Fall Motion Detection with Fall Severity Level Estimation by Mining Kinect 3D Data Stream

Orasa Patsadu[1], Bunthit Watanapa[1], Piyapat Dajpratham[2], and Chakarida Nukoolkit[1]
[1]School of Information Technology, King Mongkut's University of Technology Thonburi, Thailand
[2]Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand

Abstract: *This paper proposes an integrative model of fall motion detection and fall severity level estimation. For the fall motion detection, a continuous stream of data representing time sequential frames of fifteen body joint positions was obtained from Kinect's 3D depth camera. A set of features is then extracted and fed into the designated machine learning model. Compared with existing models that rely on the depth image inputs, the proposed scheme resolves background ambiguity of the human body. The experimental results demonstrated that the proposed fall detection method achieved accuracy of 99.97% with zero false negative and more robust when compared with the state-of-the-art approach using depth of image. Another key novelty of our approach is the framework, called Fall Severity Injury Score (FSIS), for determining the severity level of falls as a surrogate for seriousness of injury on three selected risk areas of body: head, hip and knee. The framework is based on two crucial pieces of information from the fall: 1) the velocity of the impact position and 2) the kinetic energy of the fall impact. Our proposed method is beneficial to caregivers, nurses or doctors, in giving first aid/diagnosis/treatment for the subject, especially, in cases where the subject loses consciousness or is unable to respond.*

Keywords: *Kinect 3D data stream, fall motion detection, fall severity level estimation, machine learning, smart home system.*

Received August 25, 2015; accepted December 27, 2015

1. Introduction

Falls are a leading cause of injury, disability, and accidental death. Consequential injuries could lead to worsen outcomes such as fracture [18] and prolonged hospitalization [41]. An effective fall motion detection system should be able to monitor a subject's movement [1, 45], promptly report any fall event with high accuracy and have the ability of estimating the fall severity level. This is crucial for the caregiver or healthcare personnel, especially when the subject lives alone and is unable to respond. This motivated our research work in two areas: robust real-time fall motion detection during Activity of Daily Living (ADLs) [35] and a novel framework of fall severity estimation.

For the real-time fall detection area, there are many studies of fall motion detection systems equipped with traditional hardware such as mobile, sensor, 2D camera, microphone, etc. Nevertheless, most of them are too complex, intrusive, lack of privacy, or are expensive to be practically deployed for home use. Previous researches also have two common technical issues which are the background ambiguity and the unavailability of incident information for estimation of fall severity level.

While the fall detection method is an active domain of research, fall severity estimation seems to be a laggard one. There are quite a number of measuring schemes for the Injury Severity Scale (ISS), e.g., Head Injury Criterion (HIC), Head Injury Models (HIMs), and Abbreviated Injury Scale (AIS), but mostly they are devised for discretizing the level of trauma in a car crash, sport, or pedestrian accident [31].

In response to these challenges, we propose an integrative intelligent system extending the framework of our previous work [38] for:

1. Online fall motion detection based on time-dependent body-joint data using a Kinect camera.
2. Supporting estimation of fall severity level.

For the first integral ability, we studied pattern recognition for fall detection from a huge amount of data obtained from Kinect. The proposed system relies on a machine learning algorithm which is very capable in distinguishing the various real-life fall motions from fall-like motions which may be a losing balance or ADLs such as sitting down abruptly or performing exercise on the floor. In this study, we compared two of the most prominent techniques which are Multilayer Perceptron (MLP) and Support Vector Machine (SVM) to gain technical insights into their performance. Moreover, advantages of Kinect enable us to devise a method that can resolve the background ambiguity of the human body, especially the case of a subject sitting or lying on furniture as troublesome in previous research [32]. Kinect's natural user interface [6] properties also eliminate the requirement of the subjects to wear any device on their body, which can be very cumbersome or obtrusive.

In addressing the estimation of fall severity level, this research proposes a framework for fall severity level estimation called Fall Severity Injury Score

(FSIS). The measure is based on two basic physics principles: the velocity of the impact position and the kinetic energy of the fall impact. The main risk areas of the body are defined as head, hip, and knee positions as reported by medical researches [10, 26, 30, 36, 42]. The severity level is discretized by the statistically variability analysis of inter-quartile values that is well known for being robust to the problem of outliers in the training data [9].

This paper is organized as follows: section 2 presents related works; section 3 describes the methodology of our proposed system; section 4 presents the experimental results and discussions; finally, section 5 presents the conclusion and future work directions.

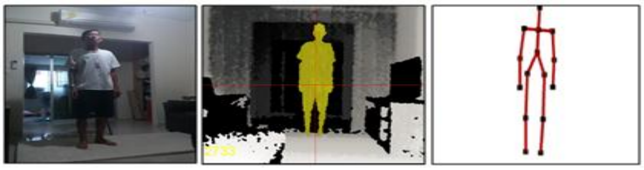
2. Related Work

The review of the two integral parts of the proposed model is presented in the order of processing stages: the fall motion detection and followed by the fall severity estimation.

2.1. Kinect and Fall Motion Detection

Considering the fall detection system in terms of time sequential data as input, most of previous works use Red, Green, and Blue (RGB) [4]. RGB requires large amount of low-level vision tasks like foreground-background segmentation, boundary detection and probably lead to a privacy issue. As technology having evolved, the depth image takes its place by the advantages of less processing and abstraction that automatically eliminates the privacy problem [27, 29, 32, 43]. However, the use of depth of image can be unreliable because of background ambiguity in the case of the subject sitting or lying on furniture. Body joint positions detection technology can overcome all mentioned issues and luckily that it is now available as a built-in feature inside commercial devices such as Kinect. Figure 1 shows the visual difference for the three technologies: RGB, depth image and body joint positions.

Many researchers are now turning to body joint position as a reliable source of input data for movement and fall analysis using specific positions such as head center, torso or other body joint positions [5, 16, 24, 34, 40, 44]. Based on our preliminary study and experiment, when compared to other positions, the torso position alone can achieve best performance for fall detection with a low degree of freedom of movement in all views and near the pelvis region. The very importance of torso is also confirmed in many reports [7, 21, 23].



a) RGB. b) Depth image. c) Skeleton. (body joint positions)

Figure 1. Different types of data for fall detection.

