

# Fall detection using History Triple Features

Georgios Goudelis  
ggoudelis@image.ntua.gr

Georgios Tsatiris  
gtsatiris@image.ntua.gr

Konstantinos Karpouzis  
kkarpou@cs.ntua.gr

Stefanos Kollias  
stefanos@cs.ntua.gr \*

## ABSTRACT

Accurate identification and timely handling of involuntary events, such as falls, plays a crucial part in effective assistive environment systems. Fall detection, in particular, is quite critical, especially in households of lonely elderly people. However, the task of visually identifying a fall is challenging as there is a variety of daily activities that can be mistakenly characterized as falls. To tackle this issue, various feature extraction methods that aim to effectively distinguish unintentional falls from other everyday activities have been proposed. In this study, we examine the capability of the *History Triple Features* technique based on Trace transform, to provide noise robust and invariant to different variations features for the spatiotemporal representation of fall occurrences. The aim is to effectively detect falls among other household-related activities that usually take place indoors. For the evaluation of the algorithm the video sequences from two realistic fall detection datasets of different nature have been used. One is constructed using a ceiling mounted depth camera and the other is constructed using an RGB camera placed on arbitrary positions in different rooms. After forming the feature vectors, we train a support vector machine using a radial basis function kernel. Results show a very good response of the algorithm achieving 100% on both datasets indicating the suitability of the technique to the specific task.

## 1. INTRODUCTION

The demographic changes in society with growing numbers of elderly subjects have already raised the alarm especially in countries of the western world. Elderly care or eldercare is nowadays a broad term that refers to the fulfilment of the special needs and requirements that are unique to senior citizens. Nursing homes, hospice care, home care, assisted living, adult day care, are some of the different services that

refer to the specific term. One of the main points on elderly care that drives research on the specific topic, is the need for independence and ageing with dignity.

Traditionally elderly care has been the responsibility of family members and was provided within the extended family home. Increasingly in modern societies, elderly care is now being provided by state or charitable institutions [24]. The reason for this change include decreasing family size, the greater life expectancy of elderly people, the geographical dispersion of families, and the tendency for women to be educated and work outside home [24]. Although these changes have affected European and North American countries first, they are now increasingly affecting Asian countries as well [22]. Based on these facts, the effort is focused on the autonomy of elderly people which tend to live alone or are not able to indulge themselves in the luxury of an attendance person. According to the Center for Research and Prevention of Injuries report, fall-caused injuries of elderly people in are five times as frequent as other injuries, a fact that reduces considerably their mobility and independence [1].

As falls are a major public health problem among older people, the number of systems aimed at detecting them has increased dramatically over recent years. According to the World Health Organization [16] approximately 28-35% of people aged 65 and over fall each year increasing to 32-42% for those over 70 years of age. The frequency of falls increases with age and frailty level. In fact, falls exponentially increase with age-related biological changes, which is leading to a high incidence of falls and fall related injuries in the ageing societies. If preventive measures are not taken in the immediate future, the number of injuries caused by falls is projected to be a 100% higher in 2030 [7]. In this context, fall detectors are being actively investigated.

Fall detectors could be simply divided into two main categories: the wearable sensor based and the vision based techniques. The first category is based on usually wearable devices such as accelerometers and gyroscopes, or on smartphones that usually contain this kind of sensors and are mainly carried continuously by subjects. The second category is based on 2D or 3D cameras and involve image analysis and pattern recognition techniques of high computational complexity, but present the advantage that a continually carryable device is not required. Researches in [13] divide fall detectors in three main categories: wearable device based, ambience sensor based and camera (vision) based,

---

\*All authors are with the Image, Video and Multimedia Systems Laboratory, National Technical University of Athens, 9 Heron Politechniou str., 15780, Athens, Greece.

while, from a different perspective, researchers in [17] detect three main categories that include methods measuring acceleration, methods that measure acceleration combined with other methods, and methods that do not measure acceleration.

Two of the first attempts to provide a general overview of the fall detection status are given in [15] and in [17]. However, as the advancement of technology on this area is rapidly growing, the reviews are mostly outdated. A newer and more extensive literature review is provided in [7] including comparisons among various kinds of studies. The article aims to serve as a reference for both clinical and biomedical engineers planning or conducting investigations on the field. The authors are mostly trying to identify challenges regarding performance under real-life conditions and the current trends on the field. A more detailed discussion of each paper is provided in [13] but it lacks references to new trends on the topic (such as smartphone based techniques etc.).

