Detection of Human Faces in Video Sequences using NGLBP Feature

A. Jainul Fathima, P. Ithaya Rani, T. Hari Prasath

Abstract: Machine analysis of face detection is an interesting topic for study in Human-Computer Interaction. The existing studies show that discovering the position and scale of the face region is difficult due to significant illumination variation, noise and appearance variation in unconstrained scenarios. This paper suggests a method to detect the location of face area using recently developed YouTube Video face database. In this work, each frame is formulated by normalization technique and separated into overlapping blocks. The Gabor filter is tuned to extract the Gabor features from the individual blocks. The averaged Gabor features are manipulated and local binary pattern histogram features are extracted. The extracted patterns are passed to the classifier with training images for face region identification. Our experimental results on YouTube video face database exhibits promising results and demonstrate a significant performance improvement when compared to the existing techniques. Furthermore, our proposed work is uncaring to head poses and sturdy to variations in illumination, appearance and noisy images

Keywords— Ensemble classifier, Gabor wavelet, Human-computer interaction, Local Binary pattern, Normalization

I. INTRODUCTION

One of the most interesting fields of image analysis is the automatic identify the area of the human faces. The major applications of finding the face areas are face recognition, facial expression recognition, gender identification, face registration, human-machine interaction, surveillance, etc. Face discovery methods identify the faces in video sequences and estimates the position and scale of all real-time faces, But finding the human face region is a difficult task. Because the human faces appear in different backgrounds (simple, clutter) and nonrigid have inconsistency in terms of poses, expressions and illuminations (good and bad) [1-2]. To overcome these problems, this paper suggested a technique to detect the face region by applying Normalized mean of Gabor LBP features (NGLBP). This methodology is not sensitive to head poses and vigorous to variations in lighting condition and noisy images. The remaining part of the paper is organized as follows: Section II describes the related works. Section III suggested details of NGLBP features. The experimental setup and implementation are discussed in next section. The last section concludes the paper and suggested future direction of work

A. Jainul Fathima*, Department of Computer Science and Engineering, kalasalingam Academy of Research and Education, Krishnankoil. Email: jainulfathima.a@klu.ac.in

P. Ithaya Rani, Department of Computer Science and Engineering, Sethu Institute of Technology, kariapatti.
Email: muhil_rani@gmail.com

T. Hari Prasath, Department of Electronics and Communication, Kamaraj College of Engineering and Technology, Virdhunagar.

II. LITERATURE SURVEY

The research on face detection by machine analysis has been carried out by many research groups. Here, we have reviewed some of the papers related to our work. The methodology for identifying the face area utilizing skin color and the Maximum Morphological Gradient Combination image was exhibited [3-4]. The system failed when it manages with skin color areas including similar color background and region of dress. H. Sagha et.al proposed a methodology for discovering sparse features using a genetic algorithm for multi-view face detection. Notwithstanding, discovering these features was time intensive and wasteful by utilizing their strategies [5]. Face detection using Gabor Filter are explained briefly in the following research [6-8]. The techniques using the LBP and Local Gradient Pattern (LGP) based features for detecting the faces was existed [9-11]. These techniques are sensitive to noise at each location and compare a central pixel with neighboring pixels. The detection of the facial components utilizing speeded up robust features presented in [12] could achieve only moderate performance. The extracts of Haar feature and a learning algorithm (Adaboost) are proposed in [13], where the methods suffer from global illumination variations. KyungjoongJeong et al [14] carried out the work Semi – LBP (SLBP) features for face detection. These features are robust against noise. However, higher detection rate could not be achieved

Most of the existing detection systems uses particular type feature extraction. However, for discovering the area of human face, a particular feature set is not enough to obtain all the information required to detect the face. The robust detection always requires appropriate information on illumination, face appearance variation, and discriminating power of the feature set demanding more than one type of feature set. Finding and fusing relevant feature sets forms an active research topic for pattern recognition. Combining the GF and LBP features for face recognition is motivated for the work reported in [15]. We propose to combine GF and LBP features for discovering the portion of face area

