Compact color descriptor for object recognition across illumination changes

Damien Muselet and Alain Trémeau
LIGIV EA 3070
Université Jean Monnet - 18 rue Benoît Lauras - 42000 Saint-Etienne - France
damien.muselet@univ-st-etienne.fr
alain.tremeau@univ-st-etienne.fr

Abstract

In this paper, we propose a compact color invariant image descriptor which characterizes both the color distribution and the spatial interactions between the pixels in the image. Our approach is based on the analysis of the pixel rank measures which are their ranks when they are sorted according to their color component levels within a color image. Indeed, we show in this paper that the correlation between the rank measures of neighbor pixels in a color image is an efficient feature to describe the content of this image. This descriptor, extracted from the chromatic co-occurrences matrices, has the advantages to be invariant to illumination changes, low-time consuming, highly discriminating and compact. The proposed rank correlation coefficient is used by our object recognition scheme whose effectiveness is assessed with a public database that contains images of objects acquired under different illuminations.

Introduction

Object recognition problem can be stated in terms of finding among all the target images of a database these which contain the same object as that represented by the query image. In this paper, we specifically address the problem of recognizing objects by analyzing their colors in digital images acquired under different lighting conditions. For this purpose, we consider images that contain one single object (see figure 1). The parameters of the camera are not modified between the different acquisitions of the images of the same object lit by different illuminations. The object can be translated or rotated in a plane perpendicular to the optical axis of the camera.

Figure 1. 9 *objects of the ALOI database (http://staff.science.uva.nl/∼aloi/).*

The color of a pixel is not only a measure of the reflectance properties of the elementary surface of the object projected onto this pixel. It is also a function of both the camera and the illumination. Therefore, the object recognition based on color analysis may fail when the images are acquired under different illumination conditions.

That's why the illumination-invariant image retrieval systems propose to characterize the color images by invariant indexes which are as less sensitive as possible to illumination changes [1, 7, 10, 5, 2]. The determination of these invariant indexes is based on illumination change models which describe the variations of colors caused by any illumination change. The problem is that in order to reduce the complexity of the computation, the classical approaches try to model these variations by linear transformations. Consequently, they use very restrictive assumptions about the camera and the illumination and lead to poor recognition results [4, 6]. Recently, Finlayson has proposed a non-linear illumination change model based on the pixel *rank measures* [3]. Thanks to the non-linearity of this model, the experimental results of this approach are better than those of the previous classical ones [3, 14].

In this paper, we exploit these rank measures, presented in the second section, in order to propose an original illumination invariant index for color images. This index is extracted from the co-occurrences matrices of the image, introduced in the third section, and characterizes the correlation between the ranks measures of neighbor pixels. The fourth section presents the Kendall's coefficient which is used to measure the rank correlation. Since it represents both the color distribution and the spatial interactions between the pixels in the image, the proposed index is highly discriminating. Its effectiveness is assessed with a public database in the fifth section.

Rank measures

A color image $\mathbf{I}$ can be separated into three color component images I^k, $k \in \{R,G,B\}$, where each pixel P is characterized by one color component level $c^k(P)$. Within each color component image, the pixels are sorted in the increasing order of their levels and are associated to a rank, so that the rank is close to 0 for the first ordered pixels, and equal to 1 for the last ordered pixels. Finlayson [3] introduces the rank $\mathscr{R}^k[\mathbf{I}](l)$ of the color component level l which is the rank of the pixels characterized by this level within the color component image I^k and is expressed as :

$$\mathscr{R}^k[\mathbf{I}](l) = \frac{Card\{P \in \mathbf{I}/c^k(P) \leq l\}}{Card\{P \in \mathbf{I}\}}. \quad (1)$$

Note that this rank can be interpreted as the normalized cumulative histogram of the image.

Finlayson assumes that the ranks of the levels within a color component image are not modified by illumination changes [3]. Thus, he proposes to characterize each pixel P by its three ranks

$\mathscr{R}^k[\mathbf{I}](c^k(P))$, $k \in \{R,G,B\}$, and to compute for each image $\mathbf{I}$, the histogram $\mathscr{H}[\mathbf{I}]$ of ranks. Each of its cells $\mathscr{H}[\mathbf{I}](\mathscr{R}^R, \mathscr{R}^G, \mathscr{R}^B)$ contains the number of pixels whose ranks are equal to $\mathscr{R}^R$, $\mathscr{R}^G$ and $\mathscr{R}^B$ in the color component images I^R, I^G and I^B, respectively. Then, Finlayson proposes to compare two images by means of the intersection between their histograms of ranks. He shows that this normalization is equivalent to three 1D-histogram equalizations.