Our research focuses on non-wearable sensor solutions and more specifically to a computer vision based approach. In the same direction, researchers in [10], placed the camera on the ceiling and analyzed the segmented silhouette and the 2D velocity of the elderly. The determination of a fall is achieved by an experienced thresholding. Authors in [25], in order to distinguish sitting from falling, added the extra information of noise. However the system cannot be very robust as most of the environments where such a system can be applied are noisy. In [19], the approach presented is based on a combination of motion history and human shape variation. To cover large areas, wall cameras have been mounted and the final decision is made by thresholding the extracted features. In [4] a technique based on a combination of integrated time motion images and eigenspace is presented. The classification of every day motions and fall events, is achieved by extracting eigen-motion and by applying multi-class Support Vector Machines.

Other techniques [26], [20] use shape-based fall detectors and separate the human silhouettes by a regular bounding box or an ellipse, and extract geometric attributes such as aspect ratio, orientation or edge points [21]. However, this approach lacks robustness and generalization of the application as it is dependent on an accurate extraction of the human silhouette and the geometrical transformations that may occur due to distance and the position of the subject from the camera. In a more recent work in [11], the method presented combines two computer vision techniques shape-based fall characterization and a learning-based classifier to distinguish falls from other daily actions while in [3], the human body is represented with ellipse fitting and the silhouette motion is modelled by an integrated normalized motion energy image. The required features for the classification of the different postures of the subject are extracted from the shape deformation quantified from the fitted silhouettes.

Recently, due to the efficiency of 3D information on partial occlusion and viewpoint problems, a number of relevant works has been published. In [12], the method presented is velocity based and takes into account the contraction or expansion of the width, height and depth of a 3D bounding box. The method does not require pre-knowledge of the

scene (i.e. floor), as the set of detected actions are adequate to complete the process of fall detection. In another approach [18] the proposed algorithm makes use of the Kinect 3D camera and creates two feature parameters, the orientation of the body and the height information of the spine, using either image or world coordinates. Kinect camera is also used in [14]. The fall detection algorithm used is based on the speed of the head that has previously been detected, the body centroid and their distance to the ground. Incorporating both the body centroid and the head, the algorithm is referred as to be less affected by the centroid fluctuation.

Again using Kinect, authors in [27], propose a statistical method that makes a decision based on information about how the human moved during the last few frames. The algorithm is based on the combination of the proposed features under a Bayesian framework. The main idea is to create a technique that, while it has been trained by data captured from a specific viewpoint, is able to classify falls that have been captured by another one.

In this paper, we propose a novel vision-based technique for automated fall detection. The method is based on a novel feature extraction technique proposed in [6] for human action recognition and it is tested on two new and challenging fall containing datasets. The method involves Trace transform [8] for the construction of a set of invariant features that represent the action sequence and can cope with variations usually appearing in video capturing. This specific method takes advantage of the natural specifications of the Trace transform, to produce noise robust features that are invariant to translation, rotation, scaling and are effective, simple and fast to create. The method is proposed for a monocular vision-based fall detection system.

The rest of the paper is organized as follows. A few words about Trace transform and the theory behind it are given in 2.1. An overview of the proposed system is provided in 2.2. History Triple feature technique is given in 2.3. The experimental procedure and the results are provided in 3 followed by conclusion in 4.

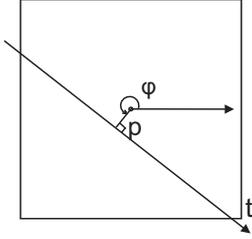
## 2. PROPOSED APPROACH

### 2.1 Trace Transform

Trace transform is a generalization of Radon [2] transform while at the same time Radon builds a sub-case of it. While Radon transform of an image is a 2D representation of the image in the  $(\phi, p)$  coordinate space with the value of the integral of the image computed along the corresponding line, placed at cell  $(\phi, p)$ , Trace calculates functional  $T$  over parameter  $t$  along the line, which is not necessarily the integral. Trace transform is created by tracing an image with straight lines where certain functionals of the image function are calculated. Different transforms having different properties can be produced from the same image. The transform produced is in fact a 2-dimensional function of the  $(\phi, p)$  parameters of each tracing line. Definition of the above parameters for an image Tracing line is given in Figure 1.

