*2009 International Conference on Machine Learning and Computing*
*IPCSIT vol.3 (2011) © (2011) IACSIT Press, Singapore*

# A Supervisory System to Detect Suspicious Behavior in Online Testing System

Syed Ahmar Qamar [1], M. Arfan Jaffar [2], Hafiz Adnan Habib [3]

[1,3] Department of Computer Engineering, University of Engineering & Technology, Taxila, Pakistan

[2] National University of Computer & Emerging Sciences, FAST, Islamabad

ahmar.qamar@nu.edu.pk, arfan.jaffar@nu.edu.pk, adnan@uettaxila.edu.pk

**Abstract.** Computer vision is one of the most emerging fields in Information Technology. We propose a unique application of computer vision in this paper. In the proposed method, we have detected the suspicious behaviour of a student in an online or computer based testing system. The work is valuable especially for test centers, schools, colleges and universities. For the development of such system, we present a three stage process to detect mysterious actions. The experiment has shown the impressive results for warning generation and activity registration on successful detection. The most occurring inexplicable actions are listed: arm movement and flipping, head shaking dual sides, wobbling around premises, revolving side-by-side or back to back.

**Keywords:** behaviour registration, online testing, feature monitoring

## 1. Introduction

The behaviour of human is being categorized into number of applications. The detection of different type of behaviours varies case to case .e.g. recognizing the symptoms of patient's disease includes the abnormal behaviour or in recreational places, the detection of sudden fall down of a person addresses abnormality and from security point of view i.e. punching, kicking, gun-pointing especially in banks are the suspicious and mistrustful behaviour. The suspiciousness also counts on the detection of bags, sacks for a long period of time on platform, airport etc. In our case the detection of specific actions of a user declare as suspicious i.e. arm movement and flipping, head shaking dual sides, wobbling around premises, revolving side-by-side or back to back. The aim is to develop a system which acts as a supervisor to monitor the human previously listed acts. The standardized procedure for the admission in well-known universities requires the computer based tests and it will also assist for an online testing system e.g. GRE, Language Evaluation, and JCP etc. Normally such establishments involve user interaction with the computers but if anything happen other than normal, treat as suspicious by isolating those frames which genuinely counts it.

To make an artificially intelligent system especially in case of human motion detection, tracking, action, events have gained widespread interest in both academic and industrial research, with emphasis on real time system [1, 2].The applied significance for human monitoring and classification focuses on real time systems, a lot of work done on sensitive domains, whether surveillance or surgical or many more. As with the passage of time, technology is increasing too fast so its adherent us to propose and work in other domains or the need of further enhancements in already implemented systems.

The initial problem to cope with is the acquiring device adjustment. The camera position is dependent on case to case. It could be giving the best view of a client whether an Ariel or front to front. The gist is to detect the listed possible suspicious behaviour from prescribed view angle. The following actions are declared suspicious for the system: Arms Flipping, Dual side head movement, half upper body motion, Full back-back rotation, wobbling around premises.

In this paper, the motion vectors are drawn through which control points are produced. Those points refer to the interested region of an object expecting some behaviour and it's been marked. After marking, it will vary with that part movement. On the basis of those variances, the magnitude is calculated. The outcome magnitude is if above the defined threshold value then the warning will be generated to the system and activity is registered.

- The system is fully automated and do not rely on manual selection.
- It strongly works on real time applications e.g. behaviour detection for online testing system

In Section 2, the previous work is studied. Section 3 explains the proposed method and results.

## 2. Related Work

The behavioural classification into normal and abnormal changes relies on the use cases need. Normally abnormal behaviour initially has been classified for security and somehow scrutiny measures. There has been a number of video surveillance systems [3, 4] .All these systems can detect and track multiple people.

