A Method for Automatic Fall Detection of Elderly People Using Floor Vibrations and Sound—Proof of Concept on Human Mimicking Doll Falls

Yaniv Zigel, *Member, IEEE*, Dima Litvak, and Israel Gannot*

***Abstract*—Falls are a major risk for the elderly people living independently. Rapid detection of fall events can reduce the rate of mortality and raise the chances to survive the event and return to independent living. In the last two decades, several technological solutions for detection of falls were published, but most of them suffer from critical limitations. In this paper, we present a proof of concept to an automatic fall detection system for elderly people. The system is based on floor vibration and sound sensing, and uses signal processing and pattern recognition algorithm to discriminate between fall events and other events. The classification is based on special features like shock response spectrum and mel frequency ceptral coefficients. For the simulation of human falls, we have used a human mimicking doll: "Rescue Randy." The proposed solution is unique, reliable, and does not require the person to wear anything. It is designed to detect fall events in critical cases in which the person is unconscious or in a stress condition. From the preliminary research, the proposed system can detect human mimicking dolls falls with a sensitivity of 97.5% and specificity of 98.6%.**

***Index Terms*—Acoustic signal processing, fall detector, feature extraction, pattern recognition, transducers.**

I. Introduction

FALLS and sustained injuries among the elderly are a major problem worldwide, and are the third cause of chronic disability according to World Health Organization [1]. The proportion of people sustaining at least one fall during 1 year varies from 28% to 35% for the age of 65 and over [2], while falls often signal the "beginning of the end" of an older person's life. The risk of fall increases with age, and in two cases out of three, it happens at home. People who experience a fall event at home, and remain on the ground for an hour or more, may suffer from many medical complications, such as dehydration, internal bleeding, and cooling, and half of them die within 6 months [3].

Manuscript received April 15, 2009; revised May 29, 2009. First published August 25, 2009; current version published November 20, 2009. *Asterisk indicates corresponding author.*

Y. Zigel is with the Biomedical Signal Processing Research Laboratory, Department of Biomedical Engineering, Faculty of Engineering, Ben-Gurion University, Beer-Sheva 84105, Israel (e-mail: yaniv@bgu.ac.il).

D. Litvak is with the Department of Biomedical Engineering, Faculty of Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel (e-mail: litvakdi@post.tau.ac.il).

*I. Gannot is with the Department of Biomedical Engineering, Faculty of Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel, and also with the Program of Biomedical Engineering, Department of Electrical and Computer Engineering, School of Engineering and Applied Sciences, George Washington University, Washington, DC 20052 USA (e-mail: gannot@eng.tau.ac.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2009.2030171

In the past two decades, there have been several commercial solutions and academic developments aimed at automatic and non-automatic detection of falls.

A. Social Alarm

The social alarm is a wristwatch with a button that is activated by the subject in case of a fall event. The main problem with this solution is that the button is often unreachable after the fall especially when the person is panicked, confused, or unconscious [4].

B. Automatic Fall Detector

The most popular solutions for automatic detection of falls are the wearable fall detectors that are based on combination of accelerometers and tilt sensors [5], [6]. Williams *et al.* [7] developed a device that is based on a combination of shock and tilt sensors. Depeursinge *et al.* [8] used three accelerometers to obtain the position, speed, and acceleration vector of the person. Noury *et al.* [9] developed a device that is placed under the armpit, and employs two accelerometers and a microcontroller to compute the orientation of the body. A critical disadvantage of these solutions is that the person has to wear the device in the shower, a place with a high occurrence rate of falling, and we know that people prefer not to wear anything while showering. Moreover, these devices produce many false alarms [10], and old people tend to forget wearing them frequently.

C. Video Analysis-Based Fall Detection System

There are few solutions from the last years, which are based on image processing of the person's movement in real-time [11], [12]. Wu [13] analyzed the vertical and horizontal speeds during a fall. Lin *et al.* [14] developed a networked video camera system that detects moving objects, extracts features such as object speed and determines if a human fall has occurred. Camera-based solutions suffer from a few disadvantages like privacy concerns (which is critical), and difficulty to monitor the entire area of a house.

Due to the disadvantages of the existing fall detection techniques, there is a need for a better solution for the detection of elderly falls. We have developed a solution that is based on the combination of floor vibrations and sound detection during a fall [15]. The proposed solution does not require the subject to wear anything. The idea of floor vibrations was also explored by Alwan *et al.* [16], however, our proposed solution is different by means of the sensors, features, and the pattern recognition algorithm.

II. The Concept of the Proposed Fall Detection Method

A human fall on the floor creates a shock signal that propagates in the floor. In a daily routine, there are a lot of sounds in the house, but when there is a special combination of fall-specific vibration event with a suitable sound event, there is a suspicion for a human fall on the floor. We decided to use a combination of vibration and sound sensors because they can supply the information about the way the fall vibrates the floor and how it sounds (this combined data will be explained shortly). The main hypothesis of this paper is that in most cases we can accurately identify human falls and discriminate them from other events using a sound and floor vibration detection in conjunction with advanced signal processing techniques. The goal of the presented preliminary research was to develop a method, which consists of an algorithm for event detection and classification, and to prove the concept of the proposed solution. The proposed algorithm that is based on pattern recognition techniques enables to distinguish between a simulated human fall event and other events such as fall of an object on the floor.

The proposed system is based on the detection of vibration and sound signals from an accelerometer and a microphone. The sensors that are used include a Crossbow CXL02LF1Z accelerometer (San-Jose, CA) and a small amplified microphone (MS-3100 W), both are attached to the floor by adhesive tapes. The acquired signals are transmitted to a portable NI USB-6210 (National Instruments, TX, USA) data acquisition device that samples the signals at 16 kHz and transmits them to a PC. The acquired digital signals from the sensors were collected by NI LabView software, and analyzed using MATLAB software. In the experiments that we have performed, the accelerometer and the microphone were located at the side of the room (close to the wall), and connected to the floor by a scotch tape; the accelerometer was attached to the floor, the microphone was placed on the accelerometer, and the scotch tape connected them to the floor from above.

As a rule of thumb, if a machine (or floor in this case) produces high amplitude vibrations (greater than 10 g rms, where g is the gravitational acceleration of earth) at the measurement point, a relatively low sensitivity (10 mV/g) sensor is preferable. If the vibration is less than 10 g rms, a high sensitivity (100 mV/g) sensor should be generally used. The acceleration of the floor in an event of a human fall is in a scope of 1–2 g with low amplitude vibrations. Therefore, we chose to use an accelerometer with a sensitivity of 1 V/g with the ability to sense accelerations up to 2 g. The chosen microphone has frequency range of 20–16 000 Hz and signal-to-noise ratio (SNR) of more than 58 dB.

