30th Annual International IEEE EMBS Conference
Vancouver, British Columbia, Canada, August 20-24, 2008

An Acoustic Fall Detector System that Uses Sound Height Information to Reduce the False Alarm Rate

Mihail Popescu, *Senior Member*, *IEEE*, Yun Li, Marjorie Skubic, *Member*, *IEEE*, Marilyn Rantz

Abstract— **More than one third of about 38 million adults 65 and older fall each year in the United States. To address the above problem we propose to develop an acoustic fa**ll **de**tection **system (FADE) that will automatically signal a fall to the monitoring caregiver. As opposed to many existent fall detection systems that require the monitored person to wear devices such as accelerometers or gyroscopes at all times, our system is completely unobtrusive by not requiring any wearable devices. To reduce the false alarm rate we employ an array of acoustic sensors to obtain sound source height information. The sound is considered a false alarm if it comes from a source located at a height higher than 2 feet. We tested our system in a pilot study that consisted of a set of 23 falls performed by a stunt actor during six sessions of about 15 minutes each (1.3 hours in total). The actor was previously trained by our nursing collaborators to fall like an elderly person. The use of height information reduced the false alarm hourly rate from 32 to 5 at a 100% fall detection rate.**

I. INTRODUCTION

More than one third of about 38 million adults 65 and older fall each year in the United States [1]. In spite of extensive fall prevention programs [1], in 2006 there were about 400,000 fall related hospitalization with an estimated direct cost of about $19 billion [2]. About 30% of people who fall suffer severe injuries such as fractures and head trauma [1] that can render them unable to raise or to ask for help. If the person lives alone in the apartment, a fall might result in a long lie on the floor which can cause hypothermia, dehydration, pressure sores or rhabdomyolosis (destruction of skeletal muscle) [3]. Moreover, the delay in hospitalization can increase the mortality risk in some clinical conditions [4]. For example, one day delay in the hip fracture surgery may increase the 30 day mortality risk from 7.3% to 8.7% [5]. Hence, it is imperative that the falls are detected and the necessary help is provided as soon as possible.

The fall detection methods found in the literature [6] are based on two types of devices: wearable and non-wearable. The wearable devices tend to be easier to deploy while the non-wearable ones tend to be less obtrusive. The wearable devices are in general rejected by older, more frail, people [6]. Among the wearable devices we cite accelerometers, gyroscopes, mercury tilt switches and velocity sensors. Among the non-wearable fall detection devices we mention floor vibration sensors [7], video cameras [8,9], infrared cameras [10], and smart carpets. The floor vibration sensor [7] was proposed to be deployed in a motion sensor network that helped reduce the false alarm rate if motion is sensed in a given time window after the fall signal [11]. The ground-level concrete slabs covered by carpet that are typical for the nursing homes currently built in the United States represent a challenge for the floor vibration sensors. The use of video camera is promising, although the computational requirements for processing present a challenge. The infrared cameras and the smart carpet technologies are still under development. The greatest (common) challenge faced by all the fall detection systems is represented by the false alarms which may lead to its rejection by the user [6].

In this paper we describe a dedicated acoustic **fa**ll **de**tection system (FADE) based on an array of acoustic sensors. The system is inexpensive and built from off-the-shelf components. This is the first part of a more complex fall detection system that will integrate a motion detector for learning new sounds and, thus, further reducing the false alarm rate.

Acoustic sensors have been previously used in habitat monitoring [12]-[16]. In [12]-[14] a set of acoustic sensors was used to differentiate between several sound classes such as breaking glass, screams, steps, door sound and human sound. A microphone was placed in each room of the apartment to indentify the location of the sound. The acoustic sensor used in the ListenIn system [16] was designed for activity monitoring (baby noise or loud noise). The alarm, together with the encrypted sound, was send to a mobile device held by a caregiver. Human falls were not included in the sound classes detected in any of the above acoustic systems.

The structure of the paper is as follows: in section II we present the architecture of the fall detection system, in section III we describe the algorithms used for fall detection, in section IV we describe the test data and methodology, in section V we show the results and in section VI we give the conclusions.

II. SYSTEM ARCHITECTURE

The architecture of the FADE system is shown in Fig. 1.

M. P. is with the Health Management and Informatics Department, University of Missouri, Columbia, MO 65211, USA (corresponding author, phone: 573-882-1266; fax: 573-882-6158; e-mail: popescum@missouri.edu).

Y. L. and M.S. are with the Electrical and Computer Engineering Department, University of Missouri, Columbia, MO 65211, USA.

