Face-Based Illuminant Estimation

Simone Bianco and Raimondo Schettini

University of Milano-Bicocca
DISCo - Department of Informatics, Systems and Comunication

Abstract. In this work we show that it is possible to use skin tones to estimate the illuminant color. We use a face detector to find faces in the scene, and the corresponding skin colors to estimate the chromaticity of the illuminant. The method, that has been presented at CVPR 2012 [1] is based on two observations: first, skin colors tend to form a cluster in the color space, making it a cue to estimate the illuminant in the scene; second, many photographic images are portraits or contain people. If no faces are detected, the input image is processed with a low-level illuminant estimation algorithm automatically selected according to [2]. The method will be demonstrated on a public dataset of images in RAW format [3], and on images acquired live during the demo.

1 Introduction

Computational color constancy aims to estimate the actual color in an acquired scene disregarding its illuminant. Many illuminant estimation solutions have been proposed in the last few years, although it is known that the problem addressed is actually ill-posed as its solution lacks uniqueness and stability. In this work, we investigate how illuminant estimation can be performed exploiting the color statistics extracted from the faces automatically detected in the image. This work is based on the following observation [1]: skin colors provide enough and reliable information to estimate the scene illuminant and suggest the use of faces to detect skin areas. Skin colors tend to form a cluster in various color spaces and we have shown that the diversity between the gamut of skin pixels of the detected faces and the skin canonical gamut can be affordably used to estimate the scene illuminant. This observation is supported by the fact that consumer photos are very often of faces, and face detection is a technology that has been already integrated in digital cameras with very affordable results.

2 The Proposed Approach

In this work we propose a new method to estimate the scene illuminant using faces: more exactly, we detect faces in the image and extract the skin color information using a rough skin detector. Our method is based on the assumption that skin colors form a sufficiently compact cluster in the color space in order to represent a valid clue for illuminant estimation [1]. Our approach applies gamut

A. Fusiello et al. (Eds.): ECCV 2012 Ws/Demos, Part III, LNCS 7585, pp. 623–626, 2012.

mapping to skin pixels only: the illuminant is estimated as the mapping that can be applied to the gamut of the skin colors in the input image, resulting in a gamut that lies completely within the skin canonical gamut. The Gamut Mapping assumes that for a given illuminant, one observes only a limited gamut of colors [4]. It has a training phase in which a canonical illuminant is chosen and the canonical gamut is computed observing as many surfaces under the canonical illuminant as possible. Given an input image with an unknown illuminant, its gamut is computed and the illuminant is estimated as the mapping that can be applied to the gamut of the input image, resulting in a gamut that lies completely within the canonical gamut and produces the most colorful scene.

The first step is the computation of the skin canonical gamut S_C: this is the convex hull of the skin colors of different people acquired under the chosen canonical illuminant. Given an input image for which we want to estimate the illuminant, in which F faces are present, the masks $u_f(x, y)$, $f = 1, \ldots, F$, are obtained:

$$u_f(x, y) = \begin{cases} 1 \text{ if } (x, y) \in \text{face no. } f \wedge \text{is a skin pixel} \\ 0 \text{ otherwise} \end{cases}, \tag{1}$$

i.e. $u_f(x, y)$ assumes the value 1 only on the pixels inside the $f-$th detected face area, being classified as skin pixels. The extracted skin colors $\boldsymbol{skin}_f$ are then computed as

$$\boldsymbol{skin}_f = \{\boldsymbol{\rho}(x, y) : u_f(x, y) = 1\}, \tag{2}$$

Once we have extracted the skin colors for each detected face, these are converted into the $\mathrm{YC_bC_r}$ color space and luminance normalized, such that the average luminance $\overline{Y} = 0.5$. The skin gamut S_I is then computed as the convex hull of the converted values of all detected faces. The set of feasible mappings $\mathcal{M}$ is then determined: it consists of all mappings that can be applied to the skin gamut S_I of the input image and that result in a gamut that lies completely within the skin canonical gamut S_C:

$$\mathcal{M} = \{\mathcal{M}_i : \mathcal{M}_i S_I \in S_C\} \tag{3}$$

In this work each transformation $\mathcal{M}_i$ is be modeled by a diagonal mapping or von Kries model. The illuminant color is then estimated as the inverse of the average of the feasible set [5].

Once we have estimated the illuminant color , given the choice of diagonal mappings, each pixel in the image is color corrected using the von Kries model.

The operation flowchart of the proposed method is reported in Figure 1. The face detector module is run on the input image to detect any faces. If no faces are detected, the input image may be processed with any other state-of-the-art illuminant estimation algorithm. If one or more faces are detected, a skin detection module is run on the detected faces to filter out any non-skin and unreliable pixels.

