# A simple and efficient eye detection method in color images

Dro Desire Sidibe, Philippe Montesinos, Stefan Janaqi

# A simple and efficient eye detection method in color images

D. Sidibe, P. Montesinos, S. Janaqi

LGI2P/EMA - Ales School of Mines

Parc scientifique G. Besse, 30035 Nimes Cedex 1, France

Email:{Desire.Sidibe, Philippe.Montesinos, Stefan.Janaqi}@ema.fr

## Abstract

In this paper we propose a simple and efficient eye detection method for face detection tasks in color images. The algorithm first detects face regions in the image using a skin color model in the normalized RGB color space. Then, eye candidates are extracted within these face regions. Finally, using the anthropological characteristics of human eyes, the pairs of eye regions are selected. The proposed method is simple and fast since it needs no template matching step for face verification. It is robust because it can deal with rotation. Experimental results clearly show the validity of our approach. A correct eye detection rate of 98.4% is achieved using a subset of the AR face database.

**Keywords**: Eye detection, Skin detection, Skin color model, Face detection

## 1 Introduction

Automatic human face analysis and recognition has received significant attention during the past decades, due to the emergence of many potential applications such as person identification, video surveillance and human computer interface. An automatic face recognition usually begins with the detection of face pattern, and then proceeds to normalize the face images using information about the location and appearance of facial features such as eyes and mouth [1], [2]. Therefore, detecting faces and facial features is a crucial step. Many methods for solving the face detection problem have been proposed in the literature (see [3] for a more detailed review) and most of them can be put into a two-stage framework [4]. The first stage focuses attention to face candidates, i.e. regions that may contain a face are marked. In the second stage, the face candidates are sent to a "*face verifier*", which will decide whether the candidates are real faces or not. Different methods put emphasis on one or other of these stages.

Eyes can be considered the most salient and stable features in a human face in comparison with other facial features. Therefore, extraction of eyes is often a crucial step in many face detection algorithms [5], [6]. A recent review on eye detection techniques can be found in [7]. The main classical methods include the template matching method, eingenspace method and Hough transform method [8], [9]. Besides these three classical methods, many other image-based eye detection techniques have been proposed recently. Han et al. [5] use morphological operations to locate eye-analogue pixels in the input image. Then a labeling process is executed to generate eye-analogue segments which are used as guides to search for potential face regions. Finally a trained backpropagation neural network is used to identify faces and their locations. Similar ideas are used by Wu and Zhou [4]. They employ size and intensity information to find eye-analogue segments from gray scale image, and exploit geometrical relationship to filter out the possible eye-analogue pairs. They also use a template matching approach for face cadidates verification. Huang and Wechsler [10] use genetic algorithms to evolve some finite state automata to discover the most likely eye locations. Then optimal features are selected and a decision tree is built to classify whether the most salient locations identified earlier where eyes. Kawaguchi and Rizon [11] use intensity and edges information to locate the iris. The main techniques they use are template matching, separability filter and Hough transform. Song et al. [7] use similar ideas to detect eyes. An improvment of their work is the extraction of binary edge images based on multi-resolution wavelet transform.

In this paper, a simple and robust eye detection method in color images is presented. The proposed method strongly depends on a good face region selector. A skin color model is used to select face regions. Then eyes are directly detected within these regions based on anthropological characteristics of human eyes. The method is simple since it needs no training examples of eyes or faces, and no face verification step. The remainder of the paper is organized as follows. The face region detection is described in Section 2. The eye detection algorithm is addressed in Section 3. Some experimental results showing the validity of the method, are given in Section 4. Finally, concluding remarks are given in Section 5.

