IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Unusual Crowd Activity Detection

[1]Dhanya K N, [2]Dr. Siddaraju

[1]Post Graduate Student,[2]HOD and Professor

[1,2] Department of Computer Science & Engineering,

[1,2] Dr. Ambedkar Institute of Technology, Bangalore, India

Abstract: *Unusual human activity detection in crowded scenes. Suspicious behavior is hazardous in public settings and can result in significant casualties. Systems that detect motion or pedestrians have been developed using video frame collection, however even in real time, those systems lack the intelligence to spot unusual activity. To quickly and effectively manage a scamper issue before any casualties occur, it is necessary to spot scamper situations in real time via video monitoring. The proposed system focuses on identifying suspicious activity and aims to provide a technique that makes use of computer vision to detect suspicious behavior automatically. The system in this instance makes use of the OpenCV library to instantly classify various activities. The examination of the motion regularly shifts the location between two locations has been represented by a motion influence map. Utilizing CNN and YOLO algorithm for real-time object detection.*

Index Terms - Unusual crowd activity detection, OpenCV, CNN, YOLO

I.INTRODUCTION

The field of visual/optical surveillance is well-known and has numerous applications in disciplines including reality, tracking human interest, automated event identification, public safety in locations such as banks, stores, and private areas, motion-based recognition, human counting, and others. Video surveillance has become increasingly popular and necessary in every industry as a result of the quick development of high-quality; low-cost recording equipment, supercomputers, and the rising need for the analysis of such footages. Though it's crucial, it's also highly vital to identify moving items and follow them from films. On the other hand, it can be exceedingly challenging but crucial to distinguish items from their surroundings. Understanding the video's content and the context of the items is crucial for this reason. Background items from other scenes become a serious issue. As a result, the main standard is now comprehension of the movie, its elements, and the circumstances that are shown in it. Using a common human behavior method, predictable intention is used to identify unanticipated activity. The technique is initially created in consideration of a standard dataset for some activity. The pattern is then contrasted with actual data and facts as part of the verification process. Ultimately, it is decided whether the behavior is anticipated or not. Due to the necessity of defining regular human activity, it is difficult to detect unexpected behavior in real-world security systems.

Figure 1: Local Common Activity

Figure 2: Local Common Activity: Bicycle in the middle of the frame

Figure 3: Global Common Activity

Figure 4: Global Common Activity: People Crossing Frame

II.LITERATURE SURVEY

[1].SoumalyaSen, MoloyDhar and Surat Banerjee, "Human Action Recognition Implementation Using Image Parsing Techniques", EDCT 2018, an IEEE conference on Computational methods and Emerging Electronic Device Trends. Human activity must be recognized for interpersonal relationships and human-to-human communication to function properly. It contains details about an individual's identity, personality, and mental health, making it difficult to extract. How humans perceive activity is one of his key topics in computer vision and ML research *[2]* Zakie the authors of this study are Hammal1, Wen-Sheng Chu1, Jeffrey F. Cohn1, 2, Carrie Heike3, and Matthew L. Speltz4. "Automatic Action Unit Detection in Infants Using Convolutional Neural Network," IEEE's 2017 the seventh World Conference on Intelligent Communication and Intelligent Systems will be held in Beijing, China (ACII).Detecting action units in newborns are more difficult than in adults. Rapid and unique facial motions are frequent, and the jaw contour and facial texture are diminished. We suggest a multi-label approach to identify facial action units in infants' spontaneous (CNN) *[3]* He Xu1,2, Chengcheng Yuan1, Peng Li1,2, Yizhuo Wang3 were the authors of "Design and implementation of an action recognition system based on RFID.".,sensors"13th International Conference on Natural Computation, Fuzzy Systems, and Knowledge Discovery" (IEEE ICNC-FSKD 2017).Robots, smart homes, video games, virtual training, and human computer interaction all make extensive use of human action recognition. However, the bulk of action identification methods rely on visual action recognition technology, whose data is difficult to manage in a big data environment and is expanding swiftly. In this study, we use (RFID) technology to design a player system, the use of an RFID sensor, and a model of how humans move to the left, right, front, and rear. [4]"PeMapNet: Action Recognition from Depth Videos Using Pyramid Energy Maps on Neural Networks," Jiahao Li, Hejun Wu, and Xinrui Zhou.," IEEE, 2017.We suggest a comprehensive method for identifying human action in a depth video. The linked deep learning neural network topologies and a brand-new feature descriptor for depth movies are the two main contributions of this methodology. The first concept we offer in this paper is pyramid energy Maps (PeMaps), a term used to describe a depth movie's succession of frames. [5]. Ling Guan, Yifeng He, and "Knowledge Fused for Human Activity Recognition Utilizing Biset/Multiset Globality Local area Retaining Canonical Pearson Correlation," Nour El Din Elmadany. TRANSACTIONS ON IMAGE PROCESSING, IEEE, 2018.In this article, we investigate the issue of recognizing human actions when each action is recorded by a number of sensors and represented by a number of multisets. For fusing the data from multisets, we provide two brand-new information fusion algorithms. Biset globality locality preserving canonical correlation analysis is the first method. The shared feature subspace between two sentences is identified (BGLPCCA). A second method (MGLPCCA) is designed to address extra than two units and combines regular correlation analysis with multiset globality and locality maintenance. By way of studying alow-dimensional not unusual subspace, the strategies presented bymeansof those two preserve the neighborhoodand global structure of the statistics sample.

