Bridging the gap between real-life data and simulated data by providing a highly realistic fall dataset for evaluating camera-based fall detection algorithms

Greet Baldewijns[1,2,3], *Glen Debard*[1,4], *Gert Mertes*[1,2,3], *Bart Vanrumste*[1,2,3] *and Tom Croonenborghs*[1,5,6]

[1]*KU Leuven Technology Campus Geel, AdvISe, Belgium*
E-mail: greet.baldewijns@kuleuven.be
[2]*KU Leuven, ESAT-STADIUS, Belgium*
[3]*iMinds Medical Information Technology department, Belgium*
[4]*Thomas More Kempen, Mobilab, Belgium*
[5]*KU Leuven, Department of Computer Science, DTAI, Belgium*
[6]*Program in Translational NeuroPsychiatric Genomics, Brigham and Women's Hospital, Harvard Medical School, Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA.*

Published in Healthcare Technology Letters; Received on xxx; Revised on xxx.

Fall incidents are an important health hazard for older adults. Automatic fall detection systems can reduce the consequences of such a fall incident by assuring that timely aid is given. The development of these systems is therefore getting a lot of research attention. Real-life data which can help evaluate the results of this research is however sparse. Moreover, research groups that have this type of data are not at liberty to share it because of ethical and privacy related issues. Most research groups thus use simulated datasets to evaluate their developed fall detection algorithms. These simulation datasets, however, often do not incorporate the challenges the fall detection system will face when implemented in an actual home setting. In this paper a more realistic simulation dataset is presented to fill this gap between real-life data and currently available datasets. This dataset was recorded while re-enacting real life falls which were recorded during previous studies. The dataset incorporates the challenges which fall detection algorithms will face when implemented in real life and which are often not incorporated in currently available datasets. A fall detection algorithm from Debard et al. [8] was evaluated on this dataset for each individual camera viewpoint. This evaluation showed that the dataset possesses some extra challenges compared to other publicly available datasets. In this paper the dataset is discussed in depth as well as the results of this preliminary evaluation of the fall detection algorithm of Debard et al. [8]. The dataset is publicly available and can be downloaded from www.kuleuven.be/advise/datasets.

1. Introduction: Approximately 424 000 individuals die yearly from fall-related injuries. Additionally, fall incidents are the cause of more than 50% of injury-related hospitalizations among people aged over 65 [1]. Besides physiological injuries, falls can also result in psychological consequences such as fear for future falls and depression [2]. Automatic fall detection systems can reduce these physiological and psychological consequences by ensuring that timely aid is given.

Another important benefit to the use of automatic fall detection systems is the fact that all fall incidents are registered and logged for future reference where now older adults often forget to report when a fall incident occurred and no physical injuries were obtained. This fall data can subsequently be used to analyse the causes of the fall incident and can possibly reveal health related problems which would otherwise remain undetected and can subsequently be treated. In this way the quality of life of the older adult would further improve.

For this purpose several research groups are focusing on the development of automatic fall detection algorithms. An overview of the current state-of-the-art fall detection systems is given in [4] and [5]. Two kinds of sensors are typically used for this research: body-worn sensors such as accelerometers [3] and contactless sensors such as cameras [8]. Body-worn sensors can pose problems for older adults who are cognitively impaired and thus might forget to wear the sensor [2]. A similar problem arises with the use of a panic button. After a fall incident the fall victim is not always capable of activating the alarm by pressing the button [2]. Camera systems on the other hand can overcome these limitations because they do not need to be worn on the body and don't require any interaction with the older adult.

Due to these advantages camera-based fall detection systems have come to the attention of several research groups [6, 7]. In spite of the fact that previous research showed that developing fall detection algorithms which work in real-life face additional challenges compared to a simulated setting [3, 9, 10], the majority of the algorithms are still evaluated using simulated falls performed in artificial surroundings [10]. The reasons for these extra challenges are twofold. Firstly, there are several environmental challenges which need to be tackled when using real-life data such as illumination changes and occlusions. These challenges are generally not present in simulation datasets [10]. Moreover, older adults typically do not fall in a textbook way. Kangas et al. observed important differences between falls included in simulation datasets and those which actually took place in real life [9].

