# Human Behavioural Analysis with Self-Organizing Map for Ambient Assisted Living

Kofi Appiah*, Andrew Hunter†, Ahmad Lotfi*, Christopher Waltham† and Patrick Dickinson†

*Department of Computing and Technology
Nottingham Trent University,Clifton Campus
Nottingham, NG11 8NS, UK
Email: {kofi.appiah, ahmad.lofti}@ntu.ac.uk

†Lincoln School of Computing
University of Lincoln, Brayford Campus
Lincoln, LN6 7TS, UK
Email: {ahunter, cwaltham, pdickinson}@lincoln.ac.uk

*Abstract*—This paper presents a system for automatically classifying the resting location of a moving object in an indoor environment. The system uses an unsupervised neural network (Self Organising Feature Map) fully implemented on a low-cost, low-power automated home-based surveillance system, capable of monitoring activity level of elders living alone independently. The proposed system runs on an embedded platform with a specialised ceiling-mounted video sensor for intelligent activity monitoring. The system has the ability to learn resting locations, to measure overall activity levels and to detect specific events such as potential falls. First order motion information, including first order moving average smoothing, is generated from the 2D image coordinates (trajectories). A novel edge-based object detection algorithm capable of running at a reasonable speed on the embedded platform has been developed. The classification is dynamic and achieved in real-time. The dynamic classifier is achieved using a SOFM and a probabilistic model. Experimental results show less than 20% classification error, showing the robustness of our approach over others in literature with minimal power consumption. The head location of the subject is also estimated by a novel approach capable of running on any resource limited platform with power constraints.

## I. INTRODUCTION

An intelligent video surveillance system should be able to keep track of objects in a camera view *(identity tracking)*, determine where the objects are in the camera view *(location tracking)* and what the people, vehicles or objects are doing in the scene *(activity tracking)* [2]. Increasing safety and security concerns have resulted in the development of complex video surveillance and traffic monitoring systems in both research and industrial communities [1]. An intelligent surveillance system should have the capability to process video streams and characterize the actions taking place, to distinguish between normal and abnormal actions, and to draw the attention of a human operator when an action poses a threat.

The United Nations Department of Economic and Social Affairs predicts a worldwide increase in life expectancy coupled with a decline in the number of children and total fertility [3]; thus a drastic reduction in support provided by young individuals working in home and health care for elders. This calls for an automated home-based system, capable of unobtrusive monitoring of the status of elders living alone independently. Ambient intelligence is an emerging discipline that brings intelligence to our everyday environments and makes those environments sensitive to us [4]. The main aim of Ambient Assisted care is to provide quality of life (mainly to the elderly) in all stages of their life.

Many elders living at home on their own may have more than one chronic disease requiring ongoing care. Earlier research into offering some form of holistic health management has focused on medical monitoring like blood glucose levels, heart rate, weight, blood pressure and other similar symptoms [5]. Over the past decade, new research trends have focused on unobtrusive monitoring of behaviours, including activity levels, falls and adherence to health behaviours. Smart home technology developed specifically for Ambient Assistive care or monitoring health behaviours is faced with three basic challenges [5]; the hardware or physical device, the software that runs on the hardware, and the social issues related to the use of technology.

Evolution of computer vision algorithms as well as Ambient intelligence (AmI) [6], [47], oriented towards ubiquitous computing and smart environments that react in an adaptive and active way to the presence and actions of objects implemented on embedded systems, has become an interesting area of research over the past decade. Such systems allow the implementation of early vision processes similar to the first neural layer in the retina, for pre-filtering conspicuous information [7]. The potential uses of embedded systems in areas like medical image processing, computational fluid dynamics, target recognition, embedded vision systems, gesture recognition and automotive infotainment have been demonstrated in [6]–[10]. Karaman et. al [48] demonstrates the essence of video data in monitoring patients with dementia. However, the structural Hidden Markov model used relies on video data recorded via a wearable camera.

In this paper we present a smart home sub-system for monitoring activity levels for elders, which seeks to address the three basic Ambient Assistive technological challenges. The system is part of a wider development; an end-to-end assistive care solution including multiple sensors. It uses specialized ceiling-mounted intelligent activity monitoring sensor to infer the level of activity and behaviour with the use of an unsupervised neural network Self Organising Feature Map (SOFM).

The system has the ability to characterise behaviour as learned resting locations, to measure overall activity levels and to detect specific events such as potential falls. The design has been accomplished on a low-cost embedded platform, which minimizes on-board power consumption.

## II. RELATED WORK

Abnormal activity detection has been divided into two categories: parametric and non-parametric by Zhou et. al [22]. The parametric approach models normal and abnormal activities using visual features like position, speed and appearance, while the non-parametric learns the normal and abnormal patterns from the statistical properties of the observed data. Zhou *et. al* [28] presented an automated activity analysis and summarization for the monitoring of the elderly. They used an adaptive learning method to estimate the physical location and moving speed of a person from a single camera view. Chan *et. al* [30] presented a multi-sensor home monitoring system to help elderly people by observing mobility changes indicative of abnormal events. The design assesses changes in occurrence, time and duration from a statistical perspective. People with dementia often have low physical activity and some sleep problems [31]. The daily life activities and sleeping conditions has been used by Suzuki *et. al* [31] for the early detection of dementia. Adami *et. al* [32] described a system for unobtrusive detection of movement in bed that uses load cells installed at the corners of a bed. The system focused on identifying when a movement occurs based on the forces sensed by the load cells.