III. PROBLEM STATEMENT

Detection and Analysis of unusual crowd activity in surroundings for personal security at localities like shopping centers, airports, bus-stands.

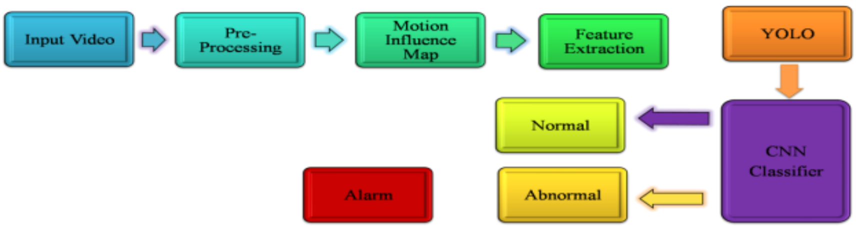

Figure 5: Overview of the proposed system

IV.METHODOLOGY

The project's primary goal is to identify any unexpected activity that occurs anywhere. The proposed project design, a video is given as an input to the device in which the video clips are converted into multiple frame sequence. From the frame sequence, feature extraction takes place where the objects present in the frame are been taken into consideration.

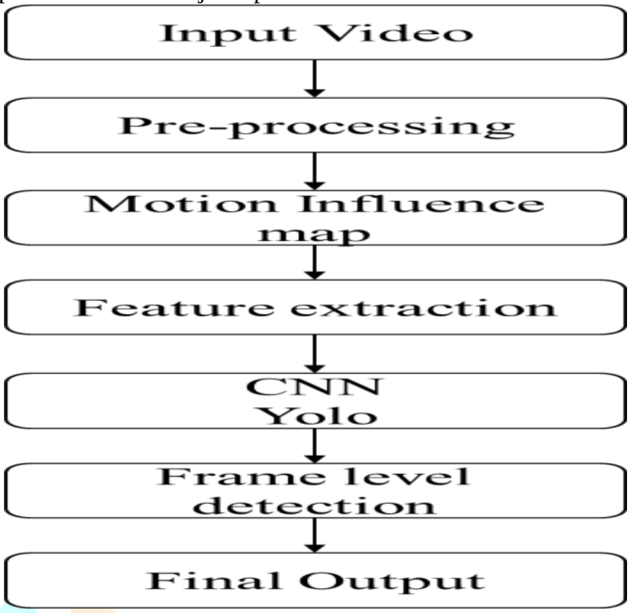

Figure 6: Proposed Methodology

The motion influence map detects any sudden movement or any unusual change from the frame sequence. The suggested work makes use of CNN to spot human behavior in a group and assess whether it's typical or remarkable. To categories activities from the dataset, we will create the deep learning system. When a photograph is entered into the system, it is assigned to one of several categories (common and unusual).The output is the action we take which maybe a siren, alarm, sending emergency call to the police or anything of our choice to take necessary action on the situation.
The code is broken up into 5 modules: training, testing, motion influence generator, choose block flow, and creating mega blocks. For the purpose of identifying and locating uncommon actions in a crowded scene, a method for describing motion characteristics is presented in this section. Here, it's important to remember that we looked at two categories of odd behavior:

1) Local 2) international, within a somewhat small space, strange local activities take place. A part of the frame may contain various motion patterns, such as the unusual look of inhuman things or a person moving quickly while the majority of the other pedestrians are moving slowly. When each traveler in a picture starts running frantically to get out of the scene, for instance, this is an example of a global uncommon activity that occurs across the frame.

1) Data Input and Pre-Processing: The system receives the video file as an input and pre-processes it. A video is processed as a series of frames, which are then, processed one after the other. First, grayscale is applied to an RGB frame. An image that has been reduced to grey scale contains only the image's intensity data, not its visible colors. Gray scaled vector is one dimensional, while RGB vector is three dimensional.

