Face detection in intelligent ambiences with colored illumination

Christina Katsimerou, Judith A. Redi, Ingrid Heynderickx

Department of Intelligent Systems
TU Delft
Delft, The Netherlands

Abstract. Human face detection is an essential step in the creation of intelligent lighting ambiences, but the constantly changing multi-color illumination makes reliable face detection more challenging. Therefore, we introduce a new face detection and localization algorithm, which retains a high performance under various indoor illumination conditions. The method is based on the creation of a robust skin mask, using general color constancy techniques, and the application of the Viola-Jones face detector on the candidate face areas. Extensive experiments, using a challenging state-of-the-art database and a new one with a wider variation in colored illumination and cluttered background, show a significantly better performance for the newly proposed algorithm than for the most widely used face detection algorithms.

Keywords: intelligent ambiences, adaptive lighting, face detection, skin segmentation, color constancy

1. Introduction

Intelligent ambiences have embedded the advances of human-technology interaction in arrays of smart sensors and actuators, in order to achieve a natural unobtrusive communication with the user [1]. A specific application of an intelligent ambience we are interested in, aims at building an automatic system which assesses the mood of a person in a room from videos and responds adaptively to it with (multi-color) light settings, which are believed to improve the well-being of the room's occupant [18]. The first and most essential step towards such a system is automatic face detection from the video input, since the face conveys highly relevant emotional information. Therefore, our research focuses on a robust face detector that can handle the challenges of head pose variation, colored illumination, cluttered background and low resolution.

Ambiences with colored artificial light present multiple challenges for face detection. First, these ambiences are created with different light sources positioned in various locations in the room. As a consequence, light on the face is never uniform, but rather unevenly distributed, producing reflections or shades. In addition, colored light sources often alter the skin color captured in the images significantly. Finally, certain ambiences are composed with low intensity illumination, making the faces hard to distinguish.

A number of robust face detection algorithms have been proposed in the literature, with the most recent ones reviewed in [24]. The main concept of the latest appearance-based approaches is that they collect a large number of positive and negative samples (face and non-face patches) from images, extract features from the intensity component of these patches and feed them as input to a classifier, which is trained to distinguish face from non-face patches. Among these approaches, the Viola-Jones (VJ) algorithm [23] has typically been preferred for detecting upright frontal faces, due to its simplicity and effectiveness.

A general shortcoming of the appearance-based methods is that a high detection rate may result in a large number of false positives, namely non-face patches that are incorrectly recognized as faces. Especially a complex background increases the chance of misclassification. To reduce false positives, complementary information, like skin color, was used in [7, 12], however, without a light compensation technique. As a result, these algorithms missed skin regions, and as such, yielded a lower detection rate than the basic VJ detector. A skin color detector was used in [8] after the Viola-Jones module, in order to filter the correct detections. In this way, the authors achieved a low false positive rate, but did not improve the detection rate.

This study proposes a refined, computationally inexpensive and human inspired face detection method from color images. The proposed algorithm combines robust skin segmentation (applied on a color-corrected image with an optimal color constancy technique) with the publicly available VJ face detection framework [11].The skin segmentation module can successfully detect skin pixels under colored, multidirectional and uneven lighting. The uniform background produced around the skin areas after segmentation can outbalance the clutter and improve the final detection rate, while the localized search of VJ only on the face-candidate regions reduces the false positives significantly. Due to a lack of suitable face databases for evaluating our face detection algorithm under ambient colored illumination, we introduce a new challenging face database, which we named "CI" for Colored Illumination.

2. Overview of the refined face detection algorithm

In order to choose a suitable face detection algorithm under colored ambience light conditions, we tested the VJ on a subset of the state-of-the-art PIE face database [20]. Preliminary results on 100 images of frontal faces under blue-flash illumination indicated that the VJ algorithm could not handle the colored illumination optimally (10% missed faces, 8% false positives). Therefore, we propose a refined approach, which enhances the basic VJ with a pre-processing step that robustly detects skin regions (Fig. 1).

