

# A REAL-TIME, MULTI-SPACE INCIDENT DETECTION SYSTEM FOR INDOOR ENVIRONMENTS

DIMITRA TRIANTAFYLLOU, STELIOS KRINIDIS,  
DIMOSTHENHS IOANNIDIS & DIMITRIOS TZOVARAS  
Information Technology Institute, Centre for Research and Technology Hellas, Greece.

## ABSTRACT

A real-time, multi-camera multi-space incident detection system for indoor environments is proposed. The system focuses on industrial environments and recognizes incidents such as human falls, intrusions to restricted areas and collisions. Building Information Models (BIM) provide all the necessary information for the set up of the system. Furthermore, once an incident is detected an alarm notifying of its type and exact location on the building is triggered so that the appropriate coping mechanism can hasten promptly to the point of the incident. Based on building information about each area particularities and special conditions, e.g. moving machinery or chemical agents, the appropriate coping measurements can be taken. Moreover, the use of the incident detection system for a long period can provide statistics about the inherent dangers of the building. Experimental results seem very promising.

*Keywords: collision, fall, incident detection, industrial environment, intrusion.*

## 1 INTRODUCTION

The fourth industrial revolution, or else Industry 4.0, denotes a new manufacturing reality towards ‘smart’ factories. Such a factory could not omit the domain of occupational health and safety that until now mainly focused on the prevention of accidents by the use of area restrictions or appropriate equipment (e.g. helmets, gloves).

In this paper, following the vision of ‘smart’ industries, a real-time, multi-space incident detection system for indoor environments is introduced. The proposed system, which constitutes an extension of previous work described in Ref. [1], detects incidents such as falls, collisions and intrusions. The geometrical information of the BIM is utilized for the set-up and calibration of the sensors in the building permitting in this way to achieve multi-space detection. Once an incident is detected the appropriate coping mechanism is notified of the incident type and exact location on the building’s architectural map so that it can immediately hasten to the point of the event and handle the situation. Combined with the knowledge of each area’s particularities, which can be extracted from BIM, e.g. chemical agents, moving machinery, the appropriate coping measures can be taken. Moreover, statistics about the incidents occurrence in each area of the building can be easily made if the system is running for a long period leading to the appropriate readjustments – measures concerning the building. At this point, it should be mentioned that all utilized cameras are depth sensors respecting the personnel’s legal and ethical rights.

The detection, recognition and modeling of the aforementioned incidents is accomplished through a Hidden Markov Model procedure that utilizes features that characterize each event, e.g. the vertical velocity of a falling human or the relative distance between two moving items that will collide.

To sum up, the main contributions of the proposed paper are:

1. The introduction of a real-time, multi-space incident detection system that produces alarms notifying for the type and exact location of the event on the building’s architectural map so that the appropriate coping mechanism can be immediately triggered.

- This information combined with knowledge on each building's area particularities can lead to the appropriate decision of ways to handle the incident.
2. The introduction of a detection system that apart from falls and intrusions, which are incidents that have preoccupied a lot the research community, it supplementary detects collisions.
  3. The incidents' modeling by a generic HMM procedure that utilizes characteristic features of each incident.
  4. A tool that in long term can provide statistics about the inherent dangers of the building's areas so that measures are taken.

## 2 RELATED WORK

Fall detection is a topic with prolific literature. In general fall detection methods can be divided in two groups: 1) vision based algorithms, 2) techniques with non-vision sensors such as accelerometers [2]. The first category is mainly preferred since wearable equipment can be irritating for users. Moreover, depth sensors gain leverage in comparison to other vision sensors since they respect the legal and ethical rights of human privacy.

A usual criterion for fall detection methods is the distance from the centroid or top of a person to the floor. In particular, in [3] a predefined threshold to the floor and a specific time of immobility on the floor specify a fall while in Ref. [4] a fall is determined by a k-NN classifier that utilizes features such as head-floor distance, person area and shape's major length to width. Moreover, other approaches [5] utilize skeleton joints tracking so as to achieve fall detection. Furthermore, in Ref. [6] an ensemble of decision trees for the fall's confidence computation is proposed, while in Ref. [7] a Bayesian framework is used. A good survey on vision based approaches for fall detection can be found in [8].

