

ABNORMAL MOTION DETECTION IN REAL TIME USING VIDEO SURVEILLANCE AND BODY SENSORS

MAYANK BARANWAL

*Department of Mechanical Engineering
University of Illinois at Urbana-Champaign, USA
baranwa2@illinois.edu*

M. TAHIR KHAN* and CLARENCE W. DE SILVA†

*Department of Mechanical Engineering
The University of British Columbia, Vancouver, Canada*

**mkhan@mail.ubc.ca*

†desilva@mech.ubc.ca

Received 13 April 2011

Accepted 3 July 2011

This paper presents a method for detecting abnormal motion in real time using a computer vision system. The method is based on the modeling of human body image, which takes into account both orientation and velocity of prominent body parts. A comparative study is made of this method with other existing algorithms based on optical flow and the use of accelerometer body sensors. From the real time experiments conducted in the present work, the developed method is found to be efficient in characterizing human motion and classifying it into basic types such as falling, sitting, and walking. The method uses a Radial Basis Function Network (RBFN) to compute the severity coefficient associated with the type of motion, based on experience. The paper evaluates the various methods and incorporates the advantages of other methods in order to develop a more reliable system for abnormal motion detection.

Keywords: Abnormal motion detection; optical flow; accelerometer body sensors; RBFN (Radial Basis Function Network).

1. Introduction

Though robots have been used in many applications including factories and rescue operations, home care robotics is a newer development which is likely to become significant with an aging population and changing demographics and household needs. In a fast-paced society,

people find it difficult to manage their daily household tasks. Human housekeepers and helpers are commonly used while the regularity and reliability of such assistance are often wanting. As the life expectancy of the population increases, the need for homecare robotic technologies will grow. For example, during the 20th

century, the US population under the age 65 has tripled while those who are 65 and older have increased by a factor of 11. According to International Population Reports [Kinsella & He, 2009], about 1 in 8 Americans were elderly in 1994, and about 1 in 5 will be elderly by the year 2030. In addition, the percentage of the population aged over 65 in European nations is anticipated to rise from 17.9% in 2007 to 53.5% in 2060. According to the 2006 Canadian Census, the number of people aged 65 and over increased by more than 446,700 compared to 2001 (+11.5%), exceeding the 4-million mark for the first time (4.3 million).

Falls, collisions, strokes, and heart attacks are among the greatest dangers faced by the elderly. For those who live alone in particular, such incidents can be extremely dangerous. The injured persons need to be provided immediate first aid or other emergency attention. A home environment that is able to automatically monitor the motion of the occupants and quickly and accurately determine abnormal motions will improve the quality of life and reduce the health care costs. Lives can be saved if immediate help is provided to injured persons.

For a robot to be able to provide help, it must first be able to differentiate between normal and abnormal motions and then identify the abnormalities. There are various approaches for detecting abnormal human motions. The most popular is the use of wearable body sensors, particularly accelerometers and gyroscopes, to measure translational and angular accelerations of the body. This method is also useful in studying medical conditions like Parkinson's disease, where hands and other body parts exhibit distinct tremors in the frequency range of 6 to 8 Hz. These sensors may be worn on different body parts. However, people may forget to wear them or may feel uncomfortable wearing them. Hence, a non-wearable or non-contact system for detecting abnormal motions in humans will be advantageous. The use of a vision system is able to provide an appropriate method to this end, as developed in this paper.

This paper presents two vision-based approaches for detecting abnormal motions in humans. The first approach uses optical flow to

determine various flow parameters such as velocities and displacements. The second approach uses modeling of human body images. A human body image is modeled as an ellipse, which has the following advantages:

- (a) Ellipse is the simplest geometric figure that can represent a human body as an image
- (b) Computationally, it is very efficient
- (c) Change in orientation and other parameters of the ellipse will help in determining the motion of human body.

Since the entire operation is performed using image pixels and not in terms of world coordinates, the method does not require camera calibration, thereby simplifying camera installation.

Both of these vision-based methods have pros and cons. The paper discusses the efficiency and usability of the methods for different types of human motion. It also proposes a weighted scheme to compare the methods in order to identify which method should be used in a particular situation. An accelerometer-based approach is also presented for comparison. Section 2 highlights previous research in the field. Section 3 describes the methodologies used in this paper and the system overview. Section 4 presents results and a comparative analysis based on them. Section 5 provides conclusions.

2. Related Work

The problem of abnormal motion detection has been studied by researchers in two main ways. These methods are primarily distinguished by the type of sensors used. The first approach makes use of wearable body sensors like accelerometers and gyroscopes. Accelerometers are MEMS-based sensors used to measure the accelerations at a particular location in a body. Gyroscopes are used to measure the orientation of a body. The orientation can also be measured by measuring the acceleration components in orthogonal directions. The second approach is a non-contact approach, which uses vision. It is based on extracting human motion data from a sequence of still images or a video

and using computer vision capabilities to detect abnormality in motion.

The first approach is relatively simple and can be easily employed. A drawback of this approach is that it is not human-independent. Abnormal motion can only be detected when the sensors are worn, and therefore the person has to wear the sensors all the time. This may be inconvenient or impractical in some situations. Hence, a human-independent approach, particularly a vision-based approach, has advantages in this regard.

Douka *et al.* [2009] proposed a scheme that combined both motion and visual information. Acceleration data from the sensors were used to indicate a fall incident, and trajectory information and the subject's visual location were used to verify the fall. For fall detection, the scheme relied on detecting a sudden change in posture followed by inactivity.

Luštrek & Kaluža [2009] used radio tags placed on various body parts to determine their locations. The use of radio tags made this scheme human-dependent, and moreover, the number of tags used was as high as 12, causing further practical difficulty.

Liu *et al.* [2010] used a kNN classification algorithm to classify postures using the ratio and difference of a human body silhouette, bounding box height and width. Though the method was considered to be efficient, it might result in false positive cases when the person stretched his hands to pick up objects. A scheme was given to reduce this effect.

Doulamis *et al.* [2010] used the trajectory of the upper boundary of the foreground object to detect a fall. However, motions like a slip seem to have a larger horizontal motion as well and are slightly different from fall scenarios, where feet displacements are small. Similarly, Tehrani *et al.* [2009] used two features, centroid and top of head, to estimate a fall. This scheme has the same drawback as that mentioned above.

