FACE RECOGNITION AND GENDER CLASSIFICATION IN PERSONAL MEMORIES

Filipe Grangeiro, Rui Jesus, Nuno Correia

Interactive Multimedia Group, CITI and DI/FCT, New University of Lisbon
Quinta da Torre, 2825 Monte da Caparica, Portugal

ABSTRACT

Face annotation is an important concept for personal memories retrieval. Using automatic face recognition to annotate and find people in those memories provides an improvement of a personal memories management system. However, its results are limited by the uncontrolled conditions inherent to personal memories. In this paper, we propose a face recognition method to address these limitations which includes techniques of skin detection and pose estimation. It is also proposed a gender classification method to provide more information about the detected faces. Experimental results are presented that show that these methods improve the overall performance of face recognition and gender classification in personal memories.

***Index Terms*—** Face Recognition, Gender Classification, Image Processing, Personal Memories.

1. INTRODUCTION

Currently the capture and storage of digital personal memories are a common practice. One of the reasons is the extended usage of digital devices, especially for capturing photos and videos, the most usual way to capture experiences. This causes an increase of the stored information requiring the use of systems that efficiently manage this information. Therefore, automated techniques for organizing this information become desirable. Identifying the people that were present is one way to remember past events but many commercial systems still leave the task of face annotation to the users. For this reason, automatic face detection and recognition are useful tools for a personal memories management system, especially in a large database, because it facilitates image and video retrieval without the need of manual annotation. In this way, it is easier to find, for example, photos with friends, family or co-workers by their name or just photos having faces. Most face recognition algorithms [1, 2] perform well in controlled environments. However, their performance in personal memories decreases because there are more occlusions and more variations in illumination, pose and facial expressions. To overcome these difficulties, Zhao et al. [3] developed a fully automatic framework to annotate people in photos using face detection and recognition, in conjunction with visual context and social context to induce the presence of persons in photos. Contextual features are also used in [4] where face annotation is formulated in a Bayesian framework. This system integrates methodologies of content-based image retrieval and face recognition. A different approach by Zhang et al. [5] describes a semi-automatic system to annotate faces. This system extracts faces from photos, uses them as a user interaction resource to assign an identity and organizes the unassigned set of extracted faces by the similarity to previously labeled faces. Another semi-automatic method was proposed in [6] using color and texture based features, in addition to the face recognition features, to define face similarity.

The gender of the face provides additional information that can improve the retrieval of past events by presenting a more accurate set of personal memories according to the user's choice, allowing a fastest and efficient retrieval. Several approaches have been proposed to address the gender classification problem like [7] or [8]. However, these proposals were evaluated on databases under controlled environment. To deal with uncontrolled databases, in [9] was proposed a system to extract face features to encode information from any zone of the face that can be used to perform gender classification.

In this paper, we propose a system for face recognition and gender classification in personal memories composed by images and videos. The system is based on a face detection method improved by two proposed techniques, skin detection and pose estimation. The rest of the paper is organized as follows. Section 2 gives an overview of the proposed system and section 3 describes the face detection and recognition methods. The next section describes the gender classification algorithm. The paper ends with experimental results and conclusions.

2. SYSTEM OVERVIEW

The proposed system is illustrated in Fig.1 and is described as follows. First, a face detection module is executed using a face detection algorithm described in [10]. To attenuate the problems related with the large variability in personal collections, we added to this module a new skin detection method and a new pose estimation technique. The proposed skin de-

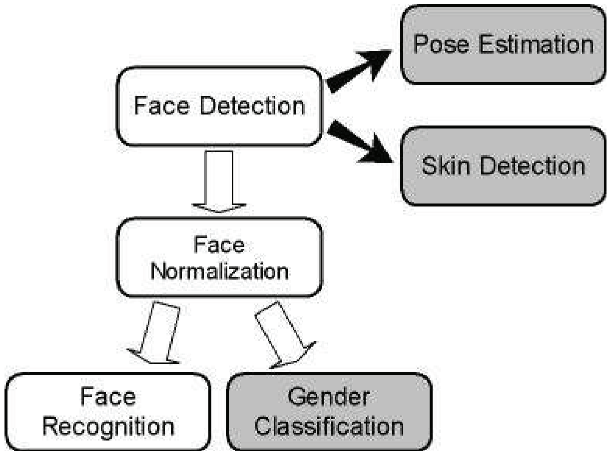

Fig. 1. System architecture.