2) Optical Flow: Following the pre-processing stage, the FarneBack the method is used to compute optical-flow for each image sequence, for every image in a frame. Flow is the trend of seemly objects moving, substrates, as well as corners in a video environment induced by the viewer's relative movement inside the scene. Optical flow is a vector with the form (r,), where r stands for each pixel's magnitude and for its direction of motion in With respect to the associated pixel in the previous frames. Using Gunnar Farneback's approach, the openCV function calcOpticalFlowFarneback() computes a dense optical flow.

3) Block optical flow: Block division: Without sacrificing generality, we divide a frame into M by N uniform blocks (C) following the computation of optical fluxes for each pixel in a frame the blocks are indexed using B1, B2, and BMN.

3.1) Calculating Optical-Flow of each block: We calculate the optical flow of each block after separating the frames into blocks by averaging the optical flows of all the pixels inside a block. Formula for figuring out a block's optical flow. Here, bi stands for the optical flow of the ith block, J stands for the block's pixel count, and f ji stands for the optical flow of the ith block's jth pixel. A block's optical-flow is a vector (r,) that shows how far and in which direction each block has moved in comparison to the corresponding block in previous frames.

4) Motion Influence Map: Various elements, such as obstructions in the way, surrounding pedestrians, and rolling carts, might affect the direction in which a pedestrian moves among a crowd. This interaction trait is referred to as the motion influence. We suppose that two factors will decide the blocks that a moving object can influence:

1. The motion direction

2. The motion speed.

The more adjacent blocks that are impacted by an object depend on how quickly it moves. Blocks that are close by have more of an impact than those that are far away...

4.1 Algorithm for creating a motion influence map

INPUT: $B \leftarrow$ motion vector set, $S \leftarrow$ block size, $K \leftarrow$ a set of blocks in a frame
OUTPUT: $H \leftarrow$ motion influence map
$H^j (j \in K)$ is set to zero at the beginning of each frame
for all $i \in K$ **do**
  $T_d = \|\mathbf{b}_i\| \times S$;
  $\frac{F_i}{2} = \angle \mathbf{b}_i + \frac{\pi}{2}$;
  $-\frac{F_i}{2} = \angle \mathbf{b}_i - \frac{\pi}{2}$;
  for all $j \in K$ **do**
    if $i \neq j$ **then**
      Calculate the Euclidean distance $D(i, j)$ between $\mathbf{b}_i$ and $\mathbf{b}_j$
      if $D(i, j) < T_d$ **then**
        Calculate the angle ϕ_{ij} between $\mathbf{b}_i$ and $\mathbf{b}_j$
        if $-\frac{F_i}{2} < \phi_{ij} < \frac{F_i}{2}$ **then**
          $H^j(\angle \mathbf{b}_i) = H^j(\angle \mathbf{b}_i) + \exp\left(-\frac{D(i,j)}{\|\mathbf{b}_i\|}\right)$
        end if
      end if
    end if
  end for
end for