Despite the fact that fall detection preoccupied plenty scientists, collision in indoor environments has not been explored, at least to our knowledge. There is work on collisions for traffic applications [9], nevertheless the environmental conditions and velocities in this situation are completely different.

## 3 INCIDENT DETECTION METHOD

The proposed incident detection method includes three steps. Firstly, the moving items in the monitoring area are detected and tracked. Subsequently, features that characterize the item's state, such as its velocity, area variance or distance from a region of interest are extracted. Finally, based on these features, an HMM method models, detects and recognizes the occurring incidents.

### 3.1 Detection and tracking of moving items

The first step of the method is detection and tracking of all the moving items in the monitoring area. In particular, it includes a robust, real-time, multi-camera, multi-space occupancy extraction system based on [10]. In this procedure, the geometrical building information extracted from the BIM is utilized for the appropriate set-up of the sensors and their calibration. Thus, all the utilized cameras' coordinates are transformed to a common global reference system located on the architectural map of the building enabling the monitoring of any area regardless its size. Furthermore, a virtual top-view camera based on the calibration data is utilized so that detection and tracking can be performed on the horizontal plane unaffected from partial occlusions. Moreover, a dual-band algorithm handles the dynamic changes of the

environment by incorporating low objects, e.g. chairs, in a small time period, while retaining higher objects like humans for a longer period.

### 3.2 Feature extraction

The second step of the incident detection method is the extraction of features that characterize the state of the tracked items. These features depend on the type of incident under detection. Therefore, in the case of falls the selected features constitute the vertical velocity, area variance and height of the tracked worker while collision depends on the values of relative distance and velocity between two moving items. Contrary, intrusion is a straightforward procedure deriving from tracking therefore no further features are needed for its detection. In the following sections the fall and collision features are analyzed in detail.

#### 3.2.1 Fall features

During a person's fall their vertical velocity and area measured from top view are augmenting drastically. In addition, their height at the end of the fall is relatively small. Therefore, these features are utilized to define a human at a falling state and are described below in detail.

1. Vertical velocity  $v$ : it constitutes the vertical velocity of the highest point of the tracked person's blob. Nevertheless, depth sensors' noise could affect the velocity's values leading to false results. Therefore, a six order, low pass FIR filter is applied on the height values while the velocity is averaged on a short constant window  $C_T$ . The velocity's formula is given below:

$$v = \frac{1}{C_T} \int_{t_c - C_T}^{t_c} h'(t) dt, \quad (1)$$

where  $h'(t)$  is the tracked person's height derivative and  $t_c$  is the current time.

Moreover, a big portion of the upper part of a human's blob could abruptly vanish due to partial occlusions resulting in false alarms. To deal with these cases the step signal created at the average velocity is detected and all the velocities that include it in their calculation are set to zero.

2. Area variance  $\sigma^2$ : it refers to the area of the tracked person from top view. It is obvious that as a person falls their area from top view drastically augments. To remain unaffected from the initial area before the fall the area variance is calculated in the same time window  $C_T$  as vertical velocity assuring relative robustness to noise. Its formula is described by the equation:

$$\sigma^2 = \frac{1}{C_T} \int_{t_c - C_T}^{t_c} \left( A(t) - \int_{t_c - C_T}^{t_c} \frac{A(t')}{C_T} dt' \right) dt, \quad (2)$$

where  $A(t)$  represents the area of the tracked person from top view.

3. Height  $h$ : it is the highest point of the tracked person's blob. The value of the final height of the person after the fall facilitates the avoidance of false alarms.

#### 3.2.2 Collision features

Contrary to fall, collision needs more than one person or moving item in order to occur. Therefore, the features that define it are relative features between two participants. In particular the relative distance and velocity between two objects are utilized.

