ICSP2008 Proceedings

An Eigenspace-Based Approach for Human Fall Detection using Integrated Time Motion Image and Neural Network

Homa Foroughi[1], Aabed Naseri[2], Alireza Saberi[2], Hadi Sadoghi Yazdi[1]
[1]Department of Computer Engineering, Ferdowsi University of Mashhad, Iran
[2]Department of Electrical and Computer Engineering, Concordia University, Canada
ho_fo46@stu-mail.um.ac.ir, anaseri1@ece.concordia.ca, saberi_alireza@ece.concordia.ca, sadoghi@sttu.ac.ir

Abstract

Falls are a major health hazard for the elderly and a serious obstacle for independent living. Since falling causes dramatic physical-psychological consequences, development of intelligent video surveillance systems is so important due to providing safe environments. To this end, this paper proposes a novel approach for human fall detection based on combination of integrated time motion images and eigenspace technique. Integrated Time Motion Image (ITMI) is a type of spatio-temporal database that includes motion and time of motion occurrence. Applying eigenspace technique to ITMIs leads in extracting eigen-motion and finally MLP Neural Network is used for precise classification of motions and determination of a fall event. Unlike existent fall detection systems only deal with limited movement patterns, we considered wide range of motions consisting normal daily life activities, abnormal behaviors and also unusual events. Reliable recognition rate of experimental results underlines satisfactory performance of our system.

1. Introduction

Fall in the elderly is a major public health problem and may lead to injury, restricted activities, fear or death. A fall incident not also causes many disabling fractures but also has dramatic psychological consequences that reduce the independence of the person. It is shown in [3] that 28-34% elderly people in the community experience at least one fall every year, and 40-60% of the falls lead to injury. The early detections and recording of fall incidents can help the elderly to obtain in-time medical treatments as well as help identify reasons of incidents while sustaining a fall. It was established that the earlier the fall is reported, the lower is the rate of morbidity-morality [1]. Thus with the population growing older and increasing number of people living alone, supportive home environments able to automatically monitor human activities are likely to widespread due to their promising ability of helping elderly people. Nowadays, the usual solution to detect falls is wearable sensors. These autonomous sensors are usually attached under the armpit, around the wrist, behind the ear's lobe or at the waist. They have gyroscopes embedded in them and detect velocity or accelerate exceed a specific threshold, vertical posture toward lying posture, and also absence of movement after fall. However the problem of such detectors is that older people often forget to wear them. Moreover in the case of noncontact sensors, the main problem is they provide fairly crude data that's difficult to interpret. Whereas computer vision systems offer a promising solution to overcome these drawbacks.

In this paper we present a novel visual fall detection system that tries to extract some considerable features from video sequences of movement patterns to detect falls. The reminder of the paper is organized as follows: in section 2 we briefly review some existing vision-based fall detection systems, in section 3 our proposed system is described in more details. Experimental results are represented in section 4 and finally we conclude in section 5 and propose some directions for future work.

2. Related Work

Recently some research has been done to detect falls using image processing techniques. A simple method was used in [5], [6] based on analyzing aspect ratio of the moving object's bounding box. The works in [2], [7] used the normalized vertical and horizontal projection of segmented object as feature vectors. To overcome occluding objects problem, some researchers have mounted the camera on the ceiling: Lee [8] detected a fall by analyzing the shape and 2D velocity

978-1-4244-2179-4/08/$25.00 ©2008 IEEE

of the person. Nait-Charif [9] tracked the person using an ellipse and inferring falling incident when target person is detected as inactive outside normal zones of inactivity like chairs or sofas. Rougier [4] used wall-mounted cameras to cover large areas and falls were detected using human shape variation. Despite the considerable achievements that has accomplished on this field in the recent years, there are some significant factors that limit the use of current systems: Most of current systems [6], [7], [9], [10] are unable to differentiate between real fall incident and an event when the person is simply lying or sitting down abruptly. This mistake is made due to either inappropriate feature extraction or poor classification. Besides, existing fall detection systems tend to deal with restricted range of movement patterns and fall incident are usually detected in contrast with limited normal scenarios like walking. To overcome these shortcomings this paper proposes a novel method, which aims not only to detect and record fall events, but also other postures in a home environment with a considerable recognition rate.