Modeling of a human body image as an ellipse is mainly motivated from the work of Nait-Charif & McKenna [2004] and Foroughi *et al.* [2008]. Nait-Charif & McKenna [2004], however, estimated unusual motions by finding

inactivity in activity zones of the room. Such an approach is highly room specific and gives a high number of false positive results. On the other hand, Foroughi *et al.* [2008] used a combination of an approximated ellipse around the human body, horizontal and vertical projection histograms, and temporal changes of the human head position as feature vectors. Fall detection was performed by the standard deviations of the orientation and the dimension ratio of the ellipse. Similar to Doulamis *et al.* [2010] and Tehrani *et al.* [2009], however, Foroughi *et al.* [2008] used a network of overlapping low sample rate cameras. Due to the small sampling rate, the work assumed that a fall was defined as an event that resulted in a person lying on the floor in a position which they were unable to quickly recover from. Diraco *et al.* [2010] computed the location of a 3D human centroid with respect to the floor. It is known, however, that distance alone cannot accurately characterize a fall.

Lin & Ling [2007] determined fall incidents by the location and the rate of change of the centroid of a human object. However, it is known that other features must be taken into account as well. Jansen & Deklerck [2006] detected fall incidents by determining the body posture along with the period of inactivity. However, the inactivity zones had to be pre-defined — i.e., calibration was required for each room. Zhong *et al.* [2004] determined unusual activity by matching the extracted features of the video with a set of pre-stored prototypes. Though the method is equivalent to using a neural network-based approach, it involved excessive computation, and the number of prototypes was a limiting factor.

The algorithm employed by Yu *et al.* [2009] uses constraints on human state change patterns to discard some patterns that do not correlate with the fall event. For instance, when a person falls from standing, the fall will be constrained by standing, bending, and lying. However, such constraints may not be helpful in identifying a wide range of abnormal motions, as the number of constraints is limited. Rougier & Meunier [2006] detected falls using only the 3D head trajectory, while Rougier *et al.* [2008] detected

falls by analyzing the human shape deformation using a video sequence.

The work described in Debard *et al.* [2009] is based on the knowledge of whether the posture is upright or not, whereas Miaou *et al.* [2006] used the change of the ratio of a person's height to width as features and obtained the associated thresholds in two ways: one took personal information into consideration and the other did not. The method described in Debard *et al.* [2009] cannot differentiate among motions such as a fall, sitting, and pickup, as the body orientation can be the same in all these cases.

Khan & Habib [2009] used a combination of motion gradients and change in shape for fall detection. Hazelho *et al.* [2008] detected falls by determining the direction of the main axis of the body and the ratio of the variances in the horizontal and vertical directions. However, if a person stretches out their hands, the variances in the vertical direction may become noticeable and may give a false positive result.

3. Methodology

Given an indoor environment, the present aim is to develop an algorithm that can detect various types of abnormal human motion, and can classify them. Although the algorithm developed here is environmentally independent in that it does not require pre-calibration based on the arrangement of the indoor environment, it uses an important condition that the largest moving object in the environment is a human. This condition helps in real time detection of abnormal motions using cameras. The other way the problem of abnormal motion detection has been approached is by the use of wearable body sensors such as accelerometers.

The specific assumption concerning the largest moving object in the indoor environment is particularly useful when the human motion is abnormal. This assumption also removes the discrepancy in the results obtained from skin color identification as the skin color depends on the person. Also, human identification based on geometric modeling of the human body is not particularly useful in the

detection of abnormal motions where a simplified geometric shape can hardly match the human shape.

Once a moving human object is identified, the body image is modeled as an ellipse. The ellipse is the simplest 2D geometric figure that can represent an image of a human body. The parameters of the ellipse are utilized in motion identification and determination of the velocity vectors. The determination of various flow vectors using optical flow can fail to deliver satisfactory results, however, for abnormal motions like falls.

The other approach is the use of wearable body sensors. Though this approach is simpler, it is not human-independent, as the person has to wear the sensors all the time. With this approach, there is no compensation for such occurrences as forgetfulness.

Use of vision systems

The present approach of abnormal motion detection using a vision system consists of the following four steps:

- (a) Identification of a human in the image
- (b) Extraction of motion parameters by using
 - (1) optical flow, or
 - (2) human body image modeling
- (c) Training the neural network
- (d) Severity value calculation and decision making

3.1. Identification of human in the image

In an indoor environment, the task is limited to identifying all the moving objects and finding the largest moving object in the image. The foreground information in the image is extracted by subtracting the background from it (see Fig. 1).

This approach has a major disadvantage. If an object changes its position in the successive images, it will always be treated as moving. Hence, the background model is adaptive in nature, as the current frame is actually the current frame along with some part of the previous

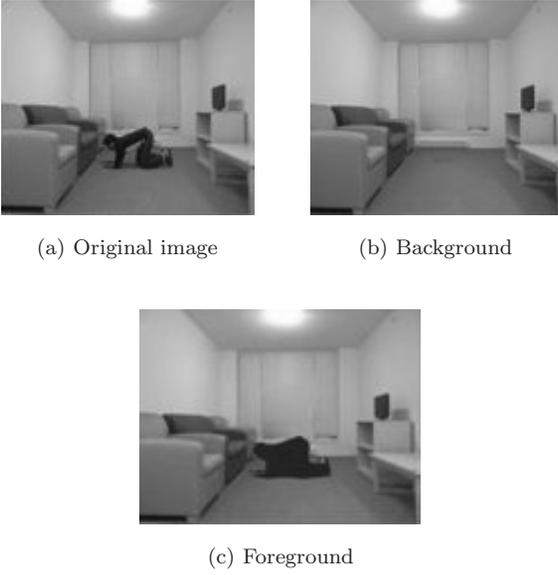


Fig. 1. Foreground extraction.

frame:

$$\begin{aligned}\mu_{\text{new}} &= (1 - \alpha) * \mu_{\text{old}} + \alpha * (\text{new image}) \\ \sigma_{\text{new}} &= (\max(\sigma_{\text{min}}, (1 - \alpha) * \sigma_{\text{old}}^2 \\ &\quad + \alpha * (\text{new image})^2))^{0.5}\end{aligned}\quad (1)$$

where

- μ_{old} – mean values of all pixels of penultimate frame
- μ_{new} – mean values of all pixels of current frame
- σ_{old} – std. deviations of all pixels of penultimate frame
- σ_{new} – std. deviations of all pixels of current frame
- σ_{min} – threshold value of std. deviations
- α – learning rate

