Color Correction for Face Detection Based on Human Visual Perception Metaphor

Krzysztof M. Kryszczuk and Andrzej Drygajło
Signal Processing Institute, Swiss Federal Institute of Technology, Lausanne (EPFL)
[krzysztof.kryszczuk, andrzej.drygajlo]@epfl.ch

Abstract

In this paper we present a method of automatic color correction of face images and its application in a face detection algorithm. The color correction method is based on the phenomenon of color constancy observed in human visual perception. This technique is further applied in a face detection system, which draws upon the analogy to the parallel organization of visual neural pathways, the magno- and parvocellular channels. Presented method proved to be efficient in diverse background and illumination conditions, including face images with background chromatically close to human skin and where prominent facial features are obscured by adverse illumination conditions.

1. Introduction

Processing of human face images is an important research area with many applications, ranging from image enhancement to automatic face recognition in security systems. Beside the face itself, most face images contain a background that must be discarded before subsequent face recognition process. Thus in most cases the first step in the image-processing task is the detection and localization of the face in the image.

A comprehensive overview of state-of-the-art face detection methods is presented by Yang et al. [1]. Particularly the knowledge-based, feature invariant, and template matching algorithms are listed as the most frequently used ones.

Human skin color can be regarded as an invariant feature and so are the skin color based methods classified by Yang et al. In fact, the skin color is an easily accessible, computationally inexpensive feature. Therefore it has been used in various face detection and recognition systems [2,3,4].

Despite the apparent skin color variations between different ethnic groups the actual skin chromaticity parameters can be clustered into a surprisingly compact set, which allows very accurate modeling [5]. The resulting skin color model can be used for color-based image segmentation focused on locating the skin-colored areas. This method of segmentation can deliver very precise distinction between the face and non-face areas of the image, provided that the background differs chromatically from the skin tone. The skin colored areas considered for further face recognition (verification or identification) can be accurately cropped out from the original image.

Skin model-based segmentation can result in precise skin area detection only if the model was created using the same spectral content of the skin illuminant as in the processed face image. Usually the information about skin illuminant is unknown for an arbitrary color image. Therefore a mismatch can occur between the model assumptions and the chromatic properties of skin depicted in the actual image. To avoid this mismatch it is necessary to normalize the image chromatically by introducing a chromatic frame of reference, common to both the model and the segmented image.

Precise retrieval of the spectral content of the illuminant in an arbitrary visual scene is an ill-posed problem [6]. Therefore a few heuristic methods have been proposed to normalize the chromaticity of the image [7].

Humans are known to cope well with the problem of color discrimination under varying illuminants thanks to the mechanism of color constancy observed in the natural visual processing [8]. In order to process a color face image acquired under unknown lighting conditions it is necessary to first employ a color correction mechanism, which would do what the phenomenon of color constancy does in humans.

The classical two assumptions that most color correction methods are based on are the "white world assumption" and the "gray world assumption" [9]. The first one assumes that there is a part of each image that is white. The second one postulates that all colors in the image should average to gray.

Hsu et al. [10] presented an interesting approach toward color correction of face images. They proposed an automatic color correction based on the localization of pixels with top 5% of luminance in the image, and assume those pixels to be ‘white’ (the “white world assumption”). Based on the chromatic distance between the white color and the actual color of the selected pixels the entire picture is being corrected. This method works with images that contain no specular reflections. However, in non-controlled environment or where the illumination control is limited, specular reflections of the face appear very frequently.

The “gray world assumption” is not applicable to face images either, taking into consideration the fact that face images normally contain large skin-colored areas.

In this paper, we propose a new method of color cast removal from face images based on the inherent chromatic features of the face itself. In order to take full advantage of the method we incorporate it into a new robust face detection algorithm inspired by the organization of the human visual pathways (magno-and parvocellular channels) [8].

The rest of the paper is organized as follows: firstly, the general assumptions and details of the proposed method are explained. Then the proposed method is employed in a face detection algorithm. Results and final remarks conclude the paper.