## 2 Face region detection

Human skin color is a very efficient feature for face detection. Although different people may have different skin color, several studies have shown that the major difference lies largely between their intensity rather than their chrominance [12], [13]. Many different color spaces have been employed. Among them one finds: RGB, normalized RGB, HSI, HSV, YCbCr, YES, YUV, CIE Lab [6]. Terrillon et al. [14] have shown that the tint-saturation-luma (TSL) space and the normalized RGB space provide best results for Gaussian models. But we can notice, following Albiol et al. [15], that if an optimal skin detector is designed for every color space, then their performance will be same. For that reason, we adopt the normalized RGB color space since it is simple and we model the skin distribution by a single Gaussian.

### 2.1 Skin color modeling

Skin color distribution can be modelled by an elliptical Gaussian probability density function (pdf), defined as:

$$f(c|skin) = \frac{1}{2\pi|\Sigma_s|^{1/2}}\, e^{-\frac{1}{2}(c-\mu_s)^T \Sigma_s^{-1}(c-\mu_s)} \quad (1)$$

where $c$ is a color vector and $(\mu_s, \Sigma_s)$ are the distribution parameters. These parameters are estimated from a training sample. We used a set of 1,158,620 skin pixels, manually selected from about 100 Internet images. The images are chosen in order to represent people belonging to several ethnic groups, and a wide range of illumination conditions.

A more sophisticated model, a mixture model, is often used in the literature [16], [14]. It is a generalization of the single Gaussian and the pdf in this case is the sum of several single Gaussians. The reason why we choose a single Gaussian model is that our experiments have shown that the performance of mixture models exceeds single model's performance only when a high true positive rate is needed (more than 80%). The same observation have been given by Caetano et al. in [17].

### 2.2 Skin detection

Once the parameters of skin color distribution in the normalized RGB color space are obtained from the training sample, we use the Mahalanobis distance from the color vector $c$ to mean vector $\mu_s$, given the covariance matrix $\Sigma_s$ to measure how "skin like" the color $c$ is:

$$\lambda_s(c) = (c - \mu_s)^T \Sigma_s^{-1}(c - \mu_s) \quad (2)$$

Given an input image, for each pixel $x$, $x = (r, g)$ in the normalized RGB color space, $x$ is considered a skin pixel if $\lambda_s(x) \le t$. In our experiments, the threshold value $t$ was chosen to obtain a true positive rate of 80%, while ensuring a false positive rate less than 15%. An example of skin detection result using an image from the AR database is shown in figure 1.

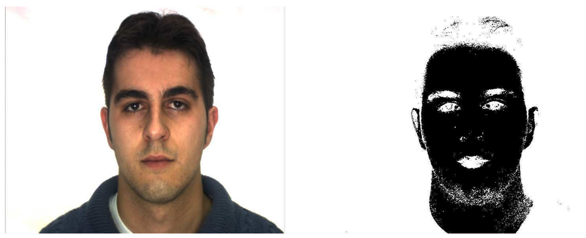

Figure 1: From left to right: original image, skin region detected.

## 3 Eye detection

In [4] and [5], eyes are detected based on the assumption that they are darker than other part of the face. Han at al. [5] use morphological operations to locate eye-analogue segments, while Wu and Zhou [4] find eye-analogue segments searching small patches in the input image that are roughly as large as an eye and are darker than their neighborhoods. In these methods, eye-analogue segments are found in the entire image resulting in a high number of possible pairs to check. On the contrary, in the proposed method, we directly search for eye-analogue segments within the face region. We consider as potential eye regions, the non-skin regions within face region. Obviously, eyes should be within a face region and eyes are not detected as skin by the skin detector. The same ideas are used by Hsu et al. [6]. Therefore, we have to find eye-analogue pairs among a reduced number of potential eye regions (see figure 2).

An ellipse is fitted to each potential eye region using a connected component analysis. Let $R_k$ be a potential eye region and $(x_k, y_k)$ its centroid. Then $R_k$, reduced to an ellipse, defines $a_k$, $b_k$ and $\theta_k$ which are, respectively, the length of the major axis, the length of the minor axis and the orientation of the major axis of the ellipse.