The higher the learning rate, the faster the adaptation of the background. There exist many noisy pixels as well. A lower learning rate slows down the process of adaptation, but with a very few noisy pixels. The condition for a pixel to be a part of the background is:

$$\begin{aligned}\mu_{\text{new}} - 0.5 * \sigma_{\text{new}} &< \text{pixel value} < \mu_{\text{new}} \\ &+ 0.5 * \sigma_{\text{new}}\end{aligned}\quad (2)$$

Foreground refinement by hole filling:

Once the foreground pixels are obtained, the foreground is further refined by expanding the

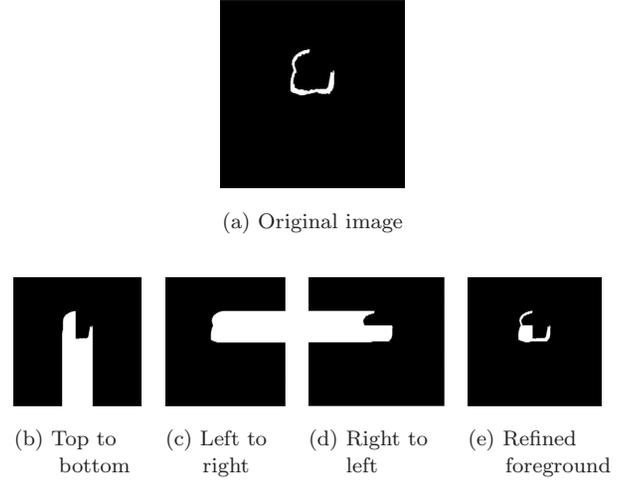


Fig. 2. Foreground refinement by hole filling.

blob in three directions and then taking their intersection. This helps in filling the voids and empty spaces in blobs (see Fig. 2).

3.2. Extraction of optical flow parameters

Optical and affine flow is based on the assumption that the velocity within the considered neighborhood is constant. It also assumes brightness constancy, which means that the brightness structures of local time-varying image regions remain unchanged under motion for a short period of time. It can be formulated as:

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (3)$$

The first-order Taylor series expansion yields the following set of linear equations:

$$\begin{Bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{Bmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{Bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{Bmatrix} \quad (4)$$

where $I(x, y)$ is the intensity of the pixel at location (x, y) .

In affine transformation [Derpanis, 2005], the system is modeled as

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_4 & a_5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_0 \\ a_3 \end{pmatrix} \quad (5)$$

where $(u, v)^T$ is the velocity vector at location $(x, y)^T$.

The affine parameters are evaluated using:

$$\begin{pmatrix} \sum I_x^2 & \sum xI_x^2 & \sum yI_x^2 & \sum I_xI_y & \sum xI_xI_y & \sum yI_xI_y \\ \sum xI_x^2 & \sum x^2I_x^2 & \sum xyI_x^2 & \sum xI_xI_y & \sum x^2I_xI_y & \sum xyI_xI_y \\ \sum yI_x^2 & \sum xyI_x^2 & \sum y^2I_x^2 & \sum yI_xI_y & \sum xyI_xI_y & \sum y^2I_xI_y \\ \sum I_xI_y & \sum xI_xI_y & \sum yI_xI_y & \sum I_y^2 & \sum xI_y^2 & \sum yI_y^2 \\ \sum xI_xI_y & \sum x^2I_xI_y & \sum xyI_xI_y & \sum xI_y^2 & \sum x^2I_y^2 & \sum xyI_y^2 \\ \sum yI_xI_y & \sum xyI_xI_y & \sum y^2I_xI_y & \sum yI_y^2 & \sum xyI_y^2 & \sum y^2I_y^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = - \begin{pmatrix} \sum I_xI_t \\ \sum xI_xI_t \\ \sum yI_xI_t \\ \sum I_yI_t \\ \sum xI_yI_t \\ \sum yI_yI_t \end{pmatrix} \tag{6}$$

In this case, at least six points are required for a unique solution. Since the Taylor series expansion approximates the solution, the precision of the estimates is enhanced by iterative procedures.

The results of the optical and affine flow model for various cases are given now. For small motions (see Figs. 3(a) and 3(b)) — i.e., normal motions — the human body may be treated as a rigid and non-deformable object, and it is

possible to get a very good estimate of the flow parameters. Fig. 3(c) indicates estimated motion (1) by the conventional method and (2) by the log polar sampling method. It is seen that for small deformations, the results are quite accurate.

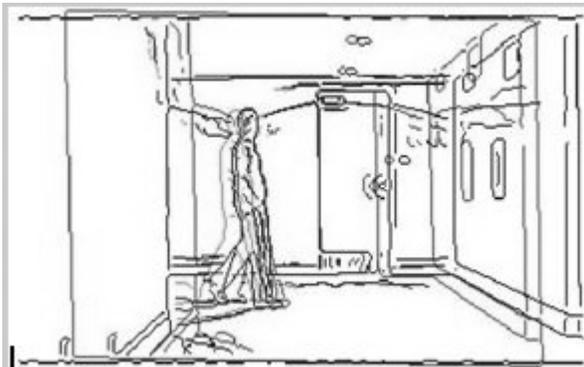
The results for large displacements show that large deformations (Fig. 4) are difficult to track by the method of optical flow. The situation



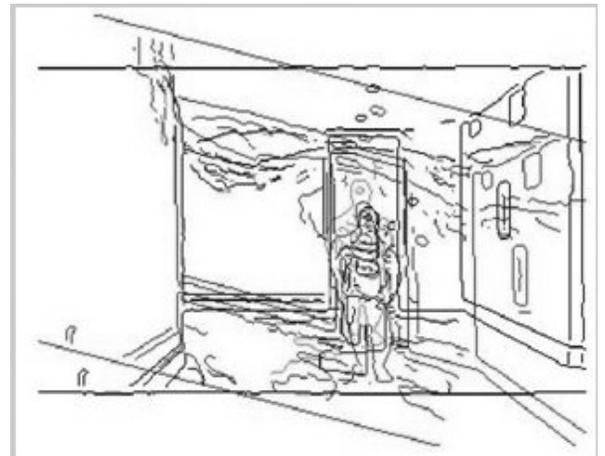
(a) Image I (b) Image II



(a) Image I (b) Image II



(c) Transformed image with estimated flow parameters



(c) Transformed image with estimated flow parameters

Fig. 3. Flow results for small deformations.

