# Object Image Retrieval with Image Compactness Vectors

Catherine Achard
*Laboratoire des Instruments et Systèmes, Paris, France*
*achard@ccr.jussieu.fr*

Jean Devars
*Laboratoire des Instruments et Systèmes, Paris, France*
*devars@ccr.jussieu.fr*

Lionel Lacassagne
*Laboratoire des Instruments et Systèmes, Paris, France*
*lionel@lis.jussieu.fr*

## Abstract

*We present in this paper a new global measure to characterise an image: the compactness vector. This measure considers both object shape and grey level distribution function and does not require any preliminary segmentation. It presents many invariants as rotation, translation, scale and luminance and is then a powerful tool for image retrieval from a query image. We present here some object retrieval examples from large database images.*

## 1. Introduction

Due to the low cost of scanners and storage devices, large images databases have been created. They are used in many applications as archives of criminal faces, archives of radiography with some pathology, and so one. These databases increase every day and a procedure is required for automatically indexing and retrieving images.

Previously, features as filenames or keywords have been used to characterise images. But when the base is very large, it is really painful for a human to determine keywords. Moreover, for complex images, it is difficult to find all keywords that characterise them.

So, we search for automatically indexing images. That's why we must use characteristics as color, texture, luminance or shape. Currently, methods use a single attribute.

For example, Swain and Ballard [1] use as index color histograms. To compare indexes, they employ a histogram intersection measure. Schiele and Crowley [2] introduce different distances between histograms and different local characteristics to compute histograms. They compare results and obtain an insensitive method according to scale and signal intensity variations (for grey level images). Other works, as [3] or [4], lead to a color indexing insensitive to changes in incident illumination. Kankanhalli et al. [5] or Gong et al. [6] use methods based on a clustering algorithm in a 3D-color space to obtain a faster indexing system.

We can too use shape indexes as [7], [8], [9], [10], [11] or [12]. A major inconvenient for these methods is that they need generally a preliminary segmentation. Quality of indexing technique is then dependent of quality of segmentation.

However, a single attribute is too restrictive for describing an image. Without color information, or with two images having the same color, if becomes essential to use shape information. On the over hand, with two objects having similar shape, it is attractive to employ color to distinguish between them. Thus, Jain and al. [13] use color histogram and edge direction histogram for indexing images. In [14], a color adjacency graph is proposed. Each node represents a single chromatic component defined as a set of pixels forming a unimodal cluster in the chromatic scattergram. Edges encode information about of colour components and their reflectance ratio. This method needs a preliminary clustering algorithm. Nastar [15] proposes the image shape spectrum witch is a shape index of the intensity surface. Good results are obtained but the method is not insensitive to change in intensity.

In this article, we propose a new features vector (the compactness vector) based on grey level distribution and shape, for object image retrieval in large databases. It is very easy to compute and does not need a preliminary segmentation. This vector is, by definition, insensitive to rotation, translation and scale. We discuss later a method to incorporate invariance to change in signal intensity.

## 2. The compactness vector

The main idea is to use level sets for characterising an object. The level set "*n*" for a grey level image I(x,y) is defined by:

$$E\{n\} = \{I(x,y) \,/\, I(x,y) > n\} \qquad (1)$$

An image defined with 256 grey levels leads to 256 level sets, each one correlated with the participation the grey levels in the object. We have now to characterise these level sets, which are binary images. For this, we have chosen the compactness defined by:

$$C = \frac{2\pi \ Area}{Perimeter^2} \tag{2}$$

The area and the perimeter are, of course, those of object in the binary image. Compactness is an interesting tool since it has several advantages as implicit invariance to rotation, translation and scale. Computing compactness for each level set leads to the Compactness Vector (CV). It will be used as index for image retrieval.

Remark: for all images, this vector has some similarities:

- its first component is the compactness of the image frame.
- the last component is always zero.

The discriminative part of the vector is then its evolution between these limits.

## 3. Implementation details

Directly implemented the CV is very expensive in computation time. However, using mathematical morphology, it can be expressed very simply.