Finally, a pair of potential eye regions is considered as eyes if it satisfies some constraints based on anthropological characteristics of human eyes. Let $R_i$ and $R_j$ be two potential eye regions. Then $(R_i, R_j)$ corresponds to a pair of eyes if the following equations are satisfied:

- $$\begin{cases} 1 < \frac{a_i}{b_i} < 3 \\ 1 < \frac{a_j}{b_j} < 3 \end{cases} \quad (3)$$

- $$|\theta_i - \theta_j| < 20^o \quad (4)$$

- $$\frac{a_i + a_j}{2} < d_{ij} < 3\,\frac{a_i + a_j}{2} \quad (5)$$

The parameters in equation (3) and equation (5) are chosen according to the fact that for human eyes, if we denote by $w_e$ and $h_e$ respectively the width and the height of an eye, the average value for $w/h$ is 2 and averagely $d_{ij} = 2w_e$ [18]. Equation (4) is based on the fact that the two major axis should have the same orientation. A final constraint is the alignment of the two major axis, i.e. for two eye regions they belong to the same line.

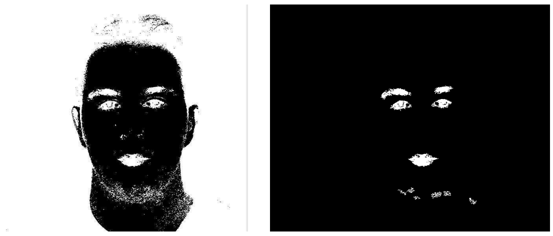

Figure 2: From left to right: skin region detected, potential eye regions.

Using these rules, the algorithm sometimes detects not only eyes, but also eyebrows. To discard regions corresponding to eyebrows, we use the fact that the center part of an eye region is darker than other parts. Then a simple histogram analysis of the region is done for selecting eye regions since an eye region should exhibit two peaks while an eyebrow region shows only one.

# 4 Experimental results

We made different experiments to evaluate the performance of the proposed algorithm. Firstly, we used the AR face database [19] to compare our results with those described by Kawaguchi and Rizon [11], and Song et al. [7]. This database contains color images of frontal view faces with different facial expressions, illumination condition and occlusions. For a direct comparison, we used the same subset of the database employed in [11] and [7]. This subset, named AR-63, contains 63 images of 21 people (12 men and 9 women) without spectacles stored in the first CD ROM. The images in AR-63 show three expressions (neutral, smile and anger) and have natural illumination condition.

Secondly, we used some images gathered from Internet for testing the robustness of the method against complex background, varying illumination condition and rotation.

## 4.1 Evaluation criterion

A commonly used criterion for the performance evaluation of an eye detection method is *the relative error* introduced by Jesorsky et al. [20]. It is defined by:

$$err = \frac{max(d_l, d_r)}{d_{lr}} \quad (6)$$

where $d_l$ is the left eye disparity, i.e. the distance between the manually detected eye position and the automatically detected position, $d_r$ is the right eye disparity, and $d_{lr}$ is the Euclidean distance between the manually detected left and right eye positions. In [4], the detection is considered to be correct if $err < 0.25$. Song et al. [7] defined an other criterion. They considered the detection to be successful if:

$$max(d_l, d_r) < \alpha.r \quad (7)$$

where $r$ is the radius of an iris and $\alpha$ is a scalar factor. Considering that the radius of an iris is about $\frac{1}{4}$ of an eye width, one can see that the criterion of equation (6) is equivalent that of equation (7) with $\alpha = 2$.

## 4.2 Results and discussion

Using the subset AR-63, the proposed method achieves a success rate of 100% based on the criterion defined in equation (6), and a success rate of 98.4% (one failed image) based on the criterion defined in equation (7) for $\alpha = 1$. Some detection results are shown in figure 3 where an eye is depicted by a small white cross.

