Position Detection with Face Recognitionusing Image Processing and Machine Learning Techniques

G.Suvarna Kumar
C.S.E Department
M.V.G.R C.E
A.P,India

P.V.G.D Prasad Reddy
CS & SE Department
Andhra University
A.P , India

R.Anil Kumar
C.S.E Department
M.V.G.R.C.E
A.P , India

Sumit Gupta
C.S.EDept
M.V.G.R.C.E
A.P,India

ABSTRACT
In this paper, an improved algorithm for detecting the position of a person in controlled environments using the face detection algorithm is proposed. This algorithm ingeniously combines different face detection, occlusion detection algorithms, EMD for facial recognition and SVM classifier. A class room environment with thirty students is used along with some constraints such as position of the camera being fixed in a way that covers all the students, the quasi-static student's position and the class environment with the fixed lighting conditions. For every class, a set of 6 attributes are derived and updated in a database. The image is given as an input to the face detection algorithm to detect some of the faces. Some faces are not detected because of occlusion, so an occlusion detection technique is implemented to detect all the occluded faces. Using the EMD based face recognition techniques, missing positions are correlated with individuals assuming a quasi-static setup. Experiments have been conducted in different classroom settings and accuracies of more than 96% have been obtained. In this paper lib SVM is used.

General Terms
Face Detection, Position Detection, EMD.

Keywords
Face detection, Occlusion detection, SVM, EMD.

1. INTRODUCTION
In the last decade, it can be observed that many algorithms were developed in image processing for face recognition [1] and face detection [2] but there was no algorithm for detecting a position in an image. The position detection in the controlled environment can be used in many environments where the positions are fixed. Some environments like seminar halls, conference halls, Auditoriums etc. And we can apply the same algorithm for some environments where the positions are reserved. With the help of this position detection algorithm many applications can be developed like automatic attendance system [3] in controlled environments. This technique can also be used for improving the face detection technique although no face detection technique will give the 100 % accuracy in its respective detection. Some of the faces in an image are blurred, not clear, occluded, the normal face detection algorithm will not detect those faces but this technique detects most of the faces in an image. First the image is given as an input to the basic face detection algorithm followed by giving the acquired result of the above as an input to the occlusion detection [4] and then this technique is applied. In some cases the students may sit in positions varying from their original static arrangement

.To overcome these kind of problems another feature is added to this system. A recognition feature is implemented after the position is detected, to validate the missing data with the neighbor. The face recognition algorithm which is implemented will help the system for a better recognition of a class, considering each face as a class.

A large number of face detection algorithms are derived from neural network approach, algorithmic approach [5] and some image morphological techniques [6]. However most of the works concentrate on single face detection, with some constrained environments. In this proposed technique, a face detection algorithm by using local SMQT features and split up snow classifier [7] and an occlusion detection algorithm is used and from the result obtained the attributes are derived and updated in the database. For recognition phase, EMD algorithm has been proposed and implemented.

A face detection algorithm is used which is implemented using the Local SMQT Features and split up Snow Classifier. Object identifier helps in identifying the shape of a face rather than any object of same shape of a face.

2. EXISTING SYSTEM
Many face detection algorithms are derived. Some of the face detection algorithms use neural network approach, algorithmic approach and some image morphological techniques. With the help of these existing face detection algorithms, 80% of accuracy is obtained. When the result is given as an input to the occlusion detection algorithm, the accuracy is improved to 95% accuracy. Then the above acquired result is given as an input to the proposed technique. Face detection algorithm using Local SMQT Features and Split up Snow classifier is explained in the next paragraph.

