A MULTI-SENSOR SYSTEM FOR FALL DETECTION IN AMBIENT ASSISTED LIVING CONTEXTS

Giovanni Diraco[1], Alessandro Leone[1], Pietro Siciliano[1], Marco Grassi[2] and Piero Malcovati[2]
[1]*Institute for Microelectronics and Microsystems, CNR, Via Monteroni, 73100, Lecce, Italy*
[2]*Dept. of Electrical Engineering, University of Pavia, 1, Via Ferrata, 27100, Pavia, Italy*

Keywords: Multi-sensor network, Ambient assisted living, Fall detection, Time-of-flight camera, Wearable accelerometer.

Abstract: The aging population represents an emerging challenge for healthcare since elderly people frequently suffer from chronic diseases requiring continuous medical care and monitoring. Sensor networks are possible enabling technologies for ambient assisted living solutions helping elderly people to be independent and to feel more secure. This paper presents a multi-sensor system for the detection of people falls in home environment. Two kinds of sensors are used: a wearable wireless accelerometer with onboard fall detection algorithms and a time-of-flight camera. A coordinator node receives data from the two sub-sensory systems with their associated level of confidence and, on the basis of a data fusion logic, it operates the validation and correlation among the two sub-systems delivered data in order to rise overall system performance with respect to each single sensor sub-system. Achieved results show the effectiveness of the suggested multi-sensor approach for improving fall detection service in ambient assisted living contexts.

1 INTRODUCTION

Among all Information and Communication Technologies Sensor Networks (SNs) are very promising as possible enablers of Ambient Assisted Living (AAL) solutions towards more secure and independent living. Traditional in-home monitoring tasks can be drastically improved thanks to the increasing availability of low-cost, low-power, small embedded devices able to sense, process and transmit data via wired/wireless communications. The ubiquitous deployment of various kinds of sensor nodes, such as cameras, accelerometers, gyroscopes and so on, including placement of sensors on the body, ensures the constant in-home monitoring of person's health. This paper presents a multi-sensor system for people fall detection with particular interest in protecting older people living alone. Since one of the major causes of injury and fear for older people is fall, SNs should be exploited to automatically provide as fast as possible call for assistance when needed, minimizing of course false alarms to improve the performance of the system and thus of the provided service. It has in fact been demonstrated that the delivery of assistance after a fall may reduce the risk of hospitalization by over 25% and of major injury or death by over 80% (Shumway-Cook, 2009). On the other hand, although the problem of in-home monitoring is socially important, nonetheless the challenge is to determine an acceptable trade-off between safety and privacy intrusion. At this purpose, the suggested system includes two different kinds of sensor nodes: a privacy-preserving Time-Of-Flight (TOF) camera and a wearable wireless accelerometer. The use of two complementary sensors coordinated by a central node (coordinator), provided with a data fusion logic, allows to prevent false alarms or missed falls. In case of emergency, the coordinator node communicates with care holders and relatives of the assisted person through a gateway node. At the same time the privacy is guaranteed both within the house, because the TOF camera processes only distance information (appearance information is neither processed nor recorded) providing only discrete high-level features and outside the house since the gateway may deliver to caregivers only falls alarms together with their level of confidence.

2 SYSTEM OVERVIEW

The fall-detector system shown in Figure 1 includes two different available commercial sensors: a TOF

Diraco G., Leone A., Siciliano P., Grassi M. and Malcovati P..
A MULTI-SENSOR SYSTEM FOR FALL DETECTION IN AMBIENT ASSISTED LIVING CONTEXTS.
DOI: 10.5220/0003834202130219
In *Proceedings of the 1st International Conference on Sensor Networks* (SENSORNETS-2012), pages 213-219
ISBN: 978-989-8565-01-3

camera and a MEMS three-axial accelerometer. Each of the two sensors is connected to an embedded PC acting as coordinator node which communicates with an ADSL gateway to ask for assistance in case of fall. The coordinator includes a fuzzy rule-based data fusion logic that aggregates information from the two sensor nodes in order to produces a single data. The TOF camera is connected to the coordinator node through USB 2.0 connection, while the wearable wireless accelerometer communicates with the coordinator node by means of a ZigBee radio link with a serial connected transceiver. The employed TOF camera, a MESA SwissRanger4000 (MESA Imaging AG, 2011), allows the description of the acquired scene in terms of 3D distance measurements (depth maps). The wearable wireless accelerometer sub-system is made up of three main blocks: the 3-axial MEMS ST-LIS3LV02DL sensor (STMicroelectronics, 2008) with I^{2}C/SPI digital output, a low power XILINX Spartan-6 FPGA with embedded fall detection routines and a ZigBee radio module to deliver potential fall alarms together with their level of confidence to the coordinator.