### 2.2 Overview of the proposed system

Exploring the capabilities of Trace transform we extended the method based on History Trace Templates initially proposed in [5] creating even more effective features for human



**Figure 1: Definition of the parameters of an image tracing line.**

action recognition [6]. The new features consist of a set of triple features divisions and are invariant to different distortions.

For each video sequence, background and silhouettes are extracted (more details on that are provided in section 3). In this case, using a number of different functionals, a number of different transformations is calculated for each frame. From these transforms, a vector that is composed of a series of invariant features calculated for each frame of one period of an action is produced. The whole sequence is represented by a new vector named History Triple Features (HTFs) and is a set of real numbers containing important discriminant information for fall detection. Principal Component Analysis (PCA) can be used at this point, to reduce vector dimensionality. A more comprehensive description of the specific feature extraction technique is given below, in section 2.3.

### 2.3 Constructing HTFs

In this section we introduce a novel human action representation using features derived from the Trace transform, hereafter simply called *History Triple Features* (HTFs). The Trace transform is a global transform that can be applied to full images. It is known to be able to pick up shape as well as texture characteristics of the object it is used to describe and offers an alternative representation of an image of it [23].

Authors in [8] have proved that using extracted triple features, robust features for the classification of different but very similar to each other image classes (e.g. different kinds of fish) can be produced. In [8] the theory behind the construction of triple features is presented. In the following, we demonstrate the construction of HTFs proposed in [6].

Having the extracted silhouettes, we first transform the silhouette containing image space, to Trace transform space. For each frame of the sequence, a set of Trace transforms is calculated. Following the procedure described in ?? for the extraction of the triple feature, a set of such features is extracted. The ratio of a pair of such features, as it has been shown, may be invariant to different kind of distortions, depending on the functionals used. These functionals may be chosen to be sensitive or relatively insensitive to the possible variations that occur in fall sequences, while maintaining discriminability.

Let  $f(p, \vartheta, t)$  be a human action sequence or a fall. Applying a Trace functional  $T$ , along lines tracing the  $n$  frame referring to  $s_n(p, \theta)$  silhouette, where  $n = 1 \dots N$  and  $N$  is the number of frames, a Trace transform  $\check{g}_n(p, \theta)$  is produced. Applying different  $T$ s to every silhouette  $s_n(p, \theta)$  a set of  $\check{g}_{n_i}(p, \theta)$  transforms is produced. Where  $i = 1 \dots L$  and  $L$  is the the number of transforms one chooses to calculate. For every  $\check{g}_{n_i}(p, \theta)$  a set of  $\Pi_{norm}(F, C)$  normalized triple features is computed.

In a simple way, a triple feature is constructed as follows:  
a) Trace transform is produced by applying a Trace functional  $T$  along lines tracing the image.  
b) The circus function of the image is produced by applying a diametric functional  $P$  along columns of the Trace transform.  
c) The triple feature is finally produced by applying a circus functional  $\Phi$  along the string of number produced in step b. The procedure is illustrated in Figure 2.

Dividing all  $\Pi_{norm}(F, C)$  by each other, a set of independent features is produced. So, the whole action sequence is finally represented by a vector  $\mathbf{v}$  comprised by the set of all triple feature ratios calculated for every frame of the action sequence.

$$\mathbf{v} = (\Pi_{rat_1}, \Pi_{rat_2}, \dots, \Pi_{rat_{g-1}}, \Pi_{rat_g}) \quad (1)$$

where  $\Pi_{rat}$  is the ratio of two normalized triple features and  $g$  the number of calculated ratios.

