

Face Recognition and Tracking in Videos

Swapnil Vitthal Tathe^{*,1}, Abhilasha Sandipan Narote², Sandipan Pralhad Narote³

¹ Department of Electronics & Telecommunication, Sinhgad College of Engineering, Pune, India

² Department of Information Technology, SKN College of Engineering, Pune, India

³ Department of Electronics & Telecommunication, Government Residence Women Polytechnic, Tasgaon, India

ARTICLE INFO

Article history:

Received: 03 April, 2017

Accepted: 01 June, 2017

Online: 23 July, 2017

Keywords :

Face Detection

Face Recognition

Tracking

ABSTRACT

Advancement in computer vision technology and availability of video capturing devices such as surveillance cameras has evoked new video processing applications. The research in video face recognition is mostly biased towards law enforcement applications. Applications involves human recognition based on face and iris, human computer interaction, behavior analysis, video surveillance etc. This paper presents face tracking framework that is capable of face detection using Haar features, recognition using Gabor feature extraction, matching using correlation score and tracking using Kalman filter. The method has good recognition rate for real-life videos and robust performance to changes due to illumination, environmental factors, scale, pose and orientations.

1 Introduction

In recent years, there has been rapid growth in research and development of video sensors and analyzing technologies. Cameras are vital part of security systems used to acquire information from environment. It includes video, audio, thermal, vibration and other sensors for civilian and military applications. The intelligent video acquisition system is a convergence technology including detecting and tracking objects, analyzing their movements and responding to them.

Face recognition system is used to automatically identify a person from an image or a video source. The recognition task is performed by obtaining facial features from an image of the subject's face. The main objective of video-based face recognition and tracking is to identify a video face-track of unknown individuals. It identifies the facial features by extracting features from image of subjects face and analyzing the relative position, size, shape etc. These obtained features are utilized to search corresponding matching features in other images or subsequent frames.

The face recognition system can be divided into three main parts: detection, representation and classification. The face detection from video is challenging task as it needs compensation to distortions arising from motion. The most popular algorithm for last decade is Viola Jones algorithm based on cascaded Haar feature detector [1]. Due to its high computa-

tional speed and robustness it is widely used to detect face in still images as well as video frames. Skin color of human being has limited range of chrominance that can be modeled to represent skin color pixels [2]. To overcome the limitation of complex background with skin color like objects it detects feature like shape and appearance to validate face detection [3]. The representation stage extracts distinct and discriminative features that represents face texture and appearance for face classification. The feature extraction stage is an important as the final matching decision is biased to the quality of features. Recognition stage consists of generating reference template by averaging multiple faces and matching with the detected faces. Eigenface method is simple and popular method widely used in this field [4]. Other methods like LBP [5], LDA [6], Gabor Filter [7], etc. Neural Network is efficient method used for classification [8]. It requires large training samples suitable for face recognition in images. The computation time of recognition system depends on the size of feature vector and number of database images. The state-of-art methods are mostly used for still image face recognition. Research is in progress to develop robust methods for face recognition in real-time.

This paper is an extension of work originally presented in IEEE Annual India Conference [9]. The system proposed here is intended to detect and identify the person. The first stage captures the video segments. Pre-processing algorithms like histogram

*Swapnil Vitthal Tathe, Sinhgad College of Engineering, Pune, India & swapniltathe7@gmail.com

equalization are employed to remove the unwanted artifacts present in the image [10]. The second stage detects faces in complex backgrounds using Viola-Jones [1] face detector. Next stage is representation of face features using Gabor filter [11] that inculcates directional features. The Gabor features are compared with the database images using correlation coefficients [12]. The later stage tracks the detected face in subsequent frames. Tracking utilizes popularly used Kalman filter [13] that makes use of statistical analysis to predict the location of face. The proposed system identify a person in high security premises to restrict the movements of unauthorized person.

The paper is organized as follows: Section 2 presents state-of-art technologies of the field. Section 3 describes methodologies used. Section 4 presents Results and Section 5 gives Conclusions.

2 Related Work

Recognition of individuals is of high importance for various reasons. The real-time algorithms has hard limitations to overcome as any delay in processing may result in loss of important information. Facial features are used for describing individuals as it has unique nature i.e. no two persons can have same faces. The facial features commonly used in many of the studies are skin color [14], spatio-temporal [15], geometry[16] and texture pattern [17]. Face detection algorithms are computationally intensive, which makes it is difficult to perform face detection task in real-time. As human face changes with factors like facial expression, mustache, beard, glasses, size, lightening conditions and orientation, detection algorithms must be adaptive and robust[18].

