International Journal of Web Technology
Volume: 03 Issue: 01 June 2014 Page No.9-14
ISSN: 2278-2389

Innovative Analytic and Holistic Combined Face Recognition and Verification Method Using Artificial Neural Network

Usha Mary Sharma
Department of Computer Science &Engineeringand IT
Assam Don Bosco University, Assam, India
Email:ushamary.sharma@gmail.com

Abstract - Automatic recognition and verification of human faces is a significant problem in the development and application of Human Computer Interaction (HCI).In addition, the demand for reliable personal identification in computerized access control has resulted in an increased interest in biometrics to replace password and identification (ID) card. Over the last couple of years, face recognition researchers have been developing new techniques fuelled by the advances in computer vision techniques, Design of computers, sensors and in fast emerging face recognition systems. In this paper, a Face Recognition and Verification System has been designed which is robust to variations of illumination, pose and facial expression but very sensitive to variations of the features of the face. This design reckons in the holistic or global as well as the analyticor geometric features of the face of the human beings. The global structure of the human face is analysed by Principal Component Analysis while the features of the local structure are computed considering the geometric features of the face such as the eyes, nose and the mouth. The extracted local features of the face are trained and later tested using Artificial Neural Network (ANN). This combined approach of the global and the local structure of the face image is proved very effective in the system we have designed as it has a correct recognition rate of over 90%.

Keywords: Human Computer Interaction;face recognition;faceverification;biometrics;featureextraction;dimensionalityreduction;global features;geometricfeatures;artificial neural network;Gaussian noise;salt and pepper noise

I. INTRODUCTION

Face recognitionsystem is a computer based security system capable of automatically validating the identity of a person. It is one of the well-known techniquesof Biometrics. Biometrics identifies or verifies a person based on one's physical characteristics by matching the real time patterns against the enrolled ones.The problem of automatic face recognition involves three key steps/subtasks:
(1) Detection andrough normalization of faces,
(2) Featureextraction and accurate normalization offaces,
(3) Identification and/or verification [1]. Here, however, we assume that the detection precedes this recognition process.Face recognition is influenced by many complications, such as the differentfacial expressions, the angleand light directions falling on the face, the cultural diversity, age, illness, differentposturesand size.Sometimes, the images captured for the same person in different surroundings may be unlike.By considering the feature extraction methods and dimensionality reduction techniques in the application of pattern recognition, a number of face recognition systems has been developed with distinct degrees of success. In this paper an artificial neural network(ANN) based face recognition is presented.

The rest of the paper is organized as follows:

Section II states the problem formulation of the paper while Section III furnishes a survey on the recent developments of the researchers of Face Recognition System followed by the essential theoretical notes in Section IV. Section V gives a brief note on the Architecture of the Neural Network. Section VI comprises the training model followed by a discussion on the system performance and the results of the testing of the neural network in Section VII.

II. PROBLEM FORMULATION

An Artificial Neural Network is to be designed so that, it can recognize and classify a human face relative to the faces of the training database. Hence, the designed system classifies the face and verifies as well.Further, the face recognitionsystem needs to solve the problem concerning with different facial expressions as well. The system must be able to knowthat two images of the same person with different facial expressions actually are the same person.Some of the additionalissue that a face recognition systemmust solve is:makeup, posing positions, illumination conditions and effect of noise.

III. LITERATURE SURVEY

Face recognition using image data (2D face recognition) is a mature field that has been researched for many decades. Researchers have shown increased interest in developing reliable and efficient face recognition techniques. The various systems compete on each other in reducing computation time while improving the correct recognition rate. One of the commonly employed techniques involves representing the image by a vector in a dimensional space of size similar to the image. However, the large dimensional space of the image reduces the speed and robustness of face recognition. This problem has been overcome rather effectively by dimensionality reduction techniques such as the holistic features based face recognition approaches. Initially, researchers made use of Principal Component Analysis (PCA), the 2D Discrete Cosine Transform (2D-DCT) and morphable models [2], amongothers. The de facto standard for

International Journal of Web Technology
Volume: 03 Issue: 01 June 2014 Page No.9-14
ISSN: 2278-2389

2D face recognition iseigenfaces which is a technique that applies PCA to an image and was introduced to the field by Turk and Pentland [3]. The 2D-DCT based technique proposed in [4] models the 2D-DCT features using Gaussian Mixture Models.In the beginning of the 1970's, face recognition was treatedas a 2D pattern recognition problem[5]. The distancesbetween important points were used to recognize knownfaces. E.g measuring the distance between the eyes or otherimportant points or measuring different angles of facial components.

