



PERGAMON

Available at  
[www.ElsevierComputerScience.com](http://www.ElsevierComputerScience.com)  
POWERED BY SCIENCE @ DIRECT®

Pattern Recognition 37 (2004) 1201–1217

**PATTERN  
RECOGNITION**

THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

[www.elsevier.com/locate/patcog](http://www.elsevier.com/locate/patcog)

# Color balancing of digital photos using simple image statistics

Francesca Gasparini, Raimondo Schettini\*

*DISCO (Dipartimento di Informatica, Sistemistica e Comunicazione), Università degli Studi di Milano-Bicocca, Edificio 7,  
Via Bicocca degli Arcimboldi 8, Milano 20126, Italy*

Received 16 May 2003; accepted 19 December 2003

## Abstract

The great diffusion of digital cameras and the widespread use of the internet have produced a mass of digital images depicting a huge variety of subjects, generally acquired by unknown imaging systems under unknown lighting conditions. This makes color balancing, recovery of the color characteristics of the original scene, increasingly difficult. In this paper, we describe a method for detecting and removing a color cast (i.e. a superimposed color due to lighting conditions, or to the characteristics of the capturing device), from a digital photo without any a priori knowledge of its semantic content. First a cast detector, using simple image statistics, classifies the input images as presenting no cast, evident cast, ambiguous cast, a predominant color that must be preserved (such as in underwater images or single color close-ups) or as unclassifiable. A cast remover, a modified version of the white balance algorithm, is then applied in cases of evident or ambiguous cast. The method we propose has been tested with positive results on a data set of some 750 photos.

© 2003 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.

*Keywords:* Color constancy; Cast detection; Cast removal; Von Kries; White balance; Gray world algorithm

## 1. Introduction

The great diffusion of digital cameras and the widespread use of the internet have produced a mass of digital images depicting a huge variety of subjects, generally acquired by non-professional photographers using unknown imaging systems under unknown lighting conditions. The quality of these real-world photographs can often be considerably improved by digital image processing. At present color and contrast corrections are usually manually performed within the framework of specific software packages. Since these interactive processes may prove difficult and tedious, especially for amateur users, an automatic image enhancement tool would be most desirable.

Typical image properties requiring correction are color, contrast and sharpness. We have approached the open issues

of designing reliable automatic tools that can improve the overall quality of digital photographs pragmatically by designing a modular enhancing procedure (Fig. 1). Each module can be considered as an autonomous element, that can be combined in more or less complex algorithms.

This paper focuses on unsupervised color correction. The problem is that of automatically and reliably removing a color cast (a superimposed color due to lighting conditions, or to the characteristics of the capturing device). The solution we have designed exploits only simple image statistics to drive the color procedure, preventing artifacts.

Section 2 illustrates the problem addressed, and the related methods available in the literature. In Section 3, we describe the automatic procedure for color balancing, which we have structured in two main parts: a cast detector and a cast remover. The detector, described in Section 3.1, classifies the input images with respect to their chromaticity as: (i) no-cast images, (ii) evident cast images, (iii) ambiguous cast images (images with a weak cast, or for which whether or not a cast exists is a subjective opinion), (iv) images with a predominant color which must be preserved (such as in underwater images, or single color close-ups), (v) unclassifiable.

\* Corresponding author. Tel.: +39-02-6448-7840; fax: +39-02-6448-7839.

E-mail addresses: [gasparini@disco.unimib.it](mailto:gasparini@disco.unimib.it) (F. Gasparini), [schettini@disco.unimib.it](mailto:schettini@disco.unimib.it) (R. Schettini).

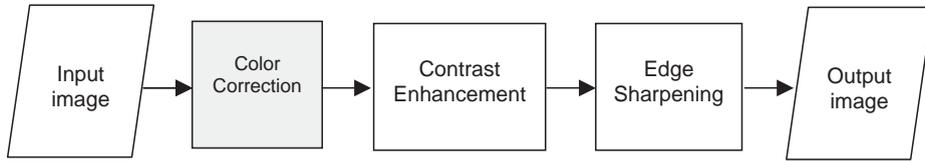


Fig. 1. Our enhancement chain constituted by independent enhancement modules.

The remover, a modified version of the white balance algorithm [1,2], piloted by the class of the cast, is applied to images labeled as having evident or ambiguous cast. This step is described in Section 3.2. Our experimental results are reported and commented in Section 4, while our conclusions are presented in Section 5.

## 2. Related work

Automatic color correction is a challenging issue as the RGB images recorded by an imaging device depend on the following elements:

1. *The response of each color channel as a function of intensity*, usually modeled as a gamma correction. For generic images the channel responses of the imaging device are usually unknown. Cardei et al. [2,3] have shown that gamma-on images (i.e. images for which gamma does not equal unity) can be color corrected without first linearizing them. Consequently, it is common practice to address color and gamma corrections independently.
2. *The device white balancing*, usually set by the device manufacturer so that it produces equal RGB values for a white patch under some chosen illuminant. Images of unknown color balance can still be corrected, but this poses an additional challenge to the color correction algorithm.
3. *The relative spectral sensitivities of the capturing device as a function of wavelength*. This information is often not available (e.g. for images downloaded from the Internet), and sensitivities can differ significantly for different cameras. As Cardei et al. [2,3] have pointed out, while two different camera models balanced for the same illuminant will have by definition the same response to white, they may have different responses to other colors.
4. *The surface properties of the objects present in the scene depicted and the lighting conditions* (lighting geometry and illuminant color). Unlike human vision, imaging devices (such as digital cameras) can not adapt their spectral responses to cope with different lighting conditions. As a result, the acquired image may have a cast, i.e. an undesirable shift in the entire color range. The ability of the human visual system to discount the illuminant, rendering the perceived colors of objects almost independent of illumination is called *color constancy*. This would be a useful property in any vision system performing tasks

that require a stable perception of an object's color, as in object recognition, image retrieval, or image reproduction, e.g. Ref. [4].

Color constancy is an under-determined problem and thus impossible to solve in the most general case [1]. Several strategies are proposed in the literature. These, in general, require some information about the camera being used, and are based on assumptions about the statistical properties of the expected illuminants and surface reflectances. From a computational perspective, color constancy is a two-stage process: the illuminant is estimated, and the image colors are then corrected on the basis of this estimate. The correction generates a new image of the scene as if taken under a known standard illuminant. The color correction step is usually based on a diagonal model of illumination change, deriving from the Von Kries hypothesis that color constancy is an independent gain regulation of the three cone signals, through three different gain coefficients [5]. This diagonal model is generally a good approximation of change in illumination, as demonstrated by Finlayson et al. in Ref. [6]. Should the model lead to large errors, its performance can still be improved with sensor sharpening [7,8]. In formula the Von Kries hypothesis can be written as

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = \begin{bmatrix} k_L & 0 & 0 \\ 0 & k_M & 0 \\ 0 & 0 & k_S \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}, \quad (1)$$

where  $L$ ,  $M$ ,  $S$  and  $L'$ ,  $M'$ ,  $S'$  are the initial and post-adaptation cone signals and  $k_{L,M,S}$  are the scaling coefficients [5]. The scaling coefficients can be expressed as the ratio between the cone responses to a white under the reference illuminant and those of the current one. A typical reference illuminant, which is also the one we have used here, is the D65 CIE standard illuminant [9]. In practical situations the  $L$ ,  $M$ ,  $S$  retinal wavebands are transformed into CIE XYZ tristimulus values by a linear transformation, or approximated by image RGB values [12].

Whatever the features used to describe the colors, we must have some criteria for estimating the illuminant and thus the scaling coefficients in Eq. (1). The gray world algorithm assumes that, given an image of sufficiently varied colors, the average surface color in a scene is gray [10]. This means that the shift from gray of the measured averages on the three channels corresponds to the color of the illuminant.

The three scaling coefficients in Eq. (1) are therefore set to compensate this shift.

The white balance algorithm, instead, looks for a white patch in the image, the chromaticity of which will then be the chromaticity of the illuminant. The white patch is evaluated as the maximum found in each of the three image bands separately [1,2]. The scaling coefficients are now obtained comparing these maxima with the values of the three channels of the chosen reference white.

Retinex algorithms, [11,12], try to simulate the adaptation mechanisms of the human vision, performing both color constancy and dynamic range enhancement. Although related to the white patch algorithms, they are not a simple diagonal model, but also take into account the spatial relationships in the scene.

Still based on the diagonal model, the gamut mapping approach [13] determines the set of all possible RGB values of world surfaces under a known *canonical* illuminant. This set is represented as a convex hull in the color space. The set of all the possible RGB values consistent with the image data under an unknown illuminant, forms another convex hull. The object is to determine the diagonal mapping of these two hulls. However, there is no unique solution to this problem, and, in particular, not all the mappings found correspond to real world illuminants. Finlayson reformulated the problem in a two-dimensional chromatic space ( $r = R/B$ ,  $g = G/B$ ) where the diagonal model is still valid, then argued that diagonal maps can be further constrained by considering only those corresponding to expected illuminants [14–17]. Once the set of possible maps had been computed, several different models for finding a unique solution were proposed in Refs. [18,19].

