Distance/Motion-Based Segmentation Under Heavy Background Noise

Yajun Fang, Ichiro Masaki & Berthold K.P. Horn

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: A new real-time object segmentation algorithm is needed to deliver satisfactory segmentation results with high reliability and accuracy under the impact of heavy background noise.

Motivation: To improve the safety of drivers and the efficiency of transportation systems, more and more automobile companies use visual cameras to detect objects of interests, such as vehicles, pedestrians, etc., and related information. For example, to provide "adaptive cruise control" function, i.e., to adaptively adjust vehicles' speed based on its distance information to the preceding vehicles, it is necessary to detect accurate distance and position information for preceding vehicles. Fast and powerful segmentation algorithms are needed.

Previous Work: Currently, it is hard to find a simple and fast segmentation algorithm that provides accurate and reliable semgentation result, especially under the impact of heavy background noise. Many static segmentation algorithms have to trade off between accuracy and speed. To reach high segmentation accuracy and reliability, usually algorithms have to be complex. Background containing non-rigid objects such as trees makes segmentation problems even harder. The ability to suppress the background noise is important to improve segmentation performance for current real-time intelligent vehicle application systems.

For video sequences with stable background, tracking algorithms are typically used to improve segmentations performance. In many situations such as indoor monitoring, reference images are available. Many algorithms [3] capture background in reference images and remove it from the following video sequences. The previous frames can also be used as the reference images when changes between successive video frames are small. However, reference images might not always be available. Tracking algorithms do not work well either when there are significant background changes between successive video frames. Furthermore, such methods do not work well if the background contains non-rigid objects such as trees.

Approach: In paper [1], we have presented a new distance-range-based segmentation method that detect targets on its separated edge layers, which simplifies the segmentation task and provides high accuracy and robustness. There still exist the following issues to be further pursued.

– The ability to identify objects other than vehicles.

– The ability to segment multiple objects of different sizes within the same distance range.

– The ability to tolerate background noise.

We adopt the following steps to address the previous issues:

1. We first apply the target detection algorithm introduced in paper [2] [1] to decide disparity ranges corresponding to both background and targets, and then separate the binocular stereo edge map into several edge layers corresponding to all targets and background at different depths. The common pixels belonging to both target edge layers and background edge layers are removed, leading to "clean target edge layer" at each specific depth range.

   For two continuous binocular stereo image as in Fig. 1(a)-(d), that background edge pixels correspond to disparity range 0-5 while the nearest targets correspond to disparity range 17-19. Thus, the edge map of original binocular stereo images, as in Fig. 1(e), can be separated into two edge layers as in Fig. 1(f) and (g). Fig. 1(h) shows clean edge layer in disparity range 17-19 after removing all pixels belong to background layer. Usually, it is not necessary to remove the common pixels as in paper [1] since it might remove a few edge pixels that belong to targets of interest and make initial segmented regions less smaller than their real size as in Fig. 1(k).

2. To compensate for the potential losses, a motion-based region expansion is necessary. While motion vectors for all edge pixels in Fig. 1(i) do not show any clear pattern as in Fig. 1(j), we notice that motion vectors for all edge pixels in Fig. 1(h) clearly differentiate the foreground from background as in Fig. 1(j), which allows us to expand the segmentation result on Fig. 1(k) to its actual segmentation region in Fig. 1(l).

The proposed object segmentation method takes advantage of both motion and distance information to overcome the limitations of traditional segmentation algorithms. It can provide ideal speed while improving segmentation reliability, accuracy and robustness. The previous steps not only suppress the background noise impact, but also make it possible to detect objects of different sizes.

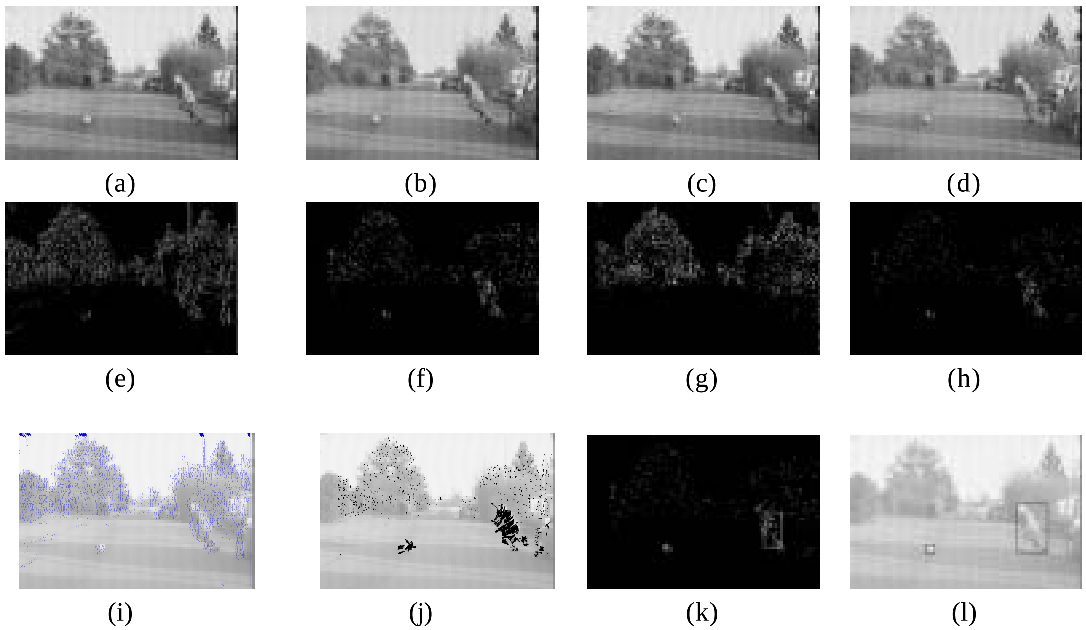

Figure 1: Segmentation results based on static box removal and motion region expansion. (a)(b) Binocular stereo image at frame 50 (left/right). (c)(d) Binocular stereo image at frame 49 (left/right). (e) Edge map of (a). (f) Separated edge layer at disparity range 17-19. (g) Separated edge layer in background region. (h) Clean edge layer in disparity range 17-19 [Remove (g) from (f)]. (i) Motion vectors for edge pixels in (e). (j) Motion vectors for edge pixels in (h). (k) Initial segmentation results for (h). (l) Motion-based expansion result based on (j)(k).

Impact: Since the algorithm does not make any assumption on target shape, it is possible to be used in many applications to detect pedestrians with any pose and to detect targets with different shapes.

Future Work: Such algorithm can be combined with other sensors to further improve the reliability and accuracy that is vital for Intelligent Transportation Systems.

Research Support: This research is supported in part by a gift account for the MIT/MTL Intelligent Transportation Research Center at Massachusetts Institute of Technology.

References:

[1] Yajun Fang, Ichiro Masaki, and Berthold Horn. Distance range based segmentation in intelligent transportation systems: Fusion of radar and binocular stereo. *Proc. of IEEE Intelligent Vehicles Symposium*, 2001.

[2] Yajun Fang, Ichiro Masaki, and Berthold Horn. Depth-based target segmentation in intelligent transportation systems: Fusion of radar and binocular stereo. *Submitted to IEEE Transactions on Intelligent Transportation Systems*, 2002.

[3] Chris Stauffer and W.E.L. Grimson. Adaptive background mixture methods for real-time tracking,. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2,* pages 246–252, 1999.