Fusion of Infrared Image and Visible Image for Fall Detection Base on Discriminant Feature

Zong-Huei, Chen[1], Mei-Yung, Chen[2]
National Taiwan Normal University, Taipei, Taiwan

Abstract

Fall detection is a very important application in surveillance systems. Therefore, it is very important to build a complete fall detection system. However, this is not suitable for environments with weak light sources since the most of the detection lenses are in a sufficient light source. Thus, in this paper, we introduce a fall-detection system based on the fusion image. It can not only solve the problem that the human cannot be detected under a dim background, but also locate the human more accurately. We use an open source deep convolutional neural network (CNN)-based approach named OpenPose to extract the discriminant features which let us to build the human centerline. Since OpenPose will detect and focus on the human activity, the other heat sources cannot be detected. In other words, OpenPose can help us remove the things that we do not interested. Finally, we define the human falls by detecting the variation of centerline. The experimental result shows that the proposed system can be applied in both bright and dark environments. Besides, the detection of the fall has 99% accuracy, higher than the other methods which use the visible image and the infrared image. The accuracy can guarantee the performance.

Keywords: fall detection, Image fusion, OpenPose, discriminant feature

1. Introduction

Falls are considered to be the most dangerous behavior among the elderly. The report indicates that falls are the second leading cause of accidental death among the elderly, with as many as 424,000 deaths in 2007. However, if appropriate countermeasures are taken, such as fall detection or fall prevention, it is considered to have positive help for the quality of life of the elderly [1]. Therefore, this study focuses on the former, responding as soon as possible when the elderly fall, so that they can get the assistance they deserve more quickly.

The fall accident might occur anytime and anywhere, especially in the dark environment. Most of the detect method base on the visible images singly [2], [3], [4]. However, they cannot detect in both bright and dark environments.

In the light of this, thermal sensors and IR sensors are good candidates [5], [6]. The above method is not limited to the strength of the light source. However, the detection area must be kept as empty as possible. It can’t allow the additional heat source simultaneously.

Consider the above situation, using fusion images to detect can combine the advantages of visible images and thermal images. The message in the fusion images can let us catch the human body in the dark environment, and help us to remove the target that we don’t interested. Therefore, developing fall detection systems with fusion images becomes a promising research direction.

This paper proposes a novel detection method for human falls. This method processes every frame captured by monitoring, which is to use the OpenPose skeleton extraction algorithm to obtain the discriminant features data of people on the screen. According to the human body centerline angle with the ground, these determine the conditions to identify falling behavior.

The remainder of this paper is organized as follows: Section II reviews the current methods of fall detection. Section III details the approach. Section IV presents the results of an experiment. Section V points to conclusion.

2. Related works

Non-wearable detection is the use of sensors to detect factors in the environment to understand whether someone has fallen around, such as sound sensing [7], position tracking [8], vibration sensing [9], visible image [10], infrared image [11]. The first three items are susceptible to interference from non-target objects or the environment itself, which can lead to false detection or classification. Therefore, it is often necessary to set a special environment during the experiment; visible light images are limited by the human body and the environment, which may cause overlap effect, the body contour is more difficult to extract; although infrared thermal images can improve this problem, it is impossible to determine that the detected object is the target to be measured, and it is necessary to have the ability to exclude the remaining objects.

In 1983, Burt and Adelson proposed Laplacian pyramid as a multi-resolution representation for images, which is commonly used in image fusion [12], [13]. First, we extract the details of the visible images. Next, we use the Laplacian pyramid transform to decompose the infrared images at multiple scales. The decomposed bottom layer image and the detailed image of the visible will fuse together. Finally, we reconstruct the image successfully.

To locate the position of the human body in the lens effectively and quickly, OpenPose [14] can be an appropriate option. It was developed in Carnegie Mellon University in 2017. It is a library used for multi-person key point detection in real-time. It can also efficiently detect hand and facial key points in images.