Generally, the trajectory data of tracked objects are recorded as a set of $(x, y)$ locations of the tracked object's centre of mass from frame to frame. In [18], they used flow vectors $f = \{x, y, \delta x, \delta y\}$ rather than sequence of positions to describe an object's movement. Thus if an object $i$ appears in $n$ frames it can be represented by a set $Q_i$ of $n$ flow vectors all lying within a unit hypercube in 4D phase space: $Q_i = \{f_1, f_2, \ldots, f_{n-1}, f_n\}$. Owens et. al in [16], used a hierarchical neural network as a novelty detector. Normal trajectories are used during training, and experiment conducted shows a high detection rate. Humphreys et. al [15] has extensively use cost functions based on SOFM to detect, track and classify object trajectories. The paper also demonstrates improved performance, by breaking down the SOFM into three parts.

Grimson et. al [24] used the $(x, y)$ location, speed/direction $(dx, dy)$ and size to develop a codebook using an on-line Vector Quantization(VQ). A co-occurrence statistic is accumulated over the codebook and a hierarchical classification performed to identify normal and abnormal activities. In [23] a Dynamic Oriented Graph (DOG) is used to structure common patterns of objects activities. The entrance, path and departure nodes are used to construct the graph. The spatial motion information, size and colour of the objects are used to classify their activity. During testing, known or normal object patterns should match an existing node, else the activity is rejected and classified as unusual or abnormal.

Jiang et. al [13] uses an information-based trajectory dissimilarity measure making use of Bayesian information criterion (BIC) to determine the number of clusters in the agglomerative hierarchical clustering algorithm. Each trajectory or feature sequence is modelled by a hidden Markov model (HMM) and the number of clusters in the BIC decreases as similar trajectories are merged. In [22] an unsupervised spectral clustering, represented by the mean and variance of their data points is used for clustering trajectories. An unusual approach is taken by Dickinson et. al [14], who use HMMs to differentiate between normal and unusual behaviour in domestic scenes. The hidden states correspond to learned inactive states, such as sitting on a particular chair, and transitions correspond to movement between them. Trajectories are therefore represented as movements between familiar static events. The learned model of normal behaviour is used in conjunction with a threshold model to classify test sequences.

A multi-sample-based similarity measure using a dynamic hierarchical clustering method has been present in [1]. Trajectory data acquired over a period of time are represented by a 5-state HMM with Gaussian emission probability and used as the training data. HMMs are learnt from clusters with large number of samples and are used for detecting abnormal trajectories. Han et. al [27] used an unsupervised fuzzy self-organising map trained with normal activities. Trajectory features $(x, y, \sqrt{\delta x^2 + \delta y^2}, \tan^{-1}(\delta y / \delta x))$ are translated into a fixed length vector of size $4N$. An object with trajectory length less than N, is padded with it last centre position $(x, y)$, zero speed and zero direction. The $4N$ vector is used to train the FSOM, which is then used for the abnormal activity detection in real-time/on-line mode.

A Spatial Occupancy Map (SpOM) built from object trajectories has been used in [12] for detection of unusual trajectories in a camera scene. Object trajectories are modelled as motion time series in [20] to train a coefficient feature space, which is subsequently used for trajectory classification. Principal Component Analysis (PCA) trained with sub trajectories, used in conjunction with HMM has been presented in [11] for classifying motion trajectories. In [19] trajectories are collected and processed on-line as a list of vectors representing the spatial position of the object. Rather than the Euclidean distance, a new distance measure has been introduced to check if a trajectory fits a given cluster.

Owens et. al in [17] used a self-organising feature map to learn and characterize normal trajectories. The 4D flow vector $(x, y, \delta x, \delta y)$ used in [24] has been extended into 8D $(x, y, s(x), s(y), s(\delta x), s(\delta y), s(\delta^2 x), s(\delta^2 y))$ to include a second order motion information. The on-line system presented in [17] is capable of detecting abnormalities in both instantaneous and whole trajectory motion. In [49] a feature vector $\dot{\gamma}_t$ which encapsulates the local curvature of the trajectory as well as the local velocity magnitude is used to represent the tracked object's position. The feature vectors are modelled with a HMM and the similarity between trajectories expressed with a quantization-based HMM. Piciarelli et. al in [25] used a support vector machine (SVM) for the classification and clustering of 2D trajectory data. Mixtures of Gaussians (MoGs) have been used in [26] to group 4D motion histogram data into coherent trajectories and used to identify events after training.

Cuddihy *et. al* [38] presented an Automatic Inactivity Detection system capable of constructing an individual models of normal activity within a home using motion sensor data. Alerts are generated when a period of inactivity exceeds a

normal length for a particular residence. Gil *et. al* [40] used data mining and visualisation to illustrate activity from sensor data that can reveal individual life patterns and changing circumstances. They measure activity to determine whether it would be possible to build a busyness model of a person's life at various levels of granularity.

Hayes et al. [41] presented a system for continuous and unobtrusive monitoring capable of observing changes in physical behaviour over time. Similarly, Shieh *et. al* [36] presented an automatic monitoring system where activities are monitored by infrared positioning sensors. Williams *et. al* [37] presented a design, implementation and evaluation of a distributed network of smart cameras whose function is to detect and localise falls. Their design demonstrates that a distributed low-power system can perform adequately for the detection of fall when monitoring an elderly.

Dickinson and Hunter [35] presented a novel method for detecting unusual modes of behaviour in video surveillance data, suitable for supporting home-based care of elderly patients. Their approach is based on learning a spatial map of normal inactivity for an observed scene to detect unusual patterns of inactivity. In [29], collected data from various sensors to measure activity duration using a hierarchical approach which locally predicts on-going activities by a Support Vector Machine and globally refines them by a Conditional Random Field focused on time segments involving related activities.