However, most systems are not designed for detecting abnormal behaviours as their major function. For the recognition of human action, different approaches were used as one is of Hidden-Markov Model [5] and representing activities using velocity vectors [6]. As well as for normal and abnormal behaviour, the blob analysis [7] was also built but the leading Ghallab [8] modeled temporal scenarios whose occurrences need to be recognized on-line. Mecocci [9] introduces architecture of automated real time video surveillance capable of detecting anomalous events. Such behaviours are detected whenever the observed trajectories deviate from the typical learned prototypes. The researchers then having a depth analysis and started working on suspicious behaviour other than an abnormal one. The methodologies were proposed as by Shoaib Khan [10]. The proposed scheme is on moment invariant based classifier by considering upon Hu moment invariants. The deviation in the attention drags the scholars to do research on other grounds like to analyze activities in streets, market, crowd areas and others. For this, the human monitoring system by Ismail Haritaoglu [11] in an outdoor environment. This employs the combination of shape analyses and tracking to locate people and their parts and to create model of people's appearances so that they can tracked through interactions.

The Jezekiel Ben-Arie [12] gives the same idea related to the aforementioned for human activity recognition using multi-dimension indexing. They developed a novel method to view-based recognition. The idea is to identify from sparsely sampled sequences of a few body poses acquired from video clips. The work of James W. Davis [13, 14, 15] highlights to analyze the human motion recognition problems more keenly. In his papers, the problems are interpreted in a very simple and explicit way. We can extend the techniques for this paper to propose the work or to build new systems. The creation of motion history image, hierarchical analyzes, motion gradients are depicted.

## 3. Proposed Method

### 3.1. Preprocessing

In the first module, the video has been recorded from the camera. The camera position must be adjusted which varies case to case but it presumes fulfil of our requirement. The field of view should include the half upper body portion for the complete detection in our case. This process will be iterative on every clip to easily cope with complexity problems.

### 3.2. Optimal Thresholding

The Thresholding technique [16] is to partition the image by using a single global threshold $T$. Any pixel with a gray level $\leq T$ is assigned as zero and any pixel value $> T$ is assigned 255. The objective is to generate a binary image. The iterative thresholding technique is used to obtain $T$ automatically.

1.Select an initial estimate for $T$

2.Segment the image using $T$ producing two regions of pixels $R0\ pixels < T$ and $R1\ pixels > T$.

3.Compute the average gray level values $\mu_0$ and $\mu_1$ for the pixels in region $R0$ and $R1$.

4.Computing a new threshold value :

$$T = \frac{1}{2}(\mu_0 + \mu_1) \quad (1)$$

5.Repeat step 2 through 4 until the difference in $T$ in successive iteration is smaller than a predefined parameter $T_0$

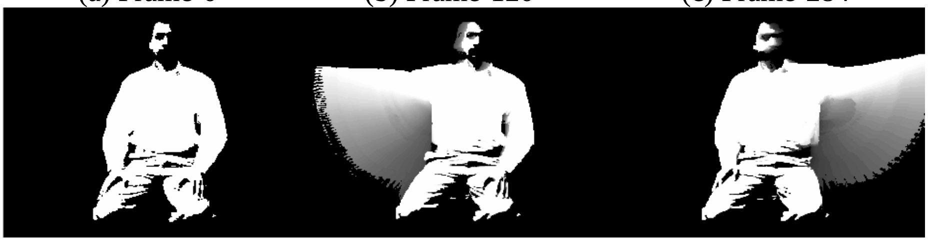
## 3.3. Template Maker Image

We used temporal methodology [13] for the history of motion in sequence of images. Instead of structural features, we prefer to gather and recognize the "motion patters". This method is focusing on representing user arbitrary movement. It records the motion in consecutive frames and pastes them onto a static single image. It has been constructed on layers of images by time to time on moving region of images.

$$TMI_\delta(x,y) = \begin{cases} \tau & if\ (I(x,y)) \neq 0 \\ 0 & else\ if\ I_\delta(x,y) < \tau - \delta \end{cases} \quad (2)$$

