Smart Camera Reconfiguration in Assisted Home Environments for Elderly Care

Krishna Reddy Konda, Andrea Rosani(✉),
Nicola Conci, and Francesco G.B. De Natale

Department of Information Engineering and Computer Science,
University of Trento, Via Sommarive 9, 38123 Trento, Italy
andrea.rosani@unitn.it

Abstract. Researchers of different fields have been involved in human behavior analysis during the last years. The successful recognition of human activities from video analysis is still a challenging problem. Within this context, applications targeting elderly care are of considerable interest both for public and industrial bodies, especially considering the aging society we are living in. Ambient intelligence (AmI) technologies, intended as the possibility of automatically detecting and reacting to the status of the environment and of the persons, is probably the major enabling factor. AmI technologies require suitable networks of sensors and actuators, as well as adequate processing and communication technologies. In this paper we propose an innovative solution based on a real time analysis of video with application in the field of elderly care. The system performs anomaly detection and proposes the automatic reconfiguration of the camera network for better monitoring of the ongoing event. The developed framework is tested on a publicly available dataset and has also been deployed and evaluated in a real environment.

Keywords: Elderly care · Real time video analysis · Automatic camera reconfiguration

1 Introduction

In the society of the developed countries there is the evidence that the number of elderly people, is rapidly increasing with respect to the past, also due to an longer life expectancy. Normal life activity becomes somehow difficult while aging, thus resulting a parameter to be monitored in order to infer the health status of a person [14,24].

Moreover, the analysis of human behavior can detect remarkable situations where people need immediate assistance to avoid major risks and severe injuries.

Notwithstanding the effort of researchers in the direction of behavior understanding, a complete and reliable description of the ongoing situation using only sensors analysis is still a challenging task.

Activities and behaviors can be detected using different approaches, and video analysis is often used among them [6] because in many situation cameras

L. Agapito et al. (Eds.): ECCV 2014 Workshops, Part IV, LNCS 8928, pp. 45–58, 2015.
DOI: 10.1007/978-3-319-16220-1_4

are already present in the site for security reasons and, if not, because of the low cost of installation and maintenance. Compared to other sensing techniques, video cameras have the advantage of not being intrusive and burdensome like wearable sensors, and they are usually more precise than ambient sensors.

Thanks to the advances in video signal processing technology and the increasing computer power available, video analysis can provide significant information for this purpose including location, posture, motion, interaction with objects, people, and environment [3].

In this paper we present a framework for elderly care using the information coming from the video motion flow directly from streaming in the compressed domain, prior to decoding. The proposed technique is able to detect events directly and to perform camera reconfiguration in order to focus the attention on the occurring event. This allows for a very fast and robust detection of the event, with the possibility of real-time processing of the information and early alerting.

For testing purposes we considered the event "fall" [17,26] as a remarkable example of a situation where there is the need of an immediate assistance. The *fall* can be described as a sudden event, which takes place in about half a second. Such a sudden event usually creates large variations in terms of pixel intensities and visual features. In addition to detection of falls, obtaining a better view of the person involved in the accident would help in taking the most appropriate action. To this extent, we propose a unique system, capable of detecting falls and also to reconfigure the camera network in real-time, in order to achieve a more accurate information of the event. Video analysis is performed by completely operating in H.264 [25] compressed domain, while reconfiguration is achieved by automatically modifying the Pan-Tilt-Zoom (PTZ) parameters of the camera.

2 State of the Art

2.1 Camera Reconfiguration

Research on camera reconfiguration is in a nascent stage. Micheloni et al. summarized the current state of the research in [16]. More in general, camera reconfiguration is performed with respect to a specific task. One of the earliest works to consider PTZ cameras is [18], where PTZ cameras are specifically used for tracking. In another paper Quaritsch et al. [20] utilize multiple cameras and reconfiguration to achieve better tracking. Scotti et al. [23] utilize a PTZ camera along with an omnidirectional camera, in order to achieve tracking of objects at higher resolution. Another work, which utilizes a combination of omni directional and PTZ cameras for tracking is presented in [5], where the authors approach the problem in terms of spatial correlation, in order to map the targets across two types of cameras.