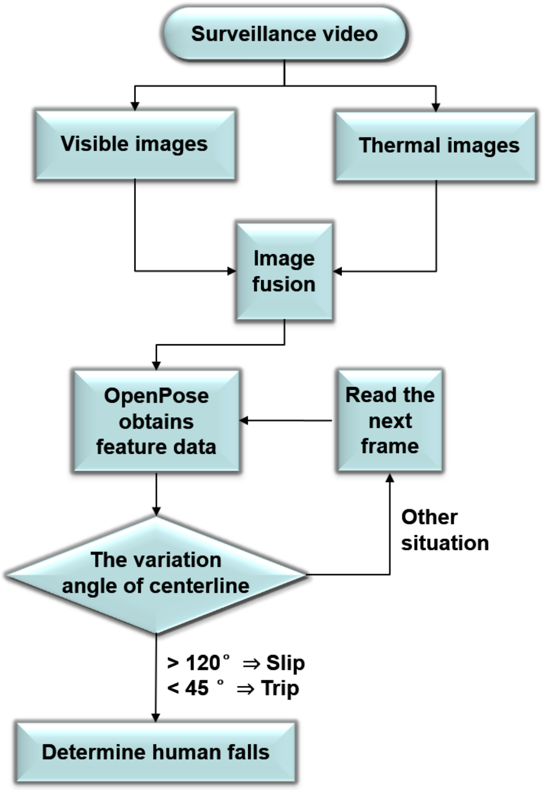
3. Methodology

In this work, we first do the image fusion, and then extract the body discriminant features by using OpenPose library. Finally, by detecting the variation of the angle, the system defines the occurrence of the fall event. The procedure of implementation of our proposed approach is as shown in Fig. 1.

3.1 Image fusion

For the same target scene, use visible light CCD sensor and infrared focal plane array IRFPA detector to collect YUV format original visible light image and 14bits original infrared image.

Figure 1: Fall Detection Block Diagram.



Next, preprocess the collected original visible image and infrared image respectively. The original infrared image preprocessing includes two-point correction, blind element compensation and median filtering. In order to improve the contrast of the infrared image, after the preprocessing, the platform histogram equalization method is used to enhance it. After the enhancement process is completed, the maximum and minimum linear mapping method is used to convert the infrared image pixel value from 14 bits data to 8 bits data. Due to the difference in the size of the visible image and the infrared image, after the infrared image enhancement processing is completed, the bilinear interpolation algorithm is used to set their sizes to be consistent. In the actual processing process, the processing of the visible image is only performed on the Y component (that is, the brightness component), and the following steps are the same.

Afterwards, register the visible image and the infrared image. Image registration is a necessary part of image fusion, which directly affects the quality of the final fused image. The visible image and the infrared image adopt a coarse registration method. Through the guarantee of optical and structural design, the imaging field of view of the visible CCD sensor and the infrared focal plane array IRFPA sensor is close to the same, that is, the coarse registration of the visible light image and the infrared image is completed.

Then, use the fusion method based on Laplacian pyramid decomposition to perform image fusion on the registered image. The process of this method is as follows: (1) Extract the details of the visible image, (2) Multi-scale decomposition of the Laplacian pyramid of the infrared image, (3) Combine the bottom layer image of the Laplacian pyramid of the infrared image with the visible image, (4) The fused image is reconstructed. Finally, display the fused image of pseudo-color.

In view of the fact that the human eye can distinguish only dozens of gray levels, but can distinguish as many as thousands of colors, this technology performs pseudo-color processing on the image after the fusion of the visible light image and the infrared image to further enhance it. Visual effects. Pseudo-color processing uses a grayscale-color conversion method, which converts the grayscale image into a color image by establishing the mapping relationship between the grayscale of the grayscale image and various colors in the color space. Fig. 2 shows the result of the image fusion.