Color by correlation has been introduced by Finlayson et al. as a successive improvement of the gamut mapping [20–22]. The basic idea is to pre-compute a correlation matrix which describes the extent to which the proposed illuminants are compatible with the occurrence of image chromaticities. In a further refinement, the correlation matrix has been set up to compute the probability that the observed chromaticities are due to each of the training illuminants. The best illuminant can then be chosen, using a maximum likelihood estimate for example, or other methods described in Ref. [18]. Funt et al. [23,24] have achieved a good performance in color constancy using a neural network. The network estimates the illuminant chromaticity based on the gamut present in the image: the input layer is binary, indicating the presence, or absence in the image of a sampled chromaticity; the output is the estimated chromaticity of the illuminant.

As we have said, color correcting images of unknown origin adds to the complexities of the already difficult problem of color constancy, because the pre-processing the image was subjected to, the camera sensors and camera balance are all unknown [1]. Digital still cameras usually employ automatic white balance techniques, based on some illumination estimation assumption, to adjust sensor amplifier gains

in order to produce an image where the white objects appear white. Moreover, an ad hoc adjustment is commonly made to force an image's maximum luminance value to pure white. Some cameras also adjust the minimum luminance to set the black point at a certain energy level. The movement of the white and black point regions of the image towards pre-defined goals typically produces pleasing peak highlights and shadows. However, the resulting images may show significant color casts in over half the image. Since the image white point is altered in a very non-linear way during image acquisition, the camera data available for post-processing retains only limited information about the illuminant [25]. For digital still cameras with an unknown white balance correction in which the brightest pixels may have been compromised, Cooper et al. developed an algorithm that analyzes the chromaticity of any large contiguous nearly gray objects found using adaptive segmentation techniques, to identify the presence of a cast, estimate the chromatic strength of the objects, and alter the image's colors to compensate for the cast [25,26]. Unsupervised segmentation is, in turn, another ill-posed problem.

### 3. Our color correction strategy

The computational strategy described here should not be interpreted as a "new color constancy algorithm"; it is instead a tool that, given as input any digital photo, produces as output a more pleasing image, that is an image the user will perceive as more natural than the original one. The success of the method is evaluated by direct comparison of the original and the processed images.

The procedure is structured in two main parts: a cast detector and a cast remover. The basic idea of our cast detector (Section 3.1) is that by analyzing the color distribution of the image in a suitable color space with simple statistical tools, it is possible not only to evaluate whether or not a cast is present, but also to classify it. The classes considered are: (i) no cast images; (ii) evident cast images; (iii) ambiguous cast images; (iv) images with a predominant color that must be preserved; and (v) unclassifiable. The estimate of the sensor scaling coefficients is assimilated within the problem of quantifying the cast. The cast remover (Section 3.2) is applied only to those images classified as having an evident or ambiguous color cast.

#### 3.1. Cast detection and classification

The algorithm is structured as follows:

As seen in Section 2, RGB images are often referred to unknown imaging devices. We assume here that the images are coded in terms of sRGB color coordinates. The sRGB is representative of the majority of devices on which color is and will be viewed [27,28]; it refers to a monitor of average characteristics, so that the images can be directly displayed by remote systems without further manipulation.

The sRGB values are mapped into the CIELAB color space [30] (see Appendix A), where the chromatic components and lightness are separated [20,25]. The CIELAB is a perceptually uniform color space in which it is possible to effectively quantify color differences as seen by the human eye; it is widely used in color imaging and is included as a standard in the international color consortium (ICC) color profiles [28].

- (i) We analyze only those pixels with a lightness in an interval that excludes the brightest and the darkest points. This because the images we consider may already have been processed during acquisition, and we assume that imaging device is unknown. Digital cameras often force the brightest image point to white and the darkest to black, altering the chroma of very light and very dark regions. Our experience on a data set of several hundred images has suggested that we consider the interval of lightness:  $30 < L^* < 95$  in identifying color cast. If the size of the considered portion of image is less than the 20% of the whole, the image statistics are not fully reliable. These images are considered unclassifiable and are not processed at all. This is the case of very dark or very light images, such as those shown in Fig. 2. Otherwise, the algorithm proceeds with the next step.
- (ii) The two-dimensional histogram,  $F(a, b)$ , of the image colors in the  $ab$ -plane is computed. For a multicolor image without cast it will present several peaks, distributed over the whole  $ab$ -plane, while for a single color image, there will be a single peak, or a few peaks in a limited region (see Fig. 3). The more concentrated the histogram and the farther from the neutral axis, the more intense the cast.

The color distribution is modeled using the following statistical measures, with  $k = a, b$ :

$$\mu_k = \int_k kF(a, b) dk, \quad (2)$$

$$\sigma_k^2 = \int_k (\mu_k - k)^2 F(a, b) dk, \quad (3)$$

respectively, the mean values and the variances of the histogram projections along the two chromatic axes  $a$ , and  $b$ .

- (iii) An *equivalent circle (EC)* with center:  $C = (\mu_a, \mu_b)$  and radius:  $\sigma = \sqrt{\sigma_a^2 + \sigma_b^2}$  is associated to each histogram. To characterize the *EC* quantitatively we introduce a distance  $D$ :

$$D = \mu - \sigma \quad (4)$$

(where  $\mu = \sqrt{(\mu_a^2 + \mu_b^2)}$ ), and the ratio:

$$D_\sigma = D/\sigma. \quad (5)$$

Because  $D$  is a measure of how far the whole histogram (identified by its *EC*) lies from the neutral axis ( $a = 0$ ,  $b = 0$ ), while  $\sigma$  is a measure of how the histogram

is spread,  $D_\sigma$  makes it possible to quantify the strength of the cast.

The algorithm analyzes the color histogram distribution in the  $ab$  chromatic plane, examining its *EC* and computing the statistical values  $D$  and  $D_\sigma$ .

1. If the histogram is concentrated and far from the neutral axis, the colors of the image are thus confined to a small region in the  $ab$  chromatic diagram (Fig. 4). The images are, instead, likely to have either an evident cast (to be removed), or a predominant color (to be preserved), if:

$$(D > 10 \text{ and } D_\sigma > 0.6) \text{ or } (D_\sigma > 1.5). \quad (6)$$

A predominant color could correspond to an intrinsic cast (widespread areas of vegetation, skin, sky, or sea), or to a single color close-up (Fig. 5).

To detect images with a predominant color corresponding to an intrinsic cast and a single color close-up, a simple classifier exploiting both color and spatial information is used [31]. A region identified as probably corresponding to skin, sky, sea, or vegetation is considered significant if it covers over 40% of the whole image; the image is classified as having an intrinsic cast, and the cast remover is not applied.

If none of the regions corresponding to skin, sky, sea, or vegetation occupies over 40% of the whole, but the image *EC* is extremely concentrated,  $D_\sigma > 6$ , and has an high average color saturation, ( $C^*/L^* > 1$ —the ratio between the chroma radius and the lightness is correlated to the color saturation), the image is classified as a single color close-up and, also in this case, the cast remover is not applied.

Images presenting a concentrated histogram which are not classified as having a predominant color, i.e. intrinsic cast images or close-ups, are held to have an evident cast and are processed for color correction as described in Section 3.2.

2. All images without a clearly concentrated color histogram are analyzed with a procedure based on the criterion that a cast has a greater influence on a neutral region than on objects with colors of high chroma. The color distribution of near neutral objects (NNO) is studied with the same statistical tools described above. A pixel of the image belongs to the NNO region if its chroma is less than an initial fixed value (set here at one fourth the maximum chroma radius of that image), and if it is not isolated, but has some neighbors that present similar chromatic characteristics, than it belongs to the NNO region. Isolated nearly gray pixels are probably due to noise. If the percentage of pixels that satisfies these requisites is less than a predefined percentage of the whole image, which experience has suggested to set at 5%, the radius of the neutral region



Fig. 2. Examples of unclassifiable images (typically very dark or very light images in which the color range evaluated is not sufficient to determine whether or not a cast exists).

