Automatic Face Emotion Recognition System

Jiequan Li and M. Oussalah

School of Engineering
Department of Electronic, Electrical and Computer Engineering
University of Birmingham
{M.Oussalah , axb517 }@bham.ac.uk

Abstract- ***Facial expression recognition has been acknowledged as an active research topic in computer vision community. The challenges include the face identification and recognition, suitable data representation, appropriate classification scheme, appropriate database, among others. In this paper, a new approach for facial emotion recognition is investigated. The proposal involves the use of Haar transform and adaptive AdaBoost algorithm for face identification and Principal Component Analysis (PCA) in conjunction with minimum distance classifier for face recognition. Two approaches have been investigated for facial expression recognition. The former relies on the use of PCA and K-nearest neighbour (KNN) classification algorithm, while the latter advocates the use of Negative Matrix Factorization (NMF) and KNN algorithms. The proposal was tested and validated using Taiwanese and Indian face databases.***

Keywords: Face recognition, PCA, NMF, facial expression recognition.

I. INTRODUCTION

Since the emergence of video surveillance applications among other security and related issues including non-intrusive biometric applications, face detection and recognition becomes an intensive research area in image processing [4, 6, 22]. Some of the challenges in face detection arise from the highly dynamic and non-rigid nature of human faces. Besides, the different face poses, facial expressions, occlusions, varying illumination conditions and cluttered backgrounds make the problem very complex and in some cases unsolvable. It becomes therefore very challenging to detect human faces in complex surveillance environments regardless of their locations, poses, scales, resolutions and lighting conditions. In the literature, one may distinguish two main streams for this purpose [6]. The first one corresponds to feature-based like approach where a set of the face image is decomposed into a set of features, while the second one is based on image-based approach which rather encompasses a classification like method where training and learning were used to identify and/or recognize faces. Obviously hybrid approaches combining feature representation and classification have also been investigated [2-4, 6]. Among the common techniques employed for feature extraction and representation, one notices the principal component employed analysis (PCA), independent component analysis (ICA), Linear discriminant analysis (LDA) [2], nonnegative matrix factorization (NMF) [5, 13, 14]. Also energy based representation like Haar transform and wavelets are commonly used in computer vision community [23]. Usually, these approaches are focused on extracting global face features, and occlusions are difficult to handle. Geometry feature-based methods analyze explicit local facial features, and their geometric relationships. Some examples of these methods are the active shape model [7], the elastic bunch graph matching algorithm for face recognition [24] and the Local Feature Analysis (LFA) [18], eigen faces [10, 15, 21].

However, many appealing proposals in this respect are still too complex for real-time processing as required in surveillance applications, which raises the problem of computational requirement challenge. In [6], it has been stated that face recognition is one of the most challenging problems to be solved in the computer vision community.

Facial expressions provide a key mechanism for understanding and conveying emotion. Meharbian [16] has suggested that the ability for humans to interpret emotions is very important to effective communication, accounting for up to 93% of communication used in a normal conversation. Consequently, automatic recognition of facial emotion has shown to be a key issue in developing more robust and reliable communications. Ekman and Freisen [9] have been pioneers in this area, helping to identify six basic emotions (anger, fear, disgust, joy, surprise, sadness) that appear to be universal across humanity. Much of expression research to date has focused on understanding how underlying muscles move to create expressions [8-10, 17]. From classification perspective, given the facial feature representation, many algorithms have been developed for the purpose of facial expression identification. This includes support vector machines, neural network, k-nearest neighbour, neural network, fuzzy logic, among others. On the other hand, one of the challenges in facial recognition is the collection of suitable database. Despite the existence of some publicly available databases [1, 19, 25], the quality of the training and classification is always limited by the quality of such databases.

In this paper, an automated system for facial expression recognition is developed. The proposal uses row images as input where a first step of face identification and recognition is conducted. This builds on the open source FAINT system [11] and OpenCV package where Haar features were used for data representation and Adaptive AdaBoost algorithm [12] for classification. While PCA and Minimum Distance Classifier (MDC) were used for face recognition. Next, for the purpose of facial expression recognition two approaches have been compared and investigated; namely PCA and NMF methods in conjunction with KNN classifier. Two databases were used: Indian face database [21] and TFEID Taiwanese face database [19].

