SturzVis

SturzVis or The MuBisA project, as it is called on the CogVis webpage, is one of the research projects of CogVis. It's goal is to detect elderly people in case they fall down and are unable to get up again. Every year about 1000 people over the age of 60 die because they collapse in their home. 30% of people over the age of 60 live alone. The MuBisA project's goal is to detect falls and alert helpers.

Figure 1 shows the project setup. Every room of the house is equipped with up to four wireless low-price webcams. Multiple webcams are needed to avoid occlusions and to be able to use 3D-based analysis algorithms. Parts of the preprocessing can be done in the cameras, if they have these features, or it's done by a central computer, which will do all the processing of the image sequences and calculate the fall detection. If an alarm is triggered an alert signal is sent to a designated assistant or a call center via SMS. Due to the fact that privacy is a big concern in this project, all the processing will be done fully automatically.

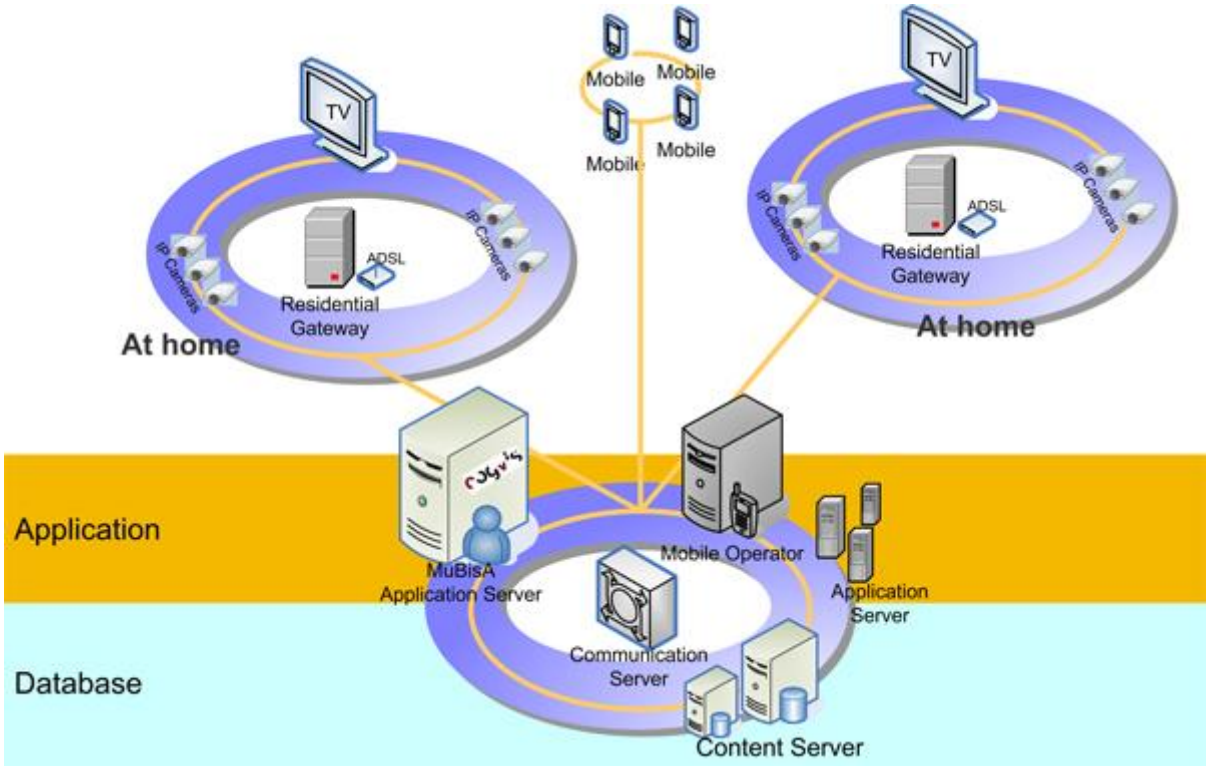

figure 1 project setup - Quelle: [1]

There are two approaches to detect a fall [2] (figure 2). The first approach is called Early Fusion. In the Early Fusion the frames of the 4 cameras are combined before the calculation. You get a 3D reconstruction of the human and the posture estimation is robust. To get the 3D reconstruction Shape-from-Silhouette is used, because the technique can be applied to

the motion images from the cameras directly and its result are good enough to detect a fall. The drawback is that the cameras have to be calibrated and the computation is more complicated.