3. Proposed System Overview

Overview of system is shown in Figure1.

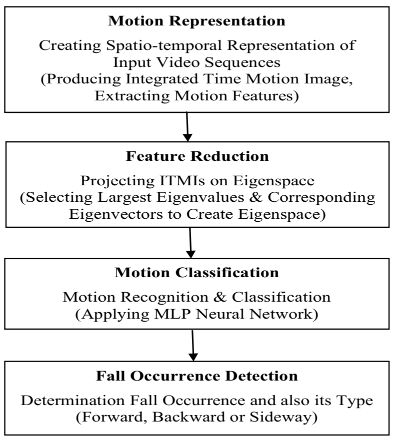


Figure 1 Proposed System Flow

Since we are interested in analyzing the motion occurring in a given window of time, we need a method that allows us to capture and represent motion directly from the video sequence. While such static representations are functions of the observed motion parameters at the corresponding spatial image location in the video sequence, in this paper we propose a spatiotemporal representation that stores occurrence time of each motion with emphasizing at final action. Based on these observations, we extract some motion information from the video sequence. Although, this motion information can be used directly in motion classification, we used eigenspace technique for feature reduction and finally reduced feature vectors are fed to the MLP Neural Network classifier that can deal with noisy data robustly.

4. Technical Details

4.1. Motion Features

The recording of human actions usually needs very large amounts of digital storage space and it is time consuming to browse the whole video to find the required information. Therefore, several motion features have been proposed to compact the whole motion sequence into one image to represent the motion. The most popular ones are the MHI and MEI [12]. The basic idea is to construct an image that can be matched against stored representations of known movements. This image is used as a temporal template. These motion features have the same size as the frame of the video, but they maintain the motion information within them.

4.1.1 MHI

A MHI is a kind of temporal template that's constructed by successively layering selected image regions over time using a simple update rule. Indeed it is the weighted sum of past successive images and the weights decay as time lapses. Normally, the MHI; $H_\tau(x,y,t)$ at time *t* and location *(x,y)* is defined by the following equation:

$$H_\tau(x,y,t)=\begin{cases}\tau & \textit{if } D(x,y,t)=1\\ Max\{0,H_\tau(x,y,t-1)-1\} & \textit{Otherwise}\end{cases} \qquad (1)$$

Where *D(x,y,t)* is a binary image obtained from subtraction of frames, and τ is the maximum duration a motion is stored. In general, τ is chosen as constant 255. So the most recent motion will generate the brightest MHI pixels and the intensity in a MHI encodes how the motion has occurred.

4.1.2 MEI

An MHI pixel can have a range of values, whereas the MEI is its binary version. This can easily be computed by thresholding $H_\tau > 0$. In the other words the MEI; $E_\tau(x,y,t)$ at time *t* and location *(x,y)* is:

$$E_\tau(x,y,t)=\bigcup_{i=0}^{\tau-1} D(x,y,t-i) \qquad (2)$$

4.1.3 ITMI

In the MHI, repeated motions in same position in different times give similar results. This is a problem in storing occurrence time of action. In this paper we propose a spatiotemporal template stores occurrence time of motion with emphasizing at final action. Integrated Time Motion Image (ITMI) at time *t* and location *(x,y)* is defined as follows:

$$ITMI_t(x,y) = ITMI_{t-\Delta t}(x,y) + tf(t) \quad (3)$$

Where *f(t)* is defined according to current image sequence $I(x,y,t)$ as follows:

$$f(t) = I(x,y,t) - I(x,y,t-\Delta t) \quad (4)$$

An example of ITMI is demonstrated in Figure 2.

