IJIRMET

ISSN (Online): 2456-0448

International Journal Of Innovative Research In Management, Engineering And Technology

Vol. 2, Issue 10, October 2017

Enhancing To Detect Face Automatically From Input Videos Using Multiple Algorithms

[1] S.Muralidharan, R.Krishnamurthy

[1] Assistant Professor, Sri Sankara Arts & Science College, Enathur, kanchipuram, Tamilnadu, India

[1] Associate Professor, Sri Sankara Arts & Science College, Enathur, kanchipuram, Tamilnadu, India

Abstract: ***This research work proposes a novel technique for Automatic face recognition (AFR) system using cascaded structures and clustering network. Human beings at a very early age are capable of recognizing the varying facial features, due to the Human Visual System (HVS). But it's difficult to depict the human visual system using computer vision system. The basic idea used in this proposed work is to use divide and conquer method, where we design a particular approach for each processing stage and then embedding the entire strategy for AFR system. In this proposed work, two important factors namely cost efficiency and applied technology for varying characteristics of input image are considered respectively, irrespective of the traditional factors such as accuracy, retrieving rate etc. For facial detection, a heterogeneous cascaded detector based on various features is designed to increase the processing capability and detecting efficiency respectively. For facial feature extraction, sparse graph, component based direct fitting and component based texture fitting methods are used to extract the various features at different orientations.***

Keywords ***: Neural Network, Cascading, Biometric,***

I. INTRODUCTION

Recognizing faces is something that people usually do effortlessly and without much conscious thought, yet it has remained a difficult problem in the area of computer vision, where some 20 years of research is just beginning to yield useful technological solutions. As a biometric technology, automated face recognition has a number of desirable properties that are driving research into practical techniques. The problem of face recognition can be stated as 'identifying an individual from images of the face' and encompasses a number of variations other than the most familiar Application of mug shot identification. One notable aspect of face recognition is the broad interdisciplinary nature of the interest in it: within computer recognition and pattern recognition; biometrics and security; multimedia processing; psychology and neuroscience. It is a field of research notable for the necessity and the richness of interaction between computer scientists and psychologists. The automatic recognition of human faces spans a variety of different technologies. At a highest level, the technologies are best distinguished by the input medium that is used, whether visible light, infra-red or 3-dimensional data from stereo or other range-finding technologies. Thus far, the field has concentrated on still, visible-light, photographic images, often black and white, though much interest is now beginning to be shown in the recognition of faces in colour video. Each input medium that is used for face recognition brings robustness to certain conditions, *e.g.* infra-red face imaging is practically invariant to lighting conditions while 3-dimensional data in theory is invariant to head pose. Imaging in the visible light spectrum, however, will remain the preeminent domain for research and application of face recognition because of the vast quantity of legacy data and the ubiquity and cheapness of photographic capture equipment.

II. AUTOMATIC FACE RECOGNITION PROCESS

The intriguing question of *how to recognize faces* has aroused great interest in scientific research for a long time. Human beings at a very early age are already capable of recognizing familiar faces without any difficulty. This ability shows remarkable *robustness*- variable face appearances. Many studies have been carried out in psychophysics and neuroscience to figure out the brain mechanisms underlying this phenomenon

A typical face-recognition system usually consists of a series of processing stages, as depicted in the following figure. These processing stages are: face detection, facial feature extraction and face identification. Each processing stage forms a sub-system of general face analysis having its own characteristics. Let us now discuss these sub-systems in more detail.