II. Face Detection and Recognition System

The face detection function is embedded in the open source Faint project –a java framework for face identification and recognition- [11]. The implementation makes use of Java Native Interface (JNI), which enables Faint to call the C++ implemented OpenCV face detection program directly from Java side. The OpenCV face detection algorithms extract Haar-like features, which encode the existence of oriented contrasts between regions of images by quantifying the difference between the total of dark region and light region. Next AdaBoost classifier [12, 26], which favours incorrectly classified patterns to be used in subsequent component classifier, was used to classify the feature as face or not. For this purpose, several positive examples (face images) and negative examples (arbitrary images) are used for training purpose. Figure 1 summarizes the face detection approach.

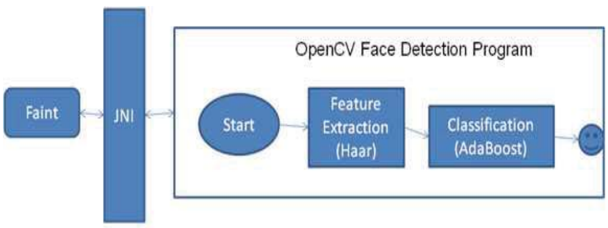

Fig. 1 General scheme of Face Detection System

The face recognition function is also embedded in Faint project. It uses the face image that is detected using face detection function as input, then uses Principal Component Analysis (PCA) to extract face-features. The latter is determined by projecting the image vector to the first **r** eigenvector space as will be detailed later on. Finally, Minimum Distance Classifier (MDC) is employed to classify the various face images. The latter calculates the distance with respect to PCA features from each face image of the training database to the request face image and the image face corresponding to the minimum distance is selected. A normalized value of the MDC constitutes the face recognition rate. Figure 2 shows the basis of the face recognition system.

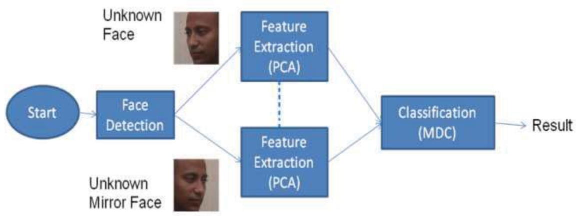

Fig. 2. General scheme of face recognition system

Figure 3 shows an example of the face recognition result using the above implementation. The dialog on the left side shows the top 9 highest recognition score of this face image. The dialog on the right is the corresponding face images of the 9 recognition score. We can see that face recognition system has found the right person.

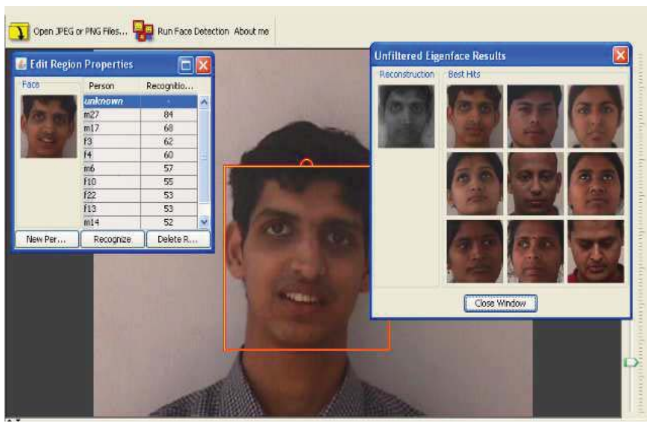

Fig. 3 Example of face recognition running screenshot

III. Facial Expression Recognition System

A. General scheme

In the same spirit as the aforementioned face identification and recognition system, the developed facial expression system encodes a feature extraction module where two distinct features are compared;

namely, the very common PCA and the negative matrix factorization (NMF) algorithm where the user can run either of the algorithms. Next, a classification algorithm using K-Nearest Neighbour algorithm is used to determine the appropriate facial expression. The latter consists of one of the five states (classes) describing emotions: Neutral, Happiness, Surprise, Sadness and Disgust. Figure 4 summarizes the different steps of the facial expression system.

