# Information Retrieval of Color Images via Wavelets

**D. Vitulano**

I.A.C. - C.N.R.
Viale del Policlinico 137
00161 Rome Italy
E-mail: vitulano@iac.rm.cnr.it

**F. Santosa**

I.M.A. - University of Minnesota
400 Lind Hall, 207 Church Street S.E.
MN 55455 USA
E-mail: santosa@ima.umn.edu

## ABSTRACT

This paper focuses on a novel framework for information retrieval of images based only on color information. Regions homogeneous in color of the input image are detected using a multiresolution analysis. In order to avoid the selection of regions having similar color, a new clustering scheme based on *repulsion between electric charges*, is proposed. The resulting model is completely automatic and achieves good results in terms of computational time and performances.

### Keywords

Color Images, Information Retrieval, Wavelet Transform

## 1 INTRODUCTION

*Information Retrieval* is a well-known problem in Digital Image Processing. It consists of searching among all images belonging to a database, the ones most similar to a given image $I$, called *query*. Many approaches have been proposed in literature and they can be coarsely split in [Nib93a]: *i)* approaches based on features extracted from raw images, such as pixel intensity, histogram, etc. (see [Alb99b]); *ii)* approaches based on the processing of coefficients in compressed transform domains [Wan97a]. Among these latter, a particular interest has been devoted to WT (Wavelet Transform) [Mal98a]. In fact, it allows us to hierarchically and directionally decompose information, exploiting the best rate-distortion performance [Lia97a].


***WSCG POSTERS proceedings***
*WSCG'2004, February 2-6, 2004, Plzen, Czech Republic.* 

Some examples of information retrieval based on WT are in [Man99a], where a technique based on the histogram of wavelet coefficients of different bands is presented. In [Lia97a], a model mainly based on a packet tree structure and subband significance is proposed while in [Wan97a] *DWT* coefficients are compared for image indexing.
The proposed model will focus on color based retrieval. Tools involved for color are: *histograms*, *moments* (first and second), *color coherence vectors* (CCV) (histogram where there is a classification between pixels belonging or not to uniformly-colored regions), *correlograms* (capturing spatial correlation between identical colors). An interesting comparison among them can be found in [Ma98a]. It is worth noting that using only color information could represent a limit since robust retrieval also needs additional information such as textures and shape. It has been proved that color information alone may lead to false positive [Seb02a]. Nonetheless, it can be seen as a component of a more general retrieval system. In fact, the commercial developed systems (like QBIC of IBM etc.) use one or more components (color, texture and so on) depending on the user's query.
The contribution of this paper is to propose a new framework exploiting only color information of im-

ages using a wavelet decomposition. The underlying idea is to determine regions homogeneous in color that should correspond to objects or semantic components of the image, such as sea, sky, grass and so on. Achieved results are very promising and robust on real world images.

## 2 THE PROPOSED MODEL

Before presenting the model, two important choices concerning *color space* and *wavelet* basis have to be made. The selection of the best *color space* is still an open problem [Liu94a]. Different color spaces have been used in literature such as HSV [Seb02a], YIQ [Pat00a], etc. In this paper we will use YES [Sab96a] since its well-known properties. As regards the basis, again, many authors prefer wavelets with few vanishing moments for computational reasons. Haar wavelets are ideal for this purpose but they show artifacts around contours [Wan97a]. In order to overcome this drawback, we will use Daubechies with 4 vanishing moments, since they are a good compromise between computational time and performances.

Let us now show our model. Starting from a color image $\{f_Y, f_E, f_S\}$, we consider just E and S components. These components are then decomposed in the *db4* basis (Matlab notation) obtaining a multiresolution analysis: $\{A_d^{2^M} f_E, D_i^{2^j} f_E\}$, $\{A_d^{2^M} f_S, D_i^{2^j} f_S\}$ where $j = 1,..,M$ is the scale level and $i = 1, 2, 3$, respectively horizontal, vertical and diagonal component. Then, we make the following hypothesis: *Objects of a pictorial scene correspond to regions homogeneus in color.* This seems to be in agreement with the human visual system, based on a low pass operation performed on both luminance and chrominance components [Gon96a]. So, we have to find out regions with: **low energy** and **wide area**. These requirements can be implemented in a computationally attractive way, using only thresholding and logical (AND) operations.

