



Fall detection using single-tree complex wavelet transform



Ahmet Yazar^{a,*}, Furkan Keskin^a, B. Uğur Töreyn^b, A. Enis Çetin^a

^aBilkent University, TR-06800 Bilkent, Ankara, Turkey

^bÇankaya University, TR-06810 Yenimahalle, Ankara, Turkey

ARTICLE INFO

Article history:

Available online 28 December 2012

Keywords:

Vibration sensor
PIR sensor
Falling person detection
Feature extraction
Single-tree complex wavelet transform
Support vector machines

ABSTRACT

The goal of Ambient Assisted Living (AAL) research is to improve the quality of life of the elderly and handicapped people and help them maintain an independent lifestyle with the use of sensors, signal processing and telecommunications infrastructure. Unusual human activity detection such as fall detection has important applications. In this paper, a fall detection algorithm for a low cost AAL system using vibration and passive infrared (PIR) sensors is proposed. The single-tree complex wavelet transform (ST-CWT) is used for feature extraction from vibration sensor signal. The proposed feature extraction scheme is compared to discrete Fourier transform and mel-frequency cepstrum coefficients based feature extraction methods. Vibration signal features are classified into “fall” and “ordinary activity” classes using Euclidean distance, Mahalanobis distance, and support vector machine (SVM) classifiers, and they are compared to each other. The PIR sensor is used for the detection of a moving person in a region of interest. The proposed system works in real-time on a standard personal computer.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Most European countries will face a large increase in the number of elderly people in the near future. The development of intelligent homes will improve the quality of life of seniors and the handicapped people. It is reported that unexpected falling is a major problem and about one-third of people over 65 falls unexpectedly each year (Hausdorff et al., 2001). Fall related health and injury costs are on the order of billions of dollars worldwide.

Fall detection systems will not only help save billions of dollars, but also provide elderly people an independent way of life (Marquis-Faulkes et al., 2004; Nait-Charif and McKenna, 2004). Currently, the commercially available fall detection systems feature body worn sensors which are connected to a wireless sensor network (Brownsell et al., 2011; Nyan et al., 2008). There are other recent studies which turn a mobile phone into a fall detector (Dai et al., 2010). The main disadvantage of these systems is that wearing and/or carrying them is not convenient all the time and elderly people may simply forget to put on these sensors.

Computer vision and sound sensor based systems can be also used for fall detection (Töreyn et al., 2005, 2008). However, most people find having a camera in their houses invasion of their privacy, even if the video and sound recordings are not transmitted elsewhere. Vibration sensor based fall detection systems are proposed by Töreyn et al. (2008) and Alwan et al. (2006). However,

pattern recognition algorithms for fall detection are not described in both articles. Without a pattern recognition algorithm the resulting system will not be a robust and practical system.

Our system is different from the currently available AAL systems. We propose to install both vibration and PIR sensors to intelligent homes to realize a robust system. Vibration and PIR sensors can be placed on the floor and the ceiling of each room of the house. These two different types of sensors complement each other and daily activities of seniors can be monitored without producing false alarms and unusual events like falling can be detected. The resulting AAL system will be a low-cost and privacy-friendly system thanks to the types of sensors used.

In this study, individual decisions of signal classifiers from vibration and PIR sensors are fused to detect a falling person. The Euclidean distance classifier, the Mahalanobis distance classifier, and the SVM classifier are trained for the vibration signals and a threshold-based classifier is used for the PIR sensor signals. A single-tree complex wavelet transform (ST-CWT) is used for feature extraction from vibration sensor signal. The feature extraction scheme is compared to discrete Fourier transform and mel-frequency cepstrum coefficients based feature extraction methods. The most important characteristic of the complex wavelet transform is that it possesses shift invariance and directional selectivity properties. Dual-tree CWT has been shown to be a promising alternative to classical DWT for a variety of signal and image processing tasks. However, DT-CWT suffers from computational burden and redundancy caused by the exploitation of two wavelet trees. In this article, we employ single-tree CWT for fall detection, which is designed using computationally efficient lifting structures in a

* Corresponding author. Tel.: +90 312 290 1525; fax: +90 312 266 4192.

E-mail addresses: yazar@ee.bilkent.edu.tr (A. Yazar), keskin@ee.bilkent.edu.tr (F. Keskin), toreyn@cankaya.edu.tr (B.U. Töreyn), cetin@bilkent.edu.tr (A.E. Çetin).

single-tree. The single-tree filter-bank producing CWT coefficients has time-varying update and prediction filters for lifting structures. Time-varying nature of the proposed filters enables the implementation of half-sample delayed filters of dual-tree CWT in a single-tree in an interchangeable manner. This leads to a time-invariant feature extraction method. Single-tree CWT is superior to both ordinary DWT because the CWT is time-invariant and dual-tree CWT which doubles the amount of data after the analysis stage. Vibration signal energies in different frequency sub-bands produced by ST-CWT are taken as feature parameters. These feature parameters are used for classifying vibration signals as “fall” or “ordinary activity” once a moving person is detected by the PIR sensor.

In Section 2, the ST-CWT method is introduced. Analysis of PIR sensor signal is presented in Section 3. In addition to ST-CWT, two different frequency based feature extraction schemes from vibration signals are studied and the three feature extraction schemes are compared to each other. The classification methods and experimental setup are described in Section 4. Finally, conclusions are drawn in the last section.