Magno *et. al* [33] presented a multi-modal video sensor node designed for low-power video surveillance able to detect changes in the environment. The system has a CMOS video camera and a Pyroelectric InfraRed (PIR) sensor exploited to reduce significantly the power consumption of the system in absence of events. The same combined sensor system has successfully been used to detect abandoned and removed objects [34]. The indoor monitoring sub-system proposed in this paper is based on a similar concept as in [33], [34], to reduce the overall power consumption, but without the use of a PIR. Our system is such that the camera will be operational only when there is movement in the observed environment. Again the behavioural subsystem is based on a neural network with extra parameters in contrast to our previous work [21].

## III. SYSTEM ARCHITECTURE

The proposed energy efficient assistive care sub-system has been developed on an ultra low cost small form factor, high performance leopardboard [43]. Figure 1 is a block-level diagram of the proposed system. The main processor is TMS320DM365 (ARM926EJ-S), an ARM9-based RISC processor offering speeds up to 300MHz. The ARM926EJ-S is a 32-bit processor core that performs 32-bit and 16-bit instructions and processes 32-bit, 16-bit, and 8-bit data. The board is also packaged with the Micron DDR2 SDRAM, it is a +3.3V powered, 8M×16×8 banks, 1Gb DDR2 SDRAM totalling 128MB. The camera board used in this development is the Aptina 1/2.5" CMOS sensor MT9P031 with 2592×1944 pixels. The kernel that runs on the leopardboard has been developed with RidgeRun Embedded Linux Software Development Kit (RRSDK) [42].

For unobtrusiveness, a high-mounted fish-eye lens with viewing angle of 105 degrees has been used. This effectively
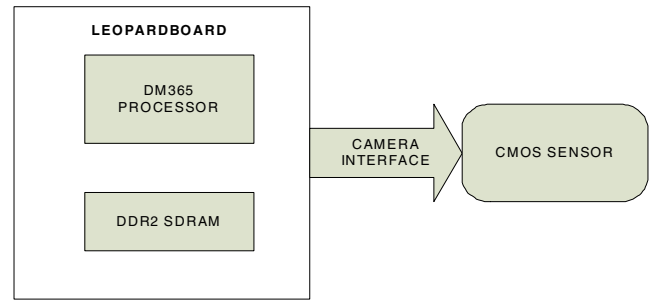


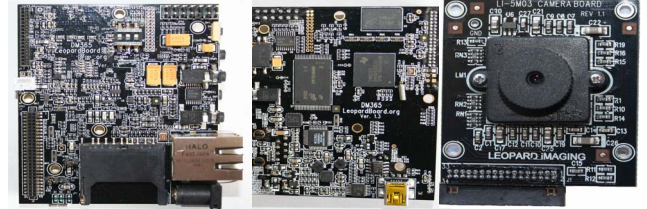Fig. 1.  An overview of the system architecture.



Fig. 2.  The leopard prototyping board with the imaging board used in the development.

reduces the number of cameras needed in any single living environment to one. The lens of the CMOS sensor in figure 2 has been replaced with a 1.9mm (fish-eye) lens. Figure 3 is the fish-eye lens used in this development with a sample output image when mounted high up in the ceiling of an office. A powerful 5-Channel Power Management IC with two step down converters and 3 low input voltage LDOs chip TPS65053 is provided on the leopardboard 365. The board is powered by +5VDC power supply and consumes less than 2W, which includes the 5 Mega-pixel camera board running at 30fps.

The main goal of our system is to be able to determine the presence of a single subject (an elder) and estimate the location in the room with minimal power consumption. Apart from the constraints (in terms of memory and processing power) of the embedded development platform chosen for this implementation, the design is also constrained by a low-power budget. This is to enable the system run for longer period when battery-powered. To satisfy these requirements, a novel edge-based differencing algorithm has been developed and implemented on the embedded ARM9 processor, rather than the standard background differencing algorithm [8], which requires the camera and the leopardboard to run continuously to update the background model. The system described in this paper uses edge-based differencing, as edges are less sensitive to illumination changes. A low power motion detector is used to trigger the camera only when activity occurs.

## IV. VIDEO-BASED ANALYSIS

This section describes the video-based analysis performed on the leopardboard for monitoring the behaviour of an elderly person. The system compared to [21] has three basic sub-systems; the detection behavioural and activity sub-systems as shown in figure 4. The detection and behavioural sub-systems are activated every time the camera switches to RUN mode
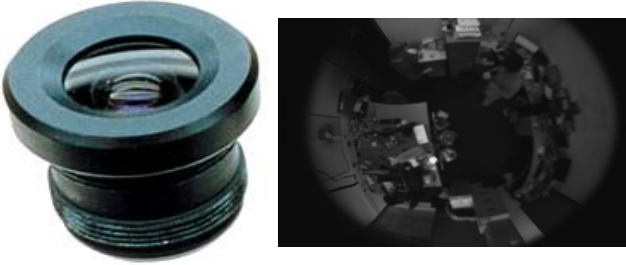
Fig. 3. The 1.9mm wide-view (fish-eye) lens used in the development with a sample output frame.
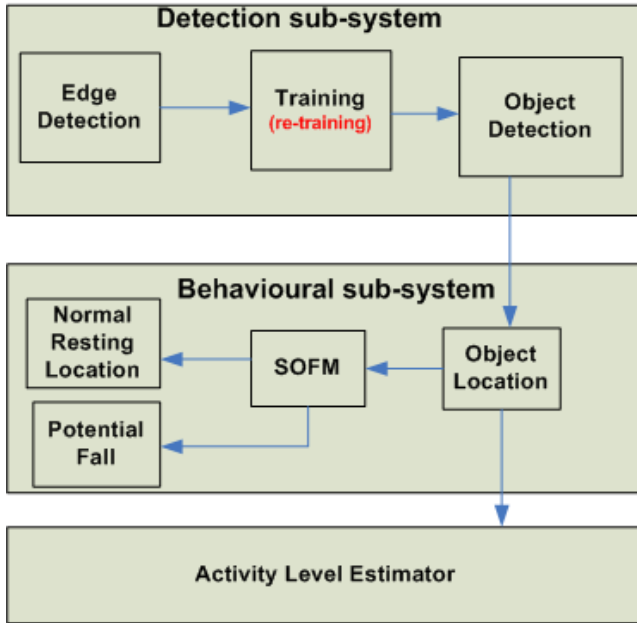


Fig. 4. A block diagram showing the various parts of the video-based analysis.