Comparing the proposed method with those described by Kawaguchi and Rizon [11], and Song et al. [7] using the same set of data, we see that the performance of our method is equivalent to that of the method of Song et al. (98.4% of correct detection), and both methods obtain slightly better results than the method of Kawaguchi and Rizon (96.8% of correct detection). The methods in [11] and [7] can deal with gray scale images but they need to detect the reflected light dots as a cue for eye localization. One main advantage of our method is that we obtain very precise eye localization without the detection of the reflected light dots.

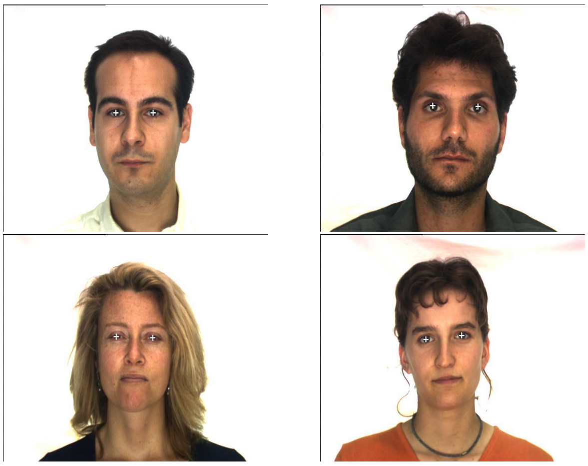

Figure 3: Examples of detected eyes by the proposed method using the subset AR-63.

Figure 4 and figure 5 show some detection results which demonstrate the robustness of the method against rotation and illumination condition. The skin detector is robust enough to deal with different illumination conditions and the algorithm is rotation invariant because we made no assumption about the face orientation for detecting eyes.

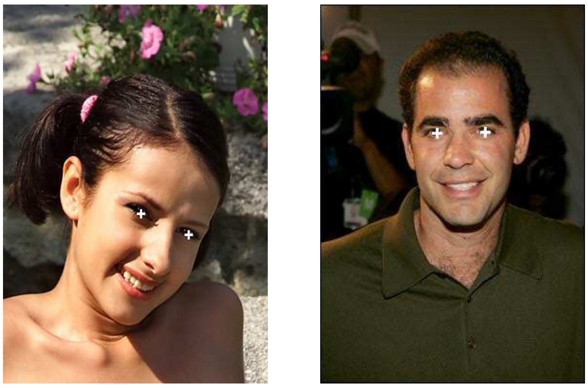

Figure 4: Other examples of eye detection.

One can also notice, figure 5, that the method can be successful when multiple faces are present. Nevertheless, they are some eyes which are not detected in that case. In particular, closed eyes can not be detected.

The most related work to ours is the work of Hsu et al. [6]. They base their face detection algorithm on a robust skin detector too. Then they extract eyes and mouth as facial features by constructing eye and mouth maps based on the luminance and the chrominace components of the image. Finally, they form an eye-mouth triangle for all possible

Figure 5: Example of multiple faces detection.

combinations of the eye candidates and one mouth candidate. Each eye-mouth triangle is verified using a score and the Hough transform. While this method gives good results and may be more robust than ours, we have found that mouth is a less stable feature than eyes since we do not use an explicit mouth or eye map. Moreover, using simple rules to detect eyes, the proposed method is faster than the one described in [6]. The average execution time, given in [6], for processing an image (size 640 x 480) on a 1.7 GHz CPU is 24.71 s. The average time for processing an image (size 768 x 576) on a 3 GHz CPU with our method is 3.8 s.

## 5 Conclusion

In this paper a simple and efficient eye detection method for detecting faces in color images is proposed. It is based on a robust skin region dectector which provides face candidates. Then using some simple rules derived from anthropological characteristics, eyes are selected within face regions. The procedure is robust enough to avoid a face verification system and it achieves a successful rate of 98.4% on a subset of the AR face database. It can also detect eyes in case of rotation and in the presence of multiple faces.

