Matching pursuit-based compressive sensing in a wearable biomedical accelerometer fall diagnosis device
Gibson, Ryan M.; Amira, Abbes; Ramzan, Naeem; Casaseca-de-la-Higuera, Pablo; Pervez, Zeeshan

Published in:
Biomedical Signal Processing and Control

DOI:
10.1016/j.bspc.2016.10.016

Publication date:
2017

Document Version
Peer reviewed version

Link to publication in ResearchOnline

Citation for published version (Harvard):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please view our takedown policy for details of how to contact us.
Matching Pursuit-Based Compressive Sensing in a Wearable Biomedical Accelerometer Fall Diagnosis Device

Ryan M. Gibson\textsuperscript{a,b,*}, Abbes Amira\textsuperscript{a,c}, Naeem Ramzan\textsuperscript{a}, Pablo Casaseca-de-la-Higuera\textsuperscript{a}, Zeeshan Pervez\textsuperscript{a}

\textsuperscript{a}University of the West of Scotland, Paisley, Scotland.
\textsuperscript{b}Present Address: Glasgow Caledonian University, Glasgow, Scotland.
\textsuperscript{c}Qatar University, Doha, Qatar.

Abstract

There is a significant high fall risk population, where individuals are susceptible to frequent falls and obtaining significant injury, where quick medical response and fall information are critical to providing efficient aid. This article presents an evaluation of compressive sensing techniques in an accelerometer-based intelligent fall detection system modelled on a wearable Shimmer biomedical embedded computing device with Matlab. The presented fall detection system utilises a database of fall and activities of daily living signals evaluated with discrete wavelet transforms and principal component analysis to obtain binary tree classifiers for fall evaluation. 14 test subjects undertook various fall and activities of daily living experiments with a Shimmer device to generate data for principal component analysis-based fall classifiers and evaluate the proposed fall analysis system. The presented system obtains highly accurate fall detection results, demonstrating significant advantages in comparison with the thresholding method presented. Additionally, the presented approach offers advantageous fall diagnostic information. Furthermore, transmitted data accounts for over 80% battery current usage of the Shimmer device, hence it is critical the acceleration data is reduced to increase transmission efficiency and in-turn improve

\*Corresponding Author

Email address: ryan.gibson@gcu.ac.uk (Ryan M. Gibson)
battery usage performance. Various Matching pursuit-based compressive sensing techniques have been utilised to significantly reduce acceleration information required for transmission.

Keywords: Fall detection, compressive sensing, multiresolution analysis, principal component analysis, acceleration signal evaluation, wearable device

1. Introduction

Studies demonstrate elderly individuals pose a significant risk to falling [1, 2], in addition to individuals with medical conditions, for example, gait [3] and neurodegenerative disorders, including epilepsy [4] and Parkinson’s disease [5]. These high-risk individuals are significantly susceptible to obtaining severe injuries from fall occurrence, including severe tissue damage and broken bones [6]. Sustained injuries are of high-importance as fall related injuries are a leading cause of death among elderly individuals [7, 8], while individuals who suffer from epilepsy are highly susceptible to significant brain injury during epileptic seizures [9]. Furthermore, these individuals may live alone as is frequent with elderly individuals and could become unable to call for help or seek medical aid due to sustained injuries during fall occurrence. Additionally, posttraumatic stress has been linked to individuals after fall occurrence [10], where adults over the age of 65 suffer additional psychological depression and stress, significantly contributing to a lower quality of life [11].

A fall detection system can be utilized to detect fall occurrence, its intensity, direction of impact and quickly raise the alarm for medical treatment if required. Additionally, an automatic fall detection system could lead to reduced fall related psychological stress and less severe head injuries occurring during epileptic seizures as the individual will receive quick medical aid appropriate to the type of fall detected.

Accelerometers can be used to automatically detect falls through different signal evaluation techniques [12]. Wearable biomedical tri-axial accelerometer-based fall detection devices transmit monitored data to a base station receiver.
for signal processing and fall detection. Accelerometer-based fall detection systems frequently detect fall occurrences through thresholding a parameter, such as absolute acceleration magnitude \[13\] or wavelet acceleration sum-vector \[14\] against an arbitrary value. This arbitrary threshold value within literature is frequently determined from analysis of preknown user-specific fall signals \[15\], where accuracy and false positives are significant issues \[16\]. Intelligent fall detection systems utilising classifiers have more recently been utilised for accelerometer-based fall detection \[17\]. Accelerometer features including the raw acceleration signal \[18\] and acceleration signal characteristics \[19\] have been evaluated for fall acceleration signal classification. Various intelligent machine learning methods have been investigated for accelerometer-based fall detection applications, including k-nearest neighbours \[20\], artificial neural networks \[21\] classifiers and more recently combined classifier systems \[22\].

