An Efficient Combination of RGB and Depth for Background Subtraction

Van-Toi Nguyen[1,2], Hai Vu[1], and Thanh-Hai Tran[1]

[1] International Research Institute MICA, HUST - CNRS/UMI-2954 - Grenoble INP and Hanoi University of Science & Technology, Vietnam

[2] University of Information and Communication Technology, Thai Nguyen University

Abstract. This paper describes a new method for background subtraction using RGB and depth data from Microsoft Kinect sensor. Our main contributions are twofolds. First, we proposed a method for noise removal from depth data. Noise suppression helps to recover missing information of the depth map, so improve the stability of background. Second, for background subtraction, instead of using traditional RGB data, we use both RGB and depth data. The depth data, once being denoised, could avoid major limitations of RGB mostly when illumination change. Our strategy of combination of RGB and Depth is that when depth measurement is reliable, the segmentation is mainly based on depth information, inversely, we use RGB as alternative. The proposed method is evaluated on a public benchmark dataset which is suffered from common problems of background subtraction such as shadows, reflections and camouflage. The experiments show better segmentation results in comparison with state of the art works. Furthermore, the proposed method is successful with a challenging task such as extracting human fall-down in a RGB-D image sequence. The foreground segmentation results is feasibility for recognition task.

Keywords: Microsoft KINECT, Background Subtractions, Color Segmentation, Depth in use, RBG-D Combinations

1 Introduction

Background subtraction, shortly named BGS, aims to separate moving/dynamic objects from static scene. This is a critical task in many vision-based applications such as object detection, tracking, and recognition. The BGS techniques in the literature are briefly surveyed in related works. One of the most common BGS techniques uses Gaussian Mixture Model (GMM) to model statistics of background pixels [4, 5]. In such works, the BGS techniques using only color features are suffered from major limitations such as camouflage, shadow or variable lighting conditions. These problems cause over segmentation results.

Recently, depth data provided by Time-of-flight cameras or Microsoft KINECT sensors [1], becomes very attractive for background subtraction, particularly, in indoor environments. Major advantages of the depth data are that it does not suffer from limitations of RGB data. However, the using sole depth data still presents some problems such as: depth sensors often raise noises at object boundaries; measurements of depth are not always available for all image pixels [3]. Therefore intuitively, utilizing both RGB and depth information will offer valuable combination schemes for pruning segmentation results. Some combination schemes are listed in related works [19, 3, 2]. However, such works still do not really exploit robust characteristics of either depth and color features. For example, [2] simply concatenates the segmentation results of color and depth features; or it requires too complicated computations [3].

To tackle these issues, we propose an efficient combination of depth and color features. Our main constributions are twofolds.

- First, we propose a method for noise removal from depth data. Noise suppression helps to recover missing information of the depth map, so improve the stability of background.
- Second, for background subtraction, instead of using traditional RGB data, we use both RGB and depth data. The depth data, once being denoised, could avoid major limitations of RGB mostly when illumination change. Our strategy of combination of RGB and Depth is that when depth measurement is reliable, the segmentation is mainly based on depth information, inversely, we use RGB as alternative.

The remaining of this paper is organized as follows. Section 2 presents a brief survey of BGS. Section 3 presents the framework of proposed method. Section 4 presents the proposed noise model of the depth features as well as identify depth in valid range. Section 5 explains our segmentation algorithm combining color and depth features. Section 6 gives the experimental results comparing the proposed method with existing ones. Section 7 concludes and suggests extension works in the future.

2 Related Works

Background subtraction is a fundamental topic in the field of computer vision. There are uncountable BGS techniques in relevance works of object detection, tracking, surveillance, robotic, so on. In this section, we briefly summary some related techniques in the literature. Readers can find good surveys on BGS techniques in [20, 21]. Based on the features, we category the BGS methods into three groups: 1 - only use color data; 2 - only use depth data; 3 - combine color and depth data.

The methods in the first group use only color features. They are traditional approaches developed in many related works. The piglets segmentation method [22] uses a reference image to model background then segment foreground objects using a threshold on the difference image. The reference image is average of a sequence of images. J. Zheng et al. [23] analyze histogram over time to extract the background image from traffic videos. C. Wren et al. [24] model background using mean color values and the distribution of the mean values as a single Gaussian. C. Stauffer and L. Grimson [5] model background by a mixture of Gaussians. Some researchers use fuzzy logic approaches for background modeling [6–9]. D. Butler et al. [10] represent each pixel by a group of clusters. Given an incoming frame, the pixels are compared against the corresponding cluster group. K. Kim et al. [11] proposed a method that quantize the sample background values utilizing the codebooks which represent a compressed form of background model in a long image sequence. A testing pixel is classified as background if the color distortion to some codewords is less than the detection threshold and its brightness lies within the brightness range of that codeword. D. Culibrk et al [12] proposed a neural network architecture to form an unsupervised Bayesian classifier for Background Modeling. S. Messelodi et al. [14] proposed an algorithm based on Kalman filtering for updating the background image within video sequences. K. Toyama et al. [13] used Wiener filtering to make probabilistic predictions of the expected background. All methods using only color features still met unexpected effects caused by illumination changes, shadows, reflections and camouflage.