Fig. 5. The convolution mask used in the Sobel edge detection.

$$|G| = \sqrt{G_x^2 + G_y^2} \qquad (1)$$

$$|G| = |G_x| + |G_y| \qquad (2)$$

To establish the background model **M** using the edge-map **E**, extracted from an image **I**, a burn-in period is used in the training phase. The burn-in period defines the minimum number of frames that can be used to model the background. During this period, the accumulated value **A**, of the number of edges that appear at every pixel location for all the frames is estimated. The burning period is also used to estimate the average number of edges $E_{avg}$ that appear in any particular frame.

$$A_{i+1} = \begin{cases} A_i + 1 & \text{if } I_i \text{ is an edge pixel} \\ A_i & \text{otherwise} \end{cases}$$

where $i$ is the $i^{th}$ pixel in the image **I**. The training phase is activated by the following three conditions:

1) when the system is powered on for the first time,
2) when the number of edges in the current edge-map **E** is greater than 25% of the average number of edges $E_{avg}$, acquired over the previous training phase, and
3) when the number of edges in **E** exceeds 6.25% of $E_{avg}$ continuously for the same number of frames as in the burn-in period.

The retraining of the systems becomes necessary when the camera drifts slightly or when too many objects appear in the scene. The retrain conditions (2 and 3) allow the system to handle major changes; e.g. resetting the camera or rearranging furniture and significant lighting changes. The percentages used in conditions 2 and 3 are empirical values determined experimentally. After the burn-in period, the background model **M** is extracted using the accumulated values at each pixel location. The use of edges rather than intensity values to model the background minimizes the need to continuously update the model to handle minimal illumination changes.

$$M_i = \begin{cases} 1 & \text{if } A_i > \text{threshold} \\ 0 & \text{otherwise} \end{cases}$$

where $M_i$ is the $i^{th}$ element of the background model **M** taken at pixel location $i$.

To extract the edges of any moving object, a simple edge-based differencing algorithm is used. The edge-map extracted from the current frame **E** is compared with the background

when there is enough activity in the scene. The detection sub-system has three major phases: edge detection, training (re-training) and object detection. The edge detection phase uses the Sobel edge [44], which calculates the gradient of the image intensity at each point.

The Sobel edge detector implemented on the leopardboard (ARM9) uses a pair of 3×3 convolution masks as shown in figure 5, $G_x$ is for estimating the gradient in the x-direction (columns) and $G_y$ for estimating the gradient in the y-direction (rows). The mask is slid over the image, manipulating 3×3 pixels at a time. The magnitude of the gradient is generally computed as the square-root of the sum of squares of the horizontal and vertical gradient as show in equation 1. To enable ease of implementation of the gradient magnitude $G$ on the ARM9 processor, an approximation has been used as shown in equation 2. For each pixel in the image (excluding the boundary pixels), the values of $G_x$ and $G_y$ are estimated using the convolution masks shown in figure 5. The sum of the two values is compared with a threshold value and used to determine if that pixel is an edge pixel or not.
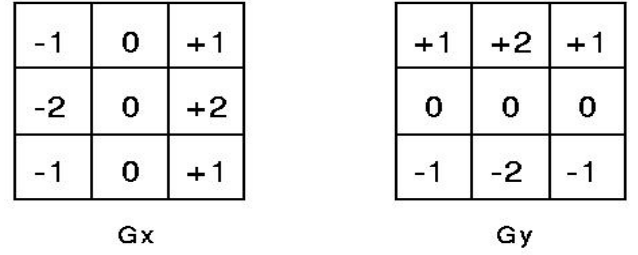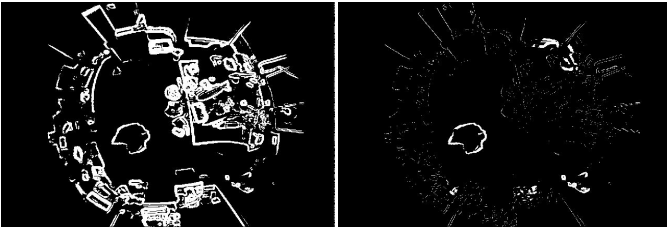
Fig. 6. Processed output from the fish-eye camera setup. The background edge-model image is shown on the left with the processed edge difference to the right.
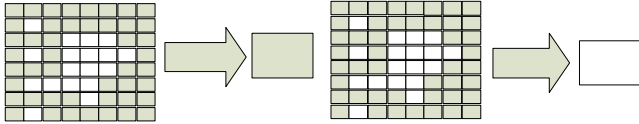


Fig. 7. A threshold value is used to determine whether a resulting block is set or cleared. The left image is a set block while the right is a cleared block



Fig. 8. The intersection of the maximum horizontal and vertical bins resulting from the block-histogram is used to estimate the object centre.

model $\mathbf{M}$ to extract any new edges that appears in the scene. Figure 6 is a sample output of the object detection phase. The edges of the input image are compared with the background model to extract the new edges, which forms the outline of the moving object. The images in figure 6 are actual processed images on the leopardboard.