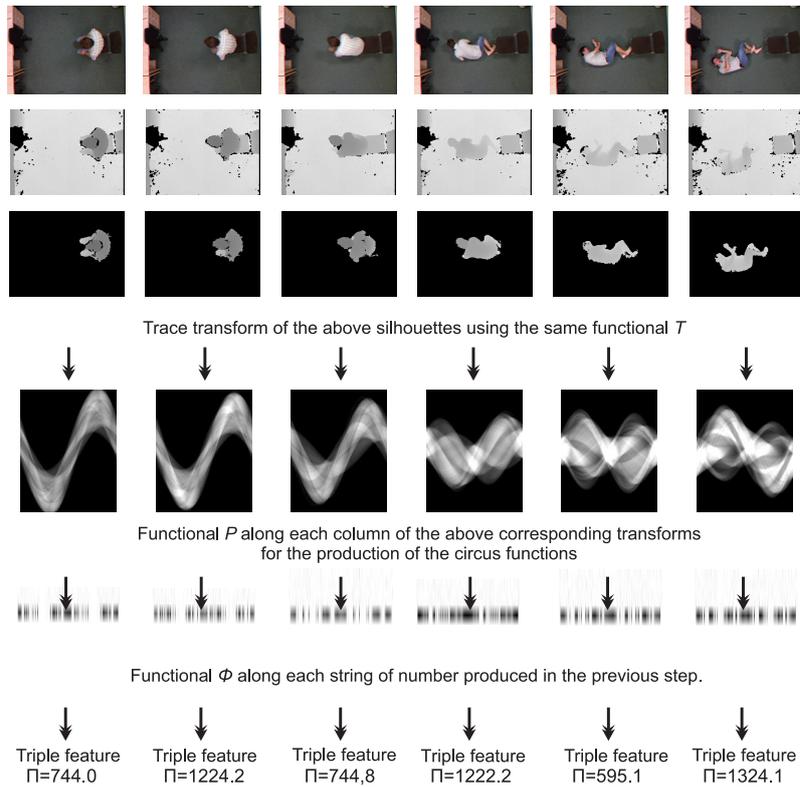
This method allows the construction of many features easily. Supposing that one makes use of 10 functionals for each stage of the construction (e.g. 10  $T$  functionals, 10  $P$  functionals and 10  $\Phi$  functionals) in a 10 frame action video, he may construct  $10 \times 10 \times 10 \times 10 = 10000$  features for one sequence. As mentioned above, these numbers may have not any physical meaning according to human perception, but they may have the required mathematical properties for classification purposes.

Since the discriminatory power of the features constructed will definitely vary, a dimensionality reduction technique could provide a selection of the most discriminant features, while, at the same time, making the problem of classification more tractable. In our scheme the HTF vectors produced as described above, become subject to PCA in order to determine an appropriate subspace that is suitable for classification. In practice we keep only a subset of the initial HTF vector that contains the most discriminant of the calculated feature capable to efficiently describe the entire action sequence. However, the performance of the algorithm without using PCA is also referred in section 3.2.

## 3. EXPERIMENTAL PROCEDURE AND RESULTS

### 3.1 Datasets and evaluation protocols

In this section we will present the experimental results and the procedure followed for the evaluation of the proposed method. Generally, there are only few available datasets

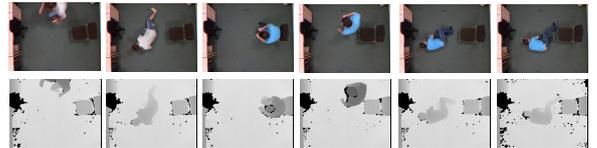


**Figure 2: Triple features extraction for a fall taken from UR fall detection dataset.**

dedicated to fall detection as most of the published techniques have been tested on their own datasets. However, in order to have a benchmark, we have evaluated our technique in two new publicly available datasets: The UR Fall Detection [9] and the Le2i Fall detection datasets [1].

UR Fall dataset contains 60 sequences recorded with 2 Microsoft Kinect cameras and corresponding accelerometric data. Sensor data was collected using PS Move (60Hz) and x-IMU (256Hz) devices. The dataset contains sequences of depth and RGB images for two differently mounted cameras (parallel to the floor and ceiling mounted, respectively), synchronization data, and raw accelerometer data. Each video stream, both RGB and depth, is stored in separate folders in the form of png image sequences. From the specific dataset we have used the depth data provided by the ceiling mounted camera following the experimental protocol given by authors in [9]. Frame samples taken from UR fall dataset are provided in figure 3.

