

Mobile Camera based Motion Segmentation by Image Resizing

Chunyu Yu

Department of Electrical Engineering and Computer Science
Syracuse University, Syracuse, NY 13244-1240
Email: cyu07@syr.edu

Abstract

The need of detecting moving object like human and vehicle by mobile camera is increasing in commerce and industry. In this paper, an image-resize methodology which can abstract motion segmentation and detect moving object from moving background is proposed. First, edges images are computed. Then movement vector between frame images are computed and the relative background motion is compensated. By adjusting the parameters of resize algorithm, human liked object or vehicle liked object can be segmented separately and the segmentation can be used for further detection. Experiments have been performed under three different environments for human detection and vehicle detection.

Keywords: mobile camera, motion segmentation, edge-based alignment, object detection.

1. Introduction

Real-time human detection is a key study field in computer vision. This technique is widely used in automotive engineering, traffic monitoring, robotics, human-computer interaction, security, military and many other applications. In literature, there are a lot of researchers focused on stationary camera based object detection. Main technique used is image difference to find the difference of two image pixel by pixel. However, when camera is moving during the recording, the whole background is changing. To resolve this problem, [1-4] developed different algorithms to against the light changes.

Another widely used method is statistical theory which classifies pixels into background and foreground. C. Rasmussen and G. D. Hager combine Maximum A posteriori Probability (MAP) with a Probabilistic Data Association Filter [5]. N. Vaswani, A. Tannenbaum, and A. Yezzi unified geometric active contours and level sets within Particle Filter [6] and R. Stolkin, A. Greig, M. Hodgetts, and J. Gilby, proposed an algorithm which

combines Expectation Maximization and extended-Markov Random Field [7]. Another method is wavelets based. The main idea is to analysis the difference part's frequency. For instance, [8] presented a motion segmentation based on Galilean wavelets. In 1980 Horn and Schunch first proposed a new method named optical flow [9]. After that, many researchers made improvement on it such as J. Zhang, F. Shi, J. Wang, and Y. Liu [10].

In this paper, an image resizing based motion segmentation method is proposed. Unlike other computational expensive method, this algorithm only relies on a new edge alignment method and fast image resizing according to object morphology. First, edges in two consecutive frame images are detected. Then the new edge-based alignment method is adopted to find the motion vector between two images. Motion compensation is performed before motion segmentation. Then the image is resized according to object morphology. After filtering pixels lower than threshold value, object is stand out of the background. The experiment was performed by MATLAB 2014a on i5-2450m and 6GB RAM in three different scenarios for

human liked object detection and vehicle liked object detection. The results illustrate that the accuracy rate can reach as high as 93.04%, and the execution time is highly speeded up (approximately 0.82s per frame).

The paper is organized as follows. The proposed method for motion segmentation and object detection is described in section 2 and section 3 presents the experiment results and the conclusion part is in section 4.

2. Proposed Method

2.1 Finding global minima

Take human liked object detection as an example. Two consecutive frame images f_1 and f_2 are loaded first. By using Canny edge detector, two binary edge images are obtained. Ideally, the camera is stable and the background is not moved, so in the difference image ($f_2 - f_1$), only moving object is visible and background is offset. Virtually, the camera may slide a very small distance in a direction, so the background and people are moved simultaneously when two consecutive images are compared. Therefore, eliminate the movement of background is essential.

Assume people are most likely appear in the central area of an image, when detecting background movement, only edges at the margin of images are compared. Specifically, fill zero into the central area of two images and find out the relative movement between the background in 8 directions (vertical, horizontal and diagonal) (Fig.1). By the algorithm in table 1, relative motion direction and distance are achieved. Base on experiments, the movement between two frames is always 3 to 6 pixels, so 10 pixels are sufficient for searching in our algorithm.

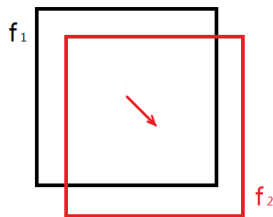


Fig. 1 One of eight directions of relative movement between two images

Table 1 Algorithm for finding relative motion direction and distance

```

for 1 to 8 directions
  for 1 to 10 pixels
     $m[i] = \text{abs}(\text{tailored}(f_1, -\text{direction}, \text{pixel}) - \text{tailored}(f_2, \text{direction}, \text{pixel}))$ 
  direction = minima(m).direction
  distance = minima(m).pixel

```

According to the direction and movement distance, tailor pixels from two images respectively. After that, beside the people, backgrounds are almost in the same position in two images (say, reach the global minima).

