Recognition of Human Unusual Activity in Surveillance Videos

Biswa Ranjan Acharya
School of Computer Engineering
KIIT University, Bhubaneswar
Odisha, India

Pradosh Kumar Gantayat
Associate Professor
DRIEMS, Cuttack, India

***Abstract*— Identification of human irregular activity is to a great degree vital for feature reconnaissance. As feature observation cameras get to be pervasive, there is a surge in studies on mechanized action understanding and abnormal occasional discovery in reconnaissance features. In any case, feature content investigation in broad daylight scenes remained a considerable test because of natural challenges, for example, serious between article impediment in swarming scene and low quality of recorded reconnaissance footage. Additionally, it is important to accomplish a strong discovery of irregular occasions, which are uncommon, uncertain, and effectively mistook for commotion. We outline a novel structure that incorporates object acknowledgment, movement estimation, and semantic-level acknowledgment for solid acknowledgment of various leveled human-object cooperation. The structure is therefore proposed to incorporate acknowledgment choices made by every part, and to probabilistically adjust for the disappointment of the segments with the utilization of the choices made by alternate segments. Therefore, human-object collaborations in an airplane terminal like environment, for example, 'a man is conveying a things', 'a man leaving his/her stuff', or 'a man grabbing another's things', are perceived.**

This paper proposes answers for determining uncertain visual perceptions and overcoming trickiness of habitual movement investigation routines be actualized calculations for perceiving abnormal exercises in a feature. The use has been made under visual studio utilizing OpenCV library. These composite managing human exercises like

i) persons shaking hands,
ii) ii) persons battling and
iii) iii) a man grabbing the sake of another.

In this work associated with a probabilistic model (HMM) for seeing human-human based activities. We have accepted static foundation for testing our usage.

***Keywords*— Unusual, surveillance, activity, recognition**

I. INTRODUCTION

Human eyes are very productive gadgets for looking over an expansive amount of low-level visual tactile information and conveying precise data to one's cerebrum for abnormal state semantic translation and increasing situational mindfulness. In the course of the most recent couple of decades, the PC vision group has tried endeavors to realize comparative perceptual capacities to manufacture visual sensors. Significant endeavors have been made in comprehension static pictures of individual articles and the comparing procedures in the human visual framework. This attempt is heightened further by the requirement for comprehension the enormous amount of feature information, with the mean to grasp different substances inside of a solitary picture as well as after some time over numerous feature outlines for comprehension their spatiotemporal relations. Noteworthy use of feature examination and comprehension is clever reconnaissance, which means to naturally translate human movement and identify unoriginal occasions that could represent a risk to open security and wellbeing.

A. Unusual Activity Understanding & Recognition

There has been a quickened extension of Shut Circuit TV (CCTV) reconnaissance as of late, to a great extent in light of rising nerves of wrongdoing and its risk to security and wellbeing [1]. Sizable quantities of observation cameras have been conveyed in broad daylight spaces running from transport buses (e.g. air terminals, underground stations), malls, sport enclosures to private avenues, serving as an instrument for wrongdoing diminishing and dangerous administration. Traditional feature observation frameworks depend vigorously on human administrators to screen exercises and focus the moves to be made upon the event of an occurrence, e.g. following Unordinary focus starting with one camera then onto the next camera or alarming influential organizations to ranges of concern.

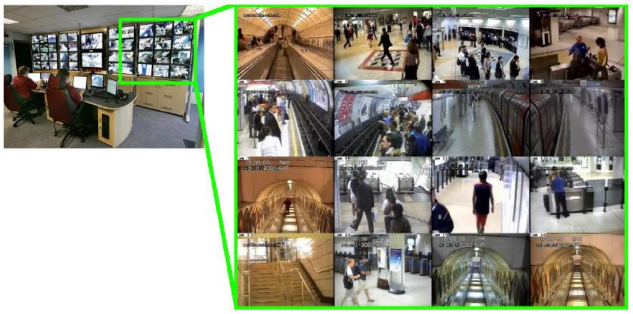

Figure 1.1: A typical CCTV control room monitoring multiple camera views. A report by Home Office [2] reveals that an operator may be required to oversee 90 cameras or more at one time. Consequently, surveillance systems are used primarily in a reactive mode, *i.e.* operators often rely on indication from external agencies to direct their surveillance, or view the recorded footage retrospectively.