The work considers the local appearance descriptors by Gabor Wavelet (GW) proposed in [6-8] and fusing it with Local Binary Pattern (LBP) features as proposed in [14] rather than working on individual feature set. The facial shape and appearance information are encoded by GF feature extraction. The detection of LBP feature captures small appearance details and tolerance to illumination
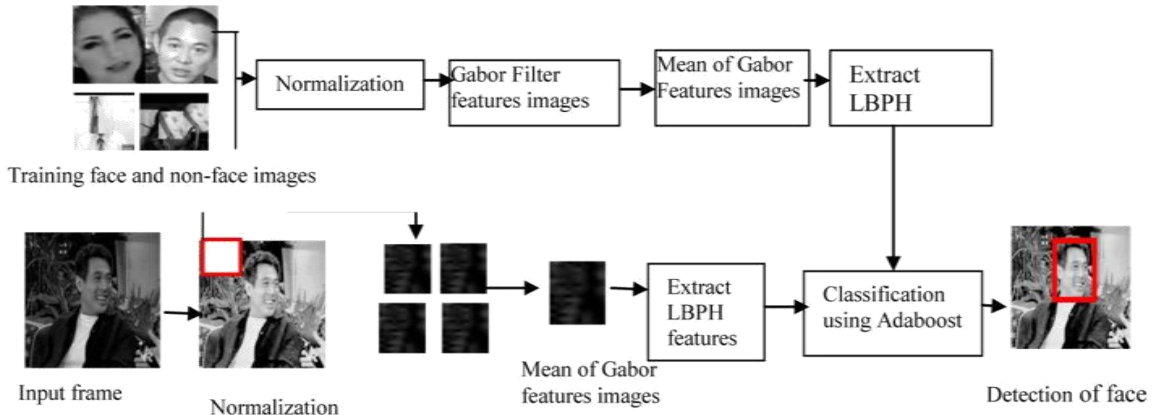
Detection of Human Faces in Video Sequences using NGLBP Feature

Fig.1 Overview of the system diagram for identifying the location of face area where red color box denote as block

ranges. Local spatial invariance is accomplished by histogramming the texture code results. The advantage of NGLBP features are used to capture the local structure corresponding to spatial frequency (scale), Spatial localization and orientation selectivity which are proved to be discriminative face, non-face and robust to illumination, noise and appearance changes

III.PROPOSED WORK

The video with single subject contains multiple frames depicting the temporal variations in different poses expressions following steps describe the proposed approaches:

1. Initially, normalization techniques are applied on each frame which adjusts the image intensity.
2. Subsequent, each frame is divided into overlapping blocks and then local features are extracted by using GF with selected scale and orientation in each block.
3. 59-LBPH features are extracted from each block contains Gabor features.
4. The Gabor LBP features are passed through the AdaBoost classifier for the pixel-wise classification with well-trained face and non-face signature.

The performance of the proposed work NGLBP features is compared with the features extracted by conventional GF [14], LBP [9] and GLBP by deploying the Ensemble Classifiers. For conducting and evaluating the work, YouTube (YT) video face databases [15] are taken. The following sub-sections describe the procedure in detail.

A. Normalization
Normalization techniques can be considered in an automatic detection system. So the histogram normalization technique [12] was applied on each frame to compensate for different lighting conditions. As the low- contrast image's histogram is narrow and centered toward the middle of the gray scale, if we distribute the histogram to a wider range, the quality of the image will be improved. So we can do it by adjusting the probability density function of the original histogram of the image so that the probability spread equally. It is used to produce an image with distributed brightness levels over the image. Initially, each frame is extracted and represented as set of frames $\{f_1, f_2, \ldots f_k\}$ from video V, where k is the number of frames. The gray levels of the kth frame is first equalized by

$$f_k = E(r_k) = \sum_{i=0}^{m} \frac{n_i}{n} m = 0, 1, \ldots \ldots, L - 1 \text{"} \qquad (1)$$

where E denote the equalization function, n is the total number of pixels, n_i is the number of pixels with gray level r_i and L is the number of discrete gray levels.

