ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 6, Issue 3, March 2017

Face Recognition in a Wild "A new approach- Fast detection and enhanced recognition rate"

Sanjay S Mishra
Research Scholar (PHD-10150)
Jain University
Bangalore, India

Dr. Manjunath Ramachandra
Guide
Jain University,
Bangalore, INDIA,

***Abstract*—In this paper we have presented anenhanced and optimized face detection and recognition approach to reduce processing time and at the same time to enhance recognition rate. We have used combined feature search for face detection and DCT based hybrid approach for face recognition. The proposed approach is implemented on FPGA and tested with Yale data base. Face normalization is used to make algorithm more robust.The proposed approach provide better accuracy and overall execution time is 2.5 Sec as compare to 25.7 sec of software implementation approach and hence provide 97x faster performance**

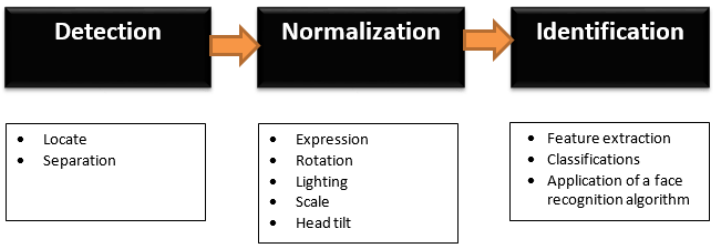
Keywords—Face recognition, Face detection, FPGA, Verilog, VHDL, Viola-Jones detector.

I. INTRODUCTION

In recent years various face recognition and detection algorithms has been proposed and many of them has provided effective results Face recognition and detection is a very wider research area and it provides immense potential to experiment new algorithms. In past many approaches has been attempted with incremental improvement. Face recognition and detection are two terms and one has to really understand the technical details of both the terms to understand actual significance. Face detection tells presence of face as an object while face recognition identifies a particular face in an image. Many face recognition algorithms are developed by considering different dynamics of environmental conditions and with some set of critical assumptions but majority of them are designed to operate on a high performance machines and realized by software approach and such implementation needs very high processing power. To run such algorithm in a real time is just next to impossible and we need to reinvent the wheel to bring new enhanced and optimized approach to run it on embedded platform or on an intelligent device. Over the years a number of successful face-image data-sets have been published to enhance and improve the performance of research on face recognition [1]. These include the Facial Recognition Technology (FERET) Database, Viola Jones object detector, the Face Recognition Grand Challenge (FRGC) facial images and depths along with companion benchmark tests and its successor, the Face Recognition Vendor Test (FRVT) database and benchmark, the CMU Pose Illumination and Expression (CMU-PIE) database and its extension, the multi-PIE database. We have focused on some of the most recent algorithms developed for real time processing and we have proposed a improved and enhance face detection and recognition technique based on modified Viola Jones detection and improved DCT based recognition algorithm, which is optimized to execute on an embedded platform with the help of FPGA and hardware accelerations with maximum feature supports and improved performance..

II. VARIOUS FACE RECOGNITION TECHNIQUES

Face recognition is basically divided into 3 segments – Detection, Normalization, and Identification.



Face Detection:

The problem of face recognition is all about face detection. This is a fact that seems quite bizarre to new researchers in this area. However, before face recognition is possible, one must be able to reliably find a face and its landmarks. This is essentially a segmentation problem and in practical systems, most of the effort goes into solving this task. In fact the actual recognition based on features extracted from these facial landmarks is only a minor last step [12].

Face detection must deal with several well-known challenges. They are usually present in images captured in uncontrolled environments, such as surveillance video systems. These challenges can be attributed to some factors.

1) Pose variation. The ideal scenario for face detection would be one in which only frontal images were involved. But, as stated, this is very unlikely in general uncontrolled conditions. Moreover, the performance of face detection algorithms drops severely when there are large pose variations. It's a major research issue. Pose variation can happen due to subject's movements or camera's angle.
2) Feature occlusion. The presence of elements like beards, glasses or hats introduces high variability. Faces can also be partially covered by objects or other faces.
3) Facial expression. Facial features also vary greatly because of different facial gestures. .
4) Imaging conditions. Different cameras and ambient conditions can affect the quality of an image, affecting the appearance of a face.