2. Single-tree CWT

In this article, single-tree CWT is used for feature extraction from sensor signals. The classical discrete wavelet transform (DWT) has several limitations hampering its effectiveness in signal and image analysis, such as time-variance and lack of directionality (Selesnick et al., 2005). It is well-known that the discrete wavelet coefficients may change significantly when the input is shifted slightly. To overcome such limitations of DWT, the dual-tree complex wavelet transform (DT-CWT) is proposed whereby two filter pairs are used in parallel to decompose a given signal (Kingsbury, 1998). In contrast to the real DWT, two sets of filters are employed in the two wavelet trees, which are called real and imaginary trees, respectively. Two different DWTs are executed in parallel in dual-tree structure where the real part of DT-CWT is provided by the first one and the imaginary part by the second one. The reasoning behind the use of dual-tree is to obtain an analytic complex wavelet $\psi_c(t)$:

$$\psi_c(t) = \psi_h(t) + j\psi_g(t) \quad (1)$$

where $\psi_h(t)$ and $\psi_g(t)$ denote wavelet functions of real and imaginary trees, respectively. If the $\psi_c(t)$ is approximately analytic (has support on only one-side of the frequency axis), the resulting transform can possess the properties such as lack of aliasing, directionality and shift invariance just like the Fourier transform whose complex basis functions are analytic (Selesnick et al., 2005). For $\psi_c(t)$ to be approximately analytic, it is required that one wavelet basis is the approximate Hilbert transform of the other wavelet basis:

$$\psi_g(t) \approx \mathcal{H}\{\psi_h(t)\} \quad (2)$$

In order to satisfy the condition in Eq. (2), low-pass analysis filters in real and imaginary trees must be offset approximately by half-sample (Selesnick, 2001):

$$g_0[n] \approx h_0[n - 0.5] \quad (3)$$

In the literature, two low-pass filters are jointly designed such that half-sample delay, perfect reconstruction and finite support conditions are simultaneously satisfied, using several filter design methods (Selesnick et al., 2005). We focus on the q -shift filter design in this paper and employ them to obtain time-varying lifting filters. Analysis q -shift filters for real and imaginary trees are shown in Table 1 (Kingsbury, 2000).

Table 1

Impulse response of Kingsbury's 8th order q -shift analysis filters for the dual-tree CWT. They are normalized so that $\sum_n h_0[n] = 1$.

Analysis filters	h_0	h_1	g_0	g_1
Q -shift filter coefficients	0.0248	-0.0808	-0.0808	-0.0248
	0	0	0	0
	-0.0624	0.4155	0.4155	0.0624
	0.1653	-0.5376	0.5376	0.1653
	0.5376	0.1653	0.1653	-0.5376
	0.4155	0.0624	-0.0624	0.4155
	0	0	0	0
	-0.0808	-0.0248	0.0248	-0.0808

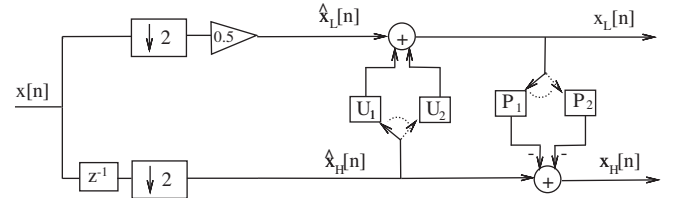


Fig. 1. Time-varying lifting scheme for single-tree CWT.

2.1. Time-varying lifting structures for single-tree CWT

We design real-valued lifting filters to be used in the single-tree context to obtain a transform that is approximately complex in the sense of DT-CWT without causing redundancy and computational burden. The implementation scheme of our real-valued ST-CWT design is shown in Fig. 1.

In Fig. 1, $U_{1,2}$ and $P_{1,2}$ denote two different sets of update and prediction filters, respectively. Since the aim is to construct an approximately complex wavelet transform using only one tree, the first update filter U_1 must correspond to the low-pass analysis filter of the real tree $h_0[n]$ and the second update filter U_2 must correspond to the low-pass analysis filter of the imaginary tree $g_0[n]$, of DT-CWT, respectively.

2.2. Lifting filter design

Even and odd samples of the subsignal $x_L[n]$ are obtained using $U_1[n]$ and $U_2[n]$, respectively. Similarly, even and odd samples of $x_H[n]$ are obtained using $P_1[n]$ and $P_2[n]$, respectively. Using U_1 and U_2 in a sequentially switched manner for low-pass filtering of the input signal, we construct a time-varying single-tree lifting structure that keeps the benefits of DT-CWT. More formally, the input signal is first divided into even-indexed samples $\hat{x}_L[n]$ and odd-indexed samples $\tilde{x}_L[n]$ through a lazy filter-bank. Even-indexed samples of $\hat{x}_L[n]$ is updated by U_1 and odd-indexed samples of $\tilde{x}_L[n]$ is updated by U_2 . Let $\tilde{h}_1[n]$ and $\tilde{h}_2[n]$ denote the effective half-band lowpass filters processing the input signal $x[n]$ before downsampling. Their z -transforms are given by

$$\tilde{H}_1(z) = 1/2 + z^{-1}U_1(z^2) \quad (4)$$

$$\tilde{H}_2(z) = 1/2 + z^{-1}U_2(z^2) \quad (5)$$

We design filters $U_1(z)$ and $U_2(z)$, or equivalently $\tilde{H}_1(z)$ and $\tilde{H}_2(z)$ using the following constraints so that the resulting transform is approximately complex:

- (i) Since $\tilde{h}_i[n]$ is a half-band filter, $\tilde{h}_i[2n] = 0$ for $n \neq 0$, $i = 1, 2$, for perfect reconstruction in a lifting structure.
- (ii) Filters $\tilde{h}_1[n]$ and $\tilde{h}_2[n]$ must have approximate group delays of $1/4$ and $3/4$, respectively so that there exist 0.5 delay difference between the two filters (Abbas and Tran, 2008).

(iii) Filters $\tilde{H}_1(z)$ and $\tilde{H}_2(z)$ must have a zero at $z = -1$, that is, $\sum_n \tilde{h}_i[n](-1)^n = 0$ for $i = 1, 2$ so that $\tilde{H}_i(e^{jw}) = 0$ at $w = \pi$.