2.2 Resizing image algorithm

Even global minima is reached, virtually the edges in two images cannot be overlapped exactly. As a result, when difference image ($f_2 - f_1$) is calculated, background edges of two images can hardly offset each other. The background may become even more complicated than human outline. As a result, the human cannot stand out of background (Fig.2).



Fig. 2 Difference image acquired without resizing algorithm

To resolve the problem, an image resizing algorithm is adopted. Table 2 is the matrix of yellow rectangle in Fig. 2 which is achieved without dilating images. Table 3 is the result acquired after dilating images. By compare two tables, we can notice that the number of pixels which value is 1 is decreased and they are more scattered in Table 3. In terms of the whole image, dilating images results to more mottled background.

Table 2 Difference image acquired without dilating
(1-value pixels: 15)

	258	259	260	261	262	263	264	265	266	267
391	0	0	0	0	0	0	0	0	0	0
392	1	1	1	1	1	0	0	0	0	0
393	1	1	1	1	1	0	0	0	0	0
394	0	0	0	0	0	0	0	0	0	0
395	0	0	0	0	0	0	0	0	0	0
396	0	0	0	0	0	0	0	0	0	1
397	0	0	0	0	0	0	0	0	0	0
398	0	0	0	0	0	0	0	0	1	0
399	0	0	0	0	0	0	0	0	1	1
400	0	0	0	0	0	0	0	0	0	1

Table 3 Difference image acquired after dilating (1-value pixels: 13)

	258	259	260	261	262	263	264	265	266	267
391	1	1	1	1	1	1	0	0	0	0
392	0	0	0	0	0	0	0	0	0	0
393	0	0	0	0	0	0	0	0	0	0
394	1	1	1	1	0	0	0	0	0	0
395	0	0	0	0	0	0	0	0	1	0
396	0	0	0	0	0	0	0	0	0	0
397	0	0	0	0	0	0	0	1	0	0
398	0	0	0	0	0	0	0	1	0	0
399	0	0	0	0	0	0	0	0	0	0
400	0	0	0	0	0	0	0	0	0	0

As moving human most likely look like vertical rectangles, they are more robust than noise (remained background) in terms of vertical image resizing. Therefore, based on the result in Table 3, the image is resized in vertical direction and then restored. After that pixels' value are between -1 and 1.

Table 4 Matrix after image resizing

	258	259	260	261	262	263	264	265	266	267
391	0.31	0.31	0.31	0.31	0.23	0.23	0.00	-0.04	0.01	0.00
392	0.34	0.34	0.34	0.34	0.23	0.23	-0.00	-0.00	0.04	0.00
393	0.31	0.31	0.31	0.31	0.16	0.17	-0.02	0.07	0.11	0.00
394	0.25	0.25	0.25	0.25	0.08	0.08	-0.04	0.17	0.18	0.00
395	0.19	0.19	0.19	0.19	0.01	0.01	-0.04	0.25	0.23	0.00
396	0.14	0.14	0.14	0.14	-0.02	-0.02	-0.00	0.29	0.23	-0.00
397	0.09	0.09	0.09	0.09	-0.02	-0.05	0.07	0.31	0.16	-0.01
398	0.03	0.03	0.03	0.03	-0.01	-0.05	0.17	0.31	0.08	-0.02
399	-0.00	-0.00	-0.00	-0.00	-0.00	-0.04	0.25	0.29	0.00	-0.02
400	-0.02	-0.02	-0.02	-0.02	0.00	-0.00	0.26	0.25	-0.02	0.01

Due to the distortion, mottled background faded dramatically compared to human outline. It can be eliminated by deleting pixels which value is lower than specific threshold value. According to the environment and object, 0.5 is adopted in this experiment. Therefore, all the pixels in Table 4 will be 0. Compare to Fig. 2, the road line in Fig. 3 is eliminated (yellow rectangle) and human is out stand of background.



Fig. 3 Difference image acquired by resizing algorithm

The following step are dilating image, filling image hole and eroding image. The largest block in the image is the object.