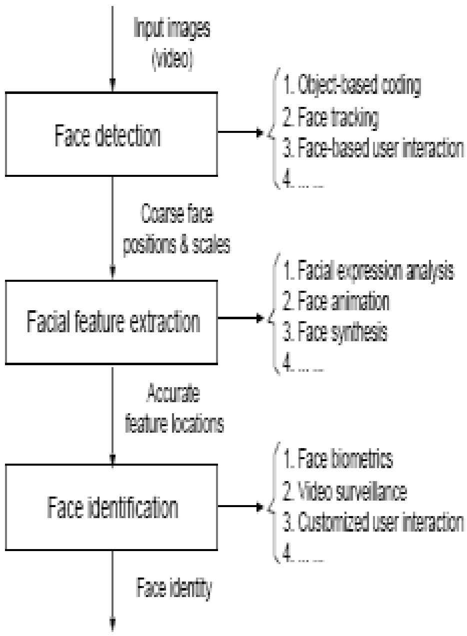

Figure 1.Framework of Face Recognition system

III. PROBLEM ANALYSIS

During the study of this technique, we have found that current techniques for face recognition usually adopt a single algorithm at each stage of the face recognition, which is insufficient to satisfy the performance requirements in real-life applications. To this end, in this thesis we aim at using *multiple* algorithms that coordinate with each other for overall enhanced performance. More specifically, for each major processing stage of face recognition (face detection, facial feature extraction and face identification), we propose a series of novel algorithms and further construct them in a *cascaded* structure.

IV. FACE DETECTION

Face detection is the first processing stage in a face-recognition system. For face detection, generally we need first to find a suitable representation (called feature representation) of the original image/video signals by using multidimensional vectors. Various image/video processing techniques can be applied at this stage to derive an effective feature representation. Based on those feature representations, the second step in face detection is to define a *decision criterion* for separating faces and non-face backgrounds, which should be able to optimally *discriminate* faces from non-face background image signals. In the past decade, extensive research has been carried out on face detection, and significant progress has been achieved to improve the detection performance with respect to two goals: detection accuracy and efficiency.

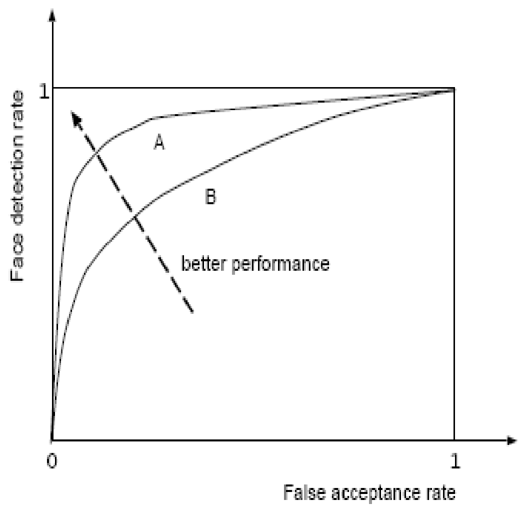

Figure 2.Face Detection Accuracy Curve

4.1. Heuristics-based detectors

Heuristics-based detectors define face characteristics. It's simple, easy to implement and usually do not require much computation cost. we review two well-known techniques that mainly make use of heuristic knowledge about face appearances: colour-based face detection and template-based face detection.

4.1.1. Colour Based detector

In colour-based detectors, an image pixel is viewed as the basic classification unit.Commonly-used colour spaces are normalized *RGB*, *Y CbCr* and *HVS*.A typical classification function used for colour-based detectors is based on the Bayes decision rule

$$\frac{p(x \mid skin)}{p(x \mid non-skin)} > \beta$$

The class-conditional probability densities can be estimated by various techniques, such as histograms, parametric modelling based on uni-modal Gaussian or mixture of Gaussians

$$p(x \mid skin) = (2\pi)^{-d/2} \mid C \mid^{-1/2} e^{-1/2(x-m)^T} C^{-1}(x-m)$$

where C denotes the covariance matrix, m is the mean colour vector and d is the dimension of the colour feature vector.

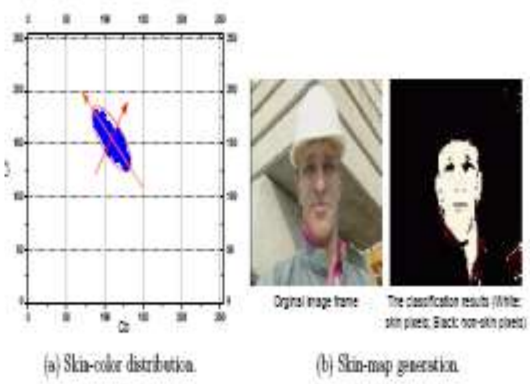

Figure 3. Skin colour classification using uni-modal Gaussian

4.1.2. Template Based detector

Template-based detectors use a pre-defined face template (model) to match possible faces from an input image. Given an image window with raw intensity $I(x; y)$, where $(x; y)$ is the coordinate within the image window, the EOM and EIM of the image window can be derived by the following steps.