The first step in the skin segmentation module performs local histogram equalization. This step is followed by color correction of the image, in order to outbalance its color bias due to the chromaticity of the illumination. We then project the color corrected image from the RGB to YCbCr space and isolate the chromaticity subspace CbCr. After that, we classify the pixels of the image as skin or non-skin and segment the image in skin and background regions. Finally we apply the VJ face detection framework on the selected regions that are likely to contain faces. More details on each step are provided in the rest of the paper.

2.1 The skin segmentation module

Skin segmentation is based on the fact that skin tones form a fairly compact cluster in various color spaces under canonical illumination [16]. Nevertheless, skin segmentation should still be regarded as very challenging due to the intrinsic variability of the skin cluster (ethnicity, indi-

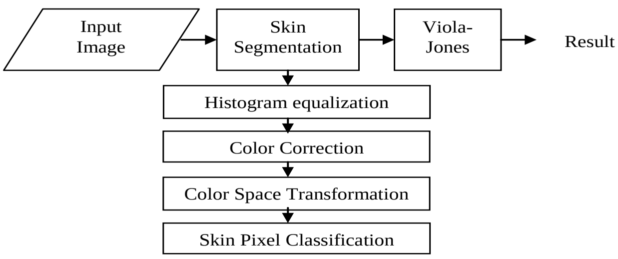

Fig. 1. Workflow of the refined face detection algorithm

vidual characteristics), as well as to the extrinsic variability of images (color overlapping with clutter background, capturing device characteristics, illumination variations), which alter the value of the colored pixels.

To start the segmentation process, histogram equalization is applied on local neighborhoods rather than on the whole image, since the latter enhances shadows on the faces, due to high intensity differences in the image.

2.2 Color Correction

Colored illumination biases the chromaticity of objects in a scene towards the color of the light source. This can severely deteriorate the performance of computer vision systems that recognize objects based on their color, like a skin color pixel classifier. In order to circumvent this issue, we use color constancy, i.e. the human visual ability to perceive the color of an object relatively similar under different illuminants. Similarly, computational color constancy aims at maintaining stable colors in an image, regardless of the color of the illuminant. This is a two-phase process and consists of (a) estimating the chromaticity of an illuminant in a scene from the input image, without prior information about the light source, and (b) adapting the image colors, so that they appear as if the image were illuminated with a canonical (neutral) light. The first problem is commonly solved by means of the general expression [21]

$$I = \frac{1}{k}\left(\iint \left|\nabla^n f_\sigma(x,y)\right|^p dxdy\right)^{\frac{1}{p}}, \tag{1}$$

where $I = (I_R, I_G, I_B)$ is the illuminant color, ∇^n is the n-order derivative, p is the Minkowski norm, and f_σ is the convolution of the image function with a Gaussian filter with scaling parameter σ. Assigning a different triplet of values to the parameters n, p, σ in (1) results in four well-known and widely used color constancy algorithms:

- *Grey World* [4] with $(n,p,\sigma)=(0,1,0)$, which is based on the assumption that the average reflectance in a scene is achromatic.
- *White Patch* [14] or Maximum RGB with $(n,p,\sigma)=(0,\infty,0)$, which is based on the assumption that the maximum reflectance in a scene is achromatic.
- *Shades of Grey* [9] with $(n,p,\sigma)=(0,p,0)$, which is based on the assumption that the pth Minkowski norm in a scene is achromatic.
- *Grey Edge* [21] with $(n,p,\sigma)=(1,p,\sigma)$, which is based on the assumption that the average reflectance of the edges in an image is achromatic.

Once we have estimated the illuminant, we color correct the image using the linear transformation of the von Kries model [22].

In adaptive lighting ambiences, however, there are multiple chromatic light sources of different intensities and in different positions, as mentioned above. This complicates the illuminant estimation. In this study, we benchmark the color constancy algorithms described above in such a multi-color light source environment.

