Range Segmentation Using Visibility Constraints

Leonid Taycher & Trevor Darrell

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: We are looking to produce a foreground segmentation mechanism that would work under widely varying illumination conditions.

Motivation: One of the most important tasks of computer vision in intelligent environments is detecting and tracking various objects (including humans). The state-of-the-art systems use foreground / background classification followed by foreground pixel clustering and cluster analysis. These systems commonly maintain a background model and label all pixels that differ significantly from this model as foreground.

Since a segmentation algorithm for person tracking should work in general environments, it should be robust to frequent and sudden illumination changes, e.g., due to outdoor lighting variations or indoor video projection. Stereo range-based methods have become popular and can satisfy this requirement when dense range data are available. However, many environments have backgrounds that are sparse, such as indoor walls. In these environments simple background range models are problematic, since there is an inherent ambiguity in classifying a valid range value in the location where the model contains an invalid range value. Even when multiple stereo views are available [3], sparse backgrounds can make it difficult to detect people.

Previous Work: Several segmentation methods have been proposed which use background models based on color / intensity [6], stereo range [3] or both [2]. Generally, non-adaptive color-based models suffer from varying illumination. Adaptive color models can be more stable to lighting changes, but can erroneously incorporate objects that stop moving into the background model. Range-based background models can be illumination invariant, but are usually sparse. To avoid ambiguity at undefined background values (and the resulting illumination dependence [1]), they are either used in conjunction with color models, or are built using observations from widely varying illumination and imaging conditions [1].

To overcome the ambiguities in a single view, information about *free space* constraints over multiple views may be used. Various voxel coloring and space carving methods [4] split the space into *voxels* and use color consistency across multiple cameras to locate opaque voxels and to detect *free space*. These methods are quite general, and work with an arbitrary set of monocular views. They also require the construction of a volumetric representation of the scene for reconstruction or segmentation. We are interested in algorithms that perform segmentation solely in the image domain, without computing a volumetric reconstruction.

Approach: Our approach to foreground segmentation is to combine *free space* constraints found from multiple stereo range views. We decide if a given pixel is "foreground" by checking whether there is any *free space* behind it, as seen from other range views. We scan the set of epipolar lines in the other views corresponding to the given pixel, and test whether there are range points on the epipolar lines which indicate empty space behind the given point (Figure 1. This is a similar computation to algorithms proposed for the rendering of the image-based visual hulls [5]. The key difference is that our method takes as input unsegmented noisy range data and evaluates 3-D visibility per ray, while the visual hull method presumes segmented color images as input and simply identifies non-empty pixels along the epipolar lines in other views. We believe ours is the first method for range image segmentation using image-based (non-voxel) freespace computation.

Future Work: We are planning to incorporate this segmentation mechanism into Intelligent Room person tracking and Visual Hull [5] computation.

Research Support: This project is funded with the generous support of the Darpa HumanID program.

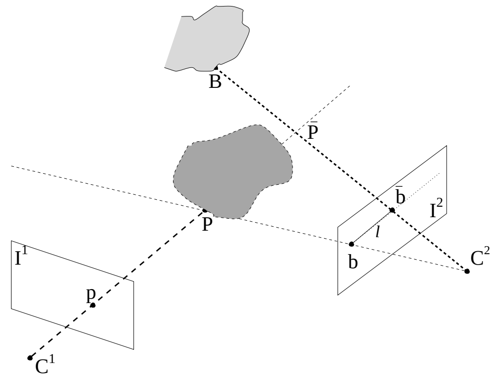

Figure 1: Visibility-based segmentation. Point $\mathbf{B}$ visible in I^2 (projecting to C^2 disparity point $\overline{\mathbf{b}}$) lies behind point $\overline{\mathbf{P}}$ relative to C^2, and thus provides evidence for existence of *free space* behind $\mathbf{P}$ (projecting to $\mathbf{b}$) by demonstrating that $\overline{\mathbf{P}}$ is transparent. Line l contains the oversilhouette for the part of an object lying along ray $[\mathbf{C^1}, \mathbf{P})$

References:

[1] T. Darrell, D. Demirdjian, N. Checka, and P. Felzenszwalb. Plan-view trajectory estimation with dense stereo background models. In *Proc. International Conference on Computer Vision (ICCV'01)*, 2001.

[2] M. Harville, G. Gordon, and J. Woodfill. Foreground segmentation using adaptive mixture models in color and depth. In *Workshop on Detection and Recognition of Events in Video.*, 2001.

[3] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer. Multi-camera multi-person tracking for easyliving. In *IEEE Workshop on Visual Surveillance*, July 2000.

[4] K. N. Kutulakos and S. M. Seitz. A theory of shape by space carving. *Int. Journal of Computer Vision*, 38(3):199–218, July 2000.

[5] Wojciech Matusik, Chris Buehler, Ramesh Raskar, Steven J. Gortler, and Leonard McMillan. Image-based visual hulls. In Kurt Akeley, editor, *Siggraph 2000, Computer Graphics Proceedings*, Annual Conference Series, pages 369–374. ACM Press / ACM SIGGRAPH / Addison Wesley Longman, 2000.

[6] K. Toyama, J. Krumm, B. Brumitt, and B. Meyes. Wallflower: Principles and practice of background maintenance. In *In Proc. International Conference on Computer Vision (ICCV'99)*, 1999.