The fall detection algorithm was developed using MATLAB software and is basically comprised of two main stages: event detection and event classification. The event detection stage uses energy features from the vibration and sound signals to locate and segment (finding the boundaries of) a suspicious fall event. The event classification stage uses temporal and spectral features, extracted from the segmented event signals, and aims to discriminate between fall events and other events.

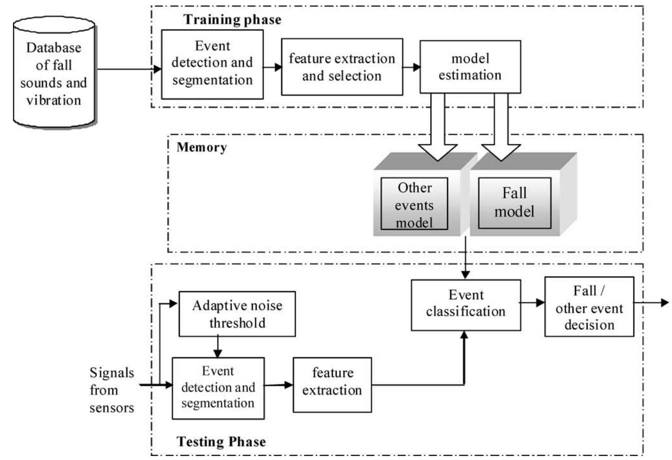

Fig. 1. Fall detection and classification block diagram.

III. Fall Detection and Classification Algorithm

The algorithm contains two phases of data analysis: training phase and testing phase that are presented in Fig. 1. Both phases use vibration and sound signals as inputs. In order to trigger the classification algorithm, there must be a significant event in the vibration signal. Once an event is detected in the vibration signal, the sound signal is analyzed.

The training phase consists of event detection and segmentation, feature extraction, feature selection, and model estimation modules. In the training phase of the algorithm, we estimate two classification models for different types of events. In the proposed system, there are two classes of classification: "fall" (in this case "Rescue Randy fall") and "other event."

A. Event Detection and Segmentation (Training Phase)

The main principle of the event detection algorithm is to detect a vibration event in the vibration signal $v(t)$ at time index t_e, and to segment it. The event detection and segmentation are performed using energy calculations $e_f(n)$ from running time frames

$$e_f(n) = \sum_{t=(n-1)L_f}^{nL_f-1} v^2(t); \qquad n = 1, \ldots, N_f - 1 \qquad (1)$$

where $e_f(n)$ represents the energy of the nth time frame, N_f is the total number of frames in the signal $v(t)$, and L_f is the length of each frame (20 ms in this case). By calculation of $e_{\max}$, which is the maximum value of the array $e_f(n)$, the event detection algorithm finds n_e (the frame index of $e_{\max}$), and calculates t_e that equals to $L_f \times n_e$.

After finding t_e by the event detection algorithm, the event segmentation algorithm extracts the event from the recorded finite vibration signal. Segmentation of the vibration event is performed by an automatic algorithm that identifies the

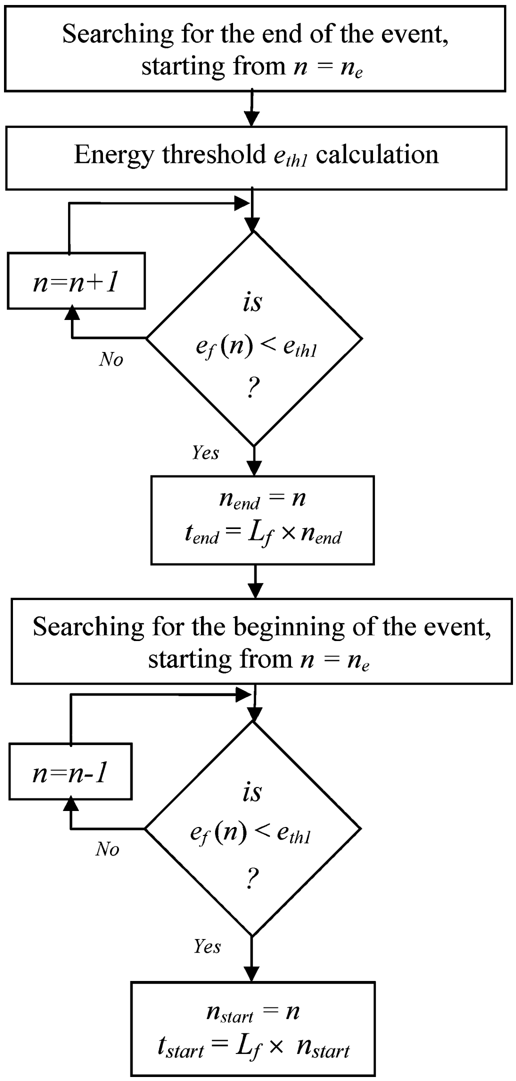

Fig. 2. Event segmentation algorithm.

boundaries of the event. The boundaries of the event t_{start} and t_{end} are calculated by an algorithm that is based on an automatic noise threshold calculation.

Fig. 2 describes the event segmentation algorithm on a recorded finite vibration signal in the training phase. In this figure, n_{end} represents the frame index of the end of the event, n_{start} is the frame index of the beginning of the event, t_{end} is the time index of the end of the event, t_{start} is the time index of the beginning of the event and e_{th1} is the energy threshold.

Calculation of energy threshold e_{th1}*:* The value of e_{th1} is calculated by assessing the e_{hist} that is the most prominent value of $e_f(n)$. e_{hist} is calculated from a binned histogram of the recorded finite vibration signal. The value of e_{th1} is calculated by

$$e_{\mathrm{th1}} = e_{\mathrm{hist}}(1 + c_1) \tag{2}$$

where c_1 is an empirical threshold coefficient (set to 0.2).

Fig. 3(a) shows an example of the event segmentation, and Fig. 3(b) shows an example of a histogram calculation of the array $e_f(n)$ with a resolution of 0.00001. As seen in Fig. 3(b), the background noise energy of the signal has a normal statistical distribution. The threshold in this example is calculated to be 0.0067 (using 20 ms frame size).

After detection and segmentation of a vibration event, segmentation of the sound event from the sound signal is performed. Starting from t_e, the algorithm finds the boundaries of the sound event using the same segmentation technique mentioned earlier. Fig. 4(a)–(d) shows an example of the results of the event detection and segmentation algorithms implemented on a vibration and sound signals of a fall event.

B. Feature Extraction (Training and Testing Phases)

The main classification problem is to distinguish between fall events and other events. Following the event detection and segmentation, features are extracted from these vibration and sound event signals (training phase trials) for model estimation. The selection of the complete feature set, the subset of features to be used from the complete feature set, and the design of the model that uses these selected features is of central importance for obtaining high classification accuracy. The complete set of features that were chosen as candidates for the model are composed of two kinds of features: temporal features and spectral features.

