IFMP – Intrinsic Face Motion Pipeline

M. Saqlain Nawaz[1], Shahab Anjum[2], Irshad Ahmed Sumra[3]

[1,2,3]Department of Computer Science, Bahria University, Lahore, Pakistan
Email [1]msaqlain.nawaz@gmail.com, [2]shahabanjum3911@gmail.com, [3]isomro28@gmail.com

Abstract: Now a days, there are a lot of security threats for which every security departments work dawn to dusk. In case, there is a person whose entrance is banned in an organization, or information about that person is required either he was there or not. For this outcome, it will be necessary to watch the CCTV footage even it is 10 hours long to sport that person. To tackle this happening, what if there is a tool or software which could do all this work in the same time. IFMP is a desktop-based video player application through which any humane face can be detected through computing algorithms of recognition and detection under consideration of framework provided by Microsoft Corporation. Its a tool that only takes the picture of the person as input and after processing all the footage, it gives the time hits as output from that footage in which the required person is being spotted.

Keywords: *Software Utility; OpenCV; Video Processor; CCTV footage; Microsoft Framework*

1. Introduction

The 21st century is known as century of IT. IT is base evolution of business environment. All worldwide companies are setting their focus on IT to improve quality of services and lower their costs. This new era of IT is uninterruptedly assisting companies which are developing themselves in IT with their hard work. The IT has changed the life of style of a modern man. Computer is dominant in all the industry and other fields of life with which humane interacts. It is very much useful in process of manufacture of machine parts [1] used to increase the efficiency of company. To improve quality, abridge distribution time, instant scheming and control of automated systems helps a lot. Computers are offering massive benefits from evolving machineries in communication areas as well as in hardware and software.The introduction of computer provided the world a new way of thoughts. The powerful, innovatory and revolutionary technologies have been changing the world and nation. Today computer is becoming necessary and introducing innovation in every fields of life. It doesn't make any difference that what this era of computer is called. What makes a difference is it is a positive revolutionary era of information and automation. The invention of computer played a native role in medical treatments and helped patients in countless ways. Its obvious that nothing is more important than life. Due to the complexity of computer based systems, information technology is necessarily used and it plays a amazing role [2]. With the use of IT, most of the routine tasks become much easier to handle. With the propagation of IT and computer based systems, many big and small companies automated their systems to enhance the productivity and efficiency of their industry. Still there are also some companies which stick to their own procedures without IT as per their own need. Software is a set of rules and procedures with association of relevant data. There are also fields of studies for management of software. Software engineering is the most common and elaborated field which contain the systematic approach and development of software. The application of these software's in science and mathematics because of which computer based system are made useful. The basic goal of software engineering is the production of efficient software at quite low cost (Hollingsworth, 2008). There are a lot of examples that denotes the importance of automated software's like an inventory control system can be used to facilitate sale-purchase process. In this system, the seller maintains the stock and keep tracks of in/out of items. The purchaser places order and seller keeps tracks of all the orders of clients with list of shipping information [3]. Now a days with the rapid progress of IT and Innovations, there are still some probable causes like lack of funds and illiterate staff. But well established companies are recommended to adopt automated systems to improve the efficiency and productivity of the business and uplift the reputation of their industry. Face detection and recognition is also very popular in field of automated systems. To overcome the security threats and monitor the in/out of organization, CCTV cameras are very useful.

The Section 2 describes the prior work done on Face detection and recognition. Section 3 elaborates the proposed methodology including flow charts and requirements. After that, Section 4, shows the results and discussions. And after analysis parameters, the conclusion lies in Section 5 with future enhancements.

2. Related Work

Face detection is a software technology of computer based system that identify the basic features of human faces in digital images. It detects facial features like eyes, nose and lips etc. In face detection all other details in a digital image which are not related to face i.e buildings, trees and bodies are ignored. Face detection can be used as a specific case of object-class detection. In object-class detection, the size and location of objects are to be found in the image under consideration like vehicles and towers.

