

3D FACE DETECTION AND RECOGNITION

Abhishek Chauhan, Shivangi Srivastava, Anamika Singh, Aman Singhal

Abstract— This paper presents real time face detection and recognition system and also an efficient technique to train the database. The algorithm used is of stereo-face detection in video sequences. A stereo face is a face of man presented by the set of images obtained from different points of views such data can be used for face of man presented by the set of images obtained from different points of views. Such data can be used for structure estimation. Our algorithm is based on computationally effective method of face detection in mono image. Information about face position is then combined using sparse stereo matching algorithm. Sparse means that stereo correspondence is estimated not for all scene points but only the point of interests. This allows obtaining the low computational cost of algorithm Then we used the ANN(Artificial Neural Network) to train the database and get the information out of that detected face and recognize it. And for effective database storage we need holographic data storage technique to store our database.

Index Terms — Image analysis, object recognition, stereo vision, Vision Access 3D Desk Cam , RAANSAC,

1 INTRODUCTION

FACE Recognition has become very active area for the research and development as the demand of security and commercial use. The last decade has shown a widely progress in this area, with emphasis on such applications as man-computer interface, multimedia applications, biometric identification in surveillance and access control systems. As the facial recognition is the task for human brain so it can be proved to be difficult task to detect it artificially. The major problem behind it is differing image qualities, facial expressions, and facial furniture, background, and illumination conditions[1],[2]. Fig1 will show the generalised approach for the face recognition system

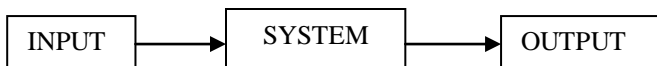


Fig1 Face recognition system

Many methods exist to facial recognition problem such as template matching, Fisher Linear Discriminant, Neural Networks, SVM, and MRC. Success has been achieved with each method to varying degrees and complexities. All these methods ensures for the better accuracy than that of 2d detection schemes.

So for that we have to use special devices to capture and feed images and those are images from video sequence or some special cameras to get stereo images. Here to acquire stereo image pairs and get the facial points we use two VisionAccess3DDeskcams. Mentioned approaches allow estimating three-dimensional face structure with appropriate quality but have disadvantages.

and complicated to implement. For the person identification system. Two major things which are required are high speed model estimation accuracy of person identified and minimum requirement of object distance estimation. Accuracy and speed estimation based on mono images view while the later are based on stereo image view. For the accuracy of the model frontal position is required to accept it.

Stereo reconstruction can give an additional information estimation. Stereo base face detection methods require additional efforts for depth estimation. It also usually need that cameras have parallel optical axes. In other case it will be necessary to perform transformation of the images to a standard parallel stereo view. Stereo reconstruction method can be used for the additional information estimation. Stereo base face detection methods require additional efforts for depth estimation.

Facial feature mainly concentrate on the eyes, nose and mouth, therefore, this paper mainly detects the characteristics of the three regions in the human face, then calculate the geometric characteristics of human face. It require that cameras have parallel optical axes. In other case it will be necessary to perform transformation of images to a standard parallel stereo view.

In this paper we present the algorithm of face detection and matching in the video sequence or can say in the frame of stereo image sequence. As our main goal in this paper is to detect and localize stereo-faces of people moving in the field of view of several video camera . Object distance estimation method does not use projective transformation of stereo planes is also considered.