The central location of the moving object is estimated in the object location phase of the behavioural sub-system. To estimate the area where most edges are concentrated, a tiling operation is adopted. Similar to a compression operation, the entire image is divided into blocks each of size $16\times16$ pixels. Thus for a VGA sized image there are $40\times30$ blocks. The number of edge pixels in each block is compared with a threshold value to determine if the resulting block is set or cleared. This is demonstrated in figure 7 with two $8\times8$ blocks, each with 46 pixels set as edge pixels. If the threshold is less than 46 (as shown to the left), the block is set. Similarly, if the threshold is greater than 46, the resulting block is cleared, as shown to the right in figure 7.

A block-based horizontal and vertical histogram is generated to estimate the area with highest edge blocks. The intersection block (blue block in figure 8) is the block (or area) with the maximum number of edges. This is then used to establish the central location of the moving object. The algorithm is very simple and well-suited for the low-powered embedded platform. The vertical block-based histogram is estimated by accumulating the number of blocks set for each row in the grid. Similar, the horizontal block-based histogram is estimated by accumulating the number of blocks set for each column in the grid. The position of the subject is estimated using the maximum vertical and horizontal block-histogram peaks. To compute the level of activity $\mathbf{T}$ between two frames, a distance measure (Manhattan distance) given as $\mathbf{T} = |x_t - x_{t-1}| + |y_t - y_{t-1}|$ is used. This is accumulated over a period of time to measure the overall activity level.

The central location $(x_t, y_t)$ of an object is collected over a period of time and used in the SOFM clustering phase to determine resting positions. The SOFM algorithm presented in [46] [39] is based on a competitive learning algorithm,
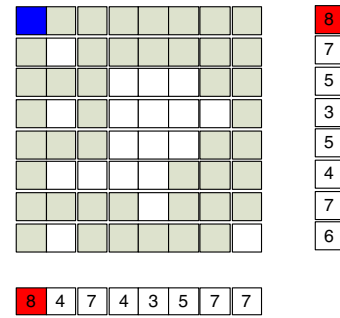
the winner-take-all (WTA) network, where an input vector is represented by the closest neuron prototype vector, which is assigned during training to a data cluster centre. The prototype vectors are stored in the "weights" of the neural network. The architecture consists of topologically organized array of neurons, each with $N-$dimensional weight vector, where $N$ is also the dimensionality of the input vector. The basic principle of the SOM is to adjust the weight vectors until the neurons represent the input data, while using a topological neighbourhood update rule to ensure that similar prototype occupy nearby positions on the topological map.

During training, the "nearest" neuron prototype vector to the input vector is identified – this is called the "winning" neuron – using a distance metric, $D$. The Euclidean distance is most frequently used as the metric.

For a given network with $M$ neurons and $N-$dimensional input vector $\mathbf{x}$, the distance for neuron with weight vector $\boldsymbol{w_j}$ $(j < M)$ is given by

$$D_j^2 = \sum_{i=0}^{N-1} (\mathbf{x_i} - w_{ji})^2. \tag{3}$$

The vector components of the winning node $\boldsymbol{w_k}$ with minimum distance $D_k$ is then updated as follows

$$\Delta w_{ji} = \eta(\mathbf{x_i} - w_{ji}). \tag{4}$$

where $\eta$ is the learning rate. The topological ordering property is imposed by also updating weight vectors of nodes in the neighbourhood of the winning node. This can be achieved by the following learning rule

$$\Delta wji = \eta N_j(\mathbf{x_i} - w_{ji}). \tag{5}$$

where $N_j$ is a neighbourhood function (defining the region around $\boldsymbol{w_k}$ ) based on the topological displacement of neighbouring neuron from the winning neuron. The size of $N_j$ decreases as training progresses.

The collected central positions are used as input to train the SOFM. To efficiently implement a *resting location* discriminator in hardware using Self Organising Feature Map and Gaussian distribution, we have conducted two basic analyses. First, we analyse the minimal dimension that can be used to represent the point-to-point location data $(x_t, y_t)$ without losing any behavioural information. Intuitively, the minimum dimension is 2D, yet in [49] the $(x_t, y_t)$ coordinate information

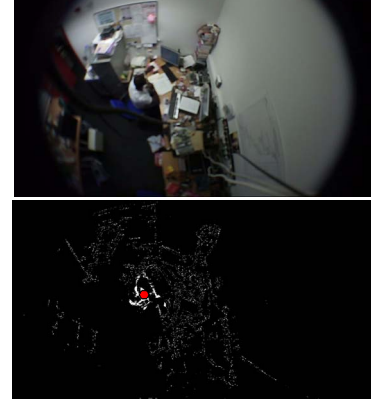Fig. 9. 6 images captured by the system before retraining.



Fig. 10. Extracted central location of an object.
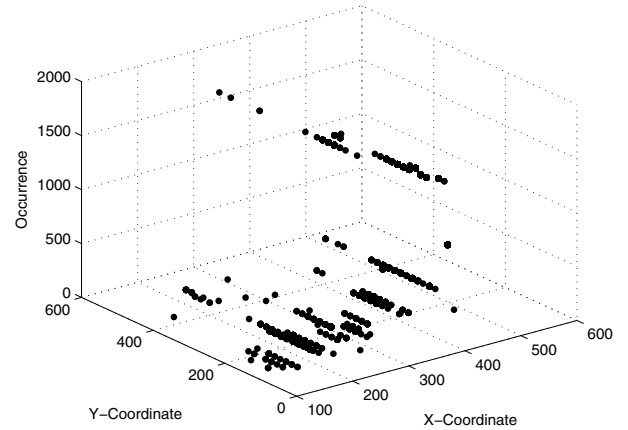


Fig. 11. A 3D graph showing the regions of the image that an object has traversed and their frequencies in the image plane.

has been reduced to a single value $\dot{\gamma}_t$ encoding the local curvature and velocity information. The penalty for the model is the high dimensional vector used in the HMM. Secondly, we analyse the most efficient way to represent the location data in the SOFM and yet able to identify outline location data. By reducing the dimension of the location data we are able to implement the SOFM on the embedded device using minimal resource. During training, we first count a group of input vectors that are associated with each neuron, and then calculate mean value $\mu$ and standard deviation $\sigma$ of the group of distances, alongside the minimum and maximum distances ($D_{min}$ and $D_{max}$). Thus using these extra parameters for each neuron in the network, an alarm is raised if an object stays outside the defined region for more than a specified period (ie. beyond a threshold distance from the closest neuron).