Where each pixel $(x, y)$ in TMI (Template Maker Image) is denoted with current time $\tau$ unit if the function $\Psi$ signals that object is present in the running video Image $I(x,y)$. The procedure for assigning the values is of prescribed order that the oldest frame will assign the least value and the last frame assigned with the latest values. The middle ones will be having the values according to the increment adjusted for that sequence of frames. The values we would be assigning to only those pixel locations having different intensities from the previous ones. The detailed flow chart for the algorithm and results are as follows:

(a) Frame-0 (b) Frame-120 (c) Frame-284



(d) TMI-0 (e) TMI-120 (f) TMI-284

(g) Frame-580 (h) Frame-824 (i) Frame-1120

(j) TMI-580 (k) TMI-824 (l) TMI-1120

Fig.1:(a)-(c), (g)-(i) extracted frames and (d)-(f) Template Maker Images of arm waving to drag other attention in Exam and (j)-(l) are of head shaking, wobbling and revolving.

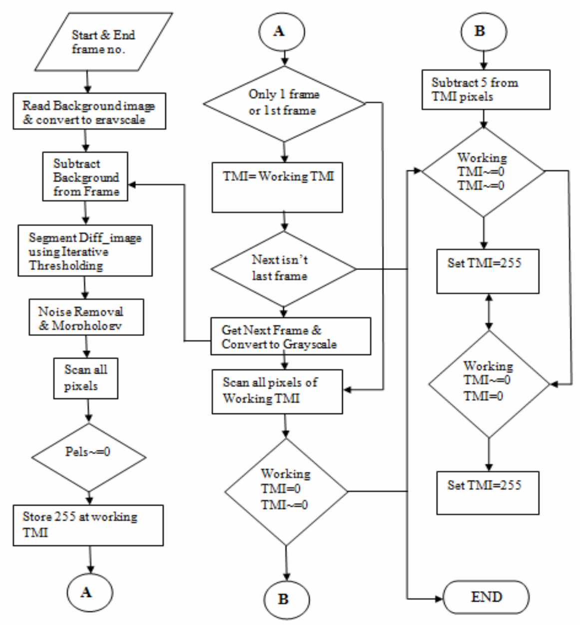



Fig.2:Flow chart of Template Maker Image

## 3.4. Motion Gradients:

The visually appearing gradients in sequence of images in the layered region over time to time can be seen in one template as in the upper figures. All successions are very much apparent from dark to light gradient. The time to time gradient values in history image to extract a motion vector.

The motion vectors are to be calculated using fixed size masks, whether it's bias to speed, the displacement of moving pixels shouldn't be approach outside. For example the standard prewitt 3x3 gradients mask:

$$F_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} / 8 \qquad F_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} / 8 \qquad (3)$$

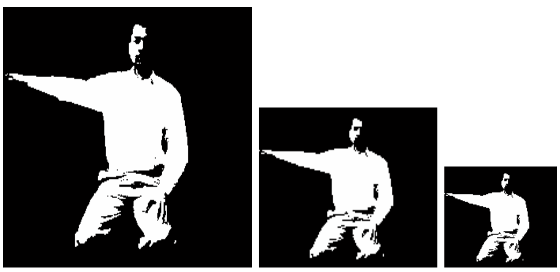
This could not be much supportive at fast moving regions so do with the motion vector problem. Let's suppose to employ the larger mask; this would again be problematic in smoothness case. No one can predict about the speed motion of an object whether it's slow or fast so both cases are being considered and it's been comprehended . For our system, the two approaches have been adopted. First the Davis [13] to use the Hierarchical representation using fixed size gradient operator for fast motion and for the small one, the second for the use of Horn –Schunk [17] implementation.

## 3.5. Hierarchical Analysis

To extrapolate the gradient in multiple image speed, the silhouette based module is further extended to hierarchical strategy. The objective for this portion is all dependent on user entered spatial resolution. These are used to compute large optical flow vector. This could be achieved by having low pass filtering and sub sampling. The investigation for two images having displacement will be easier because if it is large then for increasing levels it become smaller and quite closely co-located.

