# Visual Querying by Color Perceptive Regions

A. Del Bimbo, P. Pala
Dipartimento di Sistemi e Informatica
Università di Firenze
50139 Firenze, Italy

Extended Abstract

## Introduction

Image databases are now currently employed in an eclectic range of different areas such as entertainment, art history, advertising, medicine and industry among others. In all these contexts, a key problem regards the modality of accessing visual image content. The visual content of an image should be intended as what humans remember after looking at it for a couple of minutes. It may include shapes of relevant objects, color distribution or color patches, textured surfaces, or the arrangement of visual elements on the image plane. In image databases, unlike textual and numerical databases, the visual content by itself has no ordering rules, unless textual labels are associated with images. Conventional attempts to cast visual features into textual keywords [Srihari] has been far recognized to be inadequate in indexing pictures. The minor expressiveness of text with respect to visual features doesn't allow to fully exploit capabilities of human memory. Items retrieved through a textual query could not be relevant at all for user's expectation.
This is the reason why retrieval based on visual content has been identified as the means to overcome this modal clash. Pattern recognition and image analysis algorithms are used to extract parameters of the visual content and filter out unrelevant images. The relevance of visual elements for the final user depends on his subjectivity (which is not known in advance, when the database is created) and on the context of application (which is known, instead). On the one hand, user's subjectivity implies that modules are employed which are able to cope with imprecision and lack of knowledge about image content. On the other hand, the fact that the context of application is known in advance, can drive the choice of pattern recognition algorithms and functions that must be included in the system. For example, spatial arrangement is relevant to a doctor examining a heart section; color arrangement is important when looking at paintings, but the presence of particular color tones or the disposition of color patches on the canvas will be better remembered by an expert than by an occasional user.

Generally speaking, an efficient system for retrieval by visual content should be able to:
- provide a query paradigm that allows users to naturally specify both selective and imprecise queries;
- define retrieval facilities that are meaningful for the context of application;
- define similarity metrics which are satisfactory for user perception.

Querying by visual example is the interaction paradigm that exploits human natural capabilities in picture analysis and interpretation. It requires that visual features are extracted from images and used as indexes in the retrieval phase.
Querying by visual example allows imprecision and incompleteness of expression, by letting users draw a sketch of their memory of the image (for example colored shapes arranged in a specified pattern). It must not be expected that the first answer be already satisfactory. Rather, a continuous interaction must be foreseen, through which
the original query of the user can be refined or changed on the basis of the retrieved pictures.

In recent years, prototype systems have been proposed which implement retrieval by content from image databases by addressing different facets of the informative contents of pictorial data, such as object texture organization, [Picard] shape similarity with visual sketches [Kato] , [Mehrotra], [DelBimbo97], semantic relationships between imaged objects [Chang-84], [Chang-87], [DelBimbo-95].
With the increasing availability of devices supporting acquisition and visualization of color images, a growing attention is also being focused on chromatic features as a key facet in image content.

In this presentation we discuss the PICASSO system, developed at the Visual Information Processing Lab of the University of Florence, Italy. The system is a complete framework providing facilities for image indexing and retrieval based on shapes, colors, and spatial relationships. Only color retrieval facilities are discussed in this paper. For shape and spatial relationship based retrieval the reader can refer to [DelBimbo-97] , [DelBimbo-96].

Concerning colors, the system supports both retrieval by global color similarity and retrieval by similarity of color regions. Shape, size, and position of color regions are considered as optional features that the user can select in the query.

## Color-based Image Retrieval Experiences

Previous experiences in color--based retrieval essentially address two kinds of problems. The first regards finding database images whose color distribution is globally similar to that in a query image. In this problem, the user's interest lies on the chromatic content of the whole image: what is represented in a picture has no particular relevance. In painting databases, for example, this kind of query could help in finding paintings of the same author, or of a certain author's period, or perceptually similar paintings.
The second problem is finding a certain object in a complex scene, using its chromatic properties. The rest of the scene is not relevant for the user's purposes. Only the presence and location of the object are interesting.

Image retrieval based on global color distribution has been formerly proposed by Swain and Ballard [Swain]. Image color distribution is represented through color histograms, which are obtained by discretizing the color space, and counting how many pixels fall in each discrete color. Color histograms have the property of being invariant to translation and rotation. They slowly change when objects are partially occluded, or when there are changes in scale and angle of view. Retrieval is performed by evaluating the intersection between the global color histogram of the user--provided example and stored images.

The QBIC database system [QBIC], [Cody] evaluates similarity in terms of global properties of color histograms. A weighted distance measure is used to evaluate the similarity of color histograms. The distance measure is a weighted cross correlation between histogram bins. Weights represent the extent to which two histogram bins and are perceptually similar to each other.
In [Jain] A.K. Jain and A. Vailay used both color and shape features, analyzed over the whole image. In their system, the query is formulated through an example image and retrieval is accomplished by a similarity measure computed on the basis of the global color histogram and image edges.