Illumination and sensor variation are major concerns in visual object detection. It is desirable to transform the raw illumination and sensor varying image so the information only contains the structures of the object. Some techniques previously proposed to reduce this variation are Histogram Equalization (HE), variants of Local Binary Patterns (LBP) [8] and the Modified Census Transform (MCT) [9]. HE is a computationally expensive operation in comparison to LBP and MCT, however LBP and MCT are typically restricted to extract only binary patterns in a local area. The Successive Mean Quantization Transform (SMQT) [10] can be viewed as a tunable tradeoff between the number of quantization levels in the result and the computational load. In this paper the SMQT is used to extract features from the local area of an image. Derivations of the sensor and illumination insensitive properties of the local SMQT features are presented. Pattern recognition in the context of appearance based face detection can been approached in several ways [11]. Techniques proposed for this task are for example the Neural Network (NN) [12], probabilistic modeling [13], cascade of boosted features (AdaBoost) [14], Sparse Network of Winnows (SNoW) [15], combination of AdaBoost and SNoW and the Support Vector Machine (SVM) [16]. This technique proposes an extension to the SNoW classifier, the split up SNoW, for this classification task. The split up SNoW will utilize the result from the original SNoW classifier and create a cascade of classifiers to perform a more rapid detection. It will be shown that the number of splits and the number of weak classifiers can be arbitrary within the limits of the full classifier. Further, a stronger classifier will utilize all information gained from all weaker classifiers.

The SMQT uses an approach that performs an automatic structural breakdown of information. Our previous work with the SMQT can be found in. These properties will be employed on local areas in an image to extract illumination insensitive features. Local areas can be defined in several ways.
Nevertheless, once the local area is defined it will be a set of pixel values. Let x be one pixel and$D(x)$ be a set of $|D(x)| = D$ pixels from a local area in an image. Consider the SMQT transformation of the local area.

SMQT (local): D(x) ->M(x) (1)

This yields a new set of values. The resulting values are insensitive to gain and bias. These properties are desirable with regard to the formation of the whole intensity image I(x), which is a product of the reflectance R(x) and the illuminanceE(x). Additionally, the influence of the camera can be modeled as a gain factor g and a bias term b. Thus, a model of the image can be described by

I(x)=gE(x)R(x) + b (2)

In order to design a robust classifier for object detection the reflectance should be extracted since it contains the object structure. In general, the separation of the reflectance and the illuminance is an ill posed problem. A common approach to solving this problem involves assuming that E(x) is spatially smooth. Further, if the illuminance can be considered to be constant in the chosen local area then E(x) is given by

E(x) =E ∀ x ∋ D (3)

The SNoW learning architecture is a sparse network of linear units over a feature space [9]. One of the strong properties of SNoW is the possibility to create lookup-tables for classification. Consider a patch$W$of the SMQT features$M(\mathbf{x})$, then a classifier.

$$\theta=\sum_{x \epsilon W} h_x^{nonface} \left(M(X)\right) - \sum_{x \epsilon W} h_x^{face} \left(M(x)\right) \quad (4)$$

can be achieved using the non-face table $h_x^{nonface}$ $_x$, the face table h_x^{face} and defining a threshold for θ. Since both tables work on the same domain, this implies that one single lookup-table.

$$h_x = h_x^{nonface} - h_x^{face} \quad (5)$$

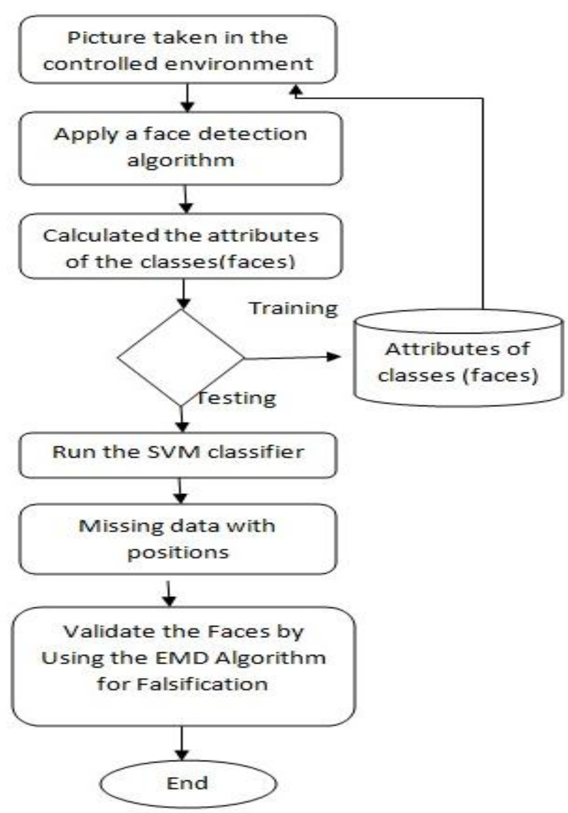
Let the training database contain $i = 1, 2. . . N$ feature patches with the SMQT features m_i(x) and the corresponding classes c_i(face or nonface). The non-face table and the face table can then be trained with the Winnow Update Rule. Initially both tables contain zeros. If an index in the table is addressed for the first time during training, the value (weight) on that index is set to one. There are three training parameters; the threshold ¥the promotion parameter α >1 and the demotion parameter $0<\beta<1$.