The simplest approach to detect faces is to find skin pixels present in frames and then apply geometrical analysis to locate exact position of face. The conventional images or frames in videos are modeled with *RGB* color space. Skin pixel values for normalized *RGB* plane are - *r* plane pixel values are in the range 0.36 to 0.456 and *g* values in the range 0.28 to 0.363. The second model in representing images is *HSV* color space. In *HSV* model pixel is classified as skin pixel if it has values $0 \leq H \leq 50$ and $0.20 \leq S \leq 0.68$ [19]. The third model in color representation is $Y C_b C_r$ color space. Skin color pixel has C_r value around 100 and C_b value around 150 [20]. Dhivakar et al. [21] used $Y C_b C_r$ space for skin pixel detection and Viola-Jones [1] method to verify correct detection. Skin color based method is sensitive to illumination variation and fails if background contains skin color like objects.

The other complex methods used for face detection are based on geometry and texture pattern that uses depth features to detect face. Face detection using edge maps and skin color segmentation is presented in [22]. Mehta et al. [6] proposed LPA and LBP methods to extract textural features with LDA dimensionality reduction and SVM classification. Lai

et al. [23] used logarithmic difference edge maps to overcome illumination variation with face verification ability of 95% true positive rate. Contour based face detection is proposed in [24]. Xiao et al. [25] used eigen faces for face detection using texture and shape information in bi view face recognition system. Viola and Jones [1] proposed haar feature based real-time method for face detection. Many researchers popularly used haar face detection with adaboost classification [8] in face detection.

Face recognition systems are popular in biometric authentication as they do not require the user intervention. Face recognition systems can be classified as sub-class of pattern recognition systems. Eigenfaces and Fisher faces were proposed in the early 1990s by Turk et al. [26]. Bayesian face recognition has better performance compared to eigenfaces [27]. Rawlinson et al. [28] presents eigen and fisher face for face recognition. Feature based methods are robust to illumination and pose variations. SIFT and SURF has robust performance in unconstrained scenarios [29]. Gangwar et al. [30] introduced Local Gabor Rank Pattern method that uses both magnitude and phase response of filters to recognize face. Neural Network is most popular tool of recent days that is used for classification which has outperformed many methods [31]. Meshgini et al. [7] proposed a method with combination of Gabor wavelets for feature extraction and support vector machine for classification. Recognition methods are time consuming as it requires accumulates training and testing phases.

Tracking is of prime importance to monitor activities of individuals in various environments [32]. The location of object in next frame is predicted based on location of object in current or preceding frames. More details for tracking methods can be found in [33]. Elafi et al [34] proposed an approach that can detect and track multiple moving objects automatically without any learning phase or prior knowledge about the size, nature or initial position. Gyorgy et al. [35] proposed Extended Kalman filter for non linear systems. Kumar et al. [36] extracted moving objects using background subtraction and applied Kalman filter for successive tracking. The other popular tracker is mean shift that employs mean shift analysis to identify a target candidate region, which has most similar appearance to target model in terms of intensity distribution. Adaptive mean shift (Camshift) algorithm is popular OpenCV algorithm for tracking [37]. In [38] combined mean shift and kalman filter is used for tracking. Particle filter tracker uses probabilistic approach to estimate position of object in next frame. The particle filter searches for histogram that best matches with reference [39]. Tracking using sparse representation [40] is gaining momentum for fast and accurate tracking.

3 Methodology

Automatic face detection is the very first step in many applications like face recognition, face retrieval, face tracking, surveillance, HCI, etc. Face is vital part of human being representing most information about the individual. Developing a representation model is difficult because face is complex, multi-dimensional visual stimuli. General architecture of proposed system is shown in Figure 1.

