

Illuminant Change Estimation via Minimization of Color Histogram Divergence

Michela Lecca and Stefano Messelodi

Fondazione Bruno Kessler, IRST - 38100 Povo, Trento - Italy
{lecca, messelod}@fbk.eu

Abstract. We present a new method for computing the change of light possibly occurring between two pictures of the same scene. We approximate the illuminant variation with the von Kries diagonal transform and estimate it by minimizing a functional that measures the divergence between the image color histograms. Our approach shows good performances in terms of accuracy of the illuminant change estimation and of robustness to pixel saturation and Gaussian noise. Moreover we illustrate how the method can be applied to solve the problem of illuminant invariant image recognition.

1 Light and Color

Color descriptors are considered among the most important features in content-based image retrieval and indexing [15]. Colors are in fact robust to noise, rescaling, rotation and image resolution. The main drawback in the use of color for object and image retrieval is the strict dependency of the color on the light in the scene. Color variations can be produced in different ways, for instance by changing the number, the position or the spectrum of the light sources. Moreover, the color of a picture often depends on the characteristics of the device used to capture the scene. The development of a device- and illuminant- invariant image representation is an old but still unsolved attractive problem in Computer Vision [15]. In this paper, we propose a method for estimating the variation of illuminant between the images of a scene taken under different light conditions. More precisely, we restrict our attention to the photometric changes induced by different kinds of lamps or by variations in the voltage of the lamps illuminating a scene. We assume that the illumination varies uniformly over the whole image and we assume the *von Kries diagonal model*, in which the responses of a camera sensor under two different illuminants are related by a diagonal linear transformation. This model has been proved to be a good approximation for illuminant changes [6], [7], especially in the case of narrow-band sensory systems [4], and it is employed in many color enhancement techniques, e.g. [2], [5], [3], [16]. Our technique estimates the von Kries transform between an image and a re-illuminated version of it by a least-squares method that minimizes a dissimilarity measure, named *divergence*, between their color histograms. The accuracy of the estimate obtained by our method has been measured on synthetic and real-world datasets,

showing good performances even in the presence of saturated pixels, Gaussian Noise and the variation of the color quantization. Moreover, we describe how our method can be applied to the illuminant invariant image retrieval task and we compare it with other image retrieval approaches.

Synopsis - Section 2 describes our approach, while its performance is discussed in Section 3, and Section 4 shows us how it can be applied to the illuminant invariant image recognition. Section 5 illustrates our future plans.

2 Diagonal Transform Computation

Let (R_0, G_0, B_0) be the response of a camera to a given illuminant and let (R, G, B) be the response of the same camera to an unknown illuminant. The von Kries diagonal model approximates the change of illuminant mapping (R, G, B) onto (R_0, G_0, B_0) by a diagonal transformation K that rescales each channel independently, i.e. $(R, G, B) \xrightarrow{K} (\alpha_0 R_0, \alpha_1 G_0, \alpha_2 B_0)$, where $\alpha_0, \alpha_1, \alpha_2$ are non-zero positive real numbers, that we refer as *von Kries parameters*.

In our method, the color of an image is described by the distributions of the values of the three channels R, G, B. Each distribution is represented by a histogram of N bins, where N is in the range in $\{1, \dots, 256\}$. Hence, the color feature of an image I is represented by a triplet $\mathbf{H} := (H^0, H^1, H^2)$ of histograms. We refer to \mathbf{H} as *color histograms*, whereas we name its components *channel histograms*.

Let I_0 and I_1 be two images, where I_1 is possibly a rescaled, rotated and differently illuminated version of I_0 . Let \mathbf{H}_0 and \mathbf{H}_1 be the color histograms of I_0 and I_1 respectively. Let H_0^i and H_1^i indicate the i th component of \mathbf{H}_0 and \mathbf{H}_1 respectively. Hereafter, we assume that each channel histogram H_j^i is normalized so that $\sum_{x=1}^N H_0^i(x) = \sum_{x=1}^N H_1^i(x)$.