However, the histogram does not take into account the spatial interactions between the colors in the image. That's why we rather propose to use the chromatic co-occurrences matrices.

Chromatic co-occurrences matrices

A chromatic co-occurrences matrix is a generalization of the grey-level co-occurrences matrix proposed by Haralick [11] to color texture analysis [15]. Let us denote $M_d^{k,k'}[\mathbf{I}]$ the chromatic co-occurrences matrix which characterizes the spatial interaction in the color image $\mathbf{I}$ between two color components k and k', $k,k' = R,G,B$, according to the distance d. The chromatic co-occurrences matrix can be considered as an array of cells indexed by color component levels. The cell $M_d^{k,k'}[\mathbf{I}](u,u')$ indicates the number of times that, in the image $\mathbf{I}$, a pixel P' whose level $c^{k'}(P')$ is equal to u', is located at the distance d from a pixel P whose level $c^k(P)$ is equal to u. Given a distance d, a color image $\mathbf{I}$ is characterized by 6 chromatic co-occurrences matrices: $M_d^{R,R}[\mathbf{I}]$, $M_d^{G,G}[\mathbf{I}]$, $M_d^{B,B}[\mathbf{I}]$, $M_d^{R,G}[\mathbf{I}]$, $M_d^{R,B}[\mathbf{I}]$ and $M_d^{G,B}[\mathbf{I}]$. According to the considered image size, the distance d is fixed by the expert. In order to be more discriminating, it can be interesting to use several distances d. Thus, if we consider n_d different distances d_i, $i = 1,...,n_d$, each image is characterized by 6xn_d co-occurrences matrices. For example, we will show in the last section, that the values $n_d = 3$ with $d_1 = 1$, $d_2 = 20$ and $d_3 = 40$ lead to good object recognition rates.

In order to measure the spatial interactions between the rank measures of the pixels in an image, we propose to extract the Kendall's rank correlation from each of the 6xn_d co-occurrences matrices of the images.

Kendall's rank correlation coefficients

Kendall's rank correlation [12] provides a measure of the strength of dependence between two variables by checking the correlation between the rank measures of these variables.

In our case, the variables are the color component levels of the pixels. As illustration, let analyze the rank correlation between the red and green components of the n_{pix} pixels P_i, $i = 1,...,n_{pix}$, of a color image. For this purpose, we have to consider each pixel pair $\{P_i, P_j\}$, $i \neq j$. If the red and green levels of the two pixels P_i and P_j are sorted in the same order, i.e. if $c^R(P_i) < c^R(P_j)$ and $c^G(P_i) < c^G(P_j)$ or if $c^R(P_i) > c^R(P_j)$ and $c^G(P_i) > c^G(P_j)$, the pair $\{P_i, P_j\}$ is called concordant. Otherwise, if these pixels are so that $c^R(P_i) < c^R(P_j)$ and $c^G(P_i) > c^G(P_j)$ or so that $c^R(P_i) > c^R(P_j)$ and $c^G(P_i) < c^G(P_j)$, the pair is called discordant. By analyzing all the pixel pairs among the n_{pix} pixels, we evaluate the measure S as the difference between the number of concordant pairs and the number of discordant pairs. In order to obtain the Kendall's rank correlation coefficient τ, we have to normalize S by the total number of pixel pairs $\frac{n_{pix}(n_{pix}-1)}{2}$ so that:

$$\tau = \frac{2S}{n_{pix}(n_{pix}-1)}. \tag{2}$$

The main advantage of this measure is that it can be directly evaluated from the red-green histogram of the image. Indeed, in the red-green histogram, the red and the green levels are sorted in the increasing order. Thus, on figure 2, we see that each cell c_i of the red-green histogram can be associated with two cell content sums $DISC_i$ (number of discordant pairs) and $CONC_i$ (number of concordant pairs) so that the contribution of a cell C_i to the measure S is $S_i = CONC_i - DISC_i$. For each cell C_i of the 2D-histogram, the sum S_i can be iteratively evaluated and the value of S is deduced from the sums S_i with $S = \sum_i S_i$. This process is very fast.