The Le2i Fall dataset has been captured in realistic video surveillance settings using a single RGB camera. The frame rate is 25 frames/s and the resolution is 320x240 pixels. The video data illustrates the main difficulties of realistic video sequences that can be found at an elderly home environment, as well as in a simple office room. Our video sequences contain variable illumination, and typical difficulties like occlusions or cluttered and textured background. The actors performed various normal daily activities and falls. The dataset contains 130 annotated videos, with extra information representing the ground-truth of the fall posi-



**Figure 3: Frame samples taken from the UR fall dataset for two falls. Upper row illustrates the RGB samples while the lower row provides the corresponding depth images.**

tion in the image sequence. The database provides different locations for testing and training, while authors in [1] have defined several protocols for the evaluation of their method. Working with the specific dataset, we have followed the protocol P1 given in the above paper, where training and test sets are built with videos from "Home" and "Coffee room" subsets. Samples from Le2i dataset are provided in Figure 4.

As described above, for the extraction of the features vectors using our approach, human silhouette extraction has been priorly achieved. For the UR dataset, human silhouette extraction was handled by computing the value difference between every depth pixel in a particular frame and its corresponding depth pixel in a precalculated reference frame. The reference frame was calculated by computing the median value of every depth pixel in a sliding window of 9 frames, in a total of 80 frames portraying an empty scene (without any human presence). Then, the mean value of



**Figure 4: Frame samples taken from the Le2i fall dataset. Upper row illustrates samples from daily actions in "Coffee room" while lower row provides samples from a fall occurred in "Home".**

every median pixel value was calculated, forming the final reference frame and eliminating a considerable amount of noise generated by the depth camera.

Human presence in a particular frame can be detected when the difference between depth pixel values of that frame and the reference, exceeds a predefined threshold. In order to add robustness in this method, we used a total of four thresholds. First, for a pixel value to be valid, we required it to be between 1100 and 3620 millimeters. It should be reminded that this represents distance from a ceiling mounted depth camera. Then, for the difference between the frame under inspection and the reference to be considered significant enough for determining human movement inside it, it was required to be between 50 and 2200 millimeters. These values were found to offer maximum tolerance against random noise.

The Le2i dataset was handled in a different way, given the fact that it consisted of RGB video files of low resolution. Furthermore, light conditions in most of the cases (especially in the "Home" subset) rendered the use of the difference between frames unreliable. In order to detect the human presence, we utilized the background-foreground segmentation approach proposed by Zivkovic in [28] and [29], which is based in the idea of performing a subtraction between the current frame and a background model. This model is constantly updated in a per-pixel fashion, using a gaussian mixture-based approach, in order to better adapt in scene changes and contain what is considered the static part of the scene.

Despite the fact that both methods mentioned above performed relatively well, effectively separating a human silhouette from the rest of the image, the presence of noisy data could not be fully eliminated. However, the feature extraction method described in this paper proved its robustness against such noise. In our experiments, the sequences have been downsampled to the spatial resolution of 320\*240 pixels and have a temporal length of 26 and 12 frames for the UR and the Le2i dataset respectively. The training examples were constructed by manually segmenting and aligning the available sequences.

At this point we should mention that there is no unified standard usually followed for the evaluation of fall detection algorithms. In our experiments, the leave-one-person-out cross validation approach was used to evaluate perfor-

mance. This specific protocol was chosen due to its popularity among researchers. It also reconstructs the real life application needs, in the closest way. Thus, the physical dynamic behavior of an unknown subject is captured by a fall detection system and thereafter processed and compared against a pre-recorded set of data that have previously trained it. The final decision is made based on the relativity of the examined sequence, with one of the data that comprise the training set, according to the system's set rule. Accordingly, the above protocol uses one person's samples for testing and the rest of the dataset is used for training. The procedure is repeated  $N$  times where  $N$  is the number of subjects within the dataset. Performance is reported as the average accuracy of  $N$  iterations.

For the construction of feature vectors from each sequence of frames, a total of seven functionals were used, according to the methodology presented in [6]. This feature extraction procedure was followed for both the UR and the Le2i datasets. Feature vectors were then used to train a SVM, using the above referred protocol. For the mapping of the training data into kernel space, we used the Gaussian Radial Basis Function kernel, with a variety of scaling factor values  $\sigma$ . We opted for soft margins and experimented with various values of the box constraint  $C$ .