2.3 Color Space Transformation

Color spaces that decouple chromaticity from intensity, such as the YCbCr color space, are considered suitable for color segmentation [6, 19]. Figs. 2 and 3 depict the correlation of Cb and Cr with Y of a skin color cluster, collected from mainly indoor images from the internet, digital cameras and web-cameras. If we look at the figures, we can deduce that the functions Cb(Y) and Cr(Y) remain roughly stable. This implies that both chromaticity components are practically luminance invariant, and as a result, the CbCr space is fairly robust for skin color segmenta-

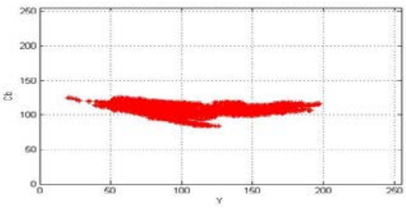

Fig. 2. Skin color cluster in YCb subspace

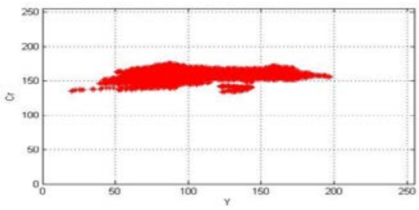

Fig. 3. Skin color cluster in YCr subspace

tion. Consequently, we choose to use only the chromaticity components of the YCbCr color space, without the non-linear transformation that Hsu adopted in [10] to remove the effect of very high or very low luminance on the chromaticity components.

2.4 Skin pixel classification

To detect skin pixels we adopt the naïve Bayesian classifier with histogram technique on the CbCr subspace. A pixel with chromaticity vector *c=(Cb,Cr)* belongs to the skin class if

$$\frac{p(c\,/\,skin)}{p(c\,/\,nskin)} \geq threshold, \qquad (2)$$

where *p(c/skin)* and *p(c/nskin)* are the class-conditional probability distribution functions (pdf) of the skin and non-skin colors, respectively. They are calculated from the occurrence frequency histogram of the skin and non-skin colors. *Threshold* is the value that minimizes the misclassifications to both classes (false positives/false negatives). The threshold theoretically depends on the a priori probabilities of the skin and non-skin classes in the training set; in practice, though, it is determined empirically.

From the output of the skin classifier we create a binary skin mask, assigning 1 to the pixels classified as skin and 0 to the pixels classified as background. Finally, we apply some morphological operators of dilation and erosion to the skin mask in order to include facial features, like the eyes and mouth, and to smooth the binary skin mask. Fig. 4 presents examples of the intensity images after the skin segmentation.

Fig. 4. Intensity images after the skin segmentation module

3. The CI database of faces

To verify the robustness of our algorithm under various illumination conditions, we needed a suitable database of faces. From the state-of-the-art face databases only PIE [20] included colored illumination, providing though only the option of blue flash lighting. To be able to test our approach more comprehensively, we produced a new database of faces, the ‘CI’ (Colored Illumination) database, which suited our particular needs better.

Table 1. Characteristics of the CI database

	Image resolution	*Pose*	*Colored Ambiences*	*No of images*
LQ	1200x1600	frontal side30 side60	neutral cozy 1 cozy 2 activating 1 activating 2 exciting relaxing	1008
HQC	2448x 3264	frontal		336
HQF	2448x 3264	frontal		336

Table 2. Characteristics of each affective ambience

Ambience	*Intensity*	*Dominant Colors*
Neutral	Medium	White
Activating 1	High	Cyan-Blue
Activating 2	High	White-Blue
Cozy 1	Low	Orange-Blue
Cozy 2	Low	Orange-White
Exciting	High	Random colors
Relaxing	Low	Green-Blue

3.1 Characteristics of the database

The face images were captured in the Experience Lab at Philips High Tech Campus under different colored lighting settings. Sixteen subjects with a variety in skin complexion posed for the database, portraying three different expressions: neutral, frown and smile. This resulted in the CI face database, consisting of 1680 RGB compressed images of faces, which can be divided in two main subgroups: 672 high quality (HQ) images, captured by two identical digital cameras, and 1008 low quality (LQ) images captured by 3 identical web cameras. Table 1 summarizes the main characteristics of each subgroup. Examples of the CI database are presented in Fig. 5.

The use of low-end cameras was motivated from real-world applications, like surveillance or human detection in a room, and offered the opportunity to study the impact of lower-quality images on the detection performance. The high-end cameras captured frontal views of the faces from different horizontal distances. The web cameras were positioned at equal distances, but at different viewing angles (i.e., 0^o, 30^o and 60^o from the frontal view). All cameras were synchronized in order to capture the same situation.