Small, low-cost and wearable biomedical embedded computer systems and devices such as the wireless Shimmer device \[23\] with tri-axial accelerometers, electrocardiography and gyroscope sensors are highly suited for various wearable biomedical and health implementations, such as fall evaluation. Shimmer devices have previously been utilised to detect falls with a reconfigurable high-throughput fall detection system \[24, 14\]. However, wearable accelerometer-based fall detection systems have been reported to obtain low-accuracy results \[12\] in addition to platform power constraints. The Shimmer biomedical device processor, sensors and Bluetooth radio require 5 mA, 4.89 mA and 45 mA current consumption respectively during active use, where the processor has a maximum possible current consumption of 6 mA. Furthermore, the consumption of total current utilised during active device operation demonstrates the processor and sensors to account for 9.11% and 8.91% of current consumption, while Bluetooth radio utilises a vastly significant 81.98% of total current consumption. Therefore it is critical, efficient data transmission occurs to improve power efficiency. Compressive sensing techniques \[25\] can be utilised to significantly reduce acceleration data transmission from the Shimmer device. Furthermore, compressive sensing has been utilised within accelerometer-based
fall detection methods to reduce the required acceleration data \[26\] in addition to significantly improving the Shimmer devices battery life through decreasing data transmission power requirements \[27\]. However, compressive sensing has only been utilised and evaluated in threshold-based systems, where currently literature indicates there has been no evaluation of compressive sensing applied to accelerometer-based features for intelligent fall detection systems.

Compressive sensing is utilised to obtain a solution from an underdetermined linear system through the advantage of signal sparsity \[25\]. Prior knowledge that the data are compressible allows for data to be acquired with fewer measurement samples than required with standard Nyquist-Shannon measurement sampling \[28\] through measuring a sparse signal of minimum samples that contain significant signal information. The resulting reduced number of samples required for signal representation is proportional to the desired compression rate and therefore obtains significant reduction in the number of measurement samples required. The corresponding recovery process of sparse signals generally involves computationally intensive and greedy algorithms which iteratively build an approximated signal solution through updating signal parameters every cycle iteration during the signal recovery process. Greedy algorithms typically do not provide the optimum solution to signal recovery, however they can outperform typical high-resolution accuracy of least-square solutions \[29\].

Matching pursuit (MP) \[30\] is a greedy algorithm that selects a suitable vector element to remove from a measurement matrix of the received sparse signal, then in each following iteration another suitable element is selected until an approximated signal is recovered. Orthogonal MP (OMP) \[29\] is an improvement of MP with reduced iterations at the cost of increased complexity, where OMP selects an optimal index vector element and updates the subspace columns every iteration cycle. OMP updates the subspace column vectors through orthogonally projecting the measurement vector. Other greedy sparse signal recovery algorithms have been proposed, which are often derived from OMP. MP and OMP select a single index vector column during each iteration cycle, resulting in an \(M\)-sparse vector requiring a minimum of \(M\) iterations for signal approx-
Figure 1: Simplified diagram of the proposed compressive sensing accelerometer-based intelligent fall detection system.

Stagewise OMP (StOMP) \cite{31} improves upon this limitation through being able to select multiple columns during each iteration step. Similarly, regularized OMP (ROMP) \cite{32} is another multiple column selection sparse signal recovery algorithm, however ROMP selects vectors with a similar magnitude value.

The proposed compressive sensing accelerometer-based intelligent fall detection system implemented on a Shimmer biomedical embedded device is presented in Fig. 1, where the system’s intelligent component is derived from analysis of a fall signal database to generate a fall detection and analysis classifier. The proposed system captures tri-axial acceleration signals from the wearable Shimmer device sensors. The acceleration signals are then transformed to the wavelet domain and made sparse within the Shimmer embedded computing device, before being transmitted for fall signal analysis. The received sparse signal is recovered through compressive sensing techniques to obtain the wavelet acceleration signal. The recovered wavelet signal is then classified, based on previous fall and activities of daily living (ADL) signals to determine fall occurrence, strength and direction. The presented article is a Matlab-based investigation, where the fall and ADL data obtained with the Shimmer device are recorded in Matlab. The wavelet transform and sparsification stages proposed for Shimmer device implementation, in addition to the base stations sparse signal recovery and fall analysis stages are modelled and investigated within Matlab.

The contributions of the work presented in this paper describes an intelligent, efficient and robust fall detection and diagnostic system utilising compressive
sensing with the wearable Shimmer accelerometer device. The presented fall
detection system implements previously evaluated ADL and fall acceleration
signals to produce robust classifiers for determining fall occurrence, strength-
type (hard and soft classes) and direction. Additionally, the intelligent fall de-
tection system is evaluated against threshold-based fall detection to determine
the proposed systems advantages. Furthermore, the presented fall detection
system utilises compressive sensing to significantly improve data transmission
efficiency within the Shimmer accelerometer device. Literature indicates com-
pressive sensing with accelerometer-based fall detection has currently only been
investigated within threshold-based techniques [26, 27]. The effect of various
compressive sensing techniques have been investigated to determine the effect
on the intelligent classifier-based fall detection system results and performance.

The importance of fall monitoring and the user-wearable Shimmer device
for accelerometer-based fall detection have been discussed. Additionally, the
proposed system for efficient fall detection and analysis with the importance of
compressive sensing for transmission efficiency have been introduced and briefly
discussed in Section 1 while the rest of the paper is organised as follows: Section
2 describes the experimental set-up and acceleration data collection, Section 3
details standard magnitude threshold fall detection and principal component
analysis (PCA) application for feature selection and analysis together with de-
cision tree-based fall detection and analysis. Section 4 describes the sparse
signal generation and compressive sensing of the presented fall detection sys-
tem. Experimental results are presented in Section 5 while Section 6 discusses
the robust and efficient fall analysis system realisation and obtained results.

2. Methodology

Fall and ADL acceleration data were obtained from an experimental test
group of $n$ healthy subjects with differing weight and height. The test group
undertook fall-based activities onto a protective region with a padded mat.
Additionally, the test group undertook real-world-based ADL tasks. All subjects
performed the required tasks, where no injuries occurred. The experimental test subject group consisted of 14 individuals: 2 female and 12 male. The subject test group data for age, weight and height with mean-average and standard deviation are demonstrated in Table 1.