To obtain the area of each level set, it suffices to compute the reverse cumulate histogram (cumulate from high to low levels). To reach the perimeter, a multi-level morphological dilatation is required. Perimeter of the level set "*n*" is then obtained by subtracting the cumulate histogram of the dilated image and the cumulate histogram of the original image, and this, for the level "*n*". So, the CV computation can be decomposed in tree steps for an original image I:

- Compute the multi-level morphological dilatation D.
- Compute the reverse cumulate histogram (from right to left) of I and D : HI and HD
- The area of the level set "*n*" is given by HI(n) while his perimeter is given by HD(n)-HI(n).

Remark: the morphological dilatation can be realised with several structuring elements (with unitary radius) without really influence results. We use, for results presented in this article, the extremal filter: $\begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{matrix}$

## 4. Measure between two CV

Several distances can be used for quantifying the similarity between two CV. Among them, we can cite:

- the L1 distance:

$$d1(V1,V2) = \sum_i \left| V1(i) - V2(i) \right| \tag{3}$$

- the L2 distance:

$$d2(V1,V2) = \sum_i \left[V1(i) - V2(i)\right]^2 \tag{4}$$

- the normalised L1 distance:

$$d1N(V1,V2) = \sum_i \frac{\left| V1(i) - V2(i) \right|}{\max\left(V1(i), V2(i)\right)} \tag{5}$$

- the normalised L2 distance:

$$d2N(V1,V2) = \sum_i \frac{\left[V1(i) - V2(i)\right]^2}{\max^2\left(V1(i), V2(i)\right)} \tag{6}$$

## 5. Invariance

### 5.1. Invariance to rotation and translation

In order to test rotation and translation invariance, we have built a base of four objects. For each one, we have taken six images in different configurations of position. The base is then composed of 24 images. A retrieval example for a multimeter is given figure 1 where the eight more similar images with the L1 distance are presented. This result shows the capacity of the CV to be a significant signature and to be insensitive to rotation.

**Figure 1- Retrieval of the top left image**

### 5.2. Invariance to lighting

Using distances presented in section 4 does not allow the illumination invariance. Indeed, signal intensity change leads to translations combined with contraction or dilatation on the CV. An example is given figure 2 where two images with different lighting and their CV are presented. For these images, we can see on the CV a translation and a contraction.

So, if we want to be insensitive to signal change, we have to introduce a distance which authorise displacements along the grey level axe. It is the "edition distance" that has been defined for computing distance between chains. Let us consider two chains $x=a_1a_2\ldots a_n$ and $y=b_1b_2\ldots b_m$. We note $x(i)=a_1a_2\ldots a_i$ and $y(j)=b_1b_2\ldots b_j$. By convention, we put $x(0)=y(0)=\lambda$ where $\lambda$ is the neutral element. We compute then the distance $D(i,j)=d(x(i),y(j))$ by recurrence to obtain the searched measure $D(x(n),y(m))=d(x,y)$. Initialisation is realised by $D(0,0)=0$. Then, we compute by iteration:

$$D(i,j)=\min\begin{cases} D(i-1,j-1)+C(a_i,b_j), \\ D(i-1,j)+C(a_i,\lambda), \\ D(i,j-1)+C(\lambda,b_j) \end{cases} \quad (7)$$

- $C(a_i,b_j)$ is the substitution cost of $a_i$ by $b_j$,
- $C(a_i, \lambda)$ is the removing cost of $a_i$,
- $C(\lambda,b_j)$ is the insertion cost of $b_j$.

Here, we have chosen $C(a_i, b_j) = C(a_i,\lambda) = C(\lambda,b_j) =$ *dist*(ai,bj) where *dist* is one of the four distances presented chapter 4.

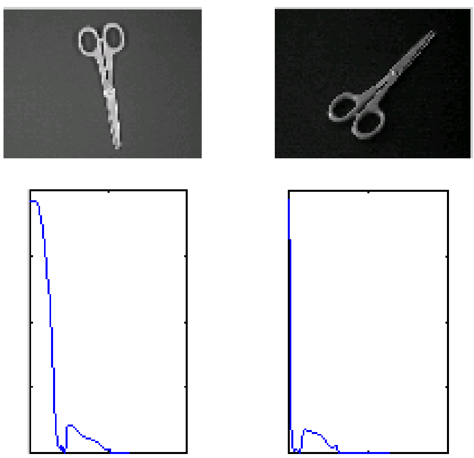

**Figure 2- Top : Original images with different illumination. Bottom : their CV.**

To test this new measure, we add to the preceding base four images per object taken with different illuminations (the base is now composed by 40 images). A retrieval result where *dist* is the L1 normalised distance is presented in figure 3. Note that we have also introduced a change in scale.