Fig. 4. Flow results for large deformations.

worsens when the object being tracked is a deformable one such as a human.

3.3. Human body image modeling

To overcome the problem in tracking deformable objects such as humans, the present paper proposes a method modeling a human body image. This method also proves to be efficient in handling complexity and hence can be easily employed in real time procedures.

A simple model for human body image is thought of as an ellipse. The advantages are as follows:

- (a) It is the simplest geometric shape that can represent the human body. Being a simple shape, the extraction of information from it is relatively less complex and time-consuming.
- (b) The rate of change of orientation of an ellipse is an indicator of severity of abnormal motion. A large rate of change of the ellipse orientation may indicate an event such as a fall or collision.
- (c) The rate of change of dimensions of an ellipse is also an indicator of the severity of abnormal motion.

Extraction of ellipse: The extraction of an ellipse from a background image is a four-step process:

- (I) Bounding box parameters are collected using an available standard algorithm.
- (II) The orientation of the blob is measured.
- (III) The larger dimension of the bounding box is treated as the major axis of the ellipse, and the smaller dimension as the minor axis. The orientation of the ellipse is decided by the orientation of the blob.
- (IV) The ellipse is constructed using the following parametric information (see Figs. 5(a) and 5(b)):

$$\begin{aligned} X(t) &= X_c + a * \cos(t) * \cos(\Phi) \\ &\quad - b * \sin(t) * \sin(\Phi) \\ Y(t) &= Y_c + a * \cos(t) * \sin(\Phi) \\ &\quad - b * \sin(t) * \cos(\Phi) \end{aligned} \quad (7)$$

where Φ is the orientation of the ellipse.



(a) Bending down

(b) Lying down

Fig. 5. Ellipse estimation.

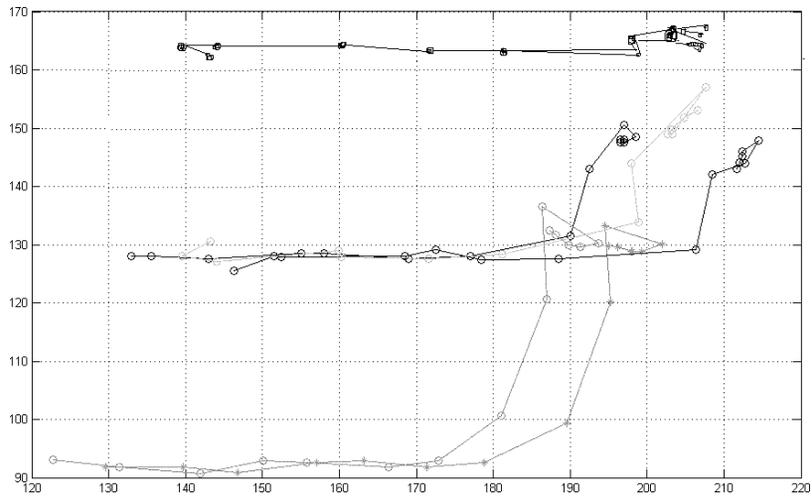
Once the matching ellipse is obtained, the next task is to track various points on the ellipse and the centroid. For convenience, the points that play a crucial role in determining the category of motion are selected. These points are located on the ellipse at angles $-15^\circ, 15^\circ, 90^\circ, 165^\circ, -165^\circ$, and -90° . The motion of the centroid is also observed here. The first two points correspond to a region near the head, and the last two points correspond to a region near the legs. The other two points correspond to a region near the waist.

These points on the ellipse are tracked in successive frames. The changes in the positions of these points are noted for various cases of motion like fall, sit, pick-up, and walk. An artificial neural network is then constructed and the network is trained for various cases.

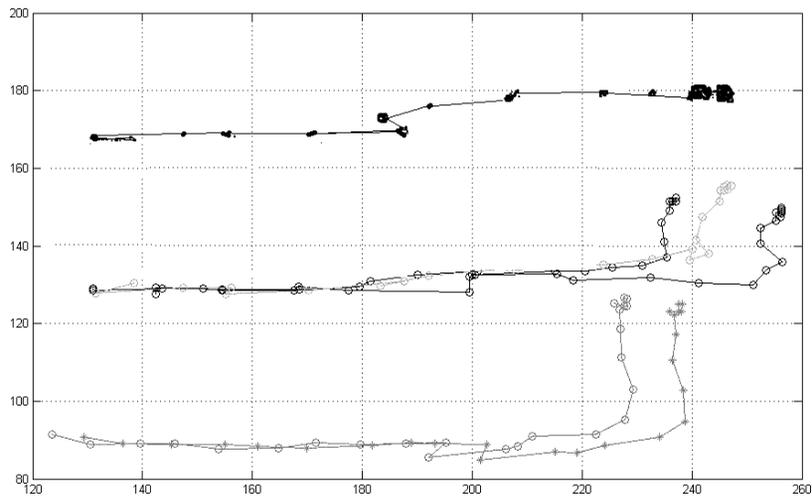
4. Results and Evaluation

Trajectories of the points on the ellipse for different motion scenarios are given in Fig. 6. Trajectories at the top indicate motion of points near head region; trajectories in the middle indicate motions of centroid and points near waist region; and the trajectories at the bottom indicate motion of points near legs. In the case of a fall (Fig. 6(a)), there is a sharp change in the trajectory of various points of the ellipse as compared to sit (Fig. 6(b)) and pick-up (Fig. 6(c)).