Qian Li et al [1] seemed to be the first article where Viola-Jones is presented on the basics of face detection algorithm. This algorithm only finds frontal upright faces, but in 2003 the very algorithm was presented in a version that also

detects profile views. Major advantages of this algorithm are:

It is efficient. The results provided by this algorithm are accurate. It has fast execution. Processing time is less. Its performance is remarkable. Face detection can also be used in face localization [2]. In this case, size and location of known number of faces must be found.

Early literature says that face-detection algorithms were only focused on the detection of front faces of humans as it only detects eyes, nose and lips etc. But with the propagation of new research, new algorithms provided more details about profile face images with enhanced efficiency and solved the problem of multi-view face detection. It also provided the information about the feature detection of rotated images along the axis (in-plane and out-of-plane rotation)[3] as well as variations in the image or video by factors like face pose , appearance and lighting etc.

For the detection of faces at multiple levels, an image pyramid [3] is calculated. The Examining of a fixed size sub-window through each layer of pyramid is part of the basic mechanism of this algorithm. The values of sub-window are related to wavelet analysis. Histograms are made for the different wavelet coefficients. The fixed size sub-window is generated and moved through each image in the pyramid. The values of sub-window are updated correctly when lightning is not uniform and subjected to histogram equalization. After that, several parallel neural networks are used that carry out the actual face detection by using the updated information provided to them. Then logical AND is used to combine the outputs and amount of false detections is reduced. In beta version of this algorithm, it also detects frontal upright faces only. A detector that shows the highest efficiency of output, is selected and used to determine the orientation of the object. as compared to the basic Viola-Jones algorithm and the algorithm presented by Rowley et al. this algorithm also detects profile view detections.

3. Methodology

IFMP is a desktop-based video player application through which any person can be identified through computing algorithms under consideration of framework provided by Microsoft Corporation.

3.1 Viola-Jones Face Detector

This algorithm examines a sub-window which has the ability to detect faces in a given image. The input image is rescaled to different sizes by using a standard image processing technique and then the detector of fixed size is run through these images. This technique has to calculate images of different size to different sizes [1]. That's why it turns out to be time consuming. Viola-Jones implies the rescaling on the detector not on input image which is a contradiction with standard approach. After that, detector is executed many times on the image with a different size each time. At first both approaches seems to be equally time consuming, but Viola-Jones have overcome this issue by devising a detector which is scale invariant that needs same number of calculations no matter what the size is. An integral image and features of Haar wavelets are used for the construction of this detector.

- **Scale Invariant Detector**

In Viola-Jones, the first step is the generation of an integral image (Figure 1) from input image. Each pixel of input image is taken equal to the total sum of all pixels left and above of the pixel under consideration.

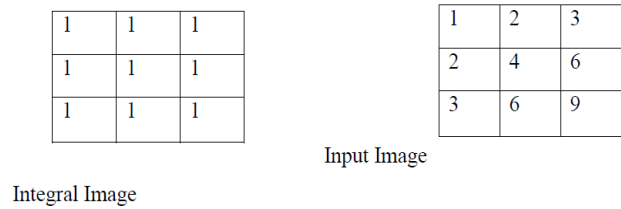Integral Image   Input Image

Figure. 1 Integral Image

By doing this, calculation of the sum of all pixels can be done by using only four values in the rectangle. These are pixels values of the integral image at corners in the input image. (Figure 2)
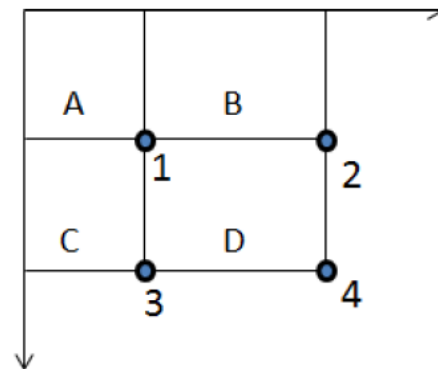

Figure. 2 Integral Image

Sum of all pixels in D = 1 + 4 - (2 + 3)

= A + (A + B + C + D) - (A + C + A + B)