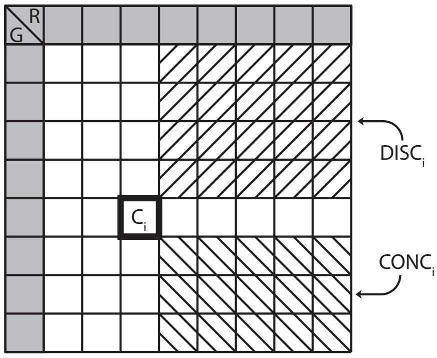

Figure 2. Evaluation of the measure $S = \sum_i S_i = \sum_i (CONC_i - DISC_i)$ from a red-green histogram.

However, this correlation measure represents only the mean rank correlation between the red and the green levels of a pixel without taking into account the spatial interaction between the pixels in the image. In order to compensate this drawback, we rather propose to measure the mean rank correlation between the red level of a pixel and the green levels of the pixels which are located at a distance d from this pixel. Using exactly the same approach as this presented on figure 2, this rank correlation can be easily and fast extracted from the red-green co-occurrences matrix computed for a distance d.

Thus, in order to measure the spatial interactions between the rank measures of the pixels in an image, we propose to extract the Kendall rank correlation coefficient from each of the 6xn_d co-occurrences matrices of the images. Thus, a color image is characterized by 6xn_d rank correlation measures. Let notice that the memory space filled by the proposed index does not depend on the size of the co-occurrences matrices, i.e. it does not depend on the color quantization step.

In order to compare the contents of two different images, we propose to evaluate the Euclidean distance between the vectors constituted by the rank correlation coefficients.

The next section presents the recognition rates obtained by this approach in the context of object recognition across illumination changes.

Experimental results

In this section, the Amsterdam Library of Object Images (ALOI) database [9] is used for testing. The Amsterdam database contains 12 sets of 1000 color images and is available at http://staff.science.uva.nl/∼aloi/. Each set contains images of one object on a uniform background under one of the 12 different illuminants having color temperatures between 2175 K to 3075 K (see figure 3). For the tests, we use the 2 extreme sets of color temperature 2175 K and 3075 K. The 1000 images of the first set are used as the query images and the 1000 images of the second set are used as the target images.

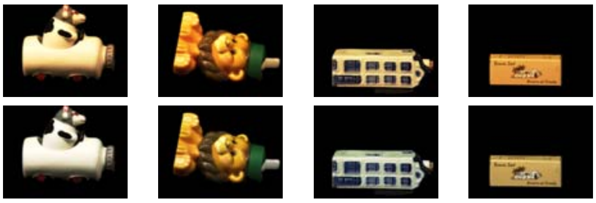

Figure 3. *4 objects of the ALOI database under illuminants having color temperatures of $2175\ K$ (first row) and $3075\ K$ (second row).*

For searching, each of the 1000 objects of the query set is compared with the 1000 objects of the target set. The 1000 target images are ordered with respect to the similarity measures between their invariant color indexes and the invariant color index of the considered query image. When the first ordered target image represents the same object as the query image, the research result is considered as perfect.

First, we propose to use one distance (i.e. $n_d = 1$) whose value d_1 ranges from 1 to 70. The images size is 384x288. Table 1 indicates the percentages of successful object recognition provided by the Kendall's coefficients with each of the distance from 1 to 70. The color component levels are quantized on 8 bits for these tests, i.e. the used co-occurrences matrices contain 256x256 cells. The index of each image is a vector constituted by the 6 Kendall's τ.

Distance d_1	Recognition rate
1	60.1
10	69.4
20	69.5
30	72.4
40	69.8
50	69.2
60	70.1
70	65.3

Table 1 : Object recognition rates obtained by the Kendall's coefficients for different values of distance d_1 when $n_d = 1$. Results obtained on the 1000 objects of the ALOI database.

Table 1 shows two things. First, the results provided by the Kendall's coefficients are very promising. Indeed, whereas the index of each image is only constituted by 6 real values, the recognition rates of 1000 different objects is around 70%. Second, we notice that the rates remain stable whatever the choice for the distace d_1.

