Real Time Recognition of Elderly Daily Activity using Fuzzy Logic through Fusion of Motion and Location Data

Shadi KHAWANDI
Phd Student

University of Angers
Angers, France

Bassam Daya
Professor

Lebanese University
Lebanon

Pierre Chauvet
Professor

University of Angers
Angers, France

ABSTRACT

One of the major problems that may encounter old people at home is falling. Approximately, one of three adults of the age of 65 or older falls every year. The World Health Organization reports that injuries due to falls are the third most common cause of chronic disability. In this paper, we proposed an approach to indoor human daily activity recognition, which combines motion and location data by using a webcam system, with a particular interest to the problem of fall detection. The proposed system identifies the face and the body in a given area, collects motion data such as face and body speeds and location data such as center of mass and aspect ratio; then the extracted parameters will be fed to a Fuzy logic classifier that classify the fall event in two classes: fall and not fall.

Keywords

Fall detection; Motion data; Face detection; Center of mass; webcam.

1. INTRODUCTION

Population all over the world is facing enormous demographic changes, especially when it comes to the category of people aged over 65 years, whose percentage is anticipated to rise from 17.9% in 2007 to 53.5% in 2060, due to the increase in life expectancy and decrease in birth rate. With the progress of medicine, people can live longer with better health, and are able to live independently in their own homes. But these home environments are often in very remote places and getting any help when needed is really difficult. One of the major problems that may encounter old people at home is falling. A fall incident may lead to severe physical injuries and cause many disabling fractures: according to the World Health Organization, falls are the third cause of chronic disabilities. Moreover, a fall has dramatic psychological consequences that reduce the independence of the person. Approximately, one of three adults of the age of 65 or older falls every year. In fact, it has been shown in [1] that 28-34% of elderly people in the community experience at least one fall every year, and 40-60% of these falls lead to serious injuries. Furthermore, the risk of falling increases with age, and in 2 out of 3 cases it happens at home. Actually, people who suffer from a fall incident at home and remain on the ground for an hour or more, usually die within 6 months. In addition to the above, the cost of hospitalization after the fall is on the rise, and hospitals are being more and more crowded, so the capability to find a place in a hospital in order to recover from the fall is never easy.

In this paper, we propose a webcam system which combines location and motion data of the elderly. The proposed system (Figure 1) detects and identifies the face and the body in a

given area based on image processing approach, extract the face speed, the body speed, the center of mass and aspect ratio; then forward the extracted parameters to a Fuzy logic classifier that classify the fall event in two classes: fall and not fall. Based on the result of the classification, the decision will be made, and in case of a fall, an alarm will be triggered indicating the need for assistance.

Fig 1: Proposed System

2. RELATED APPROACHES

The most common solution for falls detection is wearable devices, [2], [3], [4], [5] and many tools have been developed in this scope, such as social alarm which is a wrist watch with a button that is activated by the person in case he/she suffers a fall, and wearable fall detectors, which are based on combinations of accelerometers and tilt sensors [2] that can monitor velocity, acceleration and vertical posture toward lying posture. However, critical problems are associated with these devices. The main problem with social alarms is that after a fall, the person is panicked, confused or even unconscious, so the button is often unreachable. For the wearable sensors, these autonomous sensors are usually attached under the armpit, around the wrist, behind the ear's lobe or at the waist. However, the problem of such detectors is that older people often forget to wear them [6], [7], [8],[9]; indeed their efficiency relies on the person's ability and willingness to wear them. Furthermore, these devices provide fairly crude data that is difficult to interpret. They also generate many false alarms, and often mix falling and non-falling events, so generally their reliability is doubted.