2.2. Fall Severity Level Estimation

To our knowledge, the direct fall severity estimation method has never been explicitly defined to classify the probable seriousness of accidental fall that may happen in the daily routines. Such an information is valuable not only for first aid but also for supporting decision on subsequent care. Unfortunately, most of the existing models have been designed for situational context to assess trauma severity on a variety of body regions (i.e., head, hip, knee, face, neck, abdomen, and spine) [3] in several accidents such as falls from heights [2, 25, 46], falls from a bed [8, 42], falls of infants from an accident [13], and car crashes [3].

Generally, the severity of injury can be evaluated based on key influencing factors, e.g., the height of the fall, post velocity or acceleration of impacted position, and kinetic energy of the fall [2, 3, 8, 13, 25, 33, 42]. Some examples of existing models that measure injury severity are AIS [3], HIC [31, 42, 46], and HIMs [13]. For AIS, this model assesses injury severity on body more than one area based on human judgment, e.g., a physician. AIS is based on a scale of one to six, one being a minor injury and six being maximal. AIS grades all injuries of each body area and use the most severe position as a representative of severity level to represent the threat to life associated with the injury. For HIC, the model is intended to judge the head injury risk based on the acceleration. HIC does not provide the interval scale that enables comparison about kind and severity of eventual injuries, but it gives an initial orientation for an estimation of the general injury risk. Based on HIC, HIMs assesses severity of childhood head injuries from impact after a fall from a height such as jump or throw up from an elevated surface which uses maximum of acceleration (peak g).

Essentialness of the model for scaling severity of falls in ADLs is to take into account the discernable levels of morbidity using online data that, without intention to replace the specialist's judgment, warns and supports caregivers to take actions on first aids or assist specialists for deciding on subsequent care. Unfortunately, as described above, most of the existing models are off-line post-fall analyses and focusing on the seriousness of injury, threat to life or curing process.

Next, we discuss in details the designed methodology for systematically determining a fall and estimating the severity in online mode.

3. Methodology

In this section, we describe our integrative model to detect a falling of a subject and estimate fall severity level. The system is divided into four phases: data preprocessing, real-time segmentation, fall detection, and fall severity level estimation as shown in Figure 2.

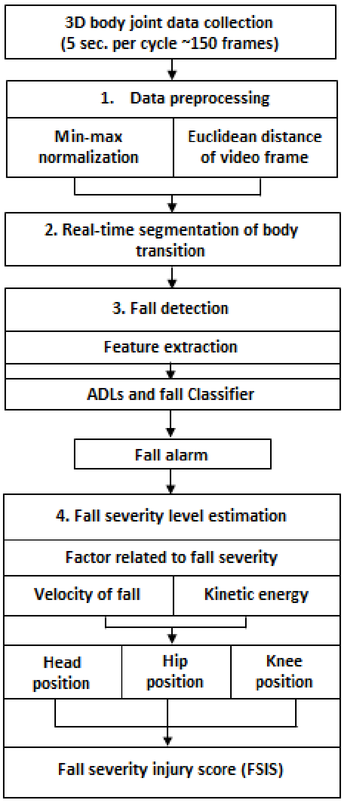


Figure 2. Integrative model of the proposed system.

Each phase of our proposed system is described in detail, next.

3.1. Data Preprocessing

In holistic view of the preprocessing, a series of 150 frames (or, a clip) of streamed 3-dimensional data of selected body joints, represented by a set of vector (X, Y, Z), is fetched and normalized. Next, a frame-to-frame Euclidean distance of the body joint position is then calculated for being input in the next phase of segmenting the body transition. Detail is presented next.

3.1.1. Selection of Body Joint Positions

We performed preliminary experiments to select representative body joint positions which could optimize both accuracy and processing time of the fall detection. We compared three alternatives:

1. All body joint positions [24].
2. Head, shoulder, torso, hip and knee position [39].
3. Torso only position.

Based on the result of our preliminary experiment (Table 1), the alternative of using only the torso position was chosen as it provides a slightly higher accuracy and a significantly faster run time when compared to the other choices.

Table 1. Comparative performances of fall detection using different set of body joint positions.

body joint position(s)	Fall motion detection Accuracy	processing time (seconds)
all body joint positions [24]	99.70%	7
head, shoulder, torso, hip and knee position [39]	99.60%	4
only torso position (our proposed feature set)	99.97%*	2*

3.1.2. Data Normalization

Due to the fact that people are with varied body sizes, the learning process for generalized model of fall detection needs the normalized measure of body movement to effectively detecting fall of any persons. For each clip, says clip *A*, each dimensional value *(d)* in the whole clip *{d}* of a body joint vector in a frame will be transformed into normalized values *(d')* in the range of [0, 1], where 0 represents minimum normalized value of that dimension in the clip and 1 represents the maximum. The min-max normalization scheme was implemented here as depicted in Equation (1) [22].

$$d' = \frac{d - \min(p)}{\max(p) - \min(p)}, \forall d \in \{d\} \in \text{clipA} \quad (1)$$

Where $\min(p)$ = minimum value of set *{d}*.
$\max(p)$ = maximum value of set *{d}*.

3.1.3. Data Transformation

In the second step, we map the normalized body-joint positions into a time series of Euclidean distance between two consecutive Kinect video frames. The larger distance the human subject moves his body, the larger value the Euclidean distance will be. The calculation of Euclidean distance [28] is depicted in Equation (2).

$$d(p,q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2 + (p_z - q_z)^2} \quad (2)$$

Where p = (p_x, p_y, p_z) and q = (q_x, q_y, q_z), are the vector (X, Y, Z) of torso position, head position, hip position, and knee position at time *t* and time *t+1* respectively.

3.2. Real-time Segmentation of Body Transition

For effective segmentation, we propose a framework that decomposes a fall event into three sequential phases:

1. The pre-fall phase (stably continuing daily life motions).

2. The body transition phase.
3. The post-fall phase (sustaining inactivity or recovered) [35] as seen in Figure 3.

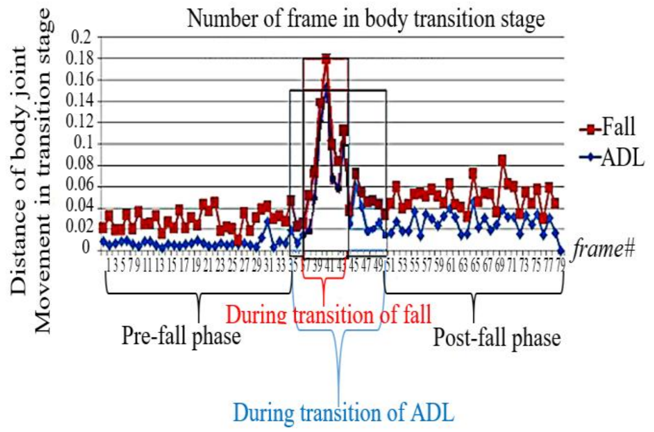


Figure 3. Comparison of number of frames in during body transitionn stage (Fall vs. ADL).