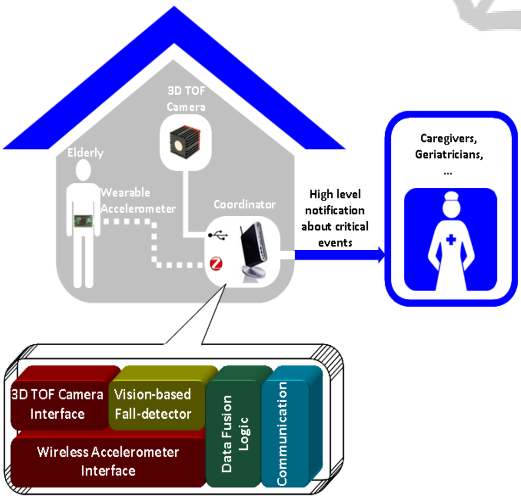

Figure 1: Overview of the multi-sensor system in which the information coming from the wearable accelerometer module and the TOF camera is combined by the coordinator node in order to reliably detect falls.

3 ACTIVE VISION SUB-SYSTEM

The 3D data coming from the TOF camera are processed by the vision algorithmic framework that provides the following functionalities: people counting, fall detection and body posture recognition. The vision algorithmic framework includes a first algorithmic level providing early vision primitives such as background modeling, people segmentation and tracking, and camera self-calibration (further details can be found in Leone, 2009). Instead, the second algorithmic level is more specifically dedicated to aspects related with feature extraction and classification. The silhouette dimensions and the gait trend are the features, extracted and analyzed, for people detection and counting functionalities. The detection of falls is based on the analysis of the trend of the person's centroid (i.e. approximation of the center-of-mass) height with respect the floor plane.

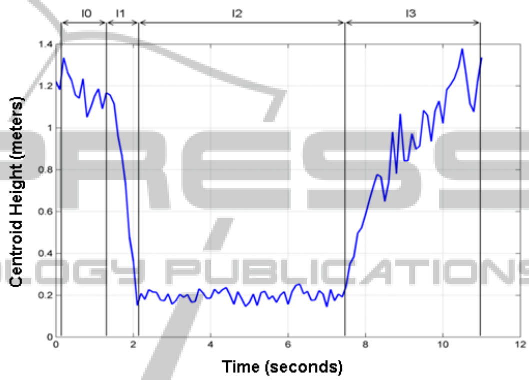

Figure 2: Typical trend of centroid height detected by the 3D vision sensor.

A typical trend of the centroid height during a fall processed by the 3D vision sub-system is reported in Figure 2 in which four phases can be distinguished during a fall: the pre fall phase (I0), the critical phase (I1), the post fall phase (I2) and the recovery phase (I3). Falls are detected by using two features: the person's centroid height and the duration of fall phases. In particular, a fall event is characterized by: 1) a centroid distance lower than a prefixed length threshold TH1 of about 40 centimeters; 2) a critical phase duration lower than a threshold TH2 of about 900 milliseconds; and 3) an unchangeable situation (negligible movements) greater than a time threshold TH3 of about 4 seconds. Thresholds TH1, TH2 and TH3 are experimentally defined according to Noury et al. (2007).

The body posture recognition functionality is able to discriminate four main postures, named Standing (ST), Bent (BE), Sitting (SI) and Lying down (LY). Postural features are extracted by analyzing the shape of two volumetric point distributions of which an example is reported in Figure 3 for the SI posture.

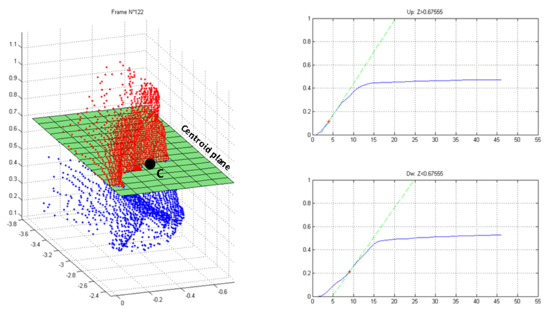

Figure 3: 3D point cloud of a posture (Sitting) and the related upper and lower volumetric point distributions.