3.2 Lighting Settings

For the experiment we used 7 different lighting settings, each encoded in the remainder of the text with an “affective” name, namely neutral, activating (1&2), cozy (1&2), exciting and relaxing. The affective state of the light settings was recognized by subjects in a separate experiment [13]. These names have no further relevance to the purposes of this study and will only serve as a reference. The different lighting settings can be described based on two main characteristics: their average intensity and dominant colors of the light sources used (see Table 2).

4. Experiments and Analysis

4.1 Experimental Setup

We tested our proposed face detection algorithm on the new CI database and on a subset of the PIE face database[20], consisting of 1165 high resolution images of faces captured in different out-of-plane poses under three different types of illumination: neutral room lighting, only blue flash, and blue flash with neutral indoor lighting. We excluded the few completely dark images of the blue flash subset of the database.

The images were first rescaled, so that the size of all faces was approximately 40 x 40 pixels. We then applied local histogram equalization in neighborhoods of 16 x 16 pixels of the downsized images. Subsequently, we color corrected the images with the four color constancy algorithms discussed above, using the Matlab implementation available online [2]. For the Shades of Grey we used p=5 and for the Grey Edge we set $(n,p,\sigma)=(1,5,2)$. We quantized the occurrence frequency histograms of skin and non-skin colors in the CbCr subspace using 64 x 64 bins. The threshold used for minimizing the error of the skin pixel classification was calculated theoretically, according to (2), and slightly adjusted upon trial and error on our training image samples.

In the final stage we applied the VJ face detector on the face-candidate areas of the gray scale images. We used the public domain Haar detectors for frontal [15] and profile face detection [3], whose implementation is available in [17] and performance thoroughly tested in [5]. The frontal classifier consisted of 22 stages of weak classifiers, with the minimum size of the face to be detected 24 x 24 pixels, while the profile classifier consisted of 26 stages, with the minimum size of the profiles 20 x 20 pixels. The response of the detector was a square containing the target with some background pixels. The detection was considered correct if the square box included all facial features and its width was not bigger than 4 times the eye distance. Otherwise, the response counted as false positive. We fused the two classifiers for frontal and profile faces in a cascade structure, only using the profile classifier if the frontal one failed, for the main reason that the former has been proven more accurate than the latter [5]. We separated our test set of images into eight subgroups (four subcategories per database), as illustrated in detail in Figures 6a-6b.

4.2 Performance of the color constancy algorithms

The performance of the algorithms is compared in terms of *accuracy*, defined as

$$accuracy = \frac{DR}{1+FP}, \tag{3}$$

where DR is the detection rate and FP the false positives rate. Table 3 summarizes the accuracy of the four color constancy algorithms, tested on the images with frontal faces of the LQ dataset, indicating the added value of color correction. Concerning the final face detection accuracy, the Shades of Grey color constancy algorithm achieves the highest average result and performs well in all different colored ambiences. The simplest Grey World performs equally well in many cases in terms of detection rate, but it has a higher number of false positives. The White Patch outperforms in ambiences with high intensity, but is unsuitable at medium and lower intensities. Hence, for the rest of the tests we applied skin segmentation after color correction with Shades of Grey**.**

4.3 Comparison of the refined algorithm with other face detection methods

Figure 6 evaluates the performance of the proposed method (Shades of Grey-based skin segmentation previous to the VJ detector) under challenging image quality, pose, and illumination conditions. For comparison, we also report the accuracies of other state of the art face detectors, namely the basic VJ algorithm [23] and Erdem et al.'s method [8], which refines the VJ results with subsequent Grey World-based skin segmentation. The proposed method seems to be more sensitive to pose rather than quality or illumination. The reason for this is that we have used a classifier mainly trained to detect frontal or nearly frontal faces. Strong out-of-plane-rotations that lead to occlusions of facial parts, like one eye, may cause the classifier to fail. Even so, the skin segmentation improves the detection performance with respect to [8, 23]. The second most signifi-

Fig. 5. Examples of the different illuminated ambiences of the CI database; from left to right: a) cozy 1, b) activating 2, c) activating 1(a-c: LQ), d) exciting (HQ F), and e) cozy2 (HQC)

cant factor is the low quality of the web-cameras, as expected. This is partly due to the false positives, resulting from the cluttered background, captured from the larger field of view of the web-cameras. Another reason is that the low-end cameras do not operate color adaptation mechanisms, as opposed to the high-end ones. Challenging multicolor illumination can decrease the accuracy of the VJ detector more than 20%. In these cases, the skin segmentation module can improve the accuracy significantly, up to 10-15%.