For each $(x; y)$, we convolve $I(x; y)$ with 3 * 3 horizontal and 3 * 3 vertical Sobel edge filters, resulting in $Sx(x; y)$ and $Sy(x; y)$, respectively. Based on these convolution results, each element of the EOM ($Eo(x; y)$) is calculated by

$$E_0(x,y) = \arctan \frac{S_y(x,y)}{S_x(x,y)} + \pi / 2 \qquad (3.4)$$

which is further normalized into the range of $[0; \pi]$. Additionally, each element of EIM ($Es(x; y)$) is generated by

$$E_s(x,y) = \sqrt{S_y{}^2(x,y) + S_x{}^2(x,y)}$$

Given a basic classification unit with edge maps $Eo(x; y)$ and $Es(x; y)$, the classification criterion is defined as

$$\sum_x \sum_y d_e(x,y) < \gamma$$

Where γ is an empirically chosen threshold. Distance $de(x; y)$ is defined by

$$d_e(x,y) = \begin{cases} \sin(|E_0(x,y) - \bar{E}_{0(x,y)}|) \rightarrow if\,(E_s(x,y) > T) and (\bar{E}_s(x,y) > T); \\ 1 \rightarrow Otherwise \end{cases}$$

where T is a threshold. Here the difference of the edge orientation is used, because it is less sensitive to variances in image contrast than the edge intensity

4.2. Learning or Statistical based detectors

A statistical or learning-based detector consists of three basic elements as presented in previous sections. The basic classification unit in these detectors is an image window of M *M pixels, which is usually represented by an $M/2$ dimensional vector x containing the raster-scanned intensities of the image window.

4.2.1. Neural-network-based detectors

Neural networks have been demonstrated to be effective tools for much pattern classification problems. Generally speaking, a neural network contains a number of interconnected units (neurons), as motivated by the human-brain structure.

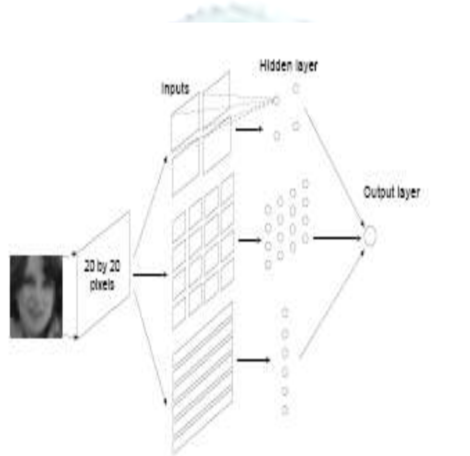

Figure 4 Neural network structures for face recognition system.

the hidden neurons of the network are partially connected to local receptive fields from the input layer, allowing the network to capture local feature structures important for face detection. Three groups of hidden neurons are adopted: 4 neurons look at 10 *10 pixel sub-regions, 16 neurons look at 5 * 5 pixel sub-regions and 6 neurons look at 20 * 5 pixel sub-regions (horizontal stripes). The first two groups of hidden neurons are supposed to capture facial features such as individual eyes, the nose and the corners of the mouth in different resolutions. The third group of neurons (the horizontal stripes) are designed to capture mouths and pairs of eyes. Experiments shows that better detection accuracy can be achieved by incorporating these heuristics to network structures. Furthermore, in order to improve the reliability of the final decision, multiple (2 to 3) networks of the same architecture can be used and the final decision is based on an arbitration of all these networks. The network as depicted in Fig.4 is trained by the back propagation algorithm using a large set of face and non-face training samples.

4.2.2. SVM-based detectors

Support Vector Machine (SVM) is another well-known statistical learning technique for generating complex decision boundaries between face and non-face patterns. SVM was first applied to face detection for face recognition. In this approach, a 19*19 window is used as the basic classification unit, and a large set of such windows containing face and non-face patterns are collected to form a training set.