1. Relative distance  $d_r$ : it refers to the distance of two participants on the horizontal plane. As it is obvious the magnitude of the distance highly relates to the probability of collision.
2. Relative velocity  $v_r$ : it represents the speed the two participants are approaching or diverging, affecting the collision probability. It is given by the following formula:

$$v_r = \frac{d(dr)}{dt}. \tag{3}$$

### 3.3 Incident recognition

A three state HMM is utilized for modeling and recognizing the incidents based on the aforementioned features. In both cases the first state (S1) defines a stable condition where the incident under detection has not occurred yet. The second state (S2) includes the actual action of the incident while the third state (S3) signifies the termination and confirmation of the event. For example, a walking worker would be in state S1. Once he started falling he would enter S2 while at the end of the fall, once he reached the floor, he would be at state S3 unless if in the meantime he regained his balanced and returned to state S1. In the same spirit, two moving objects at a big distance from each other would be in state S1. Once they started approaching each other in a close distance they would transfer to state S2 and if they finally collided they would enter state S3.

The transition probabilities of the HMM are encapsulated in the following formula:

$$P = \begin{bmatrix} (1-E) & E & 0 \\ (1-E)(1-u) & E & (1-E)u \\ r & 0 & (1-r) \end{bmatrix}, \tag{4}$$

where  $E$ ,  $u$  and  $r$  are functions that depend on the event. Nevertheless, in both cases of fall and collision, function  $E$  describes the actual incident as it occurs; function  $u$  refers to the termination and verification of the event while function  $r$  depends on the time needed to transfer to another state after the occurrence of the event. The functions are analytically described for each incident separately (depending on the event each function is represented with the corresponding subscript).

#### 3.3.1 Fall

As it is already mentioned, the characteristic features of human fall are the vertical velocity  $v$ , the area variance  $\sigma^2$  and the final height at the end of the fall  $h$ . Therefore the functions of the transition probabilities depend exclusively on these characteristics and can be described by the following formulas:

$$E_f = \frac{1}{q + e^{v\sigma^2 + T_f}}, \tag{5}$$

$$u_f = \begin{cases} 1, & H_{T_f} - h \geq 0 \\ 0, & H_{T_f} - h < 0 \end{cases}, \tag{6}$$

$$r_f = 1, \tag{7}$$

where  $T_f$  and  $H_{T_f}$  are constants.

The eqn (5) is a sigmoid function that represents the probability  $E_f$  since for cases of high vertical velocity and area variance it reaches high values close 1. Moreover, eqn (6) is a step function that defines the detection of a fall since a person cannot be considered fallen if their height is above a predefined threshold  $H_{T_r}$ . Finally, eqn (7) signifies that once the fall has ended the HMM notifies of the incident and transfers to a stable condition.

### 3.3.2 Collision

The characteristic features of collision are the relative distance and velocity of the two participants to the event. The functions that describe this incident are the following:

$$E_c = e^{-\sigma_{c_d}(d_r - T_d)} \frac{1}{1 + e^{\sigma_{c_v} v_r}}, \quad (8)$$

$$u_c = \begin{cases} 1, & D_T - d_r \geq 0 \\ 0, & D_T - d_r < 0 \end{cases}, \quad (9)$$

$$r_c = 1 - u_c, \quad (10)$$

where  $\sigma_{c_d}$  and  $\sigma_{c_v}$  define the slope of the two functions comprising the eqn (10),  $D_T$  and  $T_d$  are constants.

Equation eqn (8) comprises two functions that refer to the relative distance and the relative velocity respectively. The exponential function represents the inversely analogous relationship between the probability of collision and the distance between the two participants while the sigmoid function depicts the fact that highly negative velocities decrease the probability of collision while highly positive increase it. The step function in eqn (9) determines the distance that signifies a collision while eqn (10) declares that a collision is considered to have ended only when the distance between the two participants is above a predefined threshold.