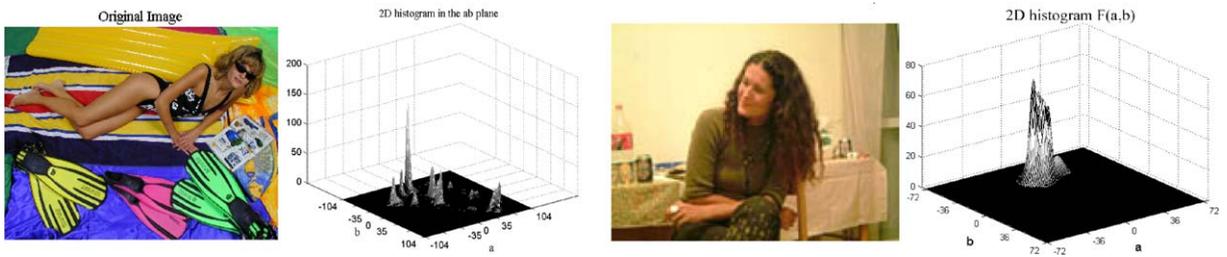


Fig. 3. Left, a multicolor image with no cast, which presents several peaks in the 2D histogram in the  $ab$ -plane; right, a cast image showing a concentrated 2D histogram.

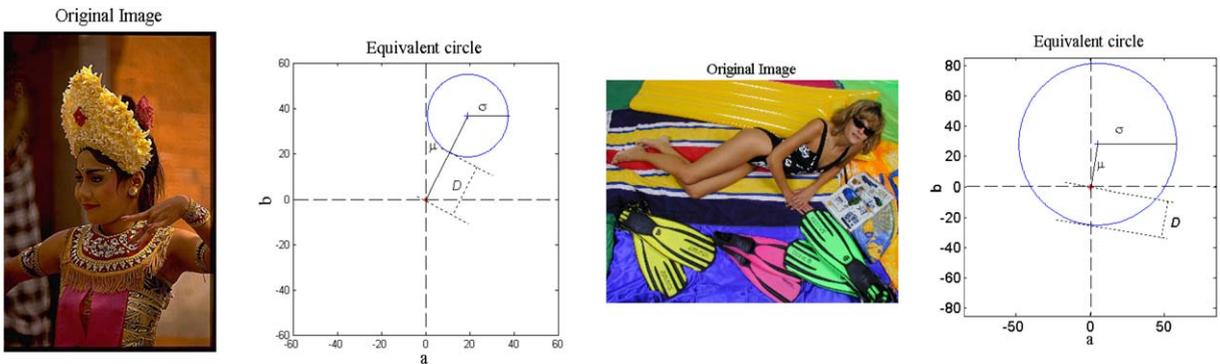


Fig. 4. Left, equivalent circles corresponding to a cast image; right, equivalent circles for a multicolor image with no cast. The parameters  $D$  and  $\sigma$  permit the characterization of the histogram distribution.

is gradually increased to the maximum chroma radius, and the region recursively evaluated.

When there is a cast, we may expect the NNO color histogram to show a single small spot, or, at most, a few spots in a limited region. NNO colors spread around the neutral axis, or distributed in several distinct spots indicate instead that there is no cast. The statistical analysis that we have performed on the whole histogram is now applied to the NNO histogram, allowing us to distinguish among three cases:

- evident cast images;
- no cast images;

- ambiguous cases (images with a weak cast, or for which the presence of a cast is a subjective opinion).

If we define, as above, the mean values and variances of the NNO histogram, with  $k = a, b$ :

$$\mu_{NNOk} = \int_k k F_{NNO}(a, b) dk, \tag{7}$$

$$\sigma_{NNOk}^2 = \int_a (\mu_{NNOk} - k)^2 F_{NNO}(a, b) dk. \tag{8}$$

we can associate with each NNO histogram a *NNO equivalent circle* ( $EC_{NNO}$ ) with center in  $C_{NNO} = (\mu_{NNOa}, \mu_{NNOb})$  and a radius of  $\sigma_{NNO} = \sqrt{\sigma_{NNOa}^2 + \sigma_{NNOb}^2}$ .

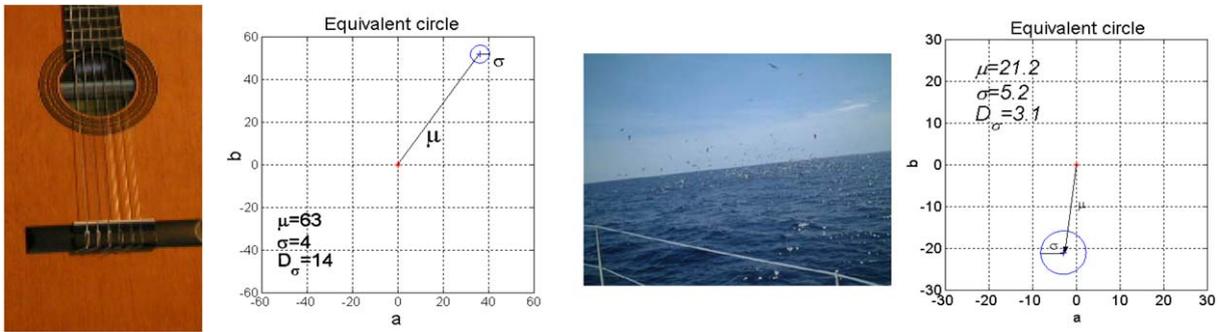


Fig. 5. Images with a predominant color and their equivalent circles. Left, single color close-up; right, intrinsic cast.

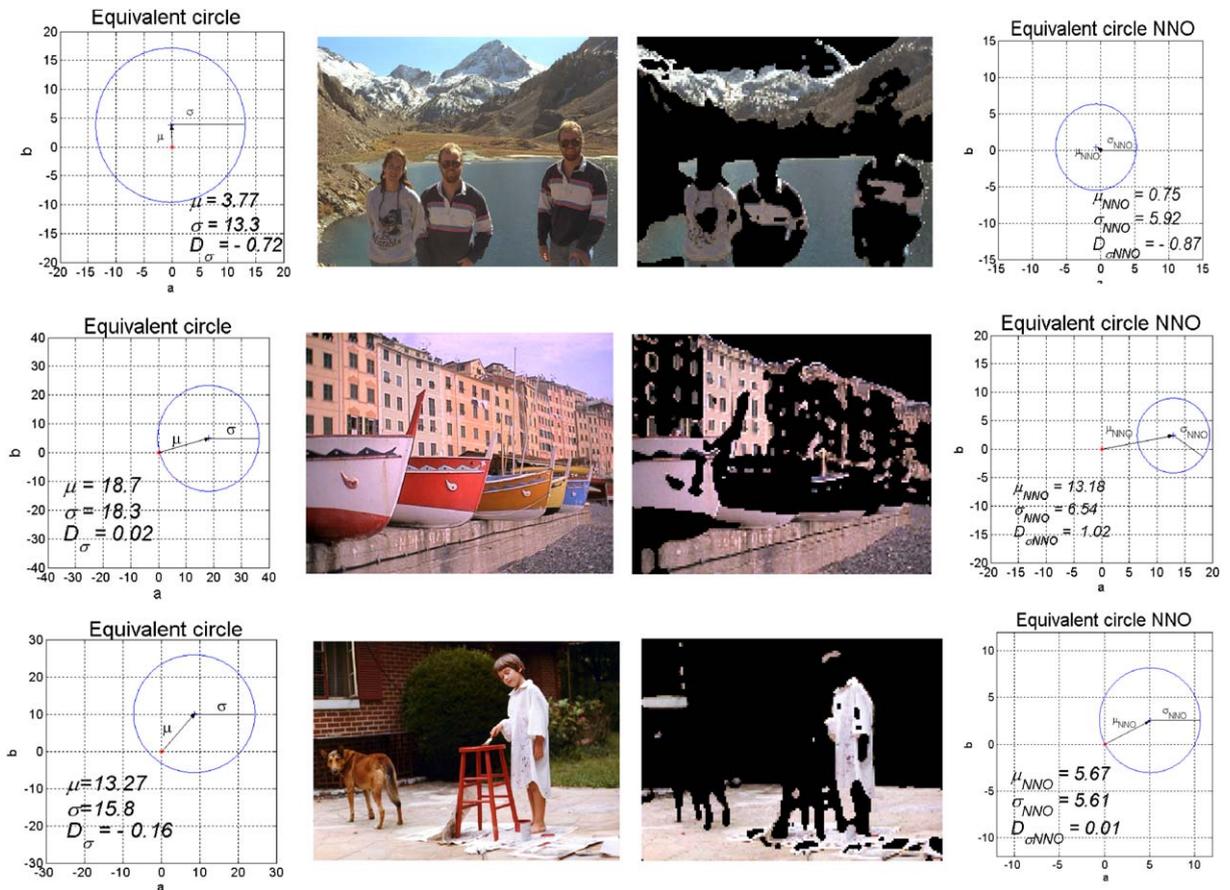


Fig. 6. If analysis of the whole image does not indicate whether or not there is a cast, the NNO must be analyzed. The three possible cases are represented in the three rows of this figure, each showing the original image and the NNO region, together with the corresponding ECs and ECNNOs. The top row shows the case of a no-cast image:  $D_{\sigma NNO} = -0.87$  and the NNO histogram is clearly spread uniformly around the neutral axis. The middle row corresponds to an image with evident cast:  $D_{\sigma NNO} = 1.02$  and shows a clear shift histogram. The bottom row presents the case of an ambiguous cast: it is not clear whether there is a cast, and any evaluation could be a subjective opinion:  $D_{\sigma NNO} = 0.01$ .