B. Activity Understanding

Diversity terms including activity and occasion are utilized as a part of the writing when alluding to action. To disambiguate these terms, a uniform wording is suggested, which will be used reliably all through the rest of this theory. This theory takes after comparable scientific categorizations proposed by Xiang & Gong [2], Poppe [3], and Moeslund et al. [4]. Specifically, activity points to a succession of primitive developments completed by a

solitary article, for example, human strolling or vehicle turning right. Movement, then again, contains a progression of successive activities, in all probability including numerous items that communicate or exist together in a regular mutual space checked by single or different cameras. Illustrations of action incorporate 'travelers are strolling on a train stage and taking a seat on a seat' and 'vehicle halting subsequent to turn comfortable movement crossing point'.

C. Unusual Event

Strange occasion is a different term that has been bringing on much disarray in the writing. The idea of a strange occasion shows up in distinctive names including anomalous, uncommon, atypical, intriguing, unoriginal, amazing occasion, or essentially a peculiarity, variation from the norm, abnormality and exception. A few studies have endeavored to characterize or recognize these terms. Xiang and Gong [5] characterize an uncommon occasion as an atypical/unusual conduct design that is not spoke to by adequate specimens in a preparation dataset already fulfills limitation of irregular example. So also, Hamid et al. [6] allude unoriginal occasions to those that are uncommon and different from standard occurrences. Breitenstein et al's. Definition [7] of uncommon occasion is all the more obvious - just occasions that have not been seen before are viewed as unoriginal occasions; those that have been seen in any event once are viewed as uncommon however not so much Strange.

As can be seen, no agreement on characterizing the term 'Abnormal occasion' has been come to so far because of subjective perspectives of distinctive studies, yet there are shared traits. Specifically, an occasion is viewed as worth further examination by a human administrator or ought to trigger an alert if:

Unoriginal occasion is another term that has been bringing about impressive perplexity in the writing. The idea of a bizarre occasion shows up in diverse names including strange, uncommon, atypical, fascinating, unoriginal, shocking occasion, or just peculiarity, anomaly, inconsistency and exception. Numerous studies have endeavored to characterize or recognize these terms. Xiang and Gong [5] characterize a bizarre occasion as an atypical/irregular conduct design that is not spoke to by adequate examples in a preparation dataset however fulfills limitation of a strange example. Additionally, Hamid et al. [6] allude strange occasions to those that are uncommon and distinct from customary occurrences. Breitenstein et al's. definition [7] for an unoriginal occasion is all the more obvious - the main occasions that have not been seen before are viewed as surprising occasions; those that have been seen in any event once are viewed as uncommon however not so much Abnormal.

As can be seen, no accord on characterizing the term 'Abnormal occasion' has been come to so far due to subjective perspectives of diverse studies, yet there are shared traits. Specifically, an occasion is viewed as worth further examination by a human administrator or ought to trigger a caution if:

1. The occasion rarely happens or has not been seen for some time recently, i.e. having low measurable representation in a dataset.

2. The occasion is obscure or flighty.

Subsequently, this paper characterizes an 'Uncommon occasion' as an action that has these regular traits.