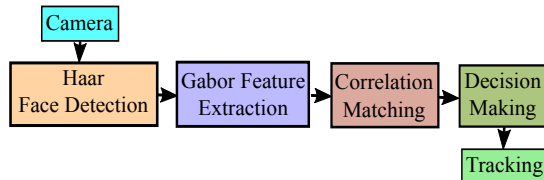


Figure 1: Human detection and tracking system

3.1 Haar Face Detection

This method is most popular method for face detection in real-time applications. Most of mobile platform uses the technique to locate faces in camera applications. Viola and Jones [1] proposed real-time Adaboost algorithm for classification of rectangular features. The detection technique is based on the idea of the wavelet template that defines the shape of an object in terms of a subset of wavelet coefficients. The different Haar features shown in Figure 2 are used to analyze the intensity variation patterns of face regions. For example the eye region appears darker than the surrounding region that can be modeled using two rectangle feature (black part corresponding to eye region and white part corresponding to surrounding region). Combination of four, six, eight and more features is used in practical face detection to make detection robust. For a given random variable X , the variance value of X is as follows:

$$Var(X) = E(X^2) - \mu^2 \quad (1)$$

where $E(X^2)$ is expected value of squared of X and μ is expected value of X .

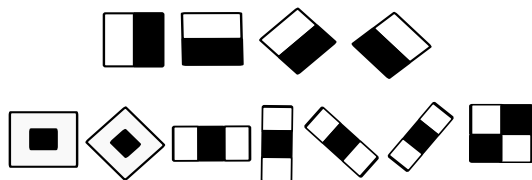


Figure 2: Haar features

Feature value is computed by subtracting the sum of variance of black region from sum of variance of white regions. To increase the computation speed, feature values can be computed using integral image [41]. Each feature is classified by a Cascaded Haar feature classifier. The Haar feature classifier multiplies the weight of each rectangle by its area and results are added together. The cascaded structure helps to speed

up the detection process by eliminating non-face regions at every stage of cascade. Figure 3 shows face detection results from a video stream capture using webcam.



Figure 3: Face detection using Haar Features

Haar based method is capable of detecting multiple faces in near real-time and robust to illumination variations.

3.2 Gabor Filter

Feature extraction is one of the most important steps in designing a pattern recognition system. The features must have features have a small within class variation and strong discriminating ability among classes. Gabor filters have band pass performance that emphasizes features in face region that are analogous to scale and orientation of filter. The characteristics of Gabor filters are appropriate for texture representation and discrimination [7]. This method is robust to face pose and illumination changes.

The Gabor filter extracts face features from gray-level images. A two-dimensional Gabor filter is a Gaussian kernel function modulated by a complex sinusoidal plane wave represented as Equation 2.

$$\Psi_{\omega,\theta}(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \exp(j\omega x') \quad (2)$$

$$x' = x\cos\theta + y\sin\theta, y = -x\sin\theta + y\cos\theta \quad (3)$$

where (x,y) is the pixel position in the spatial domain, x is the central angular frequency of the complex sinusoidal plane wave, h is the anti-clockwise rotation(orientation) of the Gabor function, r represents the sharpness of the Gaussian function along both x and y directions and $\sigma \approx \frac{\pi}{\omega}$.

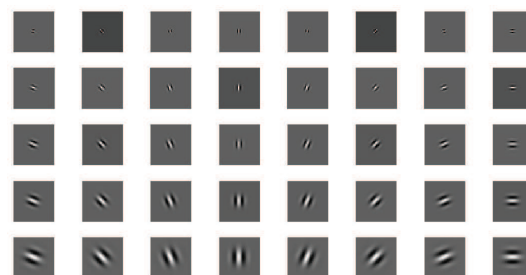


Figure 4: Gabor filters at 5 scales and 8 orientations

To extract useful features from an face image a set of Gabor filters with five frequencies (scales) and eight

orientations is used. Figure 4 shows the Gabor filter bank. The scale and orientation values are obtained using Equations 4 and 5.

$$\omega_u = \frac{\pi}{2} \times \sqrt{2^u}, u = 0, 1, \dots, 4 \quad (4)$$

$$\theta_v = \frac{\pi}{8} \times v, v = 0, 1, \dots, 7 \quad (5)$$

The Gabor representation of a face image $I(x, y)$ is obtained as:

$$G_{u,v}(x, y) = I(x, y) * \Psi_{\omega_u, \theta_v}(x, y) \quad (6)$$

where $G_{u,v}(x, y)$ denotes the two-dimensional convolution result corresponding to the Gabor filter at scale u and orientation v . This process can speed-up with the use of frequency domain transformation. Figure 6 and 7 shows the convolution results of a face image shown in Figure 5) with Gabor filters.