Fig. 4 segmentation achieved by resizing algorithm (sinister) versus without resizing algorithm (dexter)

The following assumptions are made in proposed algorithm. Firstly, there must be sufficient edges in the margin of the images as they are used to find motion vector. Secondly, there must be sufficient movement of the object. If the object is stationary, it will be considered as background. Finally, the object should large enough in terms of the image scale as the largest block is selected as object in the final step.

3. Experiment Results

Without loss of generality, the experiment was performed to test vertical resizing and horizontal resizing in three different scenarios: (1) human walk indoor, (2) human walk outdoor and (3) human walk outdoor with moving vehicle in background. The dataset of Zheng et al.'s experiment was adopted [11].

3.1. Human walk indoor without moving object

231 frames of a video were tested in this experiment. Fig. 5 as part of results illustrates that the movement segmentation is highly accurate and almost overlaps with the human body. Thus the segmentation will decrease the running time largely for further HOG detection. Even though illumination sometimes disturbs the detection, the overall successful detection rate for whole frames is 77.56%.

3.2. Human walk outdoor without moving vehicle

In this experiment, 273 frames were tested and the accurate detection rate is 93.04%. Similar to the results of indoor experiment, the segmentation part is highly overlapped with the human body (Fig. 6). Human walk outdoor with moving vehicle in the background

3.3. Human walk outdoor with moving vehicle in the background

100 frames were tested for human and vehicle. The detection rate is 86% and 83.1% respectively. As shown in Fig.7 and Fig.8, the moving vehicle can be distinguished from stopped bus.

4. Conclusions

In this paper, a fast and simple motion segmentation method for moving background is proposed. Zheng et al.'s dataset was tested in the experiment. The results show that the proposed method is highly accurate especially in complex environment. In the future, prior knowledge will be considered. The position of yellow box can be limited in a range according to its position in prior frame. Thus the algorithm can be more robust to fast light changing. Due to pretty accurate segmentation region, further detection such as HOG in this region can be speed up largely.

References

1. A. Cavallaro, O. Steiger, and T. Ebrahimi, "Tracking Video Objects in Cluttered Background," IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, no. 4, pp. 575–584, 2005.
2. F.-H. Cheng and Y.-L. Chen, "Real time multiple objects tracking and identification based on discrete wavelet transform," Pattern Recognition, vol. 39, no. 6, pp. 1126–1139, 2006.
3. R. Li, S. Yu, and X. Yang, "Efficient spatio-temporal segmentation for extracting moving objects in video sequences," IEEE Transactions on Consumer Electronics, vol. 53, no. 3, pp. 1161–1167, Aug. 2007.
4. A. Colombari, A. Fusiello, and V. Murino, "Segmentation and tracking of multiple video objects," Pattern Recognition, vol. 40, no. 4, pp. 1307–1317, 2007.
5. C. Rasmussen and G. D. Hager, "Probabilistic data association methods for tracking complex visual objects," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 560–576, 2001.
6. N. Vaswani, A. Tannenbaum, and A. Yezzi, "Tracking deforming objects using particle filtering for geometric active contours," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 8, pp. 1470–1475, 2007.
7. R. Stolkin, A. Greig, M. Hodgetts, and J. Gilby, "An em/e-mrf algorithm for adaptive model based tracking in extremely poor visibility," Image and Vision Computing, vol. 26, no. 4, pp. 480–495, 2008.
8. M. Kong, J.-P. Leduc, B. Ghosh, and V. Wickerhauser, "Spatio-temporal continuous wavelet transforms for motion-based segmentation in real image sequences," Proceedings of the International Conference on Image Processing, vol. 2, pp. 662–666 vol.2, 4-7 Oct 1998.
9. B. K. Horn and B. G. Schunck, "Determining optical flow," Cambridge, MA, USA, Tech. Rep., 1980.
10. J. Zhang, F. Shi, J. Wang, and Y. Liu, "3d motion segmentation from straight-line optical flow," in Multimedia Content Analysis and Mining, 2007, pp. 85–94.
11. Y. Zheng, "Detection of moving people with mobile cameras by fast motion segmentation", 2013 Seventh International Conference on Distributed Smart Cameras (ICDSC), pp.1-6, Oct. 29-Nov. 1 2013.

