International Journal of Scientific Research and Engineering Development-– Volume2 Issue 2, Mar –Apr 2019

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

ANALYSIS OF HUMAN CROWD BEHAVIOUR IN SURVEILLANCE VIDEOS USING SIFT TECHNIQUE

R.Abirami*, S.Manjula**

*(Master of Software Engineering, Periyar Maniyammai Institute of Science and Technology, Thanjavur)

** (Master of Software Engineering, Periyar Maniyammai Institute of Science and Technology, Thanjavur)

----------**********----------

Abstract:

Security and reconnaissance are imperative issues in this day and age. The ongoing demonstrations of psychological warfare have featured the earnest requirement for productive reconnaissance. Contemporary observation frameworks utilize computerized video recording (DVR) cameras which play host to different channels. The real downsides with this model is that it requires nonstop manual observing which is infeasible due to factors like human weariness and cost of difficult work. In addition, it is for all intents and purposes im-conceivable to scan through chronicles for imperative occasions in the past since that would require a playback of the whole span of video film. Thus, there is for sure a requirement for a computerized framework for video observation which can identify strange exercises individually. This overhead incorporates identification of the sort of exercises which would occur around there and after that surfacing with a Finite State Machine (FSM) show which precisely catches routine exercises and banners non routine ones. A framework with self learning capacity would be anything but difficult to convey and would make it conceivable to have substantial scale checking.

Keywords **— Security, video recording, Image classification, Complimentary Metal-oxide semiconductor.**

----------**********----------

I. INTRODUCTION

Since the beginning people have esteemed their very own life and the lives of their friends and family to the exclusion of everything else. Next in esteem has been their property. Throughout the hundreds of years numerous strategies have been created to ensure property against intruders or aggressors taking steps to take or decimate it. In the past as in the present, producing, modern and government associations have enlisted 'gatekeepers' to secure their offices.

As electric innovation progressed be that as it may, caution frameworks and video were presented, the most headway accompanying the presentation of the Solid State Video Camera during the 1980s. By the mid 90s, the strong state camera utilizing a

Charge Coupled Device (CCD) picture sensor was the result of decision for new security establishments. The strong state CCD sensor and the more current metal-oxide semiconductor (MOS) and Complimentary Metal-oxide semiconductor (CMOS) sensors have longer life and are steady over every single working condition. Another factor in the dangerous utilization of video in security frameworks has seen the quick improvement in hardware ability at reasonable costs.

The 1990s saw the mix of PC innovation with video security innovation. Every one of the parts were strong state. Computerized video innovation required expansive scale advanced recollections to control and store video pictures and the PC business had them. To accomplish palatable video picture transmission and capacity, the video flag must be compacted to transmit it over the current thin band telephone line systems. The video-PC industry previously had pressure for communicated, mechanical and government prerequisites. The video business required a quick and minimal effort intends to transmit the video pictures to remote area and the US government's Defense Advanced Research Projects Agency (DARPA) had officially built up the web (and intranet), the forerunner of the World Wide Web (WWW). The web (and intranet) correspondence channels and the WWW currently give the capacity to transmit and get video and sound and impart and control the information from anyplace.

2. LITERATURE SURVEY

Unsupervised image classification

The specific exercises they are modified for, however are not deployable in a general sce-One of the main effective endeavors at unsupervised characterization in the territory on com-puter vision was that of picture order. Sivic in the paper titled Discovering Object Categories in Image Collections ([8]) utilized UAI for unsupervised picture grouping. The idea utilized was basic. It included utilizing highlights like vector quantized SIFT de-scriptors processed on relative covariant locales as visual recordings portraying pictures. This decision of word on being liable to bunching utilizing UAI arranged pictures on premise of the items contained in them.

Attempts at supervised video classification

There have been various endeavors at directed learning dependent on learning customized show for exercises (as portrayed in [6], [3]). These are genuinely effective at recognizing nario. They experience the ill effects of inconveniences of non versatility regarding their organization and establishment (as referenced in the inspiration for unsupervised observation). There have been different endeavors at unsupervised action grouping (as depicted in [7], [4]). In any case, these have principally been on explicit properties of exercises like just direction or shape, and so on. They can give a sound model to unsupervised learning of the

particular trait they are intended for, yet neglect to give a conventional model for generally speaking movement examination.