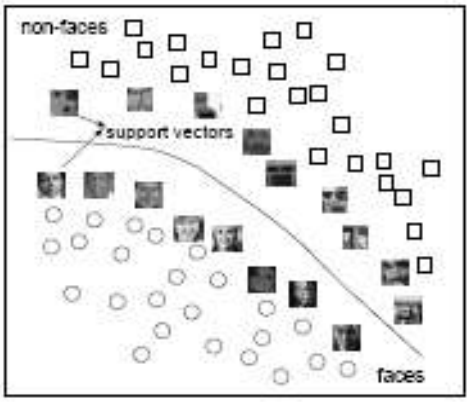

Figure 5 SVM based Face Classification

In order to capture the complex nonlinear decision boundaries between face and non-face classes, nonlinear kernel mapping (e.g. 2nd-degree polynomials) is adopted in the SVM formulation. In Fig.5 a conceptual nonlinear decision surface found by SVM is depicted, together with some example support vectors lying close to the decision surface. Compared to neural-network-based detectors, the training and classification of an SVM-based face detector demand more computation resources.

V. FACIAL FEATURE EXTRACTION

The output of a face detector usually contains coarse indications of positions and scales of existing faces from the input images (video). Given this knowledge, facial feature extraction aims at locating important facial feature structures, such as eyes, nose and mouth. Depending on specific applications, the outputs of facial feature extraction can be the centre locations of facial features. In this thesis, we focus mainly on automated model-based feature extraction, since it offers both accurate description of features and flexibility to adapt to individual feature variances with less human errors

5.1. Model-based algorithms for locating features

The performance of a model-based algorithm is characterized by the following two important aspects, *Capture range*: which includes all parameter initializations that lead to correct convergence. *Extraction accuracy*: indicate how precise the extracted feature locations are. The extraction accuracy is usually expressed by the deviation between a fitted model and the manually-labelled feature positions.

5.1.1 Deformable geometric templates

A promising technique to model common shape characteristics of objects is to use deformable models, which can adjust to individual instances with certain model constraints.

ISSN (Online): 2456-0448
International Journal Of Innovative Research In Management, Engineering And Technology
Vol. 2, Issue 10, October 2017

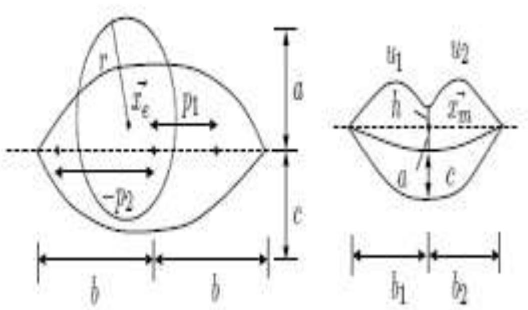

Figure 6 Templates for right eye and sectional view of mouth

In order to find the best parameters for the model to match a given image, the technique first chooses several image representations (valley map, peak map and gradient map) to extract properties such as peaks, valleys and gradients in the gray scale values of the image. Based on these representations, an energy function *E* is then defined as a weighted sum of terms reflecting the knowledge about the feature characteristics. For instance, suppose that *V* (*x; y*) is the valley map of the image, an energy term *Ev* can be defined as

$$E_v = \frac{1}{Area} \iint_{Circle-area} V(x,y) d(x,y) \qquad (4.1)$$

where the integral above is performed over the interior of the eye circle.

5.2. H-ASM: Haar-based shape model

In ASM, the face is modeled by two parts: (1) a set of *key feature points* constituting a face shape, (2) a set of *profile vectors*, each of which characterizes the local texture of a key feature-point.

Algorithm *Model fitting in ASM*
Input: An initialized shape *s* and face image *f*.
Output: Fitted shape s_f .
1. **Repeat**
2. For each feature point (x_i; *yi*) in *s*, search in its neighbourhood in *f* for a local best match (x_{0i}; y_{0i}) based on profile matching, resulting in new shape s_0.
3. Update the parameters (*x; y; r; Φ; b*) to best fit s_0.
4. Apply constraints to each *bi* to ensure that s_0 is a plausible face shape.
5. $s = s_0$
6. until convergence.
7. $s_f = s$
These steps are further illustrated as follows.