2. The concept of image color correction inspired by the color constancy phenomenon

In order to be compliant with the assumption that the skin model must be built around a common frame of reference with processed face image we propose to use the chromatic information contained in the eye area as such a reference. We use this reference to perform the chromatic correction of the entire image. This process can be interpreted as a chromatic normalization.

The vast majority of images that are otherwise suitable for face verification (frontal pose, no occlusions etc.) show the face in such a way that both or at least one of the eyes are clearly visible. The image of an open eye normally contains the pupil, the iris, the eye-white and the eyebrow. A close inspection of eye images reveals that the eye-whites and the pupil areas are the locations, which are chromatically close to gray. The concept of the chromatic normalization can be best formulated as “bringing to gray what is closest to gray”.

The proposed method is to find in the image of the eye pixels that are closest to gray. Consequently the chromatic coordinates of such pixels are modified to match gray, and same transformation is applied to the entire image. In order to perform this normalization procedure it is necessary to: localize the eye areas in the image, crop out the eye images and find the appropriate pixels for correction.

3. Color correction algorithm and creation of the skin color model

We build the skin color model using samples from face images from the VIDTIMIT database [11]. Before we take the samples, the images have to be chromatically normalized. To do that, we first locate the eye areas in the image. In our experiments we found them manually. We select for correction the area of left or right eye, whichever has the lower mean luminance. We assume that if the specular reflections are present, they will be more prominent in the overall “brighter” eye image.

For each pixel in the cropped eye image a distance from gray is calculated, using the formula:

$$D_g = \mathrm{abs}(R\text{-}G) + \mathrm{abs}(G\text{-}B) + \mathrm{abs}(\mathrm{B}\text{-}R), \qquad (1)$$

where D_g is the distance from gray and R,G,B are corresponding red, green and blue chromatic coordinates of the pixel. The pixel whose D_g is smallest is selected as the normalization reference and this pixel will be brought to gray. Next, the target gray coordinates C_g (equal for all three RGB channels) of the pixel are calculated as the rounded average of its actual coordinates:

$$C_g = \mathrm{round}[(R+G+B) / 3]. \qquad (2)$$

The difference between the original RGB coordinates of the pixel and its new target gray coordinates is calculated as follows:

$$\begin{aligned} D_R &= R - C_g, \\ D_G &= G - C_g, \\ D_B &= B - C_g. \end{aligned} \qquad (3)$$

The calculated values of D_R, D_G, *and* D_B are respectively subtracted from corresponding red, green and blue chromatic coordinates of every pixel in the original image. Should the resulting coordinate exceed the allowed range, its value is set to the extreme allowed value.

The described color correction was performed on 13 face images from the VIDTIMIT database. Then, from each image a 30 by 30 pixels patch containing skin from the face was cropped out. Each of the patches (initially in RGB format) has been converted into YCbCr color space, and the Y coordinate discarded. Resulting Cb and Cr coordinates have been clustered and their distribution modeled by a sum of two normal distributions (Figure 1).

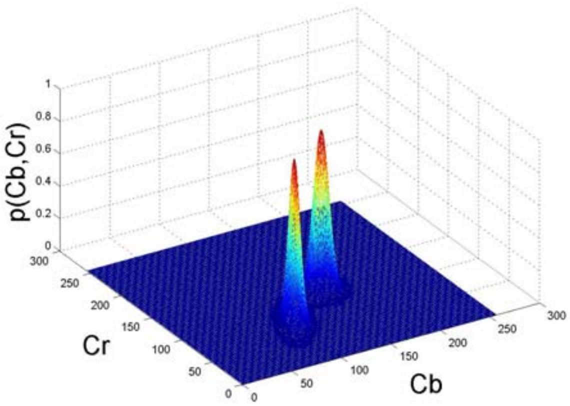

Figure 1. Skin color model in the YCbCr color space. The graph represents a probability density distribution of Cr and Cb coordinates of pixels that belong to skin-colored areas of the image.