= D

The sum of A must be added to calculation because both rectangle B and C include rectangle A. It has now been clearly shown that in constant time how the sum of rectangles pixels can be calculated which are of arbitrary in size. The Viola-Jones detector analyzes a sub-window under consideration using features (Figure 3) consisting of two or more rectangles.
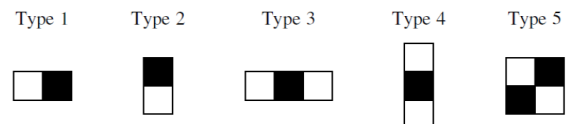

Figure. 3 Different Type of Features

Himanshu Sharma et al described in [2] that for all possible sizes and positions of the features, Viola-Jones have denote that a detector with a base resolution of 25*25 pixels provides efficient results. Total of approximately 160,000 different features can then be constructed. Thus, the number of possible features vastly outnumbers the 576 pixels contained in the detector at base resolution. These features may seem to be very simple but these perform very advanced task in face detection.

- **The Modified AdaBoost Algorithm**

In machine learning, AdaBoost is a boosting algorithm which has the ability to construct a strong classifier through a weighted combination of weak classifiers. Each feature is taken as weak classifier in order to make similarity of this terminology with the theory presented in this this section. Following is the mathematical denotion of a weak classifier:

$$h(x, f, p, \theta) = \begin{cases} 1 & if p f(x) > p\theta \\ 0 & otherwise \end{cases}$$

Where 'x' is a sub-window of pixel having order of 25*25, the polarity is 'p', the applied feature is 'f' and θ flags the threshold which makes decision about either 'x' should be classified as a negative (a non-face) or positive (a face).

The detection algorithm named as Viola-Jones examines the detector repeatedly with new window size through the same input image. In case of more than one detectable faces, the sub windows will show negative values and will be evaluated in very large amount. This understanding about the mechanism guides to formulation of a new problem [2]. Despite of figuring out faces, the discussed algorithm should reject the objects which are not related to faces. Main idea of this statement is that to reject the ich are not related to faces is quick for the sake of finding the face. Keeping this concept in mind, a detector having only one classifier (i.e a strong one), looses its efficiency because of the constant time of evaluation, independent from the input. Thus, a cascaded classifier is needed.

Matthew Day et al described in [3] that cascaded classifier (Figure 4) consists on different levels and each one contains a strong classifier. Each level checks that either the under consideration sub window is a face or some other object which is irrelevant to face. At some level, if it is observed that the sub window is not a face, it is rejected at once. In the same way, if sub window is observed as a face at some level, it is forwarded to the next stage in the stream. To improve the efficiency of sub window being as a face, it is necessary to pass as many levels.
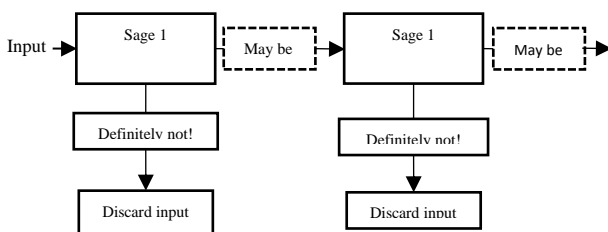

Figure. 4 Rejection Cascade

In a single stage classifier, the chances of having false positive results as output are high. To lessen this false positive rate, there should be classifier which could accept such results. However, on initial levels of staged classifier, such false positives have no impact on the final results because the next upcoming levels deal with them. Therefore Viola-Jones accepts a lot of false positives in its primary levels. After the face detection the second step starts i.e. face recognition. The detected face from the Viola Jones algorithm works as input to the face recognition.

Three algorithms were implemented for the face recognition and compared their outputs, efficiency and accuracy.

Face recognition is a technology that uses the output data of detectors and identifies the detected objects where does it belong. It has a lot of benefits like only detection of objects and face is not enough until their true name is unknown which could identify that detected object. If a vehicle is detected then there should also be some identification either it is a car or a truck. For this identification, recognition technology is used. In face recognition provides different kind of information like emotions and person's identification etc. There are a lot of algorithms used for face recognition and some of the best are discussed in this paper.