The Shimmer biomedical device was fastened to the subjects chest in the same position across all test subjects and throughout the various experimental fall and ADL activities performed. The fall activities performed with test subjects consisted of hard and soft impacts. Fall impact type examples are presented in Fig. 2. Hard impact falls consist of subjects falling from standing position, demonstrated in Fig. 2a, where subjects would not attempt to break their fall and impact directly with the ground as demonstrated in Fig. 2b. Similarly, soft impact falls consist of subjects falling from standing position, where subjects would attempt to break their fall before impact, such as falling onto their knees, as demonstrated in Fig. 2c, before impacting the ground as shown in Fig. 2d. Furthermore, subjects carried out impact type falls in multiple directions; left, forward, right and backwards.

Additionally, ADL tasks were undertaken with test subjects, including walking, sitting on a chair from standing position, standing up from chair, jumping and running. The ADL and fall databases are presented in Tables 2 and 3 detailing the activity type and the number of events obtained. Each subject performed every ADL type activity at least once to produce a minimum of 20 events within the data base. Walking ADL activities are a significantly frequent activity, hence double the amount of samples were recorded to obtain 40 monitored events. Furthermore, every subject produced at least one soft and hard

<table>
<thead>
<tr>
<th>Table 1: Subject test group data (n=14).</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>27</td>
<td>8</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>176</td>
<td>16</td>
</tr>
</tbody>
</table>
Figure 2: Fall type occurrences; (a) hard impact type fall with (b) straight fall to ground and (c) soft impact fall type with preliminary fall to subjects knee before (d) falling to ground.
impact falls in every direction to obtain 40 recorded fall events for left, forward, right and backward directions. Every fall direction-based occurrence produced 50% soft and hard impact fall types, producing 80 soft and 80 hard fall impact events. In summary, 14 test subjects obtained a database of 120 ADL and 160 fall recorded occurrences.

The obtained acceleration data from the Shimmer device was wirelessly recorded and analysed in Matlab. The Shimmer acceleration data was then processed and evaluated to model compressive sensing on the Shimmer platform and evaluate fall analysis on the received acceleration data features with the proposed fall detection system demonstrated in Fig. 3.

The proposed compressive sensing accelerometer-based intelligent fall detection system records acceleration signals in 3 dimensions (3D) on the Shimmer
device. These signals are joined together to form a tri-axial acceleration signal which is then transformed to the wavelet domain. The wavelet signal is thresholded to maintain the largest significant coefficient values while zeroing the other coefficients to obtain a sparse signal. The signals sparsity is improved through applying a preknown random permutation index function to ensure samples are randomly located throughout the vector. The sparse wavelet signal is then projected onto a random sensing matrix before being transmitted from the Shimmer device to the base station receiver for signal analysis. The received signal is reconstructed through various compressive sensing algorithms, MP, OMP, ROMP and StOMP. The reconstructed signal is then restructured in accordance with the preknown sparsity permutation function before being classified with the PCA-based decision tree to determine fall occurrence, fall direction and fall strength. The PCA-based decision tree classifier is generated
from analysis of fall and ADL signals evaluated in the wavelet domain.

3. Fall Signal Analysis

3.1. Threshold-Based Fall Detection

Current standard approaches to accelerometer-based fall detection systems utilise a wavelet-based threshold [33, 24, 26]. The discrete wavelet transform (DWT) for a Daubechies 2 level-3 mother-wavelet is applied to the time domain tri-axial x, y and z-axes acceleration components to determine the wavelet equivalent acceleration components $\alpha_x$, $\alpha_y$ and $\alpha_z$. The wavelet-domains module of acceleration vectors determine the wavelet acceleration sum vector $SV$ as follows:

$$SV = \sqrt{(\alpha_x)^2 + (\alpha_y)^2 + (\alpha_z)^2} > \tau$$  \hspace{1cm} (1)

The threshold parameter $\tau$ is applied to the wavelet domains sum vector, where values exceeding the threshold parameter determine fall occurrence. The threshold parameter is typically selected in the literature [12] by an arbitrary value based on evaluation of fall data, however an optimum threshold $O\tau$ method has been presented [13], as follows:

$$O\tau = \frac{1}{n} \sum_{i=1}^{n} \rho_i$$  \hspace{1cm} (2)

The optimum threshold is obtained through evaluating the average minimum fall acceleration values for each subject $\rho$ within a group of $n$ subjects, where each subjects minimum fall acceleration value is represented as $\rho_1, \rho_2...\rho_n$.

3.2. PCA-Based Smart Fall Detection Classifiers

The presented smart fall detection system evaluates the fall and ADL database of wavelet feature coefficients feeding a binary decision tree after PCA extraction. Initial signal features are calculated using the Daubechies 2 level-3 DWT to obtain $\alpha_x$, $\alpha_y$ and $\alpha_z$ wavelet versions of the acceleration signals, $\hat{\alpha}_x$, $\hat{\alpha}_y$ and $\hat{\alpha}_z$. ADL and fall samples were obtained from 2-second windows in the time domain, yielding an overall number of 101 three dimensional vectors ($f_s = 50Hz$),
which were reorganized to vector length 303 prior to PCA calculation so that wavelet acceleration vectors \( \hat{\alpha} = [\hat{\alpha}_z, \hat{\alpha}_y, \hat{\alpha}_z] \) of length 90 entered the PCA stage.