Figure 5: Face Image

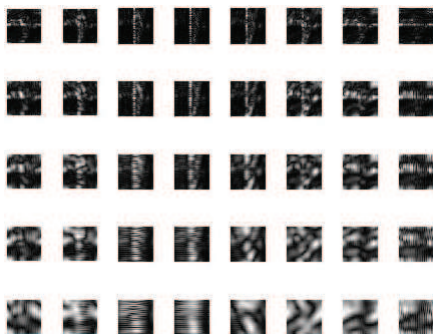


Figure 6: Magnitude of Face

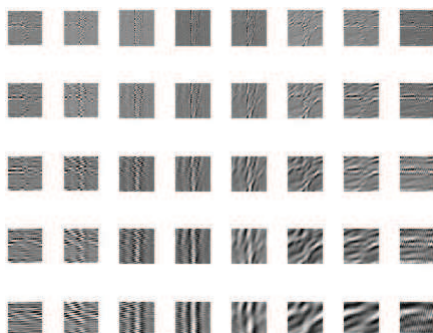


Figure 7: Real Part of Face

Each image $I(x, y)$ is represented as set of Gabor wavelet coefficients $G_{u,v}(x, y) | u = 0, 1, \dots, 4; v = 0, 1, \dots, 7$. The magnitude $G_{u,v}(x, y)$ is normalized to zero mean and unit variance and represented as a vector $z_{u,v}$ given by Equation 7 [7].

$$z = [(z_{0,0})^T (z_{0,1})^T \dots (z_{4,7})^T]^T \quad (7)$$

Gabor based recognition is robust to pose and illumination variations with additional overhead of processing time and complexity.

3.3 Correlation Score Matching

Matching is a search algorithm that compares the features from query image to the database images to find best match. The comparison is done using correlation coefficients [32] between the Gabor features of detected face and the images present in database. Prior to matching stage all images in database are applied with Gabor filter for efficient feature matching. The 2D images are represented by 1D feature vector that speeds up the matching process. The correlation score is computed as :

$$Score = \sum_{i=1}^m f_q(i) \cdot f_{db}(-i) \quad (8)$$

where m is size of feature vector, f_q represents query feature vector and f_{db} represents database feature vector. The query image best matches with the database image that results into high correlation score.

3.4 Kalman Filter Tracking

Tracking objects in video sequences is one of the most important tasks in computer vision. The approach is to search target within a region centered on last position of the object. The important facts that should be taken care for effective tracking are: speed of moving object, frame rate of video and search region. These factors are dependent on each other, higher frame rate is used to get hold of faster moving target which on other hand requires faster processing. Tracking is defined as localization and association of features across series of frames. The two major components of tracking system are object representation and data association. The first component describes object and second utilizes the information for tracking.

The location of face from face detection stage is used as apriori estimate or sometimes called as initialization for tracking. Based on this location Kalman filter estimates and updates the prediction behavior of this stage. Initialization is an important stage for the tracker.

Kalman filter predicts the next time state vector based on the movement information in the previous stages. Kalman filter equations are categorized as - time update equations and measurement update equations. The time update equations projects forward the current state and error covariance estimates to obtain the apriori estimates for the next time step. The measurement update equations are the feedback i.e. it incorporates a new measurement into the a priori estimate to obtain an improved aposteriori estimate. The time update equations are predictor equations, while the measurement update equations are

corrector equations. The process for Kalman filter estimation and update is given in Algorithm 1.

Algorithm 1 Kalman Filter

- 1: Initial Estimates for \hat{X}_{k-1} and P_{k-1}
- 2: Time Update

- Project the state

$$\hat{X}_k^- = A\hat{X}_{k-1} + Bw_k$$

- Project the error covariance

$$P_k^- = AP_{k-1}A^T + Q$$

- 3: Measurement Update

- Compute Kalman gain

$$K_k = \frac{P_k^- H^T}{HP_k^- H^T + R}$$

- Update estimate with measurement Z_k

$$\hat{X}_k = \hat{X}_k^- + K_k(z_k - H\hat{X}_k^-)$$

- Update error covariance

$$P_k = (1 - K_k H)P_k^-$$

The state equation is

$$X_k = AX_{k-1} + Bw_k \tag{9}$$

The measurement model is

$$Z_k = HX_k + v_k \tag{10}$$

A and B are the system parameters and are matrix in the multi-model system, H is a parameter in a measure system, which is matrix in the multi-measurement system. w_k and v_k represents the process and measurement noise respectively.

