

## Crowd Detection by Video Processing

Dr. S. Gokulraj<sup>1</sup>, Soundarya. B<sup>2</sup>

<sup>1</sup>M.E,Ph.D. Department of CSE, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India

<sup>2</sup>PG Scholar, Department of CSE, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India

### ABSTRACT

The video may consist of multiple shots. The target in the probe video is only annotated once with a face bounding box in a frame. Most video face identification techniques assume that the video is of single shot, and thus the bounding boxes of the target face can be extracted by tracking a face across the video frames. Nevertheless, such automatic annotation is vulnerable to the drifting of the face tracker, and the face tracking algorithm is inadequate to associate the face images of the target across multiple shots. A target face association (TFA) technique retrieves a set of representative face images in a given video that are likely to have the same identity as the target face. These face images are then utilized to construct a robust face representation of the target face for searching the corresponding subject in the gallery. Since two faces that appear in the same video frame cannot belong to the same person, such cannot-link constraints are utilized for learning a target-specific linear classifier for establishing the intra/ inter-shot face association of the target. Experimental results on the newly released JANUS challenge set 3 (JANUS CS3) dataset show that TFA method generates robust representations from target-annotated videos and demonstrates good performance for the task of video-based face identification problem.

**Keywords :** Target Face Association Technique, Facial Recognition Technique, Biometric Artificial Intelligence, Image Processing, Segmentation.

### I. INTRODUCTION

#### A. VIDEO PROCESSING

**Video Processing** uses methods where the temporal nature of video data is exploited. Image Analysis goal is to analyze the image with the purpose of first finding objects of interest and then extracting some parameters of these objects. For example, finding, an object's position and size.

Video processing is a particular case of signal processing. In particular image processing, which often employs video filters and where the input and output signals are video files or video streams. Video processing techniques are used in television sets, VCDs, DVDs, video codec's, video scalars and other devices. Video and image processing are used in many diverse applications, such as astronomy (to enhance the quality), medicine (to measure and understand some parameters of the human body, e.g., blood flow in fractured veins), image compression (to reduce the memory requirement when storing an image), sports

(to capture the motion of an athlete in order to understand and improve the performance), rehabilitation (to assess the locomotion abilities), motion pictures (to capture actors' motion in order to produce special effects based on graphics), surveillance (detect and track individuals and vehicles), production industries (to assess the quality of products), robot control (to detect objects and their pose so a robot can pick them up), TV productions (mixing graphics and live video, e.g., weather forecast), biometrics (to measure some unique parameters of a person), photo editing (improving the quality or adding effects to photographs), etc.

The different video and image processing methods are often grouped into the categories. Video and Image Compression is probably the most well defined category and contains the group of methods used for compressing video and image data. Image Manipulation category covers methods used to edit an image.

Image processing originates from the more general field of signal processing and covers methods used to segment the object of interest. Segmentation refers to methods which enhance the object while suppressing the rest of the image (for example the edges in an image). Image processing or image analysis in production industries is normally referred to as machine vision or simply vision. In Computer Vision, human has human vision and similarly a computer has computer vision. Normally computer vision also covers all methods where more than one camera is applied.

## **B. VIDEO BASED FACE IDENTIFICATION**

A facial recognition system is a technology capable of identifying or verifying a person from a digital image or a video frame from a video source. There are

multiple methods in which facial recognition systems work, but in general, they work by comparing selected facial features from given image with faces within a database. It is also described as a Biometric Artificial Intelligence based application that can uniquely identify a person by analyzing patterns based on the person's facial textures and shape.

While initially a form of computer application, it has seen wider uses in recent times on mobile platforms and in other forms of technology, such as robotics. It is typically used as access control in security systems and can be compared to other biometrics such as fingerprint and iris recognition systems. Although the accuracy of facial recognition system as a biometric technology is lower than iris recognition and fingerprint recognition, it is widely adopted due to its contactless and non-invasive process. Recently, it has also become popular as a commercial identification and marketing tool. Other applications include advanced human-computer interaction, video surveillance, automatic indexing of images, and video database, among others.