There is thus a need for publicly available real-life datasets which allow for overall evaluation of already existing algorithms. In spite of this need, research groups that acquired such datasets are not able to share this data due to ethical and privacy related reasons. To bridge this gap between currently available simulation datasets and real-life datasets a more realistic falls dataset was recorded using 2D cameras. This dataset was acquired by re-enacting real-life fall incidents that were observed during previous studies [11]. In this paper the dataset is described in depth and it is publicly available via http://www.kuleuven.be/advise/datasets. The algorithm of Debard et al. as proposed in [8] was evaluated on this dataset as a first performance benchmark of the dataset.

In the remainder of the paper the acquired dataset is discussed in-depth. This is followed by a description of the evaluated algorithm together with the presentation of the results of this algorithm. Lastly, a discussion on both the dataset and the first fall detection results is provided.

2. Related work: Although the majority of the research concerning the development of camera-based fall detection

© The Institution of Engineering and Technology 2013

Table 1 Comparison between the properties of the simulation datasets found in literature and the presented dataset

	Number of video segments	Number of fall segments	Number of normal activity segments	Segment length	Number of locations	Number of cameras	Number of actors	Realistic fall incidents	Realistic setting	Real-life Challenges	Use of walking aids
Charfi et al.	250	192	57	10 to 45 sec	4	1	11	no	yes	yes	no
Auvinet et al.	24	22	24 [a]	30 sec to 4 min	1	8	1	yes	no	no	no
Baldewijns et al.	72	55	17	Fall segments: 50 sec to 4:58 min Normal Activities: 11:38 to 35:30 min	1	5	10	yes	yes	yes	yes

Notes
[a] Auvinet et al. combined normal activities and falls in the same video segments

algorithms uses simulated datasets recorded with 2D cameras, to the best of our knowledge only two of these datasets are publicly available: the dataset from Auvinet et al. [12] and the dataset from Charfi et al. [13].

Auvinet et al. [12] recorded 24 video segments ranging from 30 seconds to 4 minutes using eight calibrated cameras. The video segments contain 22 fall events and 24 other events such as crouching, sitting and lying on a sofa. Auvinet et al. used previously published research to create realistic fall scenarios. In contrast a very artificial environment to record the video segments with a limited set of furniture is used and did not include any of the challenges which are specific to a real-life setting such as occlusions, changes in illumination, etc.

Charfi et al. [13], on the other hand, used a single camera set-up to record 249 video segments. These video segments ranged from 10 to 45 seconds. 192 of these video segments contained a fall incident and 57 contained normal activities such as walking, sitting down, standing up and housekeeping. To better fit real-life circumstances Chafri et al. incorporated some of the challenges listed in [10] in their dataset such as changes in light intensities, shadows and moving objects. To further improve the realism of their dataset they recorded the data in four different locations (home, coffee room, office, lecture room). However, a set-up in which an older adult has only one room to perform all his or her activities of daily living in, such as is often the case in a nursing home, was not included. Moreover, the recordings always start when a person is already in the room. This combined with the fact that the segments are very short makes it difficult to use background subtraction methods to isolate the person in the video frames.

While both studies took certain steps to make the data as realistic as possible, there are still some shortcomings to these datasets. Firstly, they use very short video segments, in contrast with the fact that real-life fall detection systems will need to work continuously. Moreover, due to these very short segments the post-fall activity is very limited, whereas this is often used in combination with the fall data to determine if a fall has occurred. Next, the number of non-fall activities is very limited compared to the amount of non-fall data which will be generated when implementing a fall detection system in an actual home setting. It is thus very difficult to verify the number of false alarms the fall detection algorithms will generate based on these datasets. Lastly, no walking aids are included in both datasets. As these are frequently used by older adults and can present their own challenges when a fall is occurring (e.g. walker rolls away while the person is falling), they should be included in both fall and non-fall scenarios.

Debard et al. [14] compared the performance of a state-of-the-art fall detector on the dataset of Auvinet et al. and a real-life dataset. This clearly showed the need of more realistic datasets. An overview of the properties of these simulation datasets and the in this paper presented dataset is given in Table 1.

Besides these 2D camera-based fall simulation datasets there are some datasets available which are recorded with the Microsoft Kinect, which provides a colour image, such as the datasets from Bogdan and Kepski [15], Anderson et al. [16] and from Gasparrini et al. [17]. Next to these datasets there are also some action recognition datasets available which include falls as an action such as is done in the dataset from Kuehne et al. [18]. Both the Kinect datasets as well as the action recognition datasets are however subject to the same shortcomings as other video-based datasets such as: very short fall and Activities of Daily Living (ADL) segments, limited number of fall scenarios, non-realistic falls and ADL scenarios, no furniture or a limited number of furniture in the room, etc.