The channel histograms of two images which differ by illumination are stretched each to other by the diagonal model, hence for each i we have that

$$\sum_{k=1}^x H_1^i(k) = \sum_{k=1}^{\alpha_i x} H_0^i(k), \quad (1)$$

where, as the data is discrete, the value $\alpha_i x$ is cast to an integer in the range $[1, 256]$.

Our estimate of α_i consists of two phases: firstly, for each x in $[1, 256]$ we compute the point y in $[1, 256]$ such that

$$\sum_{k=1}^x H_0^i(k) = \sum_{k=1}^y H_1^i(k). \quad (2)$$

Then we obtain the coefficient α_i as the slope of the *best* line fitting the pairs (x, y) . The *best* line is defined by means of a least squares method. The computation of the pairs (x, y) satisfying (2) is done using the following algorithm consisting of two steps:

Initialization: Let R_0 and R_1 indicate the left- and right- side of (2) respectively. Firstly, we compute the minimum values of x and y such that R_0 and R_1 are greater than zero. Then we set $M := \min(R_0, R_1)$. Let \mathcal{L} be a list of points and let \mathcal{W} be a list of weights, i.e. real numbers, with \mathcal{L} and \mathcal{W} initially empty.

Iterations: Iteratively,

1. we push the pair (x, y) in \mathcal{L} and M in \mathcal{W} ;
2. in order to satisfy equation (2), if M is equal to R_0 (R_1 resp.), i.e. $R_0 < R_1$ ($R_0 > R_1$ resp.), we increment x (y resp.) by one until $M = R_1$ ($M = R_0$ resp.). We update M and then R_0 and R_1 by

$$R_0 := R_0 - M; \quad R_1 := R_1 - M. \quad (3)$$

Note that except for the initialization step, M could be null: in this case, both x and y are incremented until M becomes strictly positive and then R_0 and R_1 are updated as in (3);

3. we repeat steps 1 and 2 until x or y is 255.

We estimate the value of α_i by minimizing with respect to α the following functional, that we call *divergence*:

$$d_\alpha(H_0^i, H_1^i) := \sum_k M_k d((x_k, y_k), \mathcal{A})^2 = \sum_k \frac{M_k}{\alpha^2 + 1} (\alpha x_k - y_k)^2. \quad (4)$$

Here M_k and (x_k, y_k) indicate the k th items of the lists \mathcal{W} and \mathcal{L} respectively, while $d((x_k, y_k), \mathcal{A})$ is the Euclidean distance between the point (x_k, y_k) and the line $\mathcal{A}: y = \alpha x$. The weights M_k are introduced to make the estimate robust to color quantization and possible noise affecting the images.

We observe that: (i) $d_\alpha(H_0^i, H_1^i) = 0 \Leftrightarrow H_0^i(\alpha p) = H_1^i(p)$, for each p in $\{1, \dots, N\}$; and (ii) $d_\alpha(H_0^i, H_1^i) = d_\perp(H_1^i, H_0^i)$.

From these properties follows that d_α is a measure of dissimilarity (*divergence*) between the channel histograms stretched each to other. In particular, if d_α is zero, then the two histograms are related by a stretching of the x axis.

Note that, since the values of R, G, B are in $[1, 256]$, the values of $\alpha_0 R_0$, $\alpha_1 G_0$, $\alpha_2 B_0$ possibly greater than 256 are truncated to 256 (*saturated pixels*). Therefore, to make the estimate robust (as much as possible) with respect to pixel saturation, the N th bin of the histograms H_0^i and H_1^i are not considered when determining the von Kries transform. This explains why in step 3. of the iterative phase the algorithm stops when x or y are 255. However, performances decrease by incrementing the number of saturated pixels (see Section 3).

Figure 1 shows an example, where a same scene has been acquired under two illuminants (a) and (b), while (c) is obtained by remapping (b) onto (a) by our estimated von Kries transform between (a) and (b). As can be seen, (a) and (c) look very similar.