Finally, it should be stressed that the moment at which the skin segmentation filter is applied is crucial (Fig. 6a): applying VJ on detected skin regions yields better performance than using a post-VJ skin segmentation filter [8] to refine the results (in terms of FP).

5. Conclusions

We have presented a face detection algorithm that retains performance under colored illumination, using a skin segmentation module before the VJ algorithm. The major contribution of the skin mask is the refinement of the detection by eliminating false positives. A pure appearance based system cannot cope with the false positives very efficiently, because it only uses information from the grey-scale image.

Table 3. Final accuracy of the refined face detection algorithm for different color constancy algorithms. Tests on 336 LQ images with frontal faces of size 40x40

	GW	WP	SG	GE	No CC	VJ
neutral	0,75	0,82	0,78	0,72	0,77	0,53
cosy1	0,51	0,35	0,56	0,35	0,47	0,47
cosy2	0,80	0,71	0,80	0,72	0,74	0,69
act1	0,94	1	0,96	1	0,92	0,77
act2	0,87	0,87	0,89	0,89	0,87	0,66
exc	0,69	0,69	0,69	0,70	0,67	0,61
relax	0,43	0	0,63	0,10	0,02	0,46
average	**0,71**	**0,63**	**0,76**	**0,64**	**0,64**	**0,60**

GW: Grey World, **WP**: White Patch, **SG**: Shades of Grey, **GE**: Grey Edge, **No CC**: no color correction, **VJ**: Viola-Jones without the skin segmentation module

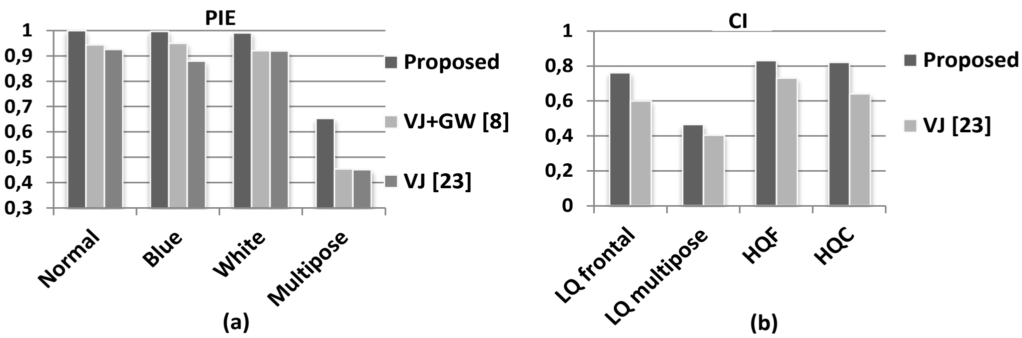

Fig. 6. Accuracy of the VJ algorithm, VJ after skin segmentation module with SG (proposed method) and VJ before skin segmentation with GW [8]: tests on the PIE and CI (comparison only of the first two methods) databases: (a) from the Pie: 867 with frontal faces under indoor illumination, flash and indoor illumination and only blue flash, and 298 with multi-pose faces (b) from the CI: 336 LQ with frontal faces, 672 LQ with multi-pose faces, 672 HQ with frontal faces from the closest and furthest cameras. Faces of size 40x40

This study can be generalized to every appearance-based face detection method in ambiences with colored illumination. More effort can be put on refining the skin mask, taking into account pair-wise dependencies between adjacent pixels. In addition, shape constraints for the face can exclude non-face skin areas (e.g. hands) or skin-like regions (e.g. wooden furniture).

Finally, the parameters of the color constancy algorithms can be adjusted automatically to the different colors of illumination, exploiting prior information and intrinsic properties of the image. Even so, this study highlights the importance of color correction even with an overall optimal color constancy algorithm, before skin segmentation, and quantifies the added value of the latter to the accuracy of an appearance based face detector.