5) Feature Extraction: A frame with an unusual activity, as well as its neighboring blocks, have unique movement impact feature vector. Additionally, because a single activity is recorded by a series of related frames, an extracted features is extracted from a cuboid composed of n frames spanning a latest frames.

5.1) Creating Mega blocks: Each of the non-overlapping mega blocks that make up a frame is composed of several motion influence blocks. The aggregate of the motion influence values of all the smaller blocks that make up a larger block is what determines a mega block's motion influence value.

5.2) Extracting Features : Following the subdivision of the most recent 't' frames into Mega blocks, an 8 t-dimensional concatenated feature vector across all frames for each Mega block is retrieved. To create a concatenated feature vector for block, for example, we concatenate the feature vectors of giant block (1, 1) from all frames ('t' number of frames) (1, 1).

6) Clustering: We do using the spatiotemporal properties to cluster each mega block and using the centers as code words. As a result, the (i, j) the mega block has K code words, where w (i, j) k K k=1. We ought to point out that at this stage of instruction; we solely use videos of everyday events. As a result, the mega block's code words simulate the kind of routine activities that might take place in the area.

7) Testing phase: After creating the terms for common activities, it is now time to assess the model. That was created using a test dataset that includes unexpected behaviors...

7.1) Minimum Distance Matrix: In terms of removing the spatial and temporal feature vectors for each mega block, the minimum Euclidean distance between a feature vector from the current test frame and the code words in the corresponding mega block determines the value of an element in the testing state's minimum distance matrix over the mega blocks.

7.2) Unusual activity detection at the frame level: The lower the value of an element in a minimum-distance matrix, the less likely it is that an unusual action will take place in that block. On the other hand, if the minimum-distance matrix contains a greater value, we can infer that there are uncommon activities in to successive frames. As a result, we conclude that the frame representative characteristic relates to the lowest matrix's highest value. If the maximum value of the minimum distance matrix exceeds the threshold, the frame is classified as unusual.

7.3) Unusual activity detection at the pixel level: When an odd frame is found, the minimum distance matrix of each mega block is compared to the threshold value. We label a block as odd if the value exceeds the threshold.

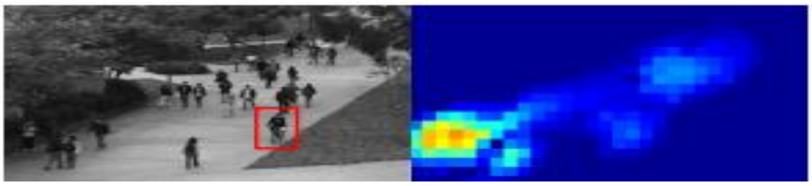

Figure 7: An example of a failure case in anomaly detection.
A frame from the original (left) and its motion influence map (right).

8) Yolo for object detection: A computer vision task called object detection entails both localizing one or more things inside an image and classifying each object that is there. This stands for the phrase "You Only Look Once," and it is an acronym. The above set of

regulations identifies as well as locates various elements in a shot (in real-time) (in real-time). The object identification technique in YOLO, which is considered as a regression issue, provides the class probabilities of the determined images. This demonstrates that a single prediction algorithm run is performed on the entire image. Several bounding boxes and class probabilities are forecast simultaneously using this method...