PCA is an unsupervised linear transformation method frequently utilized for data classification and reduced dimensionality. PCA utilizes statistical orthogonal functions to reduce signal dimensionality of the large observed data variables to obtain smaller significant data feature variables \[33\]. Let the training data set of wavelet acceleration signals \( \hat{\alpha}_\omega = [\hat{\alpha}_1, \hat{\alpha}_2, ... \hat{\alpha}_N] \), then the training set average is defined as \( \bar{\rho} = \frac{1}{N} \sum_{i=1}^{N} \hat{\alpha}_\omega_i \). The wavelet acceleration signals differ from the mean-based average through \( \Delta \hat{\alpha}_\omega = \hat{\alpha}_\omega_i - \bar{\rho} \). The set of wavelet acceleration signal mean-based differences \( \Delta \hat{\alpha}_\omega \) are evaluated to obtain the PCA covariance matrix \( CM = \frac{1}{N} \sum_{i=1}^{N} \Delta \hat{\alpha}_\omega_i \Delta^\top \hat{\alpha}_\omega_i \), where the orthogonal counterpart of the wavelet acceleration signal mean-based difference is denoted as \( \Delta^\top \hat{\alpha}_\omega \). The covariance matrix is further evaluated with eigendecomposition to determine the associated eigenvalues \( \lambda_k \) and eigenvectors \( u_k \). The eigenvalue data vector is arranged in descending order, where \( M \) principle components are selected \( \lambda_1, \lambda_2, ... \lambda_M \). The PCA implementation utilizes \( M = 45 \) principle components to determine the feature subspace dimensionality through obtaining a projection matrix \( W \). Unsupervised linear transformation with PCA is determined with \( \hat{\alpha}_\omega^{PCA} = W^\top \hat{\alpha}_\omega \). The reduced signal dimensionality output \( \hat{\alpha}_\omega^{PCA} \) of the PCA module was classified using binary decision trees \[35\] trimmed from 45 predictors to 26 for detecting falls and determining their strength and direction. The decision tree functions were thus trained using PCA coefficients \( \hat{\alpha}_\omega^{PCA} \) for fall detection, strength and direction.

4. Compressive Sensing

Compressive sensing allows an input signal in high-dimensional space to be compressed to a signal of significantly smaller dimensions, such as the wavelet-domain \[25, 36\]. The acceleration signal in the time-domain \( \Gamma \) of \( N \) samples is decomposed with an orthonormal Daubechies 2 wavelet basis \( \Psi = [\psi_1 | \psi_2 | ... | \psi_N] \) to obtain a compressive sensing suitable sparse wavelet-domain coefficient signal.
$X$ as shown:

$$X = \Psi \Gamma$$

(3)

The presented sparse wavelet signal undergoes a sparsity function $\mu$ to improve sparsity properties with thresholding and preknown indexing. The compressive sensing function can be described as follows:

$$Y = \Phi X = \Phi \mu \Psi \Gamma$$

(4)

Where the input signal vector $X$ of length $N$ is $K$-sparse, $Y$ is the observation vector of length $M$ and the Gaussian sensing matrix $\Phi$ is of size $M \times N$. Compressive sensing is significantly suited for $K$-sparse signals, where signal vector $X$ of length $N$ is represented by significant $K$ nonzero coefficients [37]. The input signal $X$ is sparse when only a low number of coefficients are nonzero. Sensing matrices are frequently implemented as random matrix functions with independent distribution such as Gaussian or Bernoulli distributions [38], which have previously been demonstrated to provide efficient signal probability recovery [39]. Additionally, compressive sensing captures $M$ signal measurements from $N$ signal samples through random linear projections [39, 25], where $M << N$, offering advantages for significantly improved data efficiency suitable for power-limited transmission platforms.

The Shimmer platforms tri-axial acceleration signals are merged to a single vector of length $N = 303$ demonstrated in Fig. 4a, a Daubechies 2 level-3 DWT is applied, obtaining a wavelet feature and detail signal $N = 90$ as shown in Fig. 4b, suitable for fall analysis with the presented intelligent PCA-based fall detection classifiers. The wavelet signals sparse properties are further improved through applying the sparsity function $\mu$. Where the wavelet signal is thresholded to obtain a wavelet signal vector $N = 90$, retaining $K = 45$ largest nonzero value coefficients, before applying a preknown random index permutation function as shown in Fig. 4c. Once the wavelet acceleration signal has been sparsed, the compression described in (4) is carried out, where a Gaussian distribution is implemented as a sensing matrix. The compressed sparse wavelet acceleration signal $Y$ is then transmitted. The greedy sparse
signal recovery algorithms MP, OMP, ROMP and StOMP previously described are applied to recover the estimated thresholded wavelet acceleration signal $\hat{X}$. The known index permutation function is applied and the wavelet coefficients $N = 90$ are extracted as demonstrated in Fig. 4d and applied to the PCA classifiers described previously in Section 3.2. The presented technique exploits the received compressed wavelet acceleration data to perform PCA-based fall detection classification without performing the inverse DWT.

MP is a technique frequently utilised for sparse signal reconstruction through a greedy iterative process to compute an approximation of the original signal.