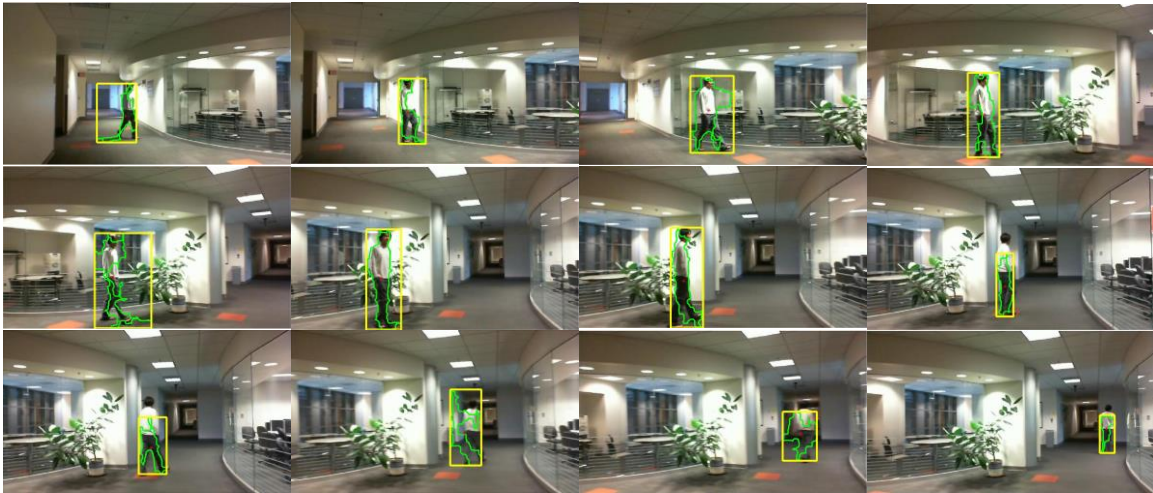


Fig. 5 Indoor sequence. The green line outlines the largest block and the yellow box indicates the possible human region

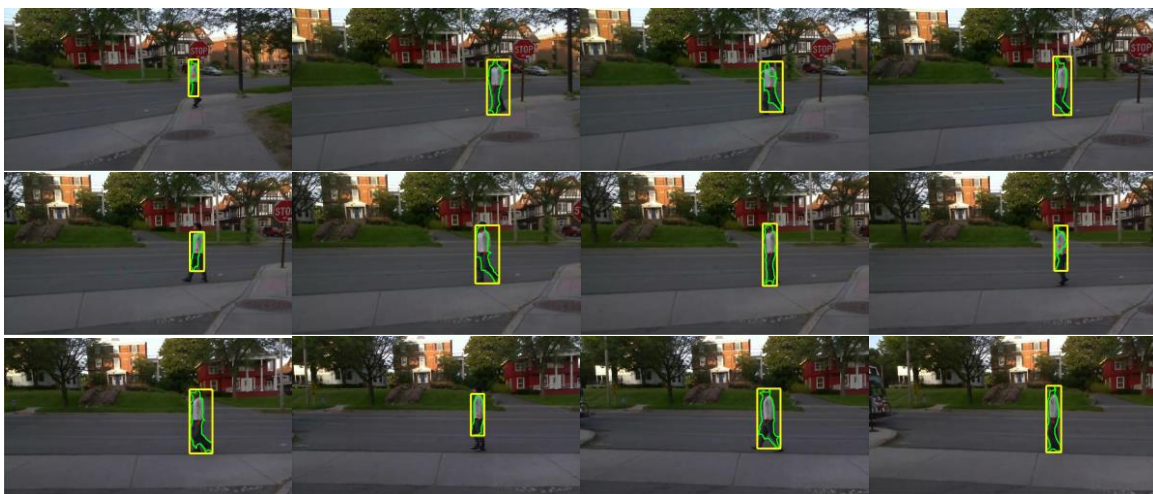


Fig. 6 Outdoor sequence. The green line outlines the largest block and yellow box indicates the possible human region