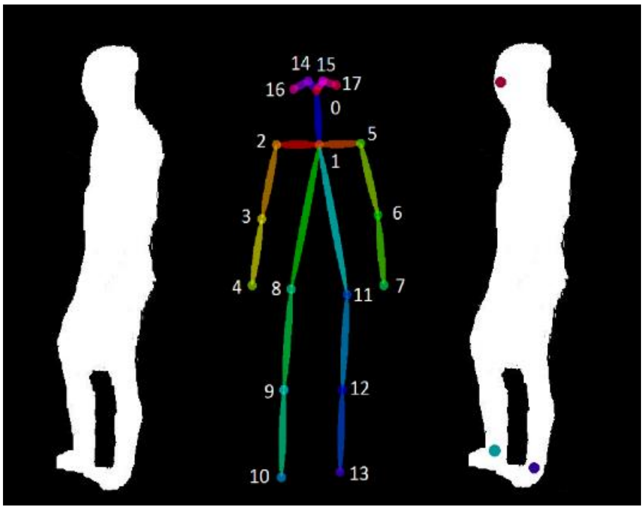
Figure 2: The example of image fusion: (a) visible; (b) infrared; (c) fusion.

(a) *(b)* *(c)*

3.2 Extract the discriminant data with OpenPose

The raw data are extracted from the video through OpenPose. However, we just care about the nose, left ankle, and right ankle, as shown in Fig. 3.

Figure 3: Extract the discriminant data with OpenPose.



For the convenience of representation, $S = \{S_0, S_1, \cdots, S_{13}\}$ represents the joint position set. We define the position of the node j at time t as $S_j(t) = (x_{tj}, y_{tj})$, $j \in \{0, 1, \cdots, 13\}$.

3.3 Define the centerline

In the process of falling, the most obvious feature of the human body is the body tilt, and tilt degree will continue to increase. In order to reflect the characteristics of the body's continuous tilt in the process of human fall, a human centerline F is defined in this paper. Let the midpoint of left ankle S13 and right ankle S10 be S, and the connection of midpoint S and nose S0 is the centerline F of the human body. Fig. 4 shows these defined elements.

Figure 4: The discriminant features. The white point is S (nose), and the black line is F (centerline).

3.4 Determine human falls

As shown in Fig. 5, θ is the angle between the centerline of the human and the ground. Through OpenPose, the data of feature points 0, 10 and 13 are $S_0(t) = (x_{t0}, y_{t0})$, $S_{10}(t) = (x_{t10}, y_{t10})$ and $S_{13}(t) = (x_{t13}, y_{t13})$ respectively. So

$$S = \frac{S_{10}+S_{13}}{2} \ , \tag{1}$$

$$S(t) = (x_t, y_t) \ . \tag{2}$$

At time t, the angle between the centerline of human body and the ground is

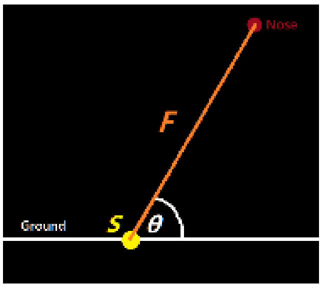
$$\tan^{-1}\left|\frac{y_{t_0} - y_t}{x_{t_0} - x_t}\right| . \tag{3}$$

when $\theta < 45°$ or $\theta > 120°$, $C = 1$. Otherwise, $C = 0$. It can be considered as satisfying the occurrence of the fall event.

$$C = \begin{Bmatrix} 1; \theta < 45° \text{ or } \theta > 120° \\ 0; \text{otherwise} \end{Bmatrix}. \quad (4)$$

Figure 5: (a) the angle between the centerline of the body and the ground; (b) the falling process

(a)



(b)

4. Experimental result

In this section, some experimental results will be provided to prove the effectiveness of the proposed fall detection method. This experiment was performed on a classroom aisle using visible images and thermal images as input images. 40 videos (including 50 human falls in the bright environment and 50 human falls in the dark environment) were used to evaluate the effectiveness of the method. The final results are summarized in Table I. As shown in Table 1, the accuracy in bright environment is very high. However, the performance of the detection base on the visible image in the dark environment is poor. It is only 4%. When we use fuse image, the final average detection rate is 99%, which illustrates the effectiveness of the proposed method applied to fall detection methods.