The horizontal component of the first scale level of E component of the input image, i.e. $D_i^{2^j} f_E(x, y)$ with $i = 1, j = 1$ can be thresholded through $t_{1,1,E} = \sqrt{\mathcal{E}(D_i^{2^j} f_E(x, y)))}$, where $\mathcal{E}(D_i^{2^j} f_E(x, y))$ is the energy of the whole horizontal component of E at the first scale level. In practice, for this component we have tuned the threshold proportionally to the average of its energy. Similarly, at the second scale level, the threshold will be computed just considering its coefficients in correspondence of the ones under $t_{1,1,E}$ of the first scale component — suitably down-sampled. The same process can be iterated for scale levels. In practice, we are following the low energy regions through the scales. The same process can be also followed for the vertical components. A further AND will determine the regions of low energy considering information of both horizontal and vertical information. Finally, another AND operation between E and S results will give $N$ compact regions $R_1, .., R_N$ — considering 8 connectivity.

We outline that the progressive lowering of the energy threshold along increasing scale level takes into account the fact that the modulus of the wavelet coefficients becomes more and more regular as the scale level increases [Mal98a].

It may happen that two different regions correspond to the same semantic region with a partial occlusion from the camera view — eg. naked arm with a watch. Then, we must impose another requirement eliminating this drawback: *selected regions have to belong to semantically different regions, i.e. they must have different color.* We will manage this problem using the repulsion between electric charges.

In order to select the $p$ regions, we have to extract the ones minimizing intersection in color besides their size (i.e. their representativeness in the image), using the following strategy.

Considering the first region $R_{h_1}$, each element $(x, y) \in R_{h_1}$ correspond to a couple of approximation values: $\{A_d^{2^M} f_E(x, y), A_d^{2^M} f_S(x, y)\}$, i.e. the approximation value of respectively E and S component in correspondence of the location $(x, y)$. Then, in the ES space it corresponds to a point that we will consider a unitary electric charge. More in general, all points $\{A_d^{2^M} f_E(x, y), A_d^{2^M} f_S(x, y)\}$ with $(x, y) \in R_{h_1}$ have the same kind of charge — suppose positive. Then, to select the $p$ regions among the $N$ available we do as follows. The first selected region $R_{h_1}$ will have the maximum size among the $N$ regions. For selecting the second one we put in the E-S space the charges of $R_{h_1}$ and, in turn, the charges of the second region, the third one and so on. We will select the one causing the least repulsion force according to the Coulomb law. In practice, the charge of each region is equivalent to a uniform charge multiplied by the number of pixels of the region itself. That is, $\forall R_l$ with $1 \leq l \leq N$: $Q_l = |R_l|\, q$, where $q$ is the unitary charge. Hence, the second region to be selected is the one minimizing the repulsion force (Colulomb law): $F \simeq \frac{Q_{h_1} Q_m}{d(\mathbf{x}_{h_1}\mathbf{x}_m)^2} \quad \forall m \neq h_1$, where $\mathbf{x}_m$ is the vector of the coordinates of the center of mass of the region $m$ and $d$ is the distance function. The corresponding energy will be $L_{h_1,m} = \frac{|R_{h_1}|q|R_m|q\ \delta s}{d(\mathbf{x}_{h_1}\mathbf{x}_m)^2\ (\delta R)^2}$, where $\delta s$ is the unitary displacement in the space of charges while $\delta R$ rep-

resents the unitary area[1].
Then, in the set of regions produced by the first minimization (involving thresholding) the quantity $E_l = \frac{\mathcal{E}(R_l)}{|R_l|} \leq \frac{\mathcal{E}(R)}{|R|} \quad \forall l$ is minimized. Hence, accounting for both the above energies, we have to minimize:

$$\begin{aligned} E_m\, L_{h_1,m} &\equiv \frac{\mathcal{E}(R)}{|R|} \frac{|R_{h_1}|q|R_m|q\ \delta s}{d(\mathbf{x}_{h_1}\mathbf{x}_m)^2\ (\delta R)^2} \\ &\propto \frac{\mathcal{E}(R)}{|R|} \frac{|R_{h_1}|q|R_m|q\ \delta s}{d(\mathbf{x}_{h_1}\mathbf{x}_m)^2\ |R_m|^2}. \end{aligned} \quad (1)$$