M. R. is with the Sinclair School of Nursing, University of Missouri, Columbia, MO 65211, USA.

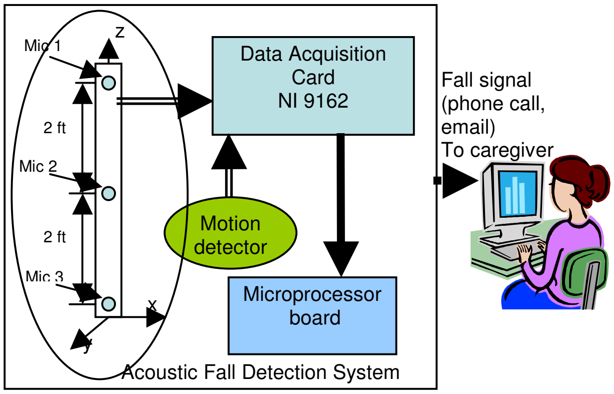


Fig. 1 The proposed fall detector architecture

The fall detector consists of a linear array of electret condenser acoustic sensors (three shown: Mic 1, Mic 2 and Mic 3) mounted on pre-amplifier boards Cana Kit CK495 (about $20 each). More acoustic sensors might be considered in the future to help improve beam forming and source separation. The acoustic sensor array was mounted vertically (along z-axis) in order to be able to capture sound height information. The FADE incorporates a motion detector for further reduction of the false alarms. The working hypothesis for FADE is that the person is alone in the department. If motion is detected during a given interval (one minute) after a fall event is computed as likely, the caregiver alert is not issued. Instead, the event that provoked the alarm is cataloged as a false alarm and stored in the internal memory. Challenges such that the existence of pets in the apartments and the movement of the hurt person after the fall incident will need to be addressed if a motion detector will be used. In order to preserve the privacy of the patient, the sound will be internally processed on a microprocessor board and only an external fall signal (email or pager) will be send to the caregiver.

In this preliminary work we were mainly interested in testing the acoustic sensors and their geometry. For this reason, we only used a reduced version of the system shown in Figure 1. The experimental architecture of the acoustic fall detection system used here is shown in Figure 2.

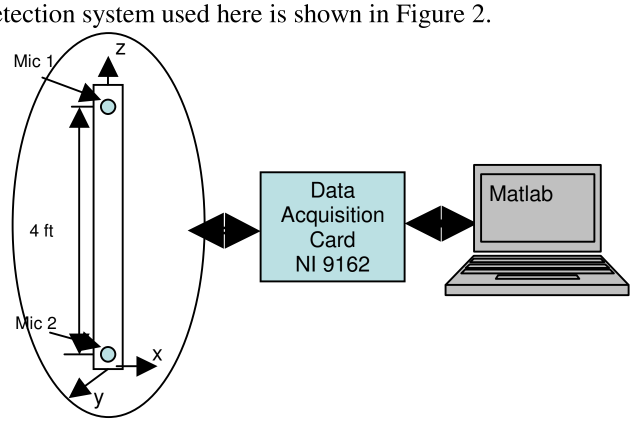


Fig. 2 The fall detector architecture used for testing in this paper.

The sound was recorded on a laptop using a National Instruments data acquisition card NI 9162 with 8 differential analog inputs. The recorded sound (as described in Section IV) was later processed using Matlab (http://www.mathworks.com). In the experiments presented here we used only 2 sound sensors (Mic 1 and Mic 2) mounted vertically 4 feet apart.

III. Algorithm Description

The main steps of the signal processing algorithm (see Figure 3) were: signal preprocessing, pattern recognition and height calculation.

A. *Signal preprocessing*

The first step in preprocessing was noise removal using a Wiener filter [17].

The next step was to compute the energy E_w of the signal s in a window k using:

$$E_w(k) = \sum_{n=Nk-(N-1)/2}^{n=Nk+(N-1)/2} s^2(n) \qquad (1)$$

where N is the number of samples in the window. The considered window was 1 second of speech signal (N=1000 samples for a sampling frequency f_s=1000 Hz) with a 50% overlap between consecutive windows. If the energy of the signal in a window was smaller than a given threshold, E_{THR}, the signal was labeled "no fall event" and no further processing was performed. The energy of the signal has to be above E_{THR} in both channels in order to proceed to the next step.