In the second category, the methods exploited only the depth features. A. Stormer et al. [16] proposed a method of background modeling and foreground objects segmentation based on Gaussian Mixture Model of depth data. The depth sensor in their work was a Time-of-flight Camera (PMD[Vision]3k-S). A. Frick et al [17] proposed an approach for 3D-TV Layered Depth Video (LDV) - Content creation using a capturing system of fourCCD - Cameras and Time-Of-Flight - Sensor (ToF - Camera). They used mean filtering to remove noise. They then also applied the GMM method for background modeling and movement detection.

The third category contains the techniques combining both color and depth features. I. Schiller and R. Koch [19] combined the segmentation of dynamic objects in depth with a segmentation in the color features using adaptive background models. They created background depth using averaging several ToF-images. They then used GMM for background modeling and foreground objects detection on color data. In such work, the authors weighted two measures depending on the actual depth values using either the variance or the amplitude of the depth image as reliability measure. G. Gordon et al. [2] proposed a method of BGS based on concatenating results of the depth and color features. They modeled each pixel as a GMM with 4 features (R,G,B,Depth) observations at each pixel over a sequence of frames in a multidimensional histogram. They used the census stereo algorithm on a pair of cameras to estimate the distance of the objects. J. Fernandez-Sanchez et al. [18] proposed a fusion method to combine color and depth (from Kinect) based on an advanced color-based algorithm. Background modeling and foreground segmentation method was based on Codebook model. They used depth cues to bias the segmentation based on color. M. Camplani et al. [3] proposed a Foureground/Background segmentation method based on a combination of two statistical classifiers using color and depth features. Their combination was obtained through a weighted average combiners. For each pixel, supporting of each classifier to the final segmentation results was obtained by considering the global edge-closeness probability and the classification labels obtained in the previous frame. The combination of depth cue and color cue in above methods allow to solve color segmentation issues such as shadows, reflections and camouflage. Although method of M. Camplani et al. [3] is state-of-the-art in the literature. However, it were too complex implementations and still did not really exploit full advantages of both depth and color information together.

3 The Framework of Background Subtraction

The framework of our proposed method is presented in the Fig.1. It composes of two main phases.

- Learning: This phase consists of modeling noise from depth map and learning the background model from depth and RGB data using GMM. This is an offline phase that takes depth and RGB sequences of background images in a duration of time.
- Background subtraction: This online phase does the background subtraction by combining the results of BGS based on RGB and depth information.

In the following, we will detail each step of the framework.

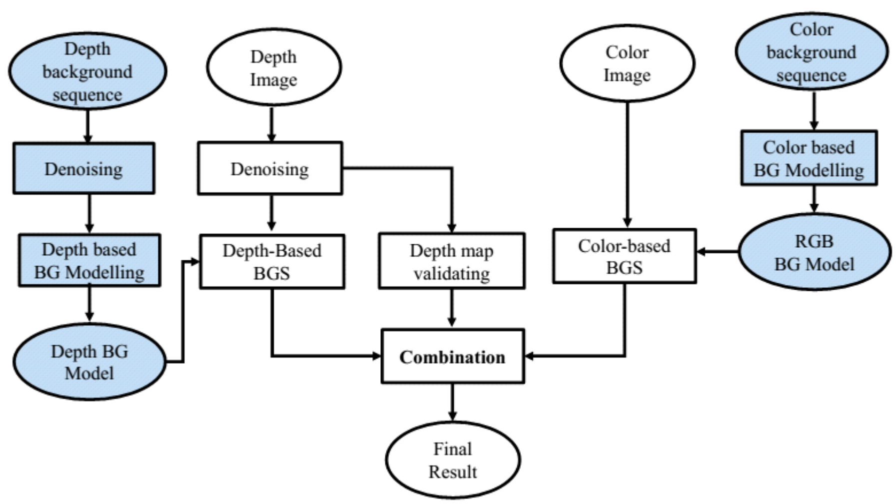

Fig. 1. The framework of proposed method. Learning steps are filled with blue color.