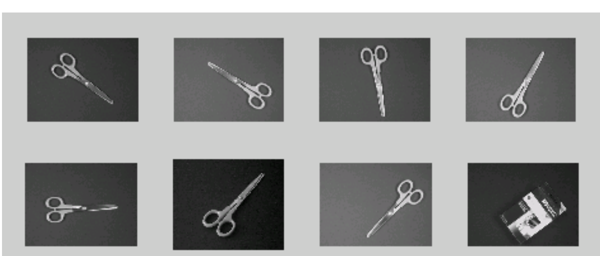

**Figure 3 - Retrieval of the top left image**

The CV seems to be insensitive to the usual desired transformations, but we have now to prove its discriminatory power. Then, we have to work on larger databases.

## 6. Experimental results

Results on two bases are presented.

The first one is the Olivetti Research Laboratory Face database that contains 400 images (40 persons and 10 images per person). Images are taken with varying lighting, pose, facial expressions and facial details.

The second one is the Columbia database that contains 1440 images of 20 3D objects. There are 72 images per object taken with 5 degrees increments direction.

It is important to keeping mind that we have none priori knowledge about the number of classes present in the base. The goal is not to recognise objects but to retrieve similar images to the query image.

In order to show some databases images, we present in figure 3 and 4 two retrieval examples, one for each base. We see on these results that the CV seems to be a good index since it can find similar images from a query image.

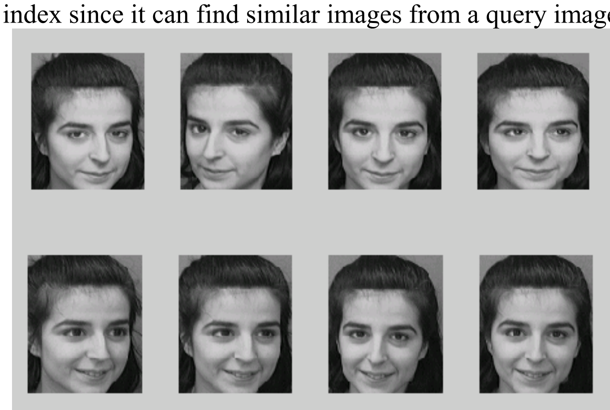

**Figure 3- Retrieval of the top left image. Olivetti database**

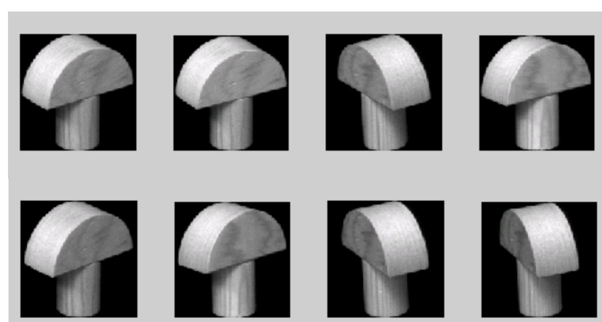

**Figure 4- Retrieval of the top left image. Columbia database.**

To present quantified results on these bases, we use a nearest neighbour rule: if the nearest image was of the same object, a correct score is given. We tabulate below summaries results obtained from the different distances.

| | Olivetti | Columbia |
|---|---|---|
| L1 | 97.25 % | 99.51 % |
| L2 | 96.75 % | 99.17 % |
| normalised L1 | 97.50 % | 99.31 % |
| normalised L2 | 97.25 % | 98.19 % |
| Edition and L1 | 81.50 % | 98.40 % |
| Edition and L2 | 75.00 % | 96.11 % |
| Edition and normalised L1 | 78.75 % | 98.47 % |
| Edition and normalised L2 | 78.75 % | 97.57 % |

**Table 1 - Recognition rate with different distances**

The distance witch leads to the best results is the "L1 distance". Recognition rate better than 95 % are obtained for both bases. When the lighting invariance is desired, scores decrease to 81 % for the first base and 98 % for the second. This consequence is expected as we have added an invariant. Also we can remark that results are better for the second base. Actually, it is because it possesses a bigger number of similar images than the first base.