The speed of the method and the robustness to rotation would be very useful for real-time applications. However, experimental results show that the method may failed if one or both eyes are closed and if the face is viewed in profile.

Further improvements can be done for the detection of multiple faces with different orientations and sizes. A multi-scale approach can be used for that.

## Acknowledgements

We would like to thank the anonymous reviewers for their useful comments and suggestions.

## References


[1] S. G. Kong, J. Heo, B. R. Abidi, J. Paik, and M. A. Abidi, "Recent advances in visual and infrared face recognition: a review," *Computer Vision and Image Understanding*, vol. 97, pp. 103–135, 2005.

[2] W. Zhao, R. Chellappa, A. Ronsenfeld, and P. J. Phillips, "Face recognition: A literature survey," *ACM Computing Surveys*, pp. 399–458, 2003.

[3] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 34–58, january 2002.

[4] J. Wu and Z. H. Zhou, "Efficient face candidates selector for face detection," *Pattern Recognition*, vol. 36, pp. 1175–1186, 2003.

[5] C. C. Han, H. Y. M. Liao, G. J. Yu, and L. H. Chen, "Fast face detection via morphology-based pre-processing," *Pattern Recognition*, vol. 33, pp. 1701–1712, 2000.

[6] R. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detecting in color images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 696–706, may 2002.

[7] J. Song, Z. Chi, and J. Liu, "A robust eye detection method using combined binary edge and intensity information," *Pattern Recognition*, vol. 39, pp. 1110–1125, 2006.

[8] R. Brunelli and T. Poggio, "Face recognition: features versus templates," *IEEE Trans. on Pattern Analysis and Machine Intel.*, vol. 15, no. 10, pp. 1042–1052, 1993.

[9] A. Pentland, B. Moghanddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in *IEEE Proc. of Int. Conf. on CVPR*, pp. 84–91, Seattle, Washington, 1994.

[10] J. Huang and H. Wechsler, "Visual routines for eye location using learning and evolution," *IEEE Trans. on Evol. Comput.*, vol. 4, no. 1, pp. 73–82, 2000.

[11] T. Kawaguchi and M. Rizon, "Iris detection using intensity and edge information," *Pattern Recognition*, vol. 36, pp. 549–562, 2003.

[12] H. Graf, T. Chen, E. Petajan, and E. Cosatto, "Locating faces and facial parts," in *Proc First Int'l Workshop Automatic Face and Resture Recognition*, pp. 41–46, 1995.

[13] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection.," *International Journal of Computer Vision*, vol. 46, no. 1, pp. 81–96, 2002.

[14] J. Terrillon, M. N. Shirazi, H. Fukamachi, and S. Akamatsu, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," in *Proc. IEEE Int'l Conf. on Face and Gesture Recognition*, pp. 54–61, 2000.

[15] A. Albiol, L. Torres, and E. J. Delp, "Optimum color spaces for skin detection," in *ICIP*, pp. 122–124, 2001.

[16] H. Greenspan, J. Goldberger, and I. Eshet, "Mixture model for face-color modeling and segmentation," *Pattern Recognition Letters*, vol. 22, pp. 1525–1536, 2001.

[17] T. S. Caetano, S. D. Olabarriaga, and D. A. C. Barone, "Do mixture models in chromaticity space improve skin detection?," *Pattern Recognition*, vol. 36, pp. 3019–3021, 2003.

[18] A. M. Alattar and S. A. Rajala, "Facial features localization in front veiw head and shoulders images," in *IEEE Proc. of ICASSP*, vol. 6, pp. 3557–3560, 1999.

[19] A. M. Martinez and R. Benavente, "The AR face database," Tech. Rep. 24, CVC, june 1998.

[20] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, "Robust face detection using the hausdorff distance," in *Proc. of the Third Int'l Conf. on Audio and Video-based Biometric Person Authentication*, pp. 90–95, 2001.