The second approach is called Late Fusion. In this approach the posture extraction is done for every camera separately and a voting algorithm combines the results of every camera. No camera calibration needs to be done with this approach. Drawbacks of this method are problems with occlusions and falls that happen in the direction of the optical axis of some cameras. In such cases a camera is not able to detect a fall and misqualifications can happen. To avoid these problems, cameras can get the right to detect a fall on their own, not evaluating the results of the other cameras.

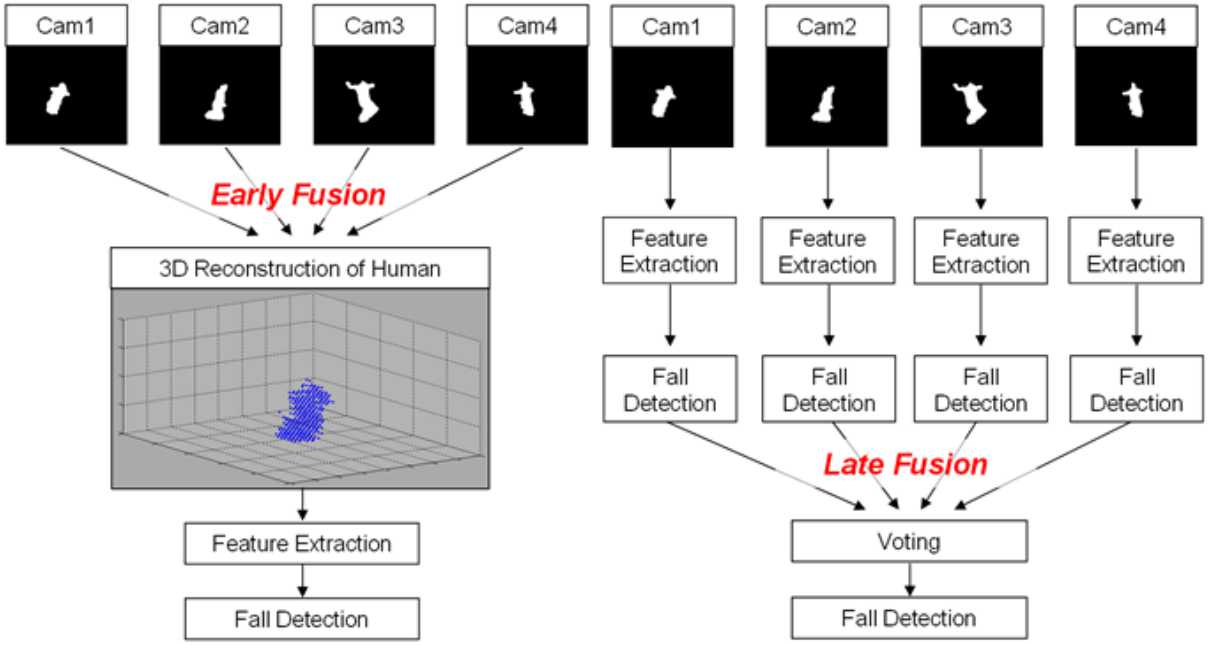

figure 2 Two ways of calculating the fall - Quelle: [1]

A short introduction of the basic processing steps of the fall detection algorithm is given in the following section. In figure 3 the workflow of the fall detection algorithm is shown.