We note that, when no changes of size or in-plane orientation occur, the von Kries map relating two images I and I' can be estimated by finding, for each color channel, the best line fitting the pairs of sensory responses (p_i, p'_i) at the



Fig. 1. (a) a picture and (b) a re-illuminated version of (a); (c) is obtained by remapping (b) onto (a) by the von Kries transform estimated. (a) and (c) appear highly similar.

i th pixels of I and I' respectively. Our approach basically applies a least square method in the space of the color histograms, and therefore makes the estimate of the von Kries coefficients robust to image rescaling and/or rotating.

3 Accuracy of Our Estimate

The accuracy of our estimate of the von Kries diagonal transform has been tested on different synthetic and real-world databases. For brevity, in this paper we report the results obtained with three public databases (TESTS51, ALOI and ECCV98). More details are available in [11].

Let I_0 be an image of a scene under a reference illuminant, and let I be an image of the same scene taken under an unknown illuminant. In the following we refer to I_0 as the *reference image*, while we refer to I as the *test image*. The accuracy of the estimate of the von Kries transform has been evaluated as

$$\mathcal{A} = 1 - L^1(I, K^{est}(I_0)), \quad (5)$$

where $L^1(I, K^{est}(I_0))$ is the L^1 distance computed on the RGB space between I and the transform $K^{est}(I_0)$ of I_0 , and K^{est} indicates the von Kries transform estimated. This distance has been normalized to range in $[0,1]$. Therefore, the closer \mathcal{A} is to 1, the better the estimate of the von Kries transform is. To evaluate the goodness of our image correction, for each pair (I_0, I) , we compared the accuracy measure (5) with the value $\mathcal{A}_0 = 1 - L^1(I, I_0)$.

In these experiments we do not consider changes of image size or orientation. The difference between the accuracy measure (5) of the best fit applied on the channel images (mentioned at the end of Section 2) and those obtained by means of our approach is negligible: about $3 \cdot 10^{-4}$ in the worst case (ECCV98).

Our algorithm has linear complexity with respect to the number of image pixels and to the color quantization N . Therefore, it is particularly efficient, also in comparison to other methods, like for instance [3]. The time for the estimation of the von Kries coefficients for a pair of images of size 150×200 is less than 40 ms on a standard Pentium4 CPU 2.8 GHz.

Tests on TESTS51 - The dataset TESTS51 has been built starting from the public dataset of Ponce and others [14] (<http://www-cvr.ai.uiuc.edu/>).

This database consists of a set of images of 8 different objects and of a set of 51 test-pictures in which the objects appear under different conditions (occluded, rescaled, rotated, differently illuminated, ...).

The 51 test-pictures have been taken as references, while the test images have been obtained by rescaling the color channels of the reference images by 20 diagonal linear functions of the form $F_{\beta_w}(R, G, B) = \beta_w(R, G, B)$, with $\beta_w = 0.2 + 0.2w$, and $w = 0, \dots, 19$. For each test image I , we estimated the 20 von Kries transforms K_w^{est} mapping the correspondent reference, estimated on I . Figure 2(left) shows the mean value of the accuracy measure (5) versus the parameter β_w , for different color quantization. The mean value of \mathcal{A}_0 is 0.22, while the mean value of \mathcal{A} is 0.9999 for $N = 256$.

The precision of our estimates β_w^{est} of β_w , $w = 0, \dots, 19$, has been measured using the error $E_w = 1 - \frac{\beta_w^{est}}{\beta_w}$. The closer E_w is to zero, the better is the accuracy on the determination of the von Kries transform. A strictly negative (positive, resp.) value of E_w indicates that the estimate is greater (smaller, resp.) than the real parameter. Figure 2(right) shows the mean value \overline{E}_w of E_w averaged over the test images, by varying β_w and the color quantization.

The best parameter estimates and the best accuracy have been obtained for the finest color quantization.