Kanade[6] was one of those whoactually managed to extract features automatic in a quitesimple manner and to recognize faces, previous systems hada manual extraction approach. Kanade's approachwas to convert a regular image into a binary image and then automatically detect the position of the eyes, mouth and nose of the face.The work done by Blanz and Vetter[7], and Huang[8] tookface recognition to a new level. By being able to use a morphable 3D model to create synthetic images has proven togive good results. It is a very appropriate approach that solvesmany of the problems of the previous works. In recent years, researchers have developed hybrid face recognition systems which take both the global features as well as the local features of the face. Weng and huang presented a face recognition model based on hierarchical neural network which is grown automatically and trained with gradient-descent. Good results for discrimination of ten distinctive subjects were reported [9].

IV. BASIC THEORY

Face recognition broadly refers to the twofold tasks of face identification and verification. Face verification consists of comparing an input image claiming an ID against the stored image of the claimed ID whilst face identification consists of finding the bestmatching image (in a database of images) given the input image.In this model, for the task of face identification the designed Network checks for the compatibility of the test image with any of the data base images while for face verification the test image is compared with the claimed image of the database for verification.These two processes are as depicted in Figure 1.

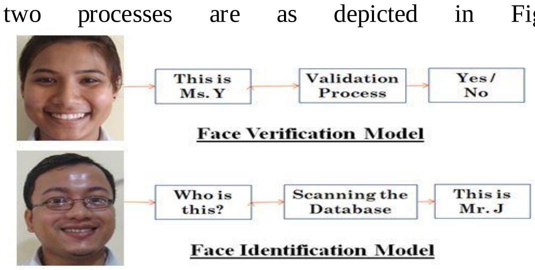


Figure 1. Block Diagram of Face Verification and Identification.

A. Different Approaches for Face Recognition

Various approaches for face recognition are discussed as follows:

***(1) Holistic matching methods*:** Thesemethods use the whole face region asthe raw input to a recognition system.One of the most widely used representations of the face region is eigen-faces which arebased on Principal Component Analysis.

(2) Feature-based (structural) matching methods: Typically, in these methods,local features such as the eyes, nose,andmouth are first extracted and theirlocations and local statistics (geometric and/or appearance) are fed into astructural classifier.These features include: Length & protrusion of ears, Lip thickness, Eye separation & shade, Nose profIle, Mouth width.

(3) Hybrid methods: Just as the humanperception system uses both local features and the whole face region to recognize a face, a machine recognitionsystem should use both. One can argue that these methods could potentially offer the best of the two types ofmethods.

B. Hybrid Feature Based Face Recognition

In the hybrid features approach, local as well as the global information of the face are extracted for recognition process. The basic steps involved in this approach could be summarized as:

a) Image Data Acquisition
b) Normalization or pre-processing of the images
c) Global and local features extraction
d)Training or testing of the images with machine learning constructs.
The issue of choosing the features to be extracted should be guided by the following concerns:

a) The features should carry enough information about the image and should not require any domain-specific knowledge for their extraction.
b) They should be easy to compute in order for the approach to be feasible for a large image collection and rapid retrieval.
c) They should relate well with the human perceptual characteristics since users will finally determine the suitability of the retrieved images.In this work, while retaining the principal component analysis of the whole image for global approach, the central moment, eigenvectors and standard deviation are employed for extracting the features of the eyes, mouth and nose used as the landmark segments of the face.

(1) Principal Component Analysis: To find principalcomponentsusing eigenfacesalgorithm we need to use the following methods:

• First of all we need to find the linear combinations of the original variables with large variance.
• The covariance matrix C or the correlation matrix R is then calculated.
•The eigenvalues and eigenvectors of C or Rarefound.
•The eigenvalues are computed in descending order (from largest to smallest), $e_1, e_2, e_3 \ldots e_p$.
•Finally the corresponding eigenvectors $a_1, a_2, a_3 \ldots, a_p$ are found , where $a_i'a_i = 1$ and $a_i'\, a_j = 0$, thus,
$y_1 = a_1'x = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p$ is the firstprincipal component,
$y_2 = a_2'x = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2p}x_p$ is the secondprincipal component,
$y_p = a_p'x = a_{p1}x_1 + a_{p2}x_2 + \ldots + a_{pp}x_p$ is the p'th principal component.

