Techniques of Face Detection

Ms. Asharani K. Jawale

Mumbai University, India.
asharanishinde@gmail.com

Abstract:- Face detection plays a vital role in any Recognition and Retrieval System in Image Processing domain. As well as face detection has remarkable effects in top priority application of recognition system for image analysis, commercial security and law enforcement.

Categories and Subject Descriptors

The goal of face detection is to identify all image regions in query image which contain a face regardless of its three-dimensional position, orientation, and lighting conditions. Such a problem becomes a challenge during face detection because faces are non-rigid and have a high degree of variability in size, shape, colour, and texture. This report is a review of different face detection techniques which tries to minimize the various challenges in face detection as provides more efficient and precise output for further processing.

General Terms

Face detection challenges, Significance of face detection, Knowledge based detection , Feature invariance , Appearance based detection, Template matching .

Keywords

Top down approach, Bottom Up approach, Multi-resolution approach, Features of human face, Histogram, feature vector, classification, verification, LBP, SURF, MCT.

*****

1. INTRODUCTION

Face detection is nothing but the localization of face image in given single image. Face and facial expression recognition have attracted much attention though they have been studied for more than 20 years by psychophysicists, neuroscientists, and engineers because many research demonstrations and commercial applications have been developed from these efforts.[1],[10]

2. CHALLENGES IN FACE DETECTION

2.1 Pose: The images of a face vary due to the relative camera-face pose (frontal, 45 degree, profile, upside down), and some facial features such as an eye or the nose may become partially or wholly occluded.[19]

2.2 Facial expression: The appearance of faces is directly

affected by a person's facial expression[10]

2.3 Occlusion: Faces may be partially occluded by other objects. In an image with a group of people, some faces may partially occlude other.[19]

2.4 Image orientation: Face images directly vary for different rotations about the camera's optical axis[28].

2.5 Illumination: The varying direction and energy distribution of the ambient illumination, together with the 3D structure of the human face, can lead to major differences in the shading and shadows on the face which causes dramatic changes in the face appearance.[19]

2.6 Structural components: Facial features such as beards, moustache's, and glasses may or may not be present and there is a great deal of variability among these components including shape, colour, and size.

2.7 Imaging conditions: When the image is formed, factors such as lighting (spectra, source distribution and intensity) and camera characteristics (sensor response, lenses) affect the appearance of a face. [28]

2.8 Low Resolution: The images taken from a surveillance camera generally consists of very small face area and so its resolution is very low. Such a low resolution face image consists of very limited information as most of the details are lost.[19]

3. SIGNIFICANCE OF FACE DETECTION

Face detection is an interesting research field in the domain of Image Processing because it creates many useful application like

- Face recognition, Facial expression recognition
- Face tracking, Facial feature extraction
- Gender classification, Identification system

- Document control and access control
- Clustering, biometric science
- Human Computer Interaction (HCI) system
- Digital Cosmetics and many more

4. BASIC STEPS IN FACE DETECTION

The process of detection and localization of face in the in4.put image involves common steps as depicted in below figure:

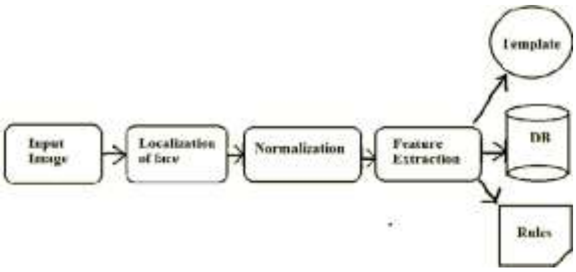


Figure 1. Basic steps in face detection[1]

First of all localization of face in input image is done by searching face candidates. Then the input image gets normalized as per the output of localization step. The normalized face part of query image used to extract different facial features for further processing as per the application of face detection in hand. Thus extracted features may get used either to form various face templates, stored in database for future verification or used to form the various face detection rules.