Tests on ALOI - ALOI [9] (<http://staff.science.uva.nl/~aloi/>) is a col-

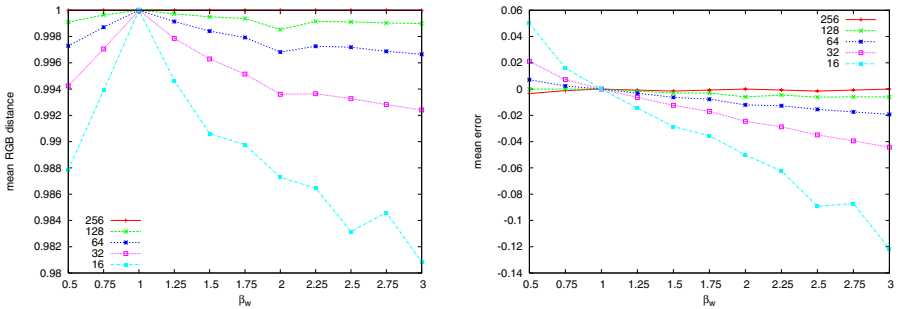


Fig. 2. TESTS51: (Left) Mean value of the accuracy measure (5) by varying the color quantization. (Right) Mean Error \overline{E}_w by varying the parameter β_w and the color quantization.

lection of 110,250 images of 1,000 objects acquired under different conditions. In ALOI, each frontal object view has been shot under 12 different light conditions, produced by varying the color temperature of five lamps illuminating the scene. More precisely, the lamp voltage was modified to be $V_j = j \times 0.047$ Volts with $j \in J = \{110, 120, 130, 140, 150, 160, 170, 180, 190, 230, 250\}$. The object images captured under the illuminant with voltage V_{110} have been taken as references, while the other object images have been used for testing. For each reference O_{110} , we estimate the von Kries transform $K_j^{est}(O_j)$ mapping O_{110} onto the test image O_j taken under the illuminant with voltage V_j , and $j \in J - \{110\}$.

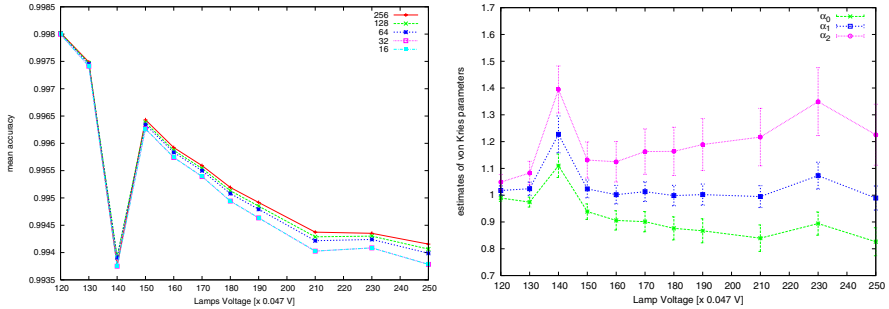


Fig. 3. ALOI: (Left) Mean accuracy (5) versus the illuminants for different color quantizations. (Right) Estimates of the von Kries parameters and their standard deviation bars for $N = 256$.

Figure 3(left) shows the mean accuracy (5) for the different lamp voltages and variations of color quantization. On average, $\mathcal{A}_0 = 0.9913$, while $\mathcal{A} = 0.9961$ for $N = 256$. We note that for $j = 140$, the accuracy is lower than for the other lamp voltages. This is because the voltage V_{140} determines a large increment of the light intensity and therefore produces a large number of saturated pixels.

In principle, the transform K mapping the object image O_{110} onto O_j should have the same parameters of the transform K' mapping a different object image O'_{110} onto O'_j , because the illuminant change is the same. In practice, since the von Kries model is only an approximation of the illuminant variation phenomenon, the parameters of K and K' differ. Hence we measure the robustness on the determination of the coefficients α_i , $i = 0, 1, 2$, by analyzing the standard deviation of their estimates. Figure 3(right) reports the averages and standard deviations of the von Kries parameters estimated with color quantization of 256. Deviations increase when the image brightness is increased, i.e. the number of saturated pixels becomes larger.