tection method is based on color and the pose estimation technique in several models for different poses. The next step is the usual face normalization. Finally, a face recognition algorithm is executed using the Principal Component Analysis (PCA) [2] method to perform the dimensionality reduction and Support Vector Machines (SVM) [1] to perform the classification. In parallel to the face recognition module, we also propose a gender classification technique to provide more information about the people in personal memories. The gray blocks represent the proposed methods.

3. FACE DETECTION AND RECOGNITION

The face detection method used in this paper is based on the method proposed by Viola and Jones [10]. The classification is performed by a cascade of classifiers, each one trained by the Adaboost algorithm which computes the strong classifier,

$$h(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where x is an image feature vector, T represents the hypotheses obtained using one single feature, $\alpha_t = \frac{1-e_t}{e_t}$ is the feature weight obtained using the classification error e_t and $h_t(x)$ represents the weak classifier with the lowest error. The cascade of classifiers starts with a simple classifier, increasing the complexity of each stage classifier until the final one that will confirm the presence of a face in a certain image sub-window. The face recognition method used is based on SVM's [1] for binary classification. Given a training set $S_m = \{(x_i, y_i)_{i=1}^m\}$ where labels $y_i \in \{0, 1\}$ and x_i is a feature vector, the decision boundary between two classes (e.g., one person against another person) is obtained by,

$$f(x) = \sum_{i=1}^{m} y_i c_i K(x_i, x) + \beta \tag{2}$$

where m is the number of training images, the coefficients c_i and β are the solutions of a quadratic programming problem, $K(x_i, x) = e^{-\gamma \|x_i - x\|^2}$ and γ is a user controlled parameter. A face is classified in a class if the following condition is satisfied,

$$f(x) \geq th \tag{3}$$

where th is a predefined threshold. If the condition is not satisfied the face belongs to the other class.

To perform multi-class face recognition, we combined two-class SVMs in a bottom-up binary tree scheme based on [1]. The recognized person is the winner of the comparisons between two people starting at the lowest level of the binary tree until the final comparison at the top of the tree. Both face detection and recognition algorithms suffer from difficulties inherent to personal memories. To attenuate them, we propose a skin detection technique to improve face detection and a pose estimation method to improve face recognition.

3.1. Skin Detection

The goal of this method is to improve face detection by developing a skin filter that will confirm or reject the presence of a face in a sub-window of an image indicated by the face detection algorithm. Assuming the face is oriented with the axes and given the coordinates (h, k) of the center point of the detected face, each face is represented by statistics measured from the pixels inside the ellipse,

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \tag{4}$$

where a and b are the semi-major and the semi-minor axes. An image is represented by the vector, $x_c = [\mu_r, \mu_g, \sigma_r, \sigma_g]^T$, where μ represents the mean color (red and green) and σ represents the standard deviation obtained from the pixels inside the ellipse. We use the normalized RGB color space (see [11] for more information about color spaces performance in skin filters) defined as,

$$c_n = \frac{C}{R+G+B} \tag{5}$$

where c_n represents the normalized color component, and C the color component. The x_c vector is used as the feature vector in equation 2 in order to train an SVM to detect the presence or absence of skin color.