4 Removing Noises in The Depth Data

4.1 Build The Noise Model of Depth Data

To build noise model of depth data, we consider the depth of static scene in a duration T. Assume that depth map of the background scene is $S = [M \times N \times T]$ with $< M, N >$ are width and height of the depth image respectively (image size usually is 640×480 pixels). A noise model of depth data aims to find positions of noise from the observed signal S and statistical parameters to filter noises from the captured depth image. Observing a depth signal s at pixel $< i, j >$ in the duration T allows evaluating stability of depth data. Intuitively, a noise pixel usually makes signal $s(i, j)$ become unstable. To measure the stability of each background pixel (i, j), we evaluate standard derivation (std) of $s(i, j)$. A pixel at location (i, j) will be defined as noise as following:

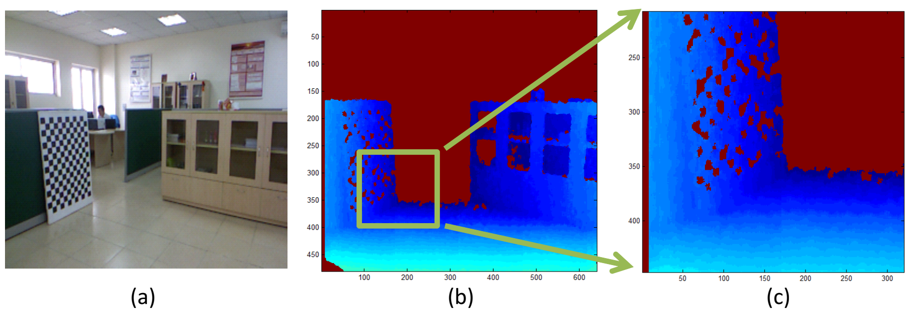

Fig. 2. An example of noise data captured by depth sensor Kinect. (a) RGB image for reference. (b) The corresponding depth image. (c) Zoom-in a region around chest board. A noise data is especially high at object boundaries.

$$Noise(i,j) = \begin{cases} 1 & \text{if } std(s(i,j)) \geq Threshold; \\ 0 & \text{if } std(s(i,j)) < Threshold; \end{cases} \quad (1)$$

The $Threshold$ is predetermined by heuristical selection. However, the empirical study shows that it is not strictly selected $Threshold$ value. A stable $s(i, j)$ always associate with a low value of std. Fig.3 shows noise pixels detected in a background scene observed in Fig.2 above. The noise signal s along time T of a pixel at coordinate (251,182), as shown in Fig.3a, is extracted. Original depth data of $s(251, 182)$ is plotted in red line in Fig.3b. It is a noise pixel according to (1). An image of noise pixels is shown in Fig.3c. As expected, the noise pixels appear high density around the chessboard.

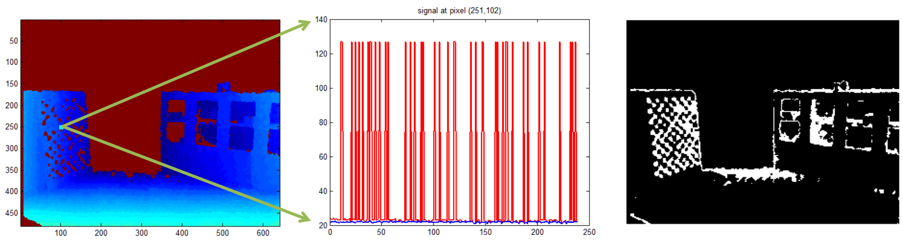

Fig. 3. Analysis of stability of noise in a duration $T = 5s$. (a) The signal s at pixel at position $(251, 182)$ is examined. (b) The corresponding signal s along T is plotted in red; the filtered signal s_f is plotted in blue. (c) Noise pixels are detected in all images. As expected, high density of noise appeared in regions of chessboard and boundary of the book cases.

4.2 Noise Reduction Using The Proposed Noise Model

The noise model supports us an effective algorithm for filtering noise pixels in the depth image. As shown in Fig.4a, identifying a pixel that is noise or not in the depth images is ensured. For such pixels, we generated new values of depth based on observation on low band data of the result of a K-mean algorithm ($K = 2$). A random value is generated to fill-in the depth pixel. Fig.4b presents results of noise detection after applying the filtering procedure. Some pixels that is still in noise is available to remove using a simple median filter on current frame. Fig.4c shows results after a median filtering with kernel size of 3×3 pixels.

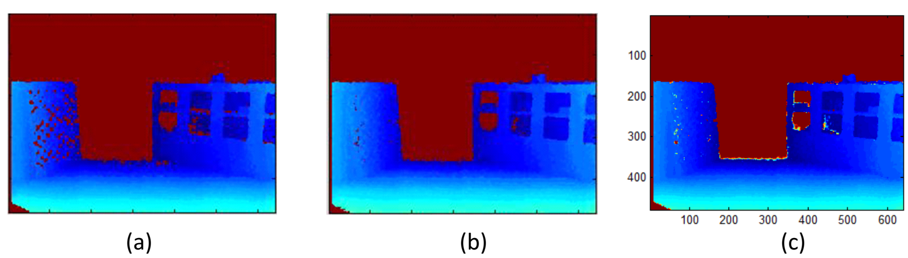

Fig. 4. Results of filtered noise on background scene. (a) An original depth frame. (b) The filtered noise depth frame. (c) Result after apply a median filtering on (b).