International Journal of Web Technology
Volume: 03 Issue: 01 June 2014 Page No.9-14
ISSN: 2278-2389

(2) Central Moment: In image processing, computer vision and related fields, an image moment is a certain particular weighted average (moment) of the image pixels' intensities, or a function of such moments, usually chosen to have some attractive property or interpretation. Image moments are useful to describeobjects after segmentation. Simple properties of the image which are found via image moments include area or total intensity, its centroid and information about its orientation [10].Central moments are translational invariant. They find their application in recognition of shape features which are independent of parameters and which cannot be controlled in an image are generated. Such features are called invariant features. There are several types of invariance. For example, if an object may occur in an arbitrary location in an image, then one needs the moments to be invariant to location. For binary connected components, this is achieved simply by using the central moments [11].

V. ARTIFICIAL NEURAL NETWORK

An artificial neural network (ANN), also called a simulated neural network (SNN) or commonly just *neural network* (NN) is an interconnected group of artificial neurons that uses a mathematical or computational model for information processing based on a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network.In more practical terms neural networks are non-linear statistical data models of decision making tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data.ANNs can identify and learn correlated patterns between input data sets and corresponding target values. This feature makes such computational models very appealing in application domains where one has little or incomplete understanding of the problem to be solved but where training data is readily available.In this work, we use a multilayered General Feed-Forward Artificial Neural Network. The architecture of this class of network, comprises three components namely, weighting function, summation function and activation function as depicted in Figure 2.

1) Weighting Function: A neuron usually receives many simultaneous inputs. Each input has its own relative weight which gives the input the impact that it needs on the processing element's summation function. These weights perform the same type of function as do the varying synaptic strengths of biological neurons.

2) Summation Function: The first step in a processingelement's operation is to compute the weighted sum of all of the inputs.The summationfunction can be more complex than just the simpleinput and weight sum of products.

3) Activation Function: An activation function f performs a mathematical operation on the signal output. The most common activation functions are:Linear function, Threshold function, Piecewise Linear function, Sigmoid function, Tangent hyperbolic function.

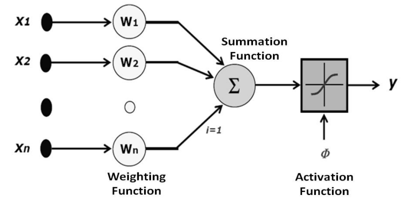


Figure 2. Multilayered General Feed-forward Network Configuration

VI. TRAINING MODEL

The block diagram of the proposed model is shown in Figure 3.

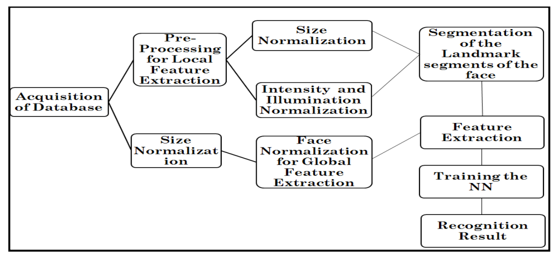


Figure 3. Block diagram of the proposed face recognition model

(1) Acquisition of Database: In this work, we have used our own image data base. The data base consists of 20 images of each of 8 persons totaling to 160 images. Of these, 120 images (15 images per person) are used for training while the other 40 images are used to testing the system. These images were captures at different instances of the day; with varying illumination level; varying distance from the camera; different light intensity; varied facial expressions. The database consists of images of both men and women. Sample of the database images of two persons with different facial expressions and different illumination level is given in Figure

4.

Figure 4. Acquired Data Base Sample of two Persons

(2) Normalization and Pre-processing of Images: The main goal of the preprocessing step in ourmethod is the reduction of the dimensionality problemby spatial normalization of the face image. The different pre-processing techniques carried out are:

a) Brightness distribution: This is carried out by histogram equalization to improve the contrast in the grayscale so as to obtain a uniform histogram. The histogram equalization method also helps the image to reorganize the intensity distributions. New intensities are not introduced into the image. Existing intensity values will be mapped to new values but the actual number of intensity pixels in the resulting image

International Journal of Web Technology
Volume: 03 Issue: 01 June 2014 Page No.9-14
ISSN: 2278-2389

will be equal or less than the original number.The problem of light intensity is thus overcome (Figure 5).

b) Noise Removal: The images of the database may be affected by noise such as Gaussian or Salt and Pepper or Poisson noise or Multiplicative noise. These are removed so that the network does not get trained with unwanted noise information (Figure 6).

c) Illumination Normalization:Images taken under different illuminations can degrade recognition performance especially for face recognition systems based on the global feature analysis in which entire face information is used for recognition. A picture can be equivalently viewed as an array of reflectivities r(x) (Figure 7).

d) *Face Normalization for Global Approach*: Face normalization is carried out by finding the mean of all the database face images followed by subtracting each image from this mean image (Figure 8).