Initially, the complete training-testing pipeline of the protocol was performed using the full feature vectors as input. Afterwards, we performed Principal Component Analysis on the each feature vector class and repeated the experimentation process. This was also repeated for a variety of vector lengths.

## 3.2 Results

Our method is shown to achieve 100% accuracy on both datasets, equal to the 0% error given by the method of Kepski and Kwlek [9] on the UR fall dataset and comparable to the 99.6% accuracy of the method in the work of Charfi et al. [1] which accompanies the Le2i database. However, these methods seem to be quite domain specific and rely on specific characteristics of the human silhouette shape and bounding box, with a strong relation to the position of the camera. The method proposed in this paper is generalized to produce features from the complete transition of the human shape between frames, as well as omit characteristics in the scene that have nothing to offer to the separability between the two classes. Results of our experiments on the UR and Le2i datasets respectively can be found in Tables 1, 2 and 3.

**Table 1: SVM performance on the UR Fall Detection Dataset (binary silhouettes)**

	using PCA	w/o PCA
Best score	100%	87.76%
Sigma	2	10
C	0.0001	0.0001
Vector length	8	1014

## 4. CONCLUSION

In this paper we examined the suitability of the History Triple Feature technique for the problem of fall detection.

**Table 2: SVM performance on the UR Fall Detection Dataset (depth silhouettes)**

	using PCA	w/o PCA
Best score	100%	87.76%
Sigma	3	10
C	0.0001	0.0001
Vector length	7	1014

**Table 3: SVM performance on the Le2i Fall Detection Dataset**

	using PCA	w/o PCA
Best score	100%	96.6%
Sigma	3	10
C	0.0001	10
Vector length	47	468

The method is based on the so-called Trace Transform and it has been shown to produce translation, rotation and scale invariant features that can cope with common noise distortions, camera placement variations and other noise inducing factors. Experimentation proved that, using HTFs, we were able to achieve 100% accuracy on fall detection problems on two different and challenging datasets of different nature and settings. Furthermore, we believe that the proposed pipeline can be easily generalized to more complicated fall detection tasks, without any dependence on the nature of the dataset under inspection. In that direction, future plans on this method include more extensive experimentation on intentional - unintentional fall discrimination, as well as deeper classification of other common activities that could lead to or follow a fall, towards better modeling of the human behavior in an elderly household. Further attempts will also be made into the integration of the proposed methodology in a real-time fall recognition module, as part of an assistive environment system.

## 5. ACKNOWLEDGMENTS

This work was funded by Greek national funds through the Operational Program \*Educational and Lifelong Learning\* of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALES. Investing in knowledge society through the European Social Fund.