If $g_0$ is an image then $g_l = REDUCE[g_{l-1}]$ defined for Gaussian and for expand we have $g_{l,1} = EXPAND[g_l]$, for computing laplacian we have:

$$L_k = g_k - EXPAND[g_{k-1}]$$



$\psi^0$ $\psi^1$ $\psi^2$

(a)

$TMI^0$ $TMI^1$ $TMI^2$

(b)

Fig.3:Level Construction of (b) as an (a)

## 3.6. Motion Estimation:

By focusing on Image pyramid, the orientation and speed step must be consider but few constraints are initially expected e.g. boundary. Now strive for imposing 4-connected or 8-connectedness on the time varying pixels for the depth analysis of motion regions. The policy has been constrained for $F_x$, $F_y$ on have some unqualified value minimum to ensure that >m object are being layered. On the basis of this, we select the level "L" of the least acceptable temporal incongruity.

$$L(x,y) = argmin_l \left(I_x^l(x,y)^2 + I_y^l(x,y)^2\right) \tag{4}$$

Suppose working on Level-0, the orientation and speed could be as follows:

$$D(x,y) = \sqrt{U_x(x,y)^2} + \sqrt{U_y(x,y)^2} \tag{5}$$

$$\theta(x,y) = arctan \frac{I_y^L(x',y')^2}{I_x^L(x',y')^2} \tag{6}$$

Where $x' = {}^{x}/{}_{m^L}$ and $y' = {}^{y}/{}_{m^L}$

If we take the gradient of above images, we will get the vector pointing out towards the direction of the movement of let say arm which gives us a normal optical flow representation as in fig .3a.

The second approach of horn-schunk for optical flow in 2-D space or estimated velocity field defined over a domain D.

$$\int_D (\nabla I.\,u + I_t)^2 + \lambda^2(\|\nabla u\|_2^2 + \|\nabla v\|_2^2)dx \tag{7}$$

We have two unknown values (u,v).The (u,v) are true flow which has two components: normal flow "d" and parallel flow "p". For normal true component, the horn-schunk defines the energy function and minimizes it:

$$E(x,y) = \left(uI_x - vI_y + I_t\right)^2 + \lambda\left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right)$$

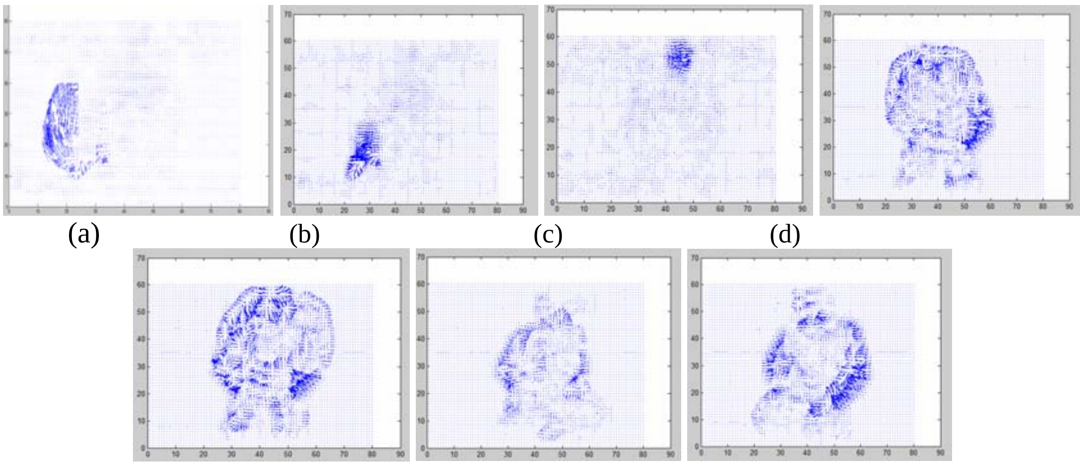
As one can examine the laplacian controls the smoothness of a particular choice can be:

$$\Delta u^2 - u - u_{avg}, \Delta v^2 - v - v_{avg} \tag{8}$$

The magnitude of $\lambda$ classifies the smoothness effect. We use $\lambda = 40$ for better result in most of our test cases. The original method uses the first order differences to produce estimated intensity derivatives and this is a crude form of numerical differentiation and possibility of errors so we implemented the method with spatiotemporal pre smoothing and 8-point central difference especially for rotation and standing cases and finally usage of Gaussian pre filter having standard deviation of 0.8. The iterative cases are used to minimize and to obtain the velocity parameters.