Looping on all the faces detected, the first step of the skin detection is the conversion of the detected face pixels in the HSV color space. Then a technique

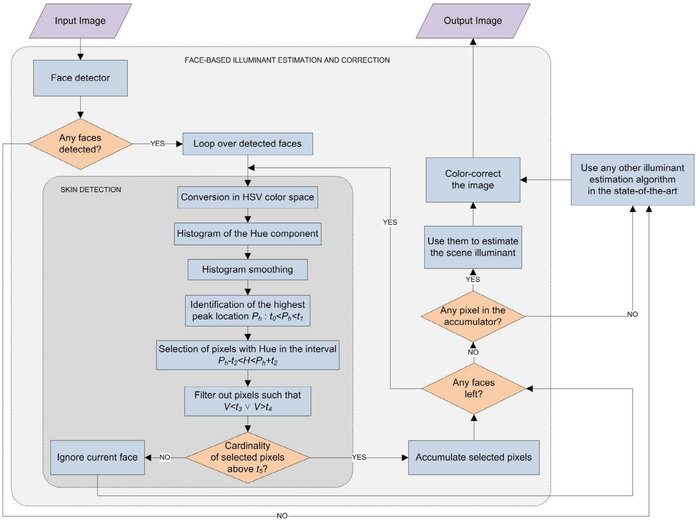

Fig. 1. The operation flowchart of the proposed method

based on scale-space histogram filtering [6] is used to identify the highest peak location and width of the histogram of the hue component, within the hue interval corresponding to feasible skin colors.

The highest peak location P_h and width $2w$ are identified within the hue interval corresponding to feasible skin colors, i.e. $P_h \in [t_0 \; t_1]$. Then only those pixels satisfying the condition $P_h - w < H < P_h + w$ are selected. This condition permits to implement a sort of adaptive skin detector, since the selected hue interval depends on the current peak location P_h. Any unreliable skin pixel as being too dark or too bright, and thus potentially clipped, is filtered out if it satisfies the condition $V < t_2 \vee V > t_3$. If the cardinality of the remaining skin pixels, normalized for the total number of pixels in the detected face, is above the threshold t_4, they are converted into the $\mathrm{YC}_b\mathrm{C}_r$ color space where they are luminance normalized as previously described, and then they are added in an accumulator.

When the looping on all the faces detected is finished, the accumulator is analyzed: if it is non-empty, the color coordinates of the skin pixels in it are used to estimate the scene illuminant by first computing the skin gamut S_I and then using the skin-based gamut mapping; otherwise the scene illuminant can be estimated using any other state-of-the-art illuminant estimation algorithm. Given the illuminant estimation, the input image is then color corrected to give the final output image.

It should be noted that the proposed method in the present form makes it possible to discard faces with an unnatural mixing illuminant condition (such as when the face is under a colored umbrella) due to the feasibility check on the hues.

3 Demonstrated Results

The method will be visually and quantitatively compared with a large number of low-level, intermediate level and high-level algorithms in the state of the art [7] on a public dataset of images in RAW format [3], and on images acquired live during the demo. The images acquired live will include the Macbeth ColorChecker (MCC) chart, which will be automatically localized [8], to accurately estimate the ground truth illuminant.

Fig. 2. Example of the image in the portrait dataset on which the proposed algorithm makes the two highest angular error. Left to right: original image; ideal correction based on the MCC; correction with the proposed algorithm; correction with the Gamut Mapping algorithm.

References

1. Bianco, S., Schettini, R.: Color constancy using faces. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 65–72 (2012)
2. Bianco, S., Ciocca, G., Cusano, C., Schettini, R.: Automatic color constancy algorithm selection and combination. Pattern Recognition 43, 695–705 (2010)
3. Gehler, P., Rother, C., Blake, A., Minka, T., Sharp, T.: Bayesian color constancy revisited. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1–8 (2008)
4. Forsyth, D.A.: A novel algorithm for color constancy. International Journal of Computer Vision 5, 5–36 (1990)
5. Barnard, K.: Improvements to Gamut Mapping Colour Constancy Algorithms. In: Vernon, D. (ed.) ECCV 2000, Part I. LNCS, vol. 1842, pp. 390–403. Springer, Heidelberg (2000)
6. Witkin, A.: Scale-space filtering: a new approach to multi-scale description. In: IEEE International Conf. on Acoustics, Speech, and Signal Processing, vol. 9, pp. 150–153 (1984)
7. Gijsenij, A., Gevers, T., van de Weijer, J.: Computational color constancy: survey and experiments. IEEE Transactions on Image Processing 20, 2475–2489 (2011)
8. Bianco, S., Cusano, C.: Color Target Localization under Varying Illumination Conditions. In: Schettini, R., Tominaga, S., Trémeau, A. (eds.) CCIW 2011. LNCS, vol. 6626, pp. 245–255. Springer, Heidelberg (2011)