1) Temporal features: The features that are extracted from the vibration and sound events are length (time) and energy (sum of square of the amplitudes over time), a total of four temporal features.
2) Spectral features: The shock response spectrum (SRS) [17] features are extracted from the vibration event signal, and mel frequency cepstral coefficients (MFCC) [18] are extracted from the sound event signal.

Table I summarizes the overall set of the candidate features.

1) SRS: For the analysis of the floor's vibration signal, we chose to use a physical approach for the description of the human–floor system. Robinovitch *et al.* [19] described the dynamics of impact to the hip during a fall event as mass–spring system. This approach justifies the use of the SRS analysis that is popular in many engineering fields for vibration signal analysis.

The SRS calculation is a transform that assumes that the fall event is a mass–spring system. The SRS is the peak acceleration responses of a large number of single degree of freedom (SDOF) systems each one with a different natural frequency (Fig. 5). It is calculated by convolution integral of the measured signal (input) with each one of the SDOF systems [17].

A typical scheme of the frequencies of the *x*-axis is based on a proportional bandwidth, such as 1/6 octave. This means that each successive natural frequency is $2^{1/6}$ times the previous natural frequency. Fig. 6 is an example of the SRS plot of one of vibration event as measured by our accelerometer. The SRS transform has in total of 133 values, but in the very low frequencies, these values are close to zero. Therefore, we have chosen 93 values of the SRS, as candidate features from the frequency bandwidth of 10.1– 2048 Hz of a specific vibration event.

2) MFCC: MFCCs represent audio signals with frequency bands that are positioned logarithmically (on the mel scale)

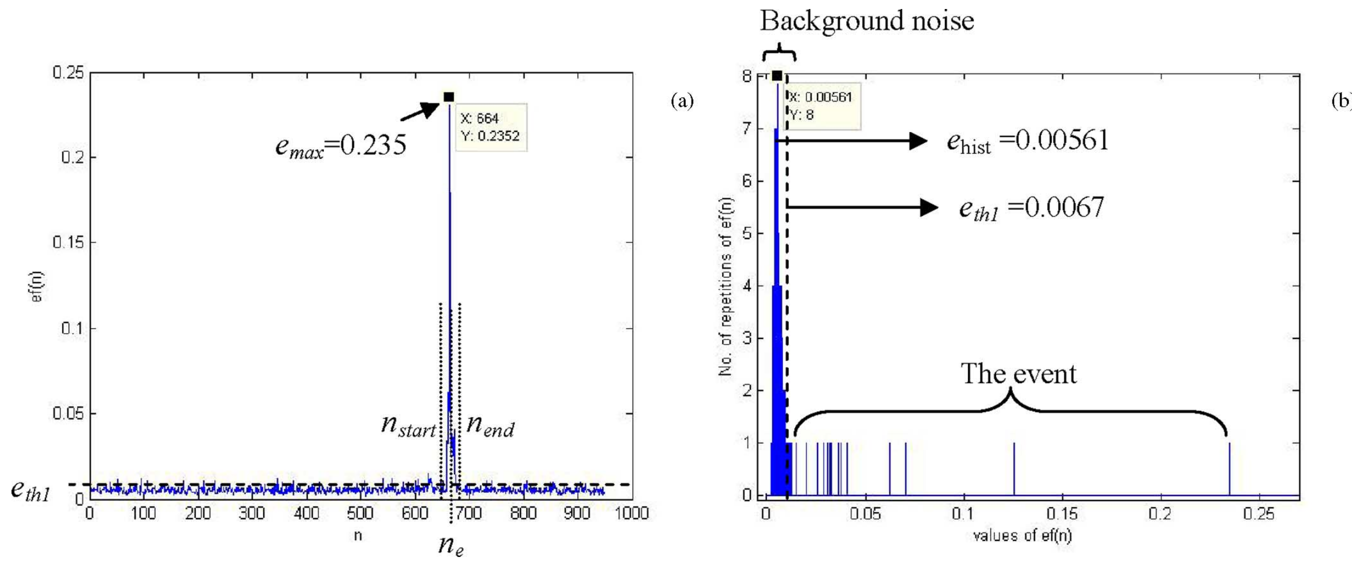

Fig. 3. (a) Energy signal and event segmentation. (b) Histogram calculation of $e_{\rm th1}$

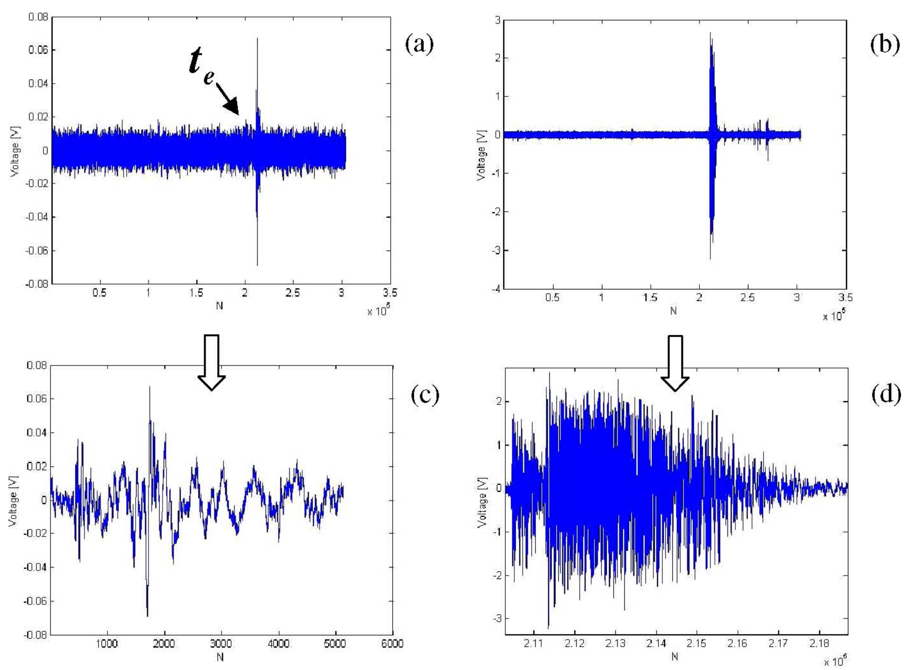

Fig. 4. (a) Vibration event. (b) Sound event. (c) Segmented vibration event. (d) Segmented sound event.