They can't be used in real time system as they are costly

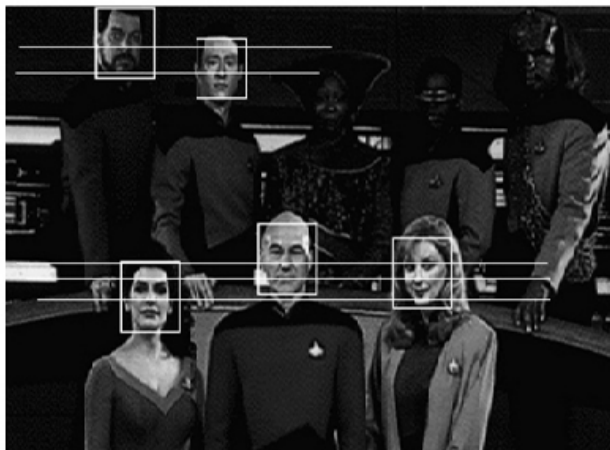


Fig2.Generation of epipolar lines

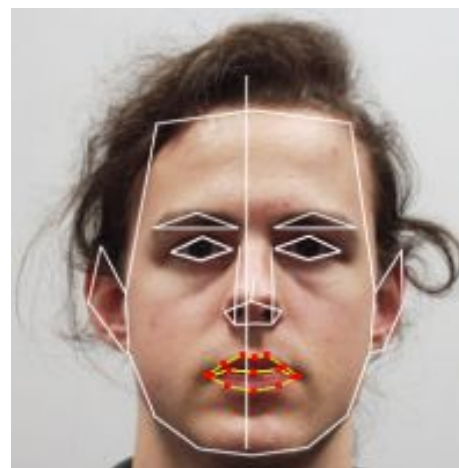


Fig4.face consideration

2 CAMERA CALIBRATION

Camera Calibration in stereo vision mainly refers to the computation of relative position of the Camera optical centers . We used Vision Access 3D DeskCam which is best suited camera for the task. It demarkates 40,000 points on a face which is acceptable to compute fundamental matrix that represents epipolar geometry of camera views. The fundamental matrix allows calculating so called epipolar lines in one frame that corresponds to given points in another frames We used RAANSAC method that is based on eight points algorithm for fundamental matrix estimation. The common points are grouped in a set of eight points and the best possible epipolar lines are drawn The user has to mar atleast eight pairs of correspondent points for calibration obtained matrix can be applied for epipolar lines estimation by using the relation.

$$l = F * x$$

Wherel-sought value, F-fundamental matrix and x-given point

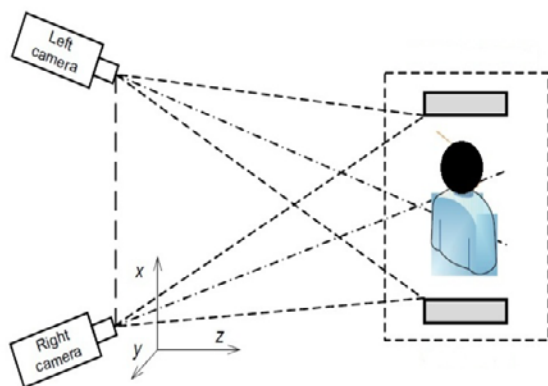


Fig 3: 3D face processing in security applications

3 STEREO FACE DETECTION

We use sparse stereo matching algorithm for stereo-faces obtaining. It works with small amount of scene points that leads to fast processing speed. In classical scheme of stereo reconstruction correspondence is estimated for each point of stereo frames. Our algorithm processes only points of interests. These points correspond to human faces found at each view. We use cascade of weak Haar classifiers for face detection in each frame.

3.1 Matching Algorithm

Applied approach for face corresponding estimation is based on geometric characteristics analysis, spatial face position and visual correlation. Faces found in each view are characterized by their geometric characteristics and pattern. The algorithm calculates correspondence degree for objects from two frames based on these characteristics. The goal of this procedure is selection of most probable pairs and rejection of faces without pair. Pair correspondence is defined by multiplication of the set of coefficients. Each coefficient corresponds to some feature correlation:

- K1- represents the faces positions accuracy in compliance with epipolar geometry;
- K2 - represents face sizes correspondence;
- K3- represents histogram correspondence of the face areas.

All calculation are performed in the coordinate system of one frame from stereo pair. Algorithm handles only pairs that comply with epipolar constraint. In other words, pair candidates for the face from one view have to lie on correspondent epilines at another view. In fig.2 each man's face in first image has two candidate faces in second image, but woman's face has only one. This constraint allows greatly reducing the amount of false pairs. Then size and position coefficients are calculated. Coefficient K1 is

defined as ratio between distance from face area centre to correspondent epipolar line and minimal facesize in pair.

It possesses the value 1 when the center of the face coincides with the line and is less than 1 otherwise. Coefficient K2 is defined as ratio between smaller and bigger area sizes. It also is less or equal to 1. It is necessary to estimate face position in the 3D space for coefficient K3 calculation[1].

3.2 Histogram matching.

Coefficients K1 and K2 allow rejecting big amount of false pairs. They are not enough because are based on geometrical features that can be obtained with essential error. That is why histogram matching based coefficient K3 is computed for the rest small amount of possible pairs. Histogrammatching is used because of its relative computationalsimplicity in comparison with other metrics. Histogram is built not for all face area to avoid influenceof background, hair, clothes pieces that can be not presented at both views. Only central part of face area is used where eyes, nose and mouth are apparently situated (see fig. 3). Median filter and area normalization are used to bring histogram tocommon form. Sum of Absolute Differences (SAD) and Sum of Squared Differences were used for histogram matching. The SAD applied to three-cannel histograms shows the best result. Error is calculated as square of Euclidian distance between each RGBcomponent's errors. Coefficient K4 then is estimated as:

$$K3 = 1 - histERR$$

wherehistERRis histogram correspondence error. Hence maximization of the product $K1 * K2 * K3$ gives us the most probable pairs. Pairs formed by faces without right correspondence are ignored by threshold. Such pairs can appear because only one face view can be visible for stereo cameras. The threshold is applied to compound coefficient (the product).Its value was estimated empirically and is equal to 0.3[1].

