MULTIPLE VIEW SEGMENTATION AND MATTING

Markus Kettern, David C. Schneider and Peter Eisert

{markus.kettern, david.schneider, peter.eisert}@hhi.fraunhofer.de – Fraunhofer Heinrich Hertz Institute, Berlin, Germany

Abstract

We propose a robust and fully automatic method for extracting a highly detailed transparency-preserving segmentation of a person's head from multiple view recordings, including a background plate for each view. At first, trimaps containing a rough segmentation into foregorund, background and unknown image regions are extracted exploiting the visual hull of an initial foreground-background segmentation. The background plates are adapted to the lighting conditions of the recordings and the trimaps are used to initialize a state-of-the-art matting method adapted to a setup with precise background plates available. From the alpha matte, foreground colours are inferred for realistic rendering of the recordings onto novel backgrounds.

Keywords: Multiview Segmentation, Visual Hull, Matting.

Initial Segmentation. An initial binary segmentation is based on the intensity and RGB variance of the difference image between image I (fig. 1a) and the respective background plate B (fig. 1b). These cues are combined into a foreground likeliness image P (fig. 1c). Segmentation is performed by thresholding P at the most stable region of its cumulative histogram such that between 20% and 90% of the image are covered by the foreground (fig. 1d).

Trimap Generation. A trimap is a common initialization for image matting and consists of three regions: The foreground region M_f will have an α-value of 1, the background M_b will have an α of zero, and M_u contains the pixels for which α has to be inferred. To obtain a trimap, the visual hull [2] of the object is created from the initial segmentations employing an efficient octree representation (fig. 1e). Only voxels visible in the initial segmentations of all views are included into the hull. The backprojection of this hull into each view yields region M_f of the trimap. A dilated version of the initial segmentation yields the region M_u, completing the trimap (fig. 1f).

Background Plate Adaption. Due to shadows, daylight and other factors arising during a recording session, the illumination of B may differ significantly from I. To make the matting process more robust and precise, we estimate the true background colors behind the object to be segmented by a Poisson interpolation [4] over region $M_u \cup M_f$ of I using the gradient of B as guidance field. The adapted background plate is called B^*.

Alpha Matting and Foreground Estimation. A matte associates each image pixel $p \in M_u$ with a transparency value α_p typically thought of as a blending factor between two unknown colours, $\hat{F}_p$ and $\hat{B}_p$. The SampleMatch algorithm [1] stochastically searches for these colors in a space spanning all combinations of image pixels located at the borders $\delta(M_u, M_f)$ and $\delta(M_u, M_b)$. Instead of using the whole border $\delta(M_u, M_b)$ as candidate background colors for each pixel p we only use the colors of a small region around p in B^*. In this way, we decrease search space size and avoid many ambiguities arising from foreground/background similarities. To obtain smoother results, the mattes are regularized with the matting Laplacian of [3] (fig. 1g). Once the transparency α_p of each pixel p is known, the foreground colour can be directly estimated by

$$F_p^* = \frac{(1-\alpha_p)B_p^* - I_p}{\alpha_p} \tag{1}$$

and used to create new composites (fig. 1h).

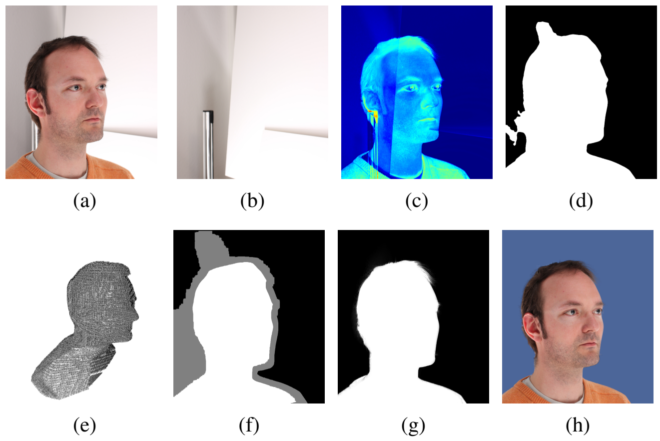

Figure 1: Processing sequence of trimap and matte generation.

References

[1] K. He, C. Rhemann, C. Rother, X. Tiang, and J. Sun. A global sampling method for alpha matting. In *CVPR*, 2011.

[2] A. Laurentini. The visual hull concept for silhouette-based image understanding. *PAMI*, 16:150–162, 1994.

[3] A. Levin, D. Lischinski, and Y. Weiss. A closed-form solution to natural image matting. *PAMI*, 30, 2008.

[4] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. In *ACM SIGGRAPH*, 2003.