Tests on ECCV98 - Here we consider a subset of the database [8] consisting of the images of 11 objects captured under 5 different illuminants (halogen, mb-5000, mb-5000+3202, syl-cwf, ph-uhl). We refer to this subset as ECCV98. We took the object images captured under the illuminant halogen as references. This data is available at <http://www.cs.sfu.ca>.

The mean accuracy for each illuminant and for different color quantizations are reported in Table 1. On average, \mathcal{A}_0 is 0.9311, while $\mathcal{A} = 0.9734$. Table 2 shows the mean values of the von Kries coefficients and their standard deviations, as for ALOI.

4 Application to Image Recognition

Let us consider a set of known images (*references*) and let I be an unknown image (*query*). The illuminant invariant image recognition consists of finding the reference I_0 that, although re-illuminated, is the most similar to the query.

Table 1. ECCV98: Mean accuracy (5) by varying the illuminant and the color quantization

Illuminant	256	128	64	32	16
mb-5000	0.9767	0.9766	0.9763	0.9755	0.9723
mb-5000+3202	0.9719	0.9716	0.9702	0.9638	0.9308
ph-ulm	0.9733	0.9733	0.9732	0.9731	0.9726
syl-cwf	0.9718	0.9717	0.9717	0.9716	0.9710

Table 2. ECCV98: Values of the von Kries parameters and their errors for $N = 256$ bins

Illuminant	$\alpha_0 \pm \Delta\alpha_0$	$\alpha_1 \pm \Delta\alpha_1$	$\alpha_2 \pm \Delta\alpha_2$
mb-5000	0.4601 ± 0.1206	0.8255 ± 0.1949	1.6004 ± 0.3787
mb-5000+3202	0.1915 ± 0.0406	0.5287 ± 0.0835	2.0121 ± 0.3177
ph-ulm	0.7405 ± 0.1059	1.0792 ± 0.1675	1.1068 ± 0.1892
syl-cwf	0.8596 ± 0.1357	0.9581 ± 0.1291	1.6924 ± 0.2772

Our solution is outlined as follows: we compute the von Kries transforms mapping each reference onto the query and we associate a dissimilarity score to each of these transforms. The solution I_0 is the image reference whose von Kries transform $T(I_0)$ has the minimum score from I .

More precisely, let \mathbf{H} be the color histogram of I . For each reference I_r of D with color histogram \mathbf{H}_r , (i) we estimate the parameters α_0 , α_1 and α_2 of the von Kries transform K mapping I_r onto I ; (ii) for each i we compute the divergence $d_{\alpha_i}(H^i, H_r^i)$ defined in (4) and the dissimilarity score

$$\delta = \sum_i d_{\alpha_i}(H^i, H_r^i). \quad (6)$$

Thus, the solution of the image recognition problem is the image I_0 of D such that the score (6) is minimized.

Due to the dependency of the divergence (4) on the values α_i ($i = 0, 1, 2$), δ does not satisfy the triangular inequality, and thus it is not a distance. Nevertheless, it is a *query-sensitive* dissimilarity measure, in the sense that it depends on the query [1]. The use of the score (6) instead of a L^p metric ($p \geq 1$) between the color histograms is justified by the major robustness of (6) to color quantization.

We say that a query I is *correctly* recognized if the reference image I_r of D minimizing (6) is a re-illuminated version of I . The performance of our approach has been evaluated using a *recognition rate*, defined as the ratio between the number of test images correctly recognized and the total number of test images. In our experiments, we considered the reference and the test sets of TESTS51 and ALOI defined in Section 3, while we excluded the test set ECCV98 because

it contains only a few images and hence it is inadequate for testing retrieval performances. By comparing the reference and the test images by using the score (6) without estimating the von Kries map or enhancing the color, we obtained the following mean recognition rates: 0.20 for TESTS51, 0.77 for ALOI.

The recognition rates obtained by means of our approach on TESTS51 and ALOI are shown in Figure 4.