Based on the constraint (i) the 7th order FIR filter should be in the following form:

$$\tilde{h}_1[n] = \{\alpha_1, 0, \alpha_2, \alpha_3, \alpha_4, 0, \alpha_5\} \quad (6)$$

where α_3 denotes the coefficient at $n = 0$. We can use the three dominant center coefficients of h_0 from Table 1 to obtain α_2 , α_3 and α_4 as follows

$$\alpha_2 = 0.1538, \quad \alpha_3 = 0.5, \quad \alpha_4 = 0.3864 \quad (7)$$

which are scaled versions of $h_0[4]$, $h_0[5]$, and $h_0[6]$, respectively. Since the filter coefficients in Eq. (6) must sum to one, we have

$$\alpha_1 + \alpha_5 = -0.0402 \quad (8)$$

To satisfy the constraint (iii), we need

$$\alpha_3 - \sum_{i \neq 0} \alpha_i = 0 \quad (9)$$

which is already satisfied by setting $\alpha_3 = 0.5$. The final constraint to satisfy is the half-sample delay the constraint (ii). The group delay of the filter $\tilde{h}_1[n]$ is given by

$$\tau_g(w) = -\frac{\partial \phi(w)}{\partial w} \quad (10)$$

where $\phi(w) = \arg\{\tilde{H}_1(e^{jw})\}$ is the phase of the DTFT of $\tilde{h}_1[n]$. The frequency response of $\tilde{h}_1[n]$ is given by:

$$\tilde{H}_1(e^{jw}; \alpha_1) = \alpha_1 e^{3jw} + \sum_{i=2}^4 \alpha_i e^{(3-i)jw} + (-0.0402 - \alpha_1) e^{-3jw} \quad (11)$$

where α_1 is the only unknown. The filter coefficient α_1 can be easily determined by one-dimensional exhaustive search in the interval $[-1, 1]$. First, for each $\alpha \in [-1, 1]$ we fit a linear model to the phase $\phi(w; \alpha) = \arg\{\tilde{H}_1(e^{jw}; \alpha)\}$. The reason is that the q -shift filters are approximately linear phase (Kingsbury, 2000) and have almost constant group delay. Fitting process is performed for low frequencies ($w \in [\frac{\pi}{8}, \frac{\pi}{2}]$) because approximately linear behavior of the phase function disappears as the w approaches $\pm\pi$. After fitting the linear model, the negative slope of the resultant line yields the group delay of the filter from Eq. (10). To have a group delay of $1/4$, it turns out that

$$\alpha_1 = -0.05, \quad \alpha_5 = 0.0098 \quad (12)$$

The second filter $\tilde{h}_2[n]$ is simply the time reversed version of the filter $\tilde{h}_1[n]$. This is similar to the time reversed design of $\{h_0, g_0\}$ filter pair in (Kingsbury, 2000). Hence, $\tilde{h}_2[n]$ is given by

$$\tilde{h}_2[n] = \{\alpha_5, 0, \alpha_4, \alpha_3, \alpha_2, 0, \alpha_1\} \quad (13)$$

Since $\tilde{h}_2[n]$ is the time-reversed version of $\tilde{h}_1[n]$ they approximately satisfy the half-sample delay condition given in Eq. (3). It is possible to implement these two filters after decimation because they can be expressed in half-band form given in Eqs. (4) and (5) where

$$U_1(z^2) = \alpha_5 z^{-2} + \alpha_4 + \alpha_2 z^2 + \alpha_1 z^4 \quad (14)$$

$$U_2(z^2) = \alpha_1 z^{-2} + \alpha_2 + \alpha_4 z^2 + \alpha_5 z^4 \quad (15)$$

Prediction filters P_1 and P_2 are designed by applying the same design strategy as in update filters. In prediction, P_1 uses only those samples of the signal $x_L[n]$ which are updated by U_1 and P_2 uses only those samples of the signal $x_L[n]$ which are updated by U_2 . From Table 1, $h_1[n] = (-1)^n h_0[N-n]$ where N is the length of the filter. Thus, effective prediction filter corresponding to P_1 is given by

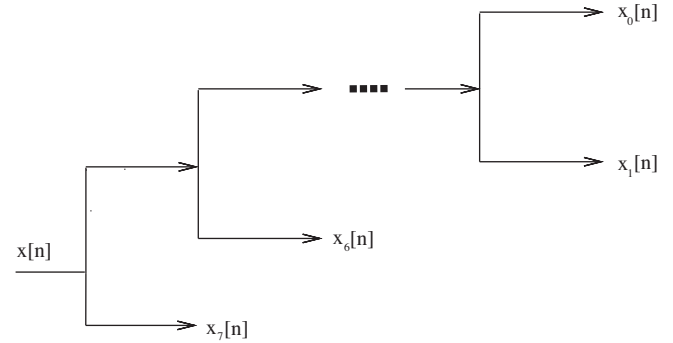


Fig. 2. Seven-level wavelet tree.

$$\tilde{g}_1[n] = \{-\alpha_5, 0, -\alpha_4, \alpha_3, -\alpha_2, 0, -\alpha_1\} \quad (16)$$

Since $g_1[n] = h_1[N-n]$ from Table 1, effective prediction filter corresponding to P_2 is given by

$$\tilde{g}_2[n] = \{-\alpha_1, 0, -\alpha_2, \alpha_3, -\alpha_4, 0, -\alpha_5\} \quad (17)$$

Update and prediction filters designed above are employed at the second decomposition level or higher. For the first level, half-sample delay condition in Eq. (3) becomes one-sample delay condition for DT-CWT to be approximately analytic at each level (Selesnick et al., 2005). Hence, simple $\{1/2, 1/2\}$ filter is used as the effective update filter at the first level, and the coefficient at $n = 0$ is changed between U_1 and U_2 . For prediction at the first level, $\{-1/4, 1, 3/4\}$ effective prediction filter is employed, which assigns weights to the samples based on their proximity to the predicted sample.