TABLE I
OVERALL SET OF CANDIDATE FEATURES

Kind of feature	*Feature name*	*No. of features*	*Vibration/Sound feature*	*Feature symbol*
Temporal features	Vibration event length	1	Vibration	L1
	Sound event length	1	Sound	L2
	Vibration event energy	1	Vibration	E1
	Sound event energy	1	Sound	E2
Spectral features	SRS	93	Vibration	S1-S93
	MFCC	13	Sound	C1-C13

[18] and approximate the human auditory system's response more closely than the linearly spaced frequency bands obtained directly from the fast Fourier transform (FFT) of the signal. The use of these coefficients is popular in speech and speaker recognition [20], and fits our goal to characterize the sound of a human fall event and other events. In our study there are different kinds of sound signals. The signal can vary from short events (such as human steps) to long events (such as a fall). We have decided to divide the sound event signals to windows with length of 0.03 s, and to calculate the MFCC coefficients for each window. The MFCC transform supplies 13 features for each window, when the first feature is the energy of the window.

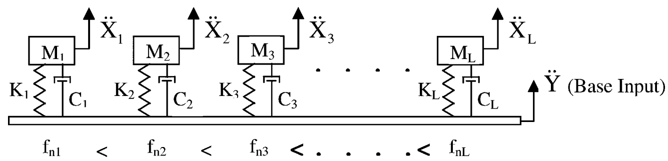

Fig. 5. SRS Model. M_i: mass; C_i: damping coefficient; K_i: stiffness; f_{n_i}: natural frequency; $\ddot{Y}$: the measured signal (input) $\ddot{X}_i$: the response of the system to the input.

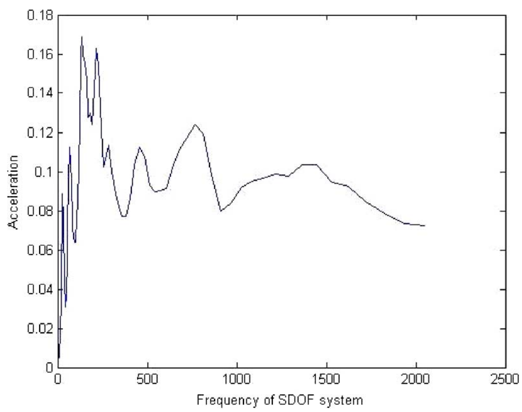

Fig. 6. SRS plot of a vibration event.

By choosing the window with the maximum value of the first feature (the most *"energetic" window), we extracted 13 MFCC features from* each sound event signal.

C. Feature Selection (Training Phase)

The problem of selecting a subset of features from *N*-dimensional feature measurement vector is called feature selection. The feature selection procedure reduces the cost of the pattern recognition process, and provides better classification accuracy due to finite training dataset size effects [21]. There are several feature selection procedures discussed in the pattern recognition literature like: sequential forward selection (SFS) [22], sequential backward selection (SBS) [23], *l–r* algorithm, sequential forward floating search (SFFS) [24], etc.

Based on the results of Zongker *et al.* [25], which evaluated the performance of various feature selection methods, the SFFS method was the most powerful algorithm for feature selection. This result was also shown in [20].

Therefore, we have chosen to use the SFFS algorithm with Mahalanobis [26] distance test criterion for performance evaluation of the features

$$D(\mathbf{Z}) = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \mathbf{C}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \tag{3}$$

where $\mathbf{Z}$ is the set of the feature vectors, $\mathbf{C}$ is the covariance matrix of the feature vectors, and $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ are the mean vectors of each class ("fall" and "other event").

The Mahalanobis distance test criterion is the appropriate one for distance measurement of features with a Gaussian distribution (discussed in Section II-F).

Three major steps are identified in the SFFS algorithm: the first is "inclusion," the second is "test," and the last is "exclusion." SFFS begins with the inclusion process to select a feature with best performance. The searching process is followed by conducting a test on every feature selected in the same iteration to specify features that will decay the whole performance. If such a feature exists, SFFS would commence the exclusion process to ignore such a feature. The algorithm will continue looking for other better features until all features are examined [24].

The algorithm for feature selection ranked the performance of 110 features (Table I), and chose a set of 17 top performing features for event classification (in Section V).

D. Model and Classifier Estimation (Training Phase)

In the training phase, we use the selected features that are extracted from the collected signals to estimate two models: "fall" model and "other event" model.

The classification algorithm used in this study is based on Bayes classification. The Bayes decision rule classifies an observation vector $\mathbf{z}$ to the class that has the highest *a posteriori* probability between the two classes [27]–[29]

$$\omega(\mathbf{z}) = \arg\max_{\omega \in \Omega} \{P(\omega|\mathbf{z})\} \tag{4}$$

where $\omega(\mathbf{z})$ is the chosen class, $\mathbf{z}$ is the observation feature vector, $P(\omega|\mathbf{z})$ is the *a posteriori* probability, and $\Omega = \{\omega_1, \omega_2\}$ is the class space.

The Bayes decision rule in terms of *a priori* probabilities and the conditional probability densities is

$$\omega(\mathbf{z}) = \arg\max_{\omega \in \Omega} \{p(\mathbf{z}|\omega)P(\omega)\} \tag{5}$$

where $p(\mathbf{z}|\omega)$ is the conditional probability density function (the likelihood function of $\mathbf{z}$ given ω), and $P(\omega)$ is the class *a priori* probability.

In this study, the *a priori* probabilities of human fall events and other events are not known. Therefore, the *a priori* probabilities for the two classes are assumed to be equal. Hence

$$P(\omega_k) = \frac{1}{2}; \qquad k = 1, 2 \tag{6}$$

where *k* is the class index.

The training dataset found to have a Gaussian distribution for each class. Therefore, Gaussian models were estimated for each class. The Gaussian conditional density function is

$$p(\mathbf{z}|\omega_k) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}_k|}} \exp\left(\frac{-(\mathbf{z} - \boldsymbol{\mu}_k)^T \mathbf{C}_k^{-1} (\mathbf{z} - \boldsymbol{\mu}_k)}{2}\right) \tag{7}$$

where $\mathbf{C}_k$ is the *k*th class covariance matrix, $\boldsymbol{\mu}_k$ is the *k*th class expectation vector, and N is the feature space dimension. The *k*th class expectation vector $\boldsymbol{\mu}_k$ is estimated by

$$\boldsymbol{\mu}_k = \frac{1}{N_k} \sum_{n=1}^{N_k} \mathbf{z}_n \tag{8}$$

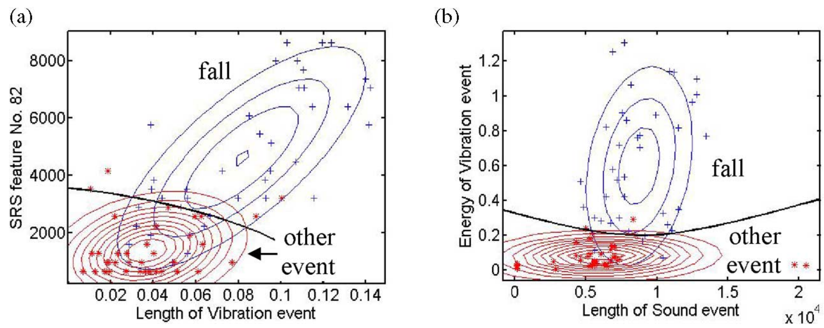

Fig. 7. Projection of the training feature vectors on a two-dimensional space and the model estimation. (a) SRS feature No. 82 and the length of vibration event features. (b) Energy of vibration event and length sound event features.