References

1. E. Aarts, Ambient Intelligence – "Visualizing the future", Proc. Conf. Smart Objects & Amb. Intelligence (2005)
2. A. Gijsenij, T. Gevers, http://colorconstancy.com
3. D. Bradley, "Profile face detection", http://www.davidbradley.info/publications/bradley-iurac-03.swf (5/11/2003)
4. G. Buchsbaum, "A spatial processor model for object color perception", J. of the Franklin Institute, 1980, 310(1):1-26
5. M. Castrillón, O. Déniz, D. Hernández, J. Lorenzo, "A comparison of face and facial feature detectors based on the Viola- Jones general Object detection Framework", Machine Vision and Applications, 2011, Vol 22, pp 481-494
6. D. Chai and K. N. Ngan, "Face segmentation using skin color map in videophone applications", IEEE Trans. Circuits and Systems for Video Technology, 1999, Vol. 9, no. 4, 551-564
7. H. Y. Chen, C. L. Huang, C. M. Fu, "Hybrid-boost learning for multi-pose face detection and facial expression recognition", Pattern Recognition, 2008, Vol. 41, no. 3, pp. 1173–1185
8. E. Erdem, S. Ulukaya, A.Karaali, "Combining Haar feature and skin color based classifiers for face detection", IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2011, pp 1497-1500
9. G. Finlayson, E. Trezzi, "Shades of gray and color constancy", in Color Image Conference, 2004, pp 37-41
10. R. L. Hsu, M. Abdel-Mottaleb et al, "Face detection in color images", IEEE transactions on pattern analysis and machine intelligence, 2002, Vol 24, No. 5
11. Intel, Intel Open Source Computer Vision Library, http://sourceforge.net/projects/opencvlibrary/
12. Kim, S. Ban, M. Lee, "Improving adaboost based face detection using face-color preferable selective attention". Lecture Notes in Computer Science, 2008, Vol. 5326, pp. 88–95
13. A. Kuijsters, J. A. Redi, B. de Ruiter I. Heynderickx, "Effects of ageing on atmosphere perception", in the process of being published in Proceedings of Experiencing Light, 2012
14. Land, "The retinex theory of color vision", Scientific American, 1977, 237 (6):pp 108-128
15. R. Lienhart, J. Maydt, "An Extended Set of Haar- like Features for Rapid Objects Detection", IEEE ICIP 2002, Vol. 1, pp. 900-903
16. S. L. Phung, A. Bouzerdoum, D. Chai, "Skin Segmentation using color pixel classification: Analysis and Comparison", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, Vol. 27, no. 1
17. A. Reimondo, Haar cascades repository, http://alereimondo.no-ip.org/OpenCV/34 (2007)
18. R. F. Riemersma-van der Lek, D. F. Swaab, et. al, "Effect of Bright Light and Melatonin on Cognitive and Non-cognitive Function in Elderly Residents of Group Care Facilities: A Randomized Controlled Trial", JAMA, 2008, 299(22):2642-2655
19. K. Sobottka and I. Pitas, "A Novel Method for Automatic Face Segmentation: Facial Feature extraction and tracking", Signal Processing: Image Comm. Vol 12, no. 3, pp. 263-281, 1998
20. T. Sim. S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression (PIE) Database", In Proc. Int'l Conf. Face and Gesture Recognition, 2002, 46-51
21. J. V. de Weijer, T. Gevens, A. Gijsenij, "Edge based color constancy", IEEE Transactions on Image Processing, 2007, 16(9):2207-2214, no2
22. J. von Kries, "Chromatic adaptation", Festschrift der Albrecht-Ludwigs-Universität, 1902 [Translation: D.L. Mac Adam, Colorimetry- Fundamentals, SPIE Milestone Series, Vol. MS 77, 1993], no 3
23. P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features", Proc. of the Conf. on Computer Vision and Pattern Recognition (CVPR 2001)
24. C. Zhang, Z. Zhang, "A survey of recent advances in face detection. Microsoft Research", 2010, 202.114.89.42, 5