5 Background/Foreground Segmentation

5.1 The Prototype of Background Subtraction

We define the background as the stationary part of a scene. We model each pixel as an independent statistical process. Gaussian Mixture Model is observed for each pixel over a sequence of frames. For ease of computation, we assume a covariance matrix of three color channels $[RGB]$ is equal. At each pixel a mixture of three Gaussian is estimated according to procedure proposed in [4]

Once we have an estimate of the background in terms of color and range, we can use this model to segment foreground from background in a subsequent image of the same scene. Ideally a pixel belongs to the foreground, F, when its current value is far from the mode of the background model relative to the standard deviation.

$$F \equiv |P_i - P_m| > k\sigma \tag{2}$$

where P_i is the pixel value at frame i (in color and range space), P_m is the mode of the background model at the same pixel, σ is the variance of the model at that pixel, and k is a threshold parameter. In our implementation, this prototype for background subtraction is implemented for both depth and color features. Foreground segmentation from depth named F_d, whereas foreground segmentation from color is named F_c. Background model of depth and color are named P_{md} and P_{mc}, respectively. We build separated model for each channel $[R, G, B, D]$. An example of the R channel for $Fall$ sequence (see details in Section 6) is shown in Fig.5.

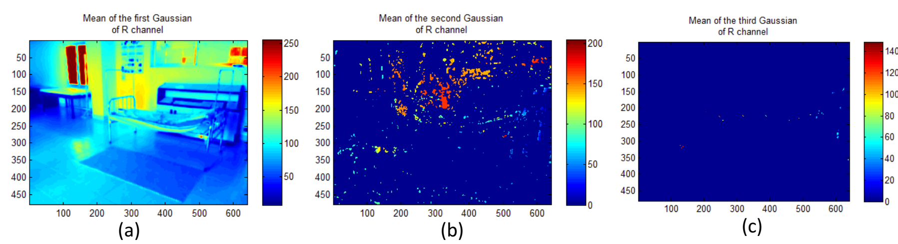

Fig. 5. GMM of R channel for $Fall$ sequence. The mean data of the first, second, and third Gaussian is visualized at (a),(b),(c), respectively.

5.2 Background Subtraction Using Depth Feature

Given a depth image as shown in Fig.6b (see Fig.6a for reference). Using background model of depth as shown in Fig.6c, we obtain different from given frame and background model. According to (2), a predetermined threshold is selected to obtain binary images including foreground regions. Further processing obtains a fine result of foreground regions (Fig.6f).

5.3 Background Subtraction Using Color Feature

Similar to BGS using depth feature, our segmentation prototype is applied to color feature. Original color frame is shown in Fig.7a. For a background model

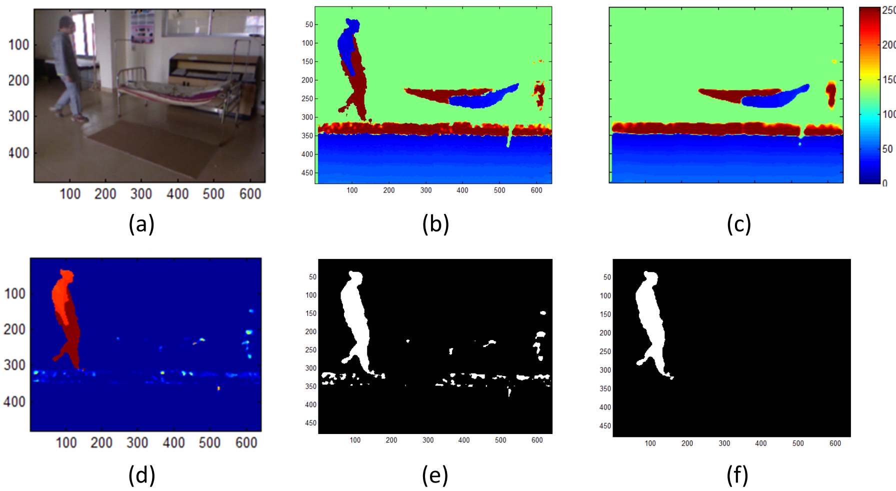

Fig. 6. BGS using Depth feature. (a)-(b) Color and depth original images, respectively. (c) Background model of depth. (d) Difference between the given frame and background model. (e) F_d segmentation. (f) Results after removing small blobs

given in Fig.5, difference from given frame and background model is shown in Fig. 7(b). Using a predetermined threshold in (2), we obtain foreground regions, as shown in Fig.7c. However, selecting a suitable threshold for BGS using color feature is more difficult than using depth feature.