By considering all above parameters we have made a detailed study of following face detection methods

1) Knowledge-based methods
   Encode human knowledge of what constitutes a typical face (usually, the relationship between facial features)
2) Feature invariant approaches
   Aim to find structural features of a face that exist even when the pose, viewpoint, or lighting conditions vary
3) Template matching methods
   Several standard patterns stored to describe the face as a whole or the facial features separately
4) Appearance-based methods
   The models (or templates) are learned from a set of training images which capture the representative variability of facial appearance

In all of above methods, te Viola jones is very popular method for face detection,. Following are the critical steps of Viola Jones method.

1) Feature extraction of a facial part
2) Train a classifies with various facial parts
3) Use a search approach to match the facial model with portions of the image and mark those with high similarity value.

Face Normalization:

Face Normalization is one of the critical steps to increase the reliability and accuracy of a face recognition algorithm. Following are the possible face normalization steps [13].

1) Image is rotated to align the eyes (eye coordinates must be known).
2) The image is scaled to make the distance between the eyes constant. The image is also cropped to a smaller size that is nearly just the face.
3) A mask is applied that zeros out pixels not in an oval that contains the typical face. The oval is generated analytically.
4) Histogram equalization is used to smooth the distribution of gray values for the non-masked pixels.
5) The image is normalized so the non-masked pixels have mean zero and standard deviation one.

We have considered intensity normalization for our suggested approach.

Face identification:

Many researchers have done surveys on various issues and challenges in face recognition and they have highlighted pros and cons of various approaches. Broadly we can classify various face recognition approaches into 3 categories.

1) Feature based approaches
2) Holistic approaches
3) Hybrid approaches

As name indicates, feature based approach make use of facial features for face recognition which is under consideration, Holistic approach makes use of global data and take full image of human face for recognition and Hybrid makes use of both the approach specified above. In addition there are few more proven face recognition algorithms which are very common like-

1) Principal component analysis (PCA)
2) Discrete cosine transform (DCT)
3) Linear discriminant analysis (LDA),
4) Local feature analysis (LFA),
5) Local binary pattern (LBP)
6) Independent component analysis (ICA)
7) Competitive feature analysis (CFA). And many more……

PCA is one of the legacy and simplest methods. It uses mathematical calculations to extract parameters and can be used in combination with different techniques. But it has certain limitations at computational costs, accuracy and memory footprint requirements. PCA is not scale invariant and requires correlated variables. The problem becomes more difficult when there is drastic and sudden change in pose, illumination and expressions [13].

Linear discriminant analysis LDA[11] is a very effective technique and generally used in reducing dimensions and works as a classifier and specially applicable in the field of machine learning and statistics. LDA utilizes both global and local facial features and hence help in reducing singularity problem to some level and also provides better accuracy. Fisherfece projection helps in resolving illumination problem and it outperforms over PCA as it projects data on low dimensional vector space. LDA is limited by singularity problem in face recognition area also it suffers small sample size problem.

Independent component analysis (ICA) basically represents a set of random variables using basis functions. In general ICA scores high over PCA due to use of multivariate signals and also because of use of higher order statistics. In ICA approach as compare to PCA, basis vectors are more spatially local and hence gives better face representations. In ICA, as name indicates, independent components are extracted through an iterative and recursive optimization procedure hence at different point of times there could be a little variation in the outcome.

LFA describes the objects in terms of statistically derived local features and retains all advantages of PCA as well. This

approach depends on the accurate detection of the face and location of the features. This approach is not efficient and may fail in case of poor quality of image where is difficult to detect

Local binary pattern (LBP) [15] is mainly used for gray scale and rotation invariant texture and is a very simple multi-resolution approach. The unique characteristics of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes due to illumination variations. Another important advantage is its computational simplicity but at the same time it is highly sensitive to glasses and also a time consuming process. Especially this approach is not suitable for low contrast images and also for shadow images.

Discrete cosine transform (DCT) [12] is an algorithm widely used in image compression to convert image from spatial domain into elementary frequency components and it represents image as a sum of sinusoids of varying magnitude and frequencies present In that image. This method is very well suited for changing illumination conditions. This approach will segment image into non overlapped blocks and apply DCT to each block to get low and high frequency sub bands and hence reduces the computational burden. Further high frequency sub bands are removed by compression and recognition information will be present in low frequency sub bands.