Another technique for recognizing fall events is acoustic based fall detection [10]. Audio signals obtained from a camera microphone or from an independent microphone are used to discriminate between a falling down event and other ordinary events. When a fall happens, the event is recognized due to the high amplitude signal produced, contrarily to ordinary actions like walking, sitting down or bending that have no distinguishable sound from the background. In practice, particular operations and transformations are applied to the audio signals in order to obtain a set of features like signal power, zero-crossing rate, signal to noise ratio, and variance. These features are compared with a threshold and if the values are higher than the threshold an alarm will be triggered indicating that a fall has occurred. Different methods are used to analyze the audio signal in order to extract the

features and use them in the fall detection approach. Such methods are analysis of the audio signal based on wavelet domain feature extraction or analysis based on Fourier domain feature extraction [11]. The problem with this type of detectors is that other sounds coming from the background may be mixed with falling events; so again, the reliability of the system is in question. Recently some research has been done to detect falls based on using image processing approach. This method is based on images obtained from installing home cameras that register the subject's movements and take pictures in multiple frames. The image from each frame is analyzed to extract different features that will be used in the decision phase, and many procedures were implemented.. A simple method used [10] [12] was based on analyzing aspect ratio of the moving object's bounding box. Other works [13],[14] used the normalized vertical and horizontal projection of segmented object as feature vectors. Nait-Charif [15] tracked the person using an ellipse and inferring falling incident when target person is detected as inactive outside normal zones of inactivity.

Sensor based approaches lack consistency when it comes to providing highly accurate automatic fall detection systems. Higher accuracy levels have been achieved to an extent using multi-dimensional combination of physiological and kinematic parameters.

Vision based approaches in comparison to others are certainly the area to look forward to.

Most of the existing vision based approaches lack flexibility. These approaches are often case specific and dependent on different scenarios. There is a need for a reliable and robust generic fall detection algorithm. Both ambience and sensor based approaches share a common disadvantage, generally, of subject data not being visually verified by the control or care service provider for accuracy. From a research perspective, there are issues relating to the availability of data sets of falls for training as well as evaluation. A comprehensive and robust fall detection system should possess both high sensitivity and good specificity. The existing approaches have not comprehensively satisfied the accuracy as well as robustness of a fall detection system. However, the existing approaches do provide a framework to further develop techniques as well as modify the existing algorithms to achieve better performance

3. PROPOSED APPROACH

The proposed approach consists of installing home cameras that will track the movement of the subject. Since our aim is to create a system that is easy to be implemented in home environments, an ordinary webcam is used in building the surveillance technique. The webcam will take the video stream and feed it to our system implemented using MATLAB environment. After being read, the video will be converted into a sequence of images. Each image will go through a series of transformations in order to extract the desired data.

The process begins by subtracting the background from the image. Then, it goes through the "cleaning" phase: noise will be removed, little holes will be eliminated and the openings will be filled. After that, the body and the face are clear and can be distinguished from other figures coming from the background. Then the exploitation of the image will become simple, and the parameters are ready to be extracted. Four

parameters are derived and studied: center of mass, aspect ratio, body speed and face speed.
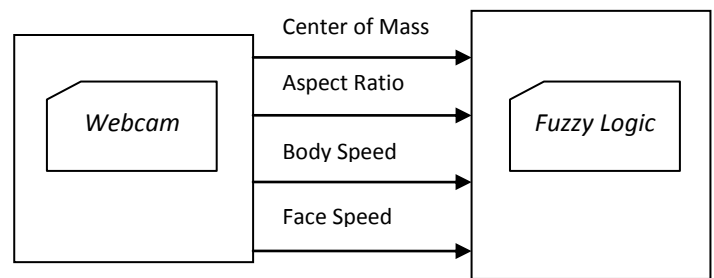

Fig 2: Proposed approach

- **Center of Mass**

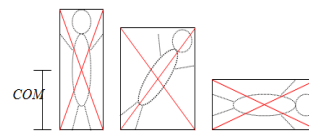The Center of Mass "COM" is defined as the height between the center of mass of the person and the floor

Fig 3: Center of Mass

Ranges:
Usually, the COM is between 15cm and 100cm. Therefore, two outstanding ranges may be taken into consideration:

   a. COM > 50: this is the normal case scenario. The person is standing, and the center of mass is far from the ground, thus chances that the subject has fallen are weak.
   b. COM<50: In this case, the center of mass is close to the ground. The subject could be sleeping or sitting, or they could be lying down due to a fall; therefore, a decision needs to be made in order to decide whether we have a falling situation or not.