6. DIFFERENT TECHNIQUES

A huge number of representation techniques are available for face detection. M. H. Yang classifies them in four categories.[2]

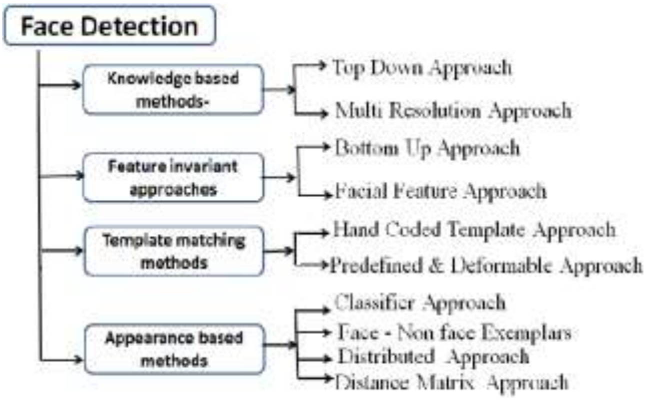


Figure.2 Different Face detection Methods

6.1.Knowledge-based methods. These rule-based methods encode human knowledge of what constitutes a typical face. Usually, the rules capture the relationships between facial features. These methods are designed mainly for face localization.

6.1.1 Top down approach:

Represent a face using a set of human coded rules. In this approach, face detection methods are developed based on the rules derived from the researcher's knowledge of human faces. It is easy to come up with simple rules to describe the features of a face and their relationships. For example, a face often appears in an image with two eyes that are symmetric to each other, a nose, and a mouth. The relationships between features can be represented by their relative distances and positions. Facial features in an input image are extracted first, and face candidates are identified based on the coded rules. A verification process is sually applied to reduce false detections.[2]

6.1.2Multi -resolution approach:

Yang and Huang used a hierarchical knowledge-based method to detect faces.[4] Attractive feature of this method is that a coarse-to-fine or focus-of-attention strategy is used to reduce the required computation. Their system consists of three levels of rules. At the highest level, all possible face candidates are found by scanning a window over the input image and applying a set of rules at each location. The rules at a higher level are general descriptions of what a face looks like while the rules at lower levels rely on details of facial features. A multi-resolution hierarchy of images is created by averaging and subsampling, and examples of the coded rules used to locate face candidates in the lowest resolution.

6.2 Feature invariant approaches. These algorithms aim to find structural features that exist even when the pose, viewpoint, or lighting conditions vary, and then use these to locate faces. These methods are designed mainly for face localization.

6.2.1 Bottom Up Approach:

This methods have been proposed to first detect facial features like eyes, eyebrows, mouth, nose, hairline and then to infer the presence of a face. Based on the extracted features, a statistical model is built to describe their relationships and to verify the existence of a face.

6.2.2 Facial Feature Based Approach:-

This approach focus on detecting the features like edges, intensity, shape, colour, texture. Sirohey proposed a localization method to segment a face from a cluttered background for face identification [5]. It uses an edge map (Canny detector) and heuristics to remove and group edges so that only the ones on the face contour are preserved.

6.3 Template matching methods. Several standard patterns of a face are stored to describe the face as a whole or the facial features separately. The correlations between an input image and the stored patterns are computed for detection. These methods have been used for both face localization and detection.

6.3.1 Hand Coded Template Approach

The standard face pattern (usually frontal), manually parameterised by function. The existence of a face is determined based on the correlation values.[2]

6.3.2 Predefined or Deformable Template Approach

Several sub-templates used for the eyes, nose, mouth, and face contour to model a face. In first phase, determines focus of attention on region of interest and in second phase, examines the details to determine the existence of a face.[2]

6.4 Appearance-based methods. These methods rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face and non-face images. Another approach in appearance-based methods is to find a discriminant function (i.e., decision surface, separating hyper plane, threshold function) between face and non-face classes.

The appearance-based approach has become the de facto standard in recent years for a variety of reasons. First, the seminal work of Viola and Jones [20] introduced a face detection framework that was sufficiently fast and accurate for many practical applications. Numerous research groups sought to modify this framework in order to improve robustness to the nuisance factors or increase its efficiency [21]. Second, continuous improvements in computing power have made it feasible to train appearance-based face detectors in reasonable amounts of time. The Viola and Jones-inspired family of face detectors and their alternatives have all benefited from this trend.

Alternative features like LBP(Local Binary Pattern). MCT(Modified Census Transfer), SURF(Speed Up Robust Feature), use of local orientation and Blob feature can be done in boosted cascade architecture of detection process.[22]

The composition and performance of an appearance-based face detection algorithm is a function of five separate components, i.e. a search strategy, a feature representation, a learning algorithm, a classifier and a detection post-processing scheme.

