Mass Security Surveillance Model Using Multiple CCTV

Mohan Vamsi A [1], V.D. Ambeth Kumar [2] C.S Manigandaa [3] , Challapalli Manikantaa[3] , P. Sindhu [2], Mayanglambam Sushilata Devi [2] and V.D.Ashok Kumar [4]

[1] Dept. of Computer Science and Engineering, Panimalar Engineering College, 600123, India
[2] Dept. of Computer Engineering, Mizoram University, Aizawl, 7960004, India
[3] Dept. of Artificial Intelligence and Data Science, Panimalar Engineering College, 600123, India
[4] Dept. of Computer Science and Engineering, Loyola Institute of Technology, 600123, India

amvamsi1308@gmail.com, ambeth@mzu.edu.in, csmanigandaa@gmail.com, cmanikantaa57@gmail.com, sindhupandi26@gmail.com , sushilata@mzu.edu.in, akumar2879@gmail.com

Abstract. In the Modern world the most important thing existing is privacy and security. Due to lack of security, people face lot of problems. In an organization it is necessary to maintain complete security from external entities. By using face recognition, the people can be authenticated, categorized and further authenticated according to their identity. To identify and allow access to the people a software based on python is used to recognize and classify the entities into known and unknown and label them in the database. And to simplify the processing of visitors and to provide access, their face data is stored in the data base which provides entry only at the specific time. The same software is used to monitor the entire premises for unknown entities all the time using distributed systems and all the face data stored in the database. If any such entity is identified without authorization in the premises, then immediately it notifies the authority.

Keywords: Multiple Camera, CCTV, Security, Social Network

1 Introduction

The applications for surveillance and monitoring workspaces are on the rise in day-to-day basis. This is because the organization is concerned about the security and privacy of its premises. The proposed system deals with the drawbacks of these existing systems by employing various software algorithms. These algorithms are used to recognize and identify different individuals in the premises. Once a person is identified, he/she is provided with an unique label and their credentials are verified and stored in the database. To monitor a person's activities as well as their whereabouts, one can search and identify him/her for monitoring.

R. Murugan et al. (eds.), *Proceedings of the International Conference on Signal Processing and Computer Vision (SIPCOV 2023)*, Advances in Engineering Research 239,
https://doi.org/10.2991/978-94-6463-529-4_10

A facial recognition application is also capable of detecting criminal suspects, locating people lost in crowds, monitoring children remotely by parents, as well as spotting terrorists and other criminals.
This paper deals with utilizing various similarly bred microprocessors and microcontrollers to access the distributed systems that control the closed circuit television (CCTV) architecture [1][2]. However, different types of video files acquired from the CCTV subsystem should be processed using the software algorithm for detection of lively objects in the environment. These entities detected can be in any state of motion. So detecting and recognizing these items in real time is a protracted process on its own[3][4]. To accomplish this there is a need to perform certain tasks such as object detection, feature extraction, training a model based on the features, and running the trained model in the real-time world to recognize and identify the trained face data[9]. This model can also be used to find specific category of people at the time of a health crisis.
The face data is stored in a distributed system which gives us an edge over the existing systems as it will be more secure and provide easy access to the stored data. It is trained in different ways based on the maximum likelihood data for each face's pixel level and shadow intensity. The process of identification of artifacts can be achieved using Haar-cascades using the Viola and Jones methods. In Viola-Jones, fast computing of features is accomplished with integral images, powerful classifiers are built with Adaboost (also referred to as feature selection), and a cascade structure is created byincorporating chosen features. Firstly, the video captures the event of the person and the contextual person is obtained as well. Scenes and images, as well as the video's structural appearance, are divided into each video. In addition to this, the flashback sequence and frame length are identified by describing the special effects and sound used during the flashbacks. As well as this, a face detection technique is used by assigning labels to the faces in order to identify multiple pose face trackers. This model has its advantages in wide range in recognizing and tracking human beings with the help of constrains provided. Once the system is implemented in an organization it increases the security.