The state vector $X_k = [x_k y_k x'_k y'_k]^T$, x_k, y_k are the object center in the x, y -axis location, x'_k, y'_k are the x, y axis speed. $Z_k = [x_{ck} y_{ck}]^T$ is the observation vector. P_k^- is apriori estimate error covariance and P_k is aposteriori estimate error covariance.

The search for object is carried in a region selected based on the estimated observation matrix Z_k . It fails if image is not properly illuminated. Kalman filter overcomes problem of Kanade-Lucas-Tomasi(KLT) [42]. It can track multiple objects with variations in pose and occlusion. The drawback is that it cannot handle rapid changes in motion of target.

Table 1: Result Comparison

Video	Frames	Face Det.	True Recog.		False Recog.	
			Eigen	Gabor	Eigen	Gabor
1	237	71	47	37	24	34
2	329	99	97	31	2	68
3	257	223	52	98	171	125
4	339	225	65	138	160	87
5	448	305	225	7	80	298
6	438	283	127	4	156	279
7	353	176	140	96	36	187
8	404	274	242	195	32	88
9	198	191	101	2	90	189
10	248	206	46	19	160	187
11	78	78	21	65	57	13
12	128	124	36	74	88	50
13	324	252	122	224	130	28
14	353	250	112	216	138	36
15	258	176	107	82	69	94
16	328	191	125	95	66	96
17	346	238	65	108	173	130
18	426	392	148	289	244	103
19	318	302	144	48	158	344
20	388	257	146	16	111	376

4 Results

The algorithms are applied on video sequences obtained from NRC-IIT Facial Video Database and our own videos recorded with webcam. Table 1 gives the comparison of proposed method with eigen face method. The values presented in table are the number of frames with true result. Table 2 gives summary of results obtained on the videos. Haar face detector used has robust performance for frontal faces. It cannot handle larger pose variations. Face detection fails if the face is very small in size compared to size of frame and in random pose changes. The second part of algorithm is face recognition that makes use of gabor filters. Gabor filter has robust performance in pose and illumination variations. Once person is recognized Kalman filter keeps track of movements the face in subsequent frames. Kalman filter has good tracking performance in partial occlusions but fails in case of sudden movements. Recognition result is shown in Figure 8. The values in table shows the playback time in sections for videos under test with face detection and recognition stage.

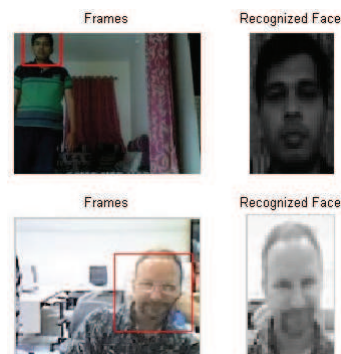


Figure 8: Face Recognition Results

Table 2: Results for Tracking

Sample Video	No. of Frames	Face Detect. (Time (sec))	Face Recog. (Time (sec))
1	237	15.5880	54.5289
2	329	20.9837	52.9032
3	257	16.4133	41.3256
4	339	21.4623	86.5154
5	448	31.4490	63.3743
6	438	30.2323	65.4514
7	353	25.3845	56.76
8	404	28.7957	113.2256
9	198	13.7490	31.8384
10	248	16.5841	39.8784
11	78	5.4555	5.97056
12	128	8.6597	20.5824
13	324	22.5488	59.4159
14	353	24.4202	66.3573
15	258	17.0961	41.4864
16	328	20.7186	52.7424
17	346	22.1367	55.6368
18	426	28.5429	68.5008
19	318	20.9056	51.1344
20	388	25.4525	35.1624

5 Conclusions

The proposed system detects the facial features of human being and verify identity of a person. Haar based face detection is popular and efficient face detection algorithm for real-time applications. Haar based face detector has high accuracy in frontal face detection. Recognition performance depends on the type of images in database. Larger the database more is recognition time. Gabor face recognition is better than many other methods as it compares features at different scale and orientations. Disadvantage of recognition stage is that comparison of large number of features requires more time for recognition that affects real-time performance. Reducing feature set reduces recognition time but affects accuracy of system. Kalman filter gives good performance in complex environment. Future directions can be optimizing algorithms to minimize processing time and achieve real-time performance.

References

[1] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, pp. 137–154, May 2004.

[2] C. Dong, T. Lin, X. Wang, and L. Mei-cheng, "Face detection under particular environment based on skin color model and radial basis function network," in *IEEE Fifth International Conference on Big Data and Cloud Computing*, pp. 256–259, Aug 2015.