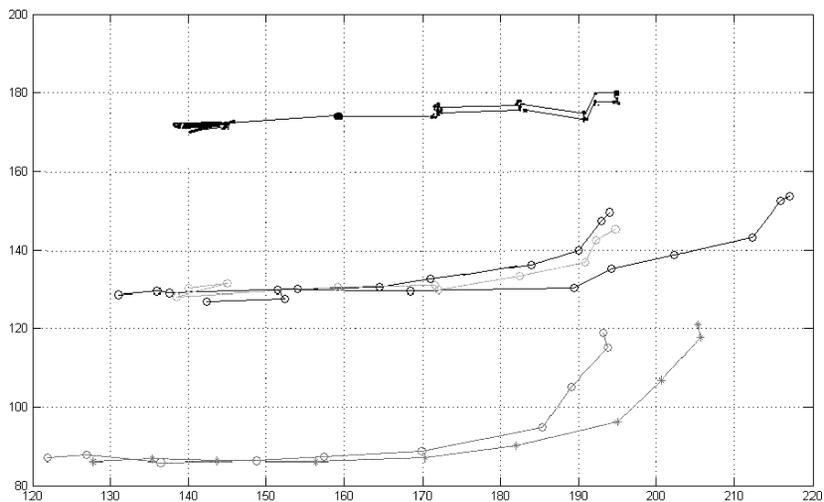
It is seen that there is a clear difference among the trajectories adopted by the points on the ellipse in different scenarios, which enables the category of motion to be characterized. From the data, it is observed that during



(a) The case of "Fall"



(b) The case of "Sit"



(c) The case of "Pick up"

Fig. 6. Trajectory of ellipse points during various motion scenarios.

Table 1. Displacement data.

Case	Displacement of Centroid	Disp. of Head (pt. 1)	Disp. of Head (pt. 2)	Disp. of Legs (pt. 1)	Disp. of Legs (pt. 2)
Fall	16.12	14.89	13.82	17.36	18.42
	16.40	19.54	21.02	15.19	12.78
	25.00	41.39	44.38	13.18	8.91
Fall	12.66	12.04	17.95	13.53	7.63
	12.78	26.35	31.88	1.34	7.52
	20.89	45.03	44.25	5.61	4.97
Fall	18.50	13.27	13.96	23.86	23.07
	20.89	21.49	21.58	26.60	24.39
	26.57	32.74	31.23	2.28	1.91
Sit	8.51	11.50	7.66	5.54	9.42
	7.83	8.44	10.47	9.18	6.07
	9.86	17.54	19.19	3.89	1.28
Sit	3.54	6.22	5.92	4.20	3.88
	4.24	7.41	7.03	4.43	4.38
	2.12	7.69	6.52	3.58	2.73
Pick	5.70	8.91	9.27	5.07	5.03
	8.20	11.99	12.66	7.84	5.30
	7.83	14.06	13.50	6.20	5.45
Pick	8.06	14.77	14.62	2.80	2.82
	6.00	5.04	5.04	6.96	6.96
	8.28	15.12	15.39	3.53	2.75

various cases of motion, the peculiarity is preserved in the form of the *ratio* between the maximum displacement and the antepenultimate to the maximum displacement.

Table 1 shows the displacements of the centroid, points near the head, and the points near the legs for various cases. All the displacements are measured in pixels along the image.

(The pixel values in the x -direction range from 120 to 160 and that in y -direction, the range is from 80 to 200)

It is observed that the ratio of the third row to the first row is the determining factor and is characteristic of every scenario.

4.1. Training of the neural network

Once all ratios for various scenarios are determined, it is necessary to find the criteria to identify various types of motions. Due to the large

Table 2. Decisions on severity values.

Severity value	Probable motion
0.8–1.0	Very high severity — fall or slip
0.6–0.8	Moderately high severity — jump
0.4–0.6	Moderate severity — pick or run
0.2–0.4	Low severity — sit
0–0.2	Non-severity — walk

amount of data and the variation among different motions of the same type, it is very difficult to visually identify the criteria. Consequently, artificial neural networks are used in the present work.

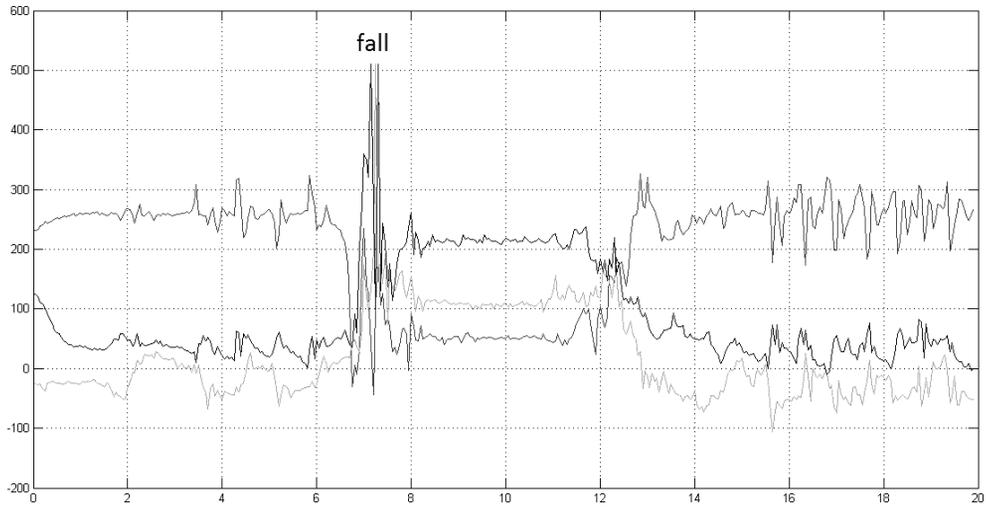
This paper uses radial basis functions networks (RBFN). The inputs to the network are the ratio values as in Table 1, and the outputs are the severity values assigned for various scenarios. For example, the fall cases have severity values as high as 0.9, and the sit cases have severity values close to 0.2. Thus, based on the severity value obtained from the network for any given input, the category of motion can be identified (see Table 2).

The output of the network is the corresponding severity value. The decision-making process is based on the classification as indicated previously.