4. Skin-color oriented image segmentation

For the processed image, the probability that each pixel's color belongs to the skin model distribution is calculated. The calculated probability values are stored in a new grayscale image, further referred to as the "skin map".

Performance of the model has been tested on a set of images different from those used for the creation of the skin color model. For each of the images the coordinates of the eyes were found manually, like during the model training. The test images were treated using the color correction procedure as described in Section 3. The model was tested for segmentation on images with and without the proposed color correction procedure. Example results are presented in Figure 2:

5. Application of the color correction method to face detection

Color information is used in many face detection and tracking algorithms. If all of the images originate from the same camera type and the spectral content of the illuminant is known, color-based segmentation is a way to quickly and robustly localize skin-colored areas without applying any prior chromatic correction. Typically, precise shape-based face detection techniques are applied after the color-based image segmentation [1].

However, if the face in the image is illuminated with a light source of unknown spectral power distribution, or/and the illumination is highly non-uniform, this approach often produces errors. Frequently the skin area in the image is not detected, or even worse, erroneously labeled.

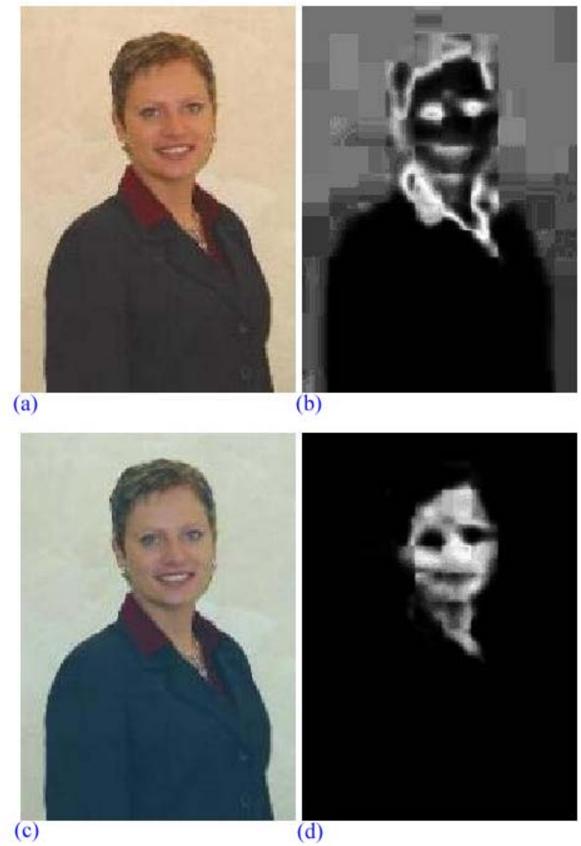

Figure 2. Results of the skin-color segmentation of the face images: (a) original image, (b) skin map of the original image, (c) original image after color correction, (d) skin map of the image after color correction.

In order to be able to use the color information to detect face in any image, we draw upon the analogy to the natural human visual system, which is known to successfully cope with the task of distinguishing colors in the presence of various illuminants.

Firstly, we revert to the idea of two separate neural pathways in the human visual system, the parvocellular and the magnocellular pathways [8] (further referred to as P-channel and M-channel, respectively). The M-channel conveys the generic shape, motion and intensity information, while the P-channel is responsible for the transmission of fine detail and color information.

As shown in numerous studies in visual search tasks, humans use the information from both neural pathways to find the desired information from a visual scene. For a given scene, the information from the channel that conveys the more discriminating data is used. If the object of interest stands out chromatically from the rest of the scene the color information is predominantly used. In a chromatically uniform scene the shape information prevails.

Therefore, we propose to use the color information simultaneously with shape-based face detection techniques for robust detection of faces in images as a high-level analogy to the M/P-channel visual processing in humans.

