LIGHTING-VARIABLE ADABOOST BASED-ON SYSTEM FOR ROBUST FACE DETECTION

R. Wood and J. I. Olszewska
School of Computing and Engineering, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK
{*j.olszewska@hud.ac.uk*}

Keywords: Face Detection; AdaBoost; Global Intensity Average Value; Illumination Variation; Lighting Measure.

Abstract: In order to detect faces in pictures presenting difficult real-world conditions such as dark background or back-lighting, we propose a new method which is robust to varying illuminations and which automatically adapts itself to these lighting changes. The proposed face detection technique is based on an efficient AdaBoost super-classifier and relies on multiple features, namely, the global intensity average value and the local intensity variations. Based on tests carried out on standards datasets, our system successfully performs in indoor as well as outdoor situations with different lighting levels.

1 INTRODUCTION

Face detection is a very important and popular field of research in computer vision as it is usually the first step of applications such as automatic human recognition, facial expression analysis, identity certification, traffic monitoring or advanced digital photography. For that, many techniques have been developed based on different features of a human face such as the color of the skin (Gundimada et al., 2004), its texture (Ahonen et al., 2006) or both (Woodward et al., 2010). Some methods rely on motion detection, since the eye are blinking or the lips are moving (Crowley, 1997). Other approaches are based on active contour (Olszewska et al., 2008) techniques to delineate the shape of the face (Yokoyama et al., 1998), the mouth (Li et al., 2006) or the hairs (Julian et al., 2010). Some works consider as facial features characteristic parts of the human face such as eyes (Lin et al., 2008) or ears (Hurley et al., 2008). The feature classification is usually done by means of neural networks (Rowley et al., 1998), Hausdorff distance measure (Guo et al., 2003), AdaBoost algorithm (Viola and Jones, 2004) or support vector machine method (Heisele et al., 2007).

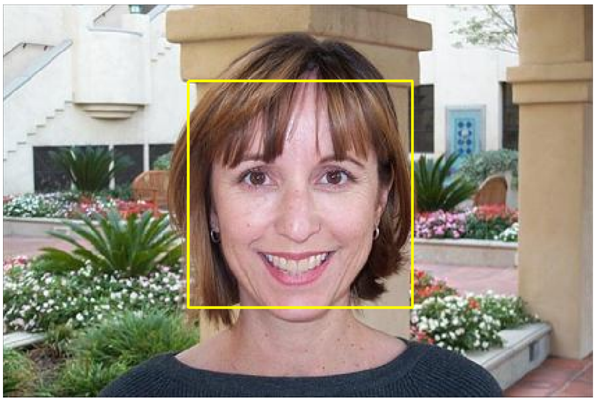
One of the main issues of the face detection techniques is their sensitivity to the lighting variations of the background and/or foreground. For example, methods based on skin color do not perform efficiently in case of foreground darkness in the picture (Zhao et al., 2003). Despite some recent works such as (Sun, 2010) or (Huang et al., 2011) which attempt to improve the automatic process of face detection in still pictures, face detection robust towards illumination changes is still a challenging task (Beveridge et al., 2010).

In this work, we have tackled with face detection in variable illumination conditions. For our purposes, we have developed an innovative approach which automatically adapts itself to these lighting changes in order to increase the robustness of the detection system.

The contribution of this paper is two-fold. Indeed, we present a new super-classifier which is based on two strong classifiers and furthermore, which allows the combined use of two variables. Hence, on one hand, the global image intensity is calculated and, on the other hand, local features such as Haar-like ones are extracted. The proposed super-classifier relying on both these global and local image features leads to the efficient detection of faces in any types of color images, especially in difficult situations with background or foreground presenting low levels of lighting.

The paper is structured as follows. In Section 2, we present our Adaboost-based face detection method which embeds two modalities, namely, the global image intensity and the Haar-like features measuring local changes in intensity. The resulting system that aims to automatically detect faces in images with different illumination conditions has been successfully tested on real-world standard images as reported and discussed in Section 3. Conclusions are drawn up in Section 4.

(a)



(b)

Figure 1: Our system's performance in the case of face detection in images with (a) artificial lighting or (b) day light.

2 LIGHTING-ADAPTABLE FACE DETECTION

Our face-detection system is based on two modalities that are the global image intensity and the local Haar-like features, described in Sections 2.1 and 2.2, respectively. The proposed AdaBoost-based method which combines these two information is explained in Section 2.3.

2.1 Global Image Intensity Average Value

Let us consider a color image $I(x,y)$ with M and N, its width and height, respectively. The conversion of the image $I(x,y)$ to the gray one $I_g(x,y)$ according to the ITU-R BT. 601 norm (Kawato and Ohya, 2000) is as follows

$$I_g(x,y) = 0.299\,R(x,y) + 0.587\,G(x,y) + 0.114\,B(x,y), \quad (1)$$

where R, G, and B are the red, green and blue channels, respectively, of the initial color image $I(x,y)$.