3.2 Principal Component Analysis (PCA)

Firstly, PCA [10] is used for face recognition algorithm. Its implementation provides small set of significant features as known as eigen vectors are used to specify face image variations [4].

The face of a human has very dynamic structure and contains specific characteristics that differ from person to person and also changes with the passage of time. It reflects the emotions, variations and plays an important role in identification. So, face recognition is very important and interesting task in field of computer vision. Face recognition is very helpful in modern era like Credit card verification, CNIC verification, identification of criminals and security systems etc. For having such amazing results, algorithms depends on different kind of features which are countless in number. All these features make algorithm complex enough to understand and execute. As the quality of results depend upon the quality of extracted features. One of the best algorithm is PCA. This algorithm follows a technique that figures out a set of projection vectors which are designed to retain as much information from the input data as possible. For this purpose, eigenvectors are used which contains the eigen values of highest level calculated from input matrix. By using this method, the dimensionality is reduced.

This algorithm is popular because it is not sensitive to time dependent changes in facial features, simplicity and speed. The features also depends upon the dimension of face I.e either it is profile, vertical or frontal etc.

The basic job of PCA is to identify the hidden patterns and information. Then by using that core information it highlight the similarities and differences. As in high dimensions, it is a big task to conduct analysis of these patterns. So, PCA becomes handy and becomes useful in analysis for the reduction of dimensionality of data. It generates an average face and represents all other faces as the sum and difference of this average face and all other Eigen faces.

3.3 Linear Discriminant Analysis (LDA)

F. Ye et al describes the Basic Principle of eigen faces method is PCA in [9]. It figure out features that provides the maximum values of total variance in data. Clearly the PCA is very good technique but there are also some issues which can't be overlooked. It does not take classes into account because of which much information may be lost. As this information relates to the discriminant analysis so, it is a hefty loss. Because of this, the collected samples become messy with each other which impose a bad effect on the classification and sometimes it becomes impossible.

Additionally, the accuracy of PCA was very less and the ratio of false recognition was very high. Hence Linear Discriminant Analysis method was implemented.

The major advantages of Linear Discriminant Analysis (LDA) are [4]

- It produces accurate results comparatively better results than PCA
- Output of code not affected by minor light variations.
- Output of algorithm is not affected by minor changes in facial expressions, hence producing correct authentication.
- It consumes less time.

In field of computer Vision and Statistics, LDA and Fisher's linear discriminant related to it are methods used recognition of pattern as well as to figure out suitable features in machine learning. Furthermore, these outcomes can be used in classification of objects. The resulting combination of these said methods may also be used as a single classifier as well as for dimensionality reduction before further classification.

3.4. System Model

The backend design is purely is C# desktop forms and the knowledge of OOP is widely used to get the specified outcome.
1. Admin login to the system
2. Admin defines the Training Set (one or more images of one or more persons which have to be recognized)
3. Admin save each picture in training set by giving it a label i.e. name
4. Admin provides the system a full-length video which is monitored by the system
5. after the monitoring part, system asks for to choose a recognition method i.e. by static image, by live camera or by the provided video which was monitored before
6. Admin print out the End results