1. *Step 1:* The position of each feature point in the current model is updated by searching for a best match in its neighbourhood. The matching is based on the quality of fit of profile vector gs of a candidate point to the mean profile vector $\overline{g}$, which is defined by the following:

$$m(g_s) = (g_s - \bar{g})^T S_g^{-1} (g_s - \bar{g}) \quad (4.4)$$

This is actually the *Mahalanobis* distance between g_s and the model mean $\overline{g}$.
2. *Step 2:* The model parameters are updated to best align to the newly generated shape s_0 after the first step. This is achieved by first estimating the parameters (*x; y; r; Φ*) and then projecting the new shape to the shape space *Φ* to obtain *b*.
3. *Step 3:* The global constraints are applied to each deformation mode b_i in the statistical shape model. By applying limits, to parameter b_i, the generated shape is constrained to a plausible face-like shape.
In Fig. 4.8, we show an example fitting procedure by ASM.
One important limitation of ASM is that only 1-D profile information is used to characterize local textures during the model deformation, which is insufficient to correct large model deviations.

Figure 7 Example model-fitting procedure of ASM.

In order to improve the robustness of the shape model, we proposed to use 2-D texture attributes to characterize each feature point in a shape model. However, since using 2-D textures inevitably increases the computation complexity, we further adopt Haar-wavelets for modelling these local texture attributes, which offers both high processing speed and fitting robustness. In the next subsection, we illustrate the proposed extension of ASM (denoted as H-ASM) in more detail.

VI. FACE IDENTIFICATION

The aim of face identification is to determine the identity of the input face by searching from a database of known individuals. The best match from the database is then returned as the identified face. For some applications, the returned identity is further verified based on e.g. a threshold criterion to determine whether it should be rejected as an ‘unknown’ identity. In this paper, we only consider the strategy to locate the correct best-match. Note that in literature, *face identification* and *face recognition* are usually used interchangeably. In this study, we explicitly use *face identification* to refer to the last

stage of face recognition, excluding the pre-processing steps such as face detection and facial feature extraction.
A considerable part of existing work on face identification employs a fixed classification function to distinguish all individuals (e.g. PCA, LDA, Bayesian or neural network) with a fixed feature representation (e.g. holistic or structural local features). However, using a *fixed* classification function and a *fixed* feature representation is not optimal for *completely* discriminating all individuals, especially for large face databases. In this chapter, we propose a cascaded face-identification technique which aims at higher identification accuracy. At each stage in the identification cascade, we select a subset of the most promising candidates and *dynamically* adapt classification for these selected candidates. Based on the first stage of the identification, the algorithm selects e.g. 50% best-ranked face candidates, and the ground-truth face has a very high probability to be among these selected faces.

6.1. Cascaded face-identification scheme

The cascaded face identification aims at enhanced identification by adopting a coarse-to-fine multistage identification procedure, where a series of classifiers are used to gradually reject unlikely candidates in the gallery database until the best match is found. During each stage, according to the outputs of the current classification, a similarity ranking of the face candidates is derived and promising candidates with high rankings are passed on to the next stage for refined classification

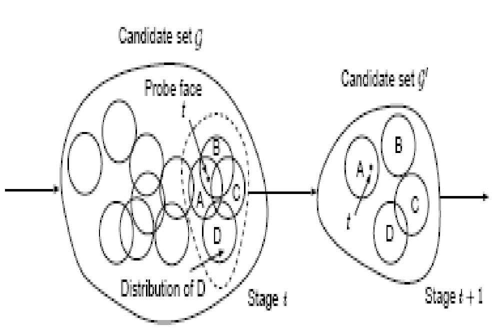

Figure 8 Two intermediate stages in the cascaded face identification.