In order to increase the discriminating power, we propose now to simultaneously use three distances (i.e. $n_d = 3$). We have tested a lot of combinations of distances $\{d_1, d_2, d_3\}$, $d_i \in [1;70]$ and we have noticed that the choice for the values of d_1, d_2 and d_3 is not crucial. Indeed, since the recognition rates are coarsely stable across variations of the distance d_1 when $n_d = 1$ (see table 1), we propose to choose distances d_1, d_2 and d_3 such as they are well spread in the range $[1;70]$, i.e. such as their values are not too close from each other. For example, in table 2, we present the results provided with $n_d = 3$ and $d_1 = 1$, $d_2 = 20$ and $d_3 = 40$. The results provided with with $n_d = 1$, $d_1 = 1$, $n_d = 1$, $d_1 = 20$ and $n_d = 1$, $d_1 = 40$ are also presented in order to evaluate the impact of the quantization process on these values.

Index	$M = 16$	$M = 64$	$M = 256$	Memory space
Histograms of ranks	**89.1**	83.2	63.1	MxMxM real values
Greyworld histograms	**73.0**	62.2	31.1	MxMxM real values
Kendall's τ of 6 co-occ. mat. ($n_d = 1$, $d_1 = 1$)	38.3	56.8	**60.1**	6 real values
Kendall's τ of 6 co-occ. mat. ($n_d = 1$, $d_1 = 20$)	51.3	68.4	**69.5**	6 real values
Kendall's τ of 6 co-occ. mat. ($n_d = 1$, $d_1 = 40$)	52.5	69.0	**69.8**	6 real values
Kendall's τ of 18 co-occ. mat. ($n_d = 3$, $d_1 = 1$, $d_2 = 20$, $d_3 = 40$)	82.3	**93.8**	93.7	18 real values

Table 2 : Object recognition results obtained by different invariant indexes with the ALOI database. The column "Memory space" indicates the memory space filled by each index in number of real values.

In order to assess the efficiency of our approach, we propose to compare its recognition rates with those of classical approaches. Since the greyworld normalization has been shown to provide the best results among 12 invariant methods [4] and since the histograms of ranks provide better results than the greyworld histograms [3], we compare our results with those obtained by these two approaches.

Columns 2, 3 and 4 of table 2 indicate the percentages of successful object recognition according to different values of M, the number of levels used to quantize each color component. Thus, the histogram of ranks and the greyworld histogram contain MxMxM cells while the co-occurrences matrix contains MxM cells. However, the memory space filled by the Kendall's coefficients extracted from these co-occurrences matrices does not depend on M. The column 5 presents the memory space filled by the indexes.

Table 2 shows that the Kendall's rank correlation coefficients extracted from co-occurrences matrices provide good recognition results while being very compact. Indeed, when we simultaneously consider three distances ($n_d = 3$), the Kendall's index is constituted by 18 real values and leads to 93.8% as recognition rate which is the best result of table 2. This recognition rate is higher than those provided by the histograms of ranks or by the greyworld histograms whereas these indexes require much more memory space (MxMxM real values).

We notice that the recognition results obtained by the histograms of ranks or by the greyworld histograms decrease when M increases. This can be explained by the fact that these histograms are only coarsely invariant to illumination changes and, consequently, they have to be coarsely quantized so that different features appear as identical. However a coarse quantization decreases the discriminating power of these indexes and a trade-off has to be found between invariance and discriminating power. Considering the Kendall's index, we notice that there is no trade-off to find because when M increases, the recognition rates also increase. Furthermore, since the memory space filled by the Kendall's index does not depend on the value of M, we recommend to set M to the highest available value (i.e. $M = 256$ for 3x8-bit color images).

In order to assess the advantages provided by compact descriptors such as the Kendall's index, it is useful to remain the approach of object recognition systems or image retrieval systems. Indeed, these systems are commonly based on three steps: i) the first step is processed off-line and consists in extracting indexes from the target images and in saving them in memory. Then during the on-line part, ii) indexes are extracted from the query image proposed by a user and iii) compared with each previously saved target index. Thus, the advantage of using compact descriptors is twofold. First, little memory space is required to save all the target indexes and second, the computational complexity of the matching step is low [8, 13]. Consequently, the time-processing required to evaluate an index is much less important than the time-processing required to compare two indexes. And the comparison time is proportional to the size of the memory space filled by the indexes.