CFA uses competing pattern detectors to decompose the object into characteristic sub-patterns which are called as features. This approach is hardware supportable and mainly used with many machine learning techniques.

Kernel principal analysis (KPCA) is a technique which will extract the non-linear features from the image. Also Linear PCA can use kernel method to exhibit non-linear characteristic i.e. non-linear mapping of data. This method is highly computationally intensive and takes a lot more time as compared to PCA.

Support vector machine (SVM) [3] are special category of supervised learning algorithms. It will analyzes the training data and accordingly then predicts the correct output category for given data-set input. SVM can be generalized using a small training set. Application of SVM is limited by its speed, size and complexity. In addition selection of kernel function parameter is also another biggest challenges. SVM is more suitable for face space information than PCA.

Neural networks in nothing but a computational structures of a human brain. These networks works based on “units” such as I/P units, hidden units and O\P units. Multi-Layer Perceptron (Mlp) neural networks are more robust and efficient in the case of wide facial variations than Radial Basis function (rbf) neural networks. Neural networks approach helps in training a system to capture complex class conditional densities of various face patterns and provides non-linearity in the network. Only drawback with this approach is that network architecture has to be extensively tuned. There are claim of 96.2% correct recognition on orl database of 400 images of 40 individuals and it takes less than 0.5 seconds time but it will go up as the number of person’s increases [11].

Labeled faces in the wild (LFW) is a face images database specially used for studying problems of unconstrained face recognition. The dataset is huge and contains more than 13,233 images of face collected from web. Each image contains the name of the person at the center of the image.

Milti-Pie database is a well-known and most extensive database which deals with various parameters including a expressions. This database includes images of different variation and combinations. Total images 750,000 of 337 people taken under 18 illumination conditions and with six facial expressions for a span of five months [17].

TABLE I: COMPARISON OF ACCURACY OF DIFFERENT FACE RECOGNITION APPROACHES

Sr. No.	Approaches	Working domain	Hardware support	Efficiency	Failure	Speed of operation
1.	PCA	Realtime	No	87%	13%	Moderate
2.	LDA	Integrated	No	90.8%	9.2%	Moderate
3.	ICA	Integrated	No	91%	9%	Moderate
4.	LFA	Realtime	Yes	91.5%	8.5%	Low
5.	LBP	Multi-Pie database	No	99.3%	.07%	Moderate
6.	DCT	Integrated	No	97%	3%	Low
7.	CFA	Integrated	Yes	90%	10%	High
8.	KPCA	Realtime	No	98%	2%	Slow
9.	SVM	Realtime	Yes	93%	7%	Moderate
10.	Neural Networks	Realtime	Yes	96.2%	3.8%	High
11.	LFW	Realtime	Yes	95%	5%	High

In hybrid approaches various face recognition techniques are combined to get better results in terms of higher recognition rates and accuracy.

TABLE II: COMPARISON OF RECOGNITION RATE OF VARIOUS HYBRID APPROACHES

Sr. No.	Method	Recognition rate
1.	Eigenface	90.5% [3]
2.	PCA+LDA (Fisherface)	95% [3]
3.	KPCA+SVM	95.40% [4]
4.	DCT+LDA	97.5% [3]
5.	DCT+SVM	98.90% [6]

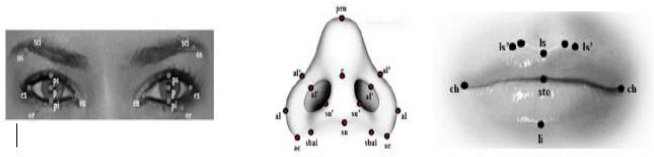
6.	LDA+SVM	95.3% [17]
7.	PCA+SVM	95.8% [17]

III. PROPOSED MODEL FOR FACE RECOGNITION

We have considered following as base for our proposed model.

1) Viola Jones for face detection
2) Intensity normalization to improve the performance
3) DCT based face recognition technique

In our enhanced proposed model the main idea is to scan a small window (feature related template) across the image and analyze the content of the template using a series of primitive features that are sensitive to facial parts; e.g., eyes, nose, and lips. In image processing/analysis, usually window-based operations are performed at fixed template (window) and on multiple scales of the input image very similar to wavelet analysis [9][10].