D. Objectives

The important objective of this work is to devise PC vision calculations for movement understanding and Abnormal occasional discovery in broad daylight scene reconnaissance features. Human exercises that these sorts of frameworks might want to perceive are not the elementary motions or activities of a solitary individual. Maybe, the objective of these frameworks is equivalent to investigate and report convoluting continuous human exercises where people and a few articles drawn from different classes to take an interest in exercises. Most human exercises out in the open spots include articles, and in this way a framework for the acknowledgment of abnormal state human-object communications is fundamental for developing robotics reconnaissance frameworks, savvy spaces, and human-PC collaboration frameworks out in the open. Case in point, on account of the reconnaissance framework for an air terminal environment, a framework needs to consider article data and in addition its developments to break down and we have built a probabilistic structure where the three principle segments (object acknowledgment, movement estimation, and semantic-level action acknowledgment) supplement one another to handle clamor and the vulnerabilities of inputs. In particular, the work is focused on the semantic layer, which perceives abnormal state human-object communications. We introduce a solid acknowledgment calculation that is in a position to adjust to protest acknowledgment or movement estimation slips. A calculation for recognizing the time interim of happening exercises and figuring the likelihood connected with that interim has been set up. Capacity to manage lapses builds the acknowledgment execution, as well as empowers the semantic layer to give criticism to alternate layers. Case in point, if the framework perceived a movement 'individual conveying his/her bag', then the article that took part in that action must be a bag. Here we have taken two issues:
(1) Tracking objects and
(2) Recognition of Suspicious events in crowded public scene.

E. Outline

This paper is organised into five sections:
Section 2 presents a review on various existing strategies and approaches for activity understanding and Unusual event detection, while providing further motivations for the proposed approaches of this paper.
Section 3 presents the mechanism of learning lessons from human feedback for on-line atypical event detection through

stream-based active learning. Here we have proposed framework to provide a solution to the existing problem.
Section 4 presents experiments are carried out to demonstrate the effectiveness of the proposed approach in balancing different active query criteria for joint unknown event discovery and decision boundary refinement, and how such an approach can lead to more robust and accurate detection of subtle Unusual events in public space.
Section 5 presents conclusions and experimented result of implemented approach.

II. RELATED WORK

As a rule, a framework for movement understanding and Uncommon occasional discovery in observation features includes a few key segments for preparing features:
1. Low-level components for background modelling, feature extraction, and object tracking.
2. Middle-level components for object and action description, e.g. object classification and action classification.
3. High-level components for semantic interpretation, e.g. activity understanding and suspicious event detection.
Extensive exertion has been spent and critical upgrades have been recorded in every segment throughout the years [1-3]. Get the wide and growing nature of this exploration zone, this survey limits down the degree by concentrating just on particular procedures inside of low and abnormal state segments that are generally embraced in visual reconnaissance frameworks. Some particular systems, for example, foundation subtraction and item following are not portrayed altogether in this survey. An overview is formed by Hu et al. [8] gives definite portrayals on each of the previously stated strategies. Whilst Turaga et al. [9] and Lavee et al. [10] condense best in class methodologies concentrating on displaying routines for comprehension activity and exercises in features, Dee and Velastin [11] highlight the difficulties in visual observation and present a rundown of diverse techniques created for visual reconnaissance. For different camera movement examination, a complete treatment of the most ordinarily utilized systems can be found as part of Javed and Shah [12], and in addition in Aghajan and Cavallaro [13]. For a general survey on dynamic learning methodology past visual observation, Settles [1] provides a complete scope on ordinary procedures and the latest improvements in the space.

A. Activity Representation

Before semantic understanding of action can be constructed up, the subject of action representation or deliberation [14] must be tended to. It is identified with the extraction, determination, and change of low-level visual properties in feature to build transitional info to a movement model. As referred to by Hongeng et al. [15] and Lavee et al. [14], movement representation ought to be sufficiently expressive to depict a mixture of exercises yet adequately discriminative in recognizing distinctive individual exercises. The determination of action representation is managed by a few variables, for example, feature casing rate, separation of items from the camera, and scene crowdedness. In the accompanying segments, the methods for selecting a fitting action representation are talked about. Points of interest and shortcomings of different representations regularly utilized for movement comprehension are likewise highlighted, including article based components, pixel-based elements, and other element representations.