In order to scan an image for faces, a patch of 32×32 pixels is applied. This patch is extracted and classified by jumping Δx= 1and Δy= 1 pixels through the whole image. In order to find faces of various sizes, the image is repeatedly downscaled and resized with a scale factor.

To overcome the illumination and sensor problem, the proposed local SMQT features are extracted. Each pixel will get one feature vector by analyzing its vicinity. This feature vector can further be recalculated to an index.
WhereV (x_i) is a value from the feature vector at position i. Thisfeature index can be calculated for all pixels which results in thefeature indices image.
A circular mask containing P = 648 pixels is applied to eachpatch to remove background pixels, avoid edge effects from possible filtering and to avoid undefined pixels at rotation operation.

3. PROPOSED SYSTEM

This proposed technique was implemented with some attributes



derived from an image and the framework is shown in Figure-1.

Figure-1: The Framework of the proposed system

Figure-2: A Sample image in a dataset

The attributes are derived for every face. In this technique, the faces are treated as a single class. For every single class, six attributes are derived.

These attributes have been derived for 4 different images taken in four different days. The attributes are length and height of the face, position of the face in coordinates, and the number of horizontal lines and vertical lines that are passed from a single class. The framework of the proposed system is shown in the figure 1. In the first step the a sample image, figure 2 is given as an input to the face detection algorithm [2] whose result is shown in figure 3, here some of the faces that are not detected are the ones that are occluded [4], this is shown in figure 4. To detect the occluded faces, the above obtained result is given as an input to the occlusion detection algorithm procedure. The output of the occlusion detection algorithm has been shown in figure 5.

Figure-4: The occluded faces are highlighted and shown

Here the overview of the database used in this technique is explained. The dataset of four different days of 30 different classes are updated and maintained in the database followed by the training phase those attributes are used for finding the results. In this technique a closed circuit camera is used to take the pictures. The lighting conditions should be constant. The respective positions of the students need to be fixed. Now the process is explained in terms of consecutive steps. In the first step, the image of the classroom is taken and given as an input to the face detection for detecting the faces and creating the bounding boxes around the faces. The sample training image is shown in the figure 2. The training is given with four different images taken in four different days and a single sample image is shown in figure 2.

Figure-3: The output of the detection algorithm

Figure-5: The result of the Occlusion Detection

3.1 PLOTTING PHASE

In the first step, a line equation [17] is calculated by using all the points which are selected from all the four training sets of images. The line equation is drawn separately for the rows and columns. A set of five points are taken from every single column of every image and with the help of those 20 points that were selected from those five images, a slope is calculated. The slope is used to plot a curve [17] on the column of every image and this curve will pass almost all the faces in a column. This line is plotted on the first column and the same process and is used to plot all the curves on all columns of an image. The same process is applied for all the rows in the image to draw the curves on all the rows. Therefore a grid is created on the image. The output of the first step is shown in step 4. A grid is formed on all the train images. Two attributes are derived from all the train images with the help of the grid on it. The attributes are the number of horizontal and vertical lines that pass through the faces. The data is derived for all the thirty classes in all the four training images. This will be used for the training stage. Proceeding to the second step, two attributes are updated in the database. The below equations are used to find the best curve which passes through all the faces in the rows and columns.