5.4 Combination of Depth and Color

Our combination takes a disjunction of the foregrounds detected by depth and color features. The final segmentation result therefore is defined by:

$$F \equiv F_d \bigcup F_c \tag{3}$$

A strategy for the combination is that where depth indicates that a pixel is in the valid range of the depth measurement, color matching is unimportant since the depth information alone is sufficient for correct segmentation. Therefore, a valid depth is proposed to obtain foreground from depth: $Valid(F_d) \equiv Depth \ll MaxVal$ where $MaxVal$ is depth value which is out of range of depth measurement. Given a depth image F, which is filtered noises using the proposed noise model in Sec.4, foreground regions F_d is able to be estimated by (2). We have been referring to presence of low confidence depth values as *invalid*. The procedure to eliminate *invalid* is:

$$\begin{aligned} B &= Valid(P_{md}) \cap (1 - F_d) \\ Valid(F_d) &= Valid(F) - B \end{aligned} \tag{4}$$

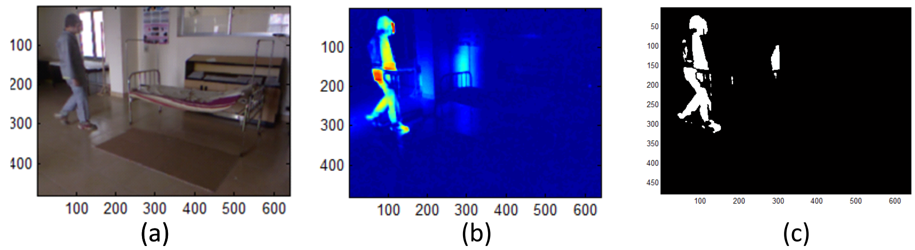

Fig. 7. (a) Original color image. (b) Difference from background model. (e). F_c segmentation results

Effectiveness of the combination scheme is shown in Fig.8. As the proposed scheme, a valid depth in use is identified first. Fig. 8b shows validated depth pixels from background models, that presents pixels in range of measurements from depth sensor. The valid depth is reduced with foreground images in Fig.8c. This sub-figure presents pixels where depth is biased than color features. Without using depth features, results of foreground segmentation is including many shadows around box, as shown in Fig.8e. Using depth information, many shadow have been removed in Fig.8f. Final results is disjunction of depth and color in Fig.8k. On the other hand, this example also present effective of color features. For pixels in out of range of depth (or invalid depth pixels), as border regions of images, foreground segmentation from color feature is utilized. Therefore, in the final results, hand of the person, who keeps the box, is included.

6 Experimental Results

The aim of our experiments is to demonstrate the performance of our proposed method. We evaluate the proposed method in two aspects: (1) Showing the effectiveness of combining RGB and depth information for background subtraction instead of using separated data; (2) Comparing the proposed method with a state-of-the-art combination method using a public dataset provided in [3]. The experimental results also confirm that the proposed method is successful for segmenting an image sequence with human fall (fall-like) actions.

6.1 Dataset and Evaluation Measurement

Dataset We will test our proposed method with five datasets. The first four are benchmark datasets that have been used in [3]. They include several indoor sequences acquired by Microsoft Kinect [1]. Each sequence contains a challenge of BGS such as shadow, camouflage, so on. The description details of this dataset can be found in [3]. A part from that, we build ourselves a dataset, named MICA-FALL, in the context of human fall detection. The main purpose is to automatically detect abnormal activities of human (patient in hospital, elderly)

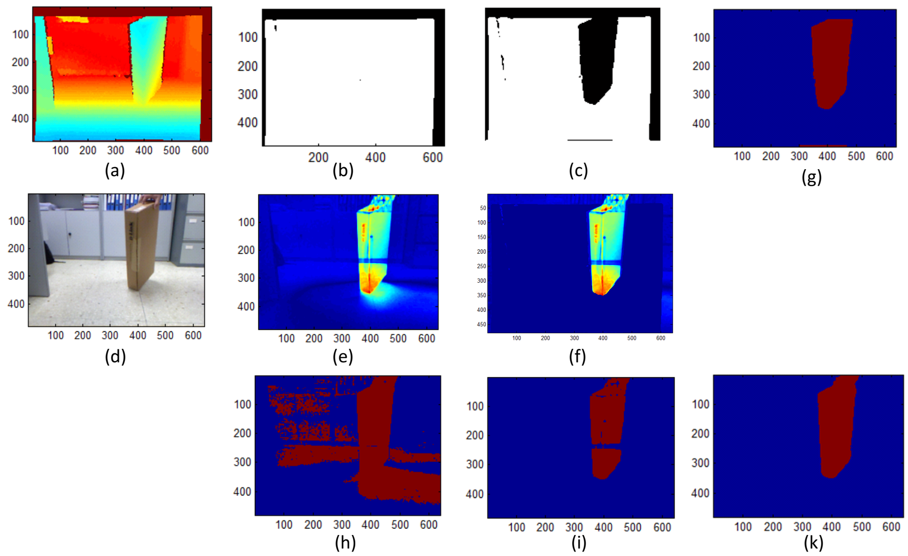

Fig. 8. First row: (a) Original depth image; (b) Valid depth measurement from background. (c) Valid depth with foreground image. (d) Original color image; (e) Difference of color from background model without depth association (c). (f) Difference of color form background model with depth association (c). (h)-(i) are foreground segmentation results from (e)-(f), respectively. (k) Final result: is disjunction of (g) and (i).

in order to alarm to assistance in hospital as soon as possible. These sequences are captured by a Kinect device in an indoor environment. This dataset is more challenge for segmenting foreground. There are big shadows on the wall when a patient moves in the room; inflection on the floor of body-patient when he falls down. The field of view in the treatment-room is quite large and patient always goes out of range of depth sensors.