Notice that the second equality allows us to give a weight to the size of the $m$-th considered subregion so that the concept of frequency of a given color is inserted in our minimization. The latter consists of maximizing: $(d(\mathbf{x}_{i_1}\mathbf{x}_j))^2\ \frac{|R_j|}{|R_{i_1}|}$. Summing up, the second good region to be selected is the one whose center of mass maximizes its distance from the center of mass of the charge already present in the space, suitably weighted by the ratio between its area and the area of the first region — the greatest one. The above process is repeated for selecting the third region and considering the charge corresponding to the two already in the space. It is iterated until all $p$ regions have been determined.

## 3 SOME RESULTS AND CONCLUSIONS

A robust retrieval based on color strongly depends on the illumination in the acquisition phase. In order to overcome this problem we adopted the model proposed in [Dre98a]. It assumes that changes in illumination can be strongly minimized via a suitable normalization of the RGB components. This model works very well on lambertian and flat images [Dre98a]. These requirements are in a complete agreement with hypotheses of model proposed in this paper, since we are interested to regular (without strong oscillations) regions. Using this normalization our performance strongly improves.
We have performed our model on many images and on different databases. We will only give here two examples of how our framework performs. In Fig. 1 we show two original images and their corresponding selected regions. We choose to keep only two coefficients, i.e. two representative regions: in this case, sky and grass. It is interesting to note that in both cases, and also in all images we dealt with, the proposed framework showed a high ability in catching semantically different regions.

[1] these two quantities have been introduced for physically dimensional reasons

Concerning objective measures of retrieval, we achieved good results in *recall* [Sal83a] but just fair in *precision*, because of false positives we have already mentioned. Therefore this approach can be a good tool to take part of an integrated framework for retrieval.
It is easy to see that the computational effort is very moderate since the method is based on multiresolution analysis, simple thresholding, and logical operations on binary matrices. The minimization concerning color overlapping is done by concentrating the charge in the center of mass. Hence we achieve the final result in a few operations and retrieval is completely automatic. Our model is invariant to camera (by which images have been acquired) rotations, invariance to limited camera translations (i.e. still framing the main objects of the scene). Finally, a fixed number of features, allowing us to exploit fast research algorithms like k-d-tree, is involved.

## REFERENCES

[Alb99a] M. G. Albanesi, M. Ferretti, A. Giancane, A Compact Wavelet Index for Retrieval in Image Database, *Proc of 10th ICIAP*, September, 1999, ISBN 0-7695-0040-4, pp. 927-932

[Alb99b] E. Albuz, E. Kocalar, A. A Khokhar, Vector-wavelet based scalable indexing and retrieval system for large color image archives, *Proc. of IEEE ICASSP*, March, 1999, vol. 6, pp. 3021-3024

[Dre98a] M. S. Drew, J. Wei, Z. N. Li, Illumination-Invariant Color Object Recognition via Compressed Chromaticity Histograms of Color-Channel-Normalized Images, *IEEE Proc. of ICCV '98*, 1998, pp. 533-540

[Gon96a] R. C. Gonzales and R. E. Woods, "Digital Image Processing", Prentice Hall, 2002

[Lia97a] K. C. Liang, C. C. J. Kuo, Progressive Image Indexing and Retrieval Based on Embedded Wavelet Coding, *Proc. ICIP*, October, 1997, vol. 1, pp. 26-29

[Liu94a] J. Liu, Y. H. Yang, Multiresolution color image segmentation, *IEEE Trans. PAMI.*, July, no. 16, 1994, pp. 689-700