Vibration sensor signals are fed to a seven-stage wavelet-tree as shown in Fig. 2 and wavelet energy values are used in the feature vector for feature extraction.

2.3. Feature extraction using ST-CWT

Seven-level single-tree complex wavelet transform is applied to 2048-sample-long signal windows using the lifting filters in Section 2.2. The seven-level ST-CWT divides the input signals into angular frequency sub-bands $[0, \pi/128]$, $[\pi/128, \pi/64]$, $[\pi/64, \pi/32]$, $[\pi/32, \pi/16]$, $[\pi/16, \pi/8]$, $[\pi/8, \pi/4]$, $[\pi/4, \pi/2]$, and $[\pi/2, \pi]$ which correspond to $[0, f_s/256]$, $[f_s/256, f_s/128]$, $[f_s/128, f_s/64]$, $[f_s/64, f_s/32]$, $[f_s/32, f_s/16]$, $[f_s/16, f_s/8]$, $[f_s/8, f_s/4]$, and $[f_s/4, f_s/2]$ Hz bands, respectively. In this study, the sampling frequency f_s is selected as 900 Hz. In a seven-level wavelet tree the sub-signal $x_0[n]$ comes from $[0, \pi/128]$, $x_1[n]$ comes from $[\pi/128, \pi/64]$, $x_2[n]$ comes from $[\pi/64, \pi/32]$, ..., and $x_7[n]$ comes from $[\pi/2, \pi]$ frequency sub-bands of the sensor signal $x[n]$, respectively. Feature parameters are extracted by finding the energies of resulting frequency sub-bands. The feature vector for a frame of 2048 sample sensor signal is given by

$$v = [\|x_0\|^2 \|x_1\|^2 \dots \|x_7\|^2]^T \quad (18)$$

The ST-CWT is compared to the DFT and the mel-cepstrum based feature extraction schemes in Section 4.

3. Analysis of PIR sensor signal

PIR sensor is used to prevent possible false alarms which may occur due to slamming a door, falling objects, or other non-human vibration signal sources (Toreyin et al., 2008).

The strength of the received signal from the PIR sensor changes with motion of a hot object within the viewing range of the sensor. Distance between the PIR sensor and hot object affects the

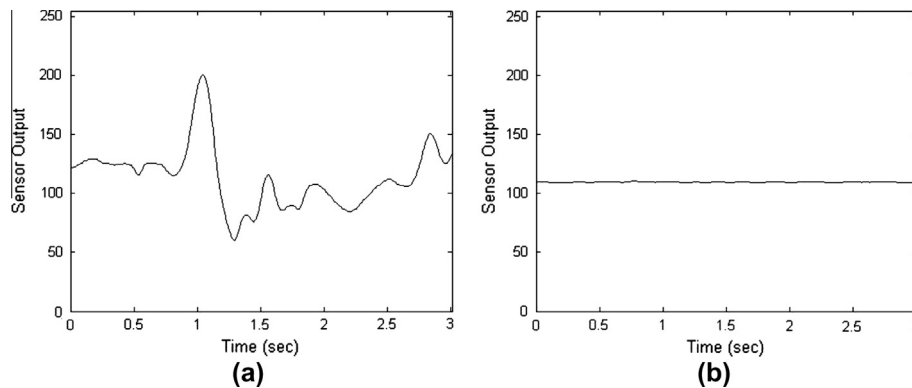


Fig. 3. Time-domain PIR sensor signal records for (a) falling and (b) no-activity cases. The y-axis is proportional to the voltage value of the analog sensors.

amplitude of the sensor signal. Typical sensor signals corresponding to falling and no-activity cases are shown in Fig. 3. The sampling rate is 100 Hz with 8-bit resolution for these records. The primary task of the PIR sensor is to detect human motion in a specific region of interest. Alarms produced by other sensors are ignored when there is no motion in a room.

The PIR sensor signal by itself is not reliable for fall detection. A walking person may also produce a signal as shown in Fig. 3(a). Therefore, we can only use PIR sensor signal as a first discriminatory step in our multi-sensor fall detection system. It is trivial to distinguish “activity” and “no-activity” cases using the PIR sensor signal. As a result, a simple thresholding method is used to detect “no-activity” cases.

If there is a pet at home then simple thresholding is not sufficient. It is possible to distinguish walking pets from human beings using the PIR sensor signal, because leg movements of pets are much frequent than human beings (Toreyin et al., 2008). In this case, a two-level wavelet tree is used for feature extraction from the PIR sensor signal and the wavelet energy vector fed to a Markov-model based classification engine. When the vibration system detects a “fall” and only pet activity exists after that an alarm is still issued in spite of the motion. Because this motion is probably due to the pet not the owner of the house.

4. Experimental setup and studies

In this article, GS-20DX vibration sensor manufactured by OYO Technologies is used. This sensor can detect vibration signals from distances up to 25 m. Therefore, it may be enough to install a single vibration sensor into each room of an intelligent house. The vibration signal dataset is composed of records containing activities such as falling, walking and running at different speeds, slamming a door, sitting on the floor, and falling objects obtained from 10 people. Analog vibration signals are sampled at a rate of 900 Hz and digitized with 8-bit resolution.

Three-seconds-long signal records from vibration sensor corresponding to falling, walking and slamming a door are shown in Fig. 4. Note that, these are raw signals without pre-processing. Note also that, amplitudes of vibration signals differ, because they depend on the distance of the person to the sensor and on the pushing off strength of the foot. Walking very close to the sensor and/or stomping may result in high-amplitude vibration signals which may resemble falling. For this reason, fall detection systems based only on the signal strength may produce false alarms. To prevent false alarms, duration of the activity should be taken into account, as well. A typical fall lasts about two seconds. Therefore, a 2048-sample-long analysis window is selected in the proposed system.