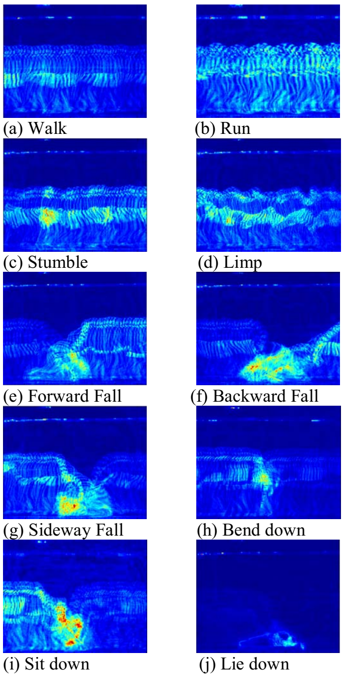


Figure 2 ITMIs of Different Scenarios

4.2. Feature Reduction

Although, motion information can be used directly in motion classification; we used eigenspace technique for feature reduction. Eigenspace-based approaches approximate the behavior images with lower dimensional feature vectors. The main supposition behind this procedure is that the behavior space has a lower dimension than the image space and that the recognition of the behaviors can be performed in this reduced space [11].

Creating eigenspace is done as follows: Pixels of each ITMI can be rearranged in a raster scan manner into a column vector Γ. So for M learning samples we have $\{\Gamma_i \mid i = 1,\ldots,M\}$ that after normalization, all of these samples are collected into a single data matrix A.

$$A = [\Phi_1,\ldots,\Phi_M] \quad (5)$$

Note that A is zero-mean, otherwise normalization, must be done as follows:

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i \quad (6)$$

$$\Phi_i = \Gamma_i - \Psi, \quad i = 1,\ldots,M \quad (7)$$

For the matrix A, the covariance matrix C is defined as follows:

$$C = AA^T \quad (8)$$

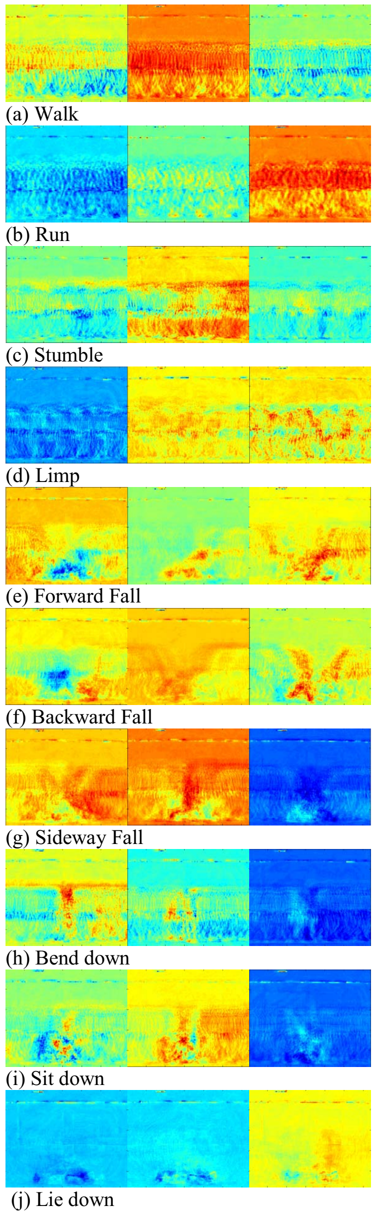


Figure 3 Reduced Eigen-motions by PCA

Let $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_M$ be the M eigenvalues of covariance matrix C arranged in decreasing order of rank, we can then select the first L eigenpairs, i.e. the eigenvectors v_i and eigenvalues λ_i to build a L-dimensional space:

$$B = \{v_i \mid i = 1,\ldots,L\} \quad (9)$$

Indeed the point in which the area under curve of eigenvalues equals to 90% of total area of under curve is selected as L. Note that the first L eigenvectors provide an adequate summary of all the images, which account for 90% of the variation. Eventually the projection matrix is defined as: The ITMI is then projected into set of points in the eigenspace by the following equation:

$$\Omega = B^T (I - \Psi) \tag{10}$$

Figure 3 shows three eigen-motions selected by PCA.