## **II. RELATED WORKS**

### **A. FACE DESCRIPTION WITH LOCAL BINARY PATTERNS: APPLICATION TO FACE RECOGNITION**

A novel and efficient facial image representation is based on local binary pattern (LBP) texture features. The face image is divided into several regions from which the LBP feature distributions are extracted and concatenated into an enhanced feature vector to be used as a face descriptor. The performance of the proposed method is assessed in the face recognition problem under different challenges. Automatic face analysis which includes face detection, face recognition, and facial expression recognition has

become a very active topic in computer vision research. A novel descriptor is based on local binary pattern texture features extracted from local facial regions.

## **B.AUTOMATIC FACE RECOGNITION FOR FILM CHARACTER RETRIEVAL IN FEATURE-LENGTH FILMS**

The objective is to recognize all the frontal faces of a character in the closed world of a movie or situation comedy, given a small number of query faces. This is challenging because faces in a feature-length film are relatively uncontrolled with a wide variability of scale, pose, illumination, and expressions, and also may be partially included. A recognition method is developed based on a cascade processing steps that normalize for the effects of the changing imaging environment.

In particular there are three areas of novelty: (i) Suppressing the background surrounding the face, enabling the maximum area of the face to be retained for recognition rather than a subset; (ii) A pose refinement step is included to optimize the registration between the test image and face exemplar; (iii) Robust distance is used to a sub space to allow for partial occlusion and expression change.

## **C.AUTOMATIC DETECTION OF PLAYER'S IDENTITY IN SOCCER VIDEOS USING FACES AND TEXT CUES**

In soccer videos, most significant actions are usually followed by close-up shots of players that take part in the action itself. Automatically annotating the identity of the players present in these shots would be considerably valuable for indexing and retrieval applications. Due to high variations in pose and illumination across shots however, current face recognition methods are not suitable for this task. The inheritant multiple media structure of soccer videos is

shown which can be exploited to understand the player's identity without relying on direct face recognition. The method is based on a combination of interest point detector to "read" textual cues that are allowed to label a player with its name, such as the number depicted on its jersey, or the super imposed text caption showing its name. The results are obtained from soccer videos taken from various recent games between national teams.

## **D.SCALABLE FACE TRACK RETRIEVAL IN VIDEOS ARCHIVES USING BAG-OF-FACES SPARES REPRESENTATION**

Huge videos archives consisting on new programs, dramas, movies and web videos (e.g. YouTube) are available in our daily life. In all this videos, human is usually one of the most important subjects. Using state of the art techniques, the videos can be efficiently detected. In order to organize large scale face tracks, containing sequence of detected faces in the videos, a efficient method is proposed to retrieve human face tracks using bag-of-faces sparse representation. A face track is encoded as a single bag-of-faces sparse representation and therefore allowing efficient indexing method to handle large scale data.

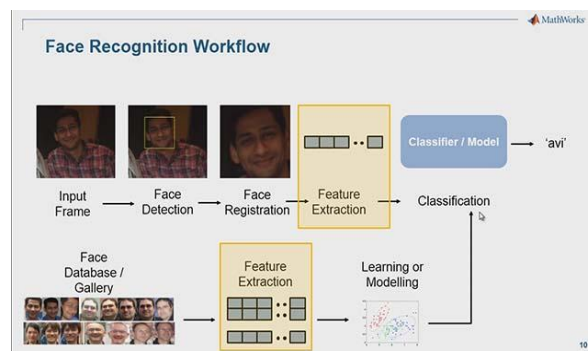
## **III. PROPOSED SYSTEM**

Intelligent video surveillance is an important branch of computer vision research, including motion detection, moving target classification and tracking, monitoring target behavior understanding and description using computer vision and image processing method. Anomaly detection can be divided into two types: global abnormal event and local abnormal event. Global abnormal event indicate that the whole frame is abnormal irrespective of the neighboring regions; local abnormal event indicate that behavior in a specific local frame area is different

from that of its neighbors in spatial-temporal terms. Global abnormal event detection aims at classifying each frame as either abnormal or normal, while in local detection it localizes the parts of the given frame which likely contains the abnormal activity. Typically, abnormal behavior analysis for people included individuals, a small number or a crowd abnormal detection.