## 4 EXPERIMENTS

### 4.1 Datasets

Two datasets from two different work environments at Centre of Research and Technology Hellas (CERTH) were acquired for the evaluation of the proposed method. The datasets comprised various incidents including falls and collisions performed by 28 different participants. Nevertheless, other events that resemble the incidents under detection, such as bending to pick up something or tie one's shoes, two persons walking to the same direction without colliding etc, are also reproduced in order to evaluate the robustness of the algorithm in incidents that affect the features that determine the HMM transitions. Screenshots from the datasets are depicted in Fig. 1.

The dataset consists of 140 falls, 60 collisions, 50 cases of people bending to tie their shoes, 100 cases of people bending to pick up an item (e.g. various sizes of boxes, chairs) and 40 cases of pairs of people walking by each other without colliding. In general, the participants were allowed to conduct the incidents in ways that felt natural to them resulting in a variety of data. Therefore, there were people falling forwards, backwards or sideways, while standing, as if they were fainting, or while walking, as if they were stumbling. Moreover, the participants simulating collision diverged and approached in various directions and velocities. Since the experiments were conducted in work environments there were many members of the personnel walking or standing in the monitored area but although they did not produce false alarms they were not taken into account for the method's evaluation.

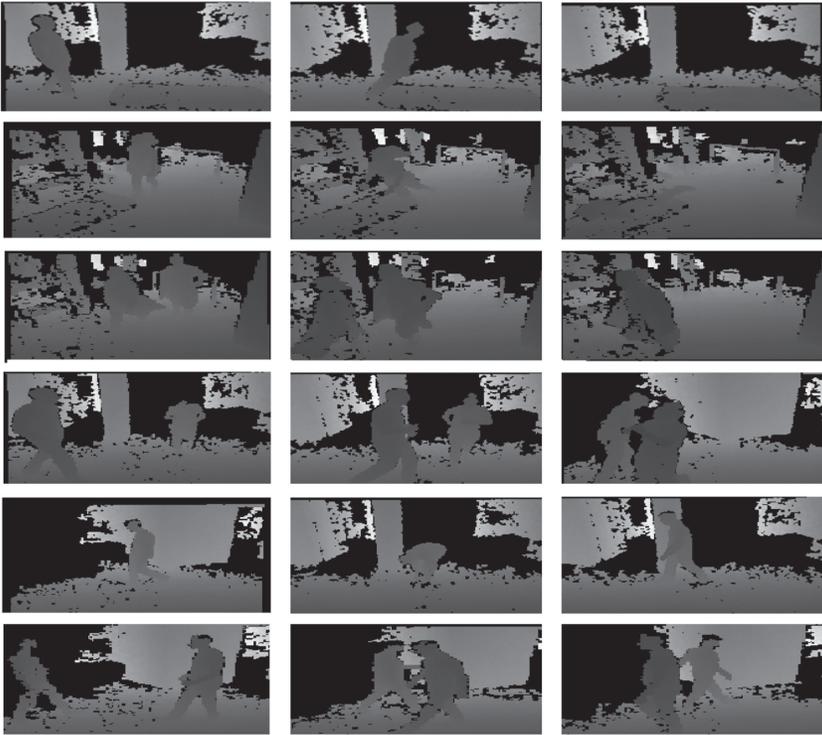


Figure 1: Incidents included in the dataset. The first two rows depict fall events, the third and fourth include collision incidents while the fifth row shows an event similar to fall, i.e. bending, and the last row an event that resembles collision, meaning two persons passing by each other in a close distance.

#### 4.2 Feature processing

In this section the features that characterize each incident and determine the HMM transition probabilities are further analysed with some examples so that their influence and importance on incident detection is better displayed. Thus, in Fig. 2 the vertical velocity, area variance and height during a fall are depicted, while Fig. 3 illustrates the relative distance and velocity during a collision. In both cases the transition probability function  $E$  (eqns (5) and (8)) that represents the actual fall is depicted.

In particular, Fig. 2 represents the height of a worker during a fall incident. As it can be observed, at the beginning the worker is in standing position, then, falls and immediately starts rising. In correspondence to the height diagram, vertical velocity (whose direction is defined upwards) starts from zero values, drastically decreases during the fall and, finally, fluctuates near zero. Similarly, the area variance highly augments during the fall. Based on these values, the probability function reaches high values during the fall while retains low values in the rest time period. The sequence of the HMM states according to the probability of values is depicted with colours on the graphs.