The equation 6 is used for finding the curve fitting for all the points which are selected in this step. The output along with the grid is shown in figure 6. The below line equation has been used to find the best fitting curve which passes through all the faces in a row and as well as the columns. The line equation p(x) determines the best fit curve by using all the coordinates of faces in the row and column.

$$P(x) = p_1x^n + p_2x^{n-1} + p_2x^{n-2} + p_3x^{n-3} + \cdots + p_nx + p_{n+1} \quad (6)$$

Figure-6: A grid on the image with horizontal and vertical lines.

3.2 ATTRIBUTES OF CLASSES

The second step is used to find the length and height [17,18] of the faces. Lengths and heights are calculated for all the classes in the image. The calculated lengths and heights of all the faces are updated in the database for training. The respective process is now used to find the lengths and heights of all the four training images. The below geometrical equation is used to find the length and height of each and every class.

In the second step, all the coordinates [18] are selected for calculating the length and height of every class. This process has been shown in the figure 7. The coordinates need to be selected for calculating the lengths and heights of the classes. Firstly, all the x coordinates are selected followed by the selection of all the y coordinates and finally, the z coordinates are selected. A sample snapshot is shown in thefigure as to how the pixels are selected. Firstly, all the x coordinates are selected followed by the selection of y coordinates and finally, the z coordinates are selected. Now, with the help of the equations 6 and 7, the lengths and heights of all the classes are derived and updated. The lengths (between x and y), heights (between x and z), positions coordinates, and the number of horizontal lines and vertical lines are passed through the single class.

The mathematical equations for calculating the length and height of the class are shown. The respective points are selected from the classes. Three different points are picked from the three different vertex points. The equation 7 and 8 has been used find the length and height of the classes respectively. The equations are derived from the basic coordinate geometry. The equations are listed below and numbered.

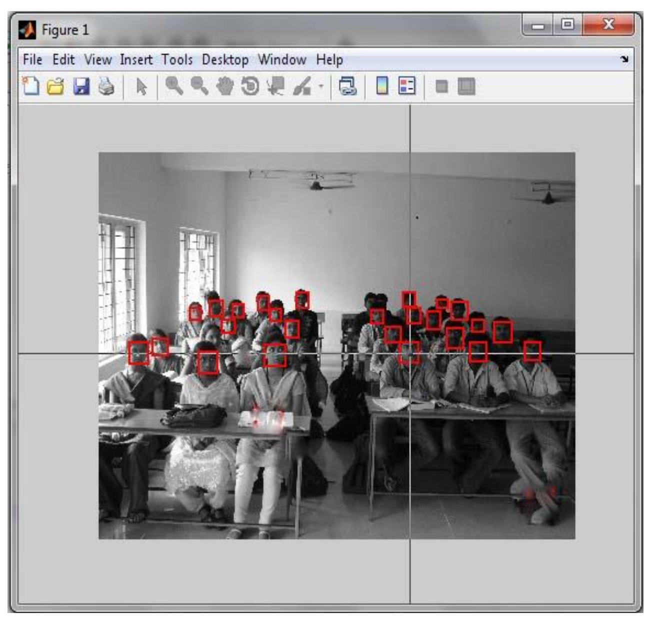


Figure-7: A snapshot to show how the coordinates are selected

$$L=\sqrt{((x(i+1)-x(i))^2+(y(i+1)-y(i))^2)} \quad (7)$$

$$H=\sqrt{((x(j+1)-x(j))^2+(y(j+1)-y(j))^2)} \quad (8)$$

Here in these equations L and H denotes the length and height of the classes(faces). The values i and j denotes number of the class(face).

3.3 SVM CLASSIFIER

SVM is a classifier derived from statistical learning theory by VapnikandChervonenkis [16,19]. SVMs are introduced by Boser, Guyon and Vapnikin COLT-92. It was initially popularized in the NIPS community but now is an important and an active field of every Machine Learning research.

Main features:

- By using the kernel trick, data is mapped onto a high-dimensional feature space without much of the computational efforts;
- Maximizing the margin achieves better generation performance;
- Soft-margin accommodates noisy data;
- Not too many parameters need to be tuned.