The critical information is in the transition phase. From our eye inspection on Kinect videos in our experiment and other videos [27], mostly, the body transition phase in fall motion occurred within 1-2 seconds, while, that in ADL motions (for example transition from sit to stand) occurred within 3-8 seconds depending on varied conditions, e.g. age, muscle strength, etc., of a subject. This observation agreed with previous findings in the medical field [47]. Based on these findings, hence, we continually process the fall detection using a shifting cache of 5 seconds of time series of torso position data.

K-means clustering model [19] configured with the number of clusters=3 is applied to detect the boundaries of the transition phase, denoted as $[t_s, ..., t_f]$. t_s states the possible start frame of transition and t_f represents the finish frame of transition.

Relying on the obtained boundaries, we can derive three unique features ($f_{duration}$, $f_{average\ distance}$, f_{slope}) for being inputs of fall motion detection as shown in Figure 4.

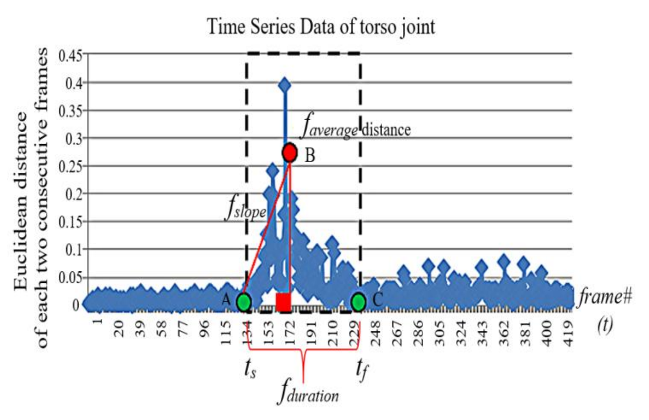


Figure 4. Three extracted features based on the segmented boundary.

$f_{duration}$ denotes the duration of the transition phase, which is distance from point *A* to point *C* in Figure 4. This can be measured by counting number of frames within the boundary of the transition phase as depicted in Equation (3).

$$f_{duration} = count\ m;\ \forall\, m \in [t_s, ..., t_f] \qquad (3)$$

Where *m* is the number of frame within the transition phase

$f_{average\ distance}$ denotes the average distance, or mean value of the frame-to-frame distance $(t_s,\ldots,t_f)$ of torso joint position, denoted as *B* in Figure 4, as depicted in Equation (4).

$$f_{average\ distance} = \frac{\sum_{i=s}^{f} \text{distance}(t_i)}{f_{\text{duration}}} \qquad (4)$$

Lastly, f_{slope} denotes the derived ratio of $f_{average\ distance}$ over the half span of transition phase. The obtained figure represents the speed of change in Euclidean distance of torso and can be calculated [12] using Equation (5).

$$Slope = \frac{2 f_{average distance}}{f_{duration}} \qquad (5)$$

3.3. Fall Motion Detection Based on a Machine Learning Approach

Relying on the real-time segmentation of body transition as illustrated in 3.2, fall motion detection can be performed as described next.

To gain insights into the machine learning performance, we compared two well known classifiers in performing fall motion: MLP [20] and SVM [14].

MLP is a kind of neural network which is a feedforward artificial neural network model that maps a set of input data onto a set of appropriate outputs. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. MLP utilizes a supervised learning technique called back propagation for training the network as seen in Figure 5.

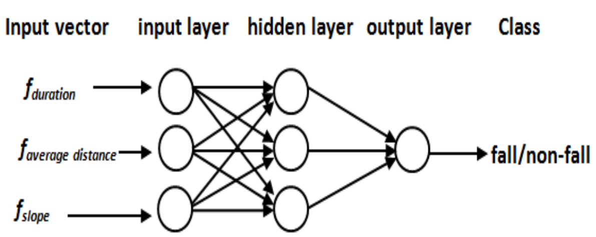


Figure 5. Architecture of MLP for fall motion detection.

Figure 5 shows the optimal architecture of the MLP classifier for fall motion detection, achieved from our empirical experiment. There are three layers (input layer, hidden layer, and output layer) with 3, 3, and 2 nodes, respectively. We set the learning rate at 0.3 and momentum at 0.2. The input layer consists of three extracted features ($f_{duration}$, $f_{average\ distance}$, f_{slope}). The hidden layer is with three nodes (as suggested by [11]: input +class/2). The output node suggests two possible values: fall or non-fall.

SVM is a novel approach that is capable of classifying linear or nonlinear data. The basic discipline of SVM is to compute a hyperplane, defined by support vectors, between each class and the rest in a way that the margin between two classes is maximized. We find a hyperplane $f(x)$, which makes category "-1" of y fall into the range of $f(x) < 0$ and category "+1" of y fall into the range of $f(x)>0$. Thus, we can distinguish the categories according to the sign of $f(x)$ as shown in Equation (6).

$$f(x) = w^T x + b \tag{6}$$

Where w is the normal vector of this hyperplane

x is the input vector

$-b\|w\|$ is the distance from origin perpendicular to hyperplane

Maximum margin and minimum square error are derived to solve SVM as seen in Equation (7).

$$E(w,b) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{m}\alpha_i(y_i - f(x_i))^2 \tag{7}$$

Where $\alpha = \{\alpha_1,...,\alpha_m\}$ is coefficient of Lagrange and $\alpha_i > 0, i = 1,2,..., m$

Our research has been tested with various kernel functions (linear, polynomial with a default order of 3, and a radial basis function with a default scaling factor of 1). Among these, radial basis function, as shown in Equation 8, yielded the best performance and had been used as kernel function in our experiment.

$$k(x, x_j) = \exp\left(-\gamma\|x - x_j\|^2\right) \tag{8}$$

Where γ is radial's size.

3.4. Fall Severity Level Estimation

Once a fall is detected, the estimation of severity level of injury on three key body joints: head, hip and knee will be preceded. Empirically, the importance of these three body positions was agreed by a domain expert (Rehabilitation Medicine) and literature review. To show a few, a report from the Center for Disease Control and prevention (CDC) [10] estimated that the number of hip fractures, a serious fall related injury, will rise from 350,000 admissions per year to over 500,000 by 2040. Magaziner *et al*. [30] indicated that hip fractures are the major category of injury produced by falls with 87% of all fractures occurring in the subject. The second highest ranked injury from falls is head trauma [36, 42], and the next one is knee fracture [26].

