first to perform image quantization, then to segment roughly the image by computing LDQC prototypes and then to finely describe regions with color shades.

### 3. COARSE SEGMENTATION

Detected regions should encompass a certain visual diversity to be visually characteristic, using a coarse segmentation. We want to stay beyond a too fine level of spatial and feature details.

To group pixels to form such regions, we perform a CA-classification of local color distributions of the image. This feature naturally integrates the diversity of colors in pixels neighbourhood. The choice of the color set to compute distributions is crucial: it must be compact to gain speed in classification and representative of a small pixel neighbourhood. Classic color histogram, computed on a uniform subsampling of a color space, are too big and imprecise. So we define the color set as the adaptive set obtained by image color quantization. Then for all neighborhoods in the image, local distributions are evaluated on this small and relevant color set. We'll refer to them as Local Distributions of Quantized Colors (LDQC's).

After classification, LDQC prototypes are back projected onto image, then small regions are either merged or discarded.

### 3.1. Image color quantization

Image is quantized in the Luv color space by CA-classification of color pixels using the Euclidean distance. The classification granularity was chosen such that big areas in images with a strong texture are represented by more than one color. At classification convergence the color prototypes define a set of  $nqc$  quantized colors. Since CA determines automatically the right number of clusters, the number of quantized colors  $nqc$  will be representative of the color diversity of the natural images.

### 3.2. Determination of LDQC prototypes

Regions are considered as collections of adjacent homogeneous LDQC's. To compute all the LDQC's in the image, we slide a window over pixels and evaluate the corresponding local distribution over the quantized color set. Let's denote  $S_{TOT}$  the image surface. LDQC's are evaluated every  $wr$  pixels, where  $wr$  is the window radius.

An appropriate distribution measure should be used for the classification.  $L^p$  distances are widely used for global color distributions on uniformly quantized color spaces but are inaccurate for small adaptive color sets.

The color quadratic form distance [5] provides a precise distance to compare any kind of color distributions by integrating the inter-bin color similarity. Its expression is given for two distributions  $\{x_i\}$  and  $\{y_i\}$  evaluated on a set of  $nqc$  colors. Then the quadratic distance is expressed as:

$$d_q(x, y)^2 = (x - y)^T A (x - y) \\ = \sum_{i=1}^{nqc} \sum_{j=1}^{nqc} (x_i - y_i)(x_j - y_j) a_{ij} \quad (3)$$

where  $A = [a_{ij}]$  gives the similarity between colors  $i$  and  $j$ :  $a_{ij} = (1 - d_{ij}/d_{max})$ , where  $d_{ij}$  is the Euclidean distance in Luv space. The color quadratic distance is used to classify the LDQC histograms with CA.

After classification, the segmented image is obtained by assigning to the  $S_{TOT}/wr$  pixels the label of the LDQC prototype minimizing the quadratic distance to the LDQC around that pixel. A maximum vote filter is applied to discard isolated pixels.

### 3.3. Adjacency information

The segmented image gives us a complete partition of the image into adjacent regions. Some regions may be too small to constitute regions of interest, so they increase needlessly the total number of regions in the database. Besides, in complex scenes, they're often located at the frontier between two big regions or inside a big region. They should be merged to improve the topology of regions of interest.

Region attributes (surface, color distribution) and region adjacency information are stored in a Region Adjacency Graph structure used to merge regions.

We want final regions of interest to cover a minimum of 1.5% of the image surface. Below this threshold a region is merged to its closest visual neighbour if it has one and is discarded otherwise. Two small regions are said to be visually close if they have close mean quantized color distributions. Small regions are discarded and not indexed.

## 4. REGION INDEXING AND RETRIEVAL

### 4.1. ADCS, a fine color variability region signature

Once regions are detected in a coarse way we have to finely describe their visual appearance. Existing visual descriptors for regions (and for images too) do not exploit the full color space: they are generally histograms evaluated on an average of 200 color bins obtained by a subsampling of the color space: uniform subsampling (e.g. in [1] and [2]) or database-dependent subsampling in [3]. Such a description forces the minimum distance between two colors to be high because the subsampling is fixed and because we only consider a few hundreds of colors among millions in a full color space. This low granularity of color description doesn't matter for complex images as they contain a wide range of different colors. But regions are by definition more homogeneous and require a higher color resolution to be compared.