Figure 4: Fall example of (a) acceleration, (b) wavelet-domain signal, (c) thresholded with random permutation and (d) OMP estimate.
as demonstrated in Algorithm 1. The MP algorithm iteratively determines the single most suitable vector element to remove from the measurement matrix $\Phi$ and repeats the subtraction process from the residual component $r$ during each iteration until the estimated original signal $\hat{X}$ is determined. It is important to note that MP will repeatedly select the same vector element to remove from the sensing matrix in order to improve the estimated signal. MP is a relatively low-complexity solution suited to high-sparsity signals [30]. Similarly, OMP determines an estimated signal through an iterative process shown in Algorithm 2. In OMP, the approximated signal $\hat{X}$ is updated every iteration with projecting $Y$ orthogonally onto columns of the sensing matrix $\Phi$ for the current index set $\Lambda$. OMP never selects a previously evaluated vector element from the sensing matrix and the residual component during any iteration is orthogonal to currently selected vector elements. OMP is a more complex technique in comparison to MP, as least squares is utilised to estimate the signal during every iteration, however it reduces iterations required for signal estimation. MP and OMP methods need to perform a minimum iteration count determined by the number of elements to be selected, which does not scale well for large dimension signals. ROMP and StOMP algorithms shown in Algorithms 3 and 4 are multiple-element selection-based techniques derived from OMP and utilise an updated residual component every iteration until signal estimation is met. ROMP groups inner elements $\lambda$ into sets $\Lambda$, where elements in each set contain a similar magnitude [40]. StOMP groups inner elements $\lambda$ above a threshold parameter $\tau$ into sets $\Lambda$. The process is then repeated on the residual vector until the estimated signal is obtained. StOMP can generally obtain good approximations through few iterations through evaluating multiple vectors [41]. For direct comparison all reconstruction algorithms were configured to operate for $K$ or less iterations.
Algorithm 1 MP algorithm for signal recovery.

**Input:** \( \Phi \): Sensing matrix; \( Y \): Observation vector; \( K \): Sparsity of signal \( X \).

**Output:** \( \hat{X} \): Original signal estimation.

**Procedure:**

1. Initialise residual \( r_0 = Y \) and set iteration counter \( i = 1 \).
2. Find index \( \lambda_i = \arg \max \| \langle r_{i-1}, \emptyset_j \rangle \| \) for \( j = 1 \ldots N \).
3. Calculate new approximation \( \hat{X}_i = \hat{X}_{i-1} + \langle \Phi_{\lambda_i}, r_{i-1} \rangle \Phi_{\lambda_i} \) and residual \( r_i = Y - \hat{X}_{i-1} = r_i - \langle \Phi_{\lambda_i}, r_{i-1} \rangle \Phi_{\lambda_i} \).
4. Increment iteration counter \( i = i + 1 \), return to step 2 if \( i < K \).
5. Output estimate \( \hat{X} \).

Algorithm 2 OMP algorithm for signal recovery.

**Input:** \( \Phi \): Sensing matrix; \( Y \): Observation vector; \( K \): Sparsity of signal \( X \).

**Output:** \( \hat{X} \): Original signal estimation.

**Procedure:**

1. Initialise residual \( r_0 = Y \) and index set \( \Lambda_0 = \{ \emptyset \} \). Iteration counter \( i = 1 \).
2. Find index \( \lambda_i = \arg \max \| \langle r_{i-1}, \emptyset_j \rangle \| \) for \( j = 1 \ldots N \).
3. Increase index set \( \Lambda_i = \Lambda_{i-1} \cup \{ \lambda_i \} \) and matrix column vectors \( \Phi_i = [\Phi_{i-1} \Phi_{\lambda_i}] \).
4. Solve least squares problem for estimate \( \hat{X} \) of signal \( X \); \( \hat{X}_i = \arg \min \| Y - \Phi_i \hat{X}_i \| \).
5. Calculate new residual \( r_i = Y - \Phi_i \hat{X}_i \).
6. Increment iteration counter \( i = i + 1 \), return to step 2 if \( i < K \).
7. Output estimate \( \hat{X} \).

5. Results

Recall \( \phi \), precision \( \varphi \) and specificity \( \delta \) metrics as shown in (5) through (7) can be utilised to evaluate fall detection, direction and strength results obtained with the proposed system.

\[
\phi = \frac{T_p}{T_p + F_n}
\]  

(5)
Algorithm 3 ROMP algorithm for signal recovery.

**Input:** \(\Phi\): Sensing matrix; \(Y\): Observation vector; \(K\): Sparsity of signal \(X\).

**Output:** \(\hat{X}\): Original signal estimation.

**Procedure:**
1. Initialise residual \(r_0 = Y\) and index set \(\Lambda_0 = \{\emptyset\}\). Iteration counter \(i = 1\).
2. Identify set \(\lambda_i\) of \(K\) largest coordinates in observation vector \(u_i = \Phi_i r_i\).
3. Regularise subsets \(\lambda_0 \subset \lambda_i\) with comparable coordinates \(|u_i| \leq 2|u_{\lambda_i}|\) for \(i, \lambda \in \lambda_0\).
4. Update index set \(\Lambda_i = \Lambda_{i-1} \cup \lambda_i\).
5. Calculate new approximation \(\hat{X}_i = \arg \min \|Y - \Phi_i \hat{X}_i\|\) and residual \(r_i = Y - \Phi_i \hat{X}_i\).
6. Increment iteration counter \(i = i + 1\), return to step 2 if \(i < K\).
7. Output estimate \(\hat{X}\).

Algorithm 4 StOMP algorithm for signal recovery.