## 6. REFERENCES

- [1] I. Charfi, J. Miteran, J. Dubois, M. Atri, and R. Tourki. Definition and performance evaluation of a robust svm based fall detection solution. In *SITIS'12*, pages 218–224, 2012.
- [2] S. R. Deans. *The Radon Transform and Some of Its Applications*. Krieger Publishing Company, 1983.
- [3] W. Feng, R. Liu, and M. Zhu. Fall detection for elderly person care in a vision-based home surveillance environment using a monocular camera. *Signal, Image and Video Processing*, 8(6):1129–1138, 2014.
- [4] H. Foroughi, H. Yazdi, H. Pourreza, and M. Javidi. An eigenspace-based approach for human fall detection using integrated time motion image and multi-class support vector machine. In *Intelligent Computer Communication and Processing, 2008. ICCP 2008. 4th International Conference on*, pages 83–90, Aug 2008.
- [5] G. Goudelis, K. Karpouzis, and S. Kollias. Robust human action recognition using history trace templates. 12th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Delft, The Netherlands, 13-15 April., 2011.
- [6] G. Goudelis, K. Karpouzis, and S. D. Kollias. Exploring trace transform for robust human action recognition. *Pattern Recognition*, 46(12):3238–3248, 2013.
- [7] R. Igual, C. Medrano, and I. Plaza. Challenges, issues and trends in fall detection systems. *BioMedical Engineering OnLine*, 12(1):66, 2013.
- [8] A. Kadyrov and M. Petrou. The Trace transform and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23:811–828, August 2001.
- [9] B. Kepski. M., Kwolek. Fall detection using ceiling-mounted 3d depth camera. In *Proc. 9th Int. Conf. on Computer Vision Theory and Applications (VISAPP)*, pages vol. 2, 640–647, 2014.
- [10] T. Lee and M. A. An intelligent emergency response system: preliminary development and testing of automated fall detection. *Journal of telemedicine and telecare*, 11(4):1947198, 2005.
- [11] X. Ma, H. Wang, B. Xue, M. Zhou, B. Ji, and Y. Li. Depth-based human fall detection via shape features and improved extreme learning machine. *IEEE J. Biomedical and Health Informatics*, 18(6):1915–1922, 2014.
- [12] G. Mastorakis and D. Makris. Fall detection system using Kinect’s infrared sensor. *Journal of Real-Time Image Processing*, 9(4):635–646, 2014.
- [13] M. Mubashir, L. Shao, and L. Seed. A survey on fall detection: Principles and approaches. *Neurocomputing*, 100:144–152, 2012.
- [14] A. Nghiem, E. Auvinet, and J. Meunier. Head detection using Kinect camera and its application to fall detection. In *11th International Conference on Information Science, Signal Processing and their Applications, ISSPA, Montreal, QC, Canada, July 2-5*, pages 164–169, 2012.
- [15] N. Noury, P. Rumeau, A. Bourke, G. Olaighin, and J. Lundy. A proposal for the classification and evaluation of fall detectors. *IRBM journal of Alliance for engineering in Biology an Medicine*, 26(6):340–349, 2008.
- [16] W. H. Organization. on falls prevention in older age who library cataloguing-in-publication data, 2007.
- [17] J. T. Perry, S. Kellog, S. M. Vaidya, J.-H. Youn, H. Ali, and H. Sharif. Survey and evaluation of real-time fall detection approaches. In *High-Capacity Optical Networks and Enabling Technologies (HONET), 2009 6th International Symposium on*, pages 158–164. IEEE, Dec. 2009.
- [18] R. Planinc and M. Kampel. Introducing the use of depth data for fall detection. *Personal Ubiquitous Comput.*, 17(6):1063–1072, Aug. 2013.
- [19] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau. Fall detection from human shape and motion history using video surveillance. In *AINA Workshops (2)*, pages 875–880. IEEE Computer Society, 2007.

- [20] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau. Fall detection from human shape and motion history using video surveillance. In *21st International Conference on Advanced Information Networking and Applications (AINA 2007), Workshops Proceedings, Volume 2, May 21-23, 2007, Niagara Falls, Canada*, pages 875–880, 2007.
- [21] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau. Robust video surveillance for fall detection based on human shape deformation. *IEEE Trans. Circuits Syst. Video Techn.*, 21(5):611–622, 2011.
- [22] H. Shirlena, T. L. Leng, and T. Mika. Transnational mobilities for care: Rethinking the dynamics of care in asia. *Global Networks*, 12(2):129–134, 2012.
- [23] S. Srisuk, M. Petrou, W. Kurutach, and A. Kadyrov. A face authentication system using the Trace transform. *Pattern Anal. Appl.*, 8(1):50–61, Sept. 2005.
- [24] G. Ting and J. Woo. Elder care: is legislation of family responsibility the solution. *Asian J Gerontol Geriatr*, 4:72–75, 2009.
- [25] B. U. Töreyn, Y. Dedeoğlu, and A. E. Çetin. Hmm based falling person detection using both audio and video. In *Proceedings of the 2005 International Conference on Computer Vision in Human-Computer Interaction, ICCV'05*, pages 211–220, Berlin, Heidelberg, 2005. Springer-Verlag.
- [26] C. M. V. Vishwakarma and S. Sural. Automatic detection of human fall in video. In *in Proc. Int. Conf. Pattern Recogn. Mach. Intell.*, page 616?623, 2007.
- [27] Z. Zhang, W. Liu, V. Metsis, and V. Athitsos. A viewpoint-independent statistical method for fall detection, 2012.
- [28] Z. Zivkovic. Improved adaptive gaussian mixture model for background subtraction. In *Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 2 - Volume 02, ICPR '04*, pages 28–31, 2004.
- [29] Z. Zivkovic and F. van der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recogn. Lett.*, 27(7):773–780, May 2006.