In support of this work, following are some diagrams explaining the overall workflow
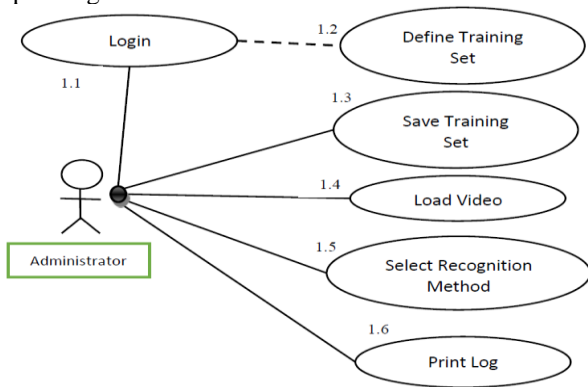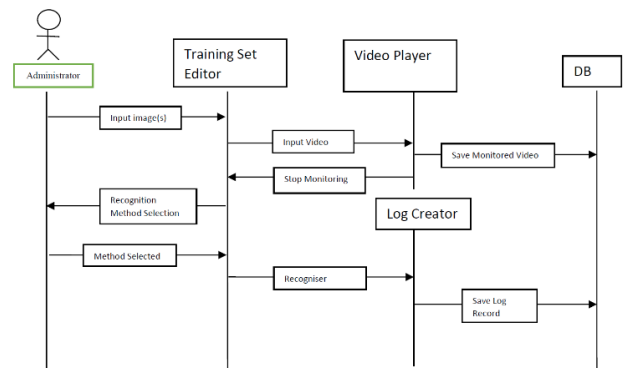

Figure.5 Use Case Diagram
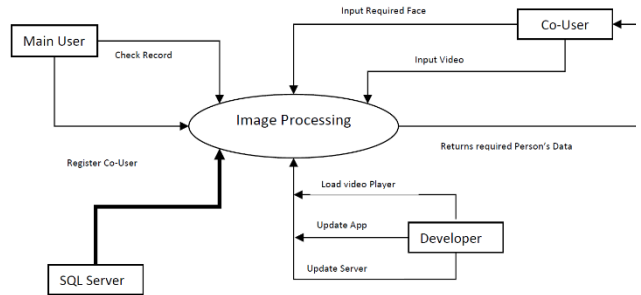

Figure.6 Sequence Diagram
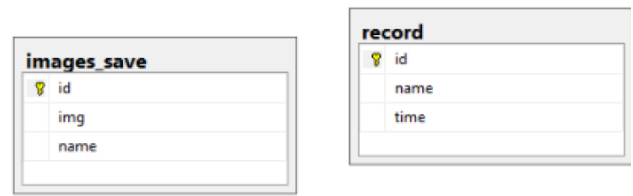

Figure.7 Data Flow Diagram
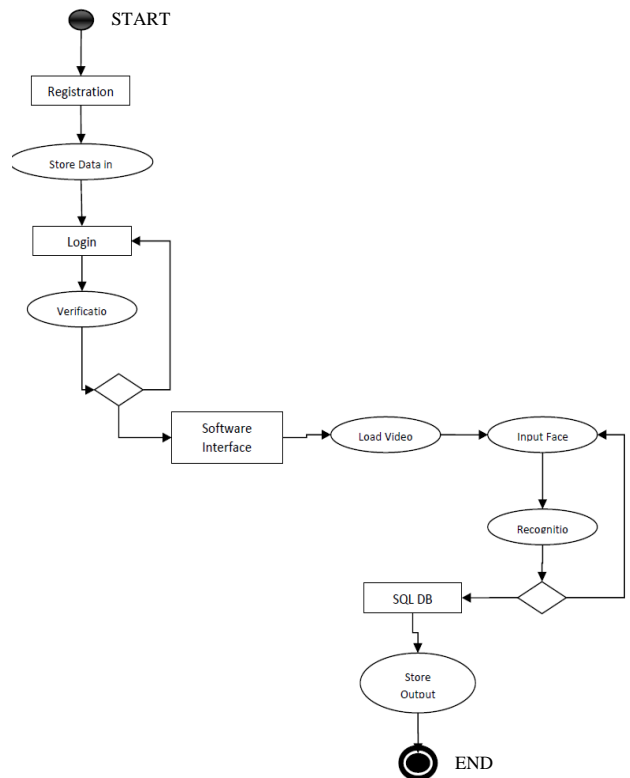

Figure.8 Database Design

Figure.9 State Transition Design

In IFMP, first of all it takes one or more, well defined (HD) image(s) of the person which is to be searched from the video. Then, that video is provided and the software watch that video completely and at the mean time the computational logic process that video at the backend. After that, the software again plays that video and as soon as any face is detected (required or not) it creates a log table which contains the name of the face and the time from the video at which the face is detected. The objectives of developed System are;