The process of severity estimation comprises two main steps, namely, the analysis of factors related to fall severity and the scaling of the severity.

3.4.1. Investigating Factor Related to Fall Severity

As seen in Figure 6, the velocity of the impacted position and the kinetic energy of the fall impact [13] will be considered when investigating body injuries possibly sustained from falls.

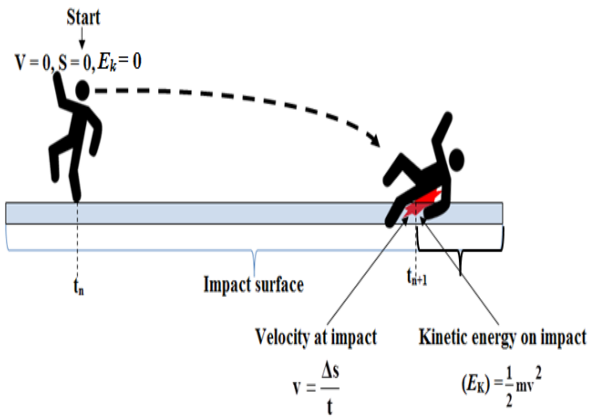


Figure 6. Velocity and Kinetic energy measured for fall severity.

The velocity *(v)*, in this context, identifies the speed of each joint position (i.e., head, hip, and knee position) hitting the floor [17]. The formulation is shown in Equation (9), where definition of notations refers to Figure 4.

$$v = \frac{\Delta s}{t} \tag{9}$$

Where Δs = distance from position *A* to position *C* (metre)

$t = f_{duration}$(second)

Kinetic energy *(E_K)* that is transmitted to and immersed by the impact surface and/or impact body [15] can be computed as defined in Equation (10).

$$\text{Kinetic energy } (E_K) = \frac{1}{2}mv^2 \tag{10}$$

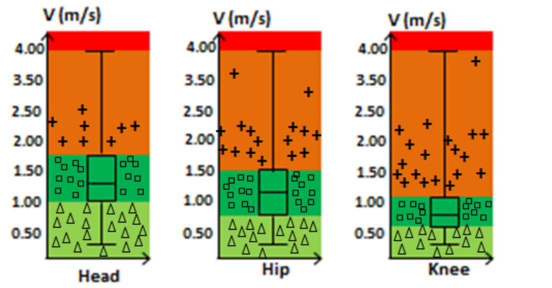
Where *m* = mass of the body (75 ± 35 kg, and height 165 ±15 cm of sample subject)

v = velocity (metre/second)

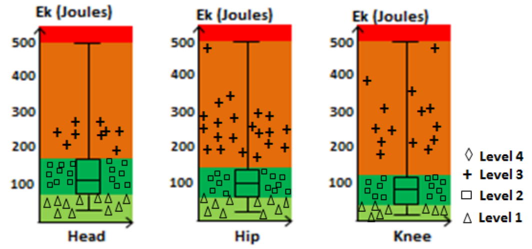
3.4.2. Division of fall severity level

We propose an anatomical-based coding system called FSIS as a framework to classify severity level of falls. The scales are derived from computed value of velocity or kinetic energy on selected body regions (head, hip, and knee).

FSIS scales, in this study, were trained using 1,320 fall video clips of eight subjects randomly selected from a total of 1,650 (10×11×15) fall video clips. Each individual subject performed a simulated fall without being specified characteristics of fall, which resulting in different types of actions performed at different speeds. The resulting FSISs are shown in Figure 7.



a) Velocity.



b) Kinetic energy.

Figure 7. Fall severity level (FSIS) on major body joint position.

At a glance, distribution of the velocity and kinetic energy of fall on each body joint position of head, hip and knee is used to divide fall severity levels into scales of four, using the inter-quartile. This approach is the fundamental method in statistics to learn the data dispersion with respect to the midpoint of the data set called the median. Using the inter-quartile analysis on the experimental data, Figure 7-a and Figure 7-b shows the Box-plot scales of velocity and kinetic energy on body joint positions of head, hip, and knee, respectively. The scale classifies fall severity into 4 levels. Level 1 is the value lower than Q1. Level 2 is the value inside the range of Q1, median, and Q3. Level 3 is the value beyond Q3 up to maximum. Level 4 is the value higher than the maximum.

Some interesting notes based on the inter-quartile analysis (exposed in Table 2) are as follows. Firstly, mean and range of velocity and kinetic energy of the head is noticeably higher than those of the hip and knee. This suggests the need of special care for the head when a fall incurred.

Table 2. Result of velocity and kinetic energy on head, hip, and knee position.

	velocity of fall (m/s)			Kinetic energy (joules)		
	Head	**Hip**	**Knee**	**Head**	**Hip**	**Knee**
max	3.93	3.86	3.83	492.76	467.7	426.98
Q3	1.66	1.51	1.19	88.71	76.73	71.56
median	1.33	1.22	0.93	58.14	47.81	47.3
Q1	1.05	0.93	0.7	34.86	27.55	24.63
min	0.02	0.02	0.02	0.01	0.01	0.01
mean	1.38	1.24	0.97	72.25	60.05	57.25
S.D.	0.47	0.45	0.39	57.54	48.96	42.14

Secondly, the experiment was conducted using only simulated falls, the most severe cases that was set up cannot be comparable with the serious injury in real life. For the time being, the training data, thus, assumes only the severity level 1 up to 3 representing mild, moderate and high chance of having injury as shown in Figure 8. In fall simulation, the subjects in level 3 had a bruise and pain from fall, although, they performed it on thick mats. Level 4 represents extremely high velocity or kinetic energy that can cause serious damage to body/bone in real life. An imaginary example of such a serious fall may be that the subject accidentally slips/falls while walking or standing and is unable to slow down the transition of fall. The body joint position hits directly to the hard floor. Whenever the data of real life fall is available, the fall severity scale can be easily adjusted reflecting the higher accurate measurement system.