2.2 Ambient Assisted Living and Fall Detection

During the last years, there has been an increasing amount of technologies dedicated to the care of the elderly, usually grouped under the umbrella of

"Ambient Intelligence". These technologies span from a variety of sensors and analysis techniques to infer the performed activity [21]. Recently, the research community developed solutions based on video analysis and computer vision systems with promising results. However, to reach maturity, several challenges still need to be faced, including the development of systems that are robust in the real-world and are accepted by users [4]. One of the major objective of Ambient Intelligence is to provide safety to the monitored people, especially those who need more attention. To this extent, the detection of falls represents a remarkable situation, where the automatic detection of the event can activate an early alert to relatives and caregivers.

Fall detection is apparently a simple task; however, a person can fall from walking or standing or while moving inside home or in a hospital. A common situation is the fall from standing on support, like ladders as an example, that could cause severe injuries and needs in many cases an immediate intervention from the carers. Moreover, especially in the care of elderly or impaired people, the falls from sleeping or lying in a bed and falls from sitting on a chair should be carefully addressed [17].

The use of cameras in home care and assistive infrastructures has widely increased, thanks to the little invasiveness (especially if compared to wearable sensors), and also because of the higher precision when compared to ambient technologies [10]. The fall detection system can be arranged with one or multiple cameras, involving also moving devices [17]. Spatio temporal analysis has been studied in [1,22], where shape modeling is performed to detect the event. Inactivity/change of shape using the information built in contextual models can be exploited to analyze human behavior and detect anomalies [11,19]. In this context an analysis in the compressed domain is considered by the authors in [15], combining global motion estimation and local motion clustering. Also posture has been used for fall detection in [9], with very high accuracy, using posture maps learned on a set of training sequences. Head position analysis has been instead considered for a three dimensional environment. The principle that considers faster vertical motion than horizontal during a fall event is applied. Thresholds should be introduced to distinguish falls from other events [12].

Although a lot of effort has been spent in the design of fall detection systems, there is plenty of room for efficient and robust development in this research direction, to achieve real time analysis, a fundamental step to achieve an effective action in case of need.

3 Motivation and Contribution

As can be noticed from the state of the art, algorithms for fall detection operate in many cases in the pixel domain, whereas most of the surveillance cameras only provide the video in the compressed domain. In order for these algorithms to be applied, the video has to be decoded, introducing an additional processing layer. Furthermore, most algorithms are not operating in real time, barely reaching 20-25 frames per second on a PC-based platform, which hampers the ability of their deployment in real scenarios. In order to respond to this need, especially in case of elderly care, it is necessary to develop low-complexity algorithms, which can deployed directly in the DSP (Digital Signal Processor) onboard of the camera and possibly in the compressed domain, thus dropping the need for decoding. In this paper we present an algorithm which completely operates in the compressed H.264 [25] domain and that requires a negligible complexity, hence it can be deployed on DSP (or similar) processor. Fall detection and reconfiguration is achieved by proposing a generic entropy measure derived using the distribution of the motion field extracted from the compressed video bit stream.

4 Framework Description

4.1 Motion Descriptors

In order to measure and monitor the movement of the objects in the camera view, we propose a descriptor based on the disorder, or entropy, of the motion vectors of the video. The standard for video coding H.264, as most of its predecessors, achieves compression through a block-based algorithm, where blocks have variable size from 4×4 to 16×16 pixels [25]. Motion vectors are calculated for individual blocks in order to remove the temporal redundancy of the video. The distribution of the motion vectors throughout the frames gives us a very accurate insight about the analytics of the video, since it tends to exhibit more disorder whenever there is any moving object in the video frame.