2 Related Work

Generally connecting a CCTV system to a microcontroller (for alarming situation) is a quite challenging task. Alarm Display and Assessment System is consisting of an integral part as the closed circuit television (CCTV) subsystem [1]. A microprocessor is used to control video by initiating record of multiple scenes, switching between live and recorded video, verifying all cameras have video, and reporting critical component failures to the host. Several essential modules are installed in a prototype surveillance system, and mobility is also an important feature. With the system, motion and

anomalies are detected and notifications are sent immediately by email and SMS[2]. A homography-based motion detection method detects and tracks moving objects. An algorithm is used to detect and track moving objects[3] without background modeling[4].

The detection of faces is considered to be one of the most challenging problems in computer vision and image analysis. Successive Mean Quantization Transform (SMQT) for illumination and sensor insensitive operation and Sparse Network of Winnow (SNoW) to speed up the original classifier based face detection technique[5]. In order to create a single high-resolution video frame, spatial and temporal information is taken into account in a short image sequence. The video frame extraction algorithm[6] is used with visual comparison and a data set for video retrieval and frame classification to enhance the resolution obtained from progressively scanned frames[7]. A post-processing tool is used to convert the frame rate from a lower number to a higher one[8].

The faces are identified in real-time in a cluttered scene using a setup for face recognition in a wide range of poses and scales. A person standing in front of the camera (although against a different background) is identified, if their image was present in the database even when the person is under poor lighting conditions[9][10][11]. An algorithm can also be employed to recognize the face and hand gestures using hybrid classifiers[12]. Face recognition software and the various tools have to be equipped with mechanisms that can employ multi-task pose invariant recognition even by using neural networks[13][14].

Matching and recognition of the labeled entities in the database with the live individual captured in the footage has to be made efficient by employing 3 degrees of freedom for the surface representation. Different methodologies like Voila Jones' algorithm are also used to identify the suspect and name them into labeled entities that are referred to when any anomaly is detected in the system[15][16][17].

3 Proposed Method

As discussed earlier, there may be many difficulties faced in maintaining secrecy, providing authority and finding any unauthorized breach.

In this system as shown in Fig 1.0 face data is collected using multiple CCTV cameras for better accuracy than using a single camera. A database is then used to store the face data collected. This is for regular people who come to the organization frequently. For temporary visitors once their request for appointment is accepted then a webcam is used as a resource to collect face data and to store that in the database for future use.

This face data is used to authorize a person to certain levels in the organization and updates the database with the Entry/Exit time. For temporary visitors, they will be authorized to enter the organization only in a particular time fixed for their appointment if they don't show up in time then the system deletes their appointment and labels didn't show up.
This system with the help of face data stored in the database monitors the organization 24/7 for unknown labels. If unknown labels are present then it means that an unauthorized person are present. If it finds any unauthorized person then it notifies the authority for immediate assistance.