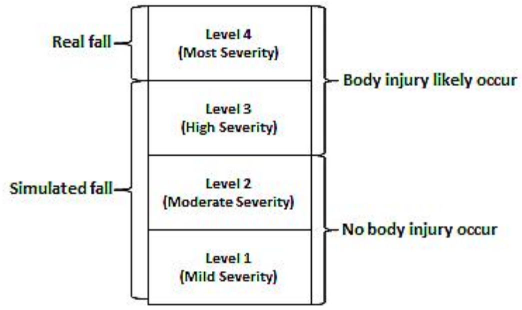


Figure 8. Four proposed levels of fall severity.

4. Experimental Results and Discussions

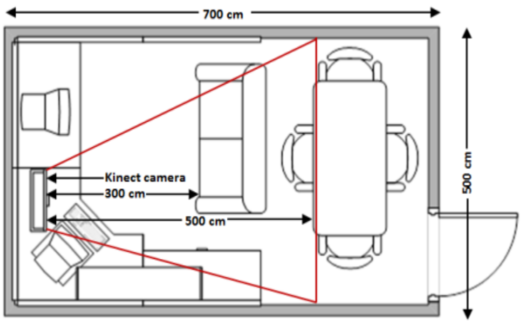
In this section, we explain the experiment setup and provide an evaluation of comparative performance of fall motion detection. In addition, we also examined the effectiveness of our proposed framework for fall severity level estimation. The case study demonstrated the effectiveness of the proposed system.

4.1. Experiment Setup and Dataset

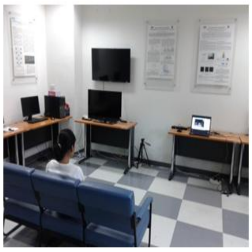
In our experiment, we established an indoor environment setting with a Kinect camera to track the movement of sample subjects in three different viewpoints *(side, frontal, and back view)* and generate a video stream with 640×480 resolution at the rate of 30 Frames Per Second (FPS). The Kinect device is set up at approximately 1 meter above the floor. The size of our room is approximately 5×7 m. as seen in Figure 9-a.

The depth sensor of Kinect uses the UserGenerator method of OpenNI [37] to extract the vectors (X, Y, Z) of fifteen body joint positions. We decided to use OpenNI in our research because it is a well-known reliable open source tool in 3D sensing middleware applications. Thus, developers can develop new routines to enhance the capabilities of the existing tool. Our research was implemented as a series of Kinect's video stream analysis modules. These modules, written in C#, read Kinect's depth information of a scene as a time series of fifteen body joint positions.

In our experiment, an additional dataset was collected, with approval of the Institutional Review Board (IRB) of our University. There were ten adult subjects (age 30 ± 8 years, body mass 75 ± 35 kg, and height 165 ±15 cm with an equal number of males and females of various weights and heights). They performed all activities on safety mats.



a) Room plan showing the room setup of our experiment.



b) Sitting on a sofa.

c) Falling down on the floor.

Figure 9. Experiment setup for fall detection system.

Our study simulated falls according to the definition of falls by Noury *et al*. [35] and Kwolek and Kepski [27], but used an extended version of scenarios on various falls and ADLs from the positions of standing, walking and sitting on a variety of seat types such as sofas, chairs with a backrest, and stools as seen in Figure 9-b and 9-c with different types of actions performed at different speeds as shown in Table 3.

Table 3. Various combinations of activities for motion detection.

Type of activities	Situations	Direction
Falls (11 types)	sit on chair with backrest/stool/sofa, stand, and walking	backward/forward/left/right
ADLs (18 types)	sit on chair with backrest/stool/sofa	backward
	lie on the floor/sofa	left/right
	bend down on the floor	left/right/forward

Each scenario was repeated 15 times per subject, the extensive number of video clips in our study was used to ensure low variance of our results and high reliability and accuracy of our system. Hence, there are a total of 4,350 (10x29x15) video clips in our experiment. To control the bias in model evaluation, our model was evaluated using 5-fold cross-validation, where the whole dataset was spilt into a training dataset for 1,320 falls and 2,160 ADLs from eight subjects, while data of the remaining subjects was used to test our model.

4.2. Discussion and Comparison of Performance of Fall Motion Detection

To gain an insightful experimental result, we conducted an experiment evaluating comparative performance of the MLP and SVM for being candidates for the fall detection algorithm on our dataset and another dataset. Furthermore, our proposed algorithm is compared with state-of-the-art approaches using depth of image.

The results of evaluating 870 video clips (330 falls and 540 ADLs) using MLP versus SVM are depicted in terms of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values as shown in Table 4.

Table 4. Results of evaluating our fall detection approaches.

	TP	TN	FP	FN
MLP	323	537	3	7
SVM	330	539	1	0

Table 4 exposes that both models are very good in detecting fall motions. SVM is better off, anyway. It has 0 FN which means all the simulated fall events in our experiment can be detected perfectly.

Moreover, we compare the performance of the two classifiers with three important performance measures: accuracy, recall, and precision as shown in Table 5.

Table 5. Measures for evaluating our fall detection approaches.

	Our dataset		dataset of Kwolek and Kepski [27]	
	MLP	**SVM**	**MLP**	**SVM**
Accuracy	99.77%	99.97%*	99.49%	99.72%
Recall	0.9974	1.000*	0.9967	0.9989
Precision	0.9988	0.9996*	0.9921	0.9945

*Perfect detect fall using SVM versus MLP

From the results reported in Table 5, fall detection using SVM achieves 99.97% accuracy, and 99.77% accuracy when using MLP. In addition, with SVM, our method achieves 1.00 recall and 0.9996 precision, While recall and precision of MLP achieve 0.9974 and 0.9988, respectively.

To reduce bias of our dataset, we further evaluated the proposed algorithm with another dataset [27]. The dataset of Kwolek and Kepski [27] contains 70 sequences of depth image using OpenNI (i.e., 30 falls and 40 ADLs). Data are recorded with 2 Kinect cameras (camera 0: parallel to the floor and camera 1: ceiling mounted). We used data on camera 0 for it setting is similar to our experiment. We have extracted torso position from depth image to test it with our algorithm. The obtained results are also shown in Table 5. The result shows that SVM also outperformed MLP.

Additionally, we have a set of activities on furniture to test the background ambiguity of the human body, especially in the case of a subject sitting or lying on furniture. All cases also detect body joint position and obviously separate between furniture and body joint positions. Therefore, the system will still be able to detect a fall or non-fall. If a fall occurs, it will raise an alarm. From our study, we found that an occlusion with the angle of human body may cause ambiguous detection using a single Kinect camera. To further enhance the performance of detection, future studies need to use multiple Kinect cameras to cover all areas of the subject's body occluded by furniture.