Support Vector Machines are an attractive approach to data modeling [19]. They combine generalization control with a technique to address the curve of dimensionality. This formulation results in a global quadratic optimization problem with box constraints, which is readily solved by interior point methods. The kernel mapping provides an unifying framework for most of the commonly employed model architectures, enabling comparisons to be performed. In classification problems, generalization control is obtained by maximizing the margin corresponding to minimization of the weight vector in a canonical framework. The solution is obtained as a set of support vectors that can be sparse. These lie on the boundary and summarize the information required to separate the data.

The SVM approach is not only well versed theoretically due to its strong base from extremely well developed machine learning theory and Statistical Learning Theory, but is also superior in practical applications. The SVM method has been successfully applied to isolated handwritten digit recognition, object recognition, text categorization, micro-array data analysis, protein secondary structure prediction, etc.

The classification problem can be restricted to consideration of the two-class problem without loss of generality. In this problem the goal is to separate the two classes by a function which is induced from available examples. The goal is to produce a classifier that will work well on unseen examples, i.e., it generalizes well. Here there are many possible linear classifiers that can separate the data, but there is only one that maximizes the margin (maximizes the distance between itself and the nearest data point of each class). This linear classifier is termed as the optimal separating hyper-plane. Intuitively, we would expect this boundary to generalize properly as opposed to the other possible boundaries.

SVM requires that each data instance is represented as a vector of real numbers.Hence, if there are categorical attributes, wemust convert them into numericdata. We recommend using m numbers to represent an m-category attribute. Onlyone of the m numbers is one, and others are zero. For example, a three-categoryattribute such as red, green, blue can be represented as (0,0,1), (0,1,0), and (1,0,0).Our experience indicates that if the number of values in an attribute is not too large, this coding might be more stable than using a single number. Scaling before applying SVM is very important. The main advantage of scaling is to avoid attributes in greater numericranges dominating those in smaller numeric ranges. Another advantage is to avoid numerical difficulties during the calculation. Due to the dependence of kernel values on the inner products of feature vectors, e.g. the linear kernel and the polynomial kernel,large attribute values might cause numerical problems. We recommend linearlyscaling each attribute to the range [□1; +1] or [0; 1].

The equations used in the SVM classifier are listed below.

Discriminate Function

the classifier is said to assign a feature vector x to class w_i if

$$g_i(x) > g_j(x) \quad \text{for all } j\neq I \quad (9)$$

For two-category case,

$$g(x)\Xi g_1(x)-g_2(x) \quad (10)$$

Linear Discriminate Function

g(x) is a linear function

$$g(x)=w^t x+b. \quad (11)$$

In the previous steps, two files are created namely, the training file and the testing file. These two files are given as an input to the SVM. Using the SVM classifier, the output created gives the missing result. These missing classes are created in another file and with the help of this file, the missing data in the image is darkened and highlighted.

3.4 EMD

In some cases the students may sit in positions different from their original assigned static positions. In this situation the proposed system will give a false output. The system will show the result with respect to the original static positions thereby misinterpreting the student who is present to be missing and hence termed absent and vice versa. This problem is the two class problemwherethe students occupying the empty front seat in the same columnthis conflict has been shown clearly in the figure 11. The class 14is occupyingclass 13, class 18 and class 17, class 22 occupying class 23, class 26 occupying class 27. This problem can be overcome by adding the recognition algorithm to the existing system

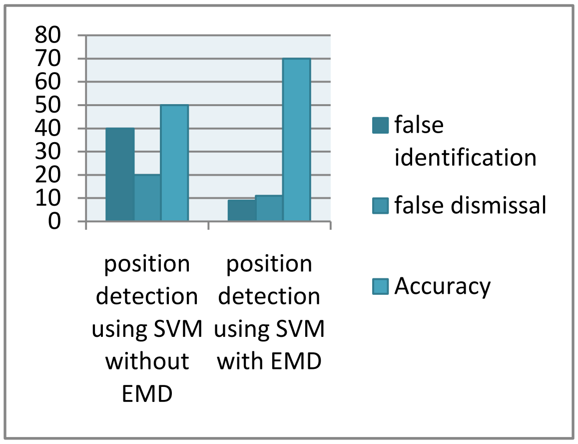


Figure-8: A bar chart to show the comparisons of Accuracy of the algorithms.