3. The dataset: While recording the dataset several goals were kept in mind. Firstly, we aimed to record the different scenarios in a realistic environment. Next, the different fall scenarios needed to resemble actual fall incidents as closely as possible. A better balance between fall segments and segments with normal activity was also envisioned. Lastly, instead of recording only short segments, the effect of long recording sessions on the performance of the data acquisition system was also assessed. The resulting dataset is publicly available via http://www.kuleuven.be/advise/datasets [1] .

In this section these different goals and how these were met are discussed.

3.1. Goal 1 - Realistic setting: A room was furnished to look like a nursing-home room. Five IP-cameras (12 fps, resolution: 640x480) were installed in this room and used to record the different fall and non-fall scenarios. Because the room was located in the basement, it did not have any windows. This was compensated by placing an additional light fixture close to the wall. This light fixture was used to provide changes in illumination. Near-infrared spots were used to facilitate the acquisition when lower-light intensities were present. A schematic overview of this room and the position of the cameras is given in Figure 3.1.

3.2. Goal 2 - Realistic fall scenarios: The real-life falls which were recorded during previous research as described in [10, 11] were watched attentively and scenarios describing these falls were written. As the number of falls recorded in these previous studies was limited, it was decided that the same fall type would be recorded several times, but with different actors so that some variations would be present. Also some fall scenarios which were not observed in real-life, but which are nevertheless present in the widely used datasets, were incorporated. These last fall scenarios are marked in the meta file.

The resulting fall scenarios were re-enacted by 10 different actors using different walking aids. In total 55 fall scenarios were

[1] The dataset will become available when the paper is published. With the submission of this paper two sample video files were included and a sample metafile.

© The Institution of Engineering and Technology 2012

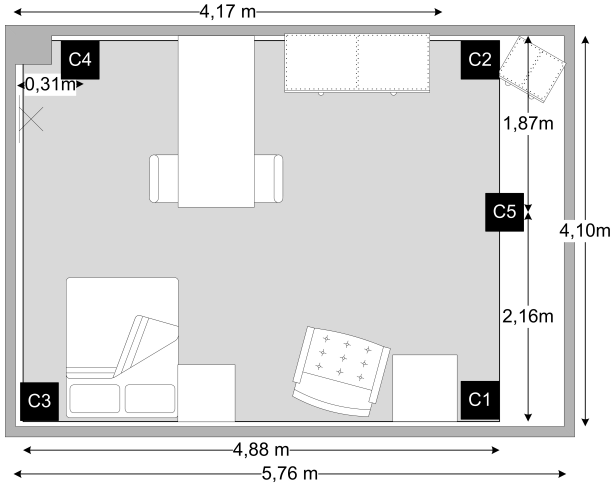


Figure 1 *Overview of the room in which the dataset was recorded (entrance to the right of C1) - The position of the different cameras is marked with C1 to C5*

recorded. The average length of the fall scenarios was 2:45 minutes with a minimum length of 50 seconds and a maximum length of 4:58 minutes. In total 2:25:54 hours of fall data were acquired by each camera. The different fall scenarios differ from each other due to the used walking aid, fall speed, possible moving objects during the fall, starting pose and stopping pose. An overview of the different types of scenarios is given in Table 2. A comparison of a real-life fall incident and a simulated fall scenario is given in Figure 2.

Table 2 Overview of the different fall scenario types

Used walking aid	
Walker:	13 scenarios
Wheelchair:	4 scenarios
No walking aid:	38 scenarios
Fall speed	
Slow falls:	22 scenarios
Fast falls:	33 scenarios
Moving objects during the fall	
Walker:	8 scenarios
Wheelchair:	4 scenarios
Blanket:	3 scenario
Chair:	10 scenarios
None:	30 scenarios
Starting pose	
Standing:	26 scenarios
Sitting:	10 scenarios
Bending over:	5 scenario
Squatting:	6 scenarios
Transitions (sit $\leftrightarrow$ stand)	8 scenarios
Ending positions	
Lying on the floor:	49 scenarios
Getting back-up after fall:	6 scenarios