NNO histograms that are well defined and concentrated far from the neutral axis will have a positive  $D_{\sigma NNO}$ : the image has an evident cast (Fig. 6). With negative values of  $D_{\sigma NNO}$  the histogram presents an almost uniformly spread

distribution around the neutral axis, indicating no cast. We set as thresholds  $D_{\sigma NNO} = 0.5$  for cast images and  $D_{\sigma NNO} = -0.5$  for no cast images. Cases falling in the interval between  $-0.5$  and  $0.5$  are considered ambiguous.

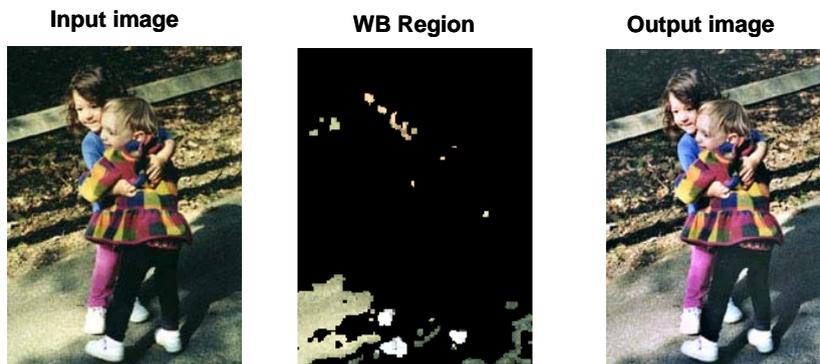


Fig. 7. The WB region and color correction for a cast image.



Fig. 8. The WB region and color correction for an ambiguous case.

### 3.2. Cast removal

The method we propose for cast removal is based on the Von Kries hypothesis with the RGB channels considered an approximation of the  $L, M, S$  retinal wavebands [12]. The estimate of the sensor scaling coefficients is assimilated within the evaluation of the color balancing coefficients. Our cast remover is not used alone, but applied after cast detection and only to those images classified as having either an evident or an ambiguous cast in order to avoid problems such as the mistaken removal of predominant color, or the distortion of the image’s chromatic balance. The diagonal transform is

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} k_R & 0 & 0 \\ 0 & k_G & 0 \\ 0 & 0 & k_B \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (9)$$

where  $RGB$  and  $R'G'B'$  are the color coordinates of the input and corrected image, respectively.

The gain coefficients,  $k_R, k_G,$  and  $k_B$  are estimated by setting at white what we have called the white balance (WB) region:

$$\begin{aligned} k_R &= WhiteR/R_{WB}, \\ k_G &= WhiteG/G_{WB}, \\ k_B &= WhiteB/B_{WB}, \end{aligned} \quad (10)$$

where  $R_{WB}, G_{WB},$  and  $B_{WB}$  are the averages of the three  $RGB$  channels over the whole selected WB region and ( $WhiteR, WhiteG, WhiteB$ ) represents the reference white chosen.

The main peculiarity of the method is that it determines the WB region on the basis of the type of cast detected. To avoid the mistaken removal of an intrinsic color, regions previously identified by the cast detector as probably corresponding to skin, sky, sea or vegetation, are temporarily removed from the analyzed image. The algorithm then looks for the WB region in the rest of the image which presents a low level of saturation. The region corresponds to the object, or, as is explained below, group of objects that will be forced to white. Forcing to white regions with high saturation usually causes a color distortion in the output image. During experimentation it became apparent that a sufficiently low level of saturation corresponds to  $C/L^* < 0.8$  (ratio of the chroma radius and the lightness in the CIELAB color space).

The algorithm now differentiates between evident cast and ambiguous cast.

In the first case, the WB region is formed by the objects corresponding to all the significant color peaks in the  $ab$  histogram of the analyzed portion of the image (Fig. 7). Only objects with a lightness greater than 30 will be accepted, because the candidates to be whitened must not be too dark.

In the case of ambiguous cast, the WB region is composed of the near neutral objects (NNO) of the analyzed portion of the original image (Fig. 8) because the influence of a

weak cast is visible only in objects near the neutral axis. By choosing the NNO region, we avoid the risk of removing the intrinsic color of an object far from the neutral axis, and the consequent chromatic distortion of the output image.

Once the scaling coefficients have been evaluated with Eq. (10), the cast remover algorithm is applied to the whole image, including regions of skin, sky, sea, and vegetation.

This color balance algorithm can be considered a mixture of the white balance and gray world procedures. The WB region is formed by several objects of different colors, which are set at white not singularly, but as an average. This prevents the chromatic distortion that follows on the wrong choice of the region to be whitened. When there is a clear cast, the colors of all the objects in the image suffer the same shift due to that specific cast. If we consider the average over several objects, introducing a kind of gray world assumption, only that shift survives and will be removed. Note that for images where the WB region consists of a single object, such as is usually the case of images with a strong cast, our method tends to correspond to the conventional white balance algorithm, while for images where the WB region includes more objects (as in images with a softer cast), the method tends to the gray world algorithm. The evaluation of the scaling coefficients can be conveniently performed on the thumbnail image or on a sub-sampled image, significantly reducing the computational time.

#### 4. Experimental results

The thresholds in our procedure were heuristically determined on a set of 40 images, indicated by our research sponsor as a good example of non-professional digital photos. These images were not used in the evaluation phase.

To verify the reliability of our method we used another data set, composed of 748 images of different sizes (from  $120 \times 160$  to  $2272 \times 1704$  pixels), resolutions and quality (in terms of jpeg compression, noise, dynamic range, etc.). These images were downloaded from personal web-pages, or acquired by our research sponsor using various digital cameras and scanners. The pre-processing of these images varied and was in the most of the cases unknown. We tried to avoid having undesirable clusters of images (i.e. images derived from the same web site and/or of the same subject), since that might have biased in some way our classification experiments.

To evaluate the goodness of the method, the performance of the cast detector and that of the cast remover must be combined. However, we first considered the performance of the cast detector alone, as this first step could also be applied in conjunction with other cast removal algorithms.

##### 4.1. Cast detector performance

Among the 748 images, 2 were considered unclassifiable by the cast detector; these images are shown in Fig. 2. For each of the remaining 746 images, the output of the cast detector was compared with the majority opinion of a panel of five experts and reported in the confusion matrix, **C**, presented in Table 1.

The diagonal terms of the matrix correspond to images for which there is an agreement between the class assigned by the detector and the opinion of the experts and cover the 85% of the cases (635/746). Examples of correctly classified images are reported in Fig. 9 (see also Figs. 10 and 11).

If we look at the misclassified elements of the matrix, we find:

- Seven predominant color images and three no-cast images erroneously classified as cast images by the cast detector. As a consequence of this misclassification, the cast remover is applied to images the colors of which should instead be preserved. This can be considered a serious error. The second and third images of Fig. 12 are examples of this type of misclassification.
- Three cast images classified by the detector as having no cast and 17 cast images as having a predominant color. These images will not be color corrected; consequently, the output images will correspond to the input ones. These cases can be considered a failure of the method as the cast is not removed; however, the images are preserved, and there is no color distortion. We do not consider this misclassification a serious error. Even less severe is the misclassification of eight ambiguous cast images as having no cast. The first image on the left of Fig. 12 is an example of a cast image erroneously classified as a predominant color image.
- A subset of 22 images classified by the detector as ambiguous cases but judged by the experts to have no cast. These cases must be evaluated individually, however the application of the cast remover to these images, does not

Table 1

	Cast detector			
	No cast	Ambiguous cast	Cast	Predominant color
Panel of experts				
No cast	<b>177</b>	22	3	2
Ambiguous cast	8	<b>133</b>	43	0
Cast	3	5	<b>219</b>	17
Predominant color	1	0	7	<b>106</b>

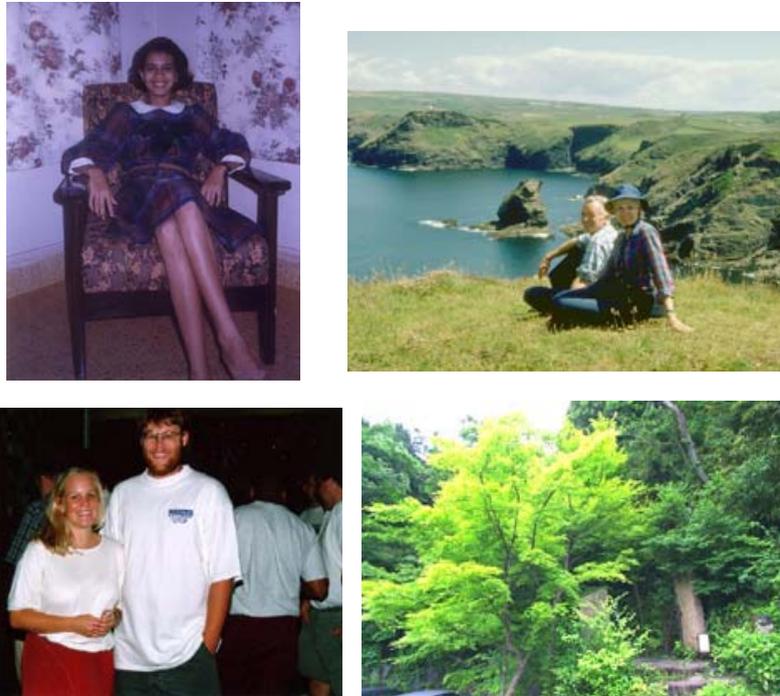


Fig. 9. Examples of images correctly classified by the detector. Above left, cast image; right, ambiguous case; below left, no-cast image; right, predominant color image.