In Fig. 3 graphs of collision features are illustrated. As the moving objects approach their relative distance diminishes and their relative velocity takes negative values. Once the relative distance reaches and remains below a threshold the relative velocity oscillates due to collision.

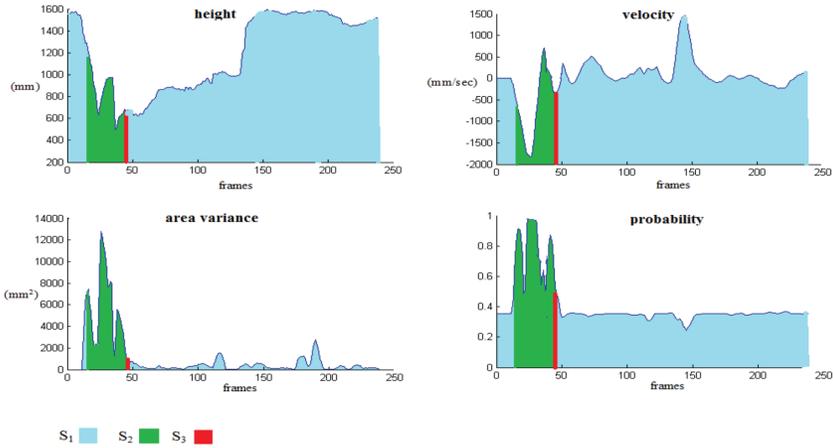


Figure 2: Fall features and probability  $E_f$ . The sequence of the HMM states according to the probability values is depicted with colours.

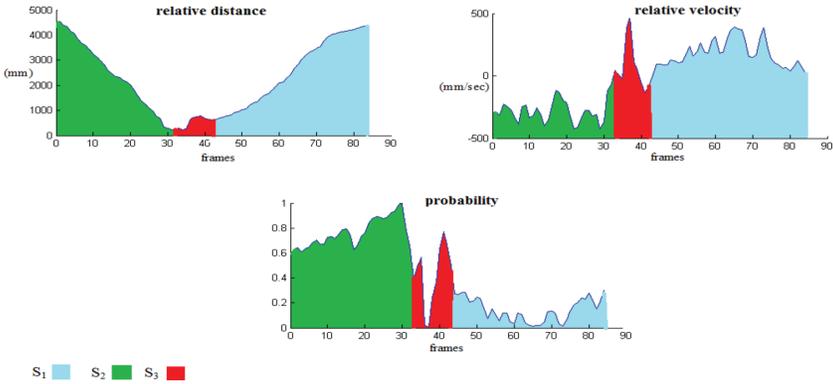


Figure 3: Collision features and probability  $E_c$ . The sequence of the HMM states according to the probability value is depicted with colours.

At this point the HMM state transitions from S2 to S3, thus probability  $E_c$  does not affect the method. When the collision terminates the state transitions from S3 to S1 and probability  $E_c$  remains below 0.5. It is obvious that although the relative distance's values are similar before and after the collision probability  $E_c$  diminishes drastically due to relative velocity.

### 4.3 Experimental results

The proposed method was tested in more than 600000 frames that included apart from the incidents of interest, i.e. falls and collisions, events with similar properties to test the algorithm's robustness. The extraction of the results is not frame-based but depends on the number of times each incident occurred.

Fall incidents were tested in 140 cases. In 136 cases the falls were detected correctly, 4 were missed due to step signal just before the fall while 12 events similar to fall produced

Table 1: Incident detection results.

	True Positives	True Negatives	False Positives
Fall	136	238	12
Collision	58	325	5

Table 2: Incident detection performance.

	Precision	Recall	F1-score
Fall	0.919	0.971	0.945
Collision	0.921	0.967	0.947

false alarms. Nevertheless, it should be mentioned that 6 of the false positive cases referred to bending to pick up a mattress used for falls. However, incidents of picking up such a voluminous object do not occur often in industrial environment or are limited to a certain area.