Surveillance video anomaly detection is essentially a pattern recognition problem, anomaly detection methods with neural networks, decision trees, cluster analysis, Bayesian theory, support vector machines (SVM), and K neighboring value algorithm were widely used. SVM is an emerging data classification technique that has been widely adopted in various fields of classification problems in recent years. SVM takes use of the principle of minimum risk-based structure to improve anomaly detection performance, with high generalization ability and classification accuracy. But SVM also has some deficiencies, the relatively large training sample will occupy a larger memory, consume more training time, parameter optimization of kernel function when selecting and classification is also a challenging problem. In the SVM, the model for classification is available from the training process with the training data. The biggest problem of the SVM model is the way to select the kernel function and its parameter values. Classification problems generally involve a number of features. However, some of them may be redundant or even irrelevant. Better performance may be obtained by discarding some features.

#### IV. MODULE DESCRIPTION



#### System Architecture

##### A.DATABASE COLLECTION

A database is an organized collection of data, generally stored and accessed electronically from a computer system. Where databases are more complex they are often developed using formal design and modeling techniques.

The database management system (DBMS) is the software that interacts with end users, applications, and the database itself to capture and analyze the data. The DBMS software additionally encompasses the core facilities provided to administer the database. The sum total of the database, the DBMS and the associated applications can be referred to as a "database system". Often the term "database" is also used to loosely refer to any of the DBMS, the database system or an application associated with the database.

##### B.DATA CREATION

A dataset is a collection of data. Datasets that are so large that traditional data processing applications are inadequate to deal with them are known as big data. In the open data discipline, data set is the unit to measure the information released in a public open data repository. A dataset is a set of objects that store data from a database in memory and support (CRUD) operations on that data without the need to be always

connected to the database. Operations are usually referred to using the acronym CRUD.

Within a database each of these operations map directly to a series of commands, however their relationship with a Restful API is slightly more complex. A workaround for this is to use a URL parsing library to dynamically download the videos from the you tube links and overwrite the videos currently in memory which were used in a previous batch size. To speed this up, a parallel computing system is used such that these batches can be loaded and pre-processed on a separate machine than the one which is training the model.

Aggregating predictions across half-second clips is a similar concept to test-time augmentation in image classification. A class prediction for an image is combined with predictions of the same image after it has been rotated, clipped, cropped, or had its color space augmented.

### **C.IMAGE PREPROCESSING**

The Input image is taken and it is resized to constant size (120x160). Preprocessing is necessary in order to correctly classify images. Pre-processing is used to enhance the chances of detecting the suspicious region. The image is converted to gray scale image to reduce the complexity of the image. Gray scale transformation does not depend on the position of the pixel. Gray scale transformation converts the image with equally distributed brightness level over the whole brightness scale. Gray scale image is converted into binary image. Binary image is a digital image that has only two possible values for each pixel. It is also called as Bi-level image. Each pixel is stored as single bit (i.e. 0 or 1). The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some images features important for further processing.

### **D.IMAGE SEGMENTATION**

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as image objects). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture.

A facial recognition system is a technology capable of identifying or verifying a person from a digital image or a video frame from a video source. There are multiple methods in which facial recognition systems work, but in general, they work by comparing selected facial features from given image with faces within a database. It is also described as a Biometric Artificial Intelligence based application that can uniquely identify a person by analyzing patterns based on the person's facial textures and shape.