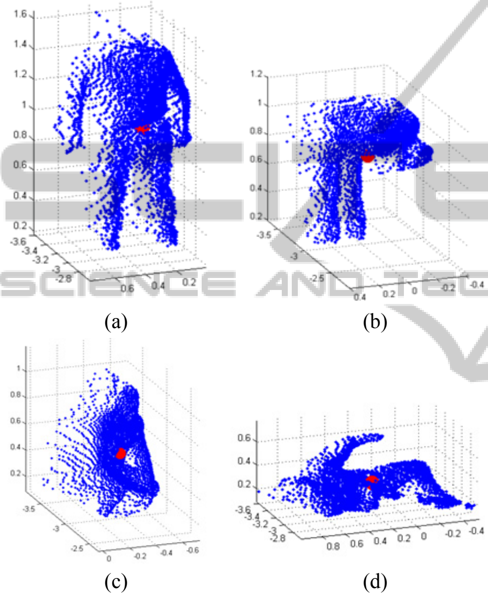

Figure 4: 3D point clouds related to the four main postures: a) Standing, b) Bent, c) Sitting, d) Lying down.

The point clouds of all four postures are reported in Figure 4. A good generalization ability during classification is relevant since postures are not perfectly repeatable: the acquisition viewpoint varies in function of subject's position and some level of variation in data range is expected due to noise effects. Therefore, a multi-class formulation of the Support Vector Machine (SVM) classifier is adopted in conjunction with rotation and scale invariant features in order to classify the four main postures. The binary nature of SVM is adapted to the multi-class nature of the posture problem by using a one-against-one strategy. Since good results are documented in literature related to posture recognition, a Radial Basis Function (RBF) kernel is used and the associated parameters are tuned according to a grid search procedure (for further details on posture refer to Leone, 2011).

4 WEARABLE SUB-SYSTEM

Regarding the hardware details for the wearable wireless accelerometer module, a simplified block diagram is shown in Figure 5.

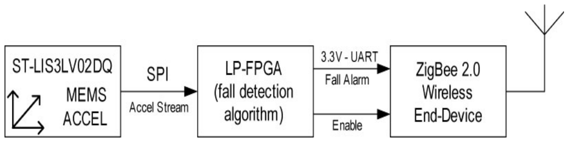

Figure 5: Block diagram of the developed wearable wireless accelerometer hardware.

The core of this system is the FPGA that controls the accelerometer,reads-out its digital output and delivers the necessary information to the ZigBee radio module, in order to have it transmitted to the coordinator if necessary and there merged with 3D vision data. Two operating modes are available for the device: in the first one, which has been very useful during test campaign and debug phase, the FPGA sends all the raw data acquired from the accelerometer to the ZigBee transmitter, and, of course, eventual alarms. By contrast, during the normal operative condition, the transmitted data are limited to the possible alarms and to a signal useful to determine if the system is working and "in range" and, of course, if the accelerometer has been worn. In both operating modes, at start-up, the control of the MEMS accelerometer is comprehensive of an initial setting of its internal registers, that allows the choice of the most suitable characteristics for fall detection purpose: among them, for example, the device full-scale, the resolution and the read-out rate. Once the setting phase is performed, the FPGA keeps querying the accelerometer, switching among the three axes. The accelerometer replies to every request with one word, containing the instantaneous acceleration experienced. This procedure is done continuously at the selected rate, which has been set by default at 10 Hz for overall acceleration query. In normal operating mode the acceleration information, when its value crosses a wake-up threshold defined within a given algorithm, is saved into suitable registers for 32 cycles, thus acquiring the amount of information necessary to the ad-hoc algorithm in order to properly reveal falls. Such algorithms are loaded into non volatile FPGA embedded core which, as underlined, has the computational task to detect falls with their associated level of confidence. The main advantage of having onboard algorithms for fall detection is to reduce the power consumption of the device, which may communicate with the coordinator using ZigBee radio only in case of a fall-like event, thus extending battery autonomy (rechargeable Li-Ion 3.7V, 1100mAh) up to a couple of weeks. The low power core runs in parallel six different routines for fall detection, providing two approaches for each axis. In order to save further energy, the sensor device wakes up if a given acceleration threshold is crossed on at least one axis. The first approach measures the *stress* in terms of fall energy, while the second checks the acceleration *shape* and the estimated orientation/behaviour of the person after the fall. In fact, the MEMS device is also able to provide static information about acceleration, thanks to its internal DC coupling and thus the orientation of the assisted person may be tracked. All relative acceleration components are referred to the gravity and an additional digital low-pass filter is used to give emphasis to reference DC information. The main idea behind this algorithm is that every time a fall event occurs, the acceleration changes significantly and after the impact the person lays down, often changing the static acceleration orientation with respect to ground. Sometimes the fall event happens fast, while other times the event is slower. In order to measure the quantity of acceleration stress or energy within an over-threshold event, observing the acceleration data over a window of about 0.5 seconds seems to be the best compromise between effectiveness, complexity and low power consumption for what concerns calculation. Further details, including the importance of the pre-fall assisted person behavior and the great difficulty of emulating real falls in controlled environment are provided by Grassi et al. (2010).