Few authors have allowed retrieval of images based on similarity of color patches or objects. This is due to the inherent major complexity of this kind of retrieval. Searching for localized color regions, eventually corresponding to relevant objects, requires to use effective algorithms to locate uniform color regions, more complex data structures to represent color image properties, and more complex algorithms to evaluate perceptual similarity.
In [Tanaka] images are partitioned into blocks of equal size, each associated with its own local histogram. Similarity matching considers adjacency conditions among blocks with similar histograms. However, blocks are created according to a static partitioning of the image, which is generally inadequate to reflect the original arrangement of colors in a complex image.
In [Vinod] an iterative technique is proposed to identify image regions which can potentially represent a given object, based on color features. Color Histogram Intersection is used to evaluate the match between image regions and query objects. Regions of potential interest are extracted considering a square window and shifting it on the image by a fixed number of pixels in one direction at a time.

A different and more reliable solution requires that the whole image is segmented into homogeneous color regions. Chromatic and geometric features of such regions are matched against corresponding regions in the query model. Segmentation is typically the hardest problem to be solved. Matching of segmented images with query color patches or segmented objects is also difficult. In a query by example, the color patches sketched by the user correspond to his approximated view of the image searched, rather than to the true image patches. Shapes of color regions of database images, as resulted from the segmentation, could not fit shapes of regions specified in the query.
The PICASSO system presented in this paper exploits a hierarchical multi-resolution segmentation and graph matching in order to support effective retrieval based on color regions. The system has been designed mainly for content based retrieval of paintings and art images, in which the assumption of smooth changes of colors within homogeneous regions does not hold.

## Hierarchical Color Image Segmentation

The number r of regions which are produced in the segmentation process determines the level of precision of the image partitionment. The value of r can be determined adaptively on the basis of statistical analysis of color distributions in the sampled color space [Corridoni] However, in the framework of an image retrieval system and in the absence of specific assumptions about images being stored, one such approach may lead to segmentations

which do not meet user's expectations about perceptual groups. Moreover, at storage time it is impossible to forecast the level of precision which will be requested in the detection of image properties expressed by users in specific queries.
If the segmentation process fails to detect a region in an image or performs a segmentation into regions which is different from that expressed by the user in the query, that image will not be retrieved in the searching process. Both these considerations evidence that, at storage time, the optimal level of precision cannot be defined.

In the PICASSO system, this hurdle is circumvented by creating multiple descriptions of each data, each one covering a different level of precision. Images are analyzed at different levels of resolution in order to obtain a pyramidal segmentation of color patches. Each region at level n is obtained by clustering adjacent regions at level n-1. A region energy is associated to each region. This energy is obtained as a weighted sum of three entries:
- the area;
- the color uniformity;
- the color contrast.

The image energy is defined as the sum of all region energies. Image segmentation is performed by minimizing a function such that the area is to be maximized, the color uniformity to be maximized and color contrast to be minimized.
Image segmentation is performed by iteratively updating region clusters at each resolution level, separately. At the lowest level of the pyramid, each region corresponds to a pixel in the image. Starting from the lowest level of the pyramid, two adjacent regions are searched whose merging would decrease the image energy. This procedure is recursively applied until the coarsest resolution is reached, such that the entire image is represented by a single region.

## Color Region Representation

In order to support effective retrieval by content of paintings, a color space has been chosen, such that close distances in the color space correspond to close distances for the user perception. This condition, which is not exhibited by the RGB space, has been accomplished with the adoption of the perceptually uniform CIE L*u*v* color space. Close colors in this space correspond to perceptually close colors.
In our approach, a uniform tessellation of the L*u*v* space has been performed and the number of colors has been reduced to 128.

## Region description

Color regions are modeled through their spatial location, area, shape, average color, and a binary 128-dimensional color vector. Entries in the color vector correspond to reference colors. If a reference color is present in the region, its corresponding entry in the color vector is set to 1 (0 otherwise).

At the coarsest resolution of its pyramidal representation, the image is represented by a single region with a color vector retaining the global color characteristics for the whole image. As the resolution increases, regions correspond to smaller area in the image and are thus characterized by the presence of a smaller number of reference colors. This yields an increasing localization of chromatic properties into smaller regions.

## Color Image Retrieval

The PICASSO system supports retrieval by visual example of images with one or several colored regions. Queries may include one or more sketched color regions of any shape and in any relative position. Position, area, elongation, and orientation attributes can be selected as relevant features of a sketched region. Color regions are sketched either by drawing a region contour and then filling it with a color selected from a color picker, or by contouring a region of an image that has been previously answered to a query, or that is selected from a set of samples. At database population time, images are automatically segmented and modeled through a pyramid structure as expounded previously. The highest node of each pyramid includes both the binary color vector associated with the whole image, and the image color histogram. Image color histograms represent images global color appearance. They are used to compute the similarity between two images in terms of their global chromatic contents.
A color index file is built by considering color vectors of the highest nodes associated with all the database images. The color index file has 128 entries, one for each reference color. Each entry storing a list of database images where that reference color is present.