6. M/P-channel inspired face detection

Since the M- and P-channel processing is responsible for processing qualitatively different information about the image we propose to reproduce this dichotomy in a face detection system. In particular, we design a shape processing routine to model the M-channel, and a color processing routine to model the P-channel.

7.1 M-channel-based search

To model the M-channel search for faces in the visual scene (image) we use a template-matching approach. As a template a general grayscale 'average face' image is used (Figure 3).

Figure 3. Average face template, resolution 115×119 (columns×rows).

The search process is performed as follows: the original image is converted into its grayscale version. Both the resulting grayscale image and the face template are high-pass filtered to reduce high contrasts in the face caused by non-uniform lighting distribution, specular reflections and self-shadows. The filtered image is divided into highly overlapping windows (5 pixels overlap) of the same size as the face template. For every window a 2D correlation coefficient with the face template is calculated. Negative correlation coefficient values are changed to null. Resulting values from the range (0,1) are regarded as probabilities of finding the face at a given window.

7.2 P-channel-based search

For each monochromatic window processed as described above, a corresponding window of identical size and location is cropped out of the original color image. Since each window is expected to contain a face image, we process them as if they would indeed contain a face. Figure 4 shows an example of this procedure. Figure 4(a) shows a chosen window before correction. Using the geometry of the average face template we automatically designate the areas that are most likely to contain the eyes in each window, presuming that the face is indeed in the window. Those areas are shown in Figure 4(c) and (d). Selection of the chromatic reference point for normalization is depicted in Figure 4(e). Consequently we perform color correction procedure described in Section 3, but the correction is applied to the current window only, rather than the entire image. The color-corrected window is shown in Figure 4(b).

Figure 4. Automatic color correction of the image window; (a) original window, (b) window after color correction, (c) right eye area, (d) left eye area, (e) selection of the chromatic point of reference for color correction. Window taken from an image acquired from an USB camera (IBM), resolution 320×240. Window resolution 60×60.

Following the correction, a skin map is calculated for each window using the skin model as described in Sections 2 and 3.

Calculating the skin map for every window is a high computational burden. In order to speed up this stage, after the color correction step every window is downsampled by a factor of 4, and the skin map is calculated on the downsampled window.

The probabilities calculated for every pixel of the window are then averaged, which gives a mean likelihood measure that the given window contains the image of human skin.

7.3 Combining the M- and P-channel information

The procedure eventually returns two probability values for every window: P_L, probability that the shape in the window has a shape of a human face, and P_S, probability that the window contains object colored like human skin.

Since the information used in shape and color processing are obtained independently we calculate the joined probability that the window contains a face $P_{S,L}$ by multiplication of probabilities:

$$P_{S,L} = P_S \cdot P_L. \tag{4}$$

The window with the highest $P_{S,L}$ is a candidate to be the actual detected face in the image. However, the exact size of the face in the image is not *a priori* known, so it is necessary to perform the face search as described above for a few scaled versions of the face template. For each run with a different template size, we obtain a new $P_{S,L}$ and the window that corresponds to it. We choose the window with the highest overall value of the probability $P_{S,L}$.

The presented method of color correction for face detection has been tested on high quality images from the VIDTIMIT database, pictures with adverse lighting conditions taken from a web-cam, and scanned photographs. Figures 5-9 show the results of the experiments. Figure 5 shows an example of a good quality picture taken from the VIDTIMIT database. Figures 6, 7 and 8 show the images acquired from a computer USB camera (IBM), taken in our laboratory, where the walls and the ceiling are chromatically close to the color of the skin. The face in Figure 6 is illuminated from its right side with daylight (coming from a window). Due to this condition the right side of the face shows strong reflections while the left side remains in the shadow. Figures 7 and 8 have the same daylight illuminant as present in Figure 6, additionally augmented by warm-white light originating from the fluorescent lamps overhead. In those pictures, top left part of the head shows highlights and the entire scene is illuminated by sources of two distinctly different spectral contents. Finally, Figure 9 shows a picture scanned from a paper photograph and saved in low resolution. In this figure, the background is chromatically very close to the skin tone.