5 FUZZY-BASED DATA FUSION

The features extracted by the two sensor sub-systems are merged in order to provide more accurate information for the critical event rather than each one of the individual source alone. The detection of fall events by means of a multi-sensor network deployed in an apartment having various rooms is a complex system to be modeled; consequently a fuzzy rule-based approach is a suitable choice. The crisp variables provided by the 3D vision sub-system are the following: 1) a Boolean variable indicates if there is any person inside a room or not; 2) a real valued confidence index in the range of (0,1) indicates the probability that the person is fallen down on the floor; 3) four real valued confidence indexes ranging in (0,1) associated to each body posture. The first variable allows to handle the situation in which the person is out of the Field-Of-View (FOV) of the TOF camera and thus the only viable information comes from the accelerometer sub-system, if worn (e.g. if the person is inside the bathroom and the nearest 3D vision system is in the bedroom). Furthermore, it is important to note that the choice to associate a confidence index to each posture (the third variable) allows to model overlapped situations in which for example a person is seated and bent at the same time. The wearable wireless accelerometer sub-system provides two crisp variables that are: 1) a Boolean variable indicating if the accelerometer is working (i.e. it is “in range” and worn) or not, and 2) a real valued confidence index ranging in (0,1) indicating the probability that the person is fallen down as detected by the worn accelerometer. We remind that the employed ZigBee modules are able to cover a standard apartment of about 100m^2 by linking to at least a couple of receivers and thus coordinators associated to the 3D vision systems of the main rooms. Triangle and trapezoid functions are the kind of membership functions used to transform the crisp values provided by each sensor sub-system into fuzzy linguistic variables. The knowledge on which basis falls should be detected gives fuzzy rules like: “*if there is no one inside the TOF camera FOV then consider only data provided by the wearable accelerometer*” and “*if there are two or more people inside a room then don’t send alarms*” and so on. The fuzzy rules are processed by using the well-known Mamdani fuzzy inference technique, producing fuzzy outputs which must be decoded or “defuzzified” in order to get an aggregated system output. The “defuzzification” process is done by using the well-known centroid technique.

6 EXPERIMENTAL RESULTS

In this section experimental results coming from each of the two different sensor sub-systems are reported. The performance of the two subsystems has been evaluated by collecting a large amount of simulated actions. It has been crucial to use several people to collect the data, in order to define an algorithm as general as possible. Moreover, in the performed recording sessions, the kind of fall has been varied several times for each person (falling forward, backward and sideward). A total amount of 450 actions including 210 falls in all directions (backward, forward and lateral, with/without recovery post fall) were simulated involving 13 healthy male equipped with knee/elbow pad protections and crash mats. The simulated falls are compliant with those categorized by Noury et al. (2009) and they can be grouped into the following categories: backward fall ending lying (FBRS), backward fall ending lying and lateral (FBRL), backward fall with recovery (FBWR), forward fall with forward arm protection (FFRA), forward fall ending lying flat (FFRS), forward fall with recovery (FFWR), lateral fall (FL). Several Activities of Daily Living (ADLs) were simulated other than falls, in order to evaluate the ability to discriminate falls from ADLs. The simulated ADL tasks belong to the following categories: sit down on a chair and stand up (SITC), sit down on the floor and stand up (SITF), lie down on a bed and stand up (LYB), lie down on floor and stand up (LYF), bend down and pick up something (BND).