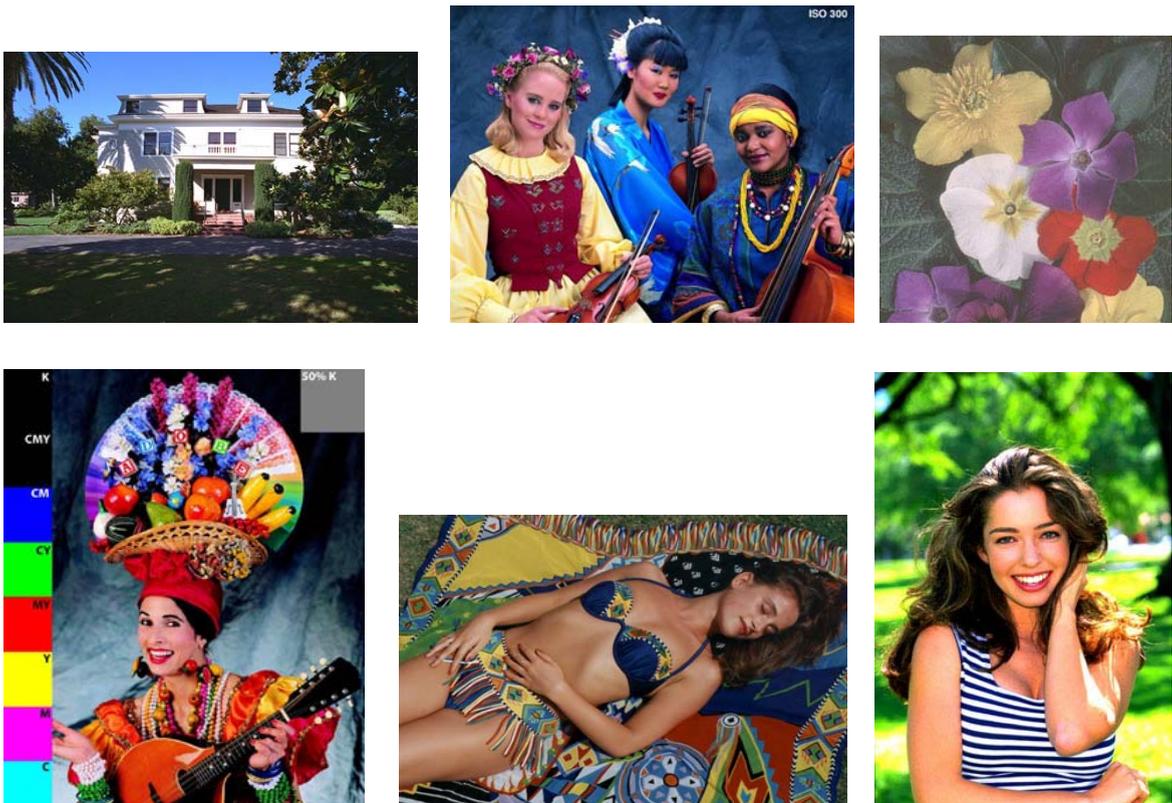


Fig. 10. Examples of images judged by the experts as no-cast images.



Fig. 11. Examples of images judged by the experts as predominant color images.



Fig. 12. Examples of misclassification. From left to right, a cast image classified as a predominant color image (not considered a serious error), a predominant color image classified as a cast image, and a no-cast image considered a cast image (the latter two, both cases of serious misclassification).

at any rate create artifacts, or substantially alter the high chroma colors.

- A subset of 48 images (5 + 43) judged ambiguous by the experts while classified by the detector as cast images, or vice versa. These images anyway have consequently been color corrected, but not according to the computational strategy designed for them. These cases must be also evaluated individually.
- Two images classified as predominant color and judged by the expert as no-cast, and one predominant color image classified as no-cast. These misclassifications have no consequences since neither no-cast nor predominant color images undergo cast removal.

#### 4.2. Cast remover performance

We consider the behavior of the cast remover only with respect to the images correctly classified as having a cast, or an ambiguous cast, (the second and third elements of the diagonal of the confusion matrix):

- Of the 219 images correctly classified as having a cast, after cast removal, 211 were judged by the panel of experts

better than the originals, six were considered equivalent, and two were judged worse. In Fig. 13, these latter 2 (left) are compared with their color balanced output (right), while Fig. 14 shows examples of cast images judged better than the originals after the color balancing.

- Among the 133 ambiguous cases, 98 images were judged by the panel of experts better than the originals after the cast removal, and the remaining equivalent. Examples of processed ambiguous cases are shown in Fig. 15 (see Fig. 16)

In the case of cast detector misclassifications, there are two possible kinds of error:

*No color balancing of cast images:* The cast images are not color balanced because classified as having no cast, or having a predominant color: the output image corresponds to the input. The error does not introduce chromatic distortion. The more frequent situation is the classification of a cast as a predominant color. It is usually the case of a strong blue, green, or pink cast which the classifier interprets as a significant region of sky, sea, vegetation, or skin, respectively, (Fig. 17). We plan to reduce this error in the future by enhancing the performance of the region classifier.



Fig. 13. The 2 cast images judged by the panel of experts to be worse after color balancing: left, the input images; right, the images after color balancing.

*Color distortion of no cast images or predominant color images:* Predominant color images or no-cast images erroneously classified as having a cast undergo color balancing. This may be the case of close-ups of a single object, or no-cast images with only one near neutral object which the cast detector considers cast images. Applying the cast remover may produce color changes in the output images. The three no-cast images, erroneously classified as cast images, are shown, before and after color balancing, in Fig. 18.

In summary, in the great majority of the experimented cases, the processed image was held to be superior to the original; only in less than 2% of the 746 images processed was the output judged worse than the original.

We have compared our results with those obtained with conventional white patch and gray world algorithms, those most often used in imaging devices. We again collected the majority opinion of the same five experts. This second analysis, gives our algorithm as consistently better than the gray world with the exception of only a few cases, (10), when it is judged equal. It does not introduce the grayish atmosphere typical of the gray world output images, as can be seen in Fig. 19 comparing the second column (output of the gray world algorithm) with the fourth (the corresponding

output of our algorithm). In the case of images with a strong cast, such as in those of the first row of Fig. 19, the gray world not only introduces a grayish overtone but at times strongly distorts the colors.

In comparison with the white patch algorithm, our method is judged better in 40% of the cases, while in the rest it is judged equal. It shows a better performance in the case of cast images with highlights, or where the acquisition system has forced the white point incorrectly. In these cases, the maxima used by the white patch corresponded to regions that had lost their chromatic information, and consequently the cast could not be removed. Examples of this kind of failure of the white patch algorithm are shown in the first and fourth rows of Fig. 19. Moreover, in the presence of a predominant color, the white patch approach generally causes severe chromatic distortion, trying to whiten the intrinsic color of the scene, as in the last row of Fig. 19.