4 ANALYSIS

To acquire stereo image pairs and get the facial points we use two VisionAccess 3D Descam.

The two cameras are mounted next to each other with a baseline distance of 20cm and the angle of convergence of 9.5 degree.

| ah* | 0 | 100 | 300 | 500 | 700 |
|-----|-------|-------|-------|-------|-------|
| 10 | 0.207 | 0.216 | 0.239 | 0.33 | 0.477 |
| 60 | 0.229 | 0.232 | 0.243 | 0.259 | 0.339 |
| 120 | 0.605 | 0.311 | 0.249 | 0.237 | 0.256 |
| 170 | 1.632 | 0.284 | 0.20 | 0.221 | 0.243 |

a-(Degree) h- (Millimeters)

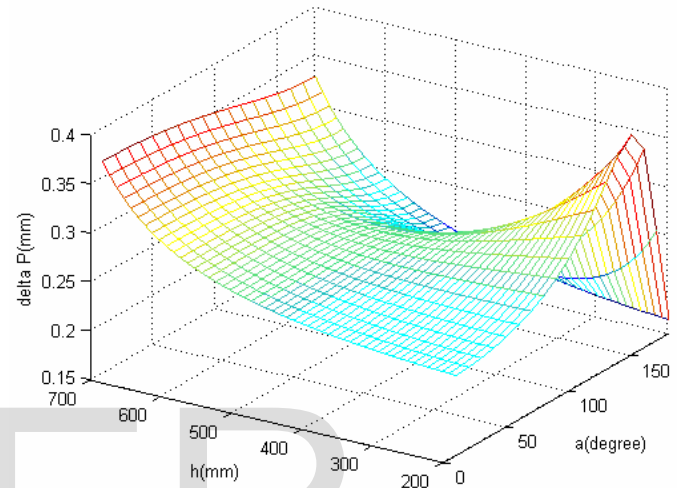


Fig.5 Histogram View of Analysed Data



Fig.6StereoImages

Table1. Variation in BaselineDistance and ConvergenceDegree

5 TRAINING DATABASE

ANN is a machine learning algorithm that has been used for various pattern classification problems such as gender classification, face recognition, and classification of facial expression. ANN classifier has advantages for classification such as incredible generalization and good learning ability. The ANN takes the features vector as input, and trains the network to learn a complex mapping for classification, which will avoid the need for simplifying the classifier. Being able to offer potentially greater generalization through learning, neural networks/learning methods have also been applied to face recognition in The ANN paradigm that is used in this application is Multilayer Feed forward Neural Networks (MFNNs).

MFNNs are a form of non linear network consisting of a set of inputs (forming the input layer), followed by one or more hidden layers of nonlinear neurons and an output layer of non-linear neurons. MFNN is an ideal means of tackling a whole range of difficult tasks in pattern recognition and regression because of its highly adaptable non-linear structure. In order to train the network to perform a given task the individual weights (w_{ij}) for each neuron are set using a supervised learning algorithm known as the error-correction back-propagation algorithm as depicted in which involves repeatedly presenting the network with samples from a training set and adjusting the neural weights in order to achieve the required output. It is essentially a gradient descent method, where when adjusting the weight matrices, the direction is moved to the greatest descent.

The learning constant, η , must be chosen with care. If it is too large, the algorithm may repeatedly overshoot the solution, which will lead to slow convergence or even no convergence at all. However, if it is too small, the algorithm will only approach the solution at a very slow rate, again leading to a slow convergence and increasing the chances of the algorithm becoming stuck in local minima. Two main methods of overcoming these problems are momentum and adaptive learning. For momentum method, if we are consistently moving in the same direction, then we want to build up some momentum in that direction.

This will help us to go through any small local minima and hopefully speed up convergence.

6 CONCLUSION:

Storage:

Holographic data storage records information throughout the volume of the medium and is capable of recording multiple images in the same area utilizing light at different angles.