B. Object Tracking

There are for two noteworthy parts of a visual following framework; Target Representation and Confinement and Separating and Information Affiliation. Target Representation and Restriction are meant for the most part a base up procedure. Ordinarily computational many-sided quality of these calculations is poor. Followings are some normal Target Representation and Confinement calculations:

• Blob Tracking: Division of item inside (blob identification, piece based connection)

• Kernel-based Tracking (Mean-shift Tracking): An iterative restriction methodology in light of the expansion of a closeness measure (Bhattacharyya coefficient).

• Contour Tracking: Location of object limit(boundary).

Separating and Information Affiliation is basically a top-down procedure, which includes joining former data about the scene or item, managing article motion, and assessment of different speculations. The computational many-sided quality of these calculations is usually much higher. Followings are some basic Sifting and Information Affiliation calculations:

Kalman Filter: An ideal recursive Bayesian channel for straight capacities subjected to Gaussian commotion.

Particle Filter: Helpful for examining the fundamental state-space conveyance of non-straight and non-Gaussian forms.

Here we utilize a Kalman Filter for tracking. This is the ideal Metho.

III. METHODOLOGY

For solid acknowledgment of human-item associations, we composed a structure made up of four layers: the segmentation layer, the object layer, the motion layer, and the semantic layer. Each of these four layers gets its functionalities. Part of the Segmentation layer comprises of portion and track protests in the scene utilizing pixel-level and blob-level preparing. The object layer recognizes classes of sectioned items, while the motion layer appraisals developments of object. The after effects of the object layer and the motion layer are given to the semantic layer, which makes utilization of identified items and their movement so as to perceive the last abnormal state action. In the semantic

layer, the exercises are perceived progressively from the most basic exercises, i.e. nuclear activities, to composite human-object communications.

.

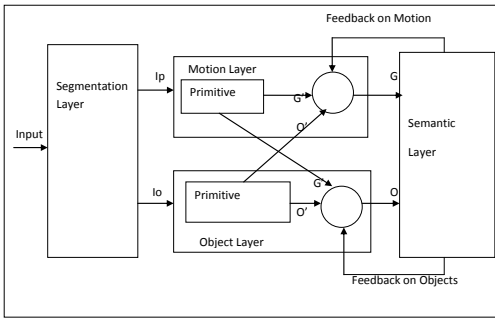

Figure 1.2: Details of the framework for recognition of high-level human-object interactions.

A. Segmentation Layer

Calculations for foundation subtraction, blob location, and blob following are used in our division layer. Our division layer fundamentally portions a group of isolated blobs as one item. Matter what it may, when a man is conveying protests, for example, bags or boxes, item blobs and human blobs frame one extensive bunch as opposed to two isolated groups. So as to fragment protests, for example, bags or boxes from the individual who is conveying it, a blob-based rendition of Haritaoglu et al's. Calculation [17] is used. Their calculation utilizes symmetry and periodicity data of individuals to portion objects from people. As a consequence of the division layer, the framework appraisals positions, shapes, and developments of the items in the scene.

1) PRE-PROCESSING

In all PC vision frameworks, straightforward transient and/or spatial smoothing is used as a part of the introductory phase of handling to decrease camera commotion. Smoothing can be utilized to evaluate transient natural commotion. For example, rain and snow caught in open air camera. For ongoing frameworks, casing size and casing rate diminishing is ordinarily used to decrease the information handling rate. On the off chance that the camera is moving or various cameras are worked as distinctive areas, picture enlistment between progressive edges or among diverse cameras is required before foundation displaying. Another key issue in preprocessing is the information organization utilized by the definite foundation subtraction calculation. The greater part of the calculations handles luminance power, which is one scalar quality for every pixel. Then again, shading picture, in either RGB or HSV shading space, is proving to be more pervasive out of sight subtraction writing. These papers contend that shading is superior to anything luminance at distinguishing questions in low-differentiate ranges and stifling shadow cast by moving articles. Notwithstanding shading, pixel-based picture components, for example, spatial and transient subsidiaries are once in a while utilized to consolidate edges and movement data. For instance, power values and spatial subsidiaries can be united to shape a solitary state space for the foundation following the Kalman channel. Pless et al. join both spatial and worldly subordinates to frame a steady speed foundation model for distinguishing speeding vehicles [18]. The principal downside of including shading or determined components in foundation displaying is the additional many-sided quality of model parameter estimation. The increment in many-sided quality is frequently decisive as most foundation demonstrating strategies keep up a target model for each pixel. For closer view identification system past scientists have utilized foundation subtraction [17].