Unsupervised Video analysis using UAI for single object using Video Epitomes

UAI has been effectively utilized for unsupervised video investigation ([1]). This utilized epit-ome subtraction to acquire space time fixes as recordings for the video archive and provided food just to single individual movement examination. Additionally, the utilization of video encapsulations made this ex-tremely computationally costly making it troublesome for arrangement in an online man-ner. The present task is essentially an augmentation of this work to multi individual action examination in a way which is possible for online arrangement.

3. METHODOLOGY:

3.1. OBJECTIVES

- To structure a model for unsupervised order of recordings of single individual exercises and thus recognize abnormal exercises.
- To broaden the methodology for single individual video grouping to multi individuals action order.

3.2. PROPOSED SYSTEM:

This proposed model, UAI there is a learning stage and an order phase (where inquiries are really replied). The learning stage learns classes and their qualities and the arrangement stage gives the probability of another report having a place with every one of the classes learnt. On the off chance that the probability does not propose that it has a place with any of the classes adapted then it is hailed as a strange action. So the motivation behind this task we center just around getting these bunches for an expansive arrangement of settings. The sort of grouping which is acquired depends to a great extent on the selection of recordings that we use to catch the data contained in the archive. Thus, a suitable selection of recordings is fundamental for acquiring great groups. We center our push to acquiring a decent arrangement of recordings, which can give us great consistent grouping

3.3. PHASES:

1. **Surveillance video input**
2. **Foreground and background classification**
3. **Crowd tracking metric**
4. **Human video clusters**
5. **Reports**

3.3.1. SURVEILLANCE VIDEO INPUT

This module bargains the video recordings for the CCTV (Closed Circuit Television) alludes to an arrangement of observation cameras that sends signs to a particular area as contribution to the proposed procedure of group reconnaissance framework. This module of the venture is to structure a video observation framework to be utilized in an establishment to fill in as a security measure, and evaluation of the viability of such a

framework as methods checking a region to find out its security

3.3.2. FOREGROUND AND BACKGROUND CLASSIFICATION

The significant work of this module is to PC vision handling it is critical (and is regularly the initial step) to isolate out articles from the picture that we are seeing. This procedure is called foundation subtraction. All the more formally, it is the procedure of expulsion of static foundation highlights from a picture. The highlights of a picture which are not some portion of the foundation are known as the closer view objects. For instance, in the event that we have a camera fixed at a traffic light, the vehicles will fill in as closer view items and highlights like zebra-crossing, light post, and so on will fill in as foundation highlights. Consider for instance the accompanying unique picture outline and the relating closer view outline acquired after foundation subtraction.

3.3.3. CROWD TRACKING METRIC

This module procedure the ordering and recovery strategy that utilizes a numerical system called SIFT(Segment Invariance fragmentary Transform) to recognize designs in the connections between the terms and ideas contained in an unstructured accumulation of content. Filter depends on the rule that recordings that are utilized in similar settings will in general have comparable implications. A key component of SIFT is its capacity to remove the calculated substance of a group of content by setting up relationship between terms that happen in comparable settings. It is called Latent Semantic Indexing on account of its capacity to associate semantically related terms that are idle in a gathering of video outlines.

3.3.4. HUMAN VIDEO CLUSTERS

The significant procedure of this module is to abuses the perception that comparative recordings happen in comparable archives and comparable reports contain comparable recordings. Thus, it is effectively ready to get the hang of moving people (distinctive recordings having comparative items, by their concurrent events in the recordings) and polysemy(videos with numerous implications, by their events in various diverse settings) in a totally unsupervised way. Filter Analysis is a novel measurable procedure for the examination of two mode and co-event information, which has applications in data recovery and separating, common recordings, CCTV recordings, AI knowledge, and in related zones. Contrasted with standard Latent Semantic Analysis which originates from direct variable based math and plays out a Singular Value Decomposition of co-event tables, this strategy depends on a blend decay got from an inactive class show. This outcome is in a progressively principled methodology which has a strong establishment in measurements**.**

REPORTS

This report demonstrates the examination of SIFT with the estimation parameters as far as time, proficiency took in classes of recordings from a given learning informational collection. Despite the fact that novel video order has been contemplated completely by this outcome module.