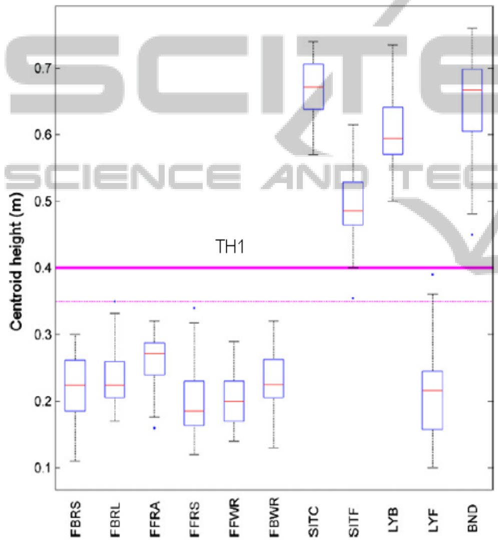

Figure 6: Statistical visualization of the minimum centroid height during the performed falls and ADLs.

6.1 Results from Vision Sub-system

As previously discussed, the vision-based fall detector is based on the tuning of three thresholds: TH1, TH2 and TH3. The first threshold TH1 alone is able to detect correctly all simulated falls, although it is not able to distinguish between a fall and a FFWR or between a fall and LYB/LYF. A statistical visualization of results related to the threshold TH1 is shown in Figure 6. The threshold TH1 alone identifies correctly 63.5% of ADLs as non-falls. By adding the second threshold TH2 the percentage of correctly detected ADLs rises to 79.4%, since the threshold TH2 allows to discriminate correctly a "voluntary lying down on floor" from an involuntary fall characterized by a shorter duration of the critical phase. The statistical visualization of TH2 discrimination capability is shown in Figure 7. By using all thresholds (TH1, TH2, TH3) simultaneously a reliability of 97.3% and an efficiency of 80.0% are achieved, since the threshold TH3 allows to detect correctly falls with recovery as non-falls by considering the duration of the post fall phase shorter than 4 seconds in case of recovery.

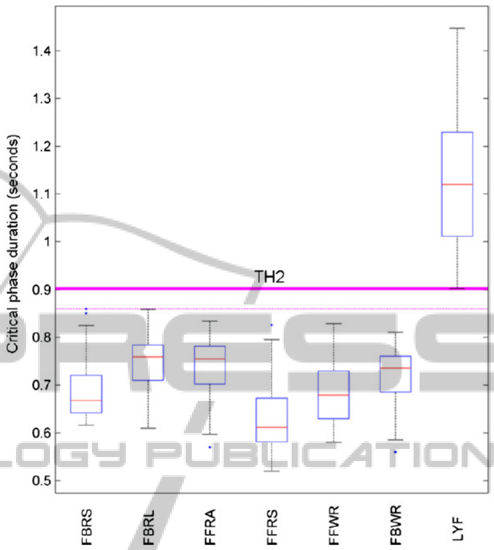

Figure 7: Statistical visualization of the critical phase duration during the performed falls and ADLs.

In the following Eq. 1 the employed definitions for efficiency and reliability scores are reported.

$$\text{Efficiency}=\frac{TP}{TP+FP}, \quad \text{Reliability}=\frac{TP}{TP+FN}. \qquad (1)$$

The best classification rates for body posture recognition are found with the optimal parameters (K;γ)=(1;32), where K is the regularization constant and γ is the kernel argument. Postures were taken at various distances from the camera, ranging from 2.5 meters to 5 meters. Classification rates at the varying of the camera distance are reported in Figure 8.

The active vision-based fall detector shows good performance also compared with other related studies. For example, a similar study was conducted by Brulin and Courtial (2010). In their work the authors investigated a multi-sensors system for fall detection in which traditional cameras (passive vision) were used for postures recognition whereas PIR detectors and thermopiles were used for presence detection and people location respectively. The authors achieved noticeable results in optimal ambient conditions, however in real-world working condition lighting variations, shadows and perspective distortions (typical issue of monocular passive vision) demoted features resulting in performance decrease. Whereas the perspective distortion can be faced by using a multiple passive camera setting as done by Cucchiara et al. (2007), the use of multiple cameras brings a number of new problems such as, just to cite a few, the stereo correspondence problem, the occlusion handling and the calibration of multiple cameras (Hu, 2006). In addition other problems such as shadows and lighting variations remain virtually unmodified.