where N_k is the number of samples with class ω_k, and $\mathbf{z}_n$ is the measurement features vector. The *k*th class covariance matrix $\mathbf{C}_k$ is estimated by

$$\mathbf{C}_k = \frac{1}{N_k - 1} \sum_{n=1}^{N_k} (\mathbf{z}_n - \boldsymbol{\mu}_k)(\mathbf{z}_n - \boldsymbol{\mu}_k)^T. \tag{9}$$

The model parameters that are stored for each class are the expectation vector $\boldsymbol{\mu}_k$ and the covariance matrix $\mathbf{C}_k$. Substitution of (6) and (7) in (5) gives the following Bayes decision rule

$$\omega(\mathbf{z}) = \omega_i \qquad \text{with}$$

$$i = \underset{k=1,2}{\operatorname{argmax}} \left\{ \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}_k|}} \cdot \exp\left(\frac{-(\mathbf{z}-\boldsymbol{\mu}_k)^T \mathbf{C}_k^{-1} (\mathbf{z}-\boldsymbol{\mu}_k)}{2} \right) \cdot \frac{1}{2} \right\}. \tag{10}$$

We can take the logarithm of the function between braces without changing the result of the argmax{} function. Therefore, (10) is equivalent to

$$\omega(\mathbf{z}) = \omega_i \qquad \text{with}$$

$$i = \underset{k=1,2}{\operatorname{argmax}} \left\{ -\frac{1}{2} \ln |\mathbf{C}_k| - \frac{1}{2} \cdot (\mathbf{z}-\boldsymbol{\mu}_k)^T \mathbf{C}_k^{-1} (\mathbf{z}-\boldsymbol{\mu}_k) \right\}. \tag{11}$$

Equation (11) calculates the maximum likelihood, and chooses class ω_i (i = 1,2) for a specific vector $\mathbf{z}$ in the N dimensional features space. That classifier is called quadratic classifier.

Following the estimation of the models, we can calculate the boundaries between the two models by using the quadratic classifier function (11) and use the estimated classifier to classify events in the testing phase of the algorithm. Fig. 7(a) and (b) shows the plots of two selected features with a quadratic classifier. The data in Fig. 7 were collected in the training phase trials.

E. Event Detection and Segmentation (Testing Phase)

In the testing phase, the signal is continuous. Therefore, the algorithm analyzes finite windows of the vibration signal by calculating the energy of the window $e_w(m)$, and finds a finite event-suspected window. When there is an event-suspected window, event detection and segmentation, feature extraction, and event classification algorithms are performed. Event classification is based on the estimated models of the events.

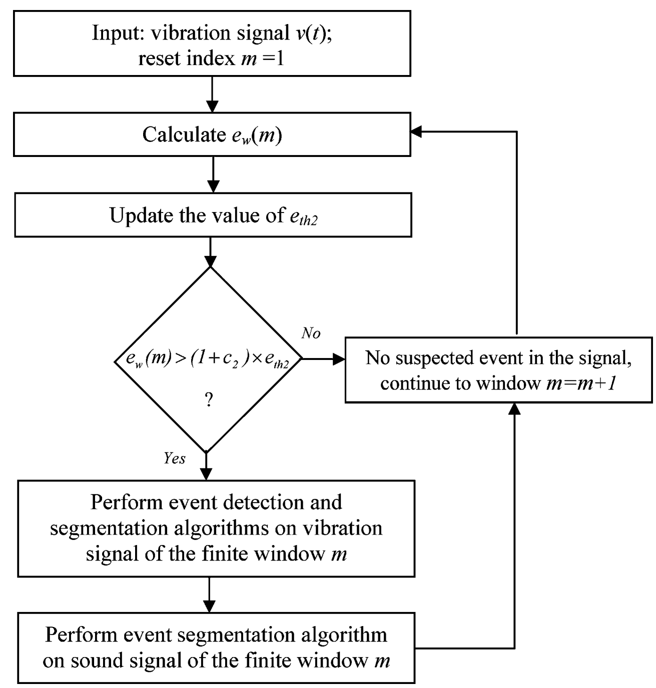

Fig. 8. Event detection and segmentation algorithms in the testing phase.

Calculation of *m*th event-suspected window energy value $e_w(m)$ is performed by

$$e_w(m) = \sum_{t=(m-1)Dw}^{(m-1)Dw+Tw} v^2(t); \qquad m = 1, 2, \ldots \tag{12}$$

where T_w is the length of the event-suspected window (20 s in this case), D_w is the time difference between event-suspected windows (10 s in this case), and $v(t)$ is the vibration signal.

Fig. 8 describes the event-detection algorithm in the testing phase. In Fig. 8, c_2 is an empirical threshold coefficient (set to 0.1), and $e_{\rm th2}$ is the time-averaged adaptive noise level.

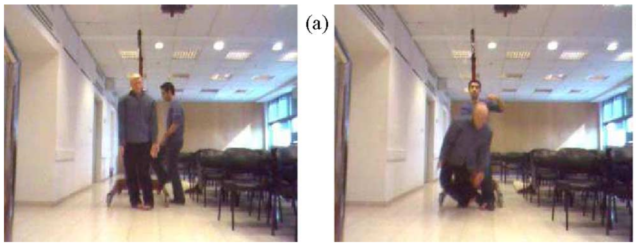

Fig. 9. (a) Randy before the drop. (b) Randy is falling.

The value of $e_{\mathrm{th}2}$ is recalculated every D_w seconds by the same technique as $e_{\mathrm{th}1}$ is calculated (by assessing the most prominent value of $e_w(k)$, which is calculated from a binned histogram of the prior 10 min).

F. Event Classification (Testing Phase)

Following the segmentation of the vibration and sound event signals from the finite event-suspected window, 17 selected features are extracted from the testing data signals (13 features from vibration signals, and four features from sound signals).

The values of the extracted features are substituted in the 17-dimentional estimated classifier (11). Following the calculation of the maximum likelihood, the classifier returns the classification result whether the event is "fall" (positive) or it is "other event" (negative).