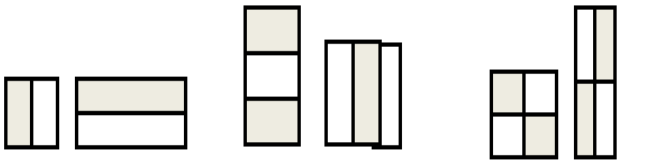


Critical facial features for face detection

It is important to understand various key aspects of Viola Jones face detection method which are described below.

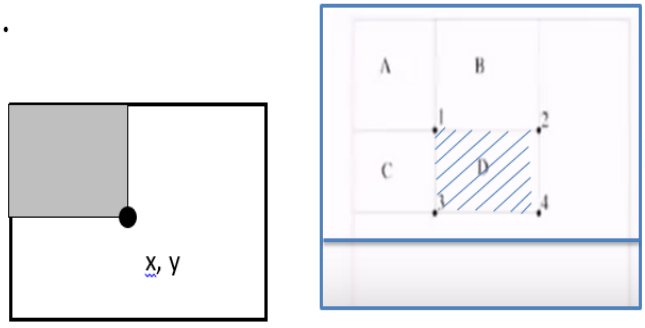
1) Haar features
2) Integral Image
3) Adaboost classification
4) Cascading

Viola-Jones use templates which are very similar to those used in the Haar Transform, known asHaar features, or Haar-like features, which are sensitive to transitions in the image (i.e., nearlyestimate the gradient). Various types are illustrated below.



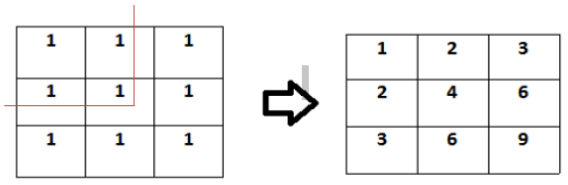
Two-rectangle features Three-rectangle features Four-rectangle features

1) Haar features are very similar to convolution kernel which are used to detect specific feature in a image
2) Harr features represents certain characteristics of face
3) Feature result is a single value which is calculated by subtracting the sum of pixels under white rectangle from the sum of pixels under black rectangle.
4) VJ algorithm will change the window size to multiple scales and rescan input image In each scale change the size of primitive features changes accordingly.
5) Base template is 24 X 24 (576 pixels) and gets enlarged to 30 X 20, 36 X 36 , 40 X 40 and 48 X48 i.e. in scale of 1.25
6) Each single type of feature is repeated all over the image in all scale, size and position so many combinations
7) If we consider all the variation of size, position of all the features , you end up calculating about 160 K + features in 24X24 window
8) To reduce time in calculating the features, they transformed the input image into a representation called the integral image, which makes the scanning invariant to scale; i.e., scans are performed at same number of operations
9) Below is demonstration of the integral image and the primitive features within a template and process which we need to follow to get integral image..
10) Image has to be gray scale image
11) Used for calculating the average intensity within a given image
12) The integral image at location (x,y), is the sum of the pixels above and to the left of (x,y), inclusive



Area under D – sample calculation is demonstrated below.

D = 1+4 – (2+3) = A + (A+B+C+D) – (A+C+A+B) = D



Original Image Integral Image

Adaboost classification:

As name indicated, Adaboost stand for adaptive boost. It is a very critical machine learning algorithm and needs good

training datasets It is used to remove redundant and irreverent features [20].

Viola and Jones run the basic classifier using templates of larger scales than the base scale of 24 X 24 , each time they enhance the quality of the decision by eliminating non-faces. The overall decision approach is known as "Cascaded Classifier". The Viola-Jones approach is trained over thousands of "face" and non-face images. In brief, the Viola-Jones algorithm performs face detection using sliding window of region at different scales, and in each scale a process of no-face elimination is performed. As the number of computations is huge (fixed per scale), the weights of the Adaboost classifier are obtained offline over tens of thousands of faces and no-face images (usually the number of no-face images is much higher than the face images). Better training results when the face images contain lots of varieties, including pose and intensity variations. An implementation of Viola-Jones existson OpenCV[19], and has been adapted in my research work.