## V. RETRAINING AND ENERGY ANALYSIS

To test the performance of the monitoring system implemented on the DM365 leopardboard, four different rooms have been used. The leopardboard has been used to collect data continuously over period of two weeks unobstructed. The robustness and retraining capabilities of the systems alongside its energy efficiency has also been evaluated. To test the ability to retrain, the system was deployed in a room with wide glass window exposed to external day light over a whole weekend. Over the period of three days, Friday to Monday, the system automatically retrained six times. It should be noted that because the room was not used over the weekend, the only cause of retraining was changes in external light and reflection in the room. Also the reflection in the room changes as the sun changes position during the day. Figure 9 shows the six images taken by the system before each retrain occurred. However, the other three rooms used for testing have at least a single occupant for a minimum of nine hours everyday.

The object detection and location capabilities of the system has also been tested and compared with a PC based implementation. As shown in figure 10, the moving object in the input

image to the top is detected and its centre extracted as shown in the bottom image. The block-based vertical and horizontal histogram is used to estimate the centre of the moving object, shown in red on figure 10. Object locations extracted from all four test rooms have been examined to determine if the resting locations are correctly identified. Segments of the processed video streams on the leopardboard are visually compared with the PC based implementation to verify the object centre is correctly labelled.

Figure 11 is a 3D graph of object locations in one of the experimental rooms with data collected over the period of 8am to 11am. The plot shows the frequency of occurrence of movement at any particular location in the 2D image plane as shown in figure 12. In 12, we show three resting locations. The location with maximum frequency is the area marked with a large green blob. This area has a kettle, which is used by other office occupants. Hence the level of movement around the kettle is high early in the morning. The area marked with a smaller green blob is the second highest in the scene followed by the area in blue.

The power efficiency of the entire system is demonstrated by the number of times the system needs to process an entire
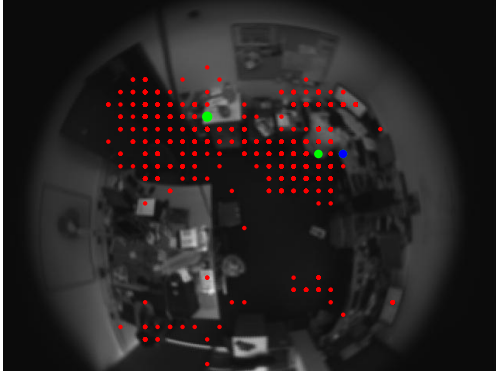
Fig. 12. An image showing in red all the locations traversed by a moving object and the area with the highest frequency.

| Seq. | Movement | Period (hr) | Power Savings |
|------|----------|-------------|---------------|
| A | 69007 | 55.33 | 57.51% |
| B | 16739 | 30 | 90% |
| C | 22366 | 216 | 97.75% |
| D | 6872 | 60 | 97% |

TABLE I.    A TABLE SHOWING THE LEVEL OF ACTIVITY RECORDER OVER A PERIOD FOR ALL THE FOUR ROOMS. THE TOTAL AMOUNT OF POWER SAVED FOR EACH PERIOD BASED ON THE LEVEL OF ACTIVITY OVER THE ENTIRE PERIOD IS ALSO SHOWN.

image frame or retrain in order to correctly identify moving objects. For a period of 3 hours 20 minutes in a normal day settings, the system retrained twice to be able to detect objects correctly. The power consumption of the leopardboard with the 5 mega-pixel camera is 2 watt (2 joules per second). Thus if the camera should run continuously for the entire 200 minutes, it will consume 24 kilowatt (kW). Within this period of time the camera recorded 5,224 activities with a significant level, thus consuming 10.4 kW of power, a 50% savings in power. This is the worse case estimate when there is fairly continuous activity in the scene. In the best case, when the camera will only run to retrain the system, the system will consume approximately 4 watt, more than 99% saving in power consumption. Table I gives an estimate of the worse case power consumption when the camera was deployed in various test rooms at different time intervals.

## VI. CONCLUSION

This paper presents a vision based system for monitoring the activity level of an individual (elder) living alone. The paper presents an embedded system based edge-base differencing algorithm capable of detecting a new or moving object. The system is also capable of training and retraining to build a map of the background using edges. A simple and easily implemented object location algorithm using block-base horizontal and vertical histograms has also been presented. The system is designed to be energy efficient so it can run longer if battery powered, unobtrusive using a fish-eye lens and require no intervention from the user during setup. The system as it is designed requires the camera to be mounted high up in the ceiling of a room. Also, to successfully locate the moving object, the assumption that only one object will be moving in

the scene is made. Single object location data collected over a period is used to train a SOFM which is furthered used to determine outline resting location new to the SOFM. Second order movement information (moving averages for the change in $x$ and $y$) are also used to estimate the activity level of an individual in the scene. Future work will involve developing the system to locate multiple subjects and locate track them in the scene. Work is ongoing to compare the proposed edge-based with background differencing algorithm in terms of speed and power consumption. Again overall activity data collated over a number of sensors will also be incorporated in our future work.

## REFERENCES

[1] Jiang, F., Wu, Y., Katsaggelos, K.A.: *Abnormal Event Detection From Surveillance Video By Dynamic Hierarchical Clustering.* in Proc. IEEE Int'l Conf. on Image Processing (ICIP'07), San Antonio, TX, Sept. 2007.