Fig. 1. Motion vectors extracted from a frame of the fall dataset. The red arrows highlight the regions in which the motion field exhibits strong disorder.

An example is shown in Figure 1, and represents a frame in a fall sequence captured from a static camera; motion vectors are overlaid on the picture. As can be seen, the motion vectors show a zero value along most of the video frame, as it is expected in case of moving camera. However, the motion vectors distribution at the edges of the person about to fall tends to have higher disorder. We propose to exploit this aspect in order to measure the amount of information in the video frame and also to use it for the detection of fall events.

4.2 Motion Entropy Measure

As mentioned in the previous section, we choose to operate in the compression domain to achieve real time operational capabilities. Motion vectors are chosen as the main features for analysis, as they are immune to changes in bit-rate and quantization parameters (QP) of the encoded H.264 video stream. The disorder in the motion field represents the information content in the video. In H.264, standard motion vectors are computed at 4×4 and the block size is based on the observed variance. Each motion vector consists of two components representing distances in pixels along X and Y direction from the best match found in the reference frame. In this context we represent the pixel difference along X and Y as $MV_x(i,j)$ and $MV_y(i,j)$, respectively, where i and j represent the location of a 4×4 block in the video frame. After reading the motion vectors from the H.264 stream, we group $MV_x(i,j)$ and $MV_y(i,j)$ into a 8×8 matrix, therefore each of these blocks represents the motion vectors of a region corresponding to an area of 32×32 pixels. On these super-blocks the 8×8 DCT (Discrete Cosine Transform) is performed according to Eqs. (1) and (2).

$$MD_x^{(c,d)}(a,b) = [\frac{1}{4}\sum_{a=0}^{7}\sum_{b=0}^{7} MV_x[(c-1)*8+a,(d-1)+b]*\cos\frac{(2a+1)*\pi}{16}*\cos\frac{(2b+1)*\pi}{16}] \quad (1)$$

$$MD_y^{(c,d)}(a,b) = [\frac{1}{4}\sum_{a=0}^{7}\sum_{b=0}^{7} MV_y[(c-1)*8+a,(d-1)+b]*\cos\frac{(2a+1)*\pi}{16}*\cos\frac{(2b+1)*\pi}{16}] \quad (2)$$

After the transform, each block describes the motion pattern of the 32×32 pixel region in X and Y directions, respectively, which becomes our motion descriptor. In the equations (c,d) represent the block location of 32×32 pixels in the frame, (a,b) represent the location of the 4×4 block within the 32×32 block.

The choice for a block size of 32×32 pixels is made to ensure minimum variability of motion vectors which occurs in the case of 16×16 mode in H.264 bit stream. The result is a 2D DCT transform of 8×8 blocks of motion vectors. Inferring from the properties of the DCT transform we can notice that DC values $MD_x^{(c,d)}(0,0)$, $MD_y^{(c,d)}(0,0)$ represent the localized global motion and AC coefficients represent the variation in motion vectors. The frequency of variation increases as we move towards the bottom-right corner. We propose to accumulate the AC coefficients to arrive at a measure of motion disorder. However, higher frequencies represent more disorder in comparison to the lower ones, hence the accumulation has to be done in a weighted manner. This is exactly the opposite of what happens in image and video compressions, where lower frequencies are usually more important. Therefore, we calculate the entropy values along X and Y as E_X and E_Y from the equations Eq. (3) and Eq. 4, respectively.

$$E_X(c,d) = \sum_{a=0}^{7}\sum_{b=0}^{7} MD_x^{(c,d)}(a,b) * [2^{a-8} + 2^{b-8}] \quad (3)$$

$$E_Y(c,d) = \sum_{a=0}^{7}\sum_{b=0}^{7} MD_y^{(c,d)}(a,b) * [2^{a-8} + 2^{b-8}] \tag{4}$$

The aggregated entropy gives us a generalized measure of information present in the video frame in case of P (predicted) macroblocks for which the motion vectors exist. However in presence of very rapid motion, which cannot be covered by motion search algorithm of the H.264, the macro blocks are typically classified as *intra*. In order to calculate the entropy measure for these blocks we utilize the number of bits that a particular macro block requires. As we know by definition, the H.264 encoder is basically a sparse encoder which assigns a variable number of bits to every macro block based on its information content. Such an assumption is perfectly valid according to Shannon's information theory.