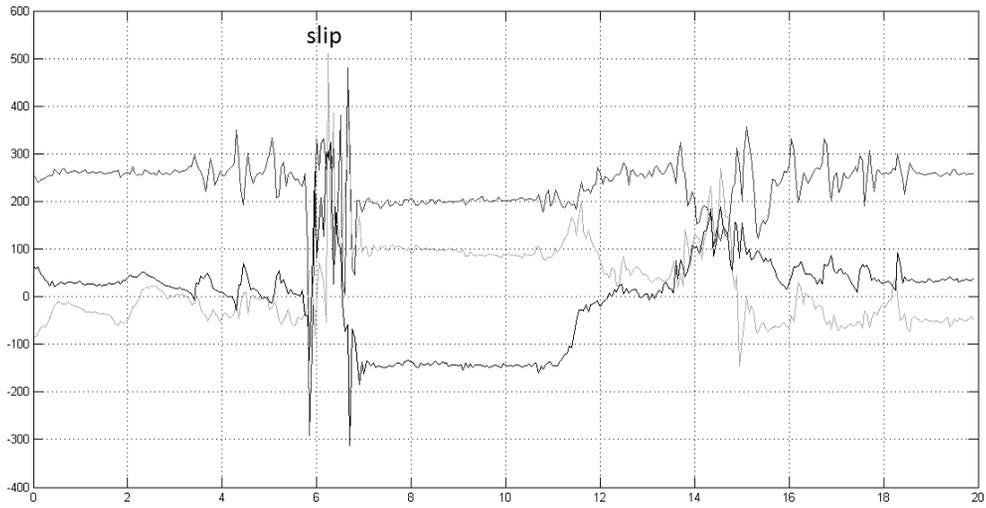
4.2. Use of wearable body sensors

The second approach presented in the present paper for the problem of detecting abnormal motions is the use of wearable body sensors — accelerometers and gyroscopes. Specifically, the linear and angular accelerations of different body parts may be used to determine the category of motion.

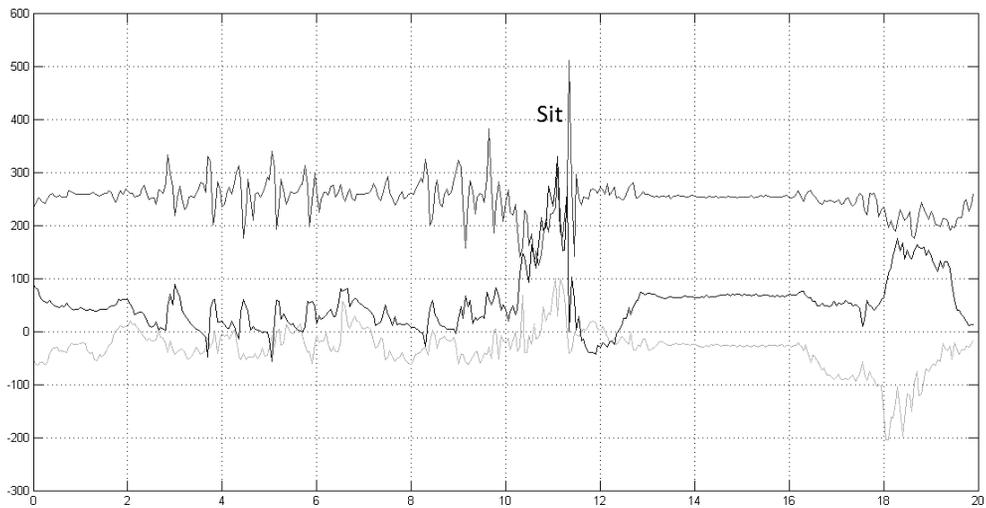
One problem with body sensors is that the system is not human-independent. Unlike vision systems, which are human-independent, the subject has to wear body sensors continuously. People with poor memory may forget to wear the sensors. The vision system approach is more appropriate in this regard. However, under inadequate lighting conditions, particularly at night, a vision system may not be effective, as the quality of the images will be deteriorated. The use



(a) "Fall" case

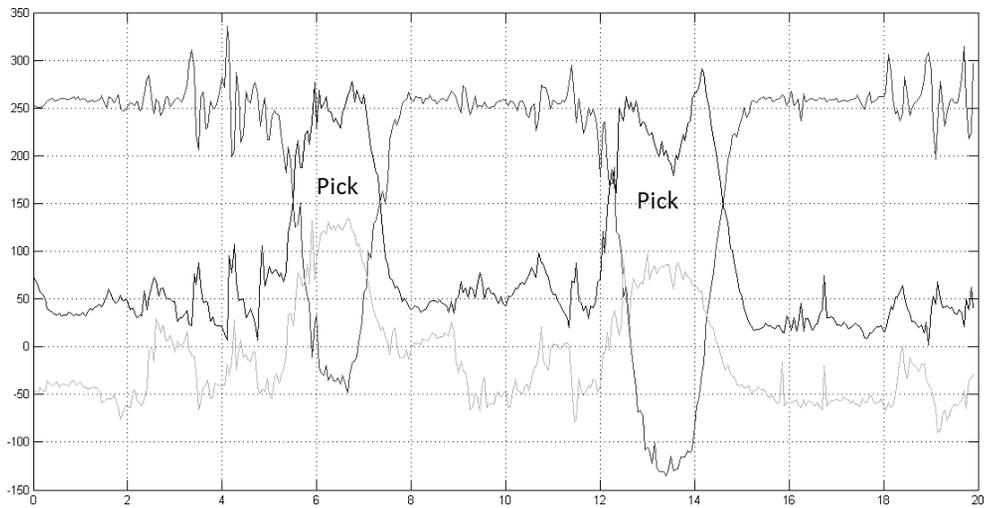


(b) "Slip" case

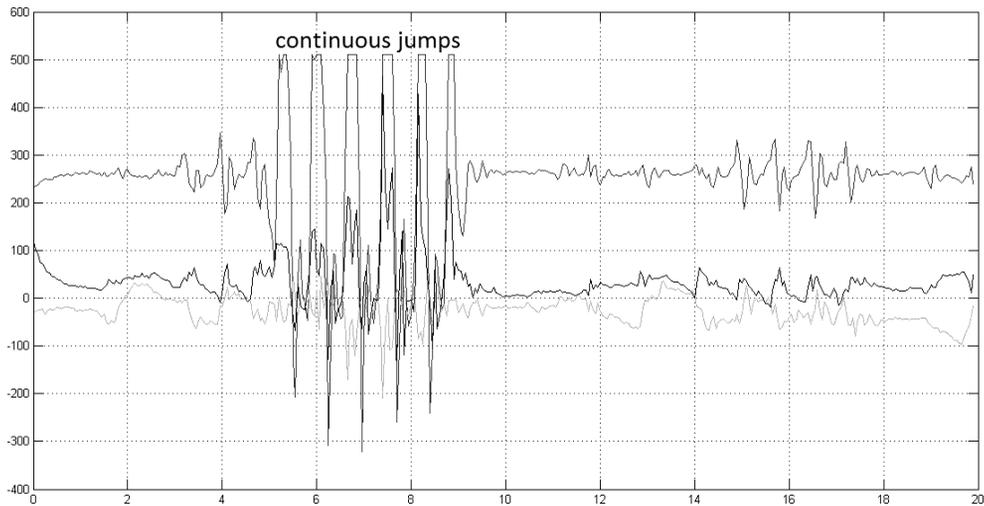


(c) "Sit" case

Fig. 7. Accelerometer response along each axis during various motion scenarios.