Table 1: Final detection rate

Human falls	Source of feature extraction		
	Visible image	Infrared image	Fused image
In bright environmet (50 videos)	76% (38/50)	86% (43/50)	100% (50/50)
In dark environmet (50 videos)	4% (2/50)	64% (32/50)	98% (49/50)
Average	40%	75%	99%

5. Conclusion

In this study, we introduce a fall-detection system based on discriminant features of fuse image. We use these extracted features to easily identify the behavior of falling. The experimental results confirmed that the proposed method was better than a single infrared thermal image or visible image. The proposed method performs better than the others, and it gets the high accuracy by 99%, proving the effectiveness of discriminant features applied to fall detection methods, especially in the dark environment.

References

For each work shown in the list of references must be a reference in the text. All citations in the text and all references must meet APA styles (American Psychological Association 7th edition – more information http://www.apastyle.org/).

[1] Kabalan Chaccour, Rony Darazi, Amir Hajjam El Hassani, Emmanuel Andrès, "From Fall Detection to Fall Prevention: A Generic Classification of Fall-Related Systems," *IEEE Sensors Journal*, 17(3):812– 822, 2017.

[2] X. Wang and K. Jia, "Human Fall Detection Algorithm Based on YOLOv3," *2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC)*, Beijing, China, 2020, pp. 50-54.

[3] J. Thummala and S. Pumrin, "Fall Detection using Motion History Image and Shape Deformation," *2020 8th International Electrical Engineering Congress (iEECON)*, Chiang Mai, Thailand, 2020, pp. 1-4.

[4] M. Inoue and R. Taguchi, "Bed Exit Action Detection Based on Patient Posture with Long Short-Term Memory," *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, 2020, pp. 4390-4393.

[5] L. Yu, H. Chen, H. He, H. Nie, X. Zhai and B. Xiong, "A Fall Detection System Based on a Thermopile Imaging Array and a Back Projection Algorithm," *2020 IEEE International Conference on Electro Information Technology (EIT)*, Chicago, IL, USA, 2020, pp. 060-065.

[6] Y. Ogawa and K. Naito, "Fall detection scheme based on temperature distribution with IR array sensor," *2020 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, 2020, pp. 1-5.

[7] D. Droghini, F. Vesperini, E. Principi, S. Squartini, F. Piazza, "Few-Shot Siamese Neural Networks Employing Audio Features for Human-Fall Detection", *Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence*, pp. 63-69, 2018.

[8] P. Loncomilla, C. Tapia, O. Daud, J. Ruiz-del-Solar, "A Novel Methodology for Assessing the Fall Risk Using Low-Cost and Off-the-Shelf Devices", *IEEE Trans. Human-Machine Syst.*, vol. 44, no. 3, pp. 406-415, Jun. 2014.

[9] V. L. Zimmer, N. Sitar, "Detection and location of rock falls using seismic and infrasound sensors", *Eng. Geol.*, vol. 193, pp. 49-60, 2015.

[10] G. Debard, G. Baldewijns, T. Goedemé, T. Tuytelaars and B. Vanrumste, "Camera-based fall detection using a particle filter," *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, 2015, pp. 6947-6950

[11] J. Quero, M. Burns, M. Razzaq, C. Nugent, M. Espinilla, "Detection of Falls from Non-Invasive Thermal Vision Sensors Using Convolutional Neural Networks", *Multidisciplinary Digital Publishing Institute Proceedings*, vol. 2, no. 19, pp. 1236, 2018.

[12] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code", *IEEE Transactions on Communications*, vol. 31, pp. 532-540, 1983.

[13] E. H. Adelson, C. H. Adelson and J. R. Bergen, "Pyramid methods in image processing", *RCA Engineer*, vol. 29, pp. 33-41, June 1984.

[14] S. E. Wei, V. Ramakrishna, T. Kanade and Y. Sheikh, "Convolutional Pose Machines", *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4724-4732, 2016.