B. Object Layer And Motion Layer

The object layer and the motion layer are intended to exploit choices made by one another. Our expectation is to make the acknowledgment procedure of items and movements more solid by considering the relationship in the middle of articles and their movements. In any case, this configuration guideline produces a cycle simultaneously. We need to know the yield of the object layer to perceive the movements. The acknowledgment procedure of items needs yields of the movement layer. The framework developed by Moore et al. [9] maintains a strategic distance from this cycle between the object layer and the motion layer by giving a need to the choice of the article layer to that of the movement layer: As a rule, articles are chosen first and movements are evaluated in light of the perceived item. Just when the article layer neglected to recognize an item, the motion layer has the capacity help the article layer. There is no cycle in this procedure, since the framework dependably tries to perceive the item first. The frame work expect that the object layer either effectively perceives an article or names it as an indistinct article. Our framework is intended to conquer this cycle of the procedure by developing a primitive article module and movement module inside the object layer and the movement layer. The primitive modules are fundamental classifiers that settle on a choice exclusively in light of visual perceptions, without yields from one another. Two primitive modules are autonomous. The object layer and the motion layer keep away from the cycle of the procedure by treating the primitive module of one another as an estimation of choice of one another. Any of the already created article acknowledgment and movements estimation methods can be securely embraced for every module.

Primitive article module: We utilize a k-closest neighbor (k-NN) classifier to perceive objects. The classifier uses six elements that have been utilized usually for the article characterization. Territory, tallness, width, edge of significant pivot, smallness, and mean shading is the components used to order objects.

Primitive movement module: A shrouded Markov model (Gee) is built for every movement. Gee have been broadly utilized for signal acknowledgment [22]. A Gee treats elements extricated from every article, for example, 'change of the focal point of mass' for our situation, as "perceptions" produced by the shrouded hubs of the model. Movement is distinguished at nearby maxima of the likelihood of a Well created current arrangement of perceptions. A forward

calculation of Gee is utilized to recognize the consummation time of a movement, and a regressive calculation is utilized to identify the beginning time. A credulous Bayesian classifier is built for every layer to settle on a definite choice. The article layer and the motion layerutilize one another's primitive module to exploit one another. The motion layer appraisals article movements in view of its primitive module, the primitive item module and, criticism from the semantic layer. The article layer groups items utilizing its primitive module, criticism from the semantic layer, and the whole history of movements assessed by the primitive movement module which is shown in the above figure 1.

1) CONNECTED COMPONENT LABELING

Connected component labeling (alternatively connected component analysis) [23] is an algorithmic application of graph theory, where subsets of connected components are uniquely labeled based on a given heuristic.

- This paper is working on binary digital images, which simplify our work.
- It would comprise of a grid of pixels.
- Each pixel represents two things Colour of a pixel and its intensity.
- So I have got a graph, in which each vertex is a pixel and edges represents the neighbours of vertex.

Goal is to find all disconnected objects in a video with respect to certain frame (reference frame extracted from videos).
Algorithm
Two pass algorithm: [23, 24]
Many algorithms in this group operate in three distinct phases.
(a) Scanning phase: In this phase, the image is scanned once to assign provisional labels to object pixels, and to record the equivalence information among provisional labels.
(b) Analysis phase: This phase analyzes the label equivalence information to determine the final labels.
(c) Labeling phase: This third phase assigns final labels for object pixels using a second pass through the image.