1. Ease of access
2. No need for watching the hole video with bare eye to find a face
3. Let the computer do all the tough work in IT section of Security Systems

3.5 Hardware & Software Details

For the sake of better results of recognition, it is necessary to use good hardware which could provide effective results and process efficiently. For example, the basic need in recognition is the camera which have the basic responsibility of providing quality images. Good resolution is required for recognition of unknown persons entering at the gates.
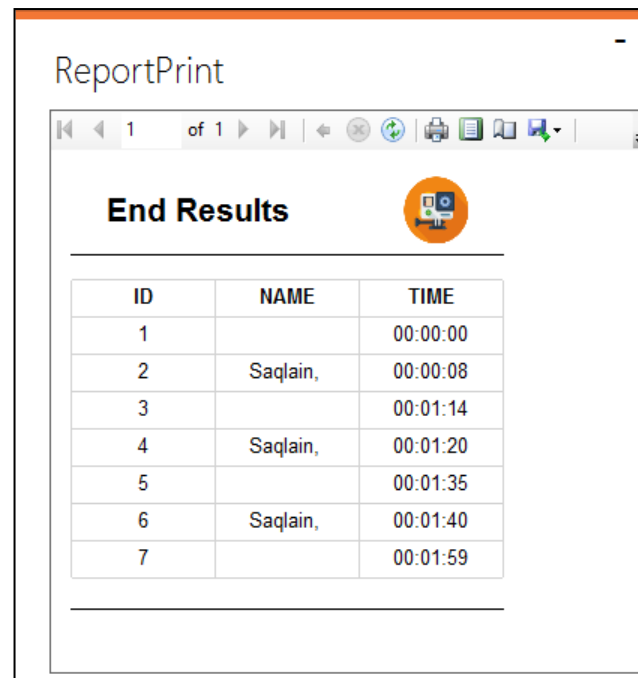
- **DSLR (Digital Single Lens Reflex)**: Initially a Digital Single Lens Reflex camera was used for capturing the face image but there were two major drawbacks of using this. Firstly, DSLR isn't economically feasible to buy. Secondly it takes pictures manually. It cannot be controlled via controller laptop or computer.
- **IP Camera**: Due to less efficiency of DSLR, IP camera was chosen for face detection but this camera also did not yield accurate results because of poor resolution. Hence facial features were not recognized accurately.
- **Webcam:** Since IP camera had poor resolution, so Logitech sphere AF webcam was chosen. It is equipped with lot of functionality and advanced features like digital zoom facility, 102-degree tilt and 189-degree field of view. It also can be integrated with visual C++ easily from its interface. It can intelligently manage the tilt in case of poor conditions for producing as best captures as possible. Because of non-availability of optical zoom it can recognize faces correctly. It can be connected to laptop or PC directly with USB interface.
- **Minimum Computer Hardware**

  1. Windows 10 OS (Latest Edition)
  2. RAM 4GB
  3. Processor 2.3GHz

- **Microsoft Visual Studio:** It is a programming management studio provided by Microsoft. It is widely used for programing and making Graphical User Interfaces as well as desktop and web applications. It has toolkits for debugging programming codes of different languages including C++ and visual C#. The proposed work is encoded by using this studio in visual C# and debugged by using its tools.
- **OpenCV Library:** In Microsoft Visual Studio, OpenCV (Open Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision, developed by Intel. It focuses

mainly on real-time image processing. If the library finds Intel's Integrated Performance Primitives on the system, it will use these proprietary optimized routines to accelerate it. For the proposed work, the researchers have used OpenCV library in Microsoft Visual Studio.
- **EMGU CV Wrapper:** A computer vision library named EMGU CV [8] is used for creation of eigen vectors [5]. It is very handy because of s has Built-in functions and commands widely used in field of face detection and recognition.
- **Microsoft SQL Server:** Microsoft SQL Server is a used as the database management studio in the proposed work to execute the SQL quires and commands. Different tables were managed to save and retrieve the data for processing (Figure 8). By using Entity Framework, the GUI was integrated with the database.