Signals acquired by vibration sensors may contain various frequency components depending on the architecture of the building in which intelligent room is located. Moreover, running machines in the building may cause additional signal activity at different frequencies. Our system checks the strength of the vibration signal before analyzing the frequency content of the signal. Signal windows with high deviations from the average value are chosen as candidate fall regions. An experimentally determined value is used as the threshold. Let $M = \max |x(n)|$ and $m = \min |x(n)|$ within an analysis window of length 2048 samples. When $M-m$ is larger than the threshold, this window of data is further processed by various feature extraction schemes for fall detection. Flowchart of the real-time algorithm is shown in Fig. 5.

In this article, three different frequency extraction schemes from vibration signals are studied and compared to each other. Feature extraction using ST-CWT is described in Section 2.3 and in the following two subsections, the DFT and the mel-cepstrum based feature extraction schemes are described.

4.1. Feature extraction using discrete Fourier transform (DFT)

Frequency content of the vibration signal is analyzed in sub-bands. It is observed that most of the vibration signal energy is concentrated in low-bands. Hence, more emphasis is given to lower frequencies by assigning more sub-bands to them as shown in Fig. 6. Boundary frequencies of sub-bands are $f_s/32$, $f_s/16$, $3f_s/32$, $f_s/8$, $3f_s/16$, $f_s/4$, $3f_s/8$, and $f_s/2$ which correspond to 28.13 Hz, 56.25 Hz, 84.38 Hz, 112.5 Hz, 168.75 Hz, 225 Hz, 337.5 Hz, and 450 Hz if the sampling frequency, f_s , is set as 900 Hz. Energies of these frequency sub-bands are computed with Eq. (19) and used as eight feature parameters.

$$G(m) = \sum_{k \in B(m)} |X(k)|^2 \quad (19)$$

where $X(k)$ is the DFT of the vibration signal and represents the m th frequency band, for $m = 1, \dots, 8$. For the 2048-point DFT, $B(m)$ represents the indices of sub-bands [1,64], [65,128], [129,192], [193,256], [257,384], [385,512], [513,768], and [769,1024].

4.2. Feature extraction using mel-frequency cepstrum coefficients (MFCC)

The other feature extraction method is based on mel-cepstrum which is widely used in speech and image processing (Jabloun et al., 1999; Cakir and Cetin, 2010; Yorulmaz et al., 2012). The frequency decomposition of the MFCC is logarithmic allocating more bands to low-frequency sub-bands compared to high-frequency sub-bands. As discussed in Section 4.1, this approach is suitable