(d) "Pick up" case



(e) "Jump" case

Fig. 7. (Continued)

of wearable body sensors can become critical in such situations.

Abnormal motion detection using a single 3-axis accelerometer is implemented in the present work and its performance is evaluated for various cases of motion. The sensor is mounted rigidly on the chest of the person. Rigid mounting is realized by using magnetic tapes and skin-tight clothes.

The specific sensor used in the present work is ADXL345 — a 3-axis digital (I2C) accelerometer from Sparkfun. The device is capable of transferring data through Bluetooth and other

wireless protocols. Experiments are conducted efficiently and easily using the setup.

Specifically, experiments are carried out for the motion cases: fall, slip, jump, walk, pick-up, and sit. In some cases, the motion trajectories are found to have some peculiar characteristics. Using these characteristics, the category of motion can be easily identified. Note that the device has been configured to work in the $\pm 2g$ sensitivity range. The value on the Y-axis is the ADC value subjected to the maximum of 512 on either side. The X-axis data represent time in seconds.

Figure 6 shows the responses from the accelerometer sensors along each axis. The different shades in the graphs along y -axis represent data corresponding to different axes: X -axis, Y -axis and the Z -axis of the accelerometer. The x -axis of the graph represents time elapsed in seconds.

It is seen that for cases like pick-up and jump, there occurs a strong characteristic through which such motions can be identified. Cases like fall and slip can also be partially identified as there occurs a sharp change in the acceleration data. However, these cases lack uniqueness because of the fact that the human body can fall in any orientation and thus the component of acceleration due to gravity does not act in the same way along each axis for different scenarios. It is very difficult to identify cases like sit as the acceleration does not change much, owing to the same body orientation. This method can be further improved by the use of an accelerometer at one of the thighs and a gyroscope.

4.3. Comparative analysis

The main aim of the present experiments is to provide a comparative study of various methods to detect abnormal motion and to show how to use the combined results to improve the overall efficiency. Table 3 provides a comparative analysis of the approaches presented in the paper.

The comparative study enables us to use the two approaches simultaneously to detect

unusual activities and to assign proper weights to individual approaches for non-erroneous detection of abnormal motion. Such fusion of multiple sensors has also been proposed in Doukas *et al.* [2009].

5. Conclusion

This paper presented a method for detecting abnormal motion in real time using a computer vision system. The method was based on the modeling of human body image, which took into account both orientation and velocity of prominent body parts. A comparative study was made of this method with other existing algorithms based on optical flow and the use of accelerometer body sensors. From the real time experiments conducted in the present work, the developed method was found to be efficient in characterizing human motion and classifying it into basic types such as falling, sitting, and walking. The method used a Radial Basis Function Network (RBFN) to compute the severity coefficient associated with the type of motion, based on experience. The efficiency of the method was found to be nearly 80%. A shortcoming of the method was in the foreground extraction, which was essentially devised for indoor detection, as it was based on the assumption that the largest moving object in the environment was a human. For implementation in other (outdoor) environments, the scheme needs to be appropriately modified. For example, the use of skin color information along with the foreground

Table 3. Summary of comparative evaluation.

Cases	Using optical flow	Using body modeling	Using body sensors
Speed	Time consuming	Sufficiently fast	Very fast
Fall detection	Poor	Very efficient	Sufficiently efficient
Walk detection	Very efficient	Very efficient	Very efficient
Sit detection	Inefficient	Sufficiently efficient	Inefficient
Jump detection	Inefficient	Very efficient	Very efficient
Slip detection	Inefficient	Very efficient	Sufficiently efficient
Pick-up	Inefficient	Sufficiently efficient	Very efficient
Availability of data	Poor in occlusions, dark environment	Poor in occlusions, dark environment	Always
Human inter-dependency	Maximum	Maximum	Minimum

identification and extraction may be appropriate. Finding moving objects (or bounding boxes) that contain human skin color or human infrared data will resolve the problem. Though the method can detect basic human motions, there are some serious abnormal motions such as, the motion of the hand in a heart attack — that cannot be tracked, as it requires motion information on various body parts obtained separately and then analyzed as a whole. The method must be modified for precise detection of individual body parts when these are visible within the camera frames.

Acknowledgment

This work has been supported by the MITACS program of Canada and the Tier 1 Canada Research Chair of Mechatronics and Industrial Automation.