- **Aspect Ratio**

   The Aspect Ratio "AR" of a person is a simple yet effective feature for differentiating normal standing poses from other abnormal poses. It can be computed by finding the bounding box which is defined as the smallest rectangle containing the moving object.

$$AR(n) = \frac{H(n)}{W(n)}$$

Where:

$H(n) =$

$Height\ of\ the\ rectangle\ for\ frame\ n$

$W(n) =$

$Width\ of\ the\ rectangle\ for\ frame\ n$

Ranges:
Usually, AR is limited between 0.15 and 6.The critical case is when AR equals 1. In this case, the height and the width of

the bounding box are even; the subject is leaning and the decision whether a fall has occurred or not is hard to be taken. Two ranges may be distinguished:

    a. AR>1: This is the normal case where height > width. In this case the person could be standing, walking, running, etc. Anyhow, the risk of a falling event is weak, and we can assume that no fall has occurred.

    b. AR<1: In this case height < width, which means that the subject is lying down. They could be simply sleeping, or a worst case could be present: the subject may have fallen.

- **Body Speed**

The speed of the body of the subject we are monitoring is defined according to the well-known relation:

$$Planar\ speed = {distance}/{time}\ (pixels/sec)$$

Where:
distance = distance between two consecutive frames (pixels)
time = processing time between two consecutive frames (sec)

Since Fall-Detection depends on both the speed and the distance face-webcam, therefore we will use the following relation:

$$Real\ Speed = {Planar\ Speed}/{Z}\ (pixels/sec)$$

Where:

$$z = \frac{X}{Y}$$

With:

$$X = {Width}/{Width0}\ ; Y = {Length}/{Length0}$$

Where:
Width: width of the detected face for each frame
Width: width of the detected face for the initial frame
Length: length of the detected face for each frame
Lenth0: length of the detected face for the initial frame

Ranges:

Usually speed is between 0 and 800 pixels/sec, and can be classified into the four different ranges that follow:

    a. Low speed: 0-200 pixels/sec: The movement is almost non present. The subject could be resting or could be lying on the floor after having suffered from a fall.

    b. Normal speed: 200-400 Pixels/sec: It corresponds to daily life activities such as walking. The probability of a fall event is low.

    c. Medium speed: 400-600 Pixels/sec: A relatively medium movement is detected. In this case the speed is considerable and the probability of a fall occurring must be taken into consideration.

    d. High speed: 600-800 Pixels/sec: In this case there's a quick and sudden movement. The risk of a fall is at its highest.