We further evaluated the proposed algorithm by comparing it with two state-of-the-art approaches as previously described in the literature [4, 32]. We re-implemented algorithms of these approaches on our data set. The result is shown in Table 6.

Table 6. Comparison of different approaches for fall detection.

	Proposed method		State-of-the-art	
	MLP	SVM	I. depth of the 3D [32]	II. depth of the 3D [4]
Accuracy	99.77%	99.97%*	97.27%	97.87%
Recall	0.9974	1.0000*	0.9727	0.9787
Precision	0.9988	0.9996*	0.9610	0.9699

* Perfect detect fall versus other methods

From Table 6, our method outperforms the two existing approaches. Most errors in the two approaches are cases of performing abrupt activities such as sitting or lying down. Our system is, therefore, very effective for applying in real-life situations because it is highly accurate, reliable, affordable, and easy to use. Our approach has been implemented in C# and is able to detect falls in real-time on an Intel core i5 Central Processing Unit (CPU) @1.7GHz and 4 GB RAM processing platform.

4.3. Demonstration and Evaluation of FSIS

Let us illustrate the interpretation of FSIS scale with a simulated fall that the subject lose body balance and fallen on the floor. The first position hitting the floor is the hip position followed by head and knee positions, respectively as seen in Figure 10.

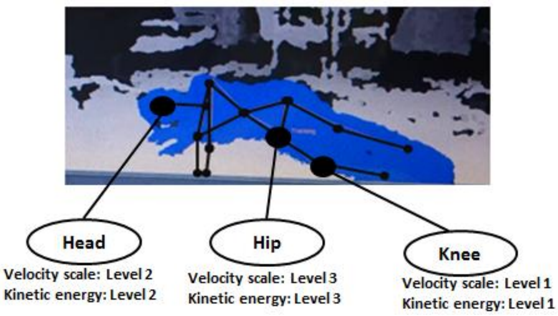


Figure 10. Sample case in our study.

In this case, just within a couple seconds after the transition phase of fall, the measured values of velocity and kinetic energy were transformed into FSIS scales and reported (possibly through LAN and social network) as shown in Figure 11. One can easily notice that only the hip position is worrisome since it is under level 3 severity. A simple, yet effective, visualization could help to grasp attention and self-explain urgency and risk of the situation. This information shall instantly urge and effectively support the caregiver or specialist to take properly care of the subject.

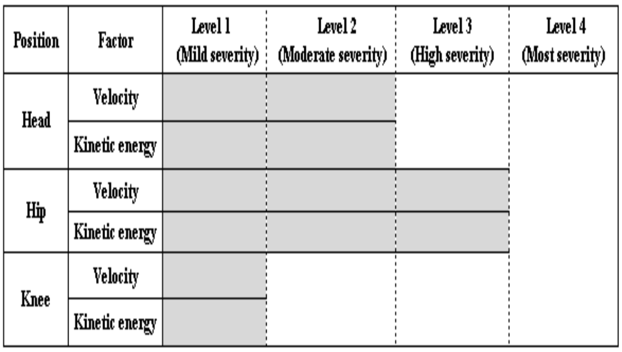


Figure 11. Summary report of fall severity level of major body joint position.

FSIS performance was evaluated using 330 video clips of various fall motions at different speeds and on a variety of seat types such as sofas, chairs with a backrest, and stools. The resulting performance of fall severity level estimation on body joint positions of head, hip, and knee, in terms of accuracy, precision, and recall are satisfactory with a range of 99.40%-99.97% accuracy, 0.9914-0.9997 precision, and 0.9930-1.000 recall. The evaluation is measured using the MLP model, 5-fold cross-validation with 1,320 fall video clips as the training dataset.

Initially, FSIS scale is beneficial to estimate severity after fall on each of body joint position for an assist due to after consequent fall. The subject should not hurry to get up immediately, but the subject should estimate body injury before obtaining help to reduce risk of supplement injury. With information of FSIS scale, the subject and the caregiver can understand the effect of fall and consequent injuries on each of body joint position. Furthermore, the FSIS scale is utilized for assessing head impact injury instead of the use of HIC and HIMs without having to wear additional devices on the body. The FSIS scale can assist AIS in diagnosis by providing supplement knowledge. So, physicians diagnose quick, accurate, and specific case.

Additionally, this keen insight of fall severity is important to caregivers and medical staff, especially, in case that the person loses consciousness and was unable to respond to caregivers. Fall severity information is used to support the decision making of caregivers in bringing the subject to see a doctor after a fall. For example, if a case of the fall severity level on the hip is in high or most severe level, the subject should see a doctor immediately to reduce risk of further injury. With this fall severity information, it would help a caregiver to make better decisions. In addition, if mild or moderate fall severity level occurs frequently, it means that the subject's body is starting to behave abnormally such as losing balance, having visual issues, or suffering side effects of medicine. Caregivers should report a doctor immediately to examine the causes of the fall and the treatment.

Finally, we have interviewed a rehabilitation physician. The physician accepts that the FSIS scale can quickly and accurately assist in diagnosis of the patient's condition in the initial stage, especially, in case of no incident information. As a result, future studies will extend the system by incorporating a decision support framework to support diagnosis, treatment, and rehabilitation of physician.

5. Conclusions and Future Work

In this paper, we propose an integrative model of fall detection and fall severity level estimation based on a live stream of joint positions obtained from Kinect camera. For fall detection, a set of three extracted features of a single joint of torso transitioning in 3 phrases, pre-fall, transition, and post-fall, is proposed as effective determinants for efficiently classifying fall out of the other ADLs. The boundaries of the three phases are continuously and automatically detected by a software agent equipped with K-mean algorithm. A set of experiments varying machine learning of choices (MLP and SVM) and data sets (owned and external sources) show very satisfactory performances. The best achievement of 99.97% accuracy with Recall of 1.0 is obtained when applying SVM. MLP shew a little laggard performance, although giving impressively effective result. The experiment also proves that the proposed method outperformed two existing models relying on depth image technology which is treated as state-of-the-art by far, esp. in terms of robustness to obscured cases, for example, a situation that subject is sitting or lying on furniture.

The last but not least, fall severity estimation is proposed to engineer the fall detection to be applicable and beneficial to the healthcare system. Velocity and kinetic energy are proposed as key features for estimating the severity and transformed into an FSIS framework that discretizes the severity into discrete level of seriousness of the fall on three important body parts: head, hip, and knee positions. This keen insight of fall severity is important to caregivers and medical staff, especially, in case the person lose consciousness and was unable to respond to caregivers. Thus, a subject could stay independent and safe within care facilities.