The entropy measures that we use in the proposed method are reported hereafter:

$$E_X(c,d) = K_x * N_b(c,d) \tag{5}$$

$$E_Y(c,d) = K_y * N_b(c,d) \tag{6}$$

$$G_{XY} = \sum_{c=0}^{\frac{Width}{32}} \sum_{d=0}^{\frac{Height}{32}} [E_X(c,d) + E_Y(c,d)] \tag{7}$$

where K_x and K_y are weighting factors and $N_b(c,d)$ is the numbers of bits for that particular macro block.

4.3 Object Detection and Segmentation

E_X and E_Y are the measure of the extent of disorder for $MV_x(i,j)$ and $MV_y(i,j)$ in case of P macro blocks, and they are considered as measure of information for intra macro blocks. Since we have defined a quantitative measure for the disorder, we now have to identify the blocks, which have high E_X and E_Y. Initially both E_X and E_Y across the frame are contributing to obtain a frame level metric for disorder, as shown in Eq (7). Then the steps shown in Algorithm 1 are followed.

The algorithm iteratively checks the values for E_X and E_Y for each block against the threshold, which varies from 100% to 50% of their respective mean values. If both conditions are met, the block is selected as a contour block and its contribution $(E_X + E_Y)$ is accumulated. The algorithm is terminated when the ratio between the disorder of the contour region and G_{XY} reaches the value K, or in case the adaptive threshold decreases beyond 0.5. In this way only the blocks with significant motion along X and Y are identified. The value of K is a user-defined parameter, and determines the size of the contour around the moving object. High values will result in extended contours around the moving objects, while low values will shrink the thickness of the contour around the object. This parameter is data dependent and should be adjusted to fit the scenario requirements.

```
input  : Entropy Measures E_X and E_Y
input  : Global Disorder Measure G_XY
input  : Segmentation Measure K
output: Contour region CR

E_X ;                                    % Variation metric for MV_x
E_Y ;                                    % Variation metric for MV_y
G_XY ;                 % Combined measure of disorder for whole frame
CR = φ ;                                 % Union of contour blocks
C = 1 ;                                                  % Gradient
Buffer = 0 ;                                      % buffer variable
while Buffer <= K * G_XY do
    for i ← 1 to width/32 do
        for j ← 1to Height/32 do
            if E_X(i, j) > C * mean(E_X)&&E_Y(i, j) > C * mean(E_Y)&&Buffer <=
            K * G_XY then
                CR = CR ⋃ Region (i,j) ;
                Buffer = Buffer + E_X (i,j)+E_X (i,j) ;
            end
        end
    end
    C = C − 0.1 ;
    if C <= 0.5 then
        Break ;
    end
end
Return CR;
```

Algorithm 1. Identification of contour blocks

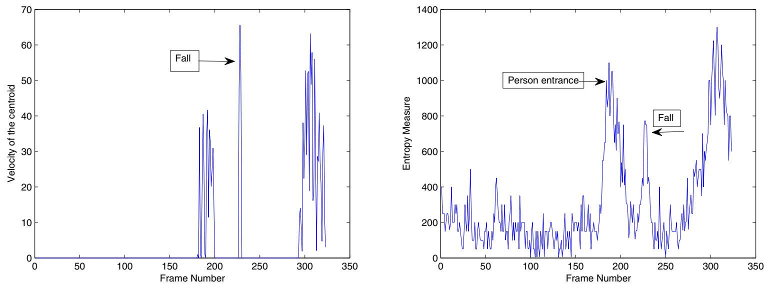
To further refine the extracted information, a 3×3 majority filter is applied across the whole frame; blocks having at least 4 neighbouring blocks, labeled as showing a high level of disorder, are classified as part of the moving objects.