$$u = u_{avg} - I_x\left(\frac{I_x u_{avg} + I_y v_{avg} + I_t}{I_x^2 + I_y^2 + \lambda}\right), v = v_{avg} - I_x\left(\frac{I_x u_{avg} + I_y v_{avg} + I_t}{I_x^2 + I_y^2 + \lambda}\right) \tag{9}$$

In equation (9), it is defined for iteration number 1 and $u_{avg}$ and $v_{avg}$ denotes the neighborhood of the u and v. The following are the motion vectors drawn captured as an output of the Block B module which produces the values for two unknown variables as discussed above.



(a) (b) (c) (d)

(e) (f) (g)

Fig.4:(a) Trajectory of motion field of Block A, (b)-(g) True flow of hand, head, standing, revolving dual sides vector captured showing mysterious behaviour

## 3.7. Marked Features Monitoring:

We have introduced a method [17] for feature selection, a tracking algorithm based on a model of affine image changes and a technique for monitoring feature during tracking. The model was basically described as a predefined input to the system for the point of interest and then track their movement but we actually imply the alteration with respect to our application by inaugurating an Automatic Shifting process. This module takes the input from the previous one in the form of control points.

These control points refers to those body parts experiencing some movement. Those specific parts are being marked with the help of some predefined geometrical shape, where as we are encircling of it. The marked point will remain unaltered for that part but will diversify for the whole lot. Here we could extend our discussion not for a single also for multiple ones. The minute flux ought to be illustrated for this which can make our system more convoluted but one of the possible solutions is available in below figures. The control points acquired in (b) could be used for (c) and (d) in the lieu of new ones from other results.

A simplified case of estimating a dense optical flow is the use of Lucas and Kanade [19] technique. It is similar to line fitting. The temporal persistence is assumed, a window around each feature point in later frame should match that of earlier frame but in a different position. The window size is to be assumed >2 so that the features are to be tracked more robustly. The care has to be taken that a large frame-to-frame pixel movement, an iterative, multi-resolution tracker is considered. The first is of energy function definition:

$$E = \sum (uI_x + vI_y + I_t)^2$$

$$Au = B \qquad A^{-1}Au = A^{-1}B \qquad Iu = A^{-1}B \qquad u = A^{-1}B$$

We setup the above procedure for three levels and iterate on solution by increasing resolution of images and by windowing new images by calculating the displacement.

The least square problem is solved using a nonlinear approach because non convergence is a common phenomenon and under the condition errors are uncorrelated with the predictor variable yields bias estimates. The minimum of sum of squares is found by setting the gradient to zero since the model contains m parameters for m gradient equations:

$$\frac{\partial S}{\partial \beta_j} = 2\sum_i r_i \frac{\partial r_i}{\partial \beta_j} = 0 \; j = 1, \ldots, m$$