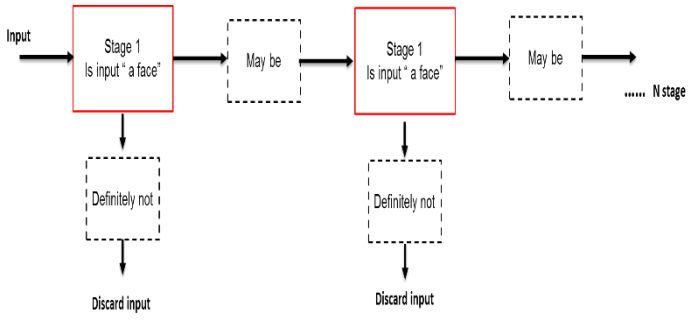
Constructs a "strong" classifier as a linear combination of weighted simple "weak" classifiers

$$F(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) + \dots$$

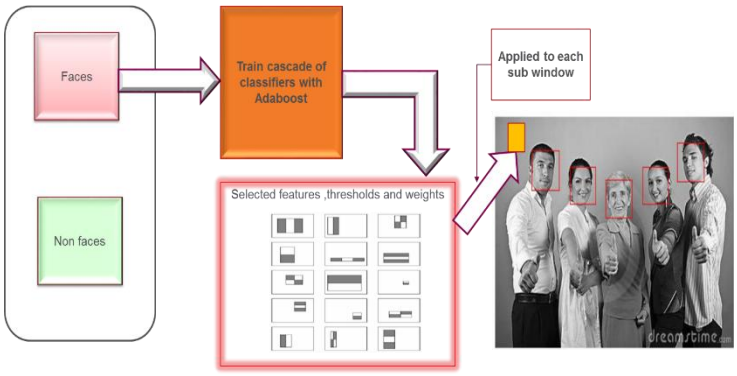
Strong classifier — Image — Weight — Weak classifier

1) Output of weak classifier – 0 (No feature identified) OR 1 (feature identified)
2) Generally 2500 features are used to define/identify strong classifier.
3) Initially, all weights set equal
4) Adaptively constructs a final strong classifier taking into account the failures of each one of the chosen weak classifiers (weight appliance)

We will arrange 2500 classifiers in cascading manner which is as shown below.

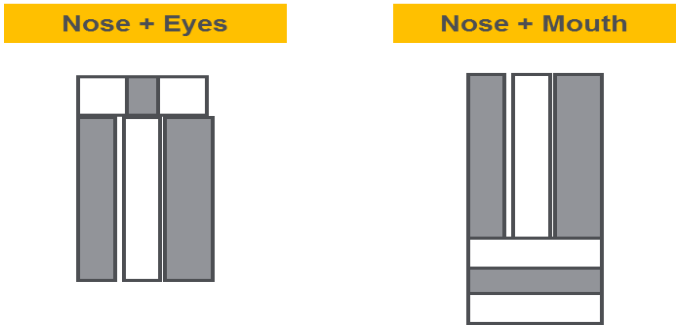


The overall process of face detection is described in below picture



IV. PROPOSED ENHANCE APPROACH:

Viola jones uses method to identify signal features at a time by using appropriate harr feature and with adaboost algorithm. To boost the face detection and to improve recognition efficiency I have introduced combined feature search. Few examples of combined feature search are as below.



We followed the below execution steps for new approach.

1) Whenever nose like feature is detected, This two features will be applied
2) Nose feature will be a center point
3) This 2 new feature window will be scaled up around center point with a factor of 1.25
4) I have reduced original classifier set from 2500 to 2000 with my new approach
5) I have added 500 classifiers based on combined feature search and tested with 3 data base.
6) It is observed that with combined feature search has improved face detection. (example, an Image where original method could not identify face, combined search has successfully identified the face).

Face normalization:

To make algorithm robust, Image can be normalized with respect to many parameters but for my research I have used intensity as a parameter to normalize the image.The average intensity of a registered image is calculated as summation of all pixel values divided by total number of pixels. Similarly average intensity of test image is calculated. We need to first calculate normalization value.

$$\text{Normalization Value} = \frac{\text{Average value of registered image}}{\text{Average value of test image}}$$

Normalized value is multiplied with each pixel of test image. Thus we get normalized image having average intensity.

Face recognition:

There are various methods to do face recognition and broadly they are classified into 3 categories.