B .Feature extraction of Gabor Filter
Following the intensity normalization, Gabor filters has been used for various image processing application as shown in the research [8]. The 2D Gabor filter is adopted and it can be mathematically expressed as:

$$\psi(a, b, \sigma, \theta) = \exp\left(-\frac{(A^2 + \gamma^2 B^2)}{2\sigma^2}\right)\left(\cos\left(\frac{2\pi}{\lambda} A\right)\right) \qquad (2)$$

$$A = a\cos\theta + b\sin\theta; \qquad B = -a\sin\theta + b\cos\theta$$

where orientation θ, the effective width σ, the wavelength λ is the spacing factor between filter in the frequency domain, the aspect ratio ϒ. We propose the GF procedure by dividing the kth frame (f_k) into overlapping blocks 'B' represented as B_0 , B_1 , ...$B_{m-30*n-30}$. The number of blocks are (m-30)*(n-30) and m and n are the number of rows and columns in each frame respectively. Typically each 'B' size is 30x30 within a frame, its convolution with a Gabor filter Ψ is defined as follows.

$$GF_{k, B, \sigma, \theta}(a, b) = f_{k, B}(a, b) \Theta \Psi(a, b, \sigma, \theta) \qquad (3)$$

where Θ is the symbol for convolution. Eight scales σ ϵ (5 to 19 with increments by 2) and four orientations θ ϵ (-45º, 90º, 45º, 0º) are used in the design of Gabor filters. The given 'B' block within a input frame f_k is filtered with the Gabor filters as in Eq. (3), resulting in a series of Gabor filtered images with features such as bars and edges usefully emphasized for better identifying the location of the face. Then extracted features are converted into Mean (M) of Gabor features in each block. The performance of the proposed mean of GF features is compared with the Standard deviation (S) and Variance (V) of GF features in each block as

$$G_{k, B, M}(a, b) = \frac{1}{\sigma * \theta}\sum_{\sigma=1}\sum_{\theta=1}^{84} \qquad (4)$$

$$G^{k,B,V}(a,b) = \left(\frac{1}{\sigma*\theta} \sum_{\sigma=1}^{8} \sum_{\theta=1}^{4} \left(GF^{k,B,\sigma,\theta}(a,b) - \left(\frac{1}{\sigma*\theta} \sum_{\sigma=1}^{8} \sum_{\theta=1}^{4} GF^{k,B,\sigma,\theta}(a,b) \right) \right) \right) \quad (5)$$

$$G^{k,B,S}(a,b) = \left(\sum_{\sigma=1}^{8} \sum_{\theta=1}^{4} \left(GF^{k,B,\sigma,\theta}(a,b) - \left(\frac{1}{\sigma*\theta} \sum_{\sigma=1}^{4} \sum_{\theta=1}^{8} GF^{k,B,\sigma,\theta}(a,b) \right) \right) / \sigma*\theta \right)^{1/2} \quad (6)$$

As a result, each block contains M, S, and V of Gabor features.

C. Feature extraction of LBP

The Gabor features of each block size are 30x30 resolutions. Then GLBP is defined by a binary coding function [29] to the obtain Gabor features in each block. Let $G_{k,B}(a,b)$ be the Gabor features in 'B' block within k_{th}

frame around pixel (a,b). The center value of 3x3 matrixes is compared with another eight values and an 8bit code is coined, which will be the value at each pixel position (a,b).

Let M to represent the matrix as:

$$M = G_{k,B}(a,b) \quad (7)$$

The value by using the proposed method GLBP is obtained as:

$$f_{GLBP}(a,b) = \sum_{a=0}^{2} \sum_{b=0}^{2} T(a,b) 2^{8} \quad \text{and}$$

$$\text{where} \quad a \neq 1 \text{ and } b \neq 1 \quad (8)$$

$$T(a,b) = \begin{cases} 1 & M(a,b) \geq M(1,1) \\ 0 & \text{else} \end{cases} \quad (9)$$

After fixing the value using GLBP method for each

pixel associated with a block, a 59-bin histogram is applied to extract the feature for the each block. A histogram (H) of the block f_{GLBP} can be defined as:

$$h_L = h_L + I(lower_L < f_{GLBP}(a,b) \geq higher_L) \quad (10)$$
$$1 \leq a \leq m - 30, 1 \leq b \leq n - 30, 1 \leq L \leq 59$$

where L is the number of bins for the values produced by

the GLBP operator. The interval of each bin is represented by the range lower and higher.

$$I(A) = \begin{cases} 1 & A \text{ is true} \\ 0 & A \text{ is false} \end{cases} \quad (11)$$

We obtained 59-GLBP histogram bins for each block in

the frame.