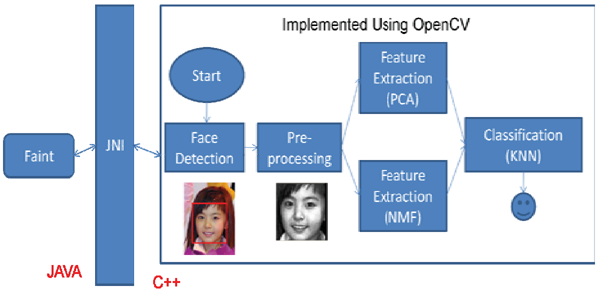

Fig. 4 General Scheme of facial expression recognition system

Figure 5 demonstrates an example of the facial expression recognition result. The user has a possibility to choose between either PCA or NMA methods and can also specify the value of the parameter k (number of neighbours) in the KNN classification algorithm. The interface also allows the user to add/delete images to (from) training database. The information of the face image's K nearest neighbors is also presented for debugging purpose.
After the facial expression is recognized, user can attached the facial expression to the image as a tag. User can also add any tags they want to attach to the image. A graphic database management tool is provided to manage all the tags stored in the database. Users can check all the images under a certain tag, thus it's a good tool for photo organization and management.

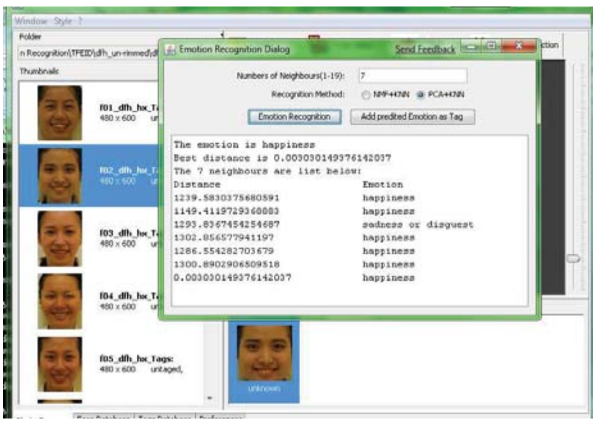

Fig. 5 Example of Running screenshot of facial expression recognition

B. PCA used as feature extractor

Principal Component Analysis is an optimal linear dimensionality reduction scheme with respect to the mean squared error (MSE) of the reconstruction. Typically, PCA can be obtained by solving the eigenvalue problem for covariance matrix of data where the first principal component is the eigenvector corresponding to the largest eigenvalue, while the second principal component is that corresponding to the next eigenvalue. More formally, given a set of samples $x = (x^1, x^2, x^3, ..., x^N)$, the projection matrix E composed of the K eigenvectors of variance with highest eigenvalues, the K-dimensional representation of an original, n-dimensional vector x, is given by the projection

$$y = E^T (x - \mu) \tag{1}$$

where

$$E = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T \tag{2}$$

and

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x^i \tag{3}$$

correspond to the variance and mean of the samples x^i (i=1 to N), respectively.
The choice of optimal value of the number of eigenvalues is still an open issue. In this course, one uses the usefulness rate of the first r eigenvalues λ_i given as

$$l = \frac{\sum_{i=1}^{r} \lambda_i}{\sum_{i=1}^{N} \lambda_i} \tag{4}$$

So, the number of eigenvalues (or equivalently eigenvectors) is chosen such that the ratio l is greater than some threshold, say 85%, which would allow us to stipulate that most of the information conveyed by the original N (training) samples or images are preserved in the new matrix constituted by the first r eigenvectors.
Therefore, the obtained first r eigenvectors of the covariance matrix form a "face space":

$U = [u_0, u_1, .., u_{r-1}] \in R^{N\,x\,r}$

A new face image Y can be projected onto the "face space" by a simple operation

$$Y = U^T X \tag{5}$$

So this is the feature vector extracted from the new face image, it's an **r** dimension vector. We can found that the dimension of the feature vector is much smaller than the dimension of the image (r < M < N), which makes the classification a lot easier.

C. Non-negative Matrix Factorization (NMF)

NMF is a method to obtain a representation of data using non-negativity constraints. It consists of a matrix factorization algorithm that decomposes the initial matrix n x m V as a product of two non-negative matrices W and H:

$$V \approx W.H \ \text{ with } V_{ij} \approx \sum_{k=1}^{r} W_{ik} H_{kj} \tag{6}$$

so that the objective function

$$J = \|V - W.H\|^2 \tag{7}$$

is minimized.