In the incident of collision, 58 out of 60 cases were detected correctly, 2 were missed while 5 false alarms were produced from similar events (3 of them referred to persons meeting and standing very close to each other and 2 to persons passing by each other).

The results and performance of the presented incident detection method are analyzed in Tables 1 and 2. In both cases precision was approximately 0.92, recall 0.97 and F1-score 0.95, which can be considered very promising results. As it is already mentioned the proposed incident detection system can also support detection of intrusion to restricted areas. Nevertheless, since the detection of the particular incident depends completely on the accuracy of the tracking algorithm, which is explored and analyzed in Ref. [11], no experimental results are provided in this paper.

All possible alarms produced by the proposed system are depicted in Fig. 4. The location of the incident on the architectural map of the building and the input from two depth cameras monitoring the scene are illustrated. Each incident is depicted with different color. In Fig. 4a–c fall, collision and intrusion to a restricted area (marked with a gray rectangle) are depicted. In Fig. 4d possible collision between two moving objects is depicted. This alarm could serve as a proactive notification in order to avoid accidents and corresponds to state S2 of the HMM.

## 5 CONCLUSIONS

This paper presents a real-time, multi-space incident detection system for indoor environments. The system focuses on industrial environments while falls, collisions and intrusions to restricted areas constitute the incidents under detection. Once an incident is detected, an alarm is produced notifying for its type and exact location so that a coping mechanism can be immediately triggered. The importance of this prompt notification increases when the incident occurs in remote or hazardous areas with chemical agents, hard climatic conditions or running machinery. Therefore, the combined knowledge of the incident area's conditions and its type provide leverage for the building's personnel safety and health since the appropriate measures can be immediately taken. Furthermore, in long term, the system could provide statistics for the occurring incidents so that adjustments are made for better incident prevention. The importance of BIM in this application is high since it constitutes an essential component for the set up of the system while it provides information about the location of the