V.TESTING

GUI layout test case:

Test Case	1
Name of Test	GUI format
Input	Resolution(width, height)
Predicted output	Resolution-based window selection
Real output	Window is shown
Result	Success

Figure 8: GUI Layout

Input button's Test case:

Test Case	1
Name of Test	Input button
Input	The quantity and location of the buttons in the window (column, row)
Predicted output	Display button's in column, row
Real Output	Buttons appear in the designated location.
Result	Success

Figure 9:Input Button's

Video frame GUI Test case:

Test Case	1
Name of Test	Video frame GUI
Input	User clicked button
Predicted	Display video frame in GUI
Real output	Display video frame in GUI
Result	Success

Figure 10: Video frame GUI

Quit button Test case:

Test Case	1
Name of Test	Quit button
Input	User input
Predicted output	To kill or exit GUI window
Real output	Window exited when user pressed exit button
Result	Success

Figure 11: Quit Button

VI.RESULTS

By using Open CV and deep learning, the proposed device is capable of accurately detecting the irregular human behaviour of the crowd, if the human behaviour is unusual automatically the alarm will turn on. Figure 12 shows the user interface ,figure 13 shows the live mode detection, in this the camera will turn on and detect weather the person is still or moving and shows along with the id, in figure 14 detects the unusual activity in the area and also gives the counter number in this it will create one line then if the person will cross that line then it will mark as one counter then last figure shows the detection of unusual activity in the crowded scene if any unusual activity will occur automatically the alarm will turn on.

User Interface:

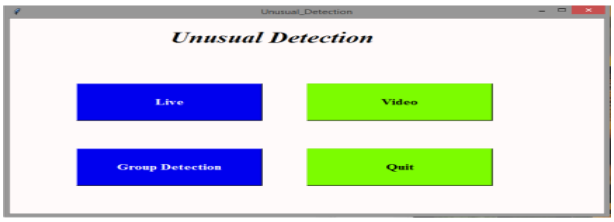

Figure 12: User Interface

Live mode:

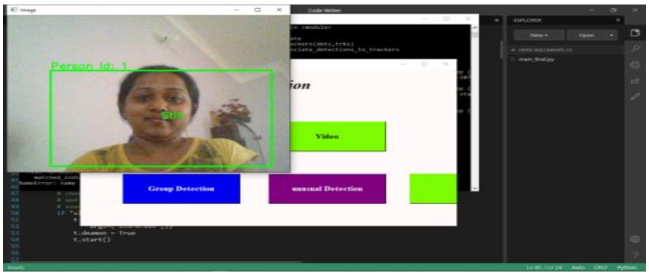

Figure 13: Live mode

Video detection with counter zero:

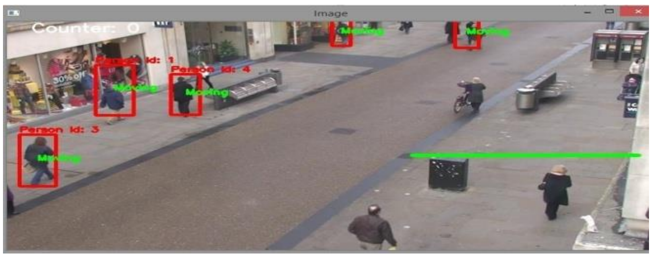

Figure 14:Video detection with counter zero

Group crowd detection:

Figure 15: Group crowd detection

VII.CONCLUSIONS

Although the structures that have been suggested thus far are intended to understand basic human behavior's like walking, running, and many other behavior's, crowded situations are not ideal for them. Using a deep learning model, the suggested device can identify unusual crowd behavior and person behavior. Less testing has been done on this theory, yet it has a slightly higher accuracy rating than other theories.