D. Classification

Adaboost algorithm is implemented for object detection to classify face/nonface region. The benefit of this algorithm is that it shows low false positive rate. The implementation and advantages of this algorithm is clearly proposed in the literature [16]. Initially the training set is formed using the NGLBP signatures from the collection of training images having both face and non-face images and stored in database (DB1). The signatures of the each frame are classified in Face /, non-face, region with, DB1, utilizing pixel-wise,

classification of boost algorithm.

IV. EXPERIMENTAL DATA, RESULTS AND DISCUSSION

The proposed method is evaluated with YT video datasets. The samples in the video datasets are noisy and do not have quality image frames due to the low resolution. The dataset consists of 1910 video clips with hundreds of frames for each clip. Out of 1910 video sequence studies, 1870 of them consists of only one person and the remaining have more than one person. For training purpose, 805 face images and 1023 non-face images are collected from ORL, Yale databases and background imaged respectively. The sample of face, non-face images are depicted as gallery images as given in Fig 2

Fig. 2 Sample of gallery images: (a) face images, (b) non-face images

A. Performance of feature extraction

The gallery of face/non-face images are collected as a data, set for the study. The gallery images are rescaled into three, types of resolution such as 25x25, 30x30 and 35x35 pixels, NGLBP, GLBP, LBP and GF are extracted from each, gallery image. Here, Gabor filter decides the effectiveness, of extracted features,

The, comparative, chart, depicts the, , detection results for, varying orientation of Gabor filter as given in Fig 3. The, various sets of orientations are (90º, 0º), (-45º, 90º, 45º, 0º),, (-45º, -22.5 º, 0º, 22.5º, 45º, 90º) and (90º, 67.5º, 45º, 22.5º,, 0º,, -22.5º,, -45º,, -67.5º),, , correspondingly. The, highest, accuracy of 89% is obtained for using the default orientation, values. The discrimination information can be captured with,
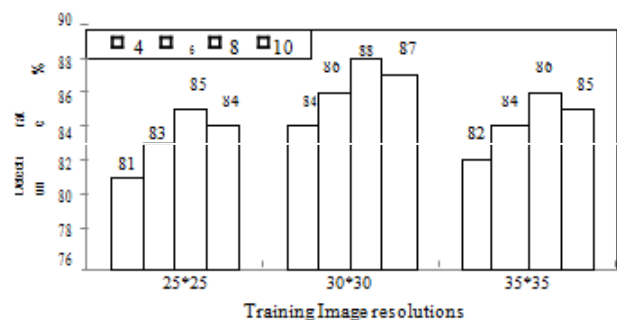
Gabor filter with 4 oritations

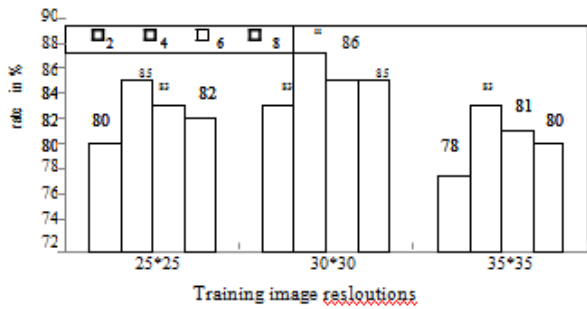**Fig.3. Four sets of orientation in YT database.**

Fig. 4 Four sets of scales in YT database.