$$r_i = y_i - f(x_i, \beta)$$

$$(J^T J)\Delta \beta = J^T \Delta y$$

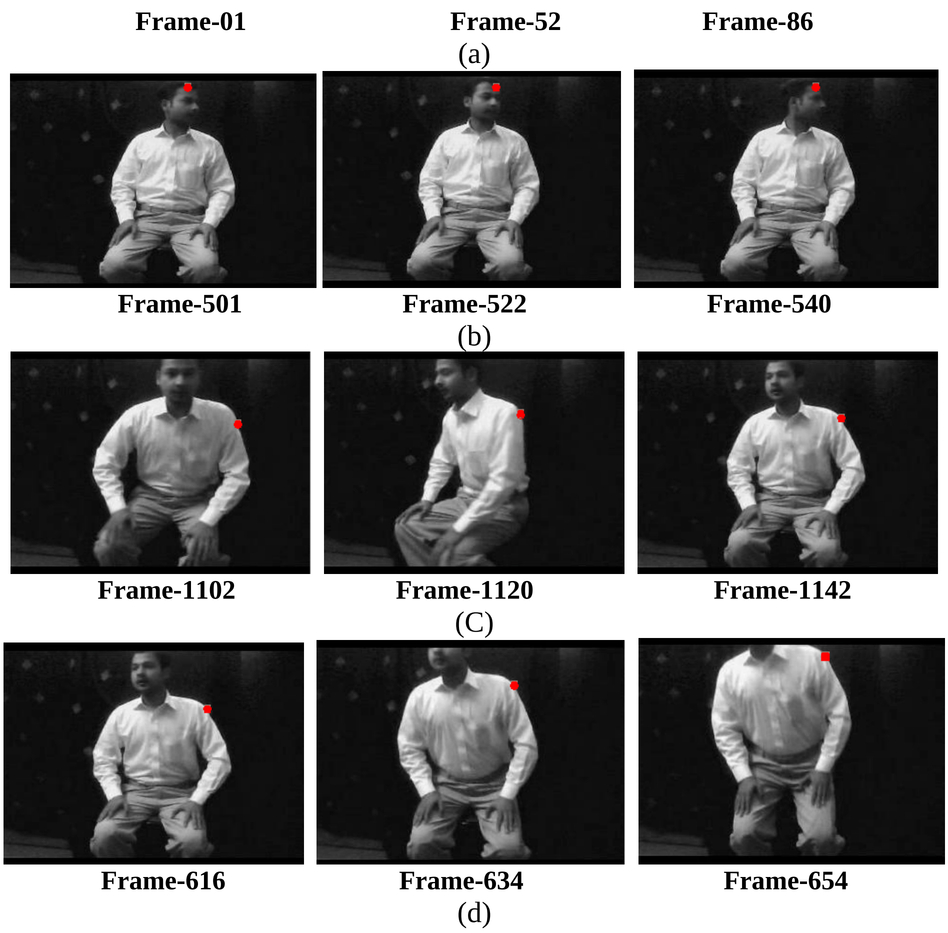



Fig.5:(a)-(d) are results of control points marking & tracking with red circle for all possible motion declare as suspicious for the proposed application

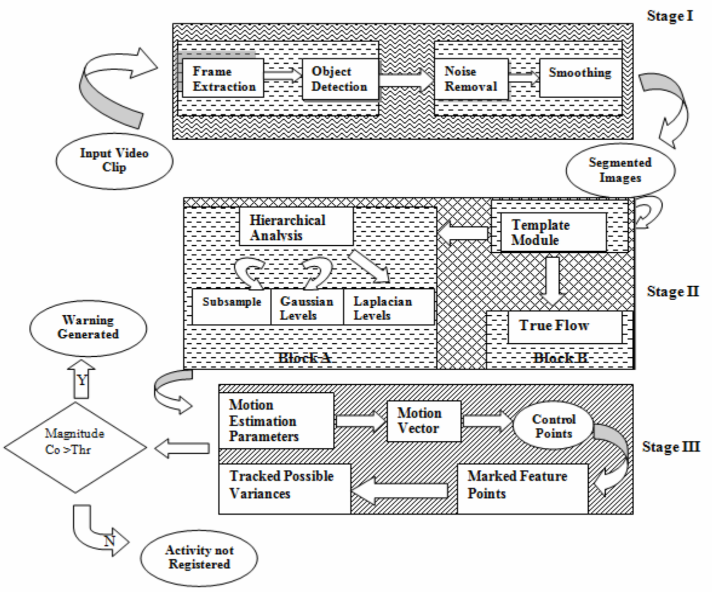



Fig.6:Block Diagram of Proposed System

## 3.8. Behavior Registration:

Mostly we observe that during online tests, the user attention and interaction probability is more than 90% or almost with the computer system so if he has some movement other than the normal one as shown in images then it drags the invigilator attention. This system also assists invigilators in such situations when he has a back towards the students.