3.3. Goal 3 - Balance between fall data and normal activity data: A better balance between the number of fall and non-fall activities was envisaged. This balance was reached by acquiring a total of 5:50:29 hours worth of normal activities divided over 17 different scenarios. The different video segments have an average length of 20:39 minutes per scenario of which the shortest segment was 11:38 minutes long and the longest was 35:30 minutes long. Each scenario consisted of a set of normal activities. Some variations were induced in the different normal activity scenarios by using different walking aids, changing the pace of the person performing the actions, changing the order in which the activities were performed, not including all activities in each scenarios and finally not performing all activities in the same positions (e.g. sleeping under the blankets, on top of the blankets or in the chair). The normal activities which were included in these scenarios were:

- Walking (with and without walking aid);
- Sit to stand and stand to sit transitions;
- Sitting;
- Eating and drinking;
- Getting into and out of bed;
- Sleeping;
- Changing clothes;
- Removing and putting on shoes;
- Reading;
- Transfers from wheelchair to chair and vice versa;
- Making the bed;
- Coughing and sneezing violently;
- Picking something up from the floor.

3.4. Goal 4 - Real-life challenges: Within the fall scenarios additional challenges were incorporated such as occlusions, falling partially or completely out of the camera view and in one scenario more than one person was in field of view of the camera. Within the normal activities the same challenges were incorporated as well as changes in illumination.

3.5. Goal 5 - Continuously recording: As in real-life environments the system needs to be recording continuously, the acquisition system was running for a whole week continuously. From this continuous stream of data we segmented those parts of the video in which a scenario was performed. Because of this continuous recording, the system suffered from some frame drops resulting in fall scenarios with a lower number of frames per second than initially intended. Since it is not unlikely that such a problem arises when installing camera-based fall detection systems in the home environment of older adults, these segments were included in the final dataset. It is important to assess the validity of the algorithms as well when the quality of the recordings is not ideal. The scenarios with a high frame drop are marked separately in the meta-file.

3.6. Data organization: The available data is organized as followed. For each scenario five video files are available corresponding with the different cameras. In the name of the video file the fall number and camera are identified (e.g. Fall1Cam1.avi). A similar approach is used for the ADL scenarios. Again scenario number and camera number are included in the naming of the file (e.g. Adl1Cam1.avi).

Meta data concerning the fall incidents is available in both pdf and txt format and can be found in the 'FallMeta.pdf' file and in the 'FallMeta.txt file'. The information available in this file includes a short description of the fall, the type of fall, length of the video segment, time of fall, used walking aid, start position and stop position. The 'AdlMeta.pdf' and 'AdlMeta.txt' file contains information concerning the different ADL scenarios such as segment length, included activities and used walking aid.

4. Fall detection: The presented dataset was used to evaluate a first camera-based fall detection algorithm from Debard et al. as discussed in [8].

Debard et al. used background subtraction to locate the person in each video frame. To increase the robustness of the background update, they used a particle filter based on three different measuring functions. The first measurement was based on the binary foreground image which was cleaned up using shadow removal and morphological operations. An ellipse that fitted the foreground well got a high score. The second measurement was

© The Institution of Engineering and Technology 2012



(a) Beginning of the real-life fall incident



(b) Beginning of the simulated fall scenario



(c) Ending of the real-life fall incident



(d) Ending of the simulated fall scenario



(f) Beginning of the real-life fall incident



(g) Beginning of the simulated fall scenario



(h) Middle of the real-life fall incident



(i) Middle of the simulated fall scenario

Figure 2. Comparison of a real-life fall incident with a simulated falls scenario

© The Institution of Engineering and Technology 2012

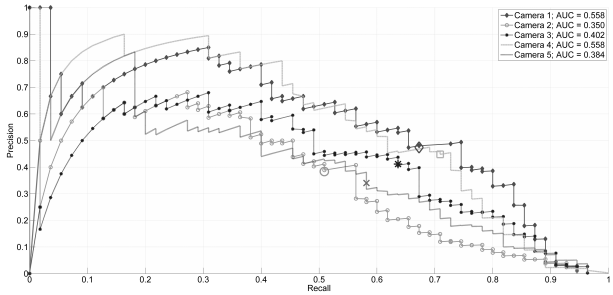


Figure 3 *Precision-Recall curves resulting from the single camera fall detection algorithm used on each camera individually (AUC = Area Under Curve). Working point chosen to calculate the precision and recall is marked individually on each curve*