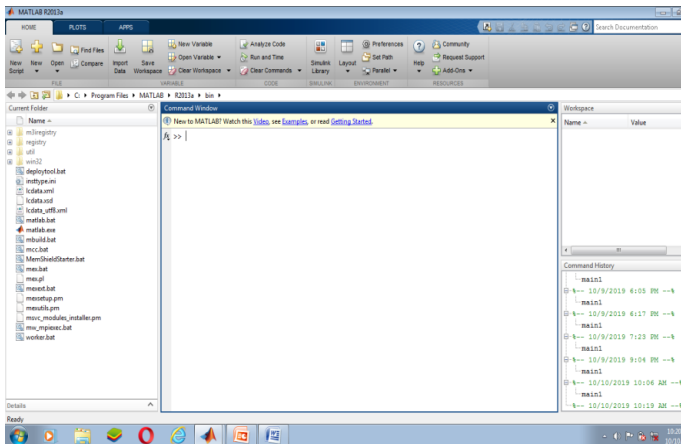
### **E.FACE IDENTIFICATION**

Face recognition is one of the many wonders that AI research has brought forward to the world. Most images of human faces have two eyes, a nose, lips, forehead, chin, ears, hair... that rarely changes. Faces are different from each other. Face of the same person changes with emotion, expression, age... In fact just change in orientation creates a different image.

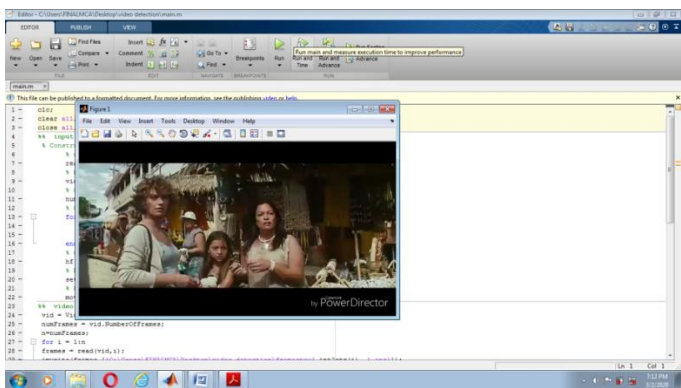
There are some components of a face that are related to age, emotion and orientation. While there are some other components that are stick to the person irrespective of the age, emotion, etc. Further, we can say that these components are not orthogonal or independent. We have all seen people who look alike sideways, but very different otherwise. Or, a kid in the family reminds people of his parents at that age, etc. So, it is not so easy to logically identify these individual components. It works in a few steps: 1. Identify a face in a given image. 2. Identify specific features in the face.

### V. SCREENSHOTS FOR OUTPUT

1.

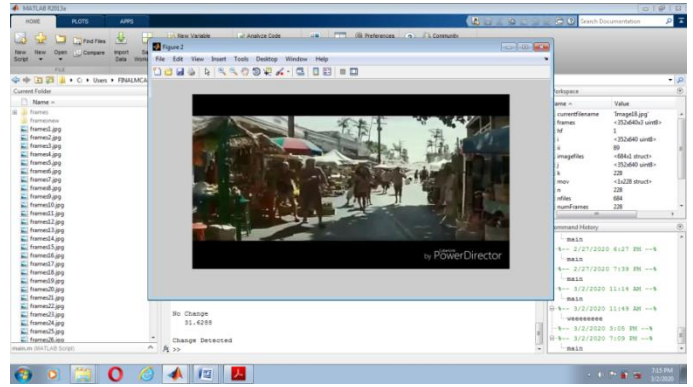
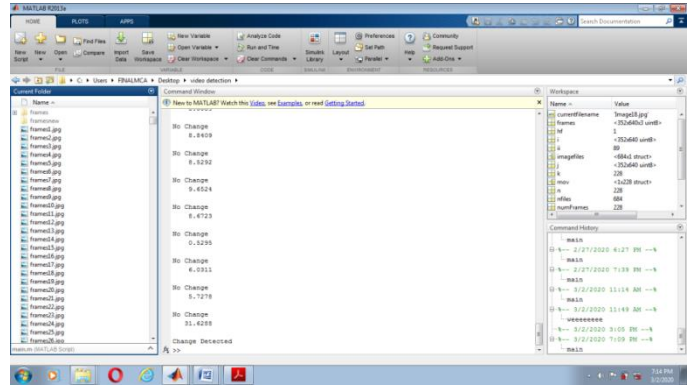