[Ma98a] W. Y. Ma, H. J. Zhang, Benchmarking of Image Features for Content-based Retrieval, *Proc. of 32-th Asilomar Conference on Signals, Systems and Computers*, 1998, vol. 1, pp. 253-257

[Mal98a] S. Mallat, "A Wavelet Tour of Signal Processing", Academic Press, 1998

[Man99a] M. K. Mandal, T. Aboulnasr, Fast Wavelet Histogram Techniques for Image Indexing, *Computer Vision and Image Understanding*, July/August", no. 1/2, 1999, vol. 75, lspp. 99-110

[MIT] http://www.cs.washington.edu/research-/imagedatabase/groundtruth/ referenced by MIT at http://www-2.cs.cmu.edu/afs/-cs/project/cil/ftp/html/v-images.html

[Nib93a] , W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasmann, D. Petkovok, P. Yanker, C. Faloutsos, G. Taubin, The QBIC project: Querying images by content using color, texture and shape, *IBM Research Report*, February, RJ 9203 (81511), 1993.

[Pat00a] A. Patel, M. Tonkelowits, WHASSUPP: A Novel Approach to Query-by-Sketch Using Wavelet Coefficients and Color Histograms, *Tech. Rep.*, Harvard University, 2000

[Sab96a] , E. Saber, A. M. Tekalp, R. Eschbach, K. Knox, Automatic Image Annotation Using Adaptive Color Classification, *Graphical Models and Image Processing*, March", 2, 1996, Vol. 58, pp. 115-126

[Sal83a] G. Salton, M. J. McGill, "Introduction to Modern Information Retrieval", McGraw-Hill, New York, 1983

[Seb02a] , N. Sebe, Q. Tian, E. Loupias, M. S. Lew, T. S. Huang", Evaluation of Salient point Techniques, *Proc. of CIVR 2002*, July, 2002, pp. 367-377

[Wan97a] J. Z. Wang, G. Wiederhold, O. Firschein, S. X. Wei, Wavelets-based Image Indexing Techniques with Partial Sketch Retrieval Capability, *Proceedings IEEE International Forum on Research and Technology Advances*, May, 1997, ADL 97, pp. 13-23

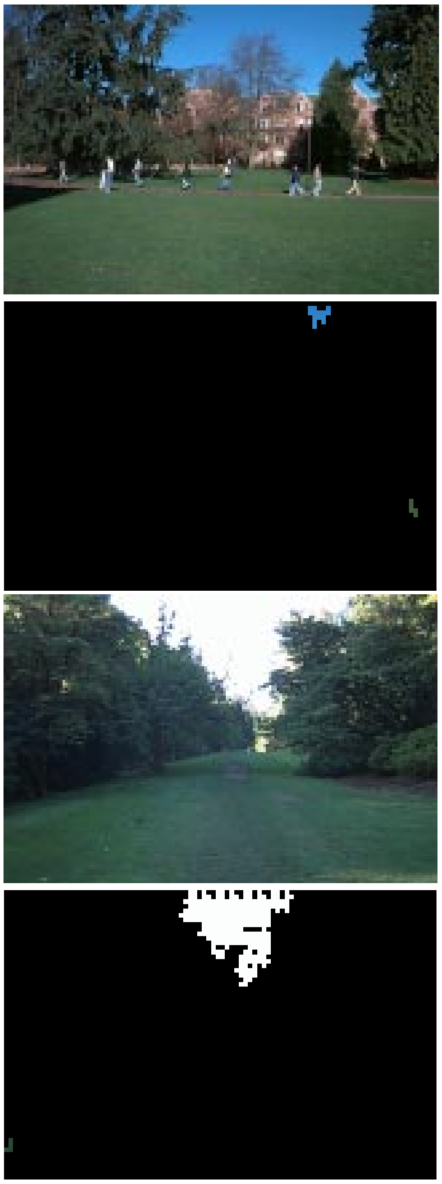

**Figure 1: Two color images [MIT] and their related selected regions (colored) on a dark background.**