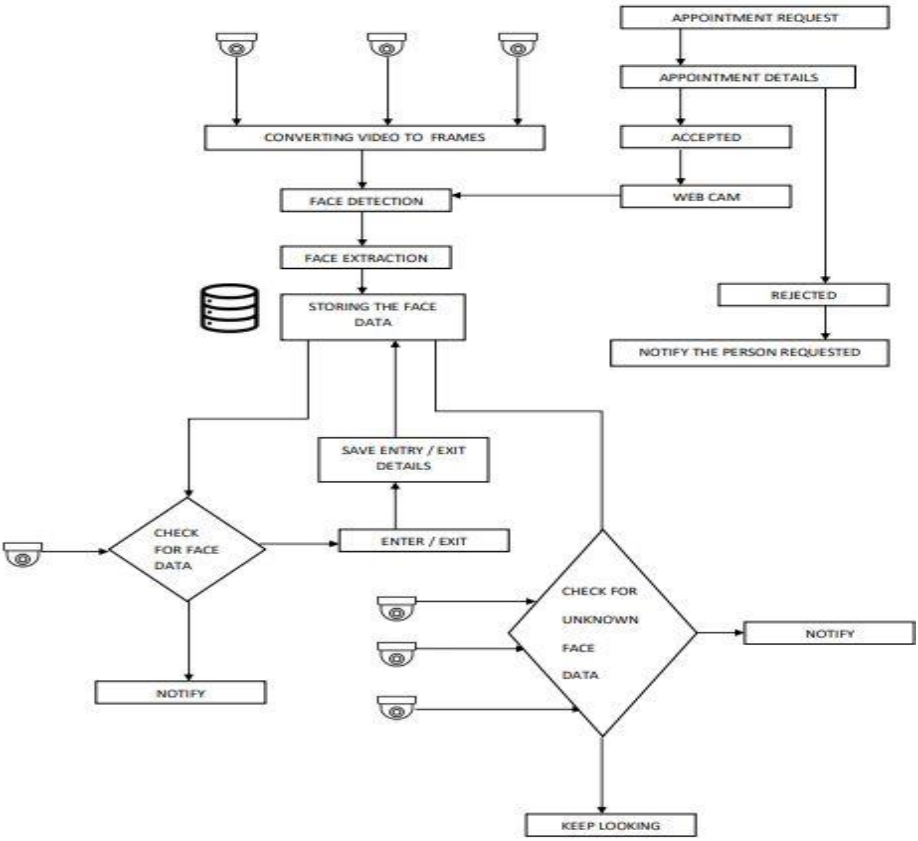

Fig. 1. Proposed System.

The system shown in the above Fig 1 is divided into 4 modules:

- Face Data
- Appointment
- Entry/Exit
- Check Unknown

3.1 Face Data

In this module, footages from cameras are acquired to process the data which enables the system to determine various forms of face data for uniquely identified individuals.

Input: Live video from multiple CCTV cameras.

Output: Storing the face data in a spreadsheet along with the details.

Declare the variables required

- storage
- frame
- video
- face
- FD-face data

```
for frame ⇐video:
    for face ⇐ frame:
            if  face ⇒ TRUE:
  FD ⇐ extract ⇐ face
            else:
 Delete ⇐ frame
storage ⇐ FD
```

Multiple cameras are used to record a live video of a person. Faces are detected from frames of the video after the video has been converted into frames. If no faces are detected, the frame is deleted if no face data was extracted from it. Until the video is complete, this process is repeated. This process saves the person's face data and details to the database for future use.

3.2 Appointment

For storing temporary data of people who are not part of the integral system of operation, webcams are to be used to initialize face data of individuals who have been granted acceptance for the request of their appointment.

Input: Live video from webcam along with other details.

Output: Storing the face data and details in database

Declare the required variables.

- frame

- video
- face
- FD-face data
- extract
- request
- accepted
- rejected
- notify

```
if request ⇾ accepted:
        for frame ⇽video:
                for face ⇽ frame:
    if  face⇾ TRUE:
          FD⇽ extract ⇽ face
                          else:
    Delete ⇽ frame
         storage ⇽ FD
                else:
        notify ⇽ request ⇽ rejected
```

In this algorithm, the face data of the people is collected to create a temporary access into the organization as per the appointment request. People who want an appointment will request the organization online by filling out the details like the purpose of visit and other personal details. Once the details are verified by the organization and accepted, the live video of a person is recorded from the webcam. Faces are detected from frames created from the video. When a face is detected, the face data is extracted from the frame and stored in the database. If no face is detected, the frame is deleted. As long as the video is not completed, this process is repeated. The face data and the person's details are stored in the database after the process is complete. Appointment requests that are rejected are notified to the person.

3.3 Entry/ Exit

People who are part of the system can be identified using this module to grant permission for entry/exit and the exact notification can be forwarded to the authority.

