High Fidelity Facial Hair Capture

Graham Fyffe
USC Institute for Creative Technologies *

Figure 1: Far left: A captured 3D facial model with hair particles. Pairs: Renderings (left) paired with reference photographs (right).

1 Introduction

Modeling human hair from photographs is a topic of ongoing interest to the graphics community. Yet, the literature is predominantly concerned with the hair volume on the scalp, and it remains difficult to capture digital characters with interesting facial hair. Recent stereo-vision-based facial capture systems (e.g. [Furukawa and Ponce 2010][Beeler et al. 2010]) are capable of capturing extremely fine facial detail from high resolution photographs, but any facial hair present on the subject is reconstructed as a blobby mass. Prior work in facial hair photo-modeling is based on learned priors and image cues [Herrera et al.], and does not reconstruct the individual hairs belonging uniquely to the subject. We propose a method for capturing the three dimensional shape of complex, multi-colored facial hair from a small number of photographs taken simultaneously under uniform illumination. The method produces a set of oriented hair particles, suitable for point-based rendering.

2 Our Approach

We base our approach on multi-view stereo using six high-resolution photographs (2600×3908 pixels) captured simultaneously from six digital SLR cameras surrounding the face. At this resolution, individual facial hairs are imaged and may be detected in 2D using a separable steerable filter quadrature pair. Reconstruction begins with the method of [Furukawa and Ponce 2010], up to the point of Poisson meshing to construct a low-resolution base mesh. We resample the base mesh into a smooth fine mesh, and solve for a vertex displacement map that minimizes a multi-view stereo cost function. To achieve consistent geometric smoothing across data sets and