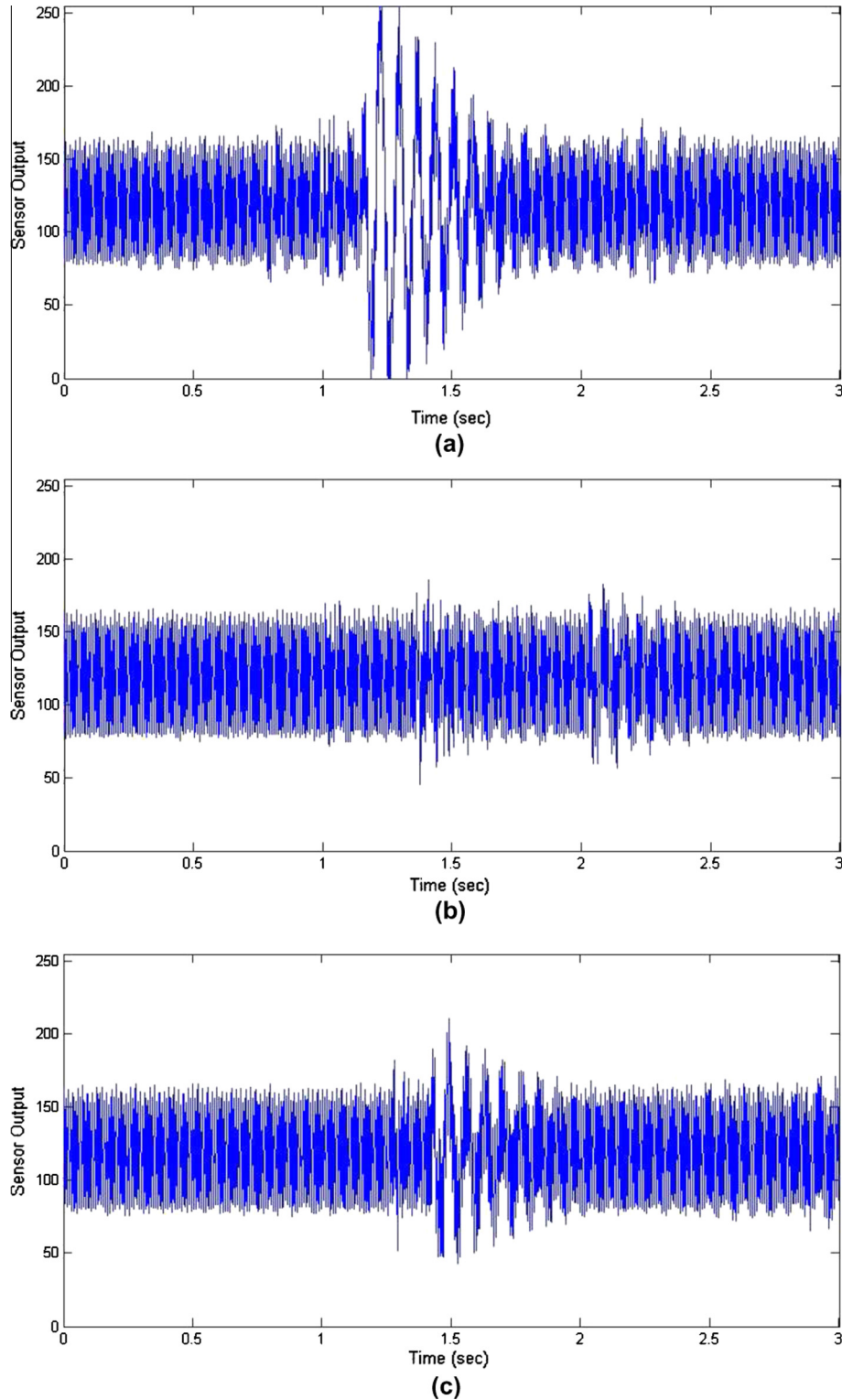


Fig. 4. Time-domain vibration sensor signals corresponding to (a) falling, (b) walking and (c) slamming a door. These are raw signals without pre-processing. The y-axis is proportional to the voltage value of the analog sensors.

to vibration signals. Most of signal energy is concentrated in low-frequency sub-bands therefore MFCC can be a good candidate for feature extraction from vibration signals.

As pointed out in Section 4.1 after computation of the DFT of the vibration signal window, energies of the frequency sub-bands are calculated with Eq. (19). Eight feature parameters are extracted by finding eight mel-frequency cepstrum coefficients, $C(u)$, with the following formula using discrete cosine transform (DCT):

$$C(u) = DCT^{-1}\{\log(G(m))\}, \quad u = 1, \dots, 8 \quad (20)$$

4.3. Computational complexity of the feature extraction algorithms

The DFT, the MFCC, and the ST-CWT are compared to each other in terms of computational complexity.

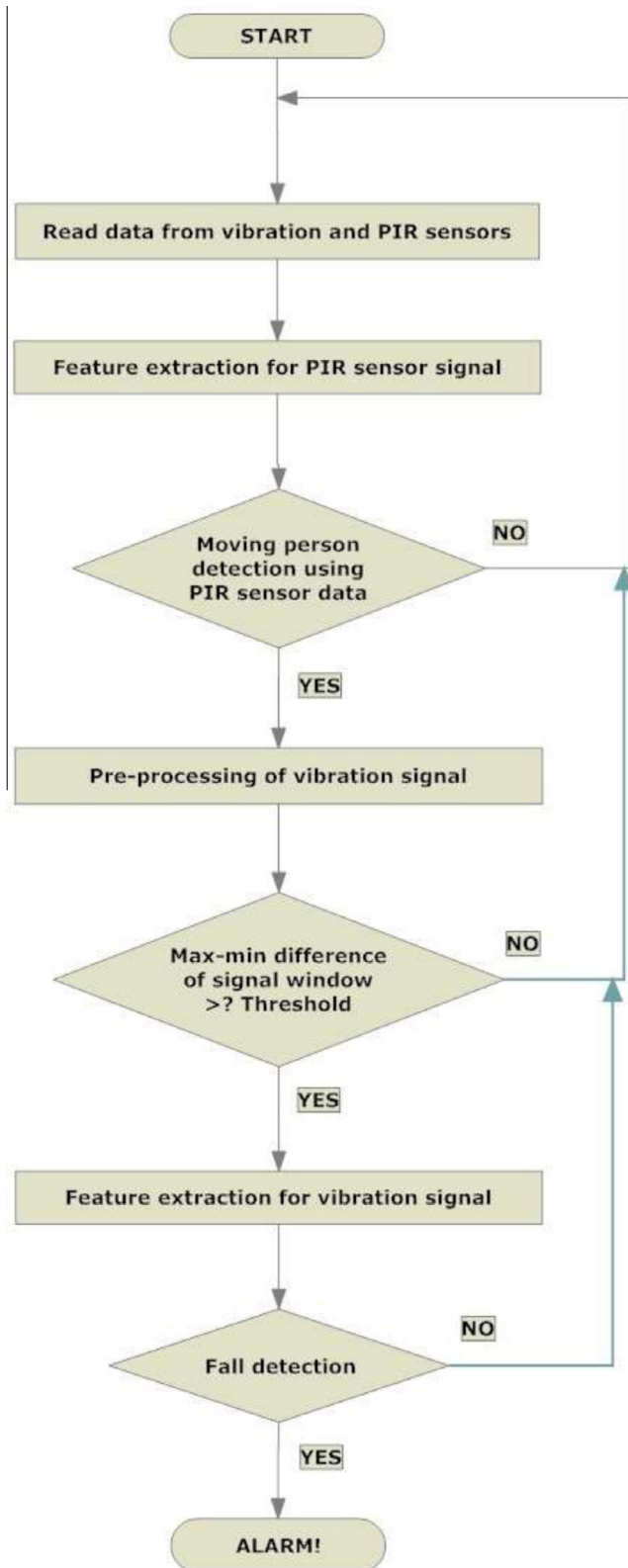


Fig. 5. Flowchart of the real-time algorithm.

Fast Fourier transform (FFT) is utilized instead of the DFT in the tests. The N -point FFT requires $N\log_2 N$ complex multiplications and additions. Besides, N real multiplications are needed to calculate the energies of the frequency sub-bands. Therefore, computational complexity of the DFT based feature extraction method is on the order of $O(N\log_2 N)$.

The mel-cepstrum based feature extraction scheme needs an extra $K \log_2 K$ real multiplications and additions to carry out the inverse discrete cosine transform, where $K = 8$ which is the number of mel-frequency cepstrum coefficients. Therefore, the computational complexity of the MFCC method is $O(N\log_2 N) + O(K \log_2 K)$. In addition, eight look-up operations are needed because the log operation is performed by a look-up table.