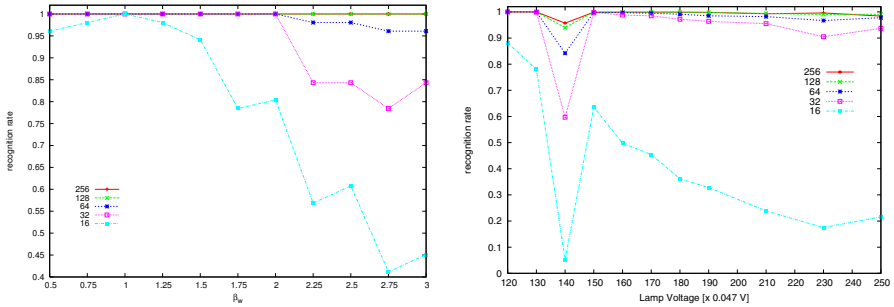


Fig. 4. Robustness of the recognition rate with respect to different color histogram quantizations for the datasets TESTS51 (left) and ALOI (right)

In the case of ALOI, we compared our results with the recognition rates obtained by two other approaches, employing the color normalization algorithms Gray World and ACE [13] (Automatic Color Equalization) respectively. The Gray World algorithm used in these tests performs the color enhancement by rescaling each channel by their mean value. The technique ACE relies on the Retinex theory [12], and combines the Gray World enhancement with a color constancy algorithm that separates the spectral distribution of the scene illuminant from the image brightness. We chose these color balancing techniques among the others because they are very popular.

In our experiments, each reference as well as each query has been normalized by ACE and by Gray-World; then their color histograms have been compared by means of the score in (6). Figure 5 (left) shows the recognition rates obtained. The performances of our approach are very similar to those given by employing the Gray World color balancing, whereas ACE gave the worst results both in terms of recognition rate and computational run time. In fact the complexity of ACE is $O(n^2)$, while that of our approach and Gray World is $O(n)$, where n is the number of image pixels.

Our approach strongly differs from the recognition methods in which the illuminant invariance is achieved by enhancing the colors of the reference and test images [8]. In these methods, color enhancement is obtained by means of a color constancy algorithm, that illuminates each input image as if it had seen under a canonical known illuminant [2]. In the image/object recognition framework, each reference in D and every query are firstly illuminated under the canonical

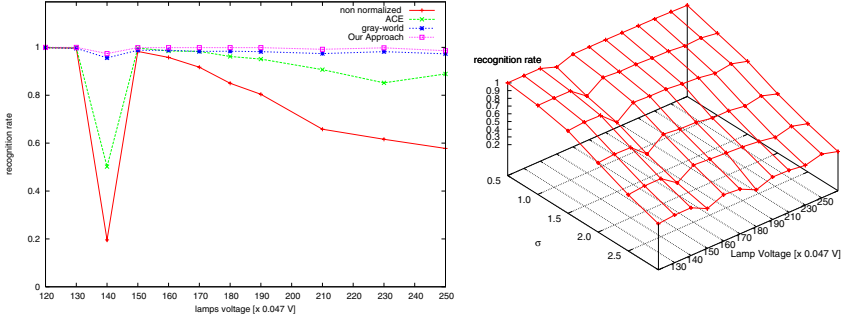


Fig. 5. ALOI: Left - Comparison among different methods for illuminant invariant image retrieval. The line labeled by *non normalized* shows the recognition rate when no color balancing is applied. Right - Recognition rate in presence of Gaussian Noise.

illuminant by a color constancy algorithm, then their re-illuminated versions are compared. If the references are captured using the canonical illuminant, the von Kries coefficients are computed as the element-wise ratios between the canonical illuminant and the estimate of the illuminant of the query.

The color enhancing of references and query are completely avoided in our approach, as we directly compare the color histograms of the query to each reference through the query-sensitive score (6). Therefore, our method is more efficient than the recognition procedures based on color normalization, because it does not require any color pre-processing of the references in the database and of the query.