C. Semantic Layer

In the semantic layer, the framework keeps up representations of human exercises which are encoded in view of their semantic structure. The shrouded Markov models introduced in the past area have the capacity to perceive signals, for example, 'arm extending' and 'arm withdrawing', however Gee themselves are not adequate to perceive abnormal state cooperations like "pushing" and 'hand shaking'. Bearing in mind the end goal to perceive abnormal state exercises in light of signal location, we take a depiction based approach in our semantic Layer. That is, our framework keeps up the representation of an action depicting how motions must be connected transiently, spatially, and coherently with a specific end goal to shape the action, and exploits them for the acknowledgment. In this area, we acquaint the fundamental ideas vital with speak to movement structures, and present formal representation language structure to depict exercises utilizing them. The acknowledgment calculation utilizing the built movement representations will be spoken of in the following area. Our representation methodology is a diverse leveled methodology, which decays an action into a few more straightforward exercises called 'sub-occasions' and depicts the essential connections among the sub-occasions. We first introduce the idea of 'time interims', which are connected with exercises and their sub-occasions to speak to their beginning and consummation times. Next, we present the "predicates" used to formally portray worldly relations between time interims, spatial relations in the middle of persons, and sensible relations connecting with different relations. At long last, programming dialect like representation language structure is given, which permits the portrayal of structures of exercises utilizing time interims and predicates.

IV. SIMULATION

In the simulation part, we have made HMM of each activity as defined in the problem. My problem activities are reproduced below.
1. Person snatching bag of another person
2. Multiple persons are fighting
3. Two persons are shaking hands
First of it set a threshold value that comes from motion and object recognition layer if the distance between 2 are more than 2 objects are less than the threshold value den that is my region of activity and we will made HMM of each activity in which we have chosen a series of frames as input and series of desired frames and calculated the probability. Best match probability will be the probability of activity and using probability I will find my activity name. This threshold value will be changed on changing the environment that means changing the distance between camera to a place of activity and it will be set using no. of observations.

A. RESULT & ANALYSIS

We have tested this software on our own dataset and some standard datasets. One snapshot is attached here.

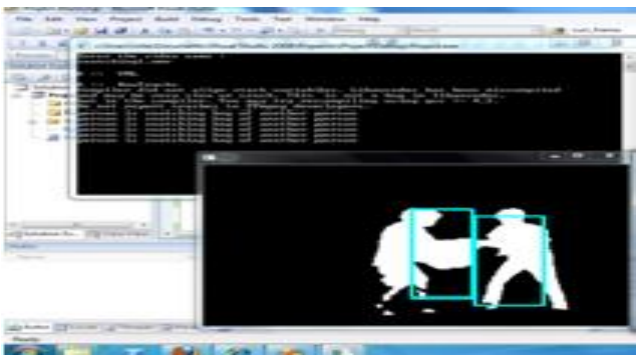

Figure 1.3: Snapshot of Person is snatching bag of another person

In this snapshot when two persons are approaching to each other and the distance between the two persons is less the threshold value and their states probabilities are matching with the HMM of Person is snatching the bag of another person. Then the message on the command prompt

"person is snatching the bag of another person" will be printed.

V. FUTURE WORK

Initially we have completed high level human-human based Unusual activity recognition in a static background and some limited no. of activities but real world are not static in the nature and activities are not limited. So this can be developed in the dynamic background (like hog descriptor or any other algorithm) and increase the human-human based activities. In the given method when the environment changes then we have to train HMM again and again, this causes an inconvenience, so this can be explored that is not training dependent.

This examination work will lay a going stone for further improvements of the programmed vehicle acknowledgment framework in a secured range. Among such expansions, apply particular changes, dynamic and versatile foundation demonstrating, a tree based progressive classifier could be analyzed in future work. More shape speaking of elements could be considered for better acknowledgment exactness for somewhat blocked items.