IV. Experimental Setup

The training and testing datasets for the algorithm were taken from experiments that have been performed on a typical concrete tile floor and a carpet using "Rescue Randy"—a human mimicking doll (74 kg) and four "popular falling" objects. These experiments were performed at distances of 2–5 m from the sensors. Moreover, we decided to drop objects and simulate events that generate significant floor vibrations close to the sensors in order to ensure the algorithm is effective in a multitude of conditions. The experiments were conducted on the ground floor, which is more challenging, since the ground floor has stronger support from the ground and therefore less vibrates than the higher levels. However, in Israel, the differences between the floors (higher or lower levels) are insignificant, since most of the floors are concrete tile floors with sand beneath the tiles. The drops of "Rescue Randy" [see Fig. 9(a) and (b)] were in a forward direction which comprise of falls on the hands, falls on the knees, and falls on the side (on the pelvis). The forward direction was chosen because of the statistics that 60% of human falls are in the forward direction [30]. Typical of human falls, no two falls of "Rescue Randy" are identical. It should be mentioned that "Rescue Randy" is a popular model to mimic the weight of a human body, and there were a few previous researches that used "Rescue Randy" to mimic a human fall on the floor [16], [31].

TABLE II
Overall Set of Selected Features

Feature name	***Feature symbol***	***Selected features***
Vibration event length	L1	L1
Sound event length	L2	L2
Vibration event energy	E1	E1
Sound event energy	E2	-
SRS	S1-S93	S2, S10, S34, S64, S68, S74, S76, S77, S82, S84, S91
MFCC	C1-C13	C3, C11, C12

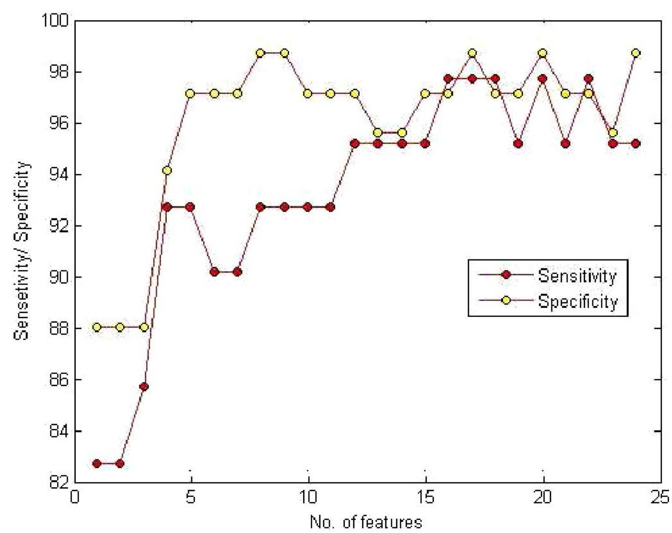

Fig. 10. Sensitivity/specificity of the algorithm versus number of features.

A. Training Phase Trials

In the training phase, "Rescue Randy" was dropped ten times from each distance (a total of 40 drop events). The objects that were dropped on the floor include a heavy bag (15 kg), a book, a plastic box and a metal box. The objects were dropped five times in each distance (a total of 20 times for each object, and a total of 80 drops of objects). Due to large distance from the sensors, only 28 of the 80 drops of the objects were passed the energy-based detection algorithm; therefore, only 28 events of dropped objects in distances of 2–5 m were included in the training set of data. The other trials that were performed close to the sensors included: walking, dropping a chair, jumping from the chair on the floor, dropping a heavy bag, dropping a plastic box, and dropping a metal box for a total of 12 events (each trial was repeated twice). In total, the training set of data included 40

TABLE III
CLASSIFICATION RESULTS

Classified as \ Real class	"Other event"		"Fall"	
	Objects	**Events close to the sensors**	**"Fall" on the floor**	**"Fall" on a carpet**
"Other event"	**44 (+4 undetected)**	**27**	1	0
"Fall"	0	1	**19**	**20**

"Rescue Randy" falls and 40 drops of objects and other events (28+12).

B. Testing Phase Trials

In the testing phase, the trials consisted of 20 drops of "Rescue Randy" (five repetitions in each distance of 2–5 m) and 48 drops of objects (four kinds of objects, each one with three repetitions in each distance). The other trials that were performed close to the sensors (the same other trials as in the training phase) included three repetitions of each event (a total of 18 events). A door slam is a popular event at home that generates vibration signals in the floor. Therefore, we have decided to test the system with door slams. By location of the sensors in a distance of 30 cm from a standard wooden door, ten repetitions of different door slams were performed. The total number of "other events" that were simulated close to the sensors was 28.

Due to the fact that apartments may be carpeted, we also performed a testing set of experiments on a carpet. There were five drops of "Rescue Randy" at distances of 2–5 m on a carpet for a total of 20 events. The testing set of data included a total of 40 drops of "Rescue Randy" and 76 drops of objects and other events.

V. RESULTS

A. Feature Selection

The algorithm for feature selection ranked the performance of 110 features that are extracted from the training data. A complete set of top performing 17 features was chosen for classification. These features are: lengths of vibration and sound event signals, energy of vibration event signal, 11 SRS features, and 3 MFCC features. Table II summarizes the overall set of the selected features for classification.

During the process of the decision on the number of features for classification, the purpose was to receive sensitivity and specificity of the classification algorithm that are higher than 97%. As shown in Fig. 10, the required sensitivity and the specificity are achieved, when the classification is performed by 17 features. Increasing the number of features for classification beyond 17, does not increase the performance of the classification algorithm, and can cause over fitting. Therefore, that number is the optimal number of features for classification.

B. Testing Phase Event Classification

Table III summarizes the results of the detection and classification algorithm that was run on the testing database of the trials. Cases in which no event detected are classified as "other event."

A positive event is classified as "fall," and a negative event is classified as "other event." From the data presented, we found the sensitivity of the system to be 97.5% and the specificity to be 98.6%. The sensitivity of the system was calculated as the number of "Rescue Randy" drops that classified as "fall" (true positive) divided to the total number of "Rescue Randy" drops (39/40). The specificity of the system was calculated as the number of objects and events close to the sensors that classified as "other event" (true negative) divided to the total number of object drops and events close to the sensors (75/76). It should be mentioned that in the case of a false negative, "Rescue Randy" fell in a distance of 5 m from the sensors. In the case of false positive, the object fell close to the sensors generating strong signals.

VI. DISCUSSION AND CONCLUSION

This paper describes a preliminary research for proof of concept of an innovative solution (method) for automatic detection of elderly people falls at home. The results of the laboratory experiments show that the proposed method has the potential to serve as a solution to the discussed problem. Assuming that "Rescue Randy" fall is a good model for simulation of real human fall events, the system can detect human falls with high precision for distance up to 5 m. The proposed solution is a low cost solution, does not require the person to wear anything, and is considerate of the person's privacy.