4.3. Motion Classification

For each activity which has to be recognized a solitary neural network has been used. So for recognition of n activities, n neural networks exist. In this study a four-layered MLP network is used for learning eigenvectors. In our network, the commonly used sigmoid function is used as the activation function for nodes in the hidden layer. The MLP utilizes the backpropagation (BP) algorithm for determining suitable weights and biases of the network using supervised training. Fig.4 shows framework of our neural network. All available data are random divided into two subsets, for training and testing respectively. The training and test sets both consist of 60 feature vectors semi-randomly selected from 120 eigen-motions.

5. Experimental Results and Discussion

In order to validate the overall system performance we applied the proposed approach to a set of videos recorded in our lab. The dataset has been collected along two weeks, by considering different light and weather conditions. 24 subjects with different height, weight and genders whose ages ranged from 20 to 30 were asked to participate in the project. We repeated

10 kinds of activities by 5 times in the experimental space and finally 50 video clips were captured and recorded in AVI format with a resolution of 320*240 pixels at 30 frames per second. Figure 4 illustrates examples of each motion. Our dataset for our experiment is composed of video sequences representing three main categories:

a) Normal Daily Activities: Walking, Running, Bending down for catching something and rising up, Sitting down and then standing up and Lying down on the floor.

b) Unusual Behavior (Fall): Although the scenarios of falling are very various, we can categorize them in three classes. Detection of various types of falls is valuable in clinical studies to determine fall incidence and costs associated with the treatment of fall injuries. As most falls occur during intentional movements initiated by the person, they happen mainly in forward or backward: stumbling on an obstacle during walking, backward slip on wet ground. But in some cases the fall occurs sideways, either during a badly controlled sit to stand transfer [1].

c) Abnormal Gait: Subjects were asked to walk in two special unusual ways: Stumbling and Limping. In the first case subjects walked as if they were suffering a balance deficiency such as dizziness or in a manner that they are unable to keep their balance or symmetry. Whereas in the case of limping subjects walked in a style that it seemed they suffer from unequal leg lengths, experiencing pain when walking, muscle weakness, disorders of proprioception, or stiffness of joints.

The experimental results show that the system has a robust recognition rate in detecting occurrence of considered events. Table 1 represents the experimental results. N_a refers to number of actions. N_c is number of correctly detected events, N_f is number of falsely detected events and R is the recognition rate.

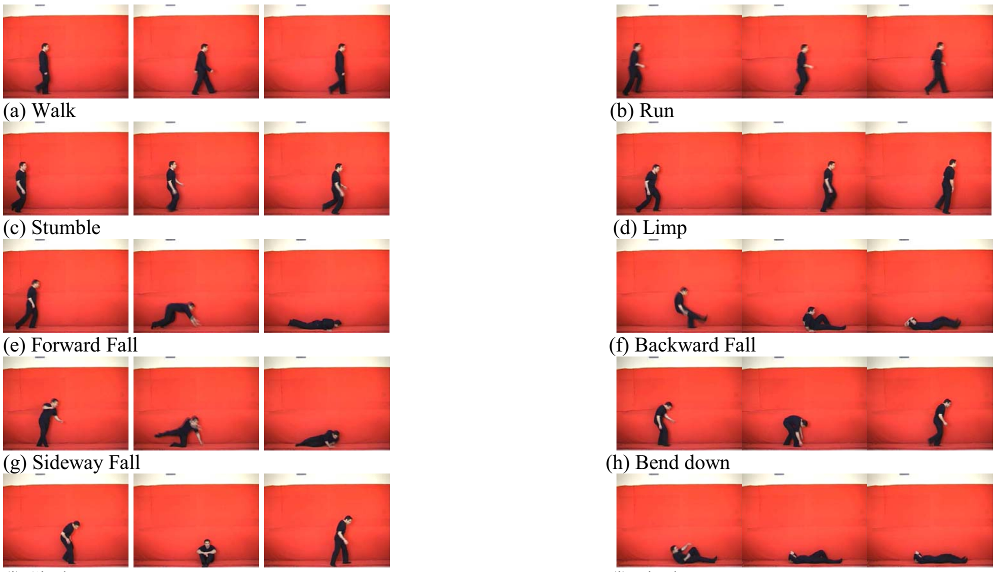
Table 1 Recognition Rate for Various Events