The dimensions of the factorized matrices W and H are $n \times r$, and $r \times m$, respectively. Usually, r is chosen so that $\boldsymbol{(n+m)r < nm}$. Each column of matrix W contains a basis vector while each column of H contains the weights needed to approximate the corresponding column in V using the basis from W.

It is has been shown [5, 13-14] that the solution to the above optimization problem is given by the following recursive rules

$$H_{ij}^{(n+1)} = H_{ij}^{(n)} \frac{\left[V^T W^{(n)}\right]_{ij}}{\left[H^{(n)} W^{(n)T} W^{(n)}\right]_{ij}}, \tag{8}$$

$$W_{ij}^{(n+1)} = W_{ij}^{(n)} \frac{\left[V^T H^{(n)}\right]_{ij}}{\left[W^{(n)} H^{(n)T} H^{(n)}\right]_{ij}}, \tag{9}$$

Besides, the uniqueness of the above is granted only when the Euclidean length of the column vector in matrix **U** is one. This requires the above update rules to follow up by a normalization step as

$$H_{ij}^{(n+1)} = H_{ij}^{(n)} \sqrt{\sum_i \left(W_{ij}^{(n)}\right)^2} \tag{10}$$

$$W_{ij}^{(n+1)} = W_{ij}^{(n)} / \sqrt{\sum_i \left(W_{ij}^{(n)}\right)^2} \tag{11}$$

Where n stands for the iteration number. Initialization is performed using positive random initial conditions for matrices **W** and **H**. Notice that in contrast to PCA where each column of matrix **W** represents an eigenvector and the factorized matrix of **H** represent the eigenvalues, NMF does not allow negative entries in the factorized matrix **W** and **H** permitting the combination of multiple bases images to represent an object.

The Matrix W is the feature space that we used to calculate image feature vector. The image vector **X** is contains only positive value. The feature vector can be calculated by projected image vector to the feature space as $W^T X$.

IV. TESTING AND EVALUATION

A. Testing procedure

We used the Taiwanese Facial Expression Image Database as a testing database (TFEID 2008). The database contains about 40 images on each expression category. The testing focussed on recognizing the four groups of facial expression: neutral, happiness, surprise, sadness and disgust. While the Indian face database [25] was used to test the face recognition system employing Faint opensource. The database contains images of 40 distinct subjects with eleven different poses for each individual. For each individual, it includes the different pose for the face, such as looking front, looking left. In addition to the variation in pose, images with four emotions – neutral, smile, laughter, sad/disgust – are also included for every individual

Here is the procedure of one single test:

i) Randomly choose half the images from each category, that's about 20 images each, and 80 images at total being used as the training set. The remaining 120 images will be used as testing set.

ii) Take some time for the program to train the image. The NMF training method usually takes a long time, up to several hours. The PCA training method is much quicker.

iii) Use the testing set to test the facial expression recognition system.

iv) Repeat 1) to 3) with difference parameter K in the KNN algorithm. K is set to odd value, and changed from 1 to 17.

B. NMF testing result

We have repeated the above testing procedure three times. For each test the parameter K changed from 1 to 17 and the outcomes are averaged, so that the effect of K to the recognition result will be minimized. The parameter r in NMF is set to 15.

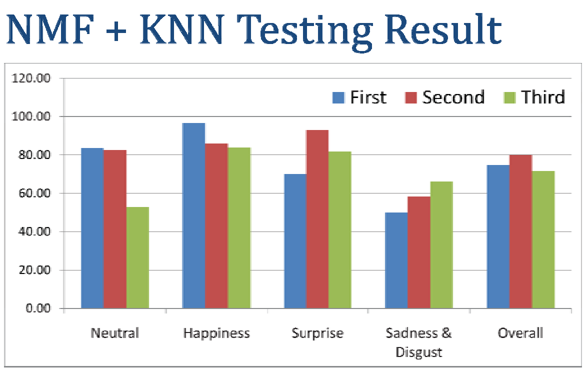

Fig. 6 NMF testing result

We can observe from Figure 6 that the NMF method has a high recognition rate on the expression category "Happiness" and "Surprise", more than 80% on average. But it has a relatively low recognition rate on “Sadness & Disgust” expression category, only 60% on average.