5 Proposed Method

5.1 Fall Detection

In the section above we have defined the motion descriptors and their usage for moving objects detection and segmentation. We approach fall detection in a similar manner. Fall detection can be described as a sudden event, which causes rapid variation of video features in the temporal domain. In line with this observation, the unified entropy measure defined earlier, also exhibits large values and also strong variations across the frames, in presence of the fall.

In order to identify potential candidate frames for the occurrence of fall, we discard the frames which have lower entropy. Lower entropy frames typically do not contain any object motion, hence the likelihood of occurrence of the fall in these frames is almost negligible. After selecting the frames with higher entropy with a cut off, which is specific to camera orientation and illumination conditions, we further analyze these frames for the detection of the fall. Figure 2 (a) shows the movement of centroid of the person per frame. As we can see, during the fall the velocity of centroid dramatically increases and then goes to zero.



(a) (b)

Fig. 2. (a) Velocity of the centroid of the person per frame. (b) Variation of entropy and the events as marked using ground truth.

Another feature that characterizes a falling event, is the sudden change in location and motion orientation of the centroid over a very short number of frames. In order to further refine the accuracy of the prediction, we also consider that after the fall occurs, the amount of motion reduces, there by decreasing the entropy measure defined in the previous section. Figure 2 (b) illustrates the variation of entropy. We can notice that any sudden event results in a spike in the entropy plot, followed by a decrease of the values after the fall.

5.2 Algorithm

Let the frame at the time instant i for a given video stream from a camera be F_i, then the entropy measure of that frame is given by $E(F_i)$ and the centroid of the segmented person as a pixel location in a video is given by $C(F_i)$, and the distance traveled by the centroid (or velocity per frame) is given by:

$$V(F_i) = Euclidean(C(F_i), C(F_{i-1})) \quad (8)$$

The proposed algorithm is shown in Figure 3. As mentioned in the previous paragraphs, we first check for the high variance in the entropy measure, the velocity of the moving object. After that we check for sudden drop of entropy in the neighborhood to check for the fall. Thresholds $Th1$, $Th2$ and $Th3$ depend on the mean and variance of entropy and velocity. The main purpose of these thresholds is to detect the peaks that occur in the entropy measure and also to

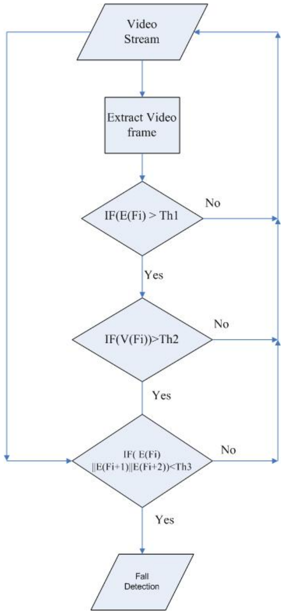


Fig. 3. Flow chart of the proposed algorithm

detect the peak changes in velocity of the centroid. Ideally, any value falling above the mean of the entropy and velocity should be considered; however, background noise also contributes significantly to motion entropy and is largely dependent on the deployed environment.

5.3 Reconfiguration

Reconfiguration of the camera is triggered by the fall detection algorithm mentioned in the above sections. The basis for reconfiguration is the fallen person. The main aim of the reconfiguration of the camera is to get the best possible view of the subject. In order to do so, we adjust the camera parameters in such a manner that the person to be observed falls at the centre of the image plane. Further precaution is also taken so that the person does not fill the entire image plane of the camera. This can be achieved in most PTZ cameras by specifying the particular area in a video frame, by using the available CGI (Commond Gateway Interface) commands. In this scenario, the segmented object is selected as the area of interest. Cameras automatically adjust alignment at the midpoint of the area specified. After the reconfiguration is complete in order to make the person fully visible, the new parameters of the camera are set using the subject segmentation information. This helps understanding the reason of fall and further monitoring of the person after the fall, that is especially relevant for elderly people living alone and monitored for their care.