We propose to index each region with its Adaptive Distribution of Color Shades (ADCS): for *each region* we classify its pixels with the CA algorithm with a high classification granularity (i.e. low agglomeration). Obtained color prototypes provide the set of color shades. The descriptor index consists of the color distribution evaluated over these colors shades: it represents the region color variability and provides an approximation of the region color texture appearance. The higher the region color diversity is, the more color shades are found by CA. The nature and the number of color shades are not fixed but specific to each region, unlike with existing descriptors. Besides color shades are picked from the *whole* Luv color space which contains 5.6 million potential colors, while existing color descriptors distinguish no more than 200 colors (on average).

### 4.2. Matching regions

For a given query region of ADCS distribution  $X$ , similar regions are such that their distribution  $Y$  minimizes  $d_q(X, Y)$ . Let's write  $X$  and  $Y$  as pairs of color/percentage:

$$X = \{(c_1^X, p_1^X), \dots, (c_{n_X}^X, p_{n_X}^X)\} \text{ and}$$

$$Y = \{(c_1^Y, p_1^Y), \dots, (c_{n_Y}^Y, p_{n_Y}^Y)\}$$

The formula (3) gives the quadratic distance between two color distributions  $x$  and  $y$  evaluated on the *same* color

set. Since two regions don't have the same set of color shades, we can rewrite the expression of the quadratic distance to discard the distributions binwise differences. We consider  $x$  as the extension of distribution  $X$  over the *entire* color space and  $y$  the extension of  $Y$ . The extension consists in setting bin values to zero for colors which are not color shades, so we have  $d_q(x, y) = d_q(X, Y)$ . By using the symmetry of  $A$  and developing expression (3), we easily obtain the following expression:

$$\begin{aligned} d_q(X, Y)^2 &= \sum_{i,j=1}^{n_X} p_i^X p_j^X a_{c_i^X c_j^X} + \sum_{i,j=1}^{n_Y} p_i^Y p_j^Y a_{c_i^Y c_j^Y} \\ &\quad - 2 \sum_{i=1}^{n_X} \sum_{j=1}^{n_Y} p_i^X p_j^Y a_{c_i^X c_j^Y} \end{aligned} \quad (4)$$

Returned regions are ranked and showed by growing quadratic distance  $d_q$ .

## 5. RESULTS

Our system was tested with a 498MHz PC on IDS database provided by courtesy of *Images Du Sud Photo Stock*. It contains 2,483 generic images of flowers, portraits, landscapes, architecture, people, fruit, gardens. Images size are about 500x500 pixels.

### 5.1. Region detection

A few segmented images are presented in figure 1. More examples can be seen at:

<http://www-rocq.inria.fr/~fauqueur/ADCS/>

Even in complex natural scenes extracted regions present a coherent visual appearance. The coarse segmentation proves its ability to integrate within regions areas formed with many shades of the same hue, strong textures, isolated spatial details, which make their specificity. 15,248 regions were automatically extracted from the 2,483 images (average of 6 regions per image). Segmenting an image took an average of 5.6s. Discarded regions (shown as small grey regions in examples) represent a very small percentage of image surfaces.

### 5.2. Region segmentation

In figure 1, the third lavender and third person images show the color shades used to describe each corresponding region. The global appearance of these quantized images shows the precision of the ADCS color variability description. A total of 261,219 colors shades from the Luv space were automatically determined to index the 15,248 regions of the database (average of 17 colors per region). 168,912 of these colors were unique (to be compared with the 200 fixed bins of a classic histogram).

Extracting ADCS index from a region took about 0.5s. A region index takes an average of 69 bytes, which makes it three times more compact than a classic color histogram.