- **Face Speed**

Let's suppose that our subject is moving one of his hands; in this case the body speed will rise, yet no fall has occurred. Thus, in order to improve the reliability of our system, we will take into consideration the speed of the face. A fall could not occur unless both Face and Body speed are medium or high. The flow diagram of the proposed system is shown below (Figure 4)
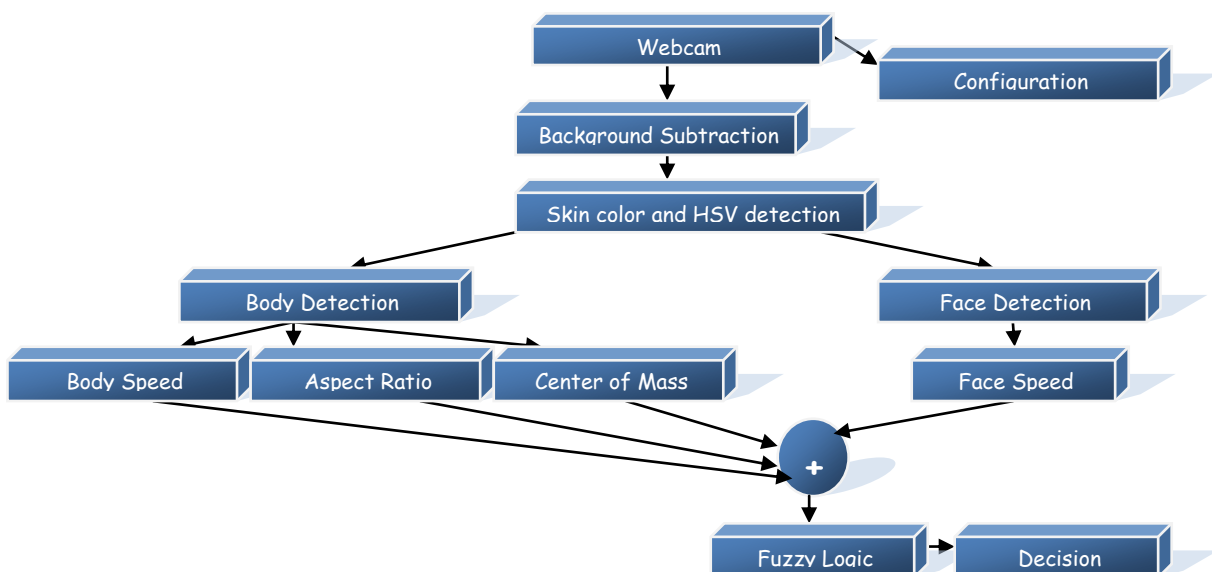
Fig 4: Flow diagram

3.1 Webcam

First, it is obvious several webcams should be used to cover the entire area to be monitored. Based on the presence of the face according to the webcam, a switch will be done to another webcam. In this paper, the proposed approach is limited to one webcam

3.2 Configuration

The is the initial phase in the system where the subject stands for few seconds towards the webcam in order to have his body and face detected as initial parameters that will be used in a later phase

3.3 Background Subtraction

Background subtraction (Figure 5) is a particularly popular method to detect moving regions in an image by differencing between current image and a reference background image in a pixel-by pixel way

Fig 5: Background Subtraction

3.4 Skin color and HSV detection

The images captured by the webcams are then processed by the system to detect skin color. This is an effective technique for determining whether an image contains a face or hands. In this technique, the appropriate threshold values are defined for all the pixels in a color space. Different color spaces are used to represent skin color pixels: RGB, RGB standard, HSV (or HSI), YCrCb, and HSV. After the detection of skin color, pixels, image filtering (erosion and dilation) is carried out.
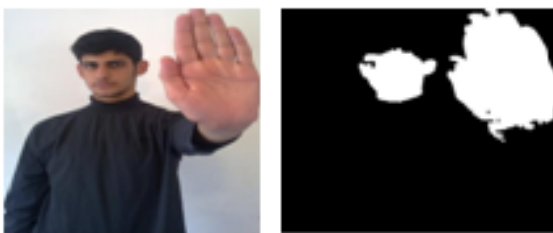

Fig 6: Image after skin color detection

3.5 Face detection

After identifying the skin areas, it is necessary to distinguish the face. For this, the shape of the detected object is compared with an ellipse. This correlation technique is very effective and efficient (Figure 7)
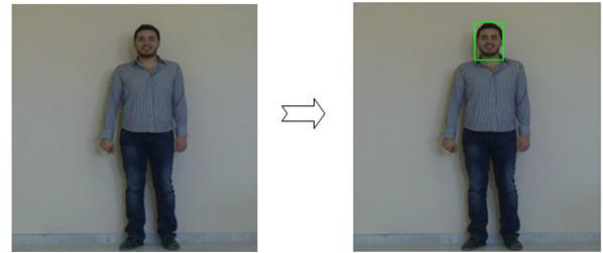