#### 4.3. Computational cost

The prototype code has been developed and tested with the Matlab 6.5 Image Processing Toolbox and with the

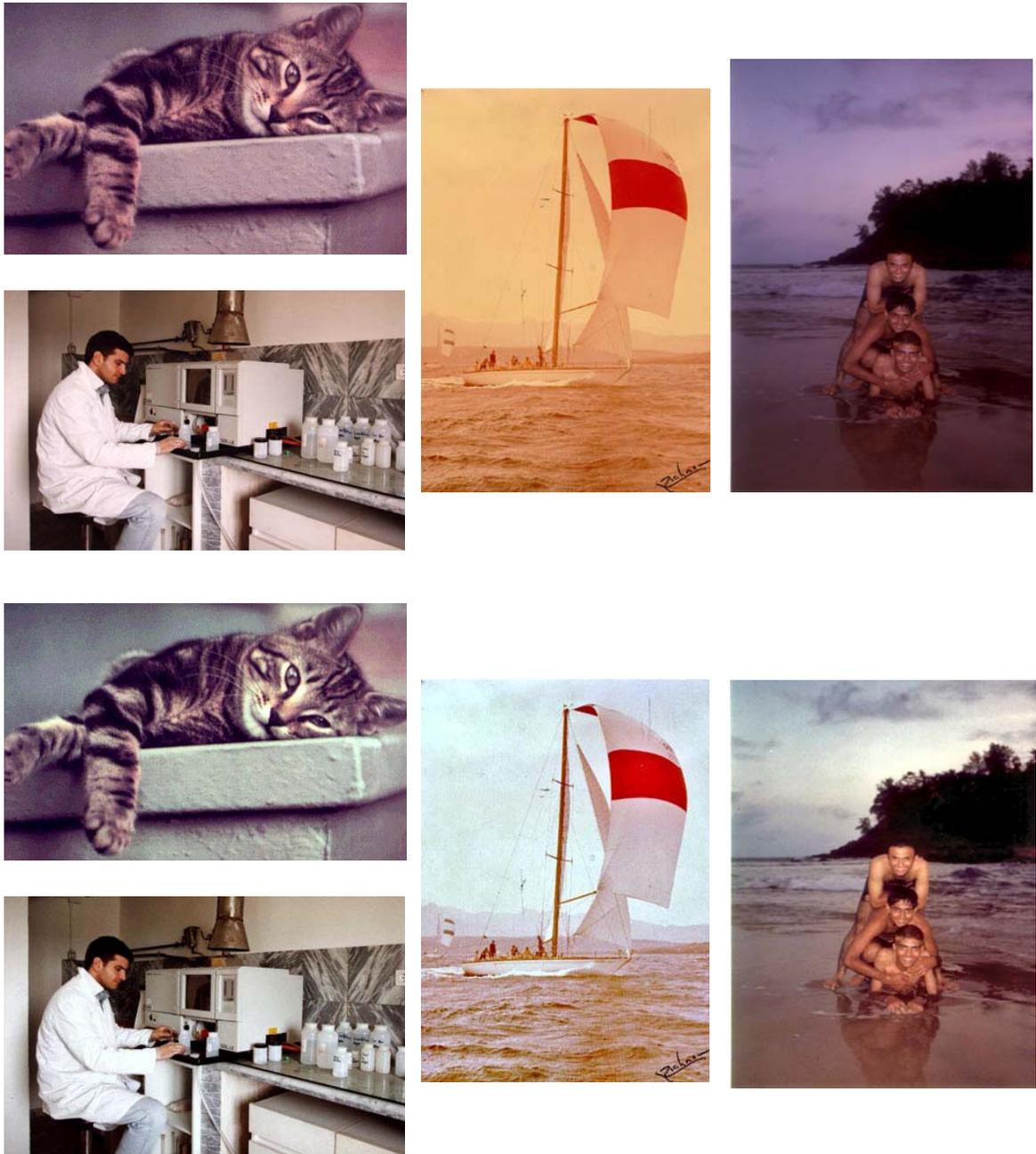


Fig. 14. Top, 4 cast images correctly classified by the cast detector. Bottom, the same 4 images after color balancing, judged by the experts better than the originals.

Borland C++ Builder 5.0, while its firmware implementation is under study.

To evaluate quantitatively the computational cost of our procedure, the two modules of cast detection and cast removal are considered separately. The image analysis in the cast detector is performed on the thumbnail image, already available in most digital cameras, or easily computable by

sub-sampling. This strategy is justified by the fact that there are no significant differences between the adopted simple statistics over the whole image and over its sub-sampled version. In the case of the C++ implementation, the average computational cost of the thumbnail analysis is about 0.17 s on a PC Intel Pentium M, 1.5 GHz, 512 MB, with Windows XP Professional as operating system. We report here

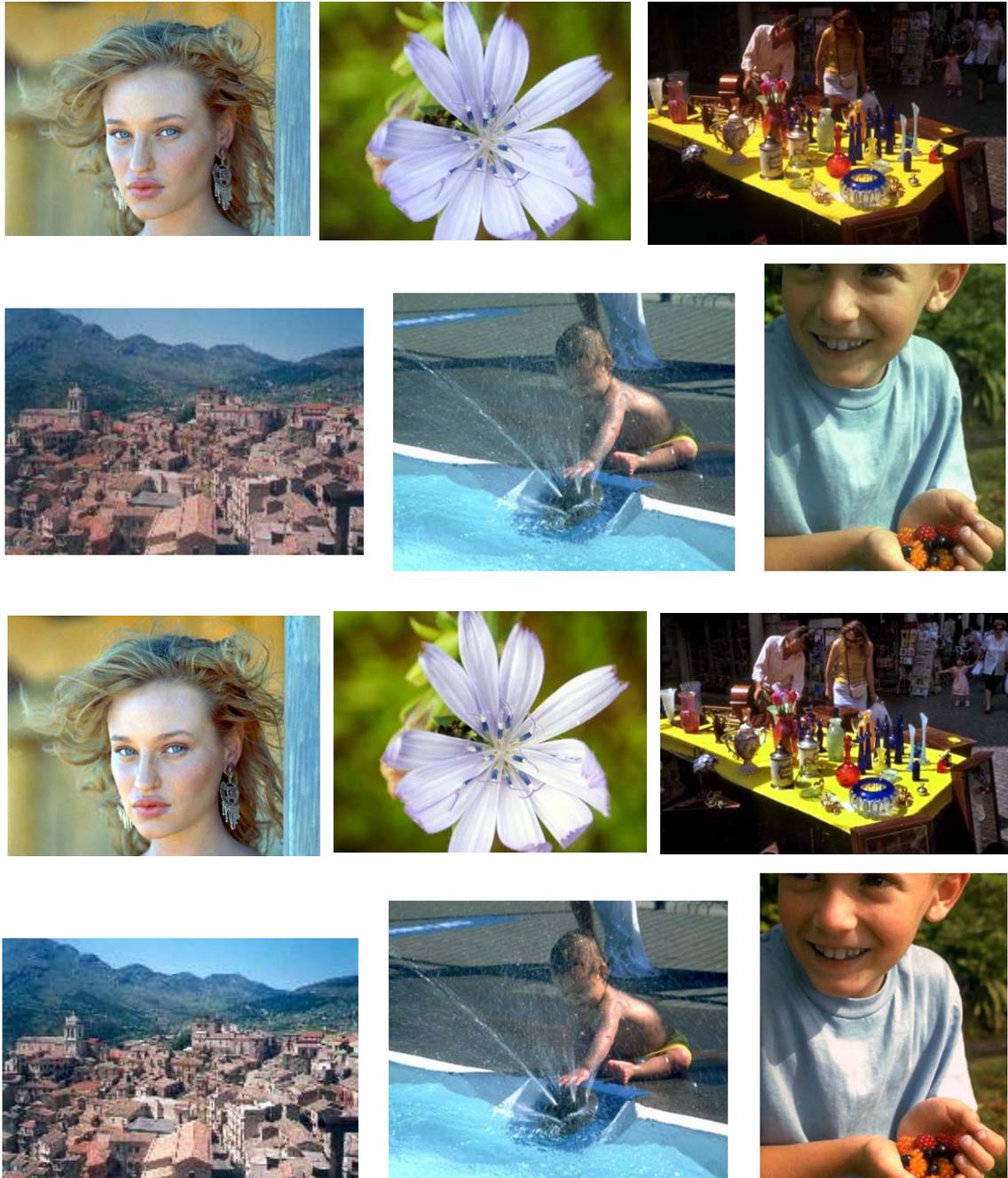


Fig. 15. Top two rows, images classified as ambiguous cases by the cast detector. Bottom two rows, the same images after color correction.

the average cost as the procedure can follow different paths depending on the characteristics of the analyzed image.

The cast remover is applied to the whole image and its cost is about  $5.6 \mu\text{s}/\text{pixel}$ . Total cost is linear with the number of image pixels as it substantially involves only a diagonal transform.

## 5. Conclusions

Color balancing is in general performed interactively. Automatic color correction, without knowledge of the image source and subject, is a very challenging task. For example, an automatic white balance adjustment applied to a sunset



Fig. 16. The cast and ambiguous case images of Fig. 9 after the color balancing. The other two images of Fig. 9 (second row) were classified as a no-cast image and a predominant color image, respectively, and did not undergo cast removal.