B. *False alarm removal using the height of the signal*

If the signal in a window had significant energy, we next checked its height. If the computed height of the signal was above 2 feet, we labeled it "no fall". If the computed height was lower than 2 feet the signal was further evaluated by the recognition procedure. The height of the sound is computed using the correlation between the sound recorded by the two sound sensors. The signal correlation was computed using the whitened spectrum cross-correlation (rather than common time domain cross-correlation) [18], that is:

$$R_{12}(t) = \sum_{n=0}^{N-1} \frac{S_1(n)S_2(n)^*}{|S_1(n)||S_2(n)|} e^{i2\pi nt/N}, \qquad (2)$$

where $S_i(n)$ is the Fourier transform of the signal $s_i(n)$ received by the i^{th}, $i \in \{1,2\}$, sound sensor and $t \in [-N,N]$. Then, we computed the delay, δ_{12}, between the signals received by two sensors as:

$$\delta_{12} = \arg\max_{t\in[-N,N]} \{R_{12}(t)\}. \qquad (3)$$

To make the search more efficient the maximum of R_{12} was searched in the $t\in$ [-20, 20] interval (each time delay interval of 0.001s represents a path difference of about 1 ft). In fact, we only used the sign of the delay δ_{12}, that is, the signal was labeled "no fall" if δ_{12}>0 (the signal reached sensor 1 faster than sensor 2).

C. *Fall recognition*

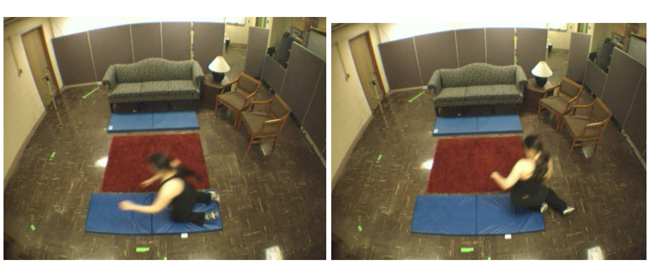
The feature extraction follows the energy calculation if the signal from a window has the energy above E_{THR}. We used the mel frequency cepstral coefficients as features. The number of coefficients (features) used was C=7. The features were extracted using the Matlab function, *mfcc*, from [19]. To make the system less dependent on the distance to the sound source, we did not use the first cepstral coefficient (proportional to the signal average) in the recognition procedure. The recognition was performed using the K-nearest neighbor (KNN) procedure with K=1. The "fall" and "no fall" training samples used in KNN were extracted from a fall session recorded by the same stunt actor but different from the six test sessions (see next section). A fall has to be detected in two consecutive windows in order to be reported as such. A summary of the algorithm if presented in Figure 3.

```
FOR each window (k)
    -Compute energy E(k) given by (1)
    -IF E(k)> ETHR
        --Compute the lag δ12 between s(1,k) and s(2,k) with (2)&(3)
        --IF δ12 <0 (sound height is 'Low')
            FOR each channel i
                Extract features from signal in channel s(i,k)
                Do KNN to detect a fall
            END
            IF (detect a fall in channel 1) AND (a fall in channel 2)
                IF previous_fall==1 =>
                    "found a fall": set fall(k)==1 and previous_fall==0
                ELSE
                    "found a possible fall": fall(k)=0, previous_fall==1.
                END
            ELSE (fall detected in only one channel)
                "no fall", fall(k)=0;
            END
        --ELSE (sound height is 'High')
            "no fall", fall(k)=0
        --END
    -ELSE (sound energy is low)
        "no fall", fall(k)=0
    -END
END FOR
```

Fig. 3. The fall detection algorithm used in this paper.

IV. Study Methodology

The falls used in this experiment were performed by a stunt actor [20]. The actor was instructed by our nursing collaborators to fall as an elderly person would fall. There were 5 types of falls: forward (Figure 4.a), backward, toward-left (Figure 4.b), toward-right and fall from a chair.