Fig 7: Face detection

3.6 Body detection

After detection of skin color, we will use some morphology as erosion and dilation for cleaning the image by eliminating the isolated pixels.
The body is detected by the smallest rectangle that contains the figure as the difference between the recent image and the background image (Figure 8).
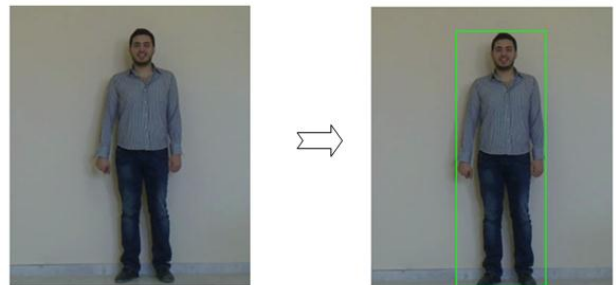

Fig 8: Body detection

3.7 Fuzzy logic classification

The approach we used in the classification phase is the fuzzy logic classifier (Figure 9). Four inputs are fed in the classifier, and one output is obtained.

Two different systems were built: the first one is the "Mamdani-type" inference, and the second one is the "Sugeno-type" inference. For the "Mamdani-type" inference, for the AND and OR operations we used the minimum and maximum methods respectively. For the "Sugeno-type" inference, for the AND and OR operations we used the product and probabilistic methods respectively
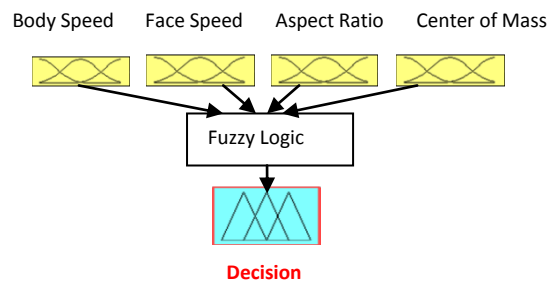

Fig 9: Fuzzy Logic classifier

For the input parameters, the following levels were applied:

- Body Speed: Low (L): 0->70 px/s; Normal (N): 70->110 px/s; High (H): 110->700px/s
- Face Speed: Low (L):0->90 px/s; Normal (N): 90->200 px/s; High (H): 200->700px/s
- For AR: Critical (C): >1(100%); Not Critical (NC); <1(100%)
- COM: Critical: 0->40%; Non Critical : 40% -> 100%

For the output parameter, Two levels were applied: Fall (F) and No Fall (NF).

For the identification of the fall event we defined the rules shown below in the figure below:

Body Speed	Face Speed	Aspect ratio	COM	Decision
L	L	NC	NC	NF
L	L	NC	C	NF
L	L	C	NC	NF
L	L	C	C	NF
L	N	NC	NC	NF
L	N	NC	C	NF
L	N	C	NC	NF
L	N	C	C	F
L	H	NC	NC	NF
L	H	NC	C	F
L	H	C	NC	F
L	H	C	C	F
N	L	NC	NC	NF
N	L	NC	C	NF
N	L	C	NC	NF
N	L	C	C	F
N	N	NC	NC	NF
N	N	NC	C	F
N	N	C	NC	F
N	N	C	C	F
N	H	NC	NC	NF
N	H	NC	C	F
N	H	C	NC	F
N	H	C	C	F
H	L	NC	NC	NF
H	L	NC	C	NF
H	L	C	NC	NF
H	L	C	C	F
H	N	NC	NC	NF
H	N	NC	C	F
H	N	C	NC	F
H	N	C	C	F
H	H	NC	NC	NF
H	H	NC	C	F
H	H	C	NC	F
H	H	C	C	F

Fig 10: Classification Rules

4. EXPERIMENTS AND RESULTS

In this section, we present the results obtained using the proposed system.