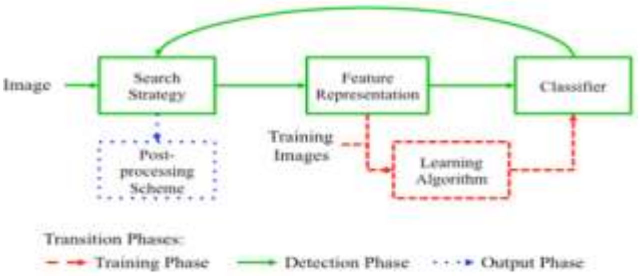


Figure3 : The interactions between the components of

an appearance-based face detector.[30]

Appearance-based face detectors typically proceed through three phases. During the training phase, a learning algorithm induces a classifier using a set of images that were transformed into a suitable feature representation for face detection. During the detection phase, the detector iterates over an input image according to a search strategy, passing transformed image patches to the trained classifier. The classifier decides whether the image patch for the current iteration contains a face. During the output phase, a post-processing scheme produces a detection output list after merging detections that appear to cover the same face and removing likely false detections.

In appearance based methods representation of input image may be either Holistic or Block based.

- **Holistic:** Each image is raster scanned and represented by a vector of intensity values or
- **Block -based:** Decompose each face image into a set of overlapping or non-overlapping blocks

6.4.1 Classification Approach

Image patterns are projected to a lower dimensional space and then a discriminant function is formed (usually based on distance metrics) for classification [11], or a nonlinear decision surface can be formed using multilayer neural networks.

These methods implicitly project patterns to a higher dimensional space and then form a decision surface between the projected face and non-face patterns.

Examples of some classifier are:

- SVM-Support Vector Machine
- HMM- Hidden Markov Model
- Neural Network, CNN-Convolution Neural Network
- Naïve Baysed Classifier
- SNoW-Spars Network Of Winnows
- Adaboost….etc

6.4.2 Face Non-Face approach:-

This is also known as Eigen Face approach, where eigenvectors of covariance matrix computed from vectorised face images in training set. Eigenvectors are then used to form the face and non-face clusters using K-means algorithm[12].

Many works on face detection, recognition, and feature extractions have adopted the idea of eigenvector decomposition and clustering.

6.4.3 Distribution Based Approach:-

Sung and Poggio developed a distribution-based system for face detection which demonstrated how the distributions of image patterns from one object class can be learned from positive and negative examples (i.e., images) of that class. Their system consists of two components, distribution-based models for face/non-face patterns and a multilayer perceptron classifier.[12]

The Eigen space decomposition as an integral part of estimating complete density functions in high-dimensional image spaces. These density estimates were then used in a maximum likelihood formulation for target detection[13]

6.4.4 Distance Matrix Approach:-

The two-value distance metric should also produce classification results that are at least as good as those obtained with a standard Mahalanobis distance metric, because both metrics are based on very similar Gaussian generative models of the local data distribution. With network-based classifiers, the two-value distance metric actually out-performs the standard Mahalanobis distance metric consistently[12]

- **Feature vector**: Each face/non-face sample is represented by a vector of these distance measurements

ISSN: 2321-8169
04 – 10

- **Train a multilayer perceptron** using the feature vectors for face detection[12]

7. REVIEW OF LITERATURE

M H Yang gives a detail review of face detection methods with examples of each category of face detection in [2].Some of them are listed below:

7.1‘Rule-based Face Detection In Frontal Views’[3]

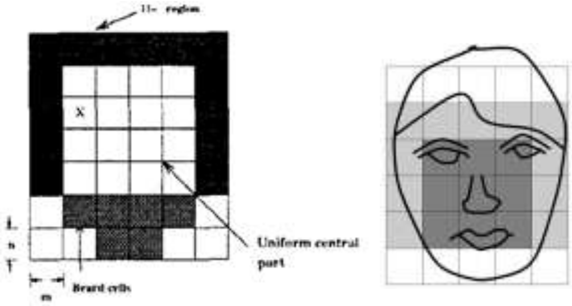


Figure 5. Rule based face detection model

The decision whether or not, a region of 4x4 cell is a facial candidate is based on:-

- the detection of a homogeneous region of 2 x 2 cells in the middle of the model
- the detection of homogeneous connected components having significant length
- the detection of a beard region

7.2 ‘Human Face Detection in Complex Background’ [4]

A multi-resolution hierarchy of images (i.e mosaic images) are created by averaging and subsampling, as shown in following figure, and the coded rules are used to locate face candidates in the lowest resolution and these are further processed at finer resolutions.