In all the test trials of "Rescue Randy" on the carpet, the events were classified as "falls," although the system was not trained for this kind of floor. The conclusion here is that the classification algorithm is robust enough to different floor types. In cases of object drops in distances of 2–5 m, in the training phase 28 out of 80 event were detected, and in the testing phase 44 out of 48 event were detected. In total, 56.25% of the object drops were detected. This is a good result because we want the system to be sensitive for detection of human falls and less sensitive to detection of other events. The difference, between the percentages of detected object drops in the training and the testing phases, can be explained by noise levels of the vibration signals; object drop events create relatively low-amplitude vibration signals that sometimes cannot be distinguished from the background noise. The training and testing trials were performed in different days, and the noise level of the system that was different in those days influenced the percentage of the detected object drop events. In cases of "Rescue Randy" drops in the training and testing phases, all the events were detected, and were sent to classification. The conclusion here is that the proposed system is sensitive enough for the detection of simulated human fall events, and less sensitive to other events. High sensitivity of the event detection algorithm proves that the chosen accelerometer was a good choice for the measurement of floor vibration signals. As mentioned previously, the false positive event was generated by an event that was close to the sensors (closer than 50 cm). Therefore, in order to avoid this kind of false alarm, we suggest putting the sensors in the corner of a room, where there is a lower chance of object fall or other event too close to the sensors.

In the testing set of experiments with door slams, the event detection algorithm detected all the events, and the event classification algorithm classified all the events as "other event." That set of experiments shows that the proposed method is good enough for detection and classification of events that can potentially trigger the fall alarm. It also proves that although the system was not trained with door slam events, it can classify them successfully (as "other events").

One limitation of the system is that it may not be sensitive to low-impact real human falls in some cases that were not tested (e.g., slow and soft human fall on the floor from chair).

However, the serious falls capable of structural damage to bone (the events in which the person cannot usually get up) cause high-amplitude floor vibrations that can be detected by the system. These problematic human fall events and the rest of different human fall events should be tested with real human fall trials in the future.

"Rescue Randy" doll is a good model for simulation of human fall events, but its main disadvantage is that its "skin" is made of plastic that is not soft as human skin; this can influence the vibration signals created in the floor following a fall event. Therefore, before a performance of trials for detection of real human falls, the classification algorithm could be improved by training the classification model, using various weights of "Rescue Randy" dolls in a wider variety of kinds of drops. Moreover, the system can be trained with drops of more objects on the floor and carpet, and other events (such as walking, environmental noise, etc.) that result in a significant floor vibration and have sufficient energy to trigger the event detection algorithm. Another future set of experiments to test the performance of the system is drops of "Rescue Randy" and the objects in distances that are larger than 5 m. Evaluation of the distance in which the sensitivity is lower than 95% could help planning the location and number of the sensors in big rooms. Generally, we think that a real system will require one accelerometer and microphone in a small room ($\sim 10\ \mathrm{m}^2$), and two sensor units in a large room ($\sim 25\ \mathrm{m}^2$). The sensors should be installed in one (or two) corners of the room.

For now, the event detection algorithm that is based on energy calculations is relatively simple, but it worked well in our experiments. In case of real human falls, that algorithm may not be good enough for the detection of "soft" human falls and events in which the individual collides with another object (table, dresser, chair, etc.). Therefore, different and more complicated algorithms for event detection might be needed.

Ambient noise such as music and TV would not influence the detection because the algorithm has to detect a floor vibration event in the first stage. In spite of this, such ambient noise can influence the classification of the events. In the future, we plan to check this issue, and if needed, to think about ways to prevent that influence on the classification results.

The system is adaptive, and can be calibrated to any kind of floor (that can be made of wood, concrete, granite, etc.). The signals that have to be calibrated are the vibration signals of that are measured by the accelerometer. It can be done efficiently by a calibration procedure that estimates the transmission function of a vibration signal that propagates in the floor by a pre determined experiments that evocates the floor in a certain point and measurement of the reaction in a certain point in a predefined distance. By knowledge of the transmission function of the signals from the training phase trials and the transmission function of the signals of the specific floor, a calibration of the vibration signals of the floor can be performed easily.

In this presented preliminary research, we chose to perform the trials on a typical concrete floor (popular in Israel). This kind of floor was preferred because we wanted the algorithm to be able to detect and classify the events in real conditions. In real life situations, there is a variety of different constructions, floor cover materials, etc.; this may potentially have a significant effect on the signals. Therefore, in the process of installment of a real system there should be a process of calibration of the algorithms to the kind of floor.

In the near future, we plan to design and produce a prototype of a fall detection system and perform a large scale experiments with real human falls in a nursing home and with stunt people (for better simulation of human falls) to train and test the classification algorithm of the system. Moreover, other events (such as object falls, steps, more door slams, etc.) can be collected in this experiment and analyzed later. The experiments with real human falls are very important for the final verification of the algorithm and the concept of the fall detection system. Following these experiments, an improvement of the classification algorithm can be performed. These future steps will hopefully lead to an ultimate solution for detection of human falls at home with high sensitivity and specificity.

References

[1] C. J. L. Murray and A. D. Lopez, "Global and regional descriptive epidemiology of disability: Incidence, prevalence, health expectancies and years lived with disability," *Global Burden Dis.*, vol. 1, pp. 201–246, 1996.

[2] T. Masud and R. O. Morris, "Epidemiology of falls," *Age Ageing*, vol. 30, no. s4, pp. 3–7, 2001.

[3] S. R. Lord, C. Sherrington, and H. B. Menz, *Falls in Older People: Risk Factors and Strategies for Prevention*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[4] J. Porteus and S. Brownsell, "Using telecare: Exploring technologies for independent living for older people," presented at the Anchor Trust, Kidlington, U.K., 2000.

[5] A. Diaz, M. Prado, L. M. Roa *et al.*, "Preliminary evaluation of a full-time falling monitor for the elderly," in *Proc. 26th Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2004, vol. 1, pp. 2180–2183.