Table 1 Results obtained for 30 frames

Frame	Body Speed	Face Speed	Aspect ratio	Center of Mass	Fall Decision
1	401.536	0.83296	0.30658	96.648	0.2
2	48.2477	1.64809	0.38425	96.9624	0.2
3	379.470	19.2494	35.4874	11.9050	0.5
4	468.838	23.8936	5.39776	0.51545	0.525
5	373.275	24.7342	7.12848	12.2281	0.5217
6	40.6910	31.0164	3.41793	20.519	0.2
7	222.271	42.6989	5.73401	15.7362	0.5204
8	315.535	44.9458	8.08624	19.4310	0.5195
9	481.935	57.8129	5.14482	7.60872	0.5227
10	108.779	98.2922	1.76922	8.60819	0.6739
11	90.3688	127.693	2.05357	23.0918	0.5233
12	453.333	128.753	6.67762	23.0672	0.8209
13	70.1276	130.065	8.68151	10.7349	0.5236
14	171.595	140.812	7.46606	22.1560	0.8197
15	94.3309	143.749	9.11134	14.4052	0.5208
16	369.032	147.770	7.48032	19.3247	0.8196
17	39.1143	228.175	4.78438	18.4564	0.5244
18	41.5779	230.868	3.03890	18.8300	0.5240
19	369.032	232.409	1.92049	12.5936	0.8238
20	69.1237	240.287	6.49145	21.6635	0.6739
21	127.649	248.995	7.08500	19.8829	0.8241
22	93.4562	342.993	0.61395	8.82842	0.5223
23	391.436	346.893	0.98022	21.0803	0.5310
24	278.841	356.783	9.12014	0.15729	0.825
25	346.176	366.738	2.37605	18.2788	0.8231
26	480.779	381.207	0.73486	17.0009	0.5208
27	49.1686	429.796	2.76290	22.4789	0.5225
28	68.3757	435.784	1.23551	18.0509	0.675
29	64.7599	439.941	4.40791	17.168	0.5205
30	128.634	490.707	4.35889	5.11024	0.8233

Below we present the fall decision variation for the 30 frames. The red zone represents the decision values greater than 0.8 which indicates fall
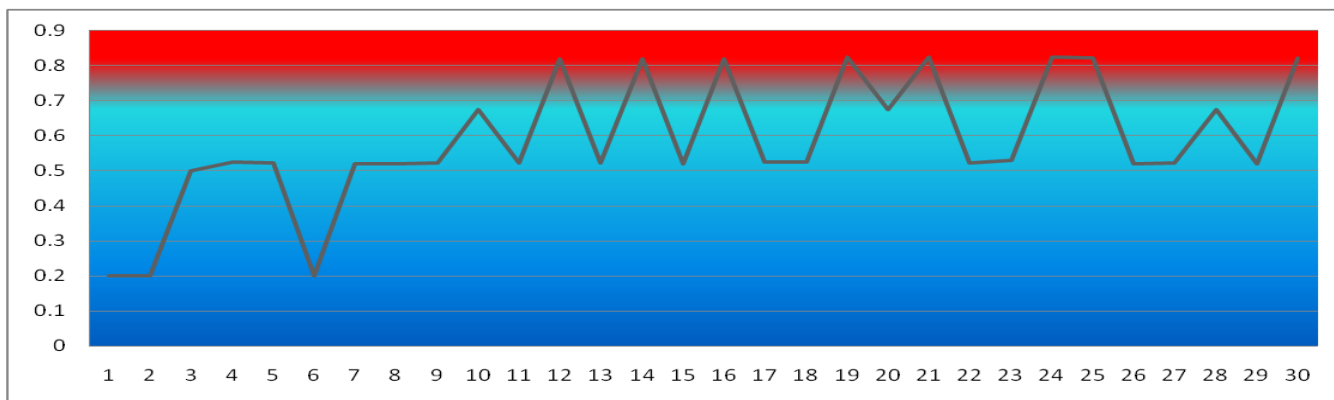
Figure 11 Fall Decision

5. CONCLUSION

The detection of a fall is one of the mst promising fields in the medical domain. The public health services institutions are interested in finding new techniques that reduce costs of taking care of the elderly. Therefore, the demand on services oriented to help elderly people live comfortably in their homes is increasing.