Given a query, the color index file is used, to select a set of candidate images that contain regions with the same colors as the query. Unrelevant images which do not contain some region with the same color as the query are quickly filtered out. The pyramid structure of each candidate image is analyzed in order to find the best matching region R_I for each query region R_Q. Given a query region Q_R, the method used to find the best matching region of an image pyramid G can be summarized as follows:

```
procedure match (R_Q, G)
> Let V be the vertex of the pyramid  G$
> Initialize the match M=0
> analyze (V, R_Q)
> return M
end
procedure} analyze(V, R_Q)
> Let C be the color of R_Q
> {\bf if} $C$ is contained in the color vector of V
> Compute the match $T$ between $R_Q$ and V, according to MATCH)
> if T>M
> update M with T
> if V is a leaf
> stop
> else
> let V_1...V_K be the children of the current node
> for k=1 to K
> analyze(V_k,R_Q)
end
```

A similarity coefficient for the whole image is evaluated as the sum of scores of the best matching image regions for each query region, PICASSO system also supports retrieval by global color similarity. Querying by global similarity supports the user when he is interested in finding images with similar global chromatic contents, without having to localize them in a specific region in the image. Retrieval by global color similarity is carried out by evaluating the correlation between color histogram of the query image and that of database images. Experimental results are presented for a database of paintings.

## Bibliography


[Chang-84] S.K. Chang and S.H. Liu.
Picture indexing and abstraction techniques for pictorial databases.
IEEE Trans. on Pattern Analysis and Machine Intelligence, 6(4), July 1984.

[Chang-87]
S.K. Chang and C.W. Shi, Q. Y.Yan.
Iconic indexing by 2-{D} strings.
IEEE Trans. on Pattern Analysis and Machine Intelligence, 9(3) pp. 413-427, July 1987.

[Cody] W.Cody.
Querying multimedia data for multiple repositories by content: The GARLIC project.
In Proc. on Visual Data Base Systems III, Lausanne, 1995.

[Corridoni] J.M. Corridoni and A. Del~Bimbo.
Multi-resolution color image segmentation.
In Dip. Sistemi e Informatica, Universita di Firenze Tech. Rep.,
no. 7-96, Jan. 1996.

[DelBimbo-95] A. Del Bimbo, E. Vicario and D. Zingoni.
Symbolic description of image sequences with spatio-temporal logic.
IEEE Trans. on Knowledge and Data Engineering,
(7)4 pp. 609--621, Aug. 1995.

[DelBimbo96] A. Del Bimbo, P. Pala.
Image Indexing Using Shape Based Visual Features
International Conference on Pattern Recognition, Wien, August 1996.

[DelBimbo97] A. Del Bimbo and P. Pala.
Visual image retrieval by elastic matching of user sketches.
IEEE Trans. on Pattern Analysis and Machine Intelligence.
(19)2, pp. 121-133, Feb. 1997

[Huang] C.L. Huang, T.Y. Cheng and C.C. Chen.
Color images' segmentation using scale space filter and Markov random field.
Pattern Recognition, (25)10 pp. 1217--1229, 1992.

[Jain] A.K. Jain and A.Vailaya.
Image retrieval using color and shape.
Pattern Recognition, 29(8) pp. 1233-1244, Aug. 1996.

[Kato] K.Hirata, T.Kato,
Query by Visual Example: Content-Based Image Retrieval.
In Advances in Database Technology - EDBT'92, A.Pirotte, C.Delobel, G.Gottlob (Eds.), Lecture Notes on Computer Science, Vol.580,

[Mehrotra] R.Mehrotra and J.E. Gary.
Similar-shape retrieval in shape data management.
IEEE Computer}, 28(9) pp. 57-62, Sept. 1995.

[Picard] F. Liu and R.W. Picard
Periodicity, directionality, and randomness: Wold features for image modeling and retrieval.
M.I.T. Tech. Rep., no. 320, 1995.

[QBIC] W.Niblack et alii,
The QBIC Project: Querying Images by Content Using Color, Texture and Shape
Res.Report 9203, IBM Res.Div. Almaden Res.Center, Feb.1993.

[Srihari] R.K. Srihari.
Automatic indexing and content-based retrieval of captioned images.
IEEE Computer, 28(9) pp. 49-56, Sept. 1995.

[Swain] M.J.Swain, D.H.Ballard,
Color Indexing
Int.Journal of Computer Vision, Vol.7, No.1, 1991.

[Tanaka] A. Nagasaka and Y. Tanaka.
Automatic video indexing and full video search for object appearances.
IFIP Trans., Visual Database Systems II, pp 113-127, Knuth, Wegner (Eds.), 1992. Elsevier.

[Vinod] V. Vinod, H. Murase,
Focussed Retrieval of Color Images
To appear on Pattern Recognition,