3.2. Pose Estimation

Most face recognition algorithms reduce their performance when the face is not in a frontal pose. In personal memories, people are not always in a frontal position to the camera. The proposed method distinguishes faces by its pose and the goal is to improve face recognition by only comparing faces with the same pose. Given a set of images $I = \{I_1, I_2, ...I_n\}$ and the set of labels $y_i = \{frontal, 45^o, profile\}$, the algorithm for pose estimation is defined by the following steps:

1. Face detection on I with AdaBoost algorithm trained only with frontal face images;

2. Face detection on I with AdaBoost algorithm trained only with profile face images;

3. Pose classification:

   - If a face is detected in step 1 and not detected in step 2 then $y_i = frontal$;
   - If a face is not detected in step 1 and detected in step 2 then $y_i = profile$;
   - If a face is detected in step 1 and in step 2 then $y_i = 45^o$.

In this way, to recognize a person instead of having just one SVM to decide if the detected face belongs to that person, our method has three SVMs, one for each pose. The input face will be classified only on the SVM that represents the same pose estimated on that face.

4. GENDER CLASSIFICATION

Gender classification is important to better recognize and categorize a person in a photo or a video. Gender classification techniques in a personal memories context suffers from similar limitations as face recognition methods in the same context. To carry out this task, a similar framework to face recognition is used. First, face detection is performed. Then, face normalization is executed. Finally, PCA is used to reduce the dimensionality of the face representation and the gender classification is given by a trained SVM (see equation 2) where $y_i \in \{male, female\}$ and the feature vector is the output of the PCA method.

5. EXPERIMENTAL RESULTS

The face detection algorithm with the skin detection method were evaluated using the personal collection of one of the authors [12] and consists of 4379 personal pictures with 1688 images having 4609 faces. This database simulates a typical personal memories collection and it is essentially composed by pictures of people, nature or urban scenes, holidays and parties. We refer to Hits as the correct face detections and False Positives as the incorrect face detections. These results will be presented in number of detected faces (#) and percentage (%).

Table 1 shows the obtained results by the face detection algorithm and by the same algorithm with a skin filter. As it is shown, the face detection algorithm has a large number of false positives. The usage of a skin filter reduced this number significantly which is an important result for face recognition and for image retrieval applications. The decrease of Hits was expected because the skin detection method only confirms or rejects already detected faces. It was impossible to increase those numbers without detecting more faces.

	Face Detection		**Face Detection with Skin Filter**	
	#	%	#	%
Detected Faces	14156	-	5037	-
Hits	3435	74.5	2819	61.1
False Positives	10721	75.7	2218	44.0

Table 1. Face detection algorithm evaluation without and with skin filter.

	Face Recognition		
	#	$\hat{\mu}$	S_E
PH	658	74.77	1.78
NH	34393	90.89	0.35

Table 2. Face recognition algorithm evaluation.

The pose estimation and the gender classification methods were evaluated on the Labeled Faces in the Wild database [13] which contains 13233 face images of 5749 different individuals (1490 female and 4259 male). The reason for selecting this database resides on its large set of unconstrained face images with a large range of variations including gender, lighting, facial expression and ethnicity. In our experiments, we used the 10-fold leave-one-out cross validation scheme [11] where each individual have 2 face images in each experience. To evaluate the face recognition and the pose estimation method, only people with 20 or more images were used totaling 54 individuals (10 were used for training) as a result of using 2 face images per person totaling 880 positives and 37840 negative face images comparisons for the 10 experiences. We refer to Positive Hits (PH) as the correct face recognitions of the person that we are trying to recognize and Negative Hits (NH) as the correct rejections of the identity of the person that we are trying to recognize. The results will be presented in number of recognized faces (#), estimated recognition accuracy ($\hat{\mu}$) and its standard error (S_E).

Face Recognition with Pose Estimation				
	#	$\#_{Total}$	$\hat{\mu}$	S_E
PH Frontal	323	657	74.66	1.57
PH Profile	1			
PH 45	333			
NH Frontal	17796	34493	91.15	0.34
NH Profile	120			
NH 45	16577			

Table 3. Face recognition algorithm evaluation modified to take advantage of facial pose information.

Gender Classification						
	#	$\hat{\mu}$	S_E	$\#_{Total}$	$\hat{\mu}_{Total}$	S_{ETotal}
HF	795	79.50	1.16	1628	81.40	0.78
HM	833	83.30	1.01			

Table 4. Gender Classification algorithm evaluation: Hits Female (HF) and Hits Male (HM).