Input: Live image from CCTV camera.
Output: Permission to Entry/Exit.
Declaring the required variables

- LI-Live image
- FD-face data

- Storage
- notify
- E_e-Entry/Exit
- t-Entry/Exit time

```
if LI==FD:
        Permission  ⇾ E_e
        storage ⇽label ⇽ t
else:
        notify ⇾ authority
```

This system is used to authenticate a person based on the information stored in the database. With the help of the face data stored in the database, comparison of it with the live image at present at the entry is done. If it matches with the face data present in the database then it provides access to the person and stores the entry/exit time along with present image creating a label for easy access to the database or else it will notify the authority.

3.4 Check Unknown

The footage from the cameras can also be used to detect any unknown persons' presence in the premises so that the authority is alerted in time to prevent any potential threats.

***Input*:** Live video from multiple CCTV cameras.
Output: Alert to authority if any unknown present.
Declare the required variables.

1. video
2. faces
3. labels
4. notify

```
while(video==1):
      for faces in video:
            if faces == labels:
                  Pass
            else:
  notify ⇾ authority
```

Checking whether any unauthorized person is present in the building using the existing face data stored in the database is done. While storing the face data in the database

a temporary label is created stating whether the person is authorized or not. This can be in the whole building or in a particular secure level only authorized to those at higher level to protect privacy. If any unauthorized person is present then the system notifies the authority.

Table 1. Comparing the Proposed System with the Existing System

Feature	Existing System	Proposed System
Microprocessor for CCTV controller	J. L. Schoeneman[1], Inefficient for controlling multiple CCTV cameras	Multiple microprocessors for a single CCTV and connect all of them to a microcontroller for more efficiency
IoT Enabled Motion CCTV Surveillance System	B. A. Sassani, A. David, X. Li and F. Mehdipour[2], IoT enabled CCTV to detect the human motion without any motion sensors and to notify	Notify the authority once a person is not identified
Moving object detection and tracking from moving camera	Won Jin Kim and In-So Kweon[3] P. Viola and M. Jones[5], Detect the moving object and to track it from a moving camera	Tracks the intruder once he/she is labelled unknown
Frame quality Enhancement from real time video	R. R. Schultz and R. L. Stevenson[6] A. Karpenko and P. Aarabi[7] Y. Yang, Y. Tung and J. Wu[8], Enhancing the quality of the frame extracted from a very large video data	Once the large set of video is stored the system enhances the quality of the frame and delete's the low quality frames to increase the storage capacity
Human detection and Face recognition	C. Nastar and M. Mitschke[9] C. Ding, C. Xu and D. Tao[4] R. Chellappa, C. L. Wilson and S. Sirohey[10] X. Tan and B. Triggs[11] B. Heo, K. Yun and J. Y. Choi[13] Detects and recognizes the face and limited conditions	Detects a human face from a well enhanced framework and recognizes the face from the database

Distributed System for CCTV	Each CCTV has no connection to one and another	CCTV's are connected through a distributed system
Capturing the Face data	Single CCTV camera is used to record the face data of a person into the database	Multiple CCTV cameras from different angels are used to record the face data
Matching the face data from the database	C. Nastar and A. Pentland[15] V D.Ambeth Kumar & Kumar, V.D. & Subramanian, Malathi & Vengatesan, K. & Ramakrishnan.M[16] Image processing is used to extract the features from the face and to match it with the face data present in the database	Feature extraction method is used in enhanced frames from multiple real time CCTV cameras to match the features
Labelling the recognized faces	J. Yang and A. G. Hauptmann[17] Naming every individual in a pre-recorded video	Labeling all the individuals in a real time video and marks unknown if not recognized

The table 1 shows that the proposed system is more efficient than existing system as it consists of all the better versions of various methods. As the existing system only recognizes the people within limited conditions whereas the proposed system satisfy all the conditions with maximum efficiency. The System tracks the unknown inside the premises as the face recognition software runs in the distributed system and as the CCTV's are connected this process is more likely faster than the existing system.
The system only permits a person only if their face data matches and the system runs continuously to detect any unauthorized person in the premises and notifies the authority. But in the existing system the face recognition is done manually by running the software along with the face data collected from training the images. The Proposed system can be used in real time without any human assistance. It saves time as it reduces the training and processing time which also increases the accuracy.