Evaluation Measurement We use a measure based-on the binary Jaccard index that is described in [25]. The measure is defined as:

$$JI = \frac{FG \cap GT}{FG \cup GT} \tag{5}$$

Where FG is foreground detection result, GT is groundtruth.

6.2 Results

First, we evaluate how the combination strategy improves the background subtraction when using RGB or Depth separately. Fig.9 shows that color information gives very poor results (57% in average). It is worst in all cases due to all challenges (camouflage, shadow, lighting change) appeared in the datasets. Depth

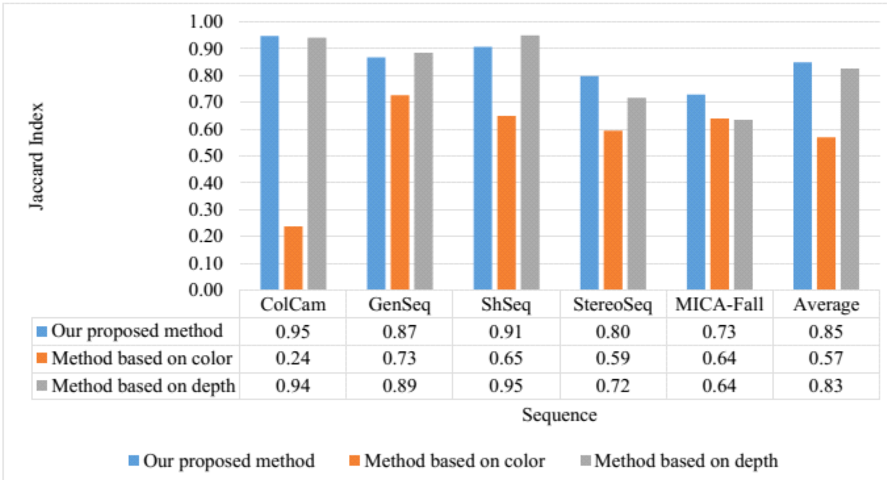

Fig. 9. Background subtration using depth, color and both: a comparison.

information gives more stable results. It provides even the best result on the sequence ShSeq. The reason is that the groundtruth does not consider hand as foreground but in reality, hand is a moving object in the scene so it can not belong to background. We can see that the proposed method that combines depth and RGB information gives the best result in overall.

We compare next our proposed method with Camplani et al. method [3]. Fig.10 shows the comparative results. We observe that our method gives the better results in 3 sequences (ColCam, GenSeq, StereoSeq) and in average that proves the effectiveness of our proposed method. The result is worse than Camplani et al. method at the sequence ShSeq due to the same reason explained previously. In the following, we will look at some examples to explain in more detail advantages of the proposed combination strategy.

Fig.11 illustrates an example with different results of background subtraction. The original image is extracted from frame 446 of the sequence ShSeq. In this sequence, the human moves the book in front of the Kinect sensor. The book and a part of the human hand are considered as foreground. Notice that the provided groundtruth of this sequence, however, does not consider the hand as foreground, but only the object of interest (the book). The Fig.11b shows depth data with red pixels are noise. Fig.11c is groundtruth. Fig.11g is the result of BGS Camplani et al. method, that we extract from the original paper. We could see the result corvers the book. Some points inside the book are missing while some outside pixels are false segmentation. As the original image shows, the problem of shadows is important in this case. Therefore, the BGS result using RGB information is very poor (Fig.11d). Depth information in this case is quite stable inside the book so the BGS using depth data is good. Fig.11f presents

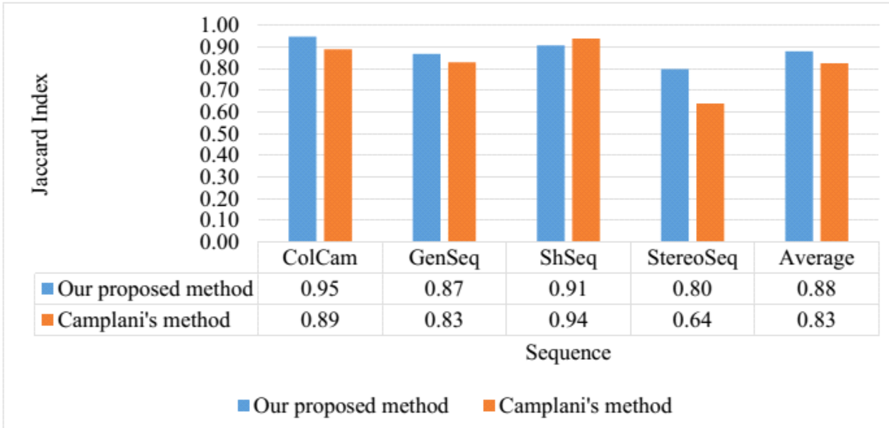

Fig. 10. Comparison with Camplani's method [3].

the result using combination strategy. It is a expected result: book and a part of hand are both segmented, without any other under / over segmentation.