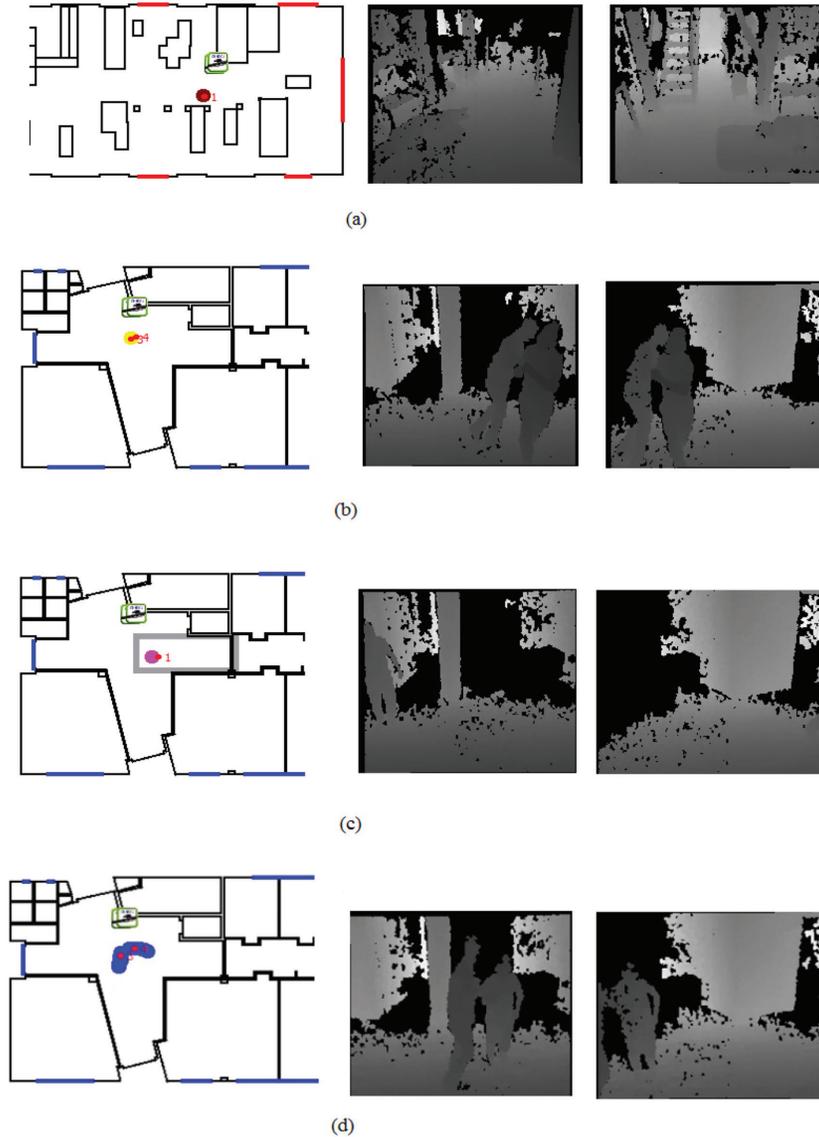


Figure 4: Incident detection and indication of their location at the building's architectural map: (a) fall, (b) collision, (c) intrusion to restricted area, (d) indication of possible collision.

incidents facilitating the immediate and appropriate coping actions according to the area's particularities.

Moreover, the proposed method for incident detection is based on a generic HMM procedure that utilizes features that characterize the incidents under detection, e.g. vertical velocity, relative distance between two moving items etc. The experimental results are very promising and conform to the vision of a 'smart' and safer industrial environment.

## REFERENCES

- [1] Triantafyllou, D., Krinidis, S., Ioannidis, D., Metaxa, I.N., Ziazios, C. & Tzouvaras, D., *A real-time fall detection system for maintenance activities in indoor environments*, 3rd IFAC AMEST Workshop, 2016.
- [2] Wu, G. & Xue, S., Portable pre-impact fall detector with wearable sensors. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, pp. 178–183, 2008.
- [3] Diraco, G., Leone, A. & Siciliano, P., An active vision system for fall detection and posture recognition in elderly health care. *Design, Automation and Test in Europe Conference and Exhibition*, pp. 1536–1541, 2010.
- [4] Kepski, M. & Kwolek, B., Fall detection using ceiling-mounted 3d depth camera. *International Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pp. 640–647, 2014.
- [5] Zhang, C., Tian, Y. & Capezuti, E., Privacy preserving automatic fall detection for elderly using RGBD cameras. *International Conference on Computer Helping People with Special Needs*, pp. 625–633, 2012.  
[https://doi.org/10.1007/978-3-642-31522-0\\_95](https://doi.org/10.1007/978-3-642-31522-0_95)
- [6] Stone, E. & Skubic, M., Fall detection in homes of older adults using the Microsoft kinect. *IEEE Journal of Biomedical and Health Informatics*, pp. 1590–1594, 2014.
- [7] Zhang, Z., Liu, W., Metsis, V. & Anthitsos, V., A viewpoint- independent statistical method for fall detection. *IEEE Conference on Pattern Recognition*, pp. 3626–3630, 2012.
- [8] Zhang, Z., Conly, C. & Anthitsos, V., A survey on vision based fall detection. *International Conference on Pervasive Technologies Related to Assistive Environments*, 2015.  
<https://doi.org/10.1145/2769493.2769540>
- [9] Ki, Y., Accident detection system using image processing and MDR. *Computer Science and Network Security*, 3(7), pp. 640–647, 2007.
- [10] Krinidis, S., Stavropoulos, G., Ioannidis, D. & Tzouvaras, D., A robust and real time multi space occupancy extraction system exploiting privacy preserving sensors. In *Communications, Control and Signal Processing (ISCCSP), 2014 6th International Symposium*, pp. 542–545, 2014.  
<https://doi.org/10.1109/isccsp.2014.6877932>
- [11] Ioannidis, D., Krinidis, S., Drosou, A., Tzouvaras, D. & Likothanassis, S., Occupant-aware indoor monitoring for enhanced building analysis. *SimAUD15*, pp. 85–92, 2015.