[6] A. K. Bourke, J. V. O'Brien, and G. M. Lyons, "Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm," *Gait Posture*, vol. 26, no. 2, pp. 194–199, 2007.
[7] G. Williams, K. Doughty, K. Cameron *et al.*, "A smart fall and activity monitor for telecare applications," in *Proc. 20th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 1998, vol. 3, pp. 1151–1154.
[8] Y. Depeursinge, J. Krauss, and M. El-Khoury, "Device for monitoring the activity of a person and/or detecting a fall, in particular with a view to providing help in the event of an incident hazardous to life or limb," U.S. Patent 6 201 476, 2001.
[9] N. Noury, A. Tarmizi, D. Savall *et al.*, "A smart sensor for the fall detection in daily routine," presented at the SICICA, Aveiro, Portugal, Jul. 2003.
[10] N. Noury, A. Fleury, P. Rumeau *et al.*, "Fall detection—Principles and methods," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2007, pp. 1663–1666.
[11] C. Rougier and J. Meunier, "Demo: Fall detection using 3D head trajectory extracted from a single camera video sequence," in *Proc. 1st Int. Workshop Video Process. Security (VP4S-06)*, Jun. 2006, pp. 7–9.
[12] T. Lee and A. Mihailidis, "An intelligent emergency response system: Preliminary development and testing of automated fall detection," *J. Telemedicine Telecare*, vol. 11, no. 4, pp. 194–198, 2005.
[13] G. Wu, "Distinguishing fall activities from normal activities by velocity characteristics," *J. Biomech.*, vol. 33, no. 11, pp. 1497–1500, 2000.
[14] C. W. Lin and Z. H. Ling, "Automatic fall incident detection in compressed video for intelligent homecare," in *Proc. 16th Int. Conf. Comput. Commun. Netw. (ICCCN 2007)*, pp. 1172–1177.
[15] D. Litvak, Y. Zigel, and I. Gannot, "Fall detection of elderly through floor vibrations and sound," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008, pp. 4632–4635.
[16] M. Alwan, P. Rajendran, S. Kell *et al.*, "A smart and passive floor-vibration based fall detector for elderly," in *Proc. 2nd IEEE Int. Conf. Inf. Commun. Technol.: From Theory Appl. (ICITTA 2006)*, pp. 23–28.
[17] T. Irvine. (2002, May 24). An introduction to the shock response spectrum. [Online]. Available: http://www.vibrationdata.com/tutorials2/srs_intr.pdf
[18] S. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 4, pp. 357–366, Aug. 1980.
[19] S. N. Robinovitch, W. C. Hayes, and T. A. McMahon, "Distribution of contact force during impact to the hip," *Ann. Biomed. Eng.*, vol. 25, no. 3, pp. 499–508, 1997.
[20] Y. Zigel and A. Cohen, "Text-dependent speaker verification using feature selection with recognition related criterion," in *Proc. Odyssey Speaker Language Recog. Workshop*, Toledo, Spain, May 2004, pp. 329–336.
[21] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.
[22] J. Kittler, "Feature set search algorithms," in *Pattern Recognition and Signal Processing*, C. H. Chert, Ed. Mphen aan den Rijn, The Netherlands: Sijthoff and Noordhoff, 1978, pp. 41–60.
[23] M. Sambur, "Selection of acoustic features for speaker identification," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 23, no. 2, pp. 176–182, Apr. 1975.
[24] P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recog. Lett.*, vol. 15, no. 11, pp. 1119–1125, 1994.
[25] D. Zongker and A. Jain, "Algorithms for feature selection: An evaluation," in *Proc. 13th Int. Conf. Pattern Recog., 1996*, vol. 2, pp. 18–22.
[26] P. C. Mahalanobis, "On the generalized distance in statistics," *Nat. Inst. Sci. India*, vol. 2, no. 1, pp. 49–55, 1936.
[27] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part 1. Detection, Estimation, and Linear Modulation Theory*. New York: Wiley, 1968.
[28] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. San Francisco, CA: Academic, 1990.
[29] R. O. Duda and E. Peter, *Hard. Pattern Classification and Scene Analysis*. New York: Wiley-Interscience, 1973.
[30] T. W. O'Neill, J. Varlow, A. J. Silman *et al.*, "Age and sex influences on fall characteristics," *Ann. Rheum. Dis.*, vol. 53, no. 11, pp. 773–775, 1994.
[31] P. Rajendran, A. Corcoran, B. Kinosian *et al.*, "Falls, fall prevention, and fall detection technologies," in *Eldercare Technology for Clinical Practitioners,* Totowa, NJ: Humana Press, 2008, pp. 187–202.

Yaniv Zigel (S'96–M'03) was born in Tel-Aviv, Israel, in 1970. He received the B.Sc. degree, in 1992, the M.Sc. degree (*cum laude*), in 1998 , and the Ph.D. degree, in 2004, all in electrical and computer engineering from Ben-Gurion University, Beer-Sheva, Israel. His M.Sc. thesis and Ph.D. research were in the field of biomedical and speech signal processing.

From 1992 to 1995, he was a Technical Officer with the Communication and Computer Corps, Israeli Defense Forces. From 2003 to 2006, he was a Senior Digital Signal Processor Algorithm Engineer with the Audio Analysis Group, NICE Systems Ltd., Ra'annana. During 2006 and 2007, he was a Research Group Leader with PuddingMedia Ltd., and a Guest Lecturer with the Department of Biomedical Engineering, Tel-Aviv University, Tel-Aviv. He is a Lecturer and Head of Biomedical Signal Processing Research Laboratory, Department of Biomedical Engineering, Faculty of Engineering, Ben-Gurion University, Beer-Sheva, since 2007. His current research interests include biomedical signal processing, audio and speech analysis, speaker and speech recognition, pattern recognition, and signal compression.

Dima Litvak was born in Moscow, Russia, in 1982. He received the B.Sc. degree in mathematics and physics, in 2006, and the M.Sc. degree in biomedical engineering, in 2009 from Tel-Aviv University, Tel-Aviv, Israel.

His M.Sc. thesis was in the field of signal processing and pattern recognition. He was a Teaching Assistant in lasers and optics in medicine, and laser–tissue interaction with Tel-Aviv University, during the M.Sc. course. He is currently with the Department of Biomedical Engineering, Faculty of Engineering, Tel-Aviv University. He is also a Patent Counsel with Dr. Eyal Bressler & Co., Ramat Gan, in 2009. His current research interests include medical devices, mechanics, and algorithms.

Israel Gannot received the B.Sc. degree in electrical engineering from the Technion-Israeli Institute of Technology, Haifa, Israel, in 1981, and the M.Sc. and Ph.D. degrees in biomedical engineering from Tel-Aviv University, Tel-Aviv, in 1989 and 1994, respectively.

During 1981 through 1987, he was with the Israeli Defense Forces Medical Corps as a Section Chief in biomedical engineering. During 1994 to 1997, he was a National Academy Sciences Postdoctoral Fellow with the Electro-optics Branch, Office of Science and Technology, US Food and Drug Administration, Silver Spring, MD. He is with the Department of Biomedical Engineering, Faculty of Engineering, Tel-Aviv University, since 1997, where he is currently an Associate Professor. During 2002 through 2005, he was a Senior Research Scientist with the National Institutes of Health, Bethesda, MD. He is currently the Chair of the Biomedical Engineering Department, Tel-Aviv University. He is also a Visiting Professor of biomedical engineering with the Program of Biomedical Engineering, Department of Electrical and Computer Engineering, School of Engineering and Applied Sciences, George Washington University, Washington, DC. He is the author of more than 45 peer-reviewed scientific papers, 85 proceeding papers, six book chapters, and 14 patents.

Dr. Gannot is a Fellow of the International Society for Optical Engineering. A fellow of the American Society for Laser Medicine and Surgery and of the American Institute of Medical and Biological Engineering. The fellow level was conferred upon for his exceptional achievements in his field of research. He holds the Chair of conferences and symposia in his field of research on regular basis.