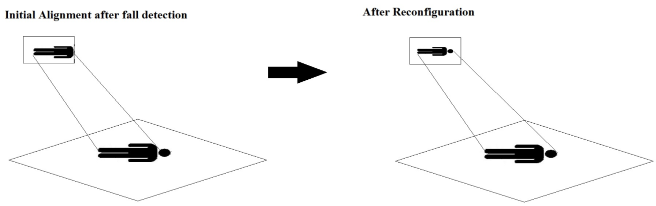


Fig. 4. Reconfiguration of the camera carried out to guarantee full visibility of the subject of interest

6 Evaluation

In order to demonstrate the utility and robustness of the algorithm, we first evaluate the performance of the fall detection algorithm by testing it against the reference fall detection dataset published by the University of Montreal [2], widely used to validate algorithms in this field.

To show the reconfiguration capability, we deployed a set up in a real environment and observe its performance during the occurrence of fall. To this extent we used two cameras "Sony SNC-EP521 indoor", day/night, with PTZ. These IP cameras are equipped with a 36x optical zoom allowing operators to cover large, open areas and zoom in for detailed close-up shots. Panning can span from 0 to 340 degrees, with max 105 degrees tilt, and their configuration can change using built in network commands. The cameras have been installed in our Department facility, and falling events have been recorded thanks to the collaboration of volunteers.

6.1 Fall Detection

Since the algorithm operates in the compressed domain, we had to convert all the videos in the dataset [2] into the H.264 format using the JM H.264 reference encoder [13], at the frame rate of 25 frames per second. The thresholds necessary for a proper operation of the algorithm are learned for each camera and are maintained constant for that particular camera for all scenarios. Fall is defined as an event lasting 5-10 seconds, starting from the momentary stop by the subject just before the fall and ending with a motion less layover of the subject. The total number of correct fall detections, as compared to the ground truth, are deemed as true positives (TP), while false detections are termed as false positives (FP). Finally, true falls which have been skipped by the detector are termed as false negatives (FN). The results obtained for the video dataset are given in Table 1 in terms of Precision, Recall and F-Measure. A comparison with respect to the state of the art techniques is provided in 2. As can be seen, the fall detection algorithm performs reasonably well especially given the fact that it operates in real time. The algorithm fails to detect the falls, when the subject is very far

(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Fig. 5. Fall detection and subsequent reconfiguration of the camera for better view

away from the camera and subsequently the motion entropy generated by the subject is very low. In such scenario noise becomes dominant thereby causing false detections. Another scenario where the algorithm fails is in case of actions, which correspond to bending down on the floor etc. However, since we also took into consideration the momentary fall entropy, just after the fall most of such false detections have been resolved.

6.2 Comparison

Our algorithm completely operates in the compressed domain. Hence it has the advantage of being very light in terms of computational and memory requirements. Nevertheless it compares very well with the other pixel domain state of the art fall detection methods as we can see from the table 1. Our method also provides a significant improvement with respect to other compressed domain methods like [15]. Most of these methods rely on the segmentation of moving object and the trajectory of its centroid, and also include other features like velocity of centroid. Present algorithm also uses these aspects, but it turns out to be more robust as it also exploits the motion disorder as one of the factors to determine fall detection. Furthermore, the compressed domain method presented in [15] uses AC and DC coefficients along with motion vectors to achieve object segmentation, which are heavily dependent on the quantization parameter used for encoding the video bit stream. The proposed method, instead is entirely based on motion vectors, which are independent with respect to changes in QP. In terms of complexity our solution offers the lowest complexity of all compressed domain methods as it operates at the level of 32×32 blocks, and the number of operations required for processing one frame are 5.2K, 16K, 48K, 106K computations for CIF, VGA, HD, full HD resolutions, respectively.