C. PCA testing result

Two separate tests are performed using PCA. For each testing the parameter K changed from 1 to 17. As it can be noticed from Figure 7 the recognition rate of PCA method for each category is always greater than 70%, which is quite stable.

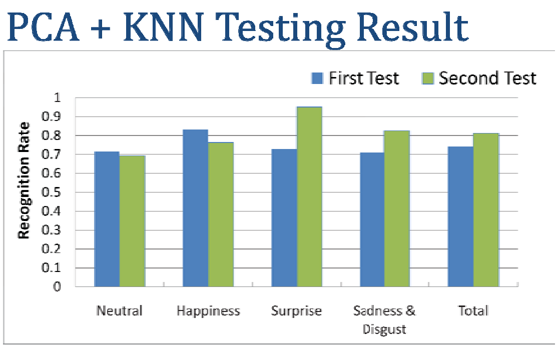

Fig. 7 PCA recognition result

By averaging the various test results, Figure 8 highlights the comparison between the NMA and PCA recognition rate for each facial expression. It can be noticed that the total recognition rates are almost the same, about 75% for both methods, while PCA outperforms NMF on “Surprise” and “Sadness & Disgust” categories. But NMF is has better recognition rate on expression category “Neutral” and “Happiness”. In terms of computational time, PCA is few order of magnitudes faster than NMF approach.

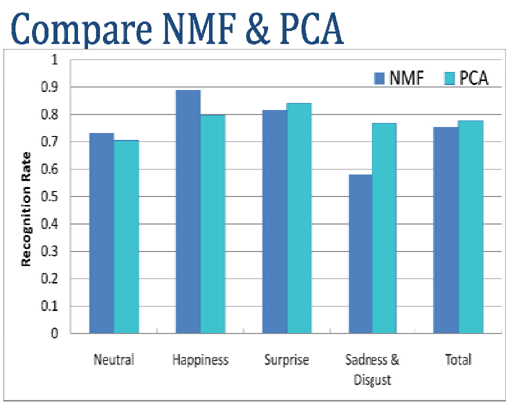

Fig. 8 Comparison between NMF and PCA recognition rates

On the other hand, in order to evaluate the effect of parameter K in KNN algorithms, the experiment is repeated for various values of K. As it can be shown in Figure 9, the best performance in terms of global recognition rate is obtained for K= 5, 6, 7 in case of PCA and 9 in case of NMF.

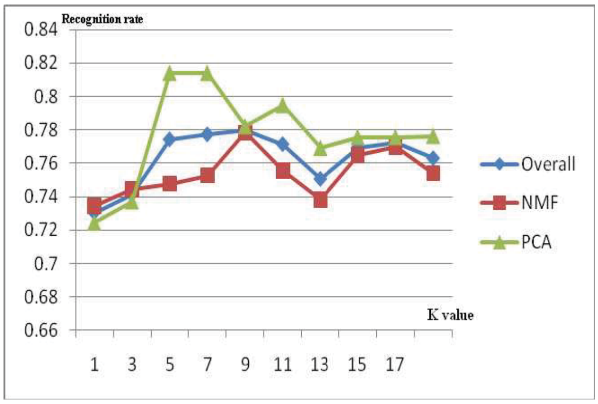

Fig. 9 KNN testing result

Typically, it was reported elsewhere that NMF has advantage on reducing data redundancy, which makes it more suitable to use with a large static training database. While PCA has advantages on the short training time, which makes more suitable to use with a dynamic training set.

Strictly speaking, the results pointed out in figures 6-9 are not meant to provide a final judgment regarding the performance of NMF versus PCA as the advantages of NMF method from theoretical and practical viewpoints are well documented elsewhere, see, for instance [13]. Rather, one aims to highlight the performance of the approaches with respect to the limited database used in this experiment. Indeed, the diversity of the face images in both Indian and Taiwanese databases is always questioned. Besides, the use of faces images with various poses arise a great challenge to the face identification process, which obviously renders the performance of the training limited. From this perspective, the overall recognition rate of 75% sounds quite a great achievement as many alternative studies reported in the literature have achieved lower recognition rate, see, for instance, [10] and references therein.