The recognition system maintains the total individual student database and then this technique will cross check whether the students and the student positions are valid or not. With this new technique the accuracy of the results is improved. In probability theory, the earth mover's distance (EMD) is a measure of the distance between two probability distributions over a region D. Informally, if the distributions are interpreted as two different ways of piling up a certain amount of dirt over the region D, the EMD [20] is the minimum cost of turning one pile into the other; where the cost is assumed to be amount of dirt moved times the distance by which it is moved . Definition is valid only if the two distributions have the same integral (informally, if the two piles have the same amount of dirt), as in normalized histograms . In that case, the EMD is equivalent to the 1st mallows distance or 1st Wasserstein distance between the two distributions. Some applications may require the comparison of distributions with different total masses. One approach is to allow for a partial match, where dirt from the most massive distribution is rearranged to make the least massive, and any leftover "dirt" is discarded at no cost.

Computing the EMD, If the domain D is discrete, the EMD can be computed by solving an distance transportation problem, which can be solved by the so-called Hungarian algorithm. In particular, if D is a one-dimensional array of "bins" [21] theEMD can be efficiently computed by scanning the array and keeping track of how much dirt needs to be transported between consecutive histogram.

$$D(I,J) = \sum_{i,j} g_{ij}\, d_{ij} \Big/ \sum_{ij} g_{ij} \qquad (12)$$

g_{ij} ground distance two bins i and j, d_{ij} distance between two bins i and j. The equation of histogram is shown in equation-12

4. RESULTS

This technique is implemented in MATLAB and some part of the technique using the c language. This new technique was implemented and run over a classroom database. The database consists of classroom images of 20 different days. The total classes (faces) present in a single image are 30. In the entire 20 days database, the positions of the thirty classes are fixed andare seated in the same location.

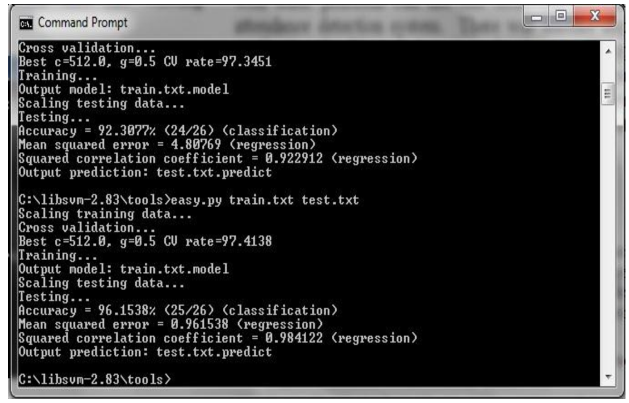


Figure-9: The accuracy of the system.

A graphical representation of an accuracy of these algorithms is shown in figure 8. Here the result is created in a text file named predict. This file will give the missing values in the training image.This resultant text file is used to highlight the missing classes in the test image. The missing data is plotted with a blue plot boxes. The intersection point of the horizontal and vertical lines will give the classes position. If the intersection does not have any class then the position will be treated as the missing position in the image.

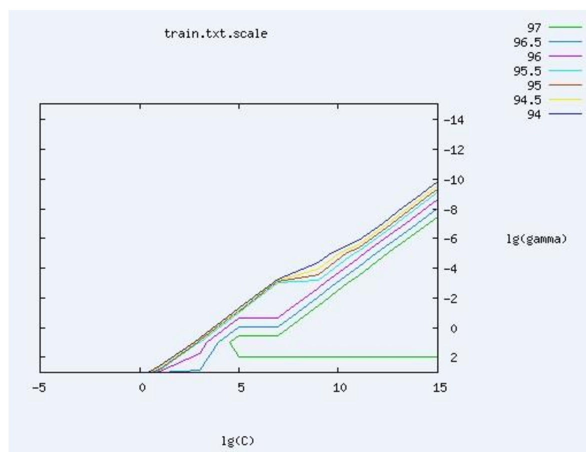


Figure-10: This plot shows the classifier training

Figure 9 and figure 10 shows the training and the training plot for the SVM classifier with the help of the Training data. Figure 13 shows the missing positions plotted in the test image. The missing positions are plotted with blue color. Figure 9 shows the accuracy of the System.