Fig. 17. Cast images erroneously classified as having a predominant color, and thus not color balanced.

image may drastically modify the nature of the scene, removing the characteristic reddish cast, while methods based on the gray world assumption will generate a wholly grayish scene in the case of a predominant color such as is found in underwater images, or in those of a forest of mostly bright green leaves. Color constancy algorithms work well only when prior assumptions are satisfied, for example that the scene is colorful enough, or uniformly illuminated by a single source of light. When these prerequisites do not hold, the result may be distorted, or grayed out colors. Moreover, traditional methods of cast removal do not discriminate between images with true cast and those with predominant colors, such as underwater images, or single color close-ups; but are applied in the same way to all images. This may result in

an unwanted distortion of the image chromatic content with respect to the original scene. A fundamental aspect of our method is that it is designed to distinguish between true cast and predominant color in a completely unsupervised way, allowing us to discriminate between images requiring color correction and those in which the chromaticity must, instead, be preserved. The correction is also calibrated on the type of cast, allowing the processing of even ambiguous images without color distortion. The whole analysis is performed by simple image statistics on the thumbnail image, already available in most digital cameras, or, easily computable by sub-sampling. This economy of size, together with the speed of the operations involved is a great advantage: the algorithm is so quick, it could easily be incorporated into more complex procedures, including object recognition and image retrieval, without a significant loss of computational time. Our cast detector can also be employed in tandem with color balance methods, to suggest which images should be color corrected, and which preserved.

#### Acknowledgements

Our investigation was performed as a part of a Olivetti I-jet's research contract. The authors thank Olivetti I-jet for the permission to present this paper.

#### Appendix A. sRGB to CIELAB color mapping

The sRGB color space was introduced in 1996 as a standard color space for image interchange, especially over the Internet. The sRGB color space complements current ICC

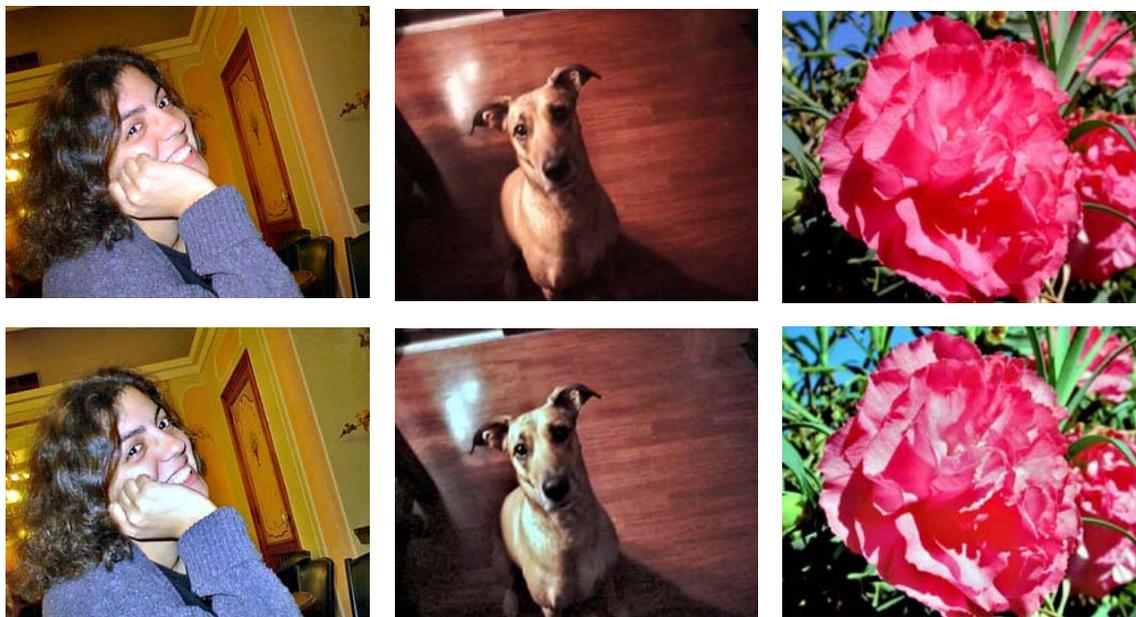


Fig. 18. Top row, the 3 no-cast images classified by the cast detector as having a cast. Bottom row, the processed images after cast removal, judged worse than the originals.

color management strategies by enabling a method of handling color in the operating systems and device drivers, using a simple and robust device independent color definition (if there does not exist an ICC profile, then the color space is sRGB. If there is an ICC profile embedded in the image, this takes priority and provides an unambiguous color space definition [28]). The sRGB color space definition is based on the average performance of typical CRT monitors under reference viewing conditions, but it is well suited to flat panel displays, television, scanners, digital cameras, and printing systems. For these reasons it has now been widely accepted in the consumer imaging industry and it has now been adopted as an IEC international standard [29]. Due to similarities of the defined reference display to real CRT monitors, often no additional color space conversion is needed to display the images. However, conversions are required to transform data into sRGB and then out to devices with different dynamic ranges, gamuts and viewing conditions. More details on the respective sRGB standard color space and the ICC profile format can be found in Refs. [28,29].

The 8 bit integer sRGB values are converted to floating point non-linear sR'G'B' values as follows:

$$R'_{sRGB} = R_{8bit}/255.0; \quad G'_{sRGB} = G_{8bit}/255.0;$$

$$B'_{sRGB} = B_{8bit}/255.0.$$

The nonlinear sR'G'B' values are transformed to linear  $R_{sRGB}, G_{sRGB}, B_{sRGB}$  values by:

$$\text{If } R'_{sRGB}, G'_{sRGB}, B'_{sRGB} \leq 0.04045,$$

$$R_{sRGB} = R'_{sRGB}/12.92; \quad G_{sRGB} = G'_{sRGB}/12.92;$$

$$B_{sRGB} = B'_{sRGB}/12.92;$$

else if  $R'_{sRGB}, G'_{sRGB}, B'_{sRGB} > 0.04045,$

$$R_{sRGB} = ((R'_{sRGB} + 0.055)/1.055)^{2.4};$$

$$G_{sRGB} = ((G'_{sRGB} + 0.055)/1.055)^{2.4};$$

$$B_{sRGB} = ((B'_{sRGB} + 0.055)/1.055)^{2.4}$$

These values are converted to XYZ (D65) by

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix}.$$

Tristimulus values are then mapped in the CIELAB color space according to the following equations [30], with the D65 reference white:  $X_n = 0.9505; Y_n = 1.00; Z_n = 1.0891;$  and  $q = Y/Y_n; p = X/X_n; r = Z/Z_n;$