1) Holistic matching method

   Whole face as raw input to the recognition system. Face region is represented by eigenvalues and it is based on principle component analysis (PCA)[5]

2) Feature based matching method
   It works on extraction of local features like nose, mouth, eyes. Local features and statistics are fed to the classifier

3) Hybrid method
   This method uses whole face as well local features to recognize a face

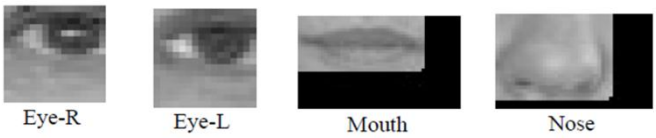
We have chosen DCT based hybrid approach for our approach and introduced a new modified approach to enhance the recognition rate.

Our proposed approach is described as described below.

1) We have considered 128X128 pixel image for our analysis.
2) Once we take a DCT of a normalized Image, Zigzag scanning – convert 128X128 to 1X16384 vector – low frequency information is present at top of vector
3) We have considered 64 coefficient for matching.
4) Based on nature of feature ,we have chosen a different dimension of window size like - Eye region is cropped with 16X16 pixel , nose 25X40 and mouth 30X50
5) Comparison is done by calculating the Euclidean distance(x) between the test image and registered image for 50 coefficients
6) Since we have 15 registered image so we get x1,x2,x3,……x15 values.
7) X are sorted in ascending order and minimum value is given a "rank 1"

$$\text{Recognition rate} = \frac{\text{Number of correctly recognized persons}}{\text{Total number of persons tested}}$$

Test image



Eye-R Eye-L Mouth Nose

V. EXPERIMENTAL RESULTS

For testing my approach, I have used following 3 data base.

1) Yale database
2) Olivetti research laboratory
3) FERRET database

Below is example image from training set.

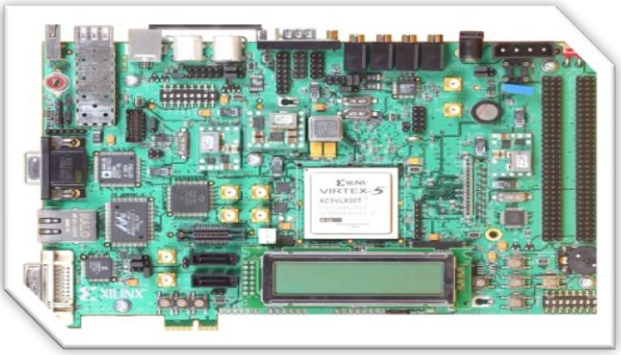
Here is image parts extracted by approach suggested by us- "Modified Viola Jones and modified hybrid DCT method" and we could find improved recognition rate which is shown below

Performance comparison between my FPGA implementation of new approach and software based approach of OpenCV is provided below. OpenCV is just a face detection and we have combined our suggested face recognition approach to compare the recognition rate.We have introduced 26 stage Harr

classified method based on suggested combined features and results are as below.

	Recognition rate	False recognition	Execution time
26 stage Haar classifiers	96.3	3.7	2.5 Sec
22 stage Haar classifiers	87.4	12.6	25.7 sec

Our new and enhance approach experiments results has shown 97 X performance improvement over software approach of openCV.



Resource consumed details of suggested approach is as below.

Resource	Available	Used	Utilization
Registers	69120	8421	12 %
LUTs	69120	18458	26 %
Block RAM	256 (36kb each block)	72	28 %
DSP 48 E	64	8	12 %

VI. CONCLUSION

The proposed approach explain about the face detection and recognition with modified viola jones and modified hybrid DCT method. Proposed method uses combination of Haar features to improve detection rate and in term improverecognition rate. The proposed approach provide better accuracy and overall execution time is 2.5 Sec as compare to 25.7 sec of software implementation approach and hence provide 97x faster performance.] We have tested our implementation against various proven data bases and could obtain consistent result.

REFERENCES

[1]. Gary B. Huang,Manu Ramesh, Tamara Berg, and Erik Learned-Miller "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments"

[2]. Sharif M., Shah J. H., Mohsin S. and Raza M., Sub-Holistic Hidden Markov Model for Face Recognition, *ResearchJournal of Recent Sciences*, **2(5)**, 10-14 **(2013)**

[3]. Zhang Yankun, and Liu Chongqing, "Efficient face recognition method based on DCT and LDA" Journal of Systems Engineering and Electronics, Vol. 15, No. 2, pp 211-216, 2004.