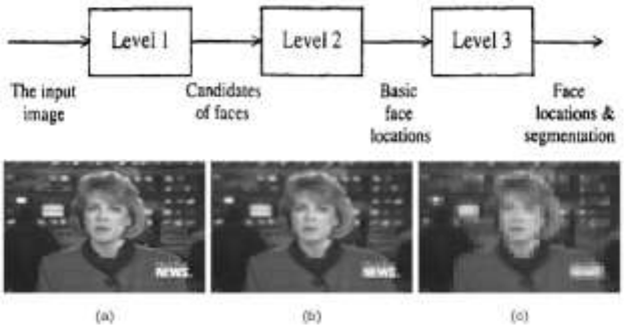


Figure 6. Multiresoulution hierarchy of image.

The lowest resolution (Level 1) image is searched for face candidates and these are further processed at finer resolutions. At Level 2, local histogram equalization is performed on the face candidates received from Level 1, followed by edge detection. Surviving candidate regions are then examined at Level 3 with another set of rules that respond to facial features such as the eyes and mouth.

7.3 ‘Low-level Features For Image Retrieval Based On Extraction Of Directional Binary Patterns And Its Oriented Gradients Histogram’ [6]

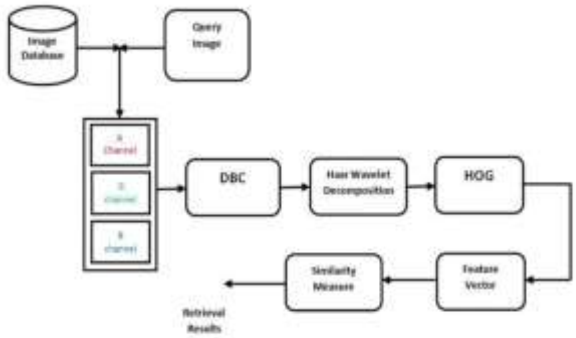
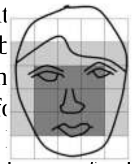


Figure 7. Flow diagram of DBP and HOG method of invariant feature extraction [6]

Above flow diagram shows how colour, texture and shape like invariant features are used for detection. The colour texture features are extracted from each channel of colour image using DBC technique. Then use of Haar wavelet transform [23] on colour texture map and its original colour image in order to extract the local information for enhancing local contrast. Another advantage of the Haar Wavelet transform is that it reduces dimensionality by preserving more texture as well as colour and shape information in the form of coefficients obtained by dividing the image into four sub bands. Haar wavelet is used due to its simplicity and computational efficiency.

Histograms of Oriented Gradients (HOG) technique proposed by N. Dalal et al. [27] on each sub-band of wavelet transformed image. The concatf computed histograms of all the four sub-b s the HOG descriptor, which is stored the sh ll as texture information, which can be used fo based image retrieval. Finally the values gain by iptor helps to calculate the feature vectors which are further used for similarity measurement and produce the final results of retrieval.

7.5Eye Finding via Face Detection for a Foveated, Active Vision System[6]

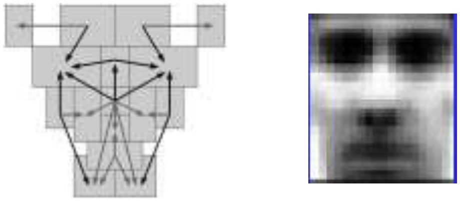


Figure 8. A 14x16 pixel ratio template for face localization based on Sinha method [6].

The template is composed of 16 regions (the grey boxes) and 23 relations (shown by arrows).[7]

Uses relative pairwise ratios of brightness in facial region like eyes are usually darker than surrounding area.

Sinha used a small set of spatial image invariants to describe the space of face patterns. His key insight for designing the invariant is that, while variations in illumination change the individual brightness of different parts of faces (such as eyes, cheeks, and forehead), remain largely unchanged. Determining pairwise ratios of the brightness of a few such regions and retaining just the "directions" of these ratios (i.e., is one region brighter or darker than the other?) provides a robust invariant. Thus, observed brightness regularities are encoded as a ratio template which is a coarse spatial template of a face with a few appropriately chosen sub regions that roughly correspond to key facial features such as the eyes, cheeks, and forehead. The brightness constraints between facial parts are captured by an appropriate set of pairwise brighter-darker relationships between sub regions. A face is located if an image satisfies all the pairwise brighter-darker constraints. The idea of using intensity differences between local adjacent regions has later been extended to a wavelet-based representation for pedestrian, car, and face detection.[2]