Events	N_a	N_c	N_f	R
Walk	60	55	5	91.66
Run	60	54	6	90.00
Stumble	60	52	8	86.66
Limp	60	50	10	83.33
Forward Fall	60	54	6	90.00
Backward Fall	60	56	4	93.33
Sideway Fall	60	54	6	90.00
Bend down	60	53	7	88.33
Sit down	60	54	6	90.00
Lie down	60	58	2	96.66

Dataset of our experiments is composed of video sequences representing 300 normal activities (walk, run, bend, sit and lie) and 180 simulated falls. Table 2 itemizes all the results obtained with our dataset.

Table 2 Recognition Results

	Detected	Not Detected
Falls	164	16
Lures	274	26



(a) Walk (b) Run (c) Stumble (d) Limp (e) Forward Fall (f) Backward Fall (g) Sideway Fall (h) Bend down (i) Sit down (j) Lie down

Figure 4 Example of Each Motion

6. Conclusions and Future Work

In this paper a novel efficient approach for activity recognition, principally dedicated to fall detection is proposed. The combination of motion and eigenspace technique, gives crucial information on human activities. Our experiments indicate that MLP neural network is suitable classifier human motion recognition. Our fall detection system has proven its robustness on realistic image sequences of ten different normal, abnormal and unusual human movement patterns. Reliable average recognition rate of experimental results (89.99%) underlines satisfactory performance and efficiency of our system. Future works will include the incorporation of multiple elderly monitoring which is able to monitor more than one person in the scene and also be able to handle occlusion. Using additional features is also a subject to be explored in the future work.

References

[1] N Noury, A Fleury, P Rumeau, A Bourke, G Laighin, V Rialle, J Lundy "Fall detection - Principles and Methods" Conf Proc IEEE Eng Med Biol Soc. 2007

[2] 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.

[3] 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.

[4] Caroline Rougier, Jean Meunier, Alain St-Arnaud, Jacqueline Rousseau, "Fall Detection from Human Shape and Motion History Using Video Surveillance" Advanced Information Networking and Applications Workshops, 2007, AINAW '07. pp. 875-880.

[5] B. T¨oreyin, Y. Dedeoglu, and A. C¸ etin. Hmm based falling person detection using both audio and video. In IEEE International Workshop on Human-Computer Interaction, Beijing, China, 2005.

[6] 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

[7] 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

[8] T. Lee and A. Mihailidis. An intelligent emergency response system: preliminary development and testing of automated fall detection. Journal of telemedicine and telecare, 11(4):194–198, 2005.

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

[10] Tao, J., M.Turjo, M.-F. Wong, M.Wang and Y.-P Tan, "Fall Incidents Detection for Intelligent Video Surveillance" , ICICS2005

[11] Takehito OGATA, JooKooi TAN Seiji ISHIKAWA, "High-Speed Human Motion Recognition Based on a Motion History Image and an Eigenspace", IEICE - Transactions on Information and Systems,Volume E89-D , Issue 1 (January 2006), Pages 281-289.

[12] Bobick, A.F., Davis, J.W.: The recognition of human movement using temporal templates. IEEE Trans. Pattern Anal. Mach. Intell.23(3), 257–267 (2001)