Another example could be seen in Fig.12. This time, we compare only the results of Camplani method and ours. The red points in Fig.12b are noisy points of depth data. Camplani et al. method gives lots of under and over segmentation while ours gives more favorable results.

Fig.13 shows the result on an image sequence of the fall-down action. Obviously, these segmentation results are feasible to implement recognizing works.

7 Conclusions

This paper proposed an efficient method for background subtraction with Kinect data. We taken into account noise of depth features. This model presents effective to eliminate noise of depth, that is attractive for identifying the valid depth pixels. Our combination scheme is based on advantages of valid depth and full view of colors. These features are complementary to obtain fine segmentation results. The proposed method was validated with benchmark dataset and shows a significant improvement with respect to a state of the art work. It was also successful for segmenting human activities in our challenging task that is to recognize a human fall-down actions. The proposed method has some limitations when object closes to a background region that is in valid range of the depth measurement. These limitations suggests direction to further researches.

Acknowledgments The research leading to this paper was supported by the National Project B2013.01.41 *"Study and develop an abnormal event recognition system based on computer vision techniques"*. We would like to thank the project and people involved in this project.

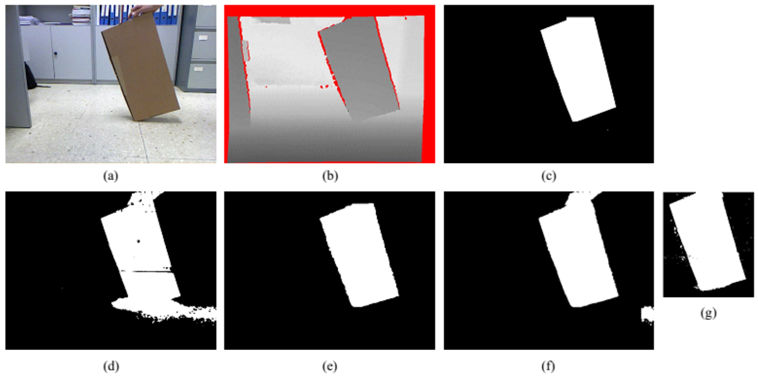

Fig. 11. Frame 446 of the ShSeq sequence: (a) Color data. (b) Depth data. (c) Groundtruth of foreground. (d) BGS result using color data. (e) BGS result using depth data. (f) BGS using combined RGB and Depth. (g) Output of Camplani's method [3] (extracted from the original paper).

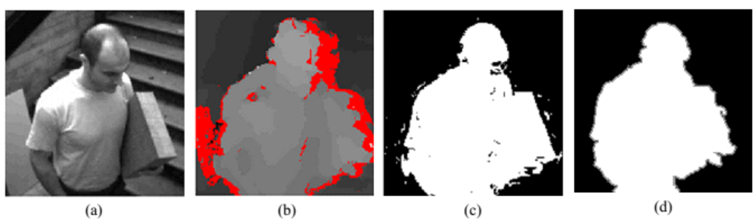

Fig. 12. Frame 139 of the Stereo sequence: (a) Color data. (b) Depth data. (c) Output of Camplani's method [3]. (d) Output of our proposed method.

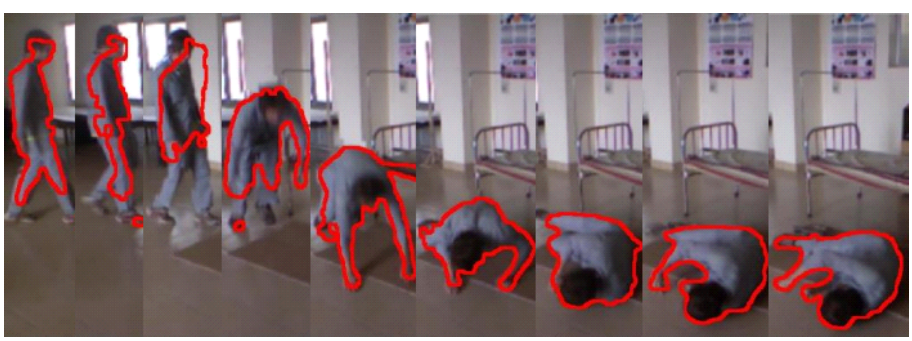

Fig. 13. The result on an example sub-sequence of MICA-FALL dataset.