V. CONCLUSION

Throughout this paper, we have investigated an automated system for face identification and recognition and then recognition of facial expressions. The system makes use of open source FAINT system for face identification. Two approaches have been compared. The former is based on PCA approach while the latter uses the NMF algorithm for feature representation. In both cases, a KNN algorithm is used for classification purpose. Two publicly available face images were used. The results show that both approaches achieve a recognition rate over 75% on average, which demonstrates the feasibility of the proposal. It has been shown that NMF better performs reduction of data redundancy while the PCA has a lower computational complexity, which makes it suitable for dynamic training databases

REFERENCES

1. AR face database http://www.isbe.man.ac.uk/~bim/data/tarfd_markup/tarfd_markup.html, viewed on April 2009
2. M. Black and A. Jepson. Eigentracking : Robust matching and tracking of articulated objects using a view-based representation. *IJCV*, 26(1):63–84, 1998.
3. I. Biederman, Recognition-by-components: a theory of human image understanding. Psychol. Rev. 94, 1987, 115–147.
4. C. M. Bishop, Pattern Recognition and Machine Learning, Information Science and Statistics, Springer, 2008
5. I. Buciu, Non-negative Matrix Factorization, A New Tool for Feature Extraction: Theory and Applications, *Int. J. of Computers, Communications & Control*, vol. III (2008). Proceedings of ICCCC 2008, pp. 67-74.
6. R. Chellappa, C. L.Wilson and S. Sirohey, Human and machine recognition of faces: A survey, Proceedings of IEEE, 83(5), 1995, 705-741.
7. G.J. Edwards, T.F. Cootes, and C.J. Taylor. Face recognition using active appearance models. In Proceedings of the European Conference on Computer Vision, 1998, 581-595
8. P. Ekman. Emotions Revealed: Recognizing Faces and Feeling to Improve Communication and Emotional Life. Holt, 2003
9. P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, 1978.
10. B. Fasel and J. Luettin. Automatic facial expression analysis: A survey. Pattern Recognition, Vol. 36(1), 259-275, 2003.
11. Faint, *faint - The Face Annotation Interface*, viewed 1 April 2009, <http://sourceforge.net/projects/faint>
12. Y. Freund and Robert E.Schapire , A decision-theoretic generalization of online learning and application to boosting, *Journal of Computer and System Sciences*, no. 55. 1997, 119-139.
13. D. Lee and H. Seung. Learning the parts of objects by nonnegative matrix factorization. *Nature*, *401:788–791, 1999.*
14. D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *Advances in Neural Information Processing Systems*, volume 13, pages 556–562, 2001.
15. A. Leonardis and H. Bischof. Multiple eigenspaces by mdl. In *Proc. ICPR*, *volume 1, pages 233–237, 2000.*
16. A. Mehrabian. Communication without words. Psychology Today, 2(4), 1968, 53-56.
17. Palmer, S. E. Hierarchical structure in perceptual representation. Cogn. Psychol. 9, 441–474 (1977).
18. P. S. Penev and J. J. Atick, Local feature analysis. A general statistical theory for object representation, Networks. Computation in Neural Systems 7(3), 1996, 477-500
19. TFEID 2008, *Taiwanese Facial Expression Image Database*, viewed 1 April 2009, <http://bml.ym.edu.tw/~download/html/>
20. F. Torre and M. Black: Robust principal component analysis for computer vision. In *Proc. of ICCV'2001*, *volume 1, pages 362–369, 2001.*
21. M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Neuroscience*, 3(1):71–86, 1991.
22. E. Wachsmuth, Oram, M. W. & Perrett, D. I. Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. Cereb., Cortex 4, 1994, 509–522.
23. A. Webb. *Statistical Pattern Recognition*. Oxford University Press, New York, 1999.
24. L. Wiscott, J. M. Fellous, N. Kruger and C. V. Malsburg, Face recognition by elastic bunch graph matching, Intelligent Biometric Techniques in Fingerprints and Face Recognition, eds LC Jain et al., CRS Press, 1999, 355-396.
25. J. Vidit, *INDIAN FACE DATABASE*, 2002, viewed on April 2009, <http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase/>.
26. P. Viola, MJ, Rapid Object Detection using a Boosted Cascade of Simple Features, in: *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, 2001, 511-518