For future work, we will extend the system by incorporating a decision support framework that enables the intelligence of information in all 3 phases to assist specialists and physicians in diagnosing and deciding proper care even in the prolonged period after post fall phase. In addition, we plan to deploy our prototype system in real use at an elderly care facility and in a hospital setting.

Acknowledgements

This research project was supported by the Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission; Data and Knowledge Engineering Lab., School of Information Technology, King Mongkut's University of Technology Thonburi; and Rajamangala University of Technology Krungthep. We thank the SIT students and staffs for their invaluable assistance in setting up the experimental environment for the capturing sessions.

References

[1] Al-Najdawi N., Tedmori S., Edirisinghe E., and Bez H., "An Automated Real-Time People Tracking System Based on KLT Features Detection," *The International Arab Journal of Information Technology*, vol. 9, no. 1, pp. 100-107, 2012.

[2] Atanasijevic T., Savic S., Nikolic S., and Djokic V., "Frequency and Severity of Injuries in Correlation with the Height of Fall," *Journal of Forensic Sciences*, vol. 50, no. 3, pp. 1-5, 2005.

[3] Baker S., O'Neill B., Haddon W., and Long W., "The Injury Severity Score: a Method for Describing Patients with Multiple Injuries and Evaluating Emergency Care," *Journal of Trauma and Acute Care Surgery*, vol. 14, no. 3, pp. 187-196, 1974.

[4] Bevilacqua V., Nuzzolese N., Barone D., Pantaleo M., Suma M., D'Ambruoso D., Loconsole C., Stroppa F., and Volpe A., "Fall Detection in Indoor Environment with Kinect Sensor," *in Proceedings of International Symposium on Innovations in Intelligent Systems and Applications*, Alberobello, pp. 319-324, 2014.

[5] Bian Z., Hou J., Chau L., and Magnenat-Thalmann N., "Fall Detection Based on Body Part Tracking Using a Depth Camera," *Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 430-439, 2015.

[6] Blake J., *The Natural User Interface Revolution*, Natural User Interfaces in .NET, Manning, 2012.

[7] Bourke A., O'Brien J., and Lyons G., "Evaluation of a Threshold-Based Tri-Axial Accelerometer Fall Detection Algorithm," *Journal of Gait and Posture*, vol. 26, no. 2, pp. 194-199, 2007.

[8] Bowers B., Lloyd J., Lee W., Powell-Cope G., and Baptiste A., "Biomechanical Evaluation of Injury Severity Associated with Patient Falls from Bed," *Rehabilitation Nursing : the Official Journal of the Association of Rehabilitation Nurses*, vol. 33, no. 6, pp. 253-259, 2008.

[9] Box Plots, available at: http://onlinestatbook.com/2/graphing_distributions/boxplots.html, Last Visited, 2013.

[10] Center for Disease Control and Prevention, "Hip Fractures among Older Adult," available at: http://www.cdc.gov/ncipc/factsheets/adulthipfx.htm, Last Visited, 2012.

[11] Class Multilayer Perceptron, available at: http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/MultilayerPerceptron.html, Last Visited, 2014.

[12] Coordinate Geometry, available at: http://gmatclub.com/forum/math-coordinate-geometry-87652.html, Last Visited, 2013.

[13] Cory C., Jones M., James D., Leadbeatter S., and Nokes L., "The Potential and Limitations of Utilising Head Impact Injury Models to Assess the Likelihood of Significant Head Injury in Infants after a Fall," *Journal of Forensic Science International*, vol. 123, no. 2-3, pp. 89-106, 2001.

[14] Cristianini N. and Shawe-Taylor J., *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, 2000.

[15] Curran G., *Homework Helpers: Physics*, The Career Press, 2012.

[16] Dai X., Wu M., Davidson B., Mahoor M., and Zhang J., "Image-Based Fall Detection with Human Posture Sequence Modelling," *in Proceedings of International Conference on Healthcare Informatics*, Philadelphia, pp. 376-381, 2013.

[17] Elert G., The Physics Hypertextbook, available at: http://www.physics.info, Last Visited, 2012.

[18] Fischer I., Krauss M., Dunagan W., Birge S., Hitcho E., Johnson S., Costantinou E., and Fraser J., "Patterns and Predictors of Inpatient Falls and Fall-Related Injuries in a Large Academic Hospital," *Infection Control and Hospital Epidemiology*, vol. 26, no. 10, pp. 822-827, 2005.

[19] Han J. and Kamber M., *Data Mining Concepts and Techniques*, Morgan Kaufmann Publishers, 2011.

[20] Haykin S., *Neural Networks and Learning Machines*, Prentice Hall, 2008.

[21] Jacob J., Nguyen T., Lie D., Zupancic S., Bishara J., Dentino A., and Banister R., "A Fall Detection Study on the Sensors Placement Location and a Rule-Based Multi-Thresholds Algorithm Using both Accelerometer and Gyroscopes," *in Proceedings of International Conference on Fuzzy Systems*, Taipei, pp. 666-671, 2011.

[22] Jain Y. and Bhandare S., "Min Max Normalization Based Data Perturbation Method for Privacy Protection," *International Journal of Computer and Communication Technology*, vol. 2, no. 8, pp. 45-50, 2011.

[23] Kangas M., Konttila A., Winblad I., and Jämsä T., "Determination of Simple Thresholds for Accelerometry-Based Parameters for Fall Detection," *in Proceedings of the 29th Annual International Conference of Engineering in Medicine and Biology Society*, Lyon, pp. 1367-1370, 2007.

[24] Kawatsu C., Li J., and Chung C., "Development of a Fall Detection System with Microsoft Kinect," *in Proceedings of the 1st International Conference on Robot Intelligence Technology and Applications*, Gwangju, pp. 623-630, 2013.

[25] Kim K. and Ashton-Miller J., "Biomechanics of Fall Arrest Using the Upper Extremity: Age Differences," *Journal of Clinical Biomechanics*, vol. 18, no. 4, pp. 311-318, 2003.

[26] Knee Pain Health Center, available at: http://www.webmd.com/pain-management/knee-pain/knee-problems-and-injuries-topic-overview, Last Visited, 2015.

[27] Kwolek B. and Kepski M., "Human Fall Detection on Embedded Platform Using Depth Maps and Wireless Accelerometer," *Computer Methods and Programs in Biomedicine*, vol. 117, no. 3, pp. 489-501, 2014.