ALGORITHM:

STEP:1: Dividing the video stream in short clips (our video documents)

//The video stream is obtained as a regular sequence of frames. The entire video is divided into short 20 frame clips which make up one document. So the entire video gives us a set of documents which we need to cluster into classes based on their similarity//

STEP: 2: Background subtraction to obtain foreground blobs

//We then perform background subtraction onto each of this clip to obtain foreground frames corresponding to each video frame. The background subtraction module that was used is based on the Gaussian Mixture Model (as presented in [9]) which is noise resistant and adapts well to changes in the backgrounds. We use the standard open CV implementation of this algorithm with appropriate learning rate parameters. Also, it apply the GMM algorithm not on the raw RBG image but on a normalized RBG as this reduces the effects of shadows and intensity changes//.

STEP: 3: Obtaining videos and other objects from foreground images(in Matlab)

//As described above, the volume of size dx ∗dy ∗dt located at (x, y, t)(x, y ∈{0, 5, 10, 15, ..., },t ∈{0, 2, 4, ...})having more than d fraction of their pixels as foreground pixels is treated as a patch and the word (x, y, t) is said to belonging to this video document. This search is done over the entire space of (x, y, t) and valid videos are added to the document//.

STEP: 4: Obtaining dictionary of humans contained in this document set (in Matlab)

//Once videos describing each individual document are identified for all the training videos, we do a clustering on this set of videos to come up with a dictionary of top k videos common to all documents. The universe of videos that we have defined is large and captures only limited amount of spatial closeness in this (x, y, t) space. So this clustering of videos into k classes groups similar videos (which are relatively nearby as indicated by their (x, y, t) coordinates) and captures common features across different documents. Consider for example the set of videos (5,10,1), (5,15,1), (10,10,1), (40,10,2), (45,15,2) and (50,10,3). A clustering on this set of videos to obtain top 2 videos would map videos (5,10,1), (5,15,1), (10,10,1) to word 1 and (40,10,2), (45,15,2), (50,10,3) to word 2. So,2 different documents containing videos (5,10,1) and (5,15,1) (which are very similar but not exactly same) would be said to be containing the same word, word //

STEP:5: This thus captures common features across different documents.

//UAI treats all videos as unrelated to one another (in fact it tries to learn the relation- ship between them). Hence without this step two documents would be classified as same only if they are exactly same. Even very similar documents would be classified as differ ent, as the videos they contain are similar but not exactly same (and UAI sees them as completely different videos).This clustering was done using the k-means function in Matlab//.

STEP:6: Generation of frame by frame document matrix (in Matlab)

//Once the dictionary of videos is obtained, raw videos in each document are mapped to one of the k common videos from the dictionary and a document × word matrix is obtained (($N_{i,j}$) entry in the matrix indicates number of occurrence of word j in document i). This matrix is much the same as the one obtained for classification of literary documents (used by any web search engine)//.

STEP:7 Running UAI on this human identified document matrix to obtain video clusters(in Matlab)

//The document word matrix thus obtained is processed through UAI to obtain n clus- ters. The algorithm gives us probability of each document and each word belonging to a particular cluster. Similar documents and videos are the ones which have high probability of belonging to the same cluster//.

4. RESULT AND DESCRIPTION:

The learning phase learns classes and their characteristics and the classification phase gives the likelihood of a new document belonging to each of the classes learnt. If the likelihood does not suggest that it belongs to any of the classes learnt then it is flagged as an unusual activity. So the purpose of this project we focus only on obtaining these clusters for a broad set of settings.

The kind of clustering which is obtained depends largely on the choice of videos that we use to capture the information contained in the document. Hence, an appropriate choice of videos is essential for obtaining good clusters. We focus our effort to obtaining a good set of videos, which is able to give us good logical clustering.