Table 1. Performance of the algorithm on the dataset [2]

Precision	Recall	F-Measure
0.89	0.86	0.88

Table 2. Comparison to the state of the art approaches described in [10]

	Our method	K-NN	C4.5	SVM	Bayes	Feng et. all
Sensitivity	0.86	0.75	0.85	0.95	0.80	0.98

6.3 Reconfiguration

In case of real evaluation the video stream obtained from the camera has a resolution of 720 $\times$ 576 pixels and a frame rate of 25 frames per second. The H.264 bit stream obtained from the camera is encoded in the baseline profile. In order to access the Network Abstraction Layer (NAL) packets from the camera we have used the functions available in the *ffmpeg* library [8]. Fall detection and moving object segmentation are implemented using the motion vectors extracted from the H.264 (JM 18.6 version) decoder [13]. In order to control the camera automatically the *curl* library functions [7] are adopted. The whole set up is implemented on an Intel i5 processor, 3.10 GHz.

Fall detection and subsequent reconfiguration is shown in Figure 5. As we can see from the images, fall of the person occurs towards the end of the image in one of the frames. However, camera instantly reconfigures to bring back the view of the fallen person. This shows that the algorithm works in real time and is robust enough to work in tricky illumination conditions.

7 Conclusions

In this paper we proposed a framework for elderly care for behavior anomaly detection from video. The analysis is performed using the information coming from the video motion flow, directly from streaming, without decoding. This allows the real time analysis of the event, introducing the possibility of implementing the proposed solution directly on-board of cameras. The described technique has been tested on a publicly available dataset, verifying its ability in detecting fall events, and compared with state of the art methods. Moreover it has been tested on a set-up developed by authors, where, besides the fall detection only, it has been proved its capability in performing real-time camera reconfiguration, in order to focus the attention of the vision system on the fallen person. The proposed framework is an instrument able to preserve the privacy of the persons monitored, since no information should be decoded and transfered before an event is detected.