Finally, we tested our recognition performances when Gaussian noise was added to the pictures. In particular, the test images of ALOI have been modified by convolving each image by a Gaussian filter with standard deviation $\sigma = 0.5, 1.0, 1.5, 2.0, 2.5, 3.0$. Figure 5(right) shows the results achieved: the recognition rate decreases with increasing levels of noise, in particular, for σ greater than 2.0 it is smaller than 0.75.

5 Conclusions and Future Directions

Our estimate of illuminant change performed impressively on the synthetic and real data, and offers high accuracy in illuminant invariant image recognition. Moreover, unlike the pixel-wise best fit mentioned in Section 2, our use of color histograms allows us to estimate the illuminant changes occurring between two images also when they have different size and orientation. Our future plans include a comparison of our approach with other methods for the estimation of the illuminant variation and its integration in the object recognizer MEMORI [10] to make this system robust to changes of light.

References

1. Athitsos, V., Hadjieleftheriou, M., Kollios, G., Sclaroff, S.: Query-sensitive embeddings. In: Proc. of SIGMOD 2005, pp. 706–717. ACM Press, New York (2005)
2. Barnard, K., Cardei, V., Funt, B.: A comparison of computational color constancy algorithms. ii: Experiments with image data. *IEEE Transactions on Image Processing* 11(9), 985–996 (2002)
3. Berens, J., Finlayson, G.: Log-opponent chromaticity coding of colour space. In: 15th IEEE Int. Conf. on Pattern Recognition, pp. 206–211 (2000)
4. Chong, H.Y., Gortler, S.J., Zickler, T.: The von kries hypothesis and a basis for color constancy. In: Proc. of IEEE ICCV (October 2007)
5. Finlayson, C., Hordley, S., Schaefer, G., Tian, G.Y.: Illuminant and device invariance using histogram equalisation. In: IS&T and SID's 11th Color Imaging Conference, pp. 205–211 (2003)
6. Finlayson, G.D., Drew, M.S., Funt, B.V.: Diagonal transforms suffice for color constancy. In: Proc. of International Conference of Computer Vision (1993)
7. Finlayson, G.D., Drew, M.S., Funt, B.V.: Color constancy: generalized diagonal transforms suffice. *J. Optical Society of America* 11(11), 3011–3019 (1994)
8. Funt, B.V., Barnard, K., Martin, L.: Is machine colour constancy good enough? In: Burkhardt, H.-J., Neumann, B. (eds.) ECCV 1998. LNCS, vol. 1406, pp. 445–459. Springer, Heidelberg (1998)
9. Geusebroek, J.M., Burghouts, G.J., Smeulders, A.W.M.: The Amsterdam library of object images. *International Journal of Computer Vision* 61(1), 103–112 (2005)
10. Lecca, M.: Object recognition in color images by the self configuring system MEM-ORI. *International Journal of Signal Processing* 3(3), 176–185 (2006)
11. Lecca, M., Messelodi, S.: Estimating illuminant changes in color images by color histogram comparison. Technical Report FBK - irst, TR 2008-04-001 (April 2008)
12. Provenzi, E., De Carli, L., Rizzi, A.: Mathematical definition and analysis of the retinex algorithm. *Journal of the Optical Society of America. Optics, Image Science, and Vision* 22(12) (2005)
13. Rizzi, A., Gatta, C., Marini, D.: From Retinex to ACE: issues in developing a new algorithm for unsupervised color equalization. *Journal of Electronic Imaging* 13(1) (2004)
14. Rothganger, F., Lazebnik, S., Schmid, C., Ponce, J.: 3D object modeling and recognition using local affine-invariant image descriptors and multi-view spatial constraints. *International Journal of Computer Vision* 66(3), 231–259 (2006)
15. Schettini, R., Ciocca, G., Zuffi, S.: A survey of methods for colour image indexing and retrieval in image databases. In: *Color Imaging Science: Exploiting Digital Media* (2001)
16. Swain, M.J., Ballard, D.H.: Color indexing. *Int. J. Comput. Vision* 7(1), 11–32 (1991)