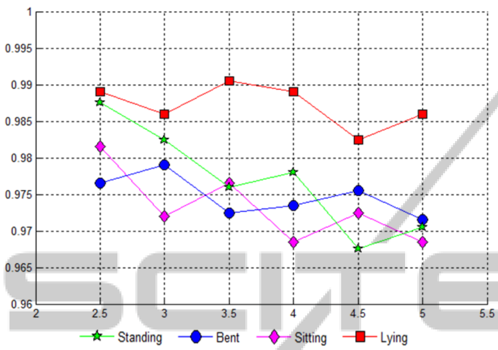

Figure 8: Classification rates at varying of camera distance from 2.5 to 5 meters.

On the other hand, the suggested solution, since it is based on active vision, presents several advantages. A TOF camera with fewer pixels than a multiple camera system can deliver more 3D information, when the multiple camera system has too many false matches due to the stereo correspondence problem. Incomplete range data are not produced by TOF camera because illumination and observation directions are collinear preventing the formation of shadows. The depth resolution does not depend on the optical arrangement (as in the case of multiple cameras) and hence extrinsic calibration is limited to the only estimation of the camera pose, whereas intrinsic calibration is not required at all. TOF cameras are fully independent of external light conditions, since they are equipped with an active light source. Finally, it is important to note that the TOF camera guarantees the person's privacy, since appearance (chromatic) information is not acquired and low-resolution depth measurements are not sufficient to reveal the person's identity or to compromise the feelings of intimacy.

6.2 Results from Wearable Sub-system

The measurements have been done setting the MEMS device to the full-scale range of ± 2g, leading to an absolute sensitivity in terms of acceleration exhibited by the wireless module of 100µg, considering 12-bit resolution. The average current consumption at 3.2-3.8V supply is of the order of 200µA when waiting for an acceleration event, about 1mA while processing the event (considering 10samples/s per axis) and 30mA when transmitting a fall-flag or in streaming mode. Since the algorithm has been developed based on the training events acquired, the data collection campaign really played a fundamental role. By exploiting the above mentioned ADLs different thresholds have been evaluated using data from several sessions for training. After that, the performance of the algorithm has been evaluated on the non-training oriented collected data and the results are shown in Table 1 for different suitable thresholds for acceleration shape approach (THL, THM, THH). An example of a plot of the acceleration along one axis (Y) has been reported in Figure 9. It is possible to verify that every fall origins a spike in the acceleration that is followed by some oscillations and then the value remains stable at a sensibly different value from the starting one.

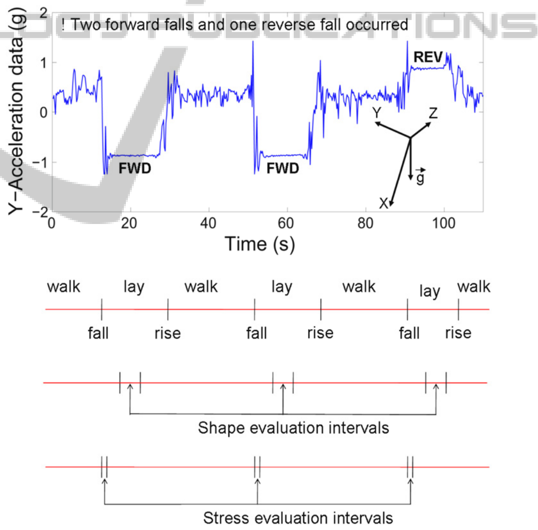

Figure 9: Acceleration data stream session measured along Y-axis with wireless device worn on the belt. The Y-axis is orthogonal to the belt.

Figure 9 also indicates the action that the actor was performing when the acceleration data was recorded and we enlightened the portion of time during which the data are evaluated by the algorithm. The FPGA runs the algorithms and sends the results to the ZigBee transmitter implementing a 3.3V-UART protocol. The timestamps related to data and axis alarm flags are added by the coordinator, which is also in charge of data fusion. Considering that efficiency and reliability scores are of course a trade-off and taking into account also the combined performance of the two sensory sub-systems after data fusion operation, the intermediate value for threshold has been exploited.

Table 1: Wearable accelerometer performance.