[28] Leskovec J., Rajaraman A., and Ullman J., *Mining of Massive Datasets*, Stanford, 2014.

[29] Ma X., Wang H., Xue B., Zhou M., Ji B., and Li Y., "Depth-Based Human Fall Detection Via Shape Features and Improved Extreme Learning Machine," *Journal of Biomedical and Health Informatics*, vol. 18, no. 6, pp. 1915-1922, 2014.

[30] Magaziner J., Hawkes W., Hebel J., Zimmerman S., Fox K., Dolan M., Felsenthal G., and Kenzora J., "Recovery from Hip Fracture in Eight Areas of Function," *Journal of Gerontology*, vol. 55, no. 9, pp. 498-507, 2000.

[31] Marjoux D., Baumgartner D., Deck C., and Willinger R., "Head Injury Prediction Capability of the HIC, HIP, SIMon and ULP Criteria," *Accident; Analysis and Prevention*, vol. 40, no. 3, pp. 1135-1148, 2008.

[32] Mastorakis G. and Makris D., "Fall Detection System Using Kinect's Infrared Sensor," *Journal of Real-Time Image Processing*, vol. 9, no. 4, pp. 635 - 646, 2014.

[33] Murthy C., Harish S., and Chandra Y., "The Study of Pattern of Injuries in Fatal Cases of Fall from Height," *Journal of Medical Sciences*, vol. 5, no. 1, pp. 45-52, 2012.

[34] Nghiem A., Auvinet E., Meunier J., "Head Detection Using Kinect Camera and Its Application to Fall Detection," *in Proceedings of the 11th International Conference on Information Science, Signal Processing and their Applications*, Montreal, pp. 164-169, 2012.

[35] Noury N., Fleury A., Rumeau P., Bourke A., Laighin G., Rialle V., and Lundy J., "Fall Detection-Principles and Methods," *in Proceedings of the 29th International Conference of the Engineering in Medicine and Biology Society*, Lyon, pp. 1663-1666, 2007.

[36] O'Keefe G. and Jurkovich G., *Injury Control a Guide to Research and Program Evaluation*, Cambridge University Press, 2000.

[37] OpenNI, available at: https://github.com/OpenNI/OpenNI, Last Visited, 2014.

[38] Patsadu O., Nukoolkit C., and Watanapa B., "Survey of Smart Technologies for Fall Motion Detection: Techniques, Algorithms and Tools," *in Proceedings of the 5th International Conference on Advances in Information Technology*, Bangkok, pp. 137-147, 2012.

[39] Planinc R. and Kampel M., "Introducing the Use of Depth Data for Fall Detection," *Journal of Personal and Ubiquitous Computing*, vol. 17, no. 6, pp. 1063-1072 , 2013.

[40] Rougier C., Auvinet E., Rousseau J., Mignotte M., and Meunier J., "Fall Detection from Depth Map Video Sequences," *in Proceedings of the 9th International Conference on Smart Homes and Health Telematics*, Montreal, pp. 121-128, 2011.

[41] Schneider E., Shubert T., and Harmon K., "Addressing the Escalating Public Health Issue of Falls Among Older Adults," *Medical Society of the State of North Carolina*, vol. 71, no. 6, pp. 547-552, 2010.

[42] Schulz B., Lee W., and Lloyd J., "Estimation, Simulation, and Experimentation of a Fall from Bed," *Journal of Rehabilitation Research and Development*, vol. 45, no. 8, pp. 1227-1236, 2008.

[43] Stone E. and Skubic M., "Fall Detection in Homes of Older Adults Using the Microsoft Kinect," *Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 290-301, 2015.

[44] Tran T., Le L., and Morel J., "An Analysis on Human Fall Detection Using Skeleton from Microsoft Kinect," *in Proceedings of the 5th International Conference on Communications and Electronics*, Danang, pp. 484-489, 2014.

[45] Uddin M., Kim D., and Kim T., "A Human Activity Recognition System Using HMMs with GDA on Enhanced Independent Component Features," *The International Arab Journal of Information Technology*, vol. 12, no. 3, pp. 304-310, 2015.

[46] Versace J., "A Review of the Severity Index," *in Proceedings of the 15th Stapp Car Crash Conference*, Coronado, pp. 771-796, 1971.

[47] Yu X., "Approaches and Principles of Fall Detection for Elderly and Patient," *in Proceedings of the 10th International Conference on E-health Networking, Applications and Services*, Singapore, pp. 42-47, 2008.

Orasa Patsadu received her B.B.A. degree in Computer Information Systems from Rajamangala University of Technology Krungthep, Thailand, in 2007 and M.Sc. in Software Engineering from King Mongkut's University of Technology Thonburi, Thailand in 2010. Since April 2011, she has been pursuing the Ph.D. in School of Information Technology, King Mongkut's University of Technology Thonburi. Her research interests include data mining and software engineering.

Bunthit Watanapa received his B.Eng. in Computer Engineering from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 1987 and M.Eng. and Ph.D. in Industrial Engineering from Asian Institute of Technology (AIT), Thailand, in 1990 and 2003, respectively. Currently he is the chairperson of the Business Information System program at the School of Information Technology, King Mongkut's University of Technology Thonburi, Thailand. His research interests are in the area of Decision Support System, Optimization, and Project Management.

Piyapat Dajpratham received her M.D. from Mahidol University, Bangkok, Thailand in 1992 and Diplomat Thai Board of Rehabilitation Medicine from Faculty of Medicine Siriraj Hospital in 1997. She had fellowship training in Neurorehabilitation & Geriatric Rehabilitation from Northwestern Healthcare Network, University of Melbourne, Australia in 1998 and 1999 respectively. Currently she is an Associate Professor in Rehabilitation Medicine, Faculty of Medicine Siriraj Hospital, Bangkok, Thailand. Her research interests are stroke rehabilitation and falling in the elderly.

Chakarida Nukoolkit received her B.Sc. in Computer Science from Thammasat Univeirisity, Thailand in 1992 and M.Sc. in Computer Science from Vanderbilt University, U.S.A. in 1995 and Ph.D. in Computer Science from University of Alabama, U.S.A. in 2001. Previously she had led the Data and Knowledge Engineering Lab at the School of Information Technology, King Mongkut's University of Technology Thonburi, Thailand. Her current research interests are in the area of data mining, data science, visualizations, creative computing, bioinformatics, and artificial intelligence.