Table 3 Precision, Recall and Area Under the Curve (AUC) per camera

	Precision	Recall	AUC
Cam 1	0.47	0.67	0.56
Cam 2	0.38	0.51	0.35
Cam 3	0.41	0.64	0.40
Cam 4	0.45	0.71	0.56
Cam 5	0.34	0.58	0.38

based on histogram correlation. The histogram was based on a structural weighted intensity histogram. The ellipse was divided in four parts representing different parts of the body. In each part, the centre was given a higher weight to reduce the impact of background pixels that could be included. The histogram model was updated every frame using the predicted state. The final measurement function was based on an upper body detector. The detections were used to update the histogram model and also to shift the weight between the first and second measurement function. The prediction of the particle filter was then used to update the background. Inside of the predicted ellipse the background was updated very slowly and the rest of the image was updated very fast. The person is located in the frame using the biggest bounding box in the frame.

When a person is located in the video frame, Debard et al. comprise a fall feature vector consisting of five fall features (aspect ratio, change in aspect ratio, fall angle, centre speed and head speed) calculated over different timeslots before, during and after the fall incident. The feature vector is subsequently used by a Support Vector Machine (SVM) to perform the final fall classification. More information concerning this fall detection algorithm can be found in [8].

As Debard et al. implemented a single camera fall detector, the data from each camera was evaluated individually. This was done by training a fall detection model for each camera and evaluating these models through 10-fold cross validation. The resulting precision-recall curves are given in figure 3. For each camera viewpoint a precision and recall were calculated as well as the area under the curve. These results are given in table 3.

5. Conclusion: In this paper a dataset containing simulation falls and normal activities is presented. The dataset distinguishes itself from other dataset for several reasons. Firstly, the simulation setting was set up to closely resemble a nursing home setting which is often the place where older adults reside. Next, real-life falls recorded during previous projects were re-enacted to ensure that the recorded fall incidents were realistic. A better balance between the amount of normal activities and fall incidents was also established. Lastly, the challenges which can be encountered when recording in real-life were also included in the dataset. Although evaluating fall detection algorithms in real-life remains necessary, a first evaluation on this more realistic dataset can reveal some challenges in the fall detection algorithm which need to be faced before installing the system in real-life.

6. Future Work: Regarding the dataset future work will include further recording of real-life data as well as recording more simulation data in different living arrangements e.g. nursing home, service flats, etc. Other research groups can also record the scenarios using different actors and settings based on the now available data. The research effort will furthermore focus on making real-life data publicly available without breaching the privacy of the participants. This could be either done by replacing the person in the video with an avatar or by extracting commonly used fall features from the data and making these features publicly available. Future work will also include the further evaluation of both single sensor camera-based fall detection algorithms as well as multi-camera fall detection algorithms. Next it is desired to remedy any shortcomings in these existing algorithms which are revealed when using this dataset, thus providing more reliable fall detection algorithms. The researchers further hope to motivate other research groups to use the dataset for the evaluation of their fall detection algorithms.

7. Acknowledgment: This work was funded through the "ingenieurs@WZC"-project which was funded by "Provincie Vlaams-Brabant". Project partners are OCMW Leuven, KU Leuven, AdvISe and InnovAGE.
The authors would like to acknowledge the following projects: iMinds FallRisk project, IWT-ERASME AMACS project, IWT Tetra Fallcam project and the ProFouND network. The authors would also like to acknowledge networking support by the COST Action IC1303 AAPELE.

8 References

[1] World Health Organization. Ageing and Life Course Unit: 'Who global report on falls prevention in older age', World Health Organization, 2008.

[2] J. Fleming and C. Brayne: 'Inability to get up after falling, subsequent time on floor, and summoning help: prospective cohort study in people over 90', *BMJ*, 2008, **337**, a2227.

[3] F. Bagalà, C. Becker, A. Cappello, L. Chiari, K. Aminian, J. M. Hausdorff, W. Zijlstra, and J. Klenk: 'Evaluation of accelerometer-based fall detection algorithms on real-world falls', *PloS one*, 2012, **5**, e37062.