7.6 An Automatic Face Identification System Using Flexible Appearance Models[8]

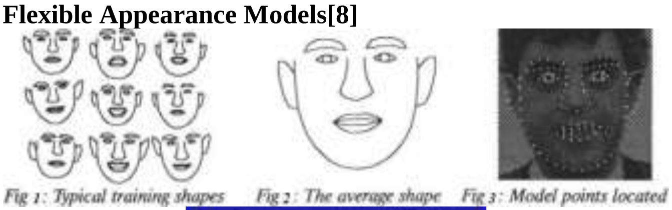


Figure 9. PDM for Template Matching Methods[8]

Point Distribution Model (PDM) is used to characterize the shape vectors over an ensemble of individuals. Yuille et al. used an energy function is defined to link edges, peaks, and valleys in the input image to corresponding parameters in the template. The best fit of the elastic model is found by minimizing an energy function the parameters.[9]

A face-shape PDM can be used to locate faces in new images by using active shape model (ASM) search to estimate the face location and shape parameters. The face patch is then deformed to the average shape, and intensity parameters are extracted. The shape and intensity parameters can be used together for classification.[2]

7.11 Face Recognition with Support Vector Machines[16]

SVMs can be considered as a new paradigm to train polynomial function, neural networks, or radial basis function (RBF) classifiers. While most methods for training a classifier (e.g., Bayesian, neural networks, and RBF) are based on of minimizing the training error, i.e., empirical risk, SVMs operates on another induction principle, called structural risk minimization, which aims to minimize an upper bound on the expected generalization error. An SVM classifier is a linear classifier where the separating hyper plane is chosen to minimize the expected classification error of the unseen test patterns. This optimal hyper plane is defined by a weighted combination of a small subset of the training vectors, called support vectors.[2]

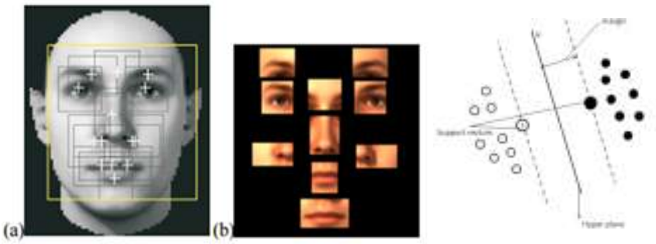


Fig20: Face detection with SVM[16]

8. ANALYSIS

8.1 Knowledge Based Method

Pros:

Simple Coded Rules

Work well on Face localization in uncluttered background

Cons:

Difficult to form rules precisely

Difficult in different Pose

Unable to detect multiple faces

8.2 Feature Invariant Method

Pros:

Invariant to pose and orientation

Work well on face localisation

Cons:

Difficult to locate facial features due to several corruption (illumination, noise, occlusion)

Difficult to detect feature in complex background

8.3 Template Matching Method

Pros:

Simple

Easy to implement

Cons:

Templates needs to initialize near face

Difficult to enumerate templates for different pose

8.4 Appearance Based Method

Pros:

Results in good empirical results

Fast and fairly robust

Able to detect face in different pose and orientation.

Cons:

Usually needs over space and scale.

Lots of –ve and +ve examples are required

Limited View based approach

9. CONCLUSION

Face images are non-rigid object and it is exciting to see face detection techniques be increasingly used in real-world applications and products.

It is a first step towards the Automatic Face recognition system in much surveillance system.

This study provides a comprehensive survey of research on face detection and to provide some structural categories for the methods.

As the Face Detection is the first step of Image Retrieval System, its accuracy and performance affects significantly on the overall computational efficiency of Retrieval System.

In competitive environment of recent trends like IOT the comparative study of face detection methods will help to improve the results of existing technique and minimize the ratio of false positive detection in many application system.

10. REFERENCES

[1] 'Face detection and its applications', by Mrs. Sunita Roy and Mr. Susanta Podder, International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 2, A[pril-May, 2013

[2] 'Detecting Faces in Images:'A Survey,'by Ming-Hsuan Yang, Member, IEEE, David J. Kriegman, and Narendra Ahuja, IEEE transactions on pattern analysis and machine intelligence, vol. 24, no. 1, January 2002

[3] C. Kotropoulos and I. Pitas, "Rule-based face detection in frontal views," Proc. Int"l Conf. Acoustics, Speech and Signal Processing, vol. 4, pp. 2537-2540, 1997

[4] G. Yang and T. S. Huang, "Human Face Detection in Complex Background," Pattern Recognition, vol. 27, no. 1, pp. 53-63, 1994.