The system includes features extraction and classification of events. We extracted four types of image data: center of mass, aspect ratio, body speed and face speed then we have used Fuzzy Logic classifier to tabulate these events according to two decisions: FALL or NON-FALL. Future work will be based on combining image and non-image data to decide whether a fall happened or not.

The experimental results proved that we could clearly discriminate a falling event from a non-fall situation by combining the different parameters. Our target from this system is to give confidence to elderly in their daily life. Such matter could be achieved in a precise detection for the fall. Elderly people do not want false detections neither they want falls not detected. The above results with sensitivity of 92.7% and specificity of 99.1% show that our system meets the requirements and can be followed by other future works in the same domain.

6. REFERENCES

[1] Lin,C.-W.,et al.,"Compressed-Domain Fall Incident Detection for Intelligent Home Surveillance". Proceedings of IEEE International Symposium on Circuits and Systems, ISCAS 2005, 2005:p.2781-3784.

[2] S. Brownsell and M. Hawley. Automatic fall detectors and the fear of falling. Journal of telemedicine and telecare, 10(5):262, 2004.

[3] N. Noury, A. Fleury, P. Rumeau, A.K. Bourke, G. O. Laighin, V. Rialle, J.E. Lundy, "Fall Detection –

[4] Sixsmith A, Johnson N. "Smart sensor to detect the falls of the elderly", IEEE Pervasive Computing, vol. 3, no. 2, pp. 42-47, April-June 2004.

[5] N. Noury, T. Hervé, V. Rialle, G. Virone, E. Mercier. "Monitoring behavior in home using a smart fall sensor and position sensors". In IEEE-EMBS. Microtechnologies in Medicine &Biology", Lyon-France,Oct 2000; 607-610.

[6] G. Coyle et al." Home telecare for the elderly "Journ. Of telemedicine and telecare 1995, 1,pp. 1183-1184.

[7] G. Williams et al. "A smart fall and activity monitor for telecare application ". Int. Conf. IEEEEMBS, Hong Kong, 1998, pp.1151- 1154 Conference on, 13(16):493 - 498, 2008.

[8] Yamaguchi. "Monitoring behavior in home using positioning sensors" Int. Conf. IEEEEMBS, Hong-Kong, 1998; 1977-79.

[9] J. Yang, S. Wang, N. Chen, X. Chen, and P. Shi. Wearable accelerometer based extendable activity recognition system. IEEE International Conference on Robotics and Automation (ICRA), pages 3641-3647,May 2010.

[10] B. T̈oreyin, Y. Dedeoglu, and A. Ç etin. HMM based Falling person detection using both audio and video. In IEEE International Workshop on Human- Computer Interaction, Beijing, China, 2005.

[11] F. Jabloun, A. E. Cetin, and E. Erzin. Teager energy based feature parameters for speech recognition in car noise. Signal Processing Letters, IEEE,, 6(10):259-261, Oct 1999.

[12] S.-G.Miaou, P.-H. Sung, and C.-Y. Huang, "A Customized Human Fall Detection System Using Omni-Camera Images and Personal Information" Proc of Distributed Diagnosis and Home Healthcare(D2H2) Conference, 2006.

[13] Arie Hans Nasution and S. Emmanuel; Intelligent Video Surveillance for Monitoring Elderly in Home Environments, International Workshop on Multimedia Signal Processing (MMSP), Greece, October 2007.

[14] R. Cucchiara, A.Pratti, and R.Vezani, "An Intelligent Surveillance System for Dangerous Situation Detection in home Environments" , in Intelligenza artificable, vol.1, n.1, pp. 11-15, 2004.

[15] Nait CH, McKenna SJ. "Activity summarisation and fall detection in a supportive home environment", Int. Conf. on Pattern Recognition (ICPR), 2004.