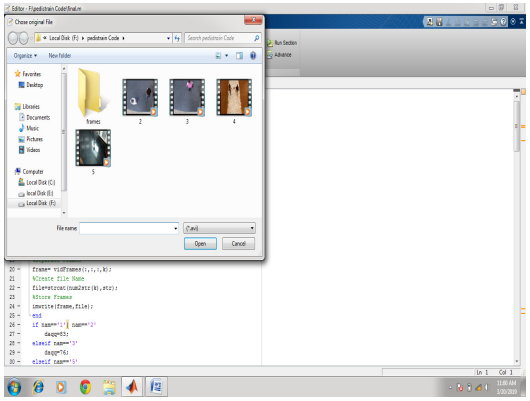

Fig.1. Output Analysis

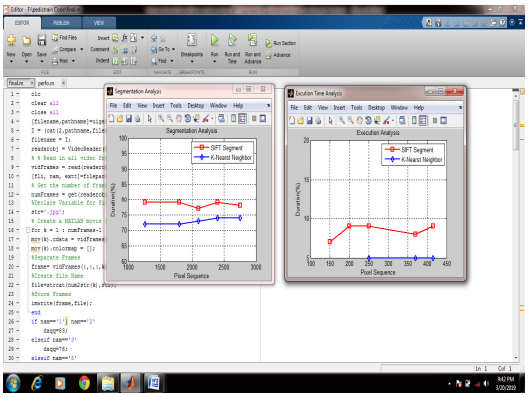

Fig.2. Output Efficiency

5. CONCLUSION:

This proposed strategy for video observation and human following demonstrates the examination of SIFT with the estimation parameters as far as time, effectiveness took in classes of recordings from a given learning informational index. In spite of the fact that novel video grouping has been considered completely by this outcome module, and likewise have displayed multi individual action arrangement utilizing SIFT, Moreover, as referenced, the probability of a video with different exercises gets

split among classes portraying the exercises contained in it. In future his may represent an issue in hailing an un-normal movement, since it depend on a confounded match of the irregular word grouping with the officially learned classes. Henceforth, we have to explore how to separate between recordings with different exercises and recordings with surprising exercises.

6. REFERENCES:

[1] Ayesha Choudhary, Manish Pal, Subhashis Banerjee, and Santanu Chaudhury. Unusual activity analysis using video epitomes and plsa. In *ICVGIP*, pages 390–397, 2008.

[2] Thomas Hofmann. Probabilistic latent semantic analysis. In *In Proc. of Uncertainty in Artificial Intelligence, UAI'99*, pages 289–296, 1999.

[3] S. Hongeng and R. Nevatia. Multi-agent event recognition. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 84– 91 vol.2, 2001.

[4] Fan Jiang, Ying Wu, and Aggelos K. Katsaggelos. A dynamic hierarchical clustering method for trajectory-based unusual video event detection. *IEEE Trans. Image Process.*, 18:907–913, 2009.

[5] B. Klare and S. Sarkar. Background subtraction in varying illuminations using an ensemble based on an enlarged feature set. *Computer Vision and Pattern Recognition Workshop*, 0:66–73, 2009.

[6] Dhruv Mahajan, Nipun Kwatra, Sumit Jain, Prem Kalra, and Subhashis Banerjee. A frame- work for activity recognition and detection of unusual activities. In *ICVGIP*, pages 15–21, 2004.

[7] C. Piciarelli, G.L. Foresti, and L. Snidara. Trajectory clustering and its applications for video surveillance. *Advanced Video and Signal Based Surveillance, IEEE Conference on*, 0:40–45, 2005.

[8] J. Sivic, B. C. Russell, A. A. Efros, A. Zisserman, and W. T. Freeman. Discovering ob- ject categories in image collections. In *Proceedings of the International Conference on Computer Vision*, 2005.

[9] Chris Stauffer and W.E.L Grimson. Adaptive background mixture models for real-time tracking. In *CVPR99, Fort Colins, CO*,1999.

[10] S. Manjula , Dr. Lakshmi " A study on Object Detection" International journal of pharmacy and technology Volume 2 ISSUE 4 2016 p.p:22875-22885.

[11] . S. Manjula , Dr. Lakshmi " A survey on Event detection in Surveillance Videos" International journal of Scientific Research ISSN 0976-2876.