As far as the accomplishment of this module is concern, the next phase is of check and balance on the magnitude values. The frame rate is of 20 fps and so far for 120 as in figure .06, the marked point will remain in constraint unlike other [20] .We have the magnitude count for the every iteration of the possible variances as it could be seen from above figures. The varying count for those pixel location if exceeds the rigid threshold value, the warning will be generated for suspicious detection and activity is said to be registered else the system will work on normal routine bases.

## 4. Future Work:

Such system must be supportive in numerous Institutions. The work up to my respect is on first the classification procedure. There could be the more identification stages e.g. the time span for the specific movement. The idea would also be extended by working on the whole exam room having number of student each experiencing various actions.

## 5. References


[1] C. Cedras and M. Shah. Motion-based recognition: a survey. *Image and Vision Comp.*, 13(2):129–155, 1995.

[2] T. Moeslund. Summaries, real time motion template using intel CVL, computer vision-based human motion capture papers. Technical Report LIA 99-01, *University of Aalborg*, 1999.

[3] T. Olson and F. Brill, Moving object detection and event recognition algorithms for smart cameras, *Proc. DARPA Image Understanding Workshop*, pp. 159-175, 1997.

[4] Tao Zhao, Ram Nevatia, Tracking multiple humans in complex situations, *IEEE Trans. on Pattern Analysis and Machine Intelligence*,Vol. 26, No. 9, September 2004.

[5] H. Ren and G. Xu, Human action recognition in smart classroom, *in IEEE Proc. Int. Conf. on Automatic Face and Gesture Recognition,* pp. 399-404, 2002.

[6] J. Ben-Arie and Z. Wang, Human activity recognition using multidimensional indexing, *IEEE Trans. on Pattern Analysis andMachine Intelligence*, vol. 24, no. 8, pp. 1091-1104, August 2002

[7] Xinyu Wu , Yongsheng Ou , Huihuan Qian, Yangsheng Xu , A detection system for human abnormal behaviour .*Dept. of Autom. & Comput.-Aided Eng.,* Chinese Univ. of Hong Kong, China, pp.1204-1208, 2-6 August, 2005

[8] M. Ghallab, On Chronicles: Representation, On-line Recognition and Learning, 5th International Conference on Principles of Knowledge Representation and Reasoning, *Cambridge*, pp. 597-606, 1996.

[9] A. Mecocci, M. Pannozzo and A. Fumarola, Automatic Detection of Anomalous Behavioural Events for Advanced Real-Time Video Surveilance, *International Symposium on Computational Intelligence for Measurement Systems and Applications, Lugano*, pp. 187-192, 2003.

[10] Yasin, hashim and Shoab Ahmed. Moment Invariants Based Human Mistrustful and Suspicious Motion Detection, Recognition and Classification. *Computer Modeling and Simulation, UKSIM 2008. Tenth International Conference on Volume*, Issue, 1-3 April 2008 Page(s):734 – 739

[11] Ismail Haritaoglu, Real-Time Surveillance of People and Their Activities, Pattern Analysis and Machine Intelligence, *Member, IEEE, David Harwood, Member, IEEE, and Larry S. Davis, Fellow, IEEE*, pp. 809-830, August 2000

[12] Jezekiel Ben-Arie, Human Activity Recognition Using Multidimensional Indexing, *Pattern Analysis and Machine Intelligence, Member, IEEE, Zhiqian Wang*, pp. 1091-1104, August 2002

[13] Davis, J.W, Hierarchical motion history images for recognizing human motion, *Detection and Recognition of Events in Video*, pp. 174-184, July 01, 2002