[5] S.A. Sirohey, "Human Face Segmentation and Identification," Technical Report CS-TR- 3176, Univ. of Maryland, 1993.

[6] Sinha, P. 1996. Perceiving and recognizing three dimensional forms. Ph.D. Dissertation, Massachusetts Institute of Technology

[7] B. Scassellati, "Eye Finding via Face Detection for a Foveated, Active Vision System," Proc. 15th Nat'l Conf. Artificial Intelligence, 1998

[8] A. Lanitis, C.J. Taylor, and T.F. Cootes, "An Automatic Face Identification System Using Flexible Appearance Models," Image and Vision Computing, vol. 13, no. 5, pp. 393-401, 1995.

[9] A. Yuille, P. Hallinan, and D. Cohen, "Feature Extraction from Faces Using Deformable Templates," Int'l J. Computer Vision, vol. 8, no. 2, pp. 99-111, 1992.

[10] Daniel Haase, Erik Rodner, and Joachim Denzler, 'Instance-weighted Transfer Learning of Active Appearance Models', IEEE Conference on Computer Vision and Pattern Recognition,2014

[11] M. Turk and A. Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[12] K.-K. Sung and T. Poggio, "Example-Based Learning for View-Based Human Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 39-51, Jan. 1998.

[13] B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for Object Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 696-710, July 1997.

[14] Ridhi Jindal, Anuj Gupta and Dr. Sonia Vatta, ' Face Detection using Digital Image Processing', International Journal of Advanced Research in Computer Science and Software Engineering 3(11), November - 2013, pp. 1572-1574

[15] "Face Recognition: A Convolutional Neural-Network Approach", Steve Lawrence IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 8, NO. 1, JANUARY 1997

[16] Face Recognition with Support Vector Machines: Global versus Component-based Approach" , by Bernd Heisele Purdy Ho Tomaso Poggio, iccv2001.pdf

[17] Ms. Varsha Gupta1, Mr. Dipesh Sharma, 'A Study of Various Face Detection Methods' International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, Issue 5, May 2014

[18] X. Zou, J. Kittler, and K. Messer, "Illumination invariant face recognition: A survey," Proc IEEE Conference on Biometrics: Theory, Applications and Systems, Sept 2007, pp.1-8.

[19] Ripal Patel, Nidhi Rathod, Ami Shah, 'Comparative Analysis of Face Recognition Approaches: A Survey', International Journal of Computer Applications (0975 – 8887) Volume 57– No.17, November 2012.

[20] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. Proc. IEEE Conference on Computer Vision and Pattern Recognition, 24(1):511 –518,2001

[21] C. Zhang and Z. Zhang. 'A survey of recent advances in face detection.' Microsoft Research Technical Report MSR-TR-2010-66, 2010.

[22] Gallery-Free Methods For Detecting And Recognizing People And Groups Of Interest "In The Wild", by Jeremiah R. Barr, Graduate Program in Computer Science and Engineering Notre Dame, Indiana, December 2014.

[23] H.A. Moghaddam, T.T. Khajoie, A.H. Rouhi and M.S. Tarzjan, "Wavelet correlogram: A "New approach for image indexing and retrieval", Pattern Recognition, vol. 38, issue 12, pp. 2506-2518, 2005..

[24] H.A. Moghaddam and M.S. Tarzjan, "Gabor Wavelet Correlogram Algorithm for Image Indexing and Retrieval", Proceedings of 18th International Conference on Pattern Recognition, ICPR 2006, vol. 2, pp. 925-928, 2006.

[25] R.M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classifcation",IEEE Transactions on Systems, Man and Cybernetics, vol. SMC-3, no. 6, pp. 610-621, 1973.

[26] T. Ojala, M. Pietikainen and D. Harwood, "A comparative study of texture measures with classification based on feature distributions", Pattern Recognition, vol. 29, no. 1, pp.51-59,1996.

[27] N. Dalal and B. Triggs, "Histograms of Oriented

ISSN: 2321-8169
04 – 10

Gradients for Human Detection", In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 2, pp. 886-893, 2005.

[28] Face detection_tutorial.pdf