**Input:** \(\Phi\): Sensing matrix; \(Y\): Observation vector; \(K\): Sparsity of signal \(X\).

**Output:** \(\hat{X}\): Original signal estimation.

**Procedure:**
1. Initialise residual \(r_0 = Y\) and index set \(\Lambda_0 = \{\emptyset\}\). Iteration counter \(i = 1\).
2. Obtain residual correlation vector \(c_s\) with matched filter, \(c_s = \Phi^T p_{i-1}\).
3. Threshold correlation residual with \(\tau\) to create index \(\lambda_i\).
4. Update index set \(\Lambda_i = \Lambda_{i-1} \cup \lambda_i\).
5. Calculate new approximation \(\hat{X}_i = \arg \min \|Y - \Phi_i \hat{X}_i\|\) and residual \(r_i = Y - \Phi_i \hat{X}_i\).
6. Increment iteration counter \(i = i + 1\), return to step 2 if \(i < K\).
7. Output estimate \(\hat{X}\).

\[
\varphi = \frac{T_p}{T_p + F_p} \quad (6)
\]

\[
\delta = \frac{T_n}{T_n + F_p} \quad (7)
\]
Recall, precision and specificity are obtained from true and false positive and negative rates: $T_p$, $F_p$, $T_n$, $F_n$. The confusion matrix defining these rates is presented in Table 4 for events and non-events detected by the system:

Table 4: Confusion matrix defining $T_p$, $F_p$, $T_n$, $F_n$ for events and non-events detected by the system

<table>
<thead>
<tr>
<th></th>
<th>Predicted Event</th>
<th>Predicted Non-event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Event</td>
<td>$T_p$</td>
<td>$F_p$</td>
</tr>
<tr>
<td>Actual Non-event</td>
<td>$F_n$</td>
<td>$T_n$</td>
</tr>
</tbody>
</table>

The events and non-events are defined depending on the measured variable:

- Fall detection: Falls are defined as events and ADLs as non-events.
- Fall strength: For hard falls, a hard fall is defined as an event. ADLs and soft falls are counted among non-events. For soft falls, the procedure is inverted, being hard falls part of the non-events, and soft ones the actual events. The overall metrics for fall strength are presented as an average of both hard and soft falls.
- Fall direction: For each of the directions considered (forward, backward, left, and right directional falls), falls in that direction are the actual events, while falls in the remaining directions and ADLs are non-events. The overall metrics are computed averaging all the directions.

The system’s recall value accounts for the actual sensitivity to events, while precision is a measurement of the predictive positive value for the events. Finally, specificity measures the robustness of the system against false positives (i.e., identifying non-events as events). An overall accuracy measure for the system can be obtained from the F-value $f$ metric, calculated as the harmonic mean of recall and precision:

$$f = 2 \frac{\phi \varphi}{\phi + \varphi}. \quad (8)$$
In this article, PCA-based decision tree classifiers are generated from a training group and evaluated against a control group in accordance with k-fold cross-validation techniques [42] for 10-fold and 5-fold methods. The 10-fold cross-validation technique separates the signal database into 10 equal sized groups, where the PCA classifier is trained utilizing 90% of acceleration signal data, while the remaining 10% is utilized for testing. Similarly, 5-fold cross-validation separates the database into 5 equal sized groups, where 80% and 20% of data is utilized for training and testing of the PCA classifier. Furthermore, the process is repeated throughout the database groups, where each instance of testing is validated against the remaining groups. This obtains a robust testing validation as the database is evaluated with different combinations for testing and training samples, where no sample appears simultaneously within training and testing groups. The obtained results for 10-fold and 5-fold cross-validation methods were recorded, where an average for each method is presented. Threshold-based fall detection with the optimum threshold parameter [14] and PCA classifier-based fall detection, strength and direction results for recall, precision, specificity and F-value without compressive sensing for 10-fold and 5-fold cross-validation methods are presented in Tables 5 and 6.

<table>
<thead>
<tr>
<th>Fall Analysis</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>Specificity (%)</th>
<th>F-Value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold-Based Fall Detection</td>
<td>87</td>
<td>93</td>
<td>81</td>
<td>90</td>
</tr>
<tr>
<td>Classifier-Based Fall Detection</td>
<td>96</td>
<td>99</td>
<td>98</td>
<td>97</td>
</tr>
<tr>
<td>Classifier-Based Fall Strength</td>
<td>91</td>
<td>99</td>
<td>96</td>
<td>95</td>
</tr>
<tr>
<td>Classifier-Based Fall Direction</td>
<td>87</td>
<td>98</td>
<td>84</td>
<td>93</td>
</tr>
</tbody>
</table>

The optimum threshold method demonstrated accuracy of 87% recall, 93% precision, 81% specificity and 90% F-value, while the presented PCA classifier method obtained significant accuracy improvements of 9%, 6%, 17% and 7% for recall, precision, specificity and F-Value respectively, obtaining 96% recall, 99% precision, 98% specificity and 97% F-value for 10-fold cross-validation.
Additionally, the presented smart PCA classifier demonstrated 91% recall, 99% precision and 95% F-value for fall strength accuracy and 87%, 98% and 93% F-value for fall direction with 10-fold cross-validation. The optimum threshold method demonstrated consistent performance decreases of approximately 10% as testing and training data increased and decreased simultaneously in accordance with 5-fold cross-validation. In contrast, PCA-based classifier techniques demonstrated a range of performance decreases from 4% to 10% in comparison with 10-fold cross-validation obtained results. Hence, the PCA-based classifier method is demonstrated to be more robust to reduced training data with k-fold cross validation techniques than the threshold fall detection method. The rest of this article will present, evaluate and describe the 10-fold cross-validation implementation.