Shape Threshold	Efficiency	Reliability
Low (THL)	98.0 %	61.5 %
Medium (THM)	88.4 %	79.3 %
High (THH)	55.1 %	97.2 %

6.3 Data Fusion Results

The data fusion process improves the detection performance thanks to the addition of both analogous and complementary information. A sample output of fall probability is shown in Table 2 in correspondence of two different kinds of falls and one normal activity. From the rows of the table, it can be seen as the merged fall probability (the ALARM column) improves the fall probability of each sub-system alone. In order to easy the performance comparison, the scores related to each standalone sub-system and the ones related to the data fusion are reported jointly in Table 3. The performance of the comprehensive framework underlines a significant raise in both efficiency and reliability.

Table 2: Sample output of fall probability.

Activity	Vision S.		Wearable S.	ALARM
	Fall	Posture	Fall	
FL	0,87	Lying	0,82	0,95
FBWR	0,64	Sitting	0,73	0,44
SITF	0,40	Sitting	0,62	0,36

Table 3: System Performance.

Sub-system	Efficiency	Reliability
TOF camera	80.0 %	97.3 %
Accelerometer	88.4 %	79.3 %
Data fusion	94.3 %	98.2 %

7 CONCLUSIONS

A multi-sensor framework for indoor people fall detection has been developed and experimental results with real actors following state-of-the art guidelines (Noury, 2009) have been performed first for each single sensor sub-system and then for the overall system. A fuzzy-based data fusion logic has been proposed able to effectively handle the uncertainty present in AAL contexts. The presented system improves the performance of people fall detection by processing multiple sensors data, showing that SNs are a very promising approach for critical event detection with low false alarm rate. In addition, fuzzy rules can be easily modified and adjusted in order to meet specific environment constraints. Furthermore, the deployment of the presented multi-sensor fall detection system in apartments dwelled by elderly people is planed in order to validate the system in real conditions.

ACKNOWLEDGEMENTS

The presented work has been carried out within the Netcarity consortium funded by the EC, FP6.

REFERENCES

Brulin, D., Courtial, E., 2010, "Multi-sensors data fusion system for fall detection," In: *Proceedings of 10th IEEE ITAB*, pp. 1-4.

Cucchiara, R., Prati, A., Vezzani, R., 2007, "A multi-camera vision system for fall detection and alarm generation," *Expert Syst J*; vol 24, no. 5, pp. 334-45.

Grassi, M., Lombardi, A., Rescio, G., Ferri, M., Malcovati, P., Leone, A., Diraco, G., Siciliano, P., Malfatti, M., Gonzo, L., 2010, "An Integrated System for People Fall-Detection with Data Fusion Capabilities Based on 3D ToF Camera and Wireless Accelerometer," In: *Proceedings of IEEE Sensors*, pp. 1016-1019.

Hu, W., Hu, M., Zhou, X., Tan, T., Lou, J., Maybank, S., 2006, "Principal axis-based correspondence between multiple cameras for people tracking," *Pattern Analysis and Machine Intelligence*, IEEE Transactions on; vol. 28, no. 4, pp. 663-671.

Leone, A., Diraco, G., Siciliano, P., 2009, "Detecting falls with 3D range camera in ambient assisted living applications: A preliminary study," *Medical Engineering & Physics*; vol. 33, no. 6, pp. 770-781.

Leone, A., Diraco, G., Siciliano, P., 2011, "Topological and volumetric posture recognition with active vision sensor in AAL contexts," In: *IEEE IWASI*, pp. 110-114

MESA Imaging AG, 2011, "*SR4000 Data Sheet Rev.5.1*," Zurich, Switzerland, 26 August 2011, <http://www.mesa-imaging.ch>.

Netcarity UE Integrated Project, *www.netcarity.org*.

Noury, N., Fleury, A., Rumeau, P., Bourke, A. K., Laighin, G. O., Rialle, V., Lundy, J. E., 2009, "Fall detection - Principles and Methods", In: *Proceedings of 29th IEEE EMBS*, pp. 1663-1666.

Shumway-Cook, A, Ciol, M. A., Hoffman, J., Dudgeon, B., Yorkston, K., Chan, L., 2009, "Falls in the Medicare population: incidence, associated factors, and impact on health care," *Physical Therapy Association*; vol. 89, no. 4, pp. 324-32.

STMicroelectronics, 2008, "*LIS3LV02DL Data Sheet Rev.2*," January 2008, <http://www.st.com>.