The computational cost of the dual-tree CWT is twice that of ordinary DWT which can be implemented in $O(N)$, where N is the number of samples in the signal (Romberg et al., 2001). The computational complexity of the ST-CWT is equivalent to the ordinary DWT. In each branch of Fig. 2, the computational cost is equivalent to a discrete convolution operation. We have a seven-level tree therefore 14 convolutional operations are performed. The number of nonzero filter coefficients is five in Eq. (6). To get a convolutional output, five real multiplications are performed and the energy values of eight sub-signals are computed as the features. The exact number of real multiplications is given in Table 2.

Feature extraction using the ST-CWT is a computationally efficient method in comparison to the DFT and the MFCC based feature extraction schemes due to the fact that the ST-CWT is designed using lifting structures in a single-tree.

4.4. Classification methods

After the feature extraction, classification for the vibration signals is done between “fall” and “ordinary activity”. For this purpose, the Euclidean distance classifier, the Mahalanobis distance classifier, and the SVM classifier are used and compared to each other.

The LIBSVM libraries are used for training and testing linear kernel based SVM classifier (Chang and Lin, 2011). It is experimentally observed that the linear kernel based SVM classifier is sufficient to classify the vibration signals as “fall” or “ordinary activity” instead of the more complex kernels like the polynomial and radial basis.

It is experimentally observed that the SVM classifier provides the best results however Mahalanobis distance classifier provides very close results as explained in the following sub-section in details.

4.5. Experimental results

The dataset contains 2048-sample-long signals corresponding to the following number of incident records: 100 falling, 1419 walking/running, 30 sitting on the floor, 30 slammed door, and 65 fallen book. Among falling and walking/running records, 40% of them are used for training classifiers. Similarly, one third of sitting on the floor, slammed door, and fallen book records are used for training. Remaining records are used as the test dataset.

Three different classifier models are formed for each of the classifiers using the single-tree complex wavelet transform, discrete Fourier transform, and mel-frequency cepstrum coefficients and the dataset is tested for each model. PIR sensor signals are analyzed with 200-sample-long windows. Variances of these signal windows are compared with a threshold which is established by test results. If the threshold value is exceeded then it is assumed that there is a human motion in a region of interest. Motion detection part works with 100% accuracy. Comparative results on the detection performance for different methods are presented in Tables 3 and 4. The results presented in the tables are obtained using the records in the test dataset.

Experimental results indicate that falling, walking, running, and sitting cases are classified into “fall” and “ordinary activity” classes with almost 100% accuracy using the ST-CWT feature extraction method with the Mahalanobis distance classifier or the SVM

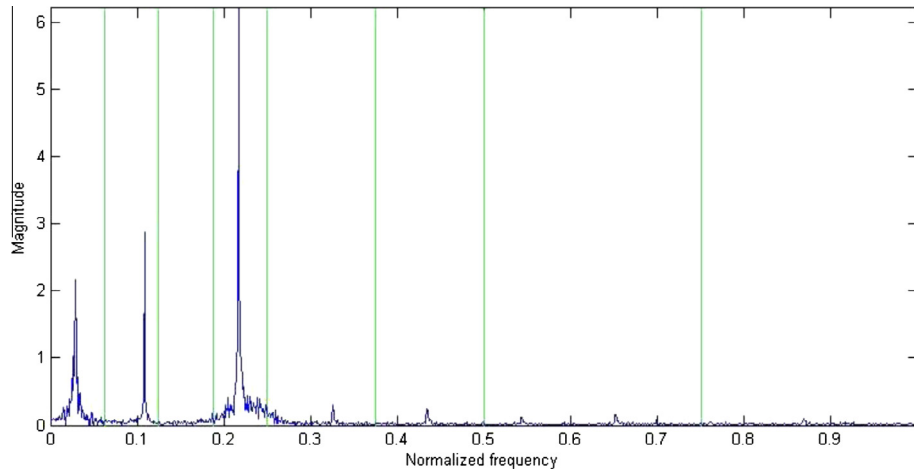


Fig. 6. Discrete Fourier transform of a vibration signal record of a falling person and separation boundary values of frequency sub-bands.

Table 2

The exact number of real multiplications in feature extraction methods. N is the number of signal samples in each window; K is the number of mel-frequency cepstrum coefficients; p is the number of nonzero filter coefficients, and r is the number of levels in a tree.

DFT	MFCC	ST-CWT
$2N \log_2 N + N$	$2N \log_2 N + N + 2K \log_2 K$	$2pN(1 - 1/2^r) + (r + 1)$
$N = 2048$	$N = 2048, K = 8$	$N = 2048, p = 5, r = 7$
47,104 multiplications	47,152 multiplications	20,328 multiplications

Table 3

Numbers of “true detection” versus “false alarm” for 2048-sample-long windows, using vibration sensor only.

		Falling	Ordinary activities		Non-human signal sources	
			Walking or running	Sitting	Slammed door	Fallen book
DFT	Euclidean	44/16	705/146	11/9	6/14	0/43
	Mahalanobis	60/0	782/69	19/1	18/2	14/29
	SVM	59/1	821/30	20/0	20/0	40/3
MFCC	Euclidean	60/0	778/73	14/6	15/5	15/28
	Mahalanobis	60/0	845/6	20/0	20/0	38/5
	SVM	60/0	842/9	20/0	20/0	40/3
ST-CWT	Euclidean	58/2	851/0	20/0	20/0	42/1
	Mahalanobis	60/0	851/0	20/0	20/0	42/1
	SVM	60/0	851/0	20/0	20/0	43/0

Table 4

Numbers of “true detection” versus “false alarm” for 2048-sample-long windows, using both PIR and vibration sensors.