The results obtained from SVM classifier may have some errors because of rear students occupying the empty front seats in the same column. This has shown in figure 12. Such changes can be captured using EMD based face recognition algorithm.

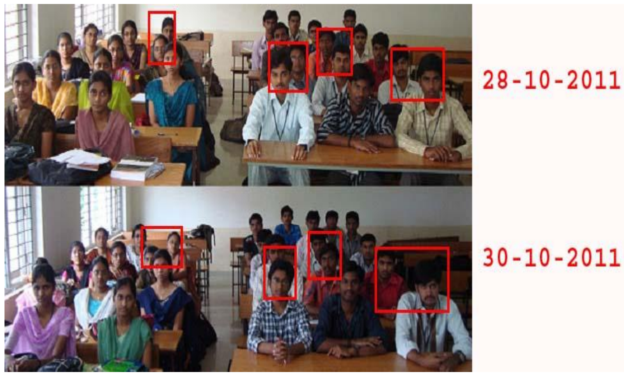


Figure-11: The Two class Problem has shown.

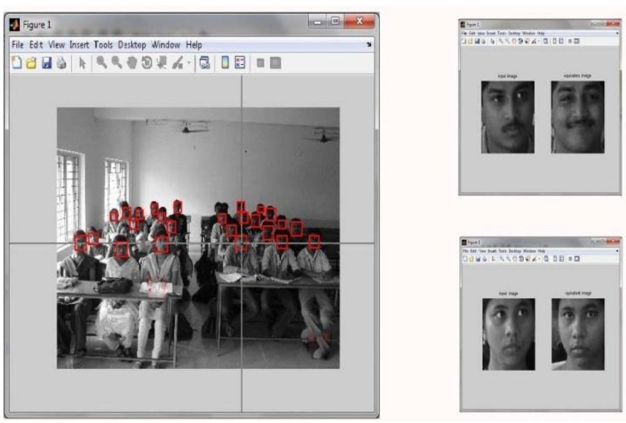


Figure-12: The result of the EMD has shown.

Figure 12 depicts EMD based facial recognition results, where the face extracted from the image is matched with the facial database.EMD is used to validate and find the two class problem. This two class problem can be validated and verified by the EMD. The results of EMD recognition have been depicted in the Figure 12.

5. CONCLUSIONS AND FUTUREWORKS

In this paper, a novel technique for detecting missing faces and properly mapping them to specific individuals has been presented. A system based on horizontal and vertical depth lines along with position coordinates has been used as input for SVM Classifier. The results are promising and a good performance is observed in spite of large number of faces and poor illumination conditions.

The problem has been solved under constraints on movement of students. In an ideal environment, any student should be able to occupy any seat and a better strategy for facial recognition has to be worked out. Also, the applications of the proposed methodology can be extended to various environments where the seating arrangements are fixed like air travel, train travel, seminar halls, laboratories, etc., and advanced curves can be employed for the same.

6. ACKNOWLEDGMENTS

Our thanks to the experts who have contributed towards development of this paper.

Prasad Reddy P.V.G.D is currently the Rector (Pro-Vice Chancellor) of Andhra University. He completed Phd in the year 1993. He has 24 years of teaching &research experience which includes more than 10 years of administrative experience with Andhra University at various senior capacities. He has so far successfully guided 12 Ph.D's, 18 M.Phil's and 259 M.Tech Projects in the field of Computer Engineering , and has to his credit, more than 104 Research papers which includes 42 publications in the referred international Journals of repute, 2 Patents filed and he has presented 19 publications in International Conferences across the world.