Table 2 shows the obtained results by the face recognition algorithm and Table 3 shows the results obtained by the face recognition algorithm modified to take advantage of facial pose information. As it is shown, the number of positive hits is similar in both methods but the number of negative hits increased when using the facial pose information.

To evaluate the gender classification method, 1490 face images from each gender were used (490 were used for training to adjust the algorithm parameters). The 10-fold leave-one-out cross validation scheme was also used. However, just one face image per person is used totaling 1000 different individuals of each gender divided in 10 experiences. In this way, we will refer to Hits (Female or Male) as the correct gender classification of the gender that we are trying to classify. Table 4 shows the obtained results by the gender classification algorithm. As it is shown, the recognition rate for males and females are relatively similar. We consider the overall classification results positive because the evaluation was performed in a database with photos captured in an unconstrained environment.

6. CONCLUSION AND FUTURE WORK

This paper proposes a system for face recognition and gender classification. There are three major contributions. A skin detection technique to improve face detection, a pose estimation method to increase face recognition and a gender classification method to provide more information about the detected faces. Experiments show that both methods, skin detection and pose estimation, improved the results of the face detection and recognition algorithms and consequently the overall performance of the system. Future work will be focused on providing more information about the detected face with new methods like facial expression analysis and age classification.

7. REFERENCES

[1] G. Guo, S. Li, and K. Chan, "Face recognition by support vector machines," *Proceedings of 4th International IEEE Conference on Automatic Face and Gesture Recognition*, pp. 196–201, 2000.

[2] M. Turk and A. Pentland, "Face recognition using eigenfaces," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586–591, 1991.

[3] M. Zhao, Y. Teo, S. Liu, T. Chua, and R. Jain, "Automatic person annotation of family photo album," *Proceedings of the 5th International Conference on Image and Video Retrieval*, pp. 163–172, 2006.

[4] A. Girgensohn, J. Adcock, and L. Wilcox, "Leveraging face recognition technology to find and organize photos," *Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrieval*, pp. 99 – 106, 2004.

[5] L. Zhang, H. Chen, M. Li, and H. Zhang, "Automated annotation of human faces in family albums," *Proceedings of the 11th ACM International Conference on Multimedia*, pp. 355–358, 2003.

[6] L. Chen, B. Hu, L. Zhang, M. Li, and Zhang H., "Face annotation for family photo album management," *International Journal of Image and Graphics*, vol. 3(1), pp. 1–14, 2003.

[7] S. Baluja and H. Rowley, "Boosting sex identification performance," *International Journal of Computer Vision*, vol. 71(1), pp. 111–119, 2007.

[8] B. Moghaddam and M. Yang, "Gender classification with support vector machines," *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition*, vol. 71(1), pp. 306–311, 2000.

[9] A. Lapedriza, M. Maryn-Jimenez, and J. Vitria, "Gender recognition in non controlled environments," *Proceedings of the 18th International Conference on Pattern Recognition*, vol. 3, pp. 834–837, 2006.

[10] P. Viola and M. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57(2), pp. 137–154, 2004.

[11] V. Vezhnevets, V. Sazonov, and Andreeva A., "A survey on pixel-based skin color detection techniques," *Proceedings of the GRAPHICON*, pp. 85–92, 2003.

[12] R. Jesus, R. Dias, R. Frias, A. Abrantes, and N. Correia, "Sharing personal experiences while navigating in physical space," *5th Workshop on Multimedia Information Retrieval, ACM SIGIR Conference on Research and Development in information Retrieval (SIGIR07)*, 2007.

[13] G. Huang, M. Ramesh, T. Berg, and T. Learned-Miller, "Labelled faces in the wild: A database for studying face recognition in unconstrained environments," *University of Massachusetts, Amherst, Technical Report*, vol. 57(2), pp. 07–49, 2007.