The comparison chart shows the accuracies attained using 4, 6, 8, and 10 scales of Gabor filters (the number of orientations of Gabor filters is fixed at four) in each resolution training images using GA classifier Fig 4.. The four sets of scales are composed of (5:2:11), (5:2:15), (5:2:19), and (5:2:23) pixels correspondingly. The results also confirm that the default 4 orientations and 8 scales of Gabor filters are the optimal parameters for detecting the face area. In both Fig 3 and Fig 4, 30x30 resolution training images shows the best result.

Table 1 Comparison of accuracies obtained using GW and NGW

Extracting Features	Detection rate in %				
	Mean	Std	Variance	Skeweness	Kurtosis
GW	80	79	81	82	81
NGW	85	84	84	84	83

Table1 compares the accuracies obtained using mean, standard deviation, skewness and variance features in training images of 30*30 resolutions. From the result, it can be observed that mean of Gabor features result in better detection of face in 30x30 resolutions training images.
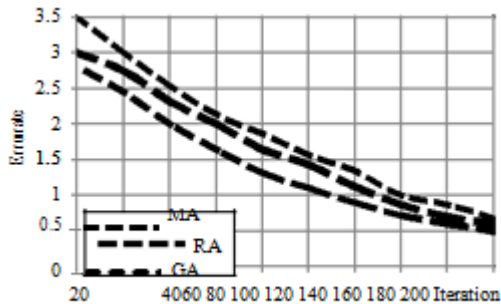
B. Performance of classifier

Fig. 5 Comparison of AdaBoost algorithms error rates

The various Adaboost classifier technique such as RA, GA, MA methods are used to train NGLBP features. The resolution of training image is 30x30. The above-said technique were compared for their performance for error checking with repeated iterations of 100 times and the results are compared in the graph as shown in Fig 5. It is inferred that GA shows lowest error rate in comparison with RA and MA methods. Hence, We preferred GA as detection algorithm for the system.

C. Performance of NGLBP

The receiver operating characteristic curves (ROC) is plotted by YT databases with GF, LBP, GLBP and NGLBP features are tested in GI, BI, N and MS as shown in Fig.6. The graph portray the relationship between a number of false positives and detection rate. The NGLBP features highlight higher performance of 3% and 2% than GLBP features in BI and remaining videos respectively. The LBP features higher performance of 3% than GF features in GI and BI videos. But The LBP features lower performance by 2% than GF features in MS videos hence sensitivity to appearance changes. The GF features lower performance by 1.5% than LBP features in N videos. However NGLBP and GLBP features have better performance of 2% than individual feature under all types of videos. The averages of all types of videos for identifying the location of face region rate significantly enhanced to about 4% by slot in the proposed NGLBP features. The higher accuracy is reported for utilizing NGLBP features over using individual feature set in different video conditions.

Detection of the face includes calculating the Sensitivity, Precision and F measure scores of these values are depicted in Table 2, which indicates the number of video sequences with GI, BI, N and MS. It can be seen that the GF and LBP features perform poorly due to their sensitivity to diverse illumination variations and common appearance respectively, while NGLBP features give much better performance. The size of the bounding box is determined using the scale on the detected face on the video sequences. The LBP and GF feature would fail to detect the face in the different poses and noise image respectively. GLBP feature would fail to detect the face in the BI variation. The Fig.7 depicts that the proposed NGLBP features are robust against noise and illumination and differences pose variations and expressions. This seems to suggest that a combination of normalization, Gabor and LBP features result in better detection of face. This work is carriedout using Intel Core i5 @ 3.20 GHz, 8 GB RAM with Matlab 2013a. Table 3 show that the time complexity for identifying the location of face region.