4. Results & Discussion

As a result, log file is constructed which contains each detected face name as well as the time from the video at which the face is recognized. Thus, end results are like this After Defining the Training Set, a video is provided as an input and allow the system to monitor.

Figure 10 Person Recognition logs

There are certain specifications of the project which are aimed to be satisfied.

- Real Time System: The project provides real time system. It can be implemented practically at the entrance and used for in effect security.

- Time Saving: This is saved by using Integrated Security System. Manual checking at the gates can be troublesome, hectic and time consuming.

- **Automated System:** A complete automated system is delivered that requires no human assistance at the gates.

- **Loss of Human Life Is Avoided:** In case of any security attack the loss is human life is avoided. Manual systems posed great threat to the security guards at the entrance of gates.

- **No Limitation on the Number of Entries in Data Base:** There are no limitations on the number of entries that can be added to the database. The length of database can be increased or decreased as per our requirements.

- **No Tampering:** As the system can only be accessed by the authorized persons who are present in the control room no tampering of database is possible by some unknown person.

5. Conclusion and Future Work

IFMP is a desktop-based video player application through which any person can be identified using face recognition and detection through computing algorithms under consideration of framework provided by Microsoft Corporation. To reduce the humane effort, IFMP effectively uses the modern computer vision techniques for face recognition which provides 90% accuracy. It provides a user-friendly interface (GUI) for the person using it. This research work can be used as a base for future research. The researchers may integrate it with other schemes in modules to provide enhance security measures. In this project face recognition for only one-person identification has been used but other researchers may also incorporate more than one persons at a time identification. A centralized database of entire institute may also be maintained by recording the entrance time and departure time of each individual in an institute. This can also be used to control access of individuals in an institute or department.

Acknowledgment

References

[1] Qian Li, Usman Niaz, Bernard Merialdo (16 August 2012), An improved algorithm on Viola-Jones object detector, DOI: 10.1109/CBMI.2012.6269796

[2] Himanshu Sharma, Sumeet Saurav, Sanjay Singh (28 September 2015), Analyzing impact of image scaling algorithms on viola-jones face detection framework, DOI: 10.1109/ICACCI.2015.7275860

[3] Matthew Day, John A. Robinson (02 February 2012), On output-processing in face detectors, DOI: 10.1109/ICSIPA.2011.6144106

[4] Samiksha Agrawal, Pallavi Khatri (06 April 2015), Facial Expression Detection Techniques: Based on Viola and Jones Algorithm and Principal Component Analysis, DOI: 10.1109/ACCT.2015.32

[5] Diego von Söhsten, Sérgio Murilo (13 October 2014), Multiple face recognition in real-time using cloud computing, Emgu CV and Windows Azure, DOI: 10.1109/ISDA.2013.6920723

[6] I. Craw (06 August 2002), Machine Storage and Recognition of Faces

[7] S.V Tathe, A. S. Narote, S. P. Narote, (03 November 2016) Human face detection and recognition in videos DOI: 10.1109/ICACCI.2016.7732378

[8] D. von Söhsten and S. Murilo, "Multiple face recognition in real-time using cloud computing, Emgu CV and Windows Azure," 2013 13th International Conference on Intellient Systems Design and Applications, Bangi, 2013, pp. 137-140. DOI: 10.1109/ISDA.2013.6920723

[9] F. Ye, Z. Shi and Z. Shi, "A Comparative Study of PCA, LDA and Kernel LDA for Image Classification," 2009 International Symposium on Ubiquitous Virtual Reality, Gwangju, 2009, pp. 51-54. doi: 10.1109/ISUVR.2009.26

[10] W. Puyati and A. Walairacht, "Efficiency Improvement for Unconstrained Face Recognition by Weightening Probability Values of Modular PCA and Wavelet PCA," 2008 10th International Conference on Advanced Communication Technology, Gangwon-Do, 2008, pp. 1449-1453. doi: 10.1109/ICACT.2008.4494037