[2] Hampapur, A., Brown, L., Connell, J., Ekin, A., Haas, N., Lu, M., Merkl, H., Pankanti, S.: *Smart video surveillance: exploring the concept of multiscale spatiotemporal tracking.* IEEE Signal Processing Magazine, Volume 22, Issue 2, March 2005 Page(s): 38 - 51

[3] UN, Department of Economic and Social Affairs, Population Division,, *World Population in 2300*, 2004.

[4] Diane J. Cook and Juan C. Augusto and Vikramaditya R. Jakkula *Ambient intelligence: Technologies, applications, and opportunities*, Pervasive and Mobile Computing, vol. 5, no. 4, pp. 277 - 298, 2009, issn. 1574-1192.

[5] Holly Jimison *A US Perspective on Research Trends in Technologies for Ambient Assisted Living*, The European Ambient Assisted Living Forum 09, 2009.

[6] Meng, H.,Freeman, M.,Pears, N., Bailey, C.: *Real-time human action recognition on an embedded, reconfigurable video processing architecture.* Special Issue of Journal on Real-Time Image Processing, 163 - 176, Volume 3, Number 3, September, 2008.

[7] Chalimbaud, P., Berry, F.: *Embedded Active Vision System Based on an FPGA Architecture* EURASIP Journal on Embedded Systems Volume 2007 (2007).

[8] Appiah, K., Hunter, A.: *A single-chip FPGA implementation of real-time adaptive background model.* IEEE International Conference on Field-Programmable Technology, pp. 95-102, December 2005.

[9] Yamaoka, K., Morimoto, T., Adachi, H., Koide, T., Mattausch, H. J.: *Image segmentation and pattern matching based FPGA/ASIC implementation architecture of real-time object tracking* In Proceedings of the 2006 Conference on Asia South Pacific Design Automation (Yokohama, Japan, January 24 - 27, 2006).

[10] Matteo Tomasi, Javier Daz, Eduardo Ros: *Real Time Architectures for Moving-Objects Tracking* Lecture Notes in Computer Science: Reconfigurable Computing: Architectures, Tools and Applications, Volume 4419, 2007.

[11] Bashir, F., Khokhar, A.,Schonfeld, D.: *Object Trajectory-Based Motion Modeling and Classification using Hidden Markov Models* IEEE Transactions on Image Processing, Vol. 16 (7), July 2007, pp. 1912-1919

[12] Jung, C.,Jacques, J.,Soldera, J., Musse, S.: *Detection of Unusual Motion Using Computer Vision* XIX Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'06),pp.349-356, 2006

[13] Jiang, F., Wu, Y., Katsaggelos, A.K.: *Abnormal event detection based on trajectory clustering by 2-depth greedy search* IEEE International Conference on Speech and Signal Processing, 2008. ICASSP 2008.

[14] Dickinson P., Hunter A.: *Using Inactivity to Detect Unusual Behaviour* Proc. of IEEE Workshop on Motion and Video Computing, Colorado, January 2008.

[15] Humphreys, J., Hunter, A.: *Multiple object tracking using a neural cost function* Image and Vision Computing, June 2008.

[16] Owens, J., Hunter, A., Fletcher, E.: *Novelty Detection in Video Surveillance Using Hierarchical Neural Networks* International Conference on Artificial Neural Networks (ICANN 2002).

[17] Owens, J. and Hunter, A.: *Application of the Self-Organizing Map to Trajectory Classification* IEEE Computer Society Proceedings of the Third IEEE international Workshop on Visual Surveillance (Vs'2000) July 01, 2000

[18] Johnson, N., Hogg, D.: *Learning the distribution of object trajectories for event recognition* Proceedings of the 6th British Conference on Machine Vision (Vol. 2) 1995, UK, 583-592.

[19] Piciarelli, C., Foresti, G.L., Snidara, L.: *Trajectory clustering and its applications for video surveillance* IEEE Conference on AVSS, 2005.

[20] Naftel, A., Anwar, F. B.: *Visual Recognition of Manual Tasks Using Object Motion Trajectories* In Proceedings of the IEEE international Conference on AVSS, 2006.

[21] K. Appiah, A. Hunter, C. Waltham *Low-power and efficient ambient assistive care system for elders,* Computer Vision and Pattern Recognition Workshops (CVPRW), pp.97,102, June 2011

[22] Yue Zhou, Y., Yan, S., Huang, T.S.: *Detecting Anomaly in Videos from Trajectory Similarity Analysis* IEEE International Conference on Multimedia and Expo, 2007.

[23] Duque, D., Santos, Henrique Dinis dos, Cortez, P.: *Prediction of abnormal behaviors for intelligent video surveillance systems* Proceedings of IEEE Symposium on computational intelligence and data mining, USA, 2007

[24] Stauffer, C., Grimson, W. E.: *Learning Patterns of Activity Using Real-Time Tracking* IEEE Trans. Pattern Anal. Mach. Intell. 22, 8 (Aug. 2000), 747-757.

[25] Piciarelli, C., Micheloni, C., Foresti, G.L. : *Trajectory-based anomalous event detection* IEEE Transactions on Circuits and Systems for Video Technology, 2008.

[26] Jung,C. R., Hennemann, L., Musse, S. R.: *Event Detection Using Trajectory Clustering and 4D Histograms* Special issue on Event Analysis in Videos in IEEE Transactions on Circuits and Systems for Video Technology, 2008.

[27] Han, C., Lin, C., Ho, G., Fan, K.: *Abnormal Event Detection Using Trajectory Features* International Computer Symposium, Taiwan December 2006.