Previous Crime Appraisal can be handled using the proposed framework. The suggested device can accurately identify the erratic crowd behavior by utilizingOpenCV and deep learning, which greatly enhances the device's accuracy and competence. This theory has not been tested much but has slightly higher accuracy ratings than other theories. Prior offense assessments can be managed using the proposed framework. The proposed device can accurately identify abnormal crowd behavior using OpenCL and deep learning. This greatly improves device accuracy and proficiency. As the use of surveillance cameras in both private and public spaces grows, so does the demand for computer-assisted automated and intelligent video sequence analysis. The detection of unusual events or pastimes in crowded scenes has sparked There is currently a lot of interest in the topic of vision-based complete surveillance. Developed a technique for expressing motion in contrast. Properties in frames. We can classify frames as commonplace or anomalous and detect areas of anomalous interest within frames due to the expressiveness the space- and time-specific motion effect maps proposed. In real-world applications, an intelligent surveillance device should find both domestic and foreign aberrant activity using a single framework. The effectiveness of the proposed method and the use of the I. H. UMN and America datasets were established through experimental research on two public datasets. Other competing methods in the literature were outperformed by this method. Experiments on the Upturn dataset confirmed that the proposed method works in a variety of scenarios. Because the direction and speed of the moving object determine the motion influence map's creation. Object's motion, the proposed method has limitations when the input video has significant perspective distortion. Of course, some approaches have failed. Figure 7 shows an example of an error in which the proposed motion influence map detected false motion in the input video due to perspective distortion. On the bottom right, the bike (represented by a red box) is moving diagonally. In this case, our method cannot tell him apart from other pedestrians.

VIII. FUTURE RESEARCH

Currently, only the suggested strategy works with stationary cameras. Using the localization results, he can easily extend to PTZ cameras. We assume that there are no significant changes in this regard. In the scene, such as pans, tilts, or zooms, which could be a limitation of the process? Addressing these characteristics is a future research topic that will be pursued by expanding on the proposed method.

REFERENCES

[1].SoumalyaSen,MoloyDharandSusrutBanerjee,"Implementation ofHumanAction Recognition using Image Parsing Techniques", 2018 Emerging Trends in Electronic Devices and Computational Techniques (EDCT),IEEE.

[2].ZakiaHammal1,WenShengChu1,JeffreyF.Cohn1,2,CarrieHeike3,andMatthewL.Speltz4,"AutomaticActionUnitDetectioninInfantsUsingConvolutionalNeuralNetwork",2017SeventhInternationalConferenceonAffectiveComputingandIntelligentInteraction(ACII),IEEE.

[3].HeXu1,2,ChengchengYuan1,PengLi1,2,YizhuoWang3,"DesignandImplementation of Action Recognition System Based on RFID Sensor", 2017 13thInternational Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery(ICNC-FSKD2017),IEEE.

[4].Jiahao Li†, Hejun Wu∗ and Xinrui Zhou‡, "PeMapNet: Action Recognition from Depth Videos Using Pyramid Energy Maps on Neural Networks", 2017InternationalConference on Tools with Artificial Intelligence, IEEE.

[5].NourElDinElmadany,YifengHeandLingGuan,"InformationFusionforHumanAction Recognition via Biset/Multiset Globality Locality Preserving Canonical Correlation Analysis", IEEETRANSACTIONSONIMAGEPROCESSING,2018.

[6] "Machine Perception of Human Activity: An Investigation," P. Turaga, R. Chellappa, and V.S. Udrea, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 18, No. 11, 2008, pp. 1473-1488.

[7] "Networks of Dynamic Probabilistic Models for Analyzing Human Interactions," A.K. Jain, H.-I. Suk, and S-W Lee, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 21, No. 7, 2011, pp. 932-945.

[8] "Spatio-Temporal Relationship Match: Video Structure Comparison for Recognition of Complex Human Activities," M. S. Ryoo and J. K. Aggarwal, Proc. 12. IEEE International Conference on Computer Vision, Kyoto, Japan, September 29-October 2, 2009, S. 1593-1600.

[9] L. Kratz and K. Nishino, "Anomaly Detection in Highly Congested Scenes Using Spatiotemporal Motion Pattern Models," Proc. IEEE Conference on laptop vision and pattern recognition, Miami, USA, 20-25 June 2009, pp. 1446-1453.