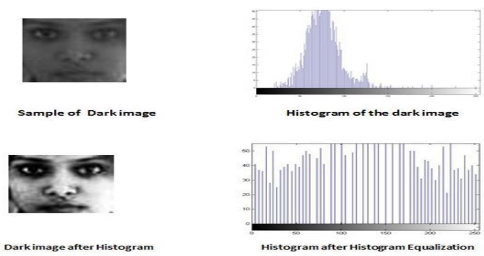


Figure 5. Histogram plot of a dark image before and after Brightness Distribution

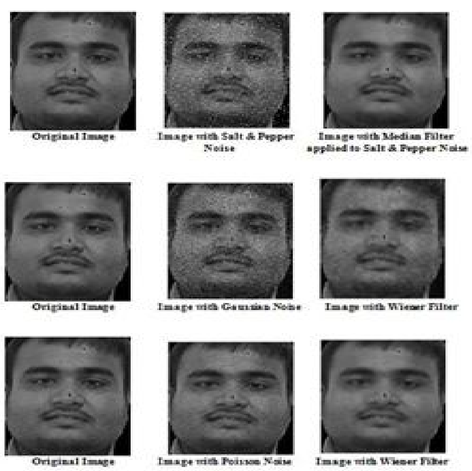


Figure 6. Sample of noise removal process

Figure 7. Sample of images with varying illumination level

Figure 8. An individual face, normalized face of the data base faces and the resulting face of the subtraction of individual face from the mean face

(3) Extraction of the Landmark Segments of the Face: From the pre-processed face images, the eyes, nose and mouth of the image are detected for procuring the local feature vector as shown in Figure 9. These segments are used to compute the Eigen Vector of the covariance, central moment and Standard Deviation.

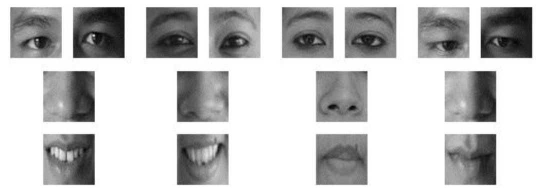


Figure 9. Sample of Face Segmentation to extract the landmark segments of the face

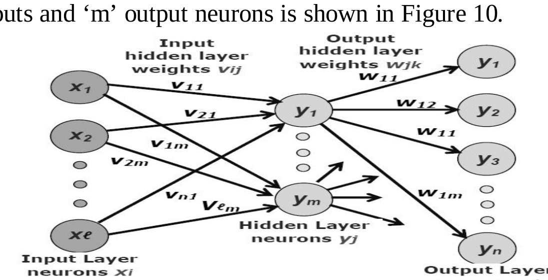
(4) Feature Extraction: The different features extracted are tabulated in Table I. For the global features, the eigenvectors corresponding to the first 4 eigenvalues are considered. The traditional motivation for selecting the Eigenvectors with the largest Eigenvalues is that the Eigenvalues represent the amount of variance along a particular Eigenvector. By selecting the Eigenvectors with the largest Eigenvalues, one selects the dimensions along which the gallery images vary the most.

TABLE I. EXTRACTED FEATURES

Sl. No	Facial Region	Extracted Features		
		Central Moment	Eigen Vector of Covariance of the Face	Standard Deviation
1	Right Eye			
2	Left Eye			
3	Nose			
4	Mouth			

VII. ARCHITECTURE OF THE NETWORK

The neural network has three layers namely the input layer, hidden layer and the output layer. The extracted features are fed into the input layer. These values arrive at each of the hidden neurons as a sum of the weighted inputs given by the suumation function in Figure 2. A neural network with ‘l’ inputs and ‘m’ output neurons is shown in Figure 10.