Based on the gray image $I_g(x,y)$, we define the average value $I_{g_{AVG}}$ of the global image intensity as

$$I_{g_{AVG}} = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} I_g(x,y). \quad (2)$$

Hence, in opposition to the other lighting compensation methods, e.g. mentioned in (Gundimada et al., 2004), (Huang et al., 2011), which directly change the pixel values of the original image, we compute a single average value $I_{g_{AVG}}$ of the global intensity of the gray image and we use it as a pivot as explained in Section 2.3. In fact, our approach shows better face detection rates as demonstrated in Section 3.

2.2 Haar-like Features

Local Haar-like features f (Viola and Jones, 2004) encode the existence of oriented contrasts between regions in the processed image. They are computed by subtracting the sum of all the pixels of a subregion of the local feature from the sum of the remaining region of the local feature using the integral image representation $II(i,j)$ which is defined as follows

$$II(i,j) = II(i-1,j) + II(i,j-1) - II(i-1,j-1) + I(i,j), \quad (3)$$

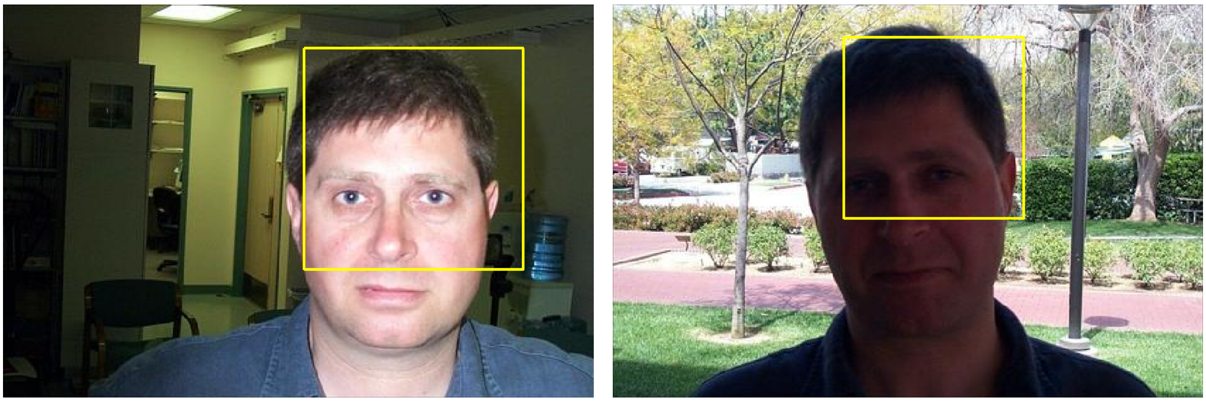
where $I(i,j)$ is the pixel value of the original image at the position (i,j).

2.3 Lighting-Variable AdaBoost Detection (LVAD) System

At first, our LVAD detection system requires a training phase during which it builds strong classifiers based on cascades of weak classifiers.

In particular, to form a T-stage cascade, T weak classifiers are selected using the AdaBoost algorithm (Viola and Jones, 2004). In fact at a t stage of this cascade, a sub-window u of an image from the training set is computed by eq. (3) and it is passed to the corresponding t^{th} weak classifier. If the region is classified as a non-face, the sub-window is rejected. If not, it is passed to the $t+1$ stage, and so forth. Consequently, more stages the cascade owns, more selective it is, i.e. less false positive detections occur. However, that could lead to the increase in the number of false negative detection.

In order to select at each t level (with $1 < t < T$) the best weak classifier, an optimum threshold θ_t is computed by minimizing the classification error due to the selection of a particular Haar-like feature value



(a) (b)

Figure 2: Our system's performance in the case of face detection in images with (a) dark background or (b) backlighting.

$f_t(u)$. The resulting weak classifier k_t is thus obtained as follows

$$k_t(u, f_t, p_t, \theta_t) = \begin{cases} 1 & \text{if } p_t f_t(u) < p_t \theta_t, \\ 0 & \text{otherwise}, \end{cases} \quad (4)$$

where p_t is the polarity indicating the direction of the inequality.

Then, a strong classifier $K_T(u)$ is constructed by taking a weighted combination of the selected weak classifiers k_t according to

$$K_T(u) = \begin{cases} 1 & \text{if } \sum_{t=1}^{T} \alpha_t k_t(u) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t, \\ 0 & \text{otherwise}, \end{cases} \quad (5)$$

where $\frac{1}{2} \sum_{t=1}^{T} \alpha_t$ is the AdaBoost threshold and α_t is a voting coefficient computed based on the classification error in each stage t of the T stages of the cascade (Viola and Jones, 2004).

Next, we introduce the lighting-adaptable strong super-classifier $\mathcal{K}(u)$ defined as

$$\mathcal{K}(u) = \begin{cases} K_L(u) & \text{if } I_{g_{AVG}} > I_{g_{th}}, \\ K_D(u) & \text{if } I_{g_{AVG}} \leq I_{g_{th}}, \end{cases} \quad (6)$$

where $I_{g_{th}}$ is the global image intensity threshold and where D and L (with $D \geq L$) are the numbers of the weak classifiers for dark and light images, respectively.

In this way, the proposed LVAD system allows to automatically select the number of stages of the AdaBoost cascade accordingly to the lighting conditions expressed in eq. (6) by $I_{g_{AVG}}$.