REFERENCES

[1]. C. Norris, M. McCahill, and D. Wood. The growth of CCTV: a global perspective on the international diffusion of video surveillance in publicly accessible space. *Surveillance and Society*, 2(2/3):110–135, 2004.

[2]. T. Xiang and S. Gong. Beyond tracking: modelling activity and understanding behaviour. *International Journal of Computer Vision*, 67(1):21–51, 2006.

[3]. R. Poppe. A survey on vision-based human action recognition. *Image and Vision Computing*, 28(6):976–990, 2010.

[4]. T. B. Moeslund, A. Hilton, and V. Kr˘uger. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2):90–126, 2006.

[5]. T. Xiang and S. Gong. Video behaviour profiling for anomaly detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5):893–908, 2008.

[6]. R. Hamid, A. Johnson, S. Batta, A. Bobick, C. Isbell, and G. Coleman. Detection and explanation of anomalous activities - representing activities as bags of event n-grams. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1031–1038, 2005.

[7]. M. D. Breitenstein, H. Grabner, and L. V. Gool. Hunting Nessie - real-time abnormality detection from webcams. In *IEEE International Workshop on Visual Surveillance*, 2009.

[8]. S. Ali and M. Shah. A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6, 2007.

[9]. M. W. Green. The appropriate and effective use of security technologies in U.S. schools. Technical Report NCJ 178265, Sandia National Laboratories, 1999.

[10]. S. Wu, B. E. Moore, and M. Shah. Chaotic invariants of Lagrangian particle trajectories for anomaly detection in crowded scenes. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2054–2060, 2010.

[11]. I. Everts, N. Sebe, and G. Jones. Cooperative object tracking with multiple PTZ cameras. In *International Conference on Image Analysis and Processing*, pages 323–330, 2007.

[12]. O. Javed, K. Shafique, and M. Shah. Appearance modeling for tracking in multiple nonoverlapping cameras. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 26–33, 2005.

[13]. G. A. Miller. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2):81–97, Mar 1956.

[14]. P. M. Gill, A. Spriggs, J. Allen, M. Hemming, P. Jessiman, D. Kara, J. Kilworth, R. Little, and D. Swain. Control room operation: findings from control room observations. Home office online report 14/05, Home Office, 2005.

[15]. M. J. V. Wert, T. S. Horowitz, and J. M. Wolfe. Even in correctable search, some types of rare targets are frequently missed. *Attention, Perception, & Psychophysics*, 71:541–553, 2009.

[16]. J. M. Wolfe, T. S. Horowitz, M. J. Van Wert, N. M. Kenner, S. S. Place, and N. Kibbi. Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General*, 136(4):623–638, 2007.

[17]. G. A. Miller. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2):81–97, Mar 1956.

[18]. T. Xiang and S. Gong. Video behaviour profiling for anomaly detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5):893–908, 2008.

[19]. M. D. Breitenstein, H. Grabner, and L. V. Gool. Hunting Nessie - real-time abnormality detection from webcams. In *IEEE International Workshop on Visual Surveillance*, 2009.

[20]. M. Bar. Visual objects in context. *Nature Reviews: Neuroscience*, 5:617–629, 2004.

[21]. J. Kim and K. Grauman. Observe locally, infer globally: a space-time MRF for detecting abnormal activities with incremental updates. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2921–2928, 2009.

[22]. V. Gouaillier and A.-E. Fleurant. Intelligent video surveillance: Promises and challenges. Technological and commercial intelligence report, CRIM and Techn˜opole Defence and Security, 2009

[23]. G. A. Miller. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2):81–97, Mar 1956.

[24]. M. J. V. Wert, T. S. Horowitz, and J. M. Wolfe. Even in correctable search, some types of rare targets are frequently missed. *Attention, Perception, & Psychophysics*, 71:541–553, 2009.

[25]. W. Hu, T. Tan, L. Wang, and S. Maybank. A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, And Cybernetics*, 34(3):334– 52, 2004.