COMPARISON CHART

| PARAMETERS | DVD | BLUERAY S | HDS S |
|----------------------------------|-------------|--------------|---------------|
| Capacity | 4.7gb | 25gb | 3.9tb |
| Laser Wavelength | 650nm (red) | 405nm (blue) | 532nm (green) |
| Disc Diameter | 120mm | 120mm | 120mm |
| Hard Coating | No | Yes | Yes |
| Data transfer rate | 11.08mbps | 36mbps | 1gbps |
| Data transfer rate (audio/video) | 10.08mbps | 54mbps | 1gbps |

For every person all possible set of images can be stored at one location at different angles. And we can retrieve the data from that location. We can provide a new slot for a new entry of different individual.

The comparison chart clearly justifies the ability and its capabilities. In summary, artificial Neural Networks and especially MLPs are one of the promises for the future in pattern recognition. They can recognize patterns within large datasets and then generalize those patterns into recommended courses of action. Yet, even though they are not traditionally programmed, the designing of neural networks does require a skill. This skill involves a strategy to acquire the necessary data to train the network. It also involves the selection of learning rules, transfer functions, data preprocessing methods and mainly how to connect the neurons within the network.

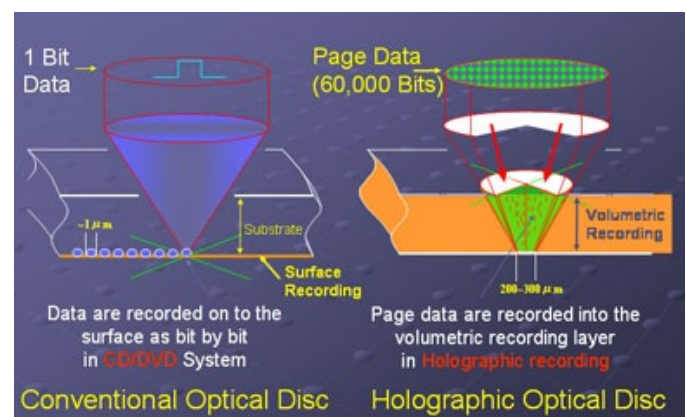


Fig7.Storage Pattern differences

References:

- [1]. Sadykhov. RaufKh.,Lamovsky. Denis V.,” **Algorithm for Real Time Faces Detection in 3D Space.**”,IMCSIT, PROCEEDINGS OF THE IMCSIT. VOLUME 3, 2008,p.g-1-6
- [2]. Li, Phung, Bouzerdom, Tivive, “**Automatic Recognition of Smiling and Neutral Facial Expressions**”, 2010 Digital Image Computing: Techniques and Applications, Australia,
- [3]. Kosov Sergey, Thorm“ahlen Thorsten, Seidel Hans-Peter “**RAPID STEREO-VISION ENHANCED FACE RECOGNITION**”
- [4]. Achim Sanjay Manikarnika,”**A GeneralFace Recognition System**”, Master Of Science In Physics And Mathematics, July 2006
- [5]. A. M. Bronstein, M. M. Bronstein, and R. Kimmel, “**Three-dimensionalface recognition,**” *International Journal of Computer Vision*,vol. 64, no. 1, 2005, pp. 5-30.
- [6]. Alvertos Nicolas, “**Resolution limitations and error analysis for stereo camera models**” Conference Proceedings - IEEE SOUTHEASTCON, p220-224, 1988.
- algorithms - A review**”, IETE Journal of Research, v 50, n 1, 176-B, p 49-61, January/February, 2004
- [8]. Richard Hartley, Andrew Zisserman , Multiple View Geometry in Computer Vision, Cambridge University Press, 2000.
- [9]. J.-G. Wang, E.T. Lim, X. Chen, and R. Venkateswarlu, “**Real-time stereo face recognition by fusing appearance and depth fisherfaces,**” J. VLSI Signal Process. Syst.,vol. 49, no. 3, pp. 409–423, 2007.
- [10]. Akihiro Hayasaka, Koichi Ito, Takafumi Aoki, Hiroshi Nakajima, and Koji Kobayashi, “**A robust 3d face recognition algorithm using passive stereo vision,**” IEICE Transactions, vol. 92-A, no. 4, pp. 1047–1055, 2009.
- [11]. C. Park, and J. Paik, "Face recognition using optimized 3D information from stereo images," *Image Analysis and Recognition*,2005,pp.1048-105

IJSER

- [7]. Kumar S. Srinivas, Chatterji B.N., “**Two-frame stereo matching**