During the testing phase, the LVAD trained system is applied to detect faces/non-faces in a test image set which does not contain any of the images of the training set. Haar-like features are thus extract from these test images according to eq. (3) and classified using the lighting-adaptable strong super-classifier $\mathcal{K}(u)$ as defined in eq. (6). The resulting face detection shows excellent performance as discussed in Section 3.

3 RESULTS AND DISCUSSION

We have tested our system on standard datasets such as (Fei-Fei et al., 2003). We have first trained our classifier on four positive and four negative images with two different numbers of weak classifiers. Then, our system was applied to detect faces in all the images of the database.

Some examples of the performance of our approach for face detection in indoor and outdoor scenes with dim light as well as in case of darkness are presented in Figs. 1 and 2, respectively.

To quantitatively assess the obtained results, the detection rate (Huang et al., 2011) is defined as

$$detection\ rate = \frac{TP}{TP+FN}, \quad (7)$$

with TP, true positive and FN, false negative.

Table 1: Face detection rates.

(Viola and Jones, 2004)	(Huang et al., 2011)	LVAD
91%	94.4%	95%

The results of these experiments are reported in Table 1. The overall detection rate of our LVAD system is 95% and our approach characterized by an adjustable number of weak classifiers is well robust for the different lighting levels present in the dataset images. Moreover, our method outperforms the state-of-the art algorithm of (Viola and Jones, 2004) which owns a fixed number of weak classifiers and the recent work of (Huang et al., 2011) for varying illumination.

4 CONCLUSIONS

In this work, we have proposed an efficient and robust face detection system that uses our strong super-classifier based on Adaboost cascades with an adaptable number of weak classifiers which is depending on the illumination conditions of the captured image.

In the future, we aim to replace in our system the global-lighting average value which computation is currently based on the processed image with a direct real-world lighting measure recorded by sensitive sensors.

REFERENCES

Ahonen, T., Hadid, A., and Pietikainen, M. (2006). Face descritpion with local binary patterns: application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041.

Beveridge, J. R., Bolme, D. S., Draper, B. A., Givens, G. H., Liu, Y. M., and Phillips, P. J. (2010). Quantifying how lighting and focus affect face recognition performance. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 74–81.

Crowley, J. L. (1997). Vision for man machine interaction. *Robotics and Autonomous Systems*, 19(3-4):347–359.

Fei-Fei, L., Andreetto, M., and Ranzato, M. A. (2003). The Caltech - 101 Object Categories dataset. Available online: `http://www.vision.caltech.edu/feifeili/Datasets.htm`.

Gundimada, S., Tao, L., and Asari, V. (2004). Face detection technique based on intensity and skin color distribution. In *IEEE International Conference in Image Processing*, volume 2, pages 1413–1416.

Guo, B., Lam, K.-M., Lin, K.-H., and Siu, W.-C. (2003). Human face recognition based on spatially weighted Hausdorff distance. *Pattern Recognition Letters*, 24(1-3):499–507.

Heisele, B., Serre, T., and Poggio, T. (2007). A component-based framework for face detection and identification. *International Journal of Computer Vision*, 74(2):167–181.

Huang, D.-Y., Lin, C.-J., and Hu, W.-C. (2011). Learning-based face detection by adaptive switching of skin color models and AdaBoost under varying illumination. *Journal of Information Hiding and Multimedia Signal Processing*, 2(3):2073–4212.

Hurley, D. J., Harbab-Zavar, B., and Nixon, M. S. (2008). *Handbook of Biometrics*, chapter The ear as a biometric, pages 131–150. Springer-Verlag.

Julian, P., Dehais, C., Lauze, F., Charvillat, V., Bartoli, A., and Choukroun, A. (2010). Automatic hair detection in the wild. In *IEEE International Conference on Pattern Recognition*, pages 4617–4620.

Kawato, S. and Ohya, J. (2000). Real-time detection of nodding and head-shaking by directly detecting and tracking the "Between-Eye". In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 40–45.

Li, Y., Lai, J. H., and Yuen, P. C. (2006). Multi-template ASM method for feature points detection of facial image with diverse expressions. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 435–440.

Lin, K., Huang, J., Chen, J., and Zhou, C. (2008). Real-time eye detection in video streams. In *IEEE International Conference on Natural Computation*, volume 6, pages 193–197.

Olszewska, J. I., DeVleeschouwer, C., and Macq, B. (2008). Multi-feature vector flow for active contour tracking. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 721–724.

Rowley, H. A., Baluja, S., and Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):23–38.

Sun, H.-M. (2010). Skin detection for single images using dynamic skin color modeling. *Pattern Recognition*, 43(4):1413–1420.

Viola, P. and Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154.

Woodward, D. L., Pundlik, S. J., Lyle, J. R., and Miller, P. E. (2010). Periocular region appearance cues for biometric identification. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 162–169.

Yokoyama, T., Yagi, Y., and Yachida, M. (1998). Active contour model for extracting human faces. In *IEEE International Conference on Pattern Recognition*, volume 1, pages 673–676.

Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458.