The CV was defined with all grey levels but we can compute it with a smaller number of grey levels. So, we have studied the behaviour of the CV when it is reduced. Results of recognition are presented in table 2. Differences between CV have been obtained with the L1distance.

| Number of level set | Olivetti | Columbia |
|---|---|---|
| 256 | 97.25 | 99.51 |
| 128 | 97.25 | 99.51 |
| 64 | 97.25 | 99.58 |
| 32 | 97.5 | 99.65 |
| 16 | 96.25 | 99.17 |
| 8 | 87 | 98.47 |
| 4 | 61.75 | 91.87 |

**Table 2 - Recognition rate by decreasing the number of level sets.**

These results show that smaller CV can be used for indexing database images. For both studied bases, 32 grey levels lead to best (or similar) results. On Columbia database, with only four numbers to characterise an image, we obtain a recognition rate of 91.87 %. Naturally, for complex images and large databases, 256 grey levels can be required.

## 7. Conclusion

We present in this article a new index for object indexing and retrieval. It considers both grey level distribution and shape and is then a discriminative tool. This index is a vector called the Compactness Vector witch is defined with one component per grey level. We show that it can be summarise at 32 components without decreasing results quality. The CV has many advantages in the indexing context: it is insensitive to rotation, translation, scale and changes in signal intensity. To establish its discriminatory power, we test it on large databases images. Good recognition rates have been obtained.

## 8. References


[1] M.J. Swain and A.H. Ballard. Color Indexing. International Journal of Computer Vision, 1991, vol. 7, n° 1, pp 11-32.

[2] B. Schiele and J.L. Crowley. Object Recognition Using Multidimensional Receptive Field Histograms. Fourth European Conference on Computer Vision, Cambridge, England, April 1996, pp. 610-619

[3] G. Healey and A. Jain. Using Physics-based Invariant Representations for the Recognition of Regions in Multispectral Satellite Images. Computer Vision and Pattern Recognition, San Francisco, June 1996, pp. 750-755.

[4] Funt and Finlayson. Color Constant Color Indexing. IEEE Transactions on Pattern Recognition and Machine Intelligence, 1995, vol. 17, n° 5, pp. 520-529.

[5] M.S. Kankanhalli, B.M. Mehtre and J.K. Wu. Cluster-based Color Matching for Image Retrieval. Pattern Recognition, 1996, vol. 29, n° 4, pp. 701-708.

[6] Y. Gong, G. Proietti and C. Faloutsos. Image Indexing and Retrieval Based on Human Perceptual Color Clustering, Computer Vision and Pattern Recognition, Santa Barbara, June 1998, pp. 578-583.

[7] B. Günsel and A.M. Tekalp. Similarity Analysis for Shape Retrieval by Example. 13th International Conference On Pattern Recognition, Vienna, Austria, 1996, pp. 330-334.

[8] S. Sclaroff. Deformable prototypes for encoding shape categories in image databases. Pattern Recognition, 1997, vol. 30, n° 4, pp. 627-641.

[9] J. Bigun, S.K. Bhattacharjee, S. Michel. Orientaion Radiograms for Image Retrieval: An Alternative to Segmentation. 13th International Conference On Pattern Recognition, Vienna, Austria, 1996, vol. C, pp. 346-350.

[10] J. Huand, A.R. Kumar, M. Mitra, W.J. Zhu, R. Zabih. Image Indexing Using Color Correlograms. Computer Vision and Pattern Recognition, San Jaun, June 1997, pp. 762-768.

[11] A.D. Bimbo and P. Pala. Effective Image Retrieval Using Deformable Templates, 13th International Conference On Pattern Recognition, Vienna, Austria, 1996, vol. C, pp. 120-124.

[12] A. Califano and R. Mohan. Multidimensional Indexing for Recognizing Visual Shapes. IEEE Transactions on Pattern Recognition and Machine Intelligence, 1994, vol. 16, n° 4, pp. 373-390.

[13] A.K. Jain and A. Vailaya. Image Retrieval using Color and Shape. Pattern Recognition, 1996, vol. 29, n° 8, pp. 1233-1244.

[14] J. Matas, R. Marik, J. Kittler. On Representation and Matching of Multi-Coloured Objects. 5th International Conference on Computer Vision, Cambridge, MA, USA, 1995, pp. 726-732.

[15] C. Nastar, The image shape spectrum for image retrieval, Research Report INRIA n° 3206, july 1997.