4 Conclusions

A machine learning model to identify and recognize a person labeled in the database stored in a distributed system is discussed in the paper. Once the system is installed in an organization it provides a secure environment with more privacy. The model even

provides an added advantage in tracking the people present inside the institution and it even checks for the intruders present inside the premises. If any unauthorized person is present inside the premises it alerts for immediate assistance.

References

1. J. L. Schoeneman, "A Microprocessor CCTV Controller for Safeguards Applications," in IEEE Transactions on Nuclear Science, vol. 29, no. 1, pp. 874-877, Feb. 1982
2. B. A. Sassani, A. David, X. Li and F. Mehdipour, "AgentPi: An IoT Enabled Motion CCTV Surveillance System," 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), Fukuoka, Japan, 2019, pp. 454-458.
3. Won Jin Kim and In-So Kweon, "Moving object detection and tracking from moving camera," 2011 8th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Incheon, 2011, pp. 758-759.
4. C. Ding, C. Xu and D. Tao, "Multi-Task Pose-Invariant Face Recognition," in IEEE Transactions on Image Processing, vol. 24, no. 3, pp. 980-993, March 2015.
5. P. Viola and M. Jones, “Rapid Object Detection Using a Boosted Cascade of Simple Features,” *Proceedings of the* 2001 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (*CVPR*), Vol. 1, 2001, pp. 511-518.
6. R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," in IEEE Transactions on Image Processing, vol. 5, no. 6, pp. 996-1011, June 1996.
7. A. Karpenko and P. Aarabi, "Tiny Videos: A Large Data Set for Nonparametric Video Retrieval and Frame Classification," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 3, pp. 618-630, March 2011.
8. Y. Yang, Y. Tung and J. Wu, "Quality Enhancement of Frame Rate Up-Converted Video by Adaptive Frame Skip and Reliable Motion Extraction," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 12, pp. 1700-1713, Dec. 2007.
9. C. Nastar and M. Mitschke, "Real-time face recognition using feature combination," Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, 1998, pp. 312-317.
10. R. Chellappa, C. L. Wilson and S. Sirohey, "Human and machine recognition of faces: a survey," in Proceedings of the IEEE, vol. 83, no. 5, pp. 705-741, May 1995.
11. X. Tan and B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions," in IEEE Transactions on Image Processing, vol. 19, no. 6, pp. 1635-1650, June 2010.

12. S. Gutta, H. Huang, F. Imam and H. Wechsler, "Face and hand gesture recognition using hybrid classifiers," Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, Killington, VT, USA, 1996, pp. 164-169.
13. B. Heo, K. Yun and J. Y. Choi, "Appearance and motion based deep learning architecture for moving object detection in moving camera," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 1827-1831.
14. X. Yin and X. Liu, "Multi-Task Convolutional Neural Network for Pose-Invariant Face Recognition," in IEEE Transactions on Image Processing, vol. 27, no. 2, pp. 964-975, Feb. 2018.
15. C. Nastar and A. Pentland, "Matching and recognition using deformable intensity surfaces," in *Computer Vision, International Symposium on*, Coral Gables, Florida, 1995 pp. 223
16. V D, Ambeth Kumar & Kumar, V.D. & Subramanian, Malathi & Vengatesan, K. & Ramakrishnan, M.. (2018). Facial Recognition System for Suspect Identification Using a Surveillance Camera. Pattern Recognition and Image Analysis. 28. 410-420