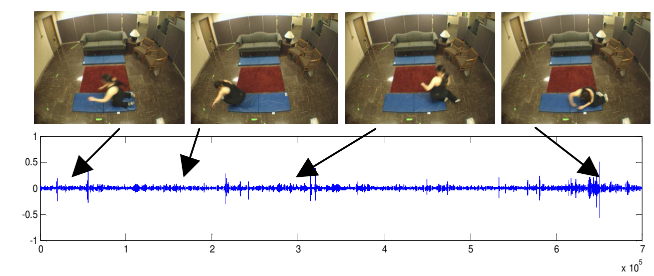
a. Forward fall b. Left-side fall

Fig. 4 Stunt actor performing falls

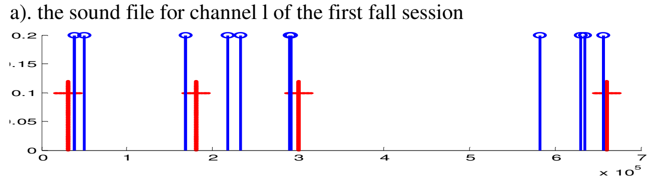
A typical fall session was about 10-15 minutes long and it contained 3-5 falls one of each type mentioned above. A nurse directed the actor during the fall session, instructing her when and how to fall and when to get up. Additionally , various other sound such as moving chairs, table knocks, feet stomping were produced by the actor and by the team members. We recorded 6 fall sessions with a total of 23 falls and a total recording time of 1.3 hours. A special 20 minute long session with 14 falls and noises was recorded and used for training. From this training session, 14 fall files and 25 false alarm files (1 second long) were extracted and used in the KNN procedure. The sound was sampled at f_s=1000 Hz. We intend to increase the sampling frequency in the future to at least 10 kHz.

V. Results

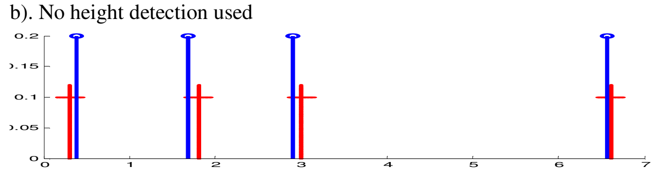
The falls and their detection results for the first fall session (4 falls in 13 minutes) are shown in Figure 5.



a). the sound file for channel l of the first fall session



b). No height detection used



c). With height detection

Fig. 5 The fall detection for the first fall session at E_{THR}=0.3 (horizontal axis is time; in b) and c) plot values are 0.1 with crossed line at a detected fall location and 0.2 with circled line at the ground truth locations)

The time position of the falls showed by the pictures in Figure 5 are marked on the sound file in Figure 5.a. In addition, in Figure 5.b the recorded falls are denoted with a short, crossed line. The false alarms in Figure 5 were produced by knocking or dropping objects such as a pen or a book on a table (not shown). Horizontal portion of the crossed line represents the uncertainly of the fall location (30 seconds), that is, if a fall event was found within 15 second of the recorded time of the fall it was declared a "detected fall". The longer lines (0.2 in length) represent the falls detected by the algorithm. Since the false alarms in Figure 5.b were produced by dropping objects on a table with height of about 3 ft, they were eliminated (at an energy threshold E_{THR}=0.3) by the introduction of the height detection algorithm (Figure 5.c).

To better describe the performance of the algorithm for all the fall sessions, we used a modified ROC curve that has on the x-axis the (number of false alarms)/hour instead of the false alarm rate (since, usually, the total number of false alarms is not known). By varying the energy threshold E_{THR}={0.1, 0.3, 0.6, 0.8, 1} we obtained 5 {detection rate, false alarm rate} pairs (see Figure 6). The detection rate for each threshold was computed as (# of falls detected)/23 and the false alarm rate as (#false alarm/hour).

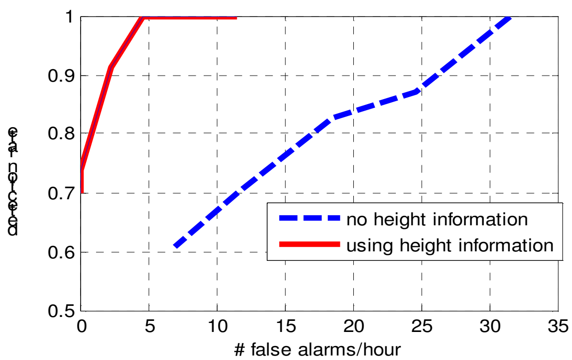


Fig. 6. The modified ROC curve for all 6 sessions used in the pilot study

We see from Figure 6 that using the height of the sound has dramatically improved the false alarm rate. For example, we detect 100% of the falls with about 30 false alarms in the "no height" case and with only about 5 if we use the height. However, our false alarm rate is still very large at about 5 alarms per hour. We believe that a target false alarm rate for an operational false detection system should be 1 in 24 hours which is about two orders of magnitude smaller than our current one.

VI. Conclusions

We presented a prototype of a dedicated fall detection system based on a linear array of audio sensors. Since the major problem with such a system is represented by false alarms we presented a possible solution in which we use the height of the signal. Different array shapes (circular) may be considered for further reduction of the false alarm rate.

In future work we intend to integrate a motion detector that will remove the detection if a motion is sensed after the event. The motion detector can help the system learn new noises by improving the false alarm library.