International Journal of Web Technology
Volume: 03 Issue: 01 June 2014 Page No.9-14
ISSN: 2278-2389

Figure 10. Multilayered feedforward network configuration

In the network designed here, the number of input neurons is 752 and the size of hidden layer is 506 and 8 neurons in its output layer to identify the faces. The size of the hidden layer was calculated as $2/3^{rd}$ of the sum of the input neurons and the output neurons.The network is a two-layer log-sigmoid/log-sigmoid network. The log-sigmoid transfer function was picked because its output range (0 to 1) is perfect for learning to output Boolean values. The specifications of the network is as tabulated in Table II.

TABLE II. SPECIFICATIONS OF THE NEURAL NETWORK

Type: Feed Forward Backpropagation Network	
Parameters	*Specifications*
Number of Layers	3 (Input layer, Hidden Layer, Output Layer)
Number of Input Unit	1 Feature Matrix consisting of feature vectors of eight persons
Number of Output Unit	1 Binary Encoded matrix consisting of eight target vectors for eight persons.
Number of Neurons in the Input Layer	752
Number of Neurons in the Output Layer	8
Number of Neurons in the Hidden Layer	(752 + 8) * (2/3) = 506
Number of Iterations	1000, 1500, 2000
Number of Validation Checks	6
Learning Rate	0.7
Momentum	0.6
Activation Functions	Log-Sigmoid and Tan-Sigmoid

The network is trained to output a 1 in the correct position of the output vector and to fill the rest of the output vector with 0's. However, noisy input images may result in the network not creating perfect 1's and 0's. After the network has been trained the output willbe passed through the competitive transfer function. This function makes sure that the output corresponding to the face most like the noisy input image takes on a value of 1 and all others have a value of 0. The result of this post-processing is the output that is actually used. Figure 11 gives a pictorial diagram of the classification process of the neural network.

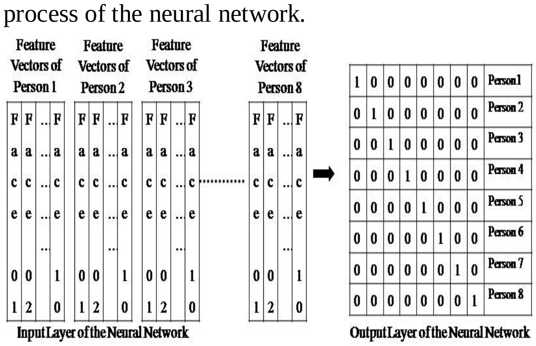


Figure 11. Pictorial representation of the network classification

VII. System Performance and Results

To obtain a network not sensitive to noise, we trained with ideal copies and noisy copies of the images in database. The noisy images have noise of mean 0.1 ("salt & pepper " noise) and 0.2 ("poisson "noise) added to them. This forces the neurons of network to learn how to properly identify noisy faces, while requiring that it can still respond well to ideal images. Further, the network was trained over many times changing the trained images and the test images. Initially, the first 15 images of each person were used for training while retaining the rest 5 images for testing. Next, images 2 through 16 were used for training with the rest of the images as test images. Thus, for each training different set of images are used for training and testing. The average of the test results are tabulated. This is to ensure that the system is not biased.
Mean squared error (MSE) of the training is used as the performance evaluation. The MSE convergence of one of the training sessions is given in Figure 12. We see that with the number of iterations of the adjusting of the weights and bias of the network, the MSE diminishes as expected.

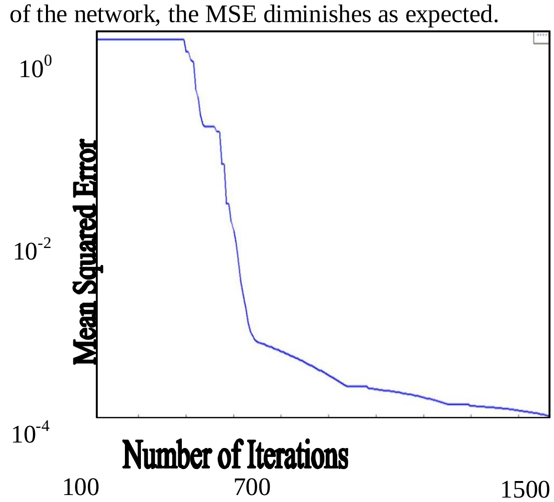


Figure 12. MSE Convergence Curve

The results of the convergence of MSE are as tabulated in Table III. We notice that for a learning rate of 0.7 the MSE convergence gives a satisfactory result. Further, for a learning rate of 0.7 when the momentum was being set to 0.6 the results are optimum. And the most satisfactory convergence of MSE is obtained when the network is trained for 1500 iterations with learning rate of 0.7 and the momentum being 0.6.