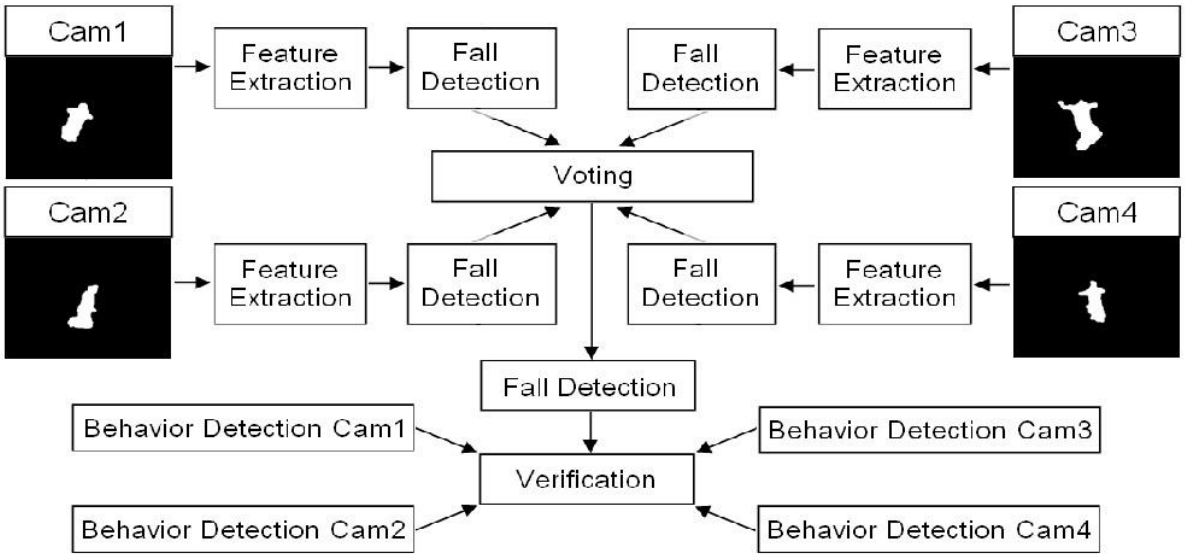

figure 3 workflow of the application - Quelle: [4]

The algorithm has two parts to detect the fall. The first part consists of the extraction of features in every frame and calculating possible posture states. The second part is the computation of a so called accumulated hitmap.

The first part starts with the segmentation of a person. Therefore a background model is calculated. In the MuBisA project the Color Mean and Variance approach is used[3]. For motion detection an approach described in [4] is used. For tracking the algorithm uses the distance of blobs in two consecutive frames computed by the centre of mass. The result of the segmentation is a silhouette of the human, so every frame has a set of pixels, which describe the human silhouette.

The next step is to estimate the posture. Three possible posture states are used in the project: "standing", "in between" and "lying". To get these states, three features are extracted out of every frame [5]:

- Bounding Box Aspect Ratio: The height of the bounding box surrounding the person by its width.
- Orientation: The orientation of the major axis of the ellipse fitted to the person, specified as the angle between the major axis and the x-axis.
- Axis Ratio: The ratio between the lengths of the major axis and the minor axis of the ellipse fitted to the person.

For every feature value every posture state gets a confidence value between 0 and 1. The sum for all three postures is 1. A membership function of the Orientation feature is shown in

figure 4. These confidence values are then summed up for all the features.

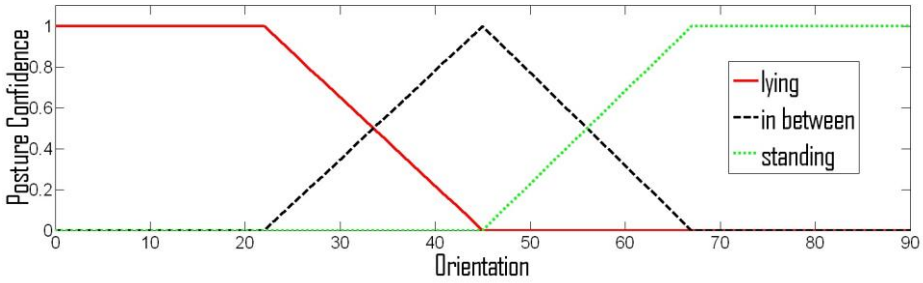

figure 4 Membership function of the Orientation feature. Source: [1]

A fall is characterized as a fall event followed by a long phase of the posture state lying. To get the fall event the motion speed of the person is taken. A fall is characterized as a high motion speed of the body of a person. The motion speed is the amount of change between to frames.

The results of the posture estimation are combined with so called learned accumulated hitmaps. These hitmaps are matrices with a value for every pixel in the picture. The value is increased by one, every time a pixel is a foreground pixel and decreased, if the pixel is no foreground pixel anymore. So the hitmap shows areas with long abidance of foreground pixels. The hitmap is trained in a training stage and then used in the classification stage to detect unexpected behavior.

The combination of the posture estimation and the hitmaps is used to determine a fall.

[1] http://www.cogvis.at/mubisa/projectsetup.html

[2] S. Zambanini, J. Machajdik, and M. Kampel. Early versus late fusion in a multiple camera network for fall detection. In *34th Annual Workshop of the Austrian Association f. Pattern Recognition (¨OAGM 2010)*, pages pp. 15–22, May 2010.

[3] Wren, C., Azarbayejani, A., Darrell, T., Pentland, A.: Pfinder: Real-time tracking of the human body. PAMI 19 (1997) p.780-785

[4] Kampel, M., Hanbury, A., Blauensteiner, P., Wildenauer, H.: Improved motion segmentation based on shadow detection. 6 (2007) 1-12

[5] Zweng, A., Zambanini, S., Kampel M.: Introducing a Statistical Behavior Model into Camera-Based Fall Detection. Lecture Notes in computer Science, 2010, Volume 6453/2010, p. 163-172