In the following, we present two implementations of the cascaded face identification scheme.*(i)cascaded LDA algorithm, (ii) cascaded feature selection and classification algorithm.*

(i)cascaded LDA algorithm

It derives different transformed spaces at each stage (adaptive classification) to better separate the most similar faces in the database.

Algorithm *Cascaded LDA*
Input: A gallery set *C*0 and probe face *u*.
Output: Identity of *t*.
1. Compute within-class scatter matrix S_w base on *C*0.

2. Let $C = C0$.
3. for $i = 1$ to *STAGES*
4. Compute between-class scatter matrix S_b^i based on c
5. Find transform matrix W to maximize the following equation as in the traditional LDA formulation

$$\frac{\det(W^T S_b^{(2)} W)}{\det(W^T S_w W)}$$

6. Select candidates with the smallest distance to u under W, which form a new candidate set $C0$.
7. Let $C = C0$.
8. Return the best match found from the last stage.

(ii)cascaded feature selection and classification algorithm.

we propose a cascaded approach with *efficient* dynamic classification. At each stage in the cascade, we select a subset of person-specific (class-specific) discriminative feature. Based on the selected features, a simple feature-matching function is defined to locate the most similar faces in the candidate set. The key advantage of this method is the *dynamic* feature selection, which is adaptive to specific candidate sets and can be carried out efficiently in a cascaded framework.

The Principle of feature selection and classification algorithm is choose a criterion based on Maximum Marginal Diversity (MMD) to evaluate the discrimination capability of each feature component. The marginal diversity refers to the average distance between the class-conditional density of each feature component and its mean.

$$MD_{wk}(x_i) = \frac{1}{C}\sum_{k=1}^{C}\int p(x_i \mid w_k)\log\frac{p(x_i \mid w_k)}{p(x_i)}dx_i$$

where $p(xi)$ is the probability density function (pdf) of xi for all samples and $p(x/\omega_k)$ is the pdf of the component xi for class $!k$. And the ratio between $p(x/\omega_k)$ and $p(xi)$ measures the difference between $p(x/\omega_k)$ and $p(xi)$, which conveys important information concerning the discriminating power of feature component x_i. Motivated by the aforementioned MMD specification, we extend this into a new scheme called *Class-specific Maximum Marginal Diversity* (CMMD). The difference between CMMD and MMD is that MMD selects feature components based on maximum discrimination between *all* classes, whereas CMMD is based on maximum discrimination between *class* ω_k currently considered and all other classes.

Algorithm Estimation procedure for CMMD.

1. Use an N-bin histogram hi to estimate $p(xi)$, where each bin is equally probable.
2. Calculate the span Nk of samples of person ω_k w.r.t h_i, i.e. the number of bins of hi between the smallest sample value and the largest sample value of ω_k.
3. Use h_k^I, consisting these N_k bins to

characterize $p(x/\omega_k)$.
4. Derive $MD_{\omega k}(xi) = \log(N=Nk)$

Complete cascaded algorithm

we present the cascaded face-recognition algorithm based on the CMMD feature selection and the corresponding matching function. The algorithm is a direct application of adaptive feature selection and classification to the cascaded identification framework.

Algorithm *Cascade Feature Set*
Input: A gallery set $C0$ and probe face u.
Output: Identity of u.
1. Let $C = C0$.
2. for $j = 1$ to $STAGES$
3. for all class ω_k in C
4. Select a set of discriminative features for ω_k with the largest CMMD based on statistics of C.
5. Compute the similarity S between u and ω_k
6. Let $C0$ contain a subset of candidates that are most similar to u.
7. Let $C = C0$.
8. Return the best-match found from the last stage.