**Fig. 1.** Illustration of coarse segmentation and fine description: original image, image of detected regions with mean color, and image of regions with color shades used for indexing. Small discarded regions are shown in grey.

### 5.3. Retrieval

As a color variability descriptor, ADCS was compared to the commonly used classic color histogram (with 216 bins in the Luv space). Retrieval performance was tested on 2 classes of regions: people skin and lavender regions for which the relevance of retrieved regions could easily be decided in IDS. Queries of 26 different lavender regions were performed using ADCS and the classic histogram. For each query, the top 30 retrieved regions were “manually” tagged as relevant/irrelevant to determine the precision (fraction of relevant images retrieved at each rank). The same was done for 26 queries of skin regions. ADCS improved retrieval precision by 25% for lavender regions (see figure 3) and 13% for skin regions (figure 2).

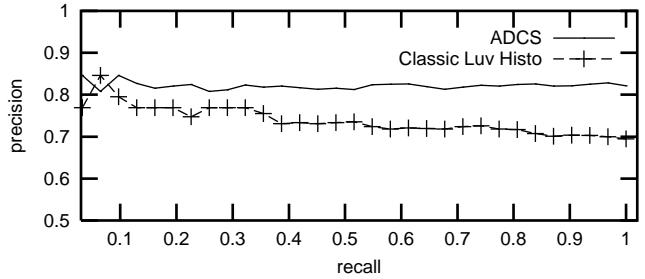
Image false positives with the classic histogram were due to the similarity inaccuracy. It could not distinguish two perceptually different regions corresponding to different objects.

The gain in retrieval precision with ADCS relatively to existing color descriptors is coherent with the gain in description accuracy of the new description and matching scheme. Retrieved regions are perceptually more similar and more consistent with the query and give an impression of perceptual continuum as ranks decrease. Regions described by many color shades returned regions with many color shades and conversely for single-colored regions. We observed that the number of color shades is also an exploited information about the color diversity of a region.

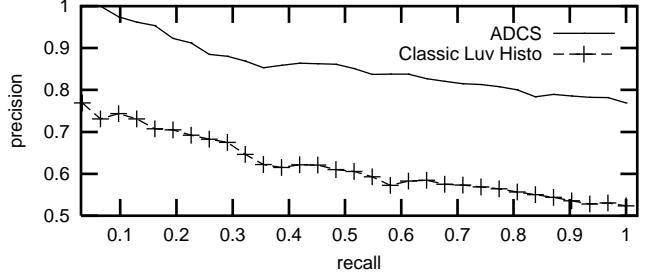
Region queries are done by exhaustive search against the 15,248 regions. Average query time is 1.3s.

## 6. CONCLUSIONS

The key idea of this paper is to detect visually specific regions of interest and match them with the fine signature to improve the retrieval results.



**Fig. 2.** ADCS index vs. classic histogram on IDS Database: Retrievals of skin regions



**Fig. 3.** ADCS index vs. classic histogram on IDS Database: Retrievals of lavender regions

We presented a novel scheme for coarse automatic image segmentation and fine region description to perform region-based queries in a generic image database. The new ADCS signature provides a representation of region color variability with more accuracy than existing descriptors.

On the description aspect, we focused here on color, but the ADCS descriptor should be combined with structural and geometrical features to provide the best retrieval performance. We also plan to investigate methods to speed up the region query process.

## 7. REFERENCES

- [1] S.F. Chang J.R. Smith, “Visualseek: A fully automated content-based image query system,” in *ACM*, 1996.
- [2] C. Carson and al., “Blobworld: A system for region-based image indexing and retrieval,” 1999.
- [3] Wei-Ying Ma and B. S. Manjunath, “Netra: A toolbox for navigating large image databases,” *Multimedia Systems*, vol. 7, no. 3, pp. 184–198, 1999.
- [4] H. Frigui and R. Krishnapuram, “Clustering by competitive agglomeration,” *Pattern Recognition*, 1997.
- [5] J. Hafner H. and al., “Efficient color histogram indexing for quadratic form distance functions,” *IEEE Trans. PAMI*, 1995.