Figure 5. Image from VIDTIMIT database (res. 512×384)

Figure 6. Image acquired from an USB camera (res. 320×240)

Figure 7. Image acquired from an USB camera (res. 320×240)

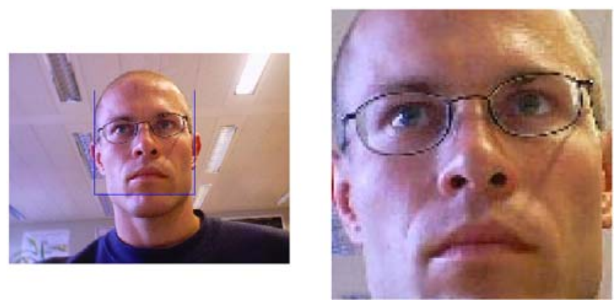

Figure 8. Image acquired from an USB camera (res. 320×240)

Figure 9. Image scanned from paper photograph, resolution of the jpeg compressed image 157×221.

7. Conclusions

In this paper we propose a method that successfully performs color correction of face images. We presented a way to incorporate this method into a generic algorithm that detects faces in images of various resolution and quality, where the face image may be distorted by adverse illumination. The advantage of the technique is that it detects a face if it is present; if it is not this fact can be inferred from the probability measures obtained during the detection process.

The algorithm may produce erroneous detection only in rare cases where neither the shape, nor the color can deliver reliable information about the location of the face. This can happen when the shape of the face is heavily distorted by adverse lighting conditions and at the same time the color of the background is indistinguishable from the skin tone. In such cases, due to lack of reliable color clues the system relies entirely on the template matching to find the best face candidate. In order to improve the system performance in such cases more appropriate filtering method than simple high-pass filter should be applied.

8. References

[1] M.H. Yang, D.J. Kriegman, N. Ahuja, "Detecting Faces in Images: A Survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, No. 1, January 2002.

[2] S. Marcel, S. Bengio, "Improving Face Verification using Skin Color Information", *Proc. 16th International Conference on Pattern Recognition*, 2002.

[3] M. Störring, T. Kočka, H.J. Andersen, E. Granum, "Tracking regions of human skin through illumination changes", *Pattern Recognition Letters,* No. 24, 2003, pp. 1715-1723.

[4] A.W. Senior, "Face and feature finding for a face recognition system", *Proc. 2nd International Conference on Audio- and Video-based Biometric Person Authentication*, Washington D.C, 1999, pp. 154-159.

[5] E. Angelopoulou, R. Molana, K. Daniilidis, "Multispectral Skin Color Modeling", Technical Report MS-CIS-01-22, June 22, 2001.

[6] R. Gross, V. Brajovic, "An Image Preprocessing Algorithm for Illumination Invariant Face Recognition", *4th Intl Conf. On Audio- and Video-Based Biometric Person Authentication*, Guilford, UK, 2003, pp.10-17.

[7] M. Störring, H.J. Andersen, E. Granum, "Estimation of the Illuminant Colour from Human Skin Colour", *4th Intl. Conf. On Automatic Face and Gesture Recognition*, Grenoble, France, 2000, pp. 64-69.

[8] M.S. Gazzaniga, *Cognitive Neuroscience, the Biology of the Mind*, W.W. Norton, New York 2002.

[9] A.C. Hurlbert. *The Computation of Color*. PhD thesis, Massachusetts Institute of Technology, Sept. 1989.

[10] R.L. Hsu, M. Abdel-Mottaleb, A.J. Jain, "Face Detection in Color Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, May 2002, pp. 696-706.

[11] C. Sanderson and K. K. Paliwal, "Polynomial Features for Robust Face Authentication", *Proc. IEEE International Conference on Image Processing*, Rochester, 2002. pp. 997-1000 (Vol. 3).