Conclusion

This paper proposes a compact color descriptor designed for object recognition across illumination changes. It is based on the analysis of the spatial distribution of the rank measures of the pixels. Since these rank measures have been shown to be coarsely insensitive to illumination variations, this index is invariant to illumination changes. Furthermore, we show that the spatial distribution of the rank measures of the pixels can be summarized by few real values which represent the correlation between the rank measures of neighbor pixels. The used correlation measure, which is proposed by Kendall, is specifically designed for rank measures.

The advantages of this Kendall's correlation measure are fourfold. First, since it is based on the rank measures of the pixels, it is invariant to illumination changes. Second, since it characterizes both the distribution of the color components of the pixels and their spatial interaction in the image, it is highly discriminating. Third, this measure can be fast computed from the co-occurrences matrices of the image and the comparison of two indexes is very fast. Fourth, we have shown that 18 real values are almost sufficient to discriminate 1000 objects and provide better results than classical invariant indexes which require much more memory space.

Future works will consist in using a corrected version of the Kendall correlation coefficient which allows to take into account the metamerism, i.e. the fact that some elementary surfaces can appear identical under one illuminant and different under another illuminant. This could be done by accounting the tied pixels for a considered level, which is not the case with the presented version of the Kendall correlation coefficient.

References

[1] G. Buchsbaum. A spatial processor model for object colour perception. *Jour. of the Franklin Institute*, 310:1–26, 1980.

[2] G. Ciocca, D. Marini, A. Rizzi, R. Schettini, and S. Zuffi. On pre-filtering with retinex in color image retrieval. In *Procs. of the SPIE Conf. on Internet Imaging II*, volume 4311, pages 140–147, 2001.

[3] G. Finlayson, S. Hordley, G. Schaefer, and G. Y. Tian. Illuminant and device invariant colour using histogram equalisation. *Pattern Recognition*, 38:179–190, 2005.

[4] G. Finlayson and G. Schaefer. Colour indexing across devices and viewing conditions. In *Procs. of the 2^{nd} Int. Workshop on Content-Based Multimedia Indexing*, pages 215–221, Brescia, Italy, 2001.

[5] G. Finlayson, B. Schiele, and J. Crowley. Comprehensive colour image normalization. *Lecture Notes in Computer Science*, 1406:475–490, 1998.

[6] B. Funt, K. Barnard, and L. Martin. Is machine colour constancy good enough ? In *Procs. of the 5^{th} European Conf. on Computer Vision*, pages 445–459, 1998.

[7] B. Funt and G. Finlayson. Color constant color indexing. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(5):522–529, 1995.

[8] J. Geusebroek. Compact object descriptors from local colour invariant histograms. In *British Machine Vision Conference*, volume 3, pages 1029–1038, 2006.

[9] J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders. The Amsterdam library of object images. *Int. J. Comput. Vision*, 61(1):103–112, 2005.

[10] T. Gevers and A. Smeulders. Pictoseek:combining color and shape invariant features for image retrieval. *IEEE Trans. on Image Processing*, 9(1):102–119, 2000.

[11] R. Haralick. Statistical and structural approaches to textures. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 67(5):786–804, 1979.

[12] M. Kendall. A new measure of rank correlation. *Biometrika*, 30:81–93, 1938.

[13] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[14] D. Muselet and L. Macaire. Combining color and spatial information for object recognition across illumination changes. *Pattern Recognition Letters*, 28 (10):1176–1185, 2007.

[15] C. Palm. Color texture classification by integrative co-occurrence matrices. *Pattern Recognition*, 37:965–976, 2004.

Author Biography

***Damien Muselet**, Engineer (Ecole des Mines de Douai,* *France, 2001), receives the PhD in Automatic Control and Computer Science from the Institute of Technology of the Université des Sciences et Technologies of Lille, France. He is a member of the LIGIV Laboratory (Laboratoire d'Informatique Graphique et d'Ingénierie de la Vision) at the University of Saint Etienne, France. His research interests are object recognition across illumination changes in color images.*

Alain Trémeau *is professor in Color Imaging at the Université Jean Monnet. Alain Trémeau is at the head of LIGIV (http://www.ligiv.org), a research laboratory working on computer Graphics, Vision Engineering and Color Imaging Science. He is currently mainly focused on mathematical imaging and color science with reference to human vision and perception. He works also in color metric with regard to color appearance and rendering measurements. He has written numerous papers or book chapters on Computational Colour Imaging and Processing.*