V. CONCLUSION & FUTURE WORK

This paper explores the NGLBP features for strong discovery of face area in video sequences under uncontrolled scenario. The experimental results illustrate that the proposed technique is successful when comparing with the existing methods. The advantage of NGLBP features is amazingly uncaring in appearance varieties through illumination, expression, and noise in the images. NGLBP features are not just robust to the varieties of image condition, additionally in encoding discriminate information, i.e. face/non-face area in spatial and frequency domains. The NGLBP test results exhibit that feature finds the best execution of revelation of face applications on the YT databases. In future work on the face regions are utilized for facial expression recognition.
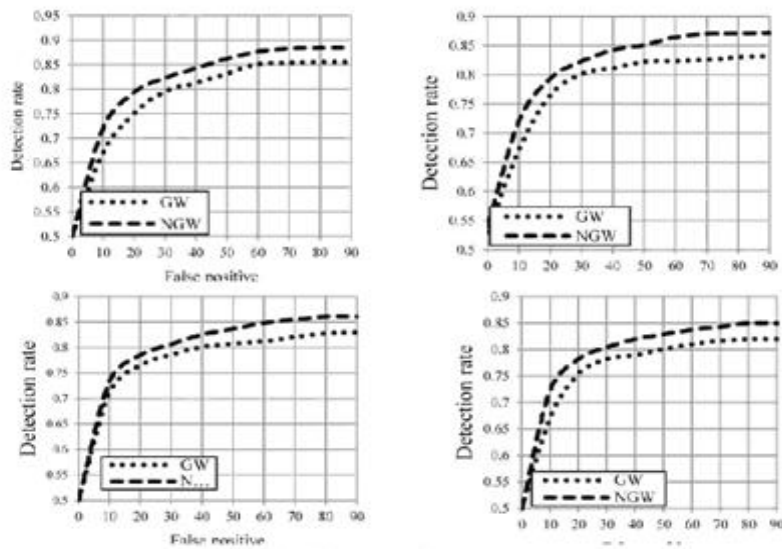
Fig.6. Roc curve false positive rate vs. Detection rate
a) Good Illumination b) Bad Illumination c) Noisy images d) Multiple Subject

Table 2 result for identifying the location of face

author	Feature Extraction	Classifier	Sensitivity in (%)				Precision in (%)				F measure in (%)			
			No. of videos				No. of videos				No. of videos			
			625	420	542	35	625	420	542	35	625	420	542	35
			GI	BI	N	MS	GI	BI	N	MS	GI	BI	N	MS
	GF	Gentleboost	84	80	78	79	85	80	82	81	84	80	79	80
	LBP	Gentleboost	85	83	80	82	87	82	83	84	85	82	82	83
proposed	GLBP	Gentleboost	87	85	83	85	88	84	85	86	87	84	85	86
	NGLBP	Gentleboost	90	88	86	87	90	85	87	88	90	86	87	87

Table 3 Time complexity for comparing our proposed approach with existing approach

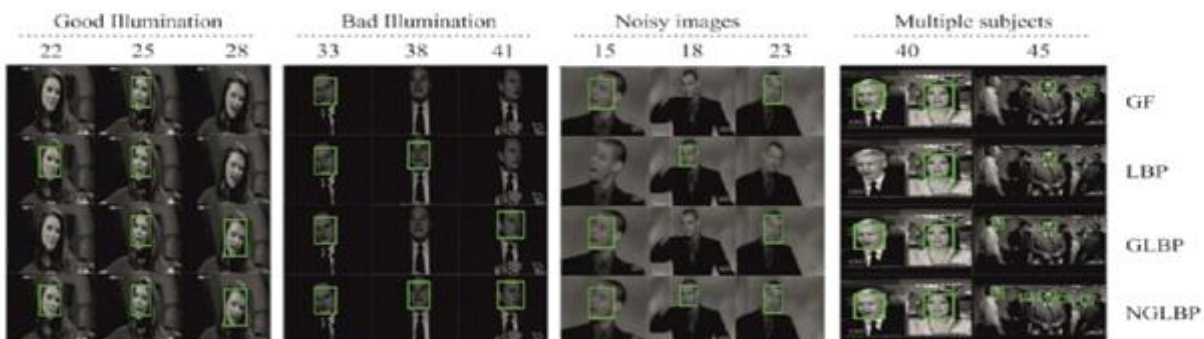| Feature Extraction	Feature extraction for gallery Images (compilation time in ms)	Testing Videos for face detection (compilation time in ms)			
		GI	BI	N	MS
LBP [9]	25	974000	789300	1213450	632211
GF [14]	28	982120	801342	1238691	637625
GLBP	28.5	983450	812420	1256302	638823
NGLBP	29.5	985629	816458	1259000	639000

Fig. 7 Sample detection of face result in YT databases

REFERENCES

[1] M.Yang, D. Kriegman and N. Ahuja, "Detecting faces in images: a survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(1), pp. 34–58.