[10]"Anomaly detection in crowded scenes," V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, Proc. IEEE conference on computer imaginative and prescient and pattern recognition, San Francisco, USA, 13.-18. Juni 2010, S. 1975-1981. [6] M. Ye, B. Wang, and X.

[11] "Detecting Anomalous Crowd Behavior Using High Frequency and Spatiotemporal Properties," B. Wang, M. Ye, X. Li, F. Zhao, and J. Ding, International Journal of Machine Vision and Applications, Vol. 23, No. 3, 2012, pp. 501-511.

[12] "Detecting Anomalous Crowd Behavior Using Size-matched Spatiotemporal Features," B. Wang, M. Ye, X. Li, and F. Zhao, International Journal of Control, Automation, and Systems, Vol. 9, Nr. 5, 2011,

[13]"Real Time Mass Kinetic Analysis," Ihaddadene and C. Djerba, Proc. 19th IEEE International Conference on Pattern Recognition, Florida, USA, 8-11 December 2008, pp. 1-4pp. "Anomaly Detection in Crowd Scenes," Wang and Z. Miao, Proc. 10th IEEE International Conference on Signal Processing, Beijing, China, 24-28 October 2010, pages 1220-1223..

[14]"Anomaly Detection in Crowd Scenes Using Historical Statistics," Proc. IEEE Global Symposium on Sensitive Signal Processing and Communication Systems, Chengdu, China, 6-8 December 2010, p. 1-4. S. Wang and Z. Miao, "Anomaly Detection in Crowd Scenes Using Historical Statistics," Proc. IEEE Global Symposium on Sensitive Signal Processing and Communication

[15] "Aberrant movement selection in bottom-up salient crowds," M. Mancus, N. Risch, J. Leroy, and B. Gosselin, Proc. 18th IEEE International Conference on Image Processing, Brussels, Belgium, 11-14 September 2011, pages 229-232.

[16] "Motion-based Abnormal Event Detection in Human Crowds," D. Y. Chen and P. C. Huang, Journal of Vision and Communication and Image Representation, Vol. 3, No.22, Nr. 2, 2011, S. 178-186. [13] "Floor Fields for Tracking in High-Density Crowd Scenes," S. Ali and M. Shah, Proc. 10th European Computer Vision Conference, Marseille, France, 12.-18. understood. 2008, S. 1-14.

[17] "Chaotic invariants of Lagrangian particle trajectories for anomaly detection in crowded scenes," S. Wu, B. Moore, and M. Shah, Proc. IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 13-18 June 2010, pp. 2054-2060.

[18]"Detecting Anomalous Crowd Behavior Using Social Force Models," RMehran, A. Oyama, and M. Shah, Proc. IEEE Conference on Computer Vision and Pattern Recognition, Miami, USA, 20-25 June 2009, pp.935-942.

[19] "Robust detection of anomalous events in real-time using multiple fixed monitors," A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 30, Issue 3, 2008, pp. 555-560.

[20]"Video Behavior Profiling for Anomaly Detection," T. Xiang and S. Gong, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 5, 2008, pp. 893-908.

[21] S. A. TsaftarisundF. JiangJ. Yuan "Detecting Anomalous Video Events Using Spatiotemporal Context," A. K. Katsaggelos, Computer Vision and Image Understanding, Vol. 3, No. 115, March 3, 2011 323-333. [22] "Iterative image registration techniques with application to stereo vision," B. D. Lucas and T. Kanade, Proc. 7th Joint International Conference on Artificial Intelligence, San Francisco, USA, 3.-9. August 1981, S. 674-679.

[23] "An Energy Model Approach to People Counting for Abnormal Crowd Behavior Detection," G. Zion, J. Cheng, X. Wu, Y. Chen, Y. Ou, and Y. Xu, Neurocomputing, Vol. 3, No.83, 2012, S. 121-135.

[24] B. Zhang, D.N. Monekosso, P. Remagnino, S. A. Velastin, and L. Q. XuInternational Journal of Machine Vision and ApplicationsVol. 19No. 5-62008pp. 345-357