Multiple compressive sensing techniques were applied to the proposed system as demonstrated in Fig. 3. The sparse received signal with $K = 45$ nonzero coefficient elements over vector length $N = 90$ is reconstructed with MP, OMP, ROMP and StOMP compressive sensing techniques. The reconstructed signals are evaluated against signal sparsity as defined in (9), where the number of significant nonzero coefficients $K$ are reduced in the transmitted signal vector length $N$ to increase sparsity.

\[
\text{Sparsity} \left( \% \right) = \left( 1 - \frac{K}{N} \right) \times 100 \quad (9)
\]

The percentage of exact coefficient recovery for MP, OMP, ROMP and StOMP compressive sensing techniques evaluated against sparsity increase are presented in Fig. 5.

### Table 6: Fall detection results with 5-fold cross validation

<table>
<thead>
<tr>
<th>Fall Analysis</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>Specificity (%)</th>
<th>F-Value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold-Based Fall Detection</td>
<td>76</td>
<td>81</td>
<td>71</td>
<td>78</td>
</tr>
<tr>
<td>Classifier-Based Fall Detection</td>
<td>86</td>
<td>90</td>
<td>89</td>
<td>88</td>
</tr>
<tr>
<td>Classifier-Based Fall Strength</td>
<td>82</td>
<td>91</td>
<td>90</td>
<td>86</td>
</tr>
<tr>
<td>Classifier-Based Fall Direction</td>
<td>83</td>
<td>90</td>
<td>79</td>
<td>86</td>
</tr>
</tbody>
</table>
The reconstructed signal initially recovers 100% exact coefficients until sparsity approaches 43%, 46%, 48% and 50% sparsity for MP, ROMP, OMP and StOMP respectively, where the percentage of exact coefficient recovery begins to decrease. The percentage of exact coefficient recovery continues to decrease as sparsity increases until 0% exact coefficients recovery occurs at 54%, 59%, 64% and 65% sparsity for MP, ROMP, OMP and StOMP respectively.

Evaluating MP, OMP, ROMP and StOMP recovered signals with the presented PCA-based decision tree classifier in terms of recall, precision, specificity and F-value for fall detection, strength and direction as sparsity increases have been presented in Fig. 6, 7 and 8 respectively. Where sparsity increase was obtained from reducing the number of largest significant coefficient values $K$ of the transmitted sparse wavelet signal.

All sparse signal recovery algorithms were initially able to obtain the same recall, precision and F-value accuracy for fall detection, strength and direction as presented in Table 5. However the general response of recall, precision, specificity and F-value results decrease as sparsity increases from 50%, until 0% accuracy occurs at 68% sparsity. Where there would frequently be a sudden significant decrease occurring at approximately 67% sparsity towards 0% accuracy. The initial decrease of recall, precision, specificity and F-value accuracy for fall detection, strength and direction occurred at approximately the same sparsity...
percentage for each compressive sensing technique. MP demonstrated accuracy decrease occurs at 50%, 53%, 50% and 50% sparsity for recall, precision, specificity and F-value parameters respectively. OMP demonstrated accuracy decrease occurs at 55%, 64%, 56% and 55% sparsity for recall, precision, specificity and F-value parameters respectively. ROMP demonstrated accuracy decrease occurs at 53%, 56%, 53% and 53% sparsity for recall, precision, specificity and F-value parameters respectively. StOMP demonstrated accuracy decrease occurs at 56%, 65%, 57% and 56% sparsity for recall, precision, specificity and F-value parameters respectively. However the responses after the initial decrease for each compressive sensing technique are significantly different across
fall analysis. For example, fall detection recall with MP and ROMP obtain 14% and 29% accuracy at 67% sparsity as shown in Fig. 6(a). While in comparison MP and ROMP obtain significantly different values of 44% and 52% accuracy for fall strength recall and 37% and 50% accuracy for fall direction with 67% sparsity as demonstrated in Fig. 7(a) and 8(a).

The various compressive sensing techniques performance can be ranked from highest to lowest obtained accuracy during sparsity increase, where StOMP demonstrated the best results, followed by OMP then ROMP and the worst performance was obtained with MP. StOMP obtained the highest recall, precision, specificity and F-value metrics throughout, followed by OMP demonstrating lit-
tle significant difference between response, where precision results obtained were similar, for example as demonstrated with fall detection precision as shown in Fig. 6b. ROMP and MP demonstrated significantly worse results than StOMP and OMP, where MP and ROMP obtained quick and high-accuracy loss as sparsity increased. Recall, specificity and F-test accuracy demonstrated a large accuracy decrease as sparsity increased, however the classifiers precision maintained relatively high until sparsity reached approximately 65%, e.g. Fig. 8b.