[3] G. Lee, K. H. B. Ghazali, J. Ma, R. Xiao, and S. A. Iubis, "An innovative face detection based on ycgcr color space," *Physics Procedia*, vol. 25, pp. 2116 – 2124, 2012.

[4] F. Crenna, G. Rossi, and L. Bovio, "Perceived similarity in face measurement," *Journal of Measurement*, vol. 50, pp. 397–406, 2014.

[5] G. Lian, J.-H. Lai, C. Y. Suen, and P. Chen, "Matching of tracked pedestrians across disjoint camera views using ci-

dlbp," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, pp. 1087–1099, July 2012.

[6] R. Mehta, J. Yuan, and K. Egiazarian, "Face recognition using scale-adaptive directional and textural features," *Pattern Recognition*, vol. 47, pp. 1846–1858, May 2014.

[7] S. Meshgini, A. Aghagolzadeh, and H. Seyedarabi, "Face recognition using gabor-based direct linear discriminant analysis and support vector machine," *Computers and Electrical Engineering*, vol. 39, no. 910, p. 727745, 2013.

[8] M. Da'san, A. Alqudah, and O. Debeir, "Face detection using viola and jones method and neural networks," in *International Conference on Information and Communication Technology Research (ICTRC)*, pp. 40–43, May 2015.

[9] S. V. Tathe, A. S. Narote, and S. P. Narote, "Face detection and recognition in videos," in *IEEE Annual India Conference (INDICON)*, pp. 1–6, Dec 2016.

[10] C. Y. Lin, J. T. Fu, S. H. Wang, and C. L. Huang, "New face detection method based on multi-scale histograms," in *2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, pp. 229–232, April 2016.

[11] F. Jiang, M. Fischer, H. K. Ekenel, and B. E. Shi, "Combining texture and stereo disparity cues for real-time face detection," *Signal Processing: Image Communication*, vol. 28, no. 9, pp. 1100 – 1113, 2013.

[12] M. B. Hisham, S. N. Yaakob, R. A. A. Raof, A. B. A. Nazren, and N. M. W. Embedded, "Template matching using sum of squared difference and normalized cross correlation," in *IEEE Student Conference on Research and Development (SCoReD)*, pp. 100–104, Dec 2015.

[13] S. Vasuhi, M. Vijayakumar, and V. Vaidehi, "Real time multiple human tracking using kalman filter," in *3rd International Conference on Signal Processing, Communication and Networking*, pp. 1–6, March 2015.

[14] B. Dahal, A. Alsadoon, P. W. C. Prasad, and A. Elchouemi, "Incorporating skin color for improved face detection and tracking system," in *IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pp. 173–176, March 2016.

[15] H. Yang, L. Shao, F. Zheng, L. Wang, and Z. Song, "Recent advances and trends in visual tracking: A review," *Neuro Computing*, vol. 74, pp. 1–9, November 2011.

[16] J. Yan, X. Zhang, Z. Lei, and S. Z. Li, "Face detection by structural models," *Image and Vision Computing*, vol. 32, no. 10, pp. 790 – 799, 2014. Best of Automatic Face and Gesture Recognition 2013.

[17] D. T. Nguyen, W. Li, and P. O. Ogunbona, "Human detection from images and videos: A survey," *Pattern Recognition*, vol. 51, pp. 148–175, March 2016.

[18] S. Liao, A. K. Jain, and S. Z. Li, "A fast and accurate unconstrained face detector," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, pp. 211–223, Feb 2016.

[19] I. Hemdan, S. Karungaru, and K. Terada, "Facial features-based method for human tracking," in *17th Korea-Japan Joint Workshop on Frontiers of Computer Vision*, pp. 1–4, 2011.

[20] M. Zhao, P. Li, and L. Wang, "A novel complete face detection method in color images," in *3rd International Congress on Image and Signal Processing*, pp. 1763–1766, 2010.

[21] B. Dhivakar, C. Sridevi, S. Selvakumar, and P. Guhan, "Face detection and recognition using skin color," in *3rd International Conference on Signal Processing, Communication and Networking*, pp. 1–7, March 2015.

[22] R. Sarkar, S. Bakshi, and P. K. Sa, "A real-time model for multiple human face tracking from low-resolution surveillance videos," *Procedia Technology*, vol. 6, pp. 1004 – 1010, 2012.