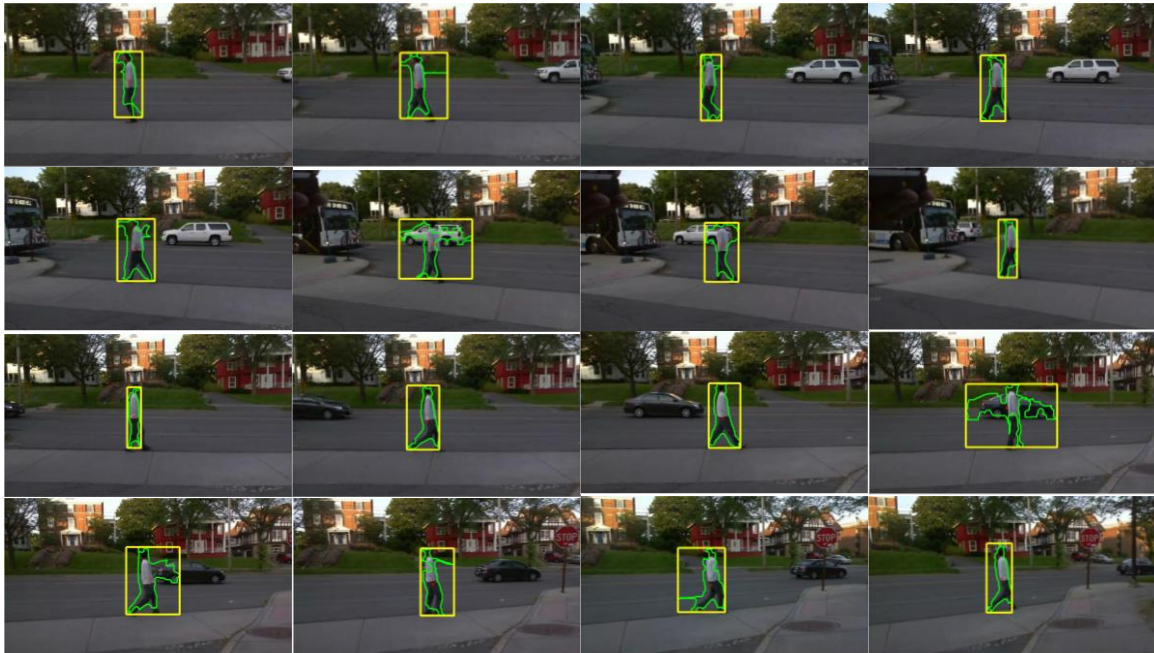


Fig. 7 Outdoor sequence with moving car in background. The green line outlines the largest block and yellow box indicates the possible human region

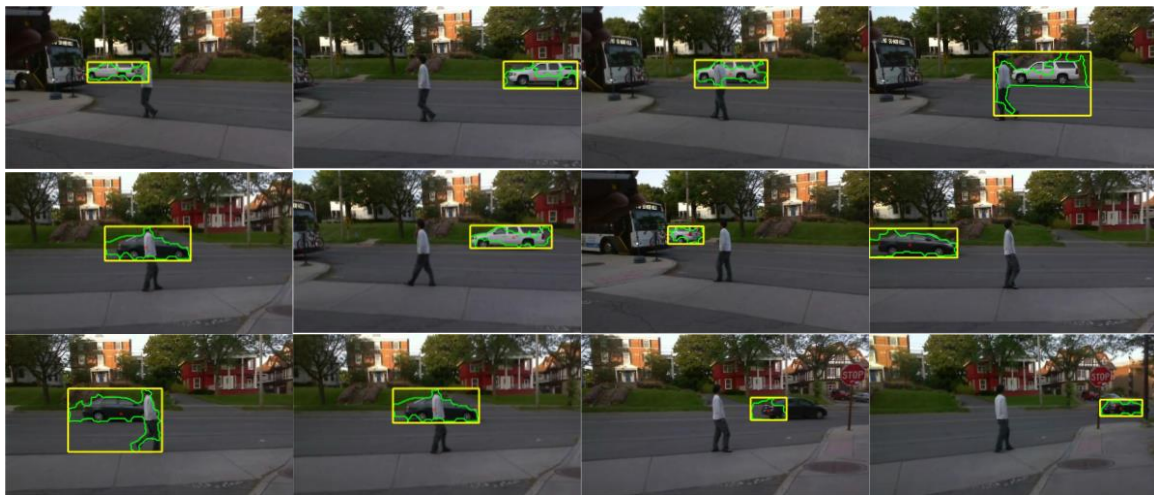


Fig. 8 Outdoor sequence with moving car in background. The green line outlines the largest block and yellow box indicates the possible vehicle region.