Sumit Gupta is currently the HOD of the CSE department in MVGR College of engineering. He Completed his Phd from NUS Singapore in Image processing and 3D Visualization research area. He has 8 years teaching experience, 11 years R & D and 3 years industry experience. He has 39 publications which include 8 national journals, 11 international journals, 11 national conferences and 9 international conferences. He was a member of professional bodies like CSI and ACM.

Suvarna Kumar G is currently working as the Senior Assistant Professor in CSE Department of MVGR College of Engineering. His Research area is Speech and video processing. He published 6 international journals and attended 2 conferences in Singapore. He has 9 years teaching experience, 2 years R &D and 1 year industrial experience. He is a member of ACM, National Science Congress, IAENG and IACSIT.

Anil Kumar R is pursuing MTech from MVGR College of engineering. His project Research Area is image processing. He published a paper in the International Conference in Machine Learning and Computing (ICMLC).

7. REFERENCES

[1] M.Turk and A. Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, Vol. 3, No. 1, Mar. 1991, pp. 71-86.

[2] The Face Detection Homepage, http://home.t-online.de/home/Robert.Frischholz/index.html.

[3] Yohei Kawaguchi, Tetsuo Shoji, Weijane Lin, Koh Kakusho,Michihiko Minoh, "Face Recognition based Attendance System" http://www.mm.media.kyoto-u.ac.jp .

[4] C. Lawrence Zitnick, Takeo Kanade, "A Cooperative Algorithm for Stereo Matching and Occlusion Detection" CMU-RI-TR-99-35.

[5] H. C. Vijaya Lakshmi, D. Patil Kulakarni "Segmentation algorithm for multiple face detection in color images with skin tone regions using color spaces and edge detection techniques," International journal of computer theory and engineering 1793-8201,2010.

[6] Rafael. C. Gonzalez, Richard. E. Woods. "Digital Image Processing. Pearson Education, Second Edition," India, 2002.

[7] Mikael Nilsson, J¨orgen Nordberg, and Ingvar Claesson, "Face detection using local SMQT features and split up SNOW classifier" 20(12):1222– 1239, November 2006.

[8] O. Lahdenoja, M. Laiho, and A. Paasio, "Reducing the feature vector length in local binary pattern based face recognition," in IEEE International Conference on Image Processing (ICIP), September 2005, vol. 2, pp. 914–917.

[9] B. Froba and A. Ernst, "Face detection with the modified census transform," in Sixth IEEE International Conference on AutomaticFace and Gesture Recognition, May 2004, pp. 91–96.

[10] M. Nilsson, M. Dahl, and I. Claesson, "The successive mean quantization transform," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), March 2005, vol. 4, pp. 429–432.

[11] M.-H. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 24, no. 1, pp. 34–58, 2002.

[12] H. Rowley, S. Baluja, and T. Kanade, "Neural network-based

face detection," in In Procedings of Computer Vision and Pattern Recognition, June 1996, pp. 203–208.

[13] H. Schneiderman and T. Kanade, "Probabilistic modeling of local appearance and spatial relationships for object recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '98), July 1998, pp. 45– 51.

[14] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in In Proceedings of the 2001IEEE Computer Society Conference on Computer Vision andPattern Recognition (CVPR), 2001, vol. 1, pp. 511–518.

[15] D. Roth, M. Yang, and N. Ahuja, "A snow-based face detector," in In Advances in Neural Information Processing Systems 12 (NIPS 12), pp. 855–861, MIT Press 2000.

[16] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines:an application to face detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '97), 1997, pp. 193–199.

[17] Math Open Reference, available at http://www.mathopenref.com.

[18] Math Open Reference, available at http://www.mathwarehouse.com

[19] Niranjan, M. "Support vector machines: a tutorial overview and critical appraisal" in Applied Statistical Pattern Recognition (Ref. No. 1999/063), IEE Colloquium on, 20 Apr 1999, ref number 1999/063.

[20] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas, "The Earth Mover's Distance as a Metric for Image Retrieval"

[21] Sylvain Boltz, Frank Nielsen, Stefano Soatto, "Earth Movers Distance On SuperPixels", Ecole Polytechnique, France,UCLA Vision Lab