[23] Z.-R. Lai, D.-Q. Dai, C.-X. Ren, and K.-K. Huang, "Multiscale logarithm difference edgemaps for face recognition against varying lighting conditions," *IEEE Transactions on Image Processing*, vol. 24, pp. 1735–1747, June 2015.

[24] A. Dey, "A contour based procedure for face detection and tracking from video," in *3rd International Conference on Recent Advances in Information Technology (RAIT)*, pp. 483–488, March 2016.

- [25] B. Xiao, X. Gao, D. Tao, and X. Li, "Biview face recognition in the shapetexture domain," *Pattern Recognition*, vol. 46, pp. 1906 – 1919, July 2013.
- [26] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 586–591, 1991.
- [27] B. Moghaddam, T. Jebara, and A. Pentland, "Bayesian face recognition," *Pattern Recognition*, vol. 33, no. 11, pp. 1771 – 1782, 2000.
- [28] T. Rawlinson, A. Bhalerao, and L. Wang, *Handbook of Research on Computational Forensics, Digital Crime and Investigation: Methods and Solutions*, pp. 53–78. Ed. C-T. Li, 2010.
- [29] A. Vinay, D. Hebbar, V. S. Shekhar, K. N. B. Murthy, and S. Natarajan, "Two novel detector-descriptor based approaches for face recognition using {SIFT} and {SURF}," *Procedia Computer Science*, vol. 70, pp. 185 – 197, 2015. Proceedings of the 4th International Conference on Eco-friendly Computing and Communication Systems.
- [30] A. Gangwar and A. Joshi, "Local gabor rank pattern (lgrp): A novel descriptor for face representation and recognition," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, pp. 1–6, Nov 2015.
- [31] N. Jindal and V. Kumar, "Enhanced face recognition algorithm using pca with artificial neural networks," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, pp. 864–872, June 2013.
- [32] G. Adhikari, S. K. Sahani, M. S. Chauhan, and B. K. Das, "Fast real time object tracking based on normalized cross correlation and importance of thresholding segmentation," in *International Conference on Recent Trends in Information Technology (ICRTIT)*, pp. 1–5, April 2016.
- [33] A. W. M. Smeulders, D. M. Chu, R. Cucchiara, S. Calderara, A. Dehghan, and M. Shah, "Visual tracking: An experimental survey," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 36, pp. 1442–1468, July 2014.
- [34] I. Elafi, M. Jedra, and N. Zahid, "Unsupervised detection and tracking of moving objects for video surveillance applications," *Pattern Recognition Letters*, vol. 84, pp. 70 – 77, 2016.
- [35] K. Gyrgy, A. Kelemen, and L. Dvid, "Unscented kalman filters and particle filter methods for nonlinear state estimation," in *The 7th International Conference Interdisciplinarity in Engineering*, vol. 12, pp. 65 – 74, 2014.
- [36] S. Kumar and J. S. Yadav, "Video object extraction and its tracking using background subtraction in complex environments," *Perspectives in Science*, vol. 8, pp. 317 – 322, 2016. Recent Trends in Engineering and Material Sciences.
- [37] M. Coskun and S. Unal, "Implementation of tracking of a moving object based on camshift approach with a uav," *Procedia Technology*, vol. 22, pp. 556 – 561, 2016.
- [38] S. V. Tathe and S. P. Narote, "Mean shift and kalman filter based human face tracking," in *Communications in Computer and Information Science, SPC*, pp. 317–324, June 2013.
- [39] M. Z. Islam, C.-M. Oh, and C.-W. Lee, "Video based moving object tracking by particle filter," *International Journal of Signal Processing, Image Processing and Pattern*, vol. 2, pp. 119 – 132, March 2009.
- [40] D. Wang, H. Lu, and C. Bo, "Online tracking via two view sparse representation," *IEEE Signal Processing Letters*, vol. 21, pp. 1031–1034, September 2014.
- [41] J. Jin, B. Xu, X. Liu, Y. Wang, L. Cao, L. Han, B. Zhou, and M. Li, "A face detection and location method based on feature binding," *Signal Processing: Image Communication*, vol. 36, pp. 179 – 189, 2015.
- [42] V. M. Arceda, K. F. Fabian, P. L. Laura, J. R. Tito, and J. G. Caceres, "Fast face detection in violent video scenes," *Electronic Notes in Theoretical Computer Science*, vol. 329, pp. 5 – 26, 2016.