[4]. Guang Dai, and Changle Zhou, "Face Recognition Using Support Vector Machines with the Robust Feature" in Proc. of IEEE International Workshop on Robot and Human Interactive Communication, pp. 49-53, 2003.

[5]. Jawad Nagi, Syed Khaleel Ahmed , and Farrukh Nagi "Pose Invariant Face Recognition using Hybrid DWT-DCT Frequency Features with Support Vector Machines" Proceedings of the 4th International Conference on 17th – 19th November 2008 Information Technology and Multimedia at UNITEN (ICIMU' 2008), Malaysia

[6]. Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. University of Massachusetts, Amherst, Technical Report 07-49, October, 2007 [8]Rabia Jafri and Hamid R. Arabnia "A Survey of Face Recognition Techniques" Journal of Information Processing Systems, Vol.5, No.2, June 2009

[7]. Daniel Hefenbrock, Jason Oberg, Nhat Tan Nguyen Thanh, Ryan Kastner, Scott B. Baden "Accelerating Viola-Jones Face Detection to FPGA-Level using GPUs"

[8]. J. Cho, B. Benson, S. Mirzaei, and R. Kastner, "Parallelized architecture of multiple classifiers for face detection," in ASAP '09: Proceedings of the 2009 20th IEEE International Conference on Application-specific Systems, Architectures and Processors. Washington, DC, USA: IEEE Computer Society, 2009, pp. 75–82.

[9]. Paul Viola, Michael J. Jones, Robust Real-Time Face Detection, International Journal of Cumputer Vision 57(2), 2004.

[10]. Henry A. Rowley, Shumeet Baluja, Takeo Kanade, Neural Network-Based Face Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 20, Issue 1, Jan 1998. Digital Object Identifier 10.1109/34.655647.

[11]. Lu J., Plataniotis K. N. and Venetsanopoulos A. N., Face recognition using LDA-based algorithms, Neural Networks, IEEE Transactions on, 14(1), 195-200 (2003)

[12]. Sharif M., Ayub M. K., Raza M. and Mohsin S., Data Reductionality Technique for Face Recognition, Proceedings of the Pakistan Academy of Sciences, 48(4), 229-234 (2011)

[13]. Shah J., Sharif M., Raza M. and Azeem A., A Survey: Linear and Nonlinear PCA Based Face Recognition Techniques, The International Arab journal of Information technology, 10(6), (2013)

[14]. Zhu X. and Ramanan D., Face detection, pose estimation, and landmark localization in the wild, Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE, 2879-2886 (2012)

[15]. Cho, J., Mirzaei, S., Oberg, J., and Kastner, R. 2009. Fpga-based face detection system using Haar classifiers. FPGA 2009.

[16]. Anissa Bouzalmat, Jamal Kharroubi, Arsalane Zarghili "Comparative Study of PCA, ICA, LDA using SVM Classifier" JOURNAL OF EMERGING TECHNOLOGIES IN WEB INTELLIGENCE, VOL. 6, NO. 1, FEBRUARY 2014

[17]. Christopher M. Bishop, Pattern Recognition and Machine Learning, first edition, Springer 2006.

[18] M. Ooi, "Hardware Implementation for Face Detection on Xilinx Virtex-II FPGA UsingThe Reversible Component Transformation Color Space," in Third IEEE International Workshop on Electronic Design, Test and Applications, Washington, DC, 2006.

[19]. OpenCV:http://sourceforge.net/projects/opencvlibrary

[20] Y. Wei, X. Bing, and C. Chareonsak, "FPGA implementation of AdaBoost algorithm for detection of face biometrics," In Proceedings of IEEE International Workshop Biomedical Circuits and Systems, page S1, 2004.

[21] M. Yang, Y. Wu, J. Crenshaw, B. Augustine,and R. Mareachen, "Face detection for automatic exposure control in handheld camera," In Proceedings of IEEE international Conference on Computer Vision System, pp.17, 2006.

[22] P. Viola and M. J. Jones, "Robust real-time face detection," Int. J. Computer Vision, vol. 57, no. 2, pp.137-154, 2004.

[23] S. Paschalakis and M. Bober, "A Low Cost FPGA System for High Speed Face Detectionand Tracking," in Proc. IEEE International Conference on Field-Programmable Technology, Tokyo, Japan, 2003.