Table Iii. Convergence Of Mse For Different Learning Rates And Momentum And Iterations Using Logarithmic Sigmoid Activation Function

Learning Rate	**Momentum**	**MSE for different Iterations**		
		1000	***1500***	***2000***
0.7	0.5	$1.5x10^{-7}$	$1.3x10^{-10}$	$1.7x10^{-4}$
	0.6	$1x10^{-7}$	$1x10^{-12}$	$1x10^{-6}$
	0.7	$1.2x10^{-6}$	$1.2x10^{-10}$	$1.1x10^{-5}$
0.8	0.5	$1.4x10^{-}$	$1.5x10^{-7}$	$1.4x10^{-}$

		4		3
	0.6	1.3×10^{-5}	1×10^{-8}	1×10^{-5}
	0.7	1.2×10^{-4}	1×10^{-7}	1.6×10^{-4}
0.9	0.5	1.7×10^{-3}	1.8×10^{-6}	1.4×10^{-2}
	0.6	1.4×10^{-5}	1×10^{-7}	1.3×10^{-4}
	0.7	1.3×10^{-3}	1.4×10^{-6}	1.7×10^{-3}

From the results of the testing done for the different training sessions, the correct recognition rate (CRR) for face identification process and the false acceptance rate (FAR) and false rejection rate (FRR) for face verification process are calculated.The false rejection rate, or FRR, is the measure of the likelihood that the biometric security system will incorrectly reject an access attempt by an authorized user.The false acceptance rate, or FAR, is the measure of the likelihood that the biometric security system will incorrectly accept an access attempt by an unauthorized user.

TABLEIV. COMPARISON OF RECOGNITION RATES OF THE PROPOSED SYSTEM WITH THE DATABASE AFFECTED BY VARIOUS TYPES OF NOISE

Database with Noise	CRR %	FAR %	FRR %
Ideal Images	93.5	02	01
Poisson	91.0	03	03
Gaussian	87.8	05	06
Speckle	85.6	06	09
Salt & Pepper	78.8	08	12

VIII. CONCLUSION

In this paper, face recognition based on ANN is proposed. ANN with Back propagation algorithm is found to be the efficient method for recognising the faces. It is observed that the proposed feature vectors are useful for proper recognition of human faces with the activation function 'log sigmoid' in the hidden layer of the Neural Network and it gives a better convergence of the MSE. It is also observed that the system's efficiency is adversely affected in the presence of Salt & Pepper noise present in the training set of images. However, for ideal images the Correct Recognition Rate is over 93%.

REFERENCES

[1] W. Zhao, R. Chellappa and A. Rosenfeld. Face recognition: A literature survey. ACM Computer Survey, 35(4):399{458, 2003.

[2] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. IEEE Transactions on Pattern Analysis and Machine Intelligence, pages 1063–1074, 2003.

[3] M. Turk and A. Pentland, (1991), 'Eigenfaces for recognition', Journal of Cognitive Neuroscience 3, 71–86.

[4] S. Lucey. The symbiotic relationship of parts and monolithic face representations in verification. Proceedings of IEEE Conference of Computer Vision and Pattern Recognition Workshop, page 89, 2004.

[5] M. D. Kelly. Visual identification of people by computer. PhD thesis, Stanford University, Stanford, CA, USA, 1971.

[6] T. Kanade. Picture Processing System by Computer Complex and Recognition of Human Faces. PhD thesis, doctoral dissertation, Kyoto University, November 1973.

[7] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. IEEE Trans. Pattern Anal. Mach. Intell., 25(9):1063-1074, 2003.

[8] J. Huang, B. Heisele, and V. Blanz. Component-based face recognition with 3d morphable models. In J. Kittler and M. S. Nixon, editors, International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA-3), volume 2688 of Lecture Notes in Computer Science, pages 27-34, Surrey, UK, 2003. Springer.

[9] J. Weng, N. Ahuja, and T. S. Huang, "Learning recognition and segmentation of 3D objects from 2D images," Proceedings of the International Conference on Computer Vision, 1993, pp 121–128.

[10] M. K. Hu, "Visual pattern recognition by moment invariants", IRE transactionson Information Theory, 1962, pp. 179–187.

[11] Sundos A. Hameed Al_azawi, "Eyes Recognition System Using Central Moment Features", Engineering& Technology Journal, 2011, vol. 29, no. 7, pp. 1400-1407.