VII. CONCLUSION

In this paper, we have addressed various topics in face recognition. These topics are divided into three key parts: face detection, facial feature extraction and face identification. For each part, we have reviewed the state-of-the art techniques in literature and proposed a set of novel techniques based on cascaded structures and designed a automatic face recognition system. During the design of each technique, we aim not only at improving the individual algorithm performance, but also at optimizing the algorithm architecture as a way to enhance the overall system performance with reference to accuracy, efficiency and robustness. More specifically, we have proposed three cascaded frameworks for the principal stages in face recognition. In this paper, the applicability of our work for embedded systems in comparison with state-of-the-art literature. First, the algorithms presented in this paper are efficient enough and of sufficiently high quality to initiate experimental embedded system implementations. Second, when looking to literature, it can be concluded that despite the many advances in face-recognition research, most of these advances are not suitable for consumer applications, because of the inherent complexity of the algorithms. For our work, we have taken *efficiency* and *accuracy* as a leading design principle from the beginning, and applied it to all stages of the face recognition. We have been able to make significant progress in improving the efficiency while maintaining the high performance. In this aspect, the work of this thesis is certainly unique and can form an important basis for further embedded application design for face recognition.

REFERENCES

1] W.Y. Zhao, R. Chellappa, A. Rosenfeld, P.J. Phillips, Face Recognition: A Literature Survey, ACM Computing Surveys,Vol:35, 2003.

[2] G. Aggarwal, A. R. Chowdhury, R. Chellappa, A system identification approach for video-based face recognition, in: 17th International Conference on PatternRecognition, Vol. 4, 2004, pp. 175- 178.
[3] B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for Object Representation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 696-710, July 1997.
[4] H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, Jan. 1998.
[5] B. Heisele, T. Poggio, and M. Pontil, "Face Detection in Still Gray Images," A.I. memo AIM-1687, Artificial Intelligence Laboratory, MIT, 2000.
[6] Ti-QiongXu; Bi-Cheng Li; Bo Wang, Face detection and recognition using neural network and hidden Markov models, Proceedings of the 2003 International Conference on Neural Networks and Signal Processing, pp. 228 - 231,2003.
[7] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," Proc. Conf. Computer Vision and Pattern Recognition, pp. 511-518, 2001.
[8] R.L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face Detection in Color Images," Proc. Int'l Conf. Image Processing, pp. 1046-1049, 2001.
[9] Z. Liu and Y. Wang, "Face Detection and Tracking in Video Using Dynamic Programming" Proc. Int'l Conf. Image Processing, 2000.
[10] L. Silva, K. Aizawa and M. Hatori. "Detection and Tracking of Facial Features". Proc. Of SPIE Visual Communications and Image Processing, Taiwan. May, 1995.
[11] M. Han, A. Sethi, and Y. Gong. A detection-based multiple object tracking method. In Proc. Int. Conf. Image Process. (ICIP), pages 3065-3068, Singapore, Oct. 2004.
[12] D. Ramanan and D. Forsyth. Using temporal coherence to build models of animals. In Proc. Int. Conf. Comp. Vision(ICCV), pages 338-346, Nice, France, Oct. 2003.
[13] D. Maio and D. Maltoni. Real-time face location on grayscale static images. Pattern Recognition, 33:1525-1539, September 2000.
[14] Mikolajczyk, K.; Choudhury, R.; Schmid, C. ,Face detection in a video sequence-a temporal approach, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol 2, pp: 96-101, 2001.
[15] Zhenqiu Zhang; Potamianos, G.; Ming Liu; Huang, T., Robust Multi-View Multi-Camera Face Detection inside Smart Rooms Using Spatio-Temporal Dynamic Programming, 7th International Conference on Automatic Face and Gesture Recognition, 2-6 April 2006 Page(s):407 - 412.
[16] J. Yang and A.Waibel. A real-time face tracker. In Proceedings of the Third IEEE Workshop on Applications of Computer Vision, pages 142-147, Sarasota, FL, 1996.
[17] G. Bradski. Computer vision face tracking for use in a perceptual user interface. Technical Report Q2, Intel Corporation, Microcomputer Research Lab, Santa Clara, CA, 1998.
[18] H. Schneiderman and T. Kanade. "A statistical method for 3d object detection applied to faces and cars". In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2000.
[19] J. Feraud, O. Bernier, and M. collobert. "A fast and accurate face detector for indexation of face images". In Proc. Fourth IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 52-59, 1998.
[20] S. Gong, S.McKenna, and J.Collins.
"An investigation into face pose distribution". In Proc. IEEE International Conference on Face and Gesture Recognition, Vermont,1996.