[28] Zhongna Zhou and Xi Chen and Yu-Chia Chung and Zhihai He and Tony X. Han and James M. Keller *Activity Analysis, Summarization, and Visualization for Indoor Human Activity Monitoring*, IEEE Transactions on Circuits and Systems for Video Technology, pp. 1489-1498, no. 11, vol. 18, 2008.

[29] Nicolini, Carlo and Lepri, Bruno and Teso, Stefano and Passerini, Andrea, *From On-going to Complete Activity Recognition Exploiting Related Activities*, Proceedings of the First International Conference on Human Behavior Understanding, 2010.

[30] Marie Chan and Eric Campo and Daniel Estve,*Assessment of activity of elderly people using a home monitoring system*, International Journal of Rehabilitation Research, 2005, March , vol. 28, no. 1, pp. 69-76.

[31] Toshiro Suzuki and Sumio Murase and Tomoyuki Tanaka and Takako Okazawa, *Assessment of activity of elderly people using a home monitoring system*, Telemedicine and e-Health, February, 2007, vol. 13, no. 1, pp. 41-44.

[32] Adami, Adriana M. and Pavel, Misha and Hayes, Tamara L. and Singer, Clifford M., *Detection of movement in bed using unobtrusive load cell sensors*, Trans. Info. Tech. Biomed.,vol. 14, issue 2, March 2010,issn. 1089-7771, pp. 481–490

[33] Magno, Michele and Tombari, Federico and Brunelli, Davide and Di Stefano, Luigi and Benini, Luca, *Multi-modal Video Surveillance Aided by Pyroelectric Infrared Sensors*, Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications - M2SFA2 2008 ,2008.

[34] Michele Magno and Federico Tombari and Davide Brunelli and Luigi Di Stefano and Luca Benini, *Multimodal Abandoned/Removed Object Detection for Low Power Video Surveillance Systems*, Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance,september 2009,isbn. 978-1-4244-4755-8, pp. 188 – 193.

[35] P. Dickinson and A. Hunter, *Using Inactivity to Detect Unusual Behaviour,*Proc. of IEEE Workshop on Motion and Video Computing, Colorado, January 2008.

[36] Jiann-Shing Shieh and Chen-Tse Chuang and Xiang Wang and Pen-Yi Kuo,*Remote Monitoring of Mobility Changes of the Elderly at Home Using Frequency Rank Order Statistics,*Journal of Medical and Biological Engineering,vol. 26, no. 2,March 2006,pp. 81–88.

[37] Williams, Adam and Ganesan, Deepak and Hanson, Allen, *Aging in place: fall detection and localization in a distributed smart camera network*, Proceedings of the 15th international conference on Multimedia, isbn. 978-1-59593-702-5,pp. 892–901,no. 10, 2007.

[38] Cuddihy, Paul and Weisenberg, Jenny and Graichen, Catherine and Ganesh, Meena,*Algorithm to automatically detect abnormally long periods of inactivity in a home,*Proceedings of the 1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments, isbn. 978-1-59593-767-4,pp. 89–94, no. 6,2007.

[39] Appiah, K. and Hunter, A. and Hongying Meng and Shigang Yue and Hobden, M. and Priestley, N. and Hobden, P. and Pettit, C., *A binary Self-Organizing Map and its FPGA implementation*, International Joint Conference on Neural Networks, IJCNN 2009. , pp. 164-171, ISSN. 1098-7576.

[40] Gil, Nubia M. and Hine, Nicolas A. and Arnott, John L. and Hanson, Julienne and Curry, Richard G. and Amaral, Telmo and Osipovic, Dorota,*Data visualisation and data mining technology for supporting care for older people,*Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility,pp. 139–146, no. 8,2007.

[41] T. L. Hayes and M. Pavel and J. A. Kaye,*An unobtrusive in-home monitoring system for detection of key motor changes preceding cognitive decline* , 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, September, 2004,pp. 2480 – 2483.

[42] RidgeRun, *SDK for Texas Instruments DaVinci DM365,*Website,2010,http://www.ridgerun.com/products/dm36x.shtml.

[43] *Leopard Imaging Inc*, Leopardboard Peripherals & Accessories, Website,2010,http://designsomething.org/leopardboard/default.aspx.

[44] N. Kazakova and M. Margala and N.G. Durdle, *Sobel edge detection processor for a real-time volume rendering system* , Proceedings of the IEEE International Symposium on Circuits and Systems, may 2004, isbn. 0-7803-8251-X, pp. 23 – 26.

[45] Kanungo, Tapas and Mount, David M. and Netanyahu, Nathan S. and Piatko, Christine D. and Silverman, Ruth and Wu, Angela Y., *An Efficient k-Means Clustering Algorithm: Analysis and Implementation,*IEEE Transactions on Pattern Analysis and Machine Intelligence,vol. 24,issue. 7,no. 12,pp. 881–892 July 2002,

[46] T. Kohonene, *Self-Organizing Maps.* Springer,New York, 1995.

[47] Alexandros Andr Chaaraoui and Pau Climent-Prez and Francisco Flrez-Revuelta, *A review on vision techniques applied to Human Behaviour Analysis for Ambient-Assisted Living*, Expert Systems with Applications, vol.39, no.12, 2012.

[48] Karaman, S. Benois-Pineau, J. Megret, R. Dovgalecs, V. Dartigues, J.-F. Gaestel, Y., *Human Daily Activities Indexing in Videos from Wearable Cameras for Monitoring of Patients with Dementia Diseases,* 20th International Conference on Pattern Recognition (ICPR), pp.4113–4116, 2010

[49] A. Hervieu, P. Bouthemy, J.-P. Le Cadre.: *A statistical video content recognition method using invariant features on object trajectories* IEEE Trans. on CSVT (Special Issue on "Event Analysis in Videos"), 2008