References

1. Anderson, D., Keller, J.M., Skubic, M., Chen, X., He, Z.: Recognizing falls from silhouettes. In: Engineering in Medicine and Biology Society, 2006. EMBS 2006. In: 28th Annual International Conference of the IEEE, pp. 6388–6391. IEEE (2006)
2. Auvinet, E., Rougier, C., Meunier, J., St-Arnaud, A., Rousseau, J.: Multiple cameras fall dataset. DIRO-Université de Montréal, Tech. Rep 1350 (2010)
3. Borges, P., Conci, N., Cavallaro, A.: Video-based human behavior understanding: A survey. IEEE Transactions on Circuits and Systems for Video Technology **23**(11), 1993–2008 (2013)
4. Cardinaux, F., Bhowmik, D., Abhayaratne, C., Hawley, M.S.: Video based technology for ambient assisted living: A review of the literature. Journal of Ambient Intelligence and Smart Environments **3**(3), 253–269 (2011)
5. Chen, C.H., Yao, Y., Page, D., Abidi, B., Koschan, A., Abidi, M.: Heterogeneous fusion of omnidirectional and ptz cameras for multiple object tracking. IEEE Transactions on Circuits and Systems for Video Technology **18**(8), 1052–1063 (2008)
6. Climent-Pérez, A., Flórez-Revuelta, P., Chaaraoui, F.: A review on vision techniques applied to human behaviour analysis for ambient-assisted living. Expert Systems with Applications **39**(12), 10873–10888 (2012)
7. Open source multiple contributions, O.S.: Command line tool for transferring data with url syntax, March 2014. http://curl.haxx.se/
8. Open source multiple contributions, O.S.: Trans standard multimedia framework for media manipulation, March 2014. http://www.ffmpeg.org/
9. Cucchiara, R., Grana, C., Prati, A., Vezzani, R.: Probabilistic posture classification for human-behavior analysis. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans **35**(1), 42–54 (2005)
10. Feng, W., Liu, R., Zhu, M.: Fall detection for elderly person care in a vision-based home surveillance environment using a monocular camera. Signal, Image and Video Processing, pp. 1–10 (2014)
11. Foroughi, H., Aski, B.S., Pourreza, H.: Intelligent video surveillance for monitoring fall detection of elderly in home environments. In: 2008 11th International Conference on Computer and Information Technology. ICCIT 2008, pp. 219–224. IEEE (2008)
12. Hazelhoff, L., Han, J., de With, P.H.N.: Video-based fall detection in the home using principal component analysis. In: Blanc-Talon, J., Bourennane, S., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2008. LNCS, vol. 5259, pp. 298–309. Springer, Heidelberg (2008)
13. HHI.: H.264 reference decoder from heinrich hertz institute, January 2014. http://iphome.hhi.de/suehring/tml/
14. Katz, S., Downs, T.D., Cash, H.R., Grotz, R.C.: Progress in development of the index of adl. The gerontologist 10(1 Part 1), pp. 20–30 (1970)
15. Lin, C.W., Ling, Z.H.: Automatic fall incident detection in compressed video for intelligent homecare. In: 2007 Proceedings of 16th International Conference on Computer Communications and Networks. ICCCN 2007, pp. 1172–1177. IEEE (2007)
16. Micheloni, C., Rinner, B., Foresti, G.L.: Video analysis in pan-tilt-zoom camera networks. IEEE Signal Processing Magazine **27**(5), 78–90 (2010)
17. Mubashir, M., Shao, L., Seed, L.: A survey on fall detection: Principles and approaches. Neurocomputing **100**(0), 144–152 (2013). (Special issue: Behaviours in video)

18. Murray, D., Basu, A.: Motion tracking with an active camera. IEEE Transactions on Pattern Analysis and Machine Intelligence **16**(5), 449–459 (1994)
19. Nait-Charif, H., McKenna, S.J.: Activity summarisation and fall detection in a supportive home environment. In: 2004 Proceedings of the 17th International Conference on Pattern Recognition. ICPR 2004. vol. 4, pp. 323–326. IEEE (2004)
20. Quaritsch, M., Kreuzthaler, M., Rinner, B., Bischof, H., Strobl, B.: Autonomous multicamera tracking on embedded smart cameras. EURASIP Journal on Embedded Systems **2007**(1), 35–35 (2007)
21. Rashidi, P., Mihailidis, A.: A survey on ambient-assisted living tools for older adults. IEEE journal of biomedical and health informatics **17**(3), 579–590 (2013)
22. Rougier, C., Meunier, J., St-Arnaud, A., Rousseau, J.: Fall detection from human shape and motion history using video surveillance. In: 21st International Conference on 2007 Advanced Information Networking and Applications Workshops, AINAW'07. vol. 2, pp. 875–880. IEEE (2007)
23. Scotti, G., Marcenaro, L., Coelho, C., Selvaggi, F., Regazzoni, C.: Dual camera intelligent sensor for high definition 360 degrees surveillance. IEE Proceedings-Vision, Image and Signal Processing **152**(2), 250–257 (2005)
24. Van Kasteren, T., Englebienne, G., Krse, B.: An activity monitoring system for elderly care using generative and discriminative models. Personal and Ubiquitous Computing **14**(6), 489–498 (2010)
25. Wiegand, T., Sullivan, G.J., Bjontegaard, G., Luthra, A.: Overview of the h. 264/avc video coding standard. IEEE Transactions on Circuits and Systems for Video Technology **13**(7), 560–576 (2003)
26. Yu, X.: Approaches and principles of fall detection for elderly and patient. In: 2008 10th International Conference on e-health Networking, Applications and Services. HealthCom 2008, pp. 42–47, July 2008