Additionally, the PCA-based classifier was able to obtain high-accuracy results while the equivalent exact coefficient recovery percentage could be relatively low. For example, StOMP obtained 70%, 96%, 82% and 80% accuracy
for fall detection recall, precision, specificity and F-value responses respectively at 65% sparsity as demonstrated in Fig. 6. While in comparison, the exact coefficient recovery was 0% at 65% sparsity as shown in Fig. 5 indicating the obtained signal error at 65% sparsity is within the PCA-based classifiers boundaries for correct fall analysis. Furthermore, the significant drop of recall, precision, specificity and F-value at 67% sparsity could be deduced to occur when signal error increases to outside the classifiers boundary limits for correct fall analysis. Recall, specificity and F-test accuracy generally demonstrated a large accuracy decrease as sparsity increased, however the classifiers precision metric maintained relatively high accuracy until sparsity reached approximately 65%. MP and ROMP techniques obtained the lowest performance results in comparison to OMP and StOMP methods which performed well in the presented fall evaluation application. Potential hypothesis for the different performance results include: MP has a convergence problem, where it can potentially reselect a previously selected vector removed from the signal. The convergence problem of MP has demonstrated much larger error rates in comparison to OMP [43], where OMP removes the potential vector reselection issue through updating its dictionary during each iteration. Furthermore OMP utilizes least squares to estimate signal approximation during every iteration further increasing the possibilities for improved performance. ROMP has been recorded in literature for demonstrating poor performance in application, where ROMP implementations are best suited to very large $N$ column data sets [41]. Possibly indicating the fall detection application presented does not contain data large enough over $N$ columns for high ROMP performance. StOMP obtained relatively high performance results, demonstrating a small performance increase over OMP. This may be due to the fall application utilizing Gaussian sensing matrices and StOMP being limited to theory on Gaussian matrices, where theoretical performance is guaranteed when applied to general Gaussian matrices [41]. Furthermore, the proposed system offers significant space savings defined with the compression ratio $\omega$ expressed in (10), evaluating the number of original data samples $N_{orig}$ and the number of data samples in the compressive sensing sparse signal $N_{comp}$. 

25
\[ \omega(\%) = \left( \frac{N_{\text{orig}} - N_{\text{comp}}}{N_{\text{orig}}} \right) \times 100 \quad (10) \]

The original Shimmer device acceleration transmitted signal vector of length \( N_{\text{orig}} = 303 \), the compressed signal obtained a sparse signal vector of length \( N_{\text{comp}} = 90 \) with 45 nonzero elements. The compressed signal vector \( N_{\text{comp}} = 90 \) demonstrated a significant compression ratio of 70.3\%, while the compression ratio in context of significant nonzero coefficients \( N_{\text{comp}} = 45 \) obtained a significant compression ratio of 85.1\% in comparison to the original acceleration signal.

The presented PCA-based decision tree classifier method significantly outperformed the threshold-based fall detection method frequently utilised within literature \[33, 24, 26\] for accelerometer-based fall detection systems. The PCA-based classifier obtained high-accuracy results with recall, specificity, precision and F-value for fall detection, strength and direction, where the classifier demonstrated significantly higher precision results. The sparsity of the proposed systems transmitted signal was increased through reducing the number of significant coefficients retained. StOMP overall performed better during evaluation than MP, OMP and ROMP, where StOMP and OMP obtained relatively similar results with fall detection, strength and direction classifiers. The classifiers accuracy failed as the transmitted signal vector reached 68\% sparsity with all compressive sensing techniques. Furthermore, the PCA-based decision tree classifier offers significant data reduction with a compression ratio over 70\% in comparison to standard acceleration signal transmission within the Shimmer device.

The obtained transmission data reduction has a direct effect on significantly improving transmission efficiency with the Shimmer device.

6. Conclusions

This article presents a compressive sensing acceleration-based intelligent fall detection and analysis system with a Shimmer embedded computing device modelled with Matlab. The proposed system utilises fall and ADL acceleration
signals within the wavelet-domain to obtain PCA-based binary decision tree fall
classifiers for fall detection, strength-type and direction. Test subjects generated
fall and ADL acceleration signals with the Shimmer device. The captured tri-
axial acceleration signals were converted to the wavelet-domain and made sparse
through selecting the most significant coefficients for wireless transmission to a
base station. The received sparse signal is then recovered through greedy com-
pressive sensing techniques: MP, OMP, ROMP and StOMP. The reconstructed
signal is then evaluated with the PCA-based decision tree classifier to deter-
mine fall detection, strength and direction. The presented system obtained
highly-accurate fall detection, strength and direction results for recall, preci-
sion, specificity and F-value accuracy. Acceleration-based fall detection systems
have reported low-accuracy response with threshold-based detection methods
[12]. The optimum threshold method presented demonstrated 87%, 93%, 81%
and 90% for fall detection recall, precision, specificity and F-value measures.
The presented PCA-based decision tree technique obtained 96%, 99%, 98% and
95% for fall detection recall, precision, specificity and F-value measures respec-
tively, demonstrating significant improvements in comparison with the optimum
threshold approach. Additionally, threshold approaches are significantly limited
as they are only capable of detecting fall occurrence. While in comparison the
presented PCA-based approach can determine fall occurrence in addition to im-
portant fall information including whether the fall was a hard or soft impact, in
addition to the direction the individual fell. Acceleration data transmission with
the Shimmer device accounts for over 80% of total current consumption during
standard use, indicating the importance of efficient data transmission on the
limited power Shimmer platform. Various compressive sensing techniques were
utilised to significantly reduce acceleration information required for transmis-
sion with the Shimmer device. StOMP obtained the best signal performance as
signal sparsity increased, where 0% exact coefficient recovery occurred at 54%,
59%, 63% and 65% for MP, OMP, ROMP and StOMP techniques. MP, OMP,
ROMP and StOMP were initially able to obtain the same recall, precision and
F-value accuracy for fall detection, strength and direction, however as sparsity
increased the accuracy decreased. Where StOMP obtained the best results for fall detection, strength and direction with OMP obtaining slightly lower accuracy results, while MP and ROMP demonstrated significantly low-accuracy results in comparison. Furthermore, the presented compressive sensing fall detection system obtains significant data reduction of over 70%, demonstrating improved transmission efficiency while maintaining the systems high accuracy results.

References