		Falling	Ordinary activities		Non-human signal sources	
			Walking or running	Sitting	Slammed door	Fallen book
DFT	Euclidean	44/16	705/146	11/9	20/0	43/0
	Mahalanobis	60/0	782/69	19/1	20/0	43/0
	SVM	59/1	821/30	20/0	20/0	43/0
MFCC	Euclidean	60/0	778/73	14/6	20/0	43/0
	Mahalanobis	60/0	845/6	20/0	20/0	43/0
	SVM	60/0	842/9	20/0	20/0	43/0
ST-CWT	Euclidean	58/2	851/0	20/0	20/0	43/0
	Mahalanobis	60/0	851/0	20/0	20/0	43/0
	SVM	60/0	851/0	20/0	20/0	43/0

classifier. Euclidean distance classifier’s performance is not satisfactory in our dataset. Slamming a door and falling a book activities do not yield any false alarms when PIR sensor information is used.

As it can be observed from Tables 3 and 4 the ST-CWT is superior to discrete Fourier transform and mel-frequency cepstrum coefficients based features.

5. Conclusion

In this paper, vibration and PIR sensors are used for fall detection within a specific region of interest. A single-tree CWT based feature extraction method is introduced for vibration sensor signal analysis. Discrete Fourier transform and mel-frequency cepstrum coefficients are compared with ST-CWT based features. ST-CWT is superior to discrete Fourier transform and mel-cepstrum feature extraction methods in our dataset.

Human activity classification is carried out by the Euclidean distance, the Mahalanobis distance, and the SVM classifiers. PIR sensor output signal is also employed to reduce possible false alarms. The proposed system works in real-time on a standard personal computer and can be used as a part of an AAL monitoring system. Since the proposed system uses both PIR and vibration sensors it is more robust than single sensor based systems. This is important because high number of false alarms cause problems in AAL call centers and inconvenience to elderly people.

Acknowledgments

This work is supported in part by the Turk Telekom with Grant No. 3015-03. Authors are grateful to Karel Electronics Corporation for granting a GS-20DX vibration sensor.

References

- Abbas, A., Tran, T., 2008. Rational coefficient dual-tree complex wavelet transform: Design and implementation. *IEEE Trans. Signal Process.* 56 (8), 3523–3534.
- Alwan, M., Rajendran, P., Kell, S., Mack, D., Dalal, S., Wolfe, M., Felder, R., 2006. A smart and passive floor-vibration based fall detector for elderly. In: 2nd Information and Communication Technologies, ICTTA '06, vol. 1, pp. 1003–1007.
- Brownsell, S., Bradley, D., Cardinaux, F., Hawley, M., 2011. Developing a systems and informatics based approach to lifestyle monitoring within ehealth. Part I – Technology and data management. In: Proc. 1st IEEE Internat. Healthcare Informatics, Imaging and Systems Biology (HISB) Conf., pp. 264–271.
- Cakir, S., Cetin, A.E., 2010. Mel-cepstral feature extraction methods for image representation. *Opt. Eng.* 49 (9), 097004+.
- Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Systems Technol.* 2, 27:1–27:27. Software available at: <<http://www.csie.ntu.edu.tw/~cjlin/libsvm>>.
- Dai, J., Bai, X., Yang, Z., Shen, Z., Xuan, D., 2010. Mobile phone-based pervasive fall detection. *Pers. Ubiquitous Comput.* 14 (7), 633–643.
- Hausdorff, J.M., Rios, D.A., Edelberg, H.K., 2001. Gait variability and fall risk in community-living older adults: A 1-year prospective study. *Arch. Phys. Med. Rehabil.* 82 (8), 1050–1056.
- Jabloun, F., Cetin, A., Erzin, E., 1999. Teager energy based feature parameters for speech recognition in car noise. *Signal Process. Lett. IEEE* 6 (10), 259–261.
- Kingsbury, N., 1998. The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters. In: Proc. IEEE DSP Workshop, pp. 319–322.
- Kingsbury, N., 2000. A dual-tree complex wavelet transform with improved orthogonality and symmetry properties. In: Internat. Conf. on Image Processing, vol. 2, pp. 375–378.
- Marquis-Faulkes, F., McKenna, S.J., Gregor, P., Newell, A.F., 2003. Scenario-based drama as a tool for investigating user requirements with application to home monitoring for elderly people. In: 10th Internat. Conf. on Human-Computer Interaction, vol. 2004, 18 February, 2003.
- Nait-Charif, H., McKenna, S.J., 2004. Activity summarisation and fall detection in a supportive home environment. In: Proc. 17th Internat. Conf. Pattern Recognition ICPR 2004, vol. 4, pp. 323–326.
- Nyan, M.N., Tay, F.E.H., Murugasu, E., 2008. A wearable system for pre-impact fall detection. *J. Biomech.* 41 (16), 3475–3481.
- Romberg, J.K., Choi, H., Baraniuk, R.G., Kingsbury, N., 2001. A hidden markov tree model for the complex wavelet transform. *IEEE Trans. Signal Process.*, 133–136.
- Selesnick, I., 2001. Hilbert transform pairs of wavelet bases. *IEEE Signal Process. Lett.* 8 (6), 170–173.
- Selesnick, I., Baraniuk, R., Kingsbury, N., 2005. The dual-tree complex wavelet transform. *IEEE Signal Process. Mag.* 22 (6), 123–151.
- Toreyin, B.U., Dedeoglu, Y., Cetin, A.E., 2005. HMM based falling person detection using both audio and video. *Comput. Vision Human-Comput. Interact.* 3766, 211–220.
- Toreyin, B.U., Soyer, E.B., Onaran, I., Cetin, A.E., 2008. Falling person detection using multisensor signal processing. *EURASIP J. Adv. Signal Process.* 2008, 1–7.
- Yorulmaz, O., Pearson, T.C., Cetin, A.E., 2012. Detection of fungal damaged popcorn using image property covariance features. *Comput. Electron. Agric.* 84, 47–52.