$$L^* = 116(Y/Y_n)^{1/3} - 16 \text{ for } (Y/Y_n) > 0.008856;$$

$$L^* = 903.3 (Y/Y_n) \text{ for } (Y/Y_n) \leq 0.008856$$

$$a^* = 500*(p_1 - q_1),$$

$$b^* = 200*(q_1 - r_1),$$

where

$$q_1 = 7.787q + 16/116 \text{ for } q \leq 0.008856;$$

$$q_1 = q^{1/3} \text{ for } q > 0.008856,$$

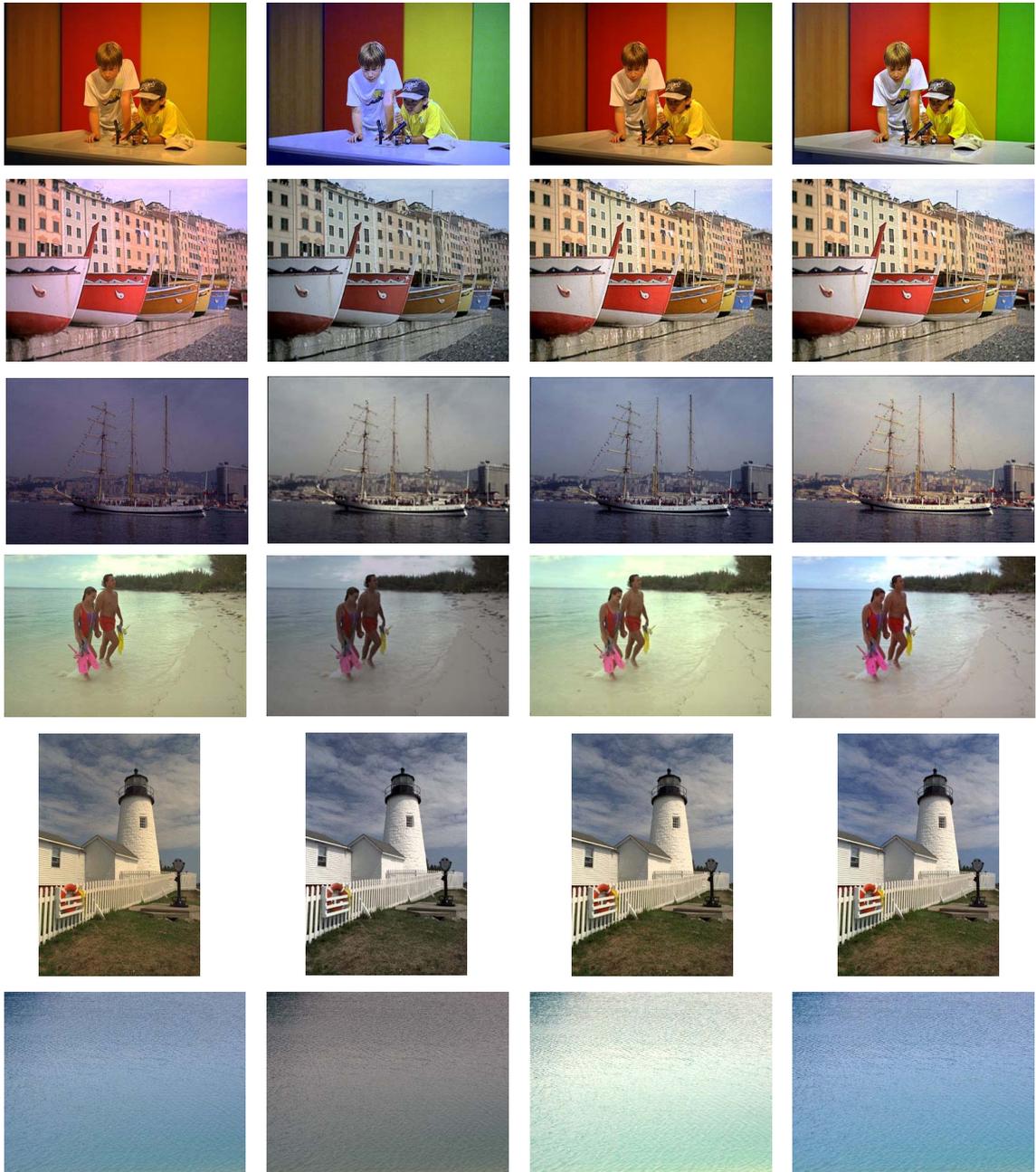


Fig. 19. Comparison of the output of our algorithm (last column), with that of the gray world (second column) and of the white patch (third column). The gray world may fail in that images without a sufficient variation in color. The white patch in the case of cast images with highlights. Our algorithm shows a greater stability.

$$p_1 = 7.787p + 16/116 \text{ for } p \leq 0.008856;$$

$$p_1 = p^{1/3} \text{ for } p > 0.008856,$$

$$r_1 = 7.787r + 16/116 \text{ for } r \leq 0.008856;$$

$$r_1 = r^{1/3} \text{ for } r > 0.008856.$$

## References

- [1] B. Funt, K. Barnard, L. Martin, Is machine colour constancy good enough?, Proceedings of the Fifth European Conference on Computer Vision, Freiburg, Germany, 1998, pp. 445–459.

- [2] V. Cardei, B. Funt, K. Barnard, White point estimation for uncalibrated images, Proceedings of the IS&T/SID Seventh Color Imaging Conference, Scottsdale, USA, 1999, pp. 97–100.
- [3] V. Cardei, B. Funt, M. Brockington, Issues in color correcting digital images of unknown origin, Proceedings of the 12th International Conference on Control Systems and Computer Science, CSCS'12, Bucharest, Romania, 26–29 May, 1999.
- [4] G. Ciocca, D. Marini, A. Rizzi, R. Schettini, S. Zuffi, Retinex pre-processing of uncalibrated images for color based image retrieval, *J. Electron. Imaging* 12 (1) (2003) 161–172.
- [5] M.D. Fairchild, *Color Appearance Models*, Addison-Wesley, Reading, MA, 1997.
- [6] G.D. Finlayson, M.S. Drew, B. Funt, Diagonal transform suffice for color constancy, Proceedings of the IEEE International Conference on Computer Vision, Berlin, 1993, pp. 164–171.
- [7] G.D. Finlayson, M.S. Drew, B. Funt, Spectral sharpening: sensor transformations for improved color constancy, *J. Opt. Soc. Am. A* 11 (1994) 1553–1563.
- [8] K. Barnard, F. Ciurea, B. Funt, Sensor sharpening for computational color constancy, *J. Opt. Soc. Am. A* 18 (2001) 2728–2743.
- [9] <http://www.cie.co.at/cie/>
- [10] G. Buchsbaum, A spatial processor model for object color perception, *J. Franklin Inst.* 310 (1980) 1–26.
- [11] E. Land, The retinex theory of color vision, *Scien. Am.* 237 (1977) 108–129.
- [12] Edwin H. Land, John McCann, Lightness and retinex theory, *J. Opt. Soc. Am.* 61 (1) (1971) 1–11.
- [13] D. Forsyth, A novel algorithm for color constancy, *Int. J. Comput. Vision* 5 (1990) 5–36.
- [14] G.D. Finlayson, Color in perspective, *IEEE Trans. Pattern Anal. Mach. Intell.* 18 (10) (1996) 1034–1038.
- [15] G.D. Finlayson, S. Hordley, Selection for gamut mapping colour constancy, Proceedings of the Eighth British Machine Vision Conference, Colchester, England, Vol. 2, 1997, pp. 630–639.
- [16] G.D. Finlayson, S. Hordley, A theory of selection for gamut mapping colour constancy, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 41, Santa Barbara, CA, USA, 1998, pp. 60–65.
- [17] G.D. Finlayson, S. Hordley, Improving gamut mapping colour constancy, *IEEE Trans. Image Process* 9 (10) (2000) 1774–1783.
- [18] K. Barnard, V. Cardei, B. Funt, A comparison of computational color constancy algorithms—Part I: methodology and experiments with synthesized data, *IEEE Trans. Image Process.* 11 (7) (2002) 972–983.
- [19] K. Barnard, L. Martin, A. Coath, B. Funt, A comparison of computational color constancy algorithms—Part II: experiments with image data, *IEEE Trans. Image Process* 11 (7) (2002) 985–995.
- [20] G.D. Finlayson, P.H. Hubel, S. Hordley, Color by correlation, Proceedings of the IS&T/SID Fifth Color Imaging Conference: Color Science, Systems and Applications, Scottsdale, USA, 1997, pp. 6–11.
- [21] G.D. Finlayson, P.H. Hubel, S. Hordley, Color by correlation: a simple unifying framework for color constancy, *IEEE Trans. Pattern Anal. Mach. Intell.* 21 (2001) 1209–1221.
- [22] G.D. Finlayson, P.H. Hubel, White point estimation using color by correlation, U.S. Patent 6038339, 2000.
- [23] K. Barnard, V. Cardei, B. Funt, Learning color constancy, Proceedings of the IS&T/SID Fourth Color Imaging Conference, Scottsdale, USA, 1996, pp. 58–60.
- [24] B. Funt, V. Cardei, Bootstrapping color constancy, Proceedings on Human Vision and Electronic Imaging IV, San Jose, CA, 1999, pp. 421–428.
- [25] T. Cooper, A novel approach to color cast detection and removal in digital images, *Proc. SPIE* 3963 (2000) 167–175.
- [26] T. Cooper, Color segmentation as an aid to white balancing for digital still cameras, *Proc. SPIE* 4300 (2001) 164–171.
- [27] <http://www.srgb.com>.
- [28] International Color Consortium (ICC). <http://www.color.org>.
- [29] International Electrochemical Commission, Colour measurement and management in multimedia systems 75 and equipment, Part 2–1: colour management. Default RGB colour space-sRGB, IEC Publication 61966-2.1, 1999.
- [30] R.W.G. Hunt, *Measuring Color*, Ellis Horwood, Chichester, 1987.
- [31] C. Cusano, G. Ciocca, R. Schettini, Image annotation using SVM, in: S. Santini, R. Schettini (eds.), Proceedings on Internet Imaging IV, Vol. SPIE 5304. pp. 330–338, San Jose, CA, USA, 2004, in press.

**About the Author**—RAIMONDO SCHETTINI is Associate Professor in the Department of Information Science, Systems Theory, and Communication (DISCo), at the University of Milano Bicocca, where he is in charge of the Imaging and Vision Lab. He has worked with Italian National Research Council (C.N.R.) since 1987. He has been team leader in several national and international research projects and published over 140 refereed papers on image processing, analysis and reproduction, and image content-based indexing and retrieval. He was General Co-Chairman of the First Workshop on Image and Video Content-based Retrieval (1998) of the First European Conference on Color in Graphics, Imaging and Vision (CGIV'2002), and of the EI Internet Imaging Conferences (2000–2004). He has been guest editor of a special issue about Internet Imaging of the *Journal of Electronic Imaging*, 2002, of another one on Color Image Processing and Analysis in *Pattern Recognition Letters*, 2003, and is currently associate editor of the *Pattern Recognition Journal*.

**About the Author**—FRANCESCA GASPARINI took her degree (Laurea) in Nuclear Engineering at the Polytechnic of Milan in 1997 and her Ph.D. in Science and Technology in Nuclear Power Plants at the Polytechnic of Milan in 2000. Since 2001 she has been a fellow at the ITC, Imaging and Vision Laboratory, of the Italian National Research Council. Francesca is currently a Post-doc Researcher at DISCo (Dipartimento di Informatica, Sistemistica e Comunicazione), University of Milano Bicocca, where her research focuses on image and video enhancement, face detection, and image quality assessment.