[4] M. Mubashir, L. Shao, and L. Seed: 'A survey on fall detection: Principles and approaches', *Neurocomputing*, 2013, **100(0)**, pp. 144-152, Special issue: Behaviours in video.

[5] G. Yongli, O. S. Yin, and P. Y. Han: 'State of the art: a study on fall detection', *Proceedings of World Academy of Science, Engineering and Technology*, 2012, pp. 62.

[6] Y. Lee and W. Chung: 'Automated abnormal behavior detection for ubiquitous healthcare application in daytime and nighttime', *Proceedings of the international conference on Biomedical and Health Informatics (BHI), IEEE-EMBS*, 2012, pp. 204-207.

[7] V.D. Nguyen, M.T. Le, A.D. Do, H.H. Duong, T.D.Thai and D.H. Tran: 'An efficient camera-based surveillance for fall detection of elderly people', *Proceedings of the 9th conference on Industrial Electronics and Applications (ICIEA) IEEE*, 2014, pp. 994-997.

[8] G. Debard, G. Baldewijns, T. Goedemé, T. Tuytelaars, and B. Vanrumste: 'Camera-based fall detection using a particle filter', *Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2015, pp. 6947-6950.

© The Institution of Engineering and Technology 2012

[9] M. Kangas, I. Vikman, L. Nyberg, R. Korpelainen, J. Lindblom, and T. Jämsä: 'Comparison of real-life accidental falls in older people with experimental falls in middle-aged test subjects', *Gait & posture*, 2012, **35(3)**, pp. 500-505.

[10] G. Debard, P. Karsmakers, M. Deschodt, E. Vlaeyen, E. Dejaeger, K. Milisen, T. Goedemé, B. Vanrumste, and T. Tuytelaars: 'Camera-based fall detection on real world data, Outdoor and large-scale real world scene analysis', *Springer*, 2012, pp. 356-375.

[11] E. Vlaeyen, M. Deschodt, G. Debard, E. Dejaeger, S. Boonen, T. Goedemé and K. Milisen: 'Fall incidents unraveled: a series of 26 video-based real-life fall events in three frail older persons', *BMC geriatrics*, 2013, **13(1)**, 103.

[12] E. Auvinet, F.Multon, A. Saint-Arnaud, J. Rousseau, and J.Meunier: 'Fall detection with multiple cameras: An occlusion resistant method based on 3-D silhouette vertical distribution', *IEEE Transactions on Information Technology in Biomedicine*, 2011, **15(2)**, pp. 290-300.

[13] I. Charfi, J. Miteran, J. Dubois, M. Atri, and R. Tourki: 'Definition and performance evaluation of a robust svm based fall detection solution', *Proceedings of the Eight International Conference on Signal Image Technology and Internet Based Systems (SITIS)*, 2012, pp. 218-224.

[14] G. Debard, M. Mertens, M. Deschodt, E. Vlaeyen, E. Devriendt, E. Dejaeger, K. Milisen, J. Tournoy, T. Croonenborghs, T. Goedemé, T. Tuytelaars and B. Vanrumste: 'Camera-based fall detection using real-world versus simulated data: how far are we from the solution?, *Journal of Ambient Intelligence and Smart Environments*, 2015, Accepted for publication.

[15] K. Bogdan and M. Kepski, 'Human fall detection on embedded platform using depth maps and wireless accelerometer.' *Computer methods and programs in biomedicine*, 2014, **117(3)**, 489-501.

[16] D. Anderson, R.H. Luke, J.M. Keller, M. Skubic, M. Rantz, and M. Aud. 'Linguistic summarization of video for fall detection using voxel person and fuzzy logic', *Computer vision and image understanding*, 2009, **113(1)**, 80-89.

[17] S. Gasparrini, E. Cippitelli, E. Gambi, S. Spinsante, J. WÃěhslÃľn, I. Orhan, and T. Lindh, 'Proposal and Experimental Evaluation of Fall Detection Solution Based on Wearable and Depth Data Fusion' *ICT Innovations 2015*, 2015, 99-108, Springer International Publishing.

[18] H. Kuehne, H. Jhuang, E. Garrote, T. Poggio and T. Serre, 'HMDB: a large video database for human motion recognition', *Proceedings of the International Conference on Computer Vision (ICCV)*, 2011, 2556-2563.

© The Institution of Engineering and Technology 2012