[2] E.Hielm and B. Low "Face detection: a survey," in Computer Vision and Image Understanding, 2001, 83 (3), pp.236–274.

[3] S.Borah, S.Konwar, T. Taichung and.Rathi, "A human face detection method based on connected component analysis ," in International Conference on Communications and Signal Processing, 2014 ,pp.1205 – 1208.

[4] O.Byung-Hun, and H. Kwang-Seok, "A study on facial components detection method for face-based emotion recognition", in International Conference on Audio, Language and Image Processing (ICALIP), 2014, pp.256 – 259

[5] H. Sagha, S. Kasaei, E. Enayati and M. Dehghani, "Finding Sparse Features in Face Detection Using Genetic Algorithms," in IEEE International Conference on Computational Cybernetics, 2008 , pp.179 - 182

[6] Li. Xiaohua, L. Kin-Man, S.Lansun and Z.Jiliu "Face detection using simplified Gabor features and hierarchical regions in a cascade of classifiers," in Pattern Recognition Letters, 2009, pp. 717–728

[7] L. Huang, A. Shimizu and H. Kobatake, "Robust face detection using Gabor filter features," in Pattern Recognition Letters, 2005, pp. 1641–1649.

[8] T. Yun and L. Guan, "Automatic face detection in video sequences using local normalization and optimal adaptive correlation techniques," Pattern Recognit., 2009, vol. 42, no. 9, pp. 1859–1868.

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

[10] B. Froba and A. Ernst, "Face detection with the modified census transform," in Proc. IEEE International. Conference. automatic Face and Gesture Recognition, 2004, pp. 91–96.

[11] B. Jun and D. Kim, "Robust face detection using local gradient patterns and evidence accumulation," in

Pattern Recognition., 2012, vol. 45, no. 9, pp. 3304–3316.

[12] D. Kim and R. Dahyot, "Face Components Detection using SURF Descriptors and SVMs," Machine Vision and Image Processing International Conference, 2008, pp. 51 –56.

[13] P. Viola and M.l Jones, "Robust real-time object detection," in IJCV, 2004, vol. 57, no. 2, pp. 137–154.

[14] J.Kyungjoong Choi and J.Gil-Jin, "Semi-Local Structure Patterns for Robust Face Detection," in IEEE Signal Processing Letters, 2015, Vol. 22, no. 9, pp 1400-1403.

[15] M.Kim, S. Kumar, V.Pavlovic and H. Rowley, "Face Tracking and Recognition with Visual Constraints in Real-World Videos," in IEEE Conference Computer Vision and Pattern Recognition, 2008.

[16] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, Vol (1), pp.511–518.

Dr. P. Ithaya Rani, Associate Professor, Department of Computer Science Engineering, Sethu Institute of Technology (SIT) College, Affiliated to Anna University Chennai, Tamil Nadu, India. I published over 25 National and International Journals including Springer, Pattern Recognition and Artificial Intelligence, IET Computer Vision, Multimedia tools and Applications, Sadhana Indian Academy of Sciences, International Journal of Computer & Mathematical Sciences and so on. Now I am a reviewer of reputed Springer journal and editor of Human Resource management system journal. I guided three research scholars under Anna University.

T. Hari Prasath, Assistant Professor, Department of Electronics and communication , Kamaraj College of Engineering and Technology, Virdunagar. He published papers in International Conferences and Journals.

AUTHORS PROFILE

A. Jainul Fathima received her B.Tech. degree in Information Technology from Anna University- Chennai in 2007and M.Tech degree in Computer Science and Engineering from Anna University – Tirunelveli in 2009 . She has 3 years of teaching experience. She is currently pursuing Ph.D. degree in Kalasalingam Academy of Research and Education , Krishnankoil. Her Research area includes Big data analytics, Computational Drug discovery, andBioinformatics.She is a Life Member of the Indian Society for Technical Education (ISTE)