References

1. Microsoft Kinect, http://www.xbox.com/en-US/xbox360/accessories/kinect, 2013
2. G. Gordon, T. Darrell, M. Harville, J. Woodfill, "Background estimation and removal based on range and color", In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition,pp. 1-6 ,Jun. 1999.
3. M. Camplani, L. Salgado, Background Foreground segmentation with RGB-D Kinect data: an efficient combination of classifiers, Journal of Visual Communication and Image Representation, Elsevier, In Press.
4. W.E.L. Grimson, Chris Stauffer, Raquel Romano, and Lily Lee, Using adaptive tracking to classify and monitor activities in a site, In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Santa Barbara, pp. 22-29, Jun. 1998.
5. C. Stauffer, W. E. L. Grimson, "Adaptive background mixture models for real-time tracking". In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. pp. 246-252, Aug. 1999.
6. Sigari, M.H., Mozayani, N., Pourreza, H.R.: Fuzzy Running Average and Fuzzy Background Subtraction : Concepts and Application. International Journal of Computer Science and Network Security 8.2, pp. 138-143, (2008).
7. Baf, F. El, Bouwmans, T., Vachon, B.: Type-2 Fuzzy Mixture of Gaussians Model : Application to Background Modeling. Advances in Visual Computing, pp. 772-781, (2008).
8. Zhang, H., Xu, D.: Fusing color and texture features for background model. Fuzzy Systems and Knowledge Discovery: Third International Conference. pp. 887-893, (2006).
9. El Baf, F., Bouwmans, T., Vachon, B.: Fuzzy integral for moving object detection. 2008 IEEE International Conference on Fuzzy Systems (IEEE World Congress on Computational Intelligence). pp. 1729-1736. (2008).
10. Butler, D., Sridharan, S., Jr., V.M.B.: Real-time adaptive background segmentation. IEEE International Conference on Acoustics, Speech, and Signal Processing. pp. III-349 (2003).
11. Kim, K., Chalidabhongse, T.H., Harwood, D., Davis, L.: Real-time foreground/background segmentation using codebook model. Real-Time Imaging. 11, pp. 172-185 (2005).
12. Culibrk, D., Marques, O., Socek, D., Kalva, H., Furht, B.: Neural network approach to background modeling for video object segmentation. IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council. 18.6, pp. 1614-1627 (2007).
13. Toyama, K., Krumm, J., Brumitt, B., Meyers, B.: Wallflower: Principles and practice of background maintenance. The Proceedings of the Seventh IEEE International Conference on Computer Vision. pp. 255-261 (1999).
14. Messelodi, S., Modena, C.M., Segata, N., Zanin, M.: A Kalman filter based background updating algorithm robust to sharp illumination changes. Image Analysis and ProcessingICIAP 2005. pp. 163-170 (2005).
15. Brutzer, S., Hoferlin, B., Heidemann, G.: Evaluation of background subtraction techniques for video surveillance. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1937-1944 (2011).
16. Stormer, A., Hofmann, M., Rigoll, G.: Depth Gradient Based Segmentation of Overlapping Foreground Objects in Range Images. 13th Conference on Information Fusion (FUSION). pp. 1-4 (2010).

17. Frick, A., Kellner, F., Bartczak, B., Koch, R.: GENERATION OF 3D-TV LDV-CONTENT WITH TIME OF FLIGHT CAMERA. 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video. pp. 1-4 (2009).
18. Fernandez-Sanchez, E.J., Diaz, J., Ros, E.: Background subtraction based on color and depth using active sensors. Sensors (Basel, Switzerland) 13, no. 7, pp. 8895-8915 (2013).
19. Schiller, I., Koch, R.: Improved Video Segmentation by Adaptive Combination of Depth Keying and. Lecture Notes in Computer Science (Image Analysis), pp. 59-68 (2011).
20. Bouwmans, T.: Recent Advanced Statistical Background Modeling for Foreground Detection - A Systematic Survey. Recent Patents on Computer Science 4.3, pp. 147-176, (2011).
21. Goyette, N., Jodoin, P., Porikli, F., Konrad, J., Ishwar, P.: Changedetection. net : A New Change Detection Benchmark Dataset. Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 1-8 (2012).
22. McFarlane, N.J.B., Schofield, C.P.: Segmentation and tracking of piglets in images. Machine Vision and Applications. 8, pp. 187-193 (1995).
23. Zheng, J., Wang, Y., Nihan, N., Hallenbeck, M.: Extracting Roadway Background Image: Mode-Based Approach. Transportation Research Record: Journal of the Transportation Research Board 1944.1, pp. 82-88 (2006).
24. Wren, C., Azarbayejani, A., Darrell, T., Pentland, A.: Pfinder: Real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence. 19.7, pp. 780-785 (1997).
25. McGuinness, K., O'Connor, N.E.: A comparative evaluation of interactive segmentation algorithms. Pattern Recognition. 43.2, pp. 434-444 (2010).