References

- Debard, G., Deschodt, M., den Bergh, J. V., Dejaeger, E., Milisen, K., Goedemé, T., Tuytelaars, T. and Vanrumste, B. [2009] “Validation of person identification for camera based fall detection,” in *4th Annual Symposium of the Benelux Chapter of the IEEE Engineering and Biology Society (EMBS)* (Enschede, Netherlands).
- Derpanis, K. G. [2005] “The computation of optical and affine flow”.
- Diraco, G., Leone, A. and Siciliano, P. [2010] “An active vision system for fall detection and posture recognition in elderly healthcare,” in *Proc. IEEE Design, Automation and Test in Europe (DATE 2010)*, 1536–1541.
- Doukas, C., Maglogiannis, I., Katsarakis, N. and Pneumatikakis, A. [2009] “Enhanced human body fall detection utilizing advanced classification of video and motion perceptual components,” *Artificial Intelligence Applications and Innovations III*, eds. Iliadis, L., Vlahavas, I., Bramer, M., “International Federation for Information Processing,” (Springer, Boston, MA) Vol. 296, pp. 185–193.
- Doulamis, A., Doulamis, N. and Kalisperakis, I. and Stentoumis, C. [2010] “A real time single-camera approach for automatic fall detection,” *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* (Commission V Symposium, Newcastle upon Tyne, UK), Vol. XXXVIII, Part 5, pp. 207–212.
- Foroughi, H., Shakeri Aski, B. S. and Pourreza, H. [2008] “Intelligent video surveillance for monitoring fall detection of elderly in home environments,” in *Proc. 11th International Conference on Computer and Information Technology* (Khulna, Bangladesh), pp. 219–224.
- Hazelho, L., Han, J. and de With, P. H. N. [2008] “Video-based fall detection in the home using principal component analysis,” in *Proc. 10th International Conference, Advanced Concepts for Intelligent Vision Systems* (Juan-les-Pins, France), pp. 298–309.
- Jansen, B. and Deklerck, R. “Context aware inactivity recognition for visual fall detection,” in *Proc. IEEE Pervasive Health Conference and Workshops 2006* (Innsbruck, Austria), pp. 1–4.
- Khan, M. J. and Habib, H. A. [2009] “Video analytic for fall detection from shape features and motion gradients,” in *Proc. World Congress on Engineering and Computer Science* (San Francisco, CA), Vol. II, pp. 1311–1316.
- Kinsella, K. and He, W. [2009] *An Aging World: 2008*, U.S. Census Bureau, International Population Reports, U.S. Government Printing Office, Washington, DC, P95/09-1.
- Lin, C. W. and Ling, Z. H. [2007] “Automatic fall incident detection in compressed video for intelligent homecare,” *Proc. 16th IEEE International Conference on Computer Communications and Networks (ICCCN 2007, Honolulu, Hawaii, USA)*, pp. 1172–1177.
- Liu, C. L., Lee, C. H. and Lin, P.-M. [2010] “A fall detection system using k -nearest neighbor classifier,” *Expert Systems with Applications, Science Direct*, Vol. 37, 7174–7181.
- Luštrek, M. and Kaluža, B. [2009] “Fall detection and activity recognition with machine learning,” *Informatica*, Vol. 33, 205–212.
- Miaou, S. G., Sung, P. H. and Huang, C. Y. [2006] “A Customized human fall detection system using omni-camera images and personal information,” in *Proc. 1st Distributed Diagnosis and Home Healthcare Conference (D2H2, Arlington, VA)*, pp. 39–42.
- Nait-Charif, H. and McKenna, S. [2004] “Activity summarisation and fall detection in a supportive home environment,” in *Proc. 17th IEEE International Conference on Pattern Recognition (ICPR’04, Cambridge, UK)*, pp. 323–326.
- Rougier, C. and Meunier, J. [2006] “Demo: Fall detection using 3D head trajectory extracted from

- a single camera video sequence,” *1st International Workshop on Video Processing for Security* (VP4S-06, Quebec City, Canada).
- Rougier, C., Meunier, J., St-Arnaud, A. and Rousseau, J. [2008] “Procrustes shape analysis for fall detection,” *The Eighth International Workshop on Visual Surveillance-VS2008* (Marseille, France).
- Tehrani, M. A., Kleihorst, R., Meijer, P. and Spaanenburg, L. [2009] “Abnormal motion detection in a real-time smart camera system,” in *Third ACM/IEEE International Conference on Distributed Smart Cameras* (Como, Italy), pp. 1–7.
- Williams, A., Ganesan, D. and Hanson, A. [2007] “Aging in place: Fall detection and localization in a distributed smart camera network,” in *Proc. 15th ACM International Conference on Multimedia* (Augsburg, Bavaria, Germany), pp. 892–901.
- Yu, X., Wang, X., Kittipanya-Ngam, P., Eng, H. L. and Cheong, L. F. [2009] “Fall detection and alert for ageing-at-home of elderly,” *Proc. International Conference on Smart Homes and Health Telematics* (ICOST’09, LNCS 5597, Springer-Verlag, Berlin, Germany), pp. 209–216.
- Zhong, H., Shi, J. and Visontai, M. [2004] “Detecting unusual activity in video,” in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (CVPR 2004, Los Alamitos, California), pp. 819–826.

Biography

Mayank Baranwal received his Bachelor’s degree from the Indian Institute of Technology, Kanpur in 2011 in Mechanical Engineering. He is now pursuing his Master’s from the University of Illinois Urbana-Champaign (UIUC) in the Department of Mechanical Engineering. His interests are mainly in the field of Control Systems, Motion Planning and Image Processing applications. He worked on this topic during Summer Internship at the University of British Columbia (UBC). He has also served as a member of the team dedicated towards the design and development of India’s nano-satellite *Jugnu*, where he worked in the ADCS (Attitude Determination and Control Systems) subsystem. He was awarded the General Proficiency Medal in his final year for securing the highest rank in his department.

M. Tahir Khan received his Bachelor’s degree from the NWFP University of Engineering and Technology, Pakistan in 1997 in Mechanical Engineering. He received his Master’s degree from the University of New South Wales, Australia in 1999 in Mechatronics. Later on, he joined the University of British Columbia (UBC), Canada where he received his Ph.D. in 2010 under Dr. Clarence W de Silva. Since then, he has been appointed as the Lab Manager and Post Doctoral Fellow in the UBC Industrial Automation Laboratory. His research interests include Industrial Robotics, Intelligent Control Systems and Neurofuzzy and Artificial Immune Systems. He has received many awards including the Ph.D. Scholarship from the Higher Education Commission, Government of Pakistan. He has publications in various international conferences and journals including the prestigious ASME/IEEE International Conferences.

Clarence W. de Silva is a Professor of Mechanical Engineering and occupies the Tier 1 Canada Research Chair Professorship in Mechatronics & Industrial Automation at the University of British Columbia, Vancouver, Canada. A Professional Engineer (P.Eng.), he is also a Fellow of: ASME, IEEE, Canadian Academy of Engineering, and Royal Society of Canada. He has authored 19 books and over 400 papers, half of which are in journals. He has received many awards including the Paynter Outstanding Investigator Award and the Takahashi Education Award of ASME Dynamic Systems and Control Division; Killam Research Prize; and Outstanding Engineering Educator Award of IEEE Canada. He has served as Editor/Associate Editor of 14 journals including ASME and IEEE transactions; and as Editor-in-Chief of *International Journal of Control and Intelligent Systems*. He received Ph.D. degrees from Massachusetts Institute of Technology, USA (1978) and the University of Cambridge, UK (1998), and an Honorary D.Eng. from the University of Waterloo, Canada (2008).