2.



3.

4.



### VI. CONCLUSION

A hybrid optimization of feature selection and the SVM model is proposed. The feature selection optimization combines the simulated annealing algorithm and genetic algorithm, and simultaneously realizes the adaptive crossover and mutation probability. The detection precision of the surveillance video abnormal events is improved; optimization of parameters for SVM training model is adopted. The scheme is compared with separate optimization based on GA and PSO. The experimental results demonstrate that our hybrid optimization can reduce the detection time effectively and improve the accuracy of classification in video surveillance abnormal detection.

### VII. REFERENCES

- [1]. T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, Dec.2006.
- [2]. O.Arandjelovic and A. Zisserman. Automatic face recognition for film character retrieval in feature-length films. In *IEEE Conferenc on Computer Vision and Pattern Recognition (CVPR)*, 2005.
- [3]. M.B Auml, M. Tapaswi, and R. Stiefelhaben. Semi-supervised learning with constraints for person identification in multimedia data. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
- [4]. L. Ballan, M. Bertini, A. D. Bimbo, and W.Nunziati. Automatic detection and recognition of players in soccer videos. In *Proceedings of International Conference on Visual Information Systems*, 2007.
- [5]. J. R. Beveridge, H. Zhang, P. J. Flynn, Y. Y. Lee, V. E. Liong, J. W. Lu, M. de Assis Angeloni, T. de Freitas Pereira, H. X. Li, G. Hua, V. Struc, J. Krizaj, and P. J. Phillips. The IJCB 2014 PaSC video face and person recognition competition. In *International Joint Conference on Biometrics (IJCB)*, 2014.
- [6]. B. C. Chen, Y.-Y. Chen, Y.-H. Kuo, T. D. Ngo, D.-D. Le, S. Satoh, and W. H. Hsu. Scalable face track retrieval in video archives using bag-of-faces sparse representation. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016.
- [7]. J.-C.Chen,R.Ranjan,S.Sankaranarayanan, A. Kumar, C.-H. Chen, V. M. Patel, C. D. Castillo, and R. Chellappa. An end-to-end system for unconstrained face verification with deep convolutional neural networks. *arXiv preprint arXiv:1605.02686*, 2016.
- [8]. R. G. Cinbis, J. Verbeek, and C. Schmid. Unsupervised metric learning for face identification in TV video. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
- [9]. F. Comaschi, S. Stuijk, T. Basten, and H. Corporaal. Online multiface detection and tracking using detector confidence and structured SVMs. In *IEEE International Conference on Advanced Video and Signal based Surveillance (AVSS)*, 2015.
- [10]. N.Crosswhite, J. Byrne, O. M. Parkhi, C. Stauffer, Q. Cao, and A. Zisserman. Template adaptation for face verification and identification. *arXiv preprint arXiv:1603.03958*, 2016.

**Cite this article as :**

Dr. S. Gokulraj, Soundarya. B, "Crowd Detection by Video Processing", *International Journal of Scientific Research in Science and Technology (IJSRST)*, Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 7 Issue 2, pp. 119-125, March-April 2020. Available at doi : <https://doi.org/10.32628/IJSRST207224>  
Journal URL : <http://ijsrst.com/IJSRST207224>