This pilot project was performed on a very limited data set. We intend to collect more data using people of different weights, various fall surfaces (carpet, hard wood floor, etc) and adding more sound sources (radio or a TV set).

References

[1] Center for Disease Control (CDC), http://www.cdc.gov/ncipc/factsheets/adultfalls.htm.

[2] J. A. Stevens, PS Corso, E. A. Finkelstein, T. R. Miller, "The costs of fatal and nonfatal falls among older adults", *Injury Prevention*, 12:290–5, 2006

[3] P. J. Ratclife, J.G.G Ledingham, P. Berman, G.K. Wilcock, J. Keenean, "Rhabdomyolysis in elderly people after collapse", *British Med. J.*, vol. 288, pp. 1877-8, 1984.

[4] R. J. Gurley, N. Lum, M. Sande, B. Lo, M. H. Katz "Persons found in their homes helpless or dead". *N Engl J Med.*; 334(26):1710-6, 1996.

[5] C. G. Moran, R.T. Wenn, M. Sikand, A.M. Taylor, Early mortality after hip fracture: is delay before surgery important", *J. of Bone and Joint Surgery*, pp. 483-9, 2005.

[6] N. Noury, A. Fleury, P. Rumeau, et al., "Fall detection-principles and methods", Proc. of the 29th IEEE EMBS, Lyon, France, Aug. 23-26, 2007.

[7] M. Alwan, P. J. Rajendran, S. Kell, D. Mack, S. Dalal, M. Wolfe, R.Felder, " A smart and passive floor-vibration based fall detector for elderly", The 2nd IEEE International Conference on Information & Communication Technologies: from Theory to Applications - ICTTA'06, April 24 - 28, 2006, Damascus, Syria.

[8] C. Rougier, J. Meunier, A. St-Arnaud, J. Russeau, "Fall detection from human shape and motion history using video surveillance", 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07), 2007.

[9] D. Anderson, R.H. Luke, J. Keller, M. Skubic, M. Rantz , Aud M, "Linguistic summarization of activities from video for fall detection using voxel person and fuzzy logic. *Computer Vision and Image Understanding*. (In Review)

[10] A. Sixsmith, N. Johnson, R. Whatmore, "Pyrolitic IR sensor arrays for fall detection in the older population", J. Phys. IV France, vol. 128, pp153-160, 2005.

[11] S. Dalal, M. Alwan, R. Seifrafi, S. Kell, D. Brown, "A Rule-Based Approach to the Analysis of Elders' Activity Data: Detection of Health and Possible Emergency Conditions", AAAI Fall 2005 Symposium (EMBC). Sep. 2005.

[12] E. Castelli, M. Vacher, D. Istrate, L. Besacier, J.F. Serignat, "Habitat telemonitoring system based on the sound surveillance", ICICTH (International Conference on Information Communication Technologies in Health), 11-13 July 2003, Samos Island, Greece.

[13] M. Vacher, D. Istrate, L. Besacier, E. Castelli, Jean-Francois Serignat, "Smart audio sensor for telemedicine" Smarts Objects Conference (SOC) 2003, 15-17 May, Grenoble, France.

[14] D. Istrate, E. Castelli, M. Vacher, L. Besacier, JF Serignat, "Information extraction from sound for medical telemonitoring", *IEEE Trans. on Information Tech. in Biomedicine*, vol. 10, no.2, April 2006.

[15] N. C. Laydrus, E. Ambikairajah, B. Celler, "Automated sound analysis system for home telemonitoring using shifted delta cepstral features", 15th International Conference on Digital Signal Processing, pp. 35-38, 2007.

[16] C. Schmandt, G. Vallejo, "'Listenin' to domestic environments from remote locations", Proc. of the 2003 Int. Conf. on Auditory Display, Boston, MA, 6-9 July 2003.

[17] J.G Proakis, D.G Manolakis, "Digital signal processing", 4th edition, 2007.

[18] J.M. Valin, F. Michaud, J. Rouat, D. Letourneau, "Robust sound source localization using microphone array on a mobile robot", Proceedings International Conference on Intelligent Robots and Systems, pp. 1228-1233, 2003.

[19] M. Slaney, "Auditory toolbox 2.0", http://www.slaney.org/malcolm.

[20] Rantz M., Aud M., Alexander G., Wakefield B., Skubic M., Luke R.H., Anderson D., Keller J.K., (2008) "Falls, Technology, and Stunt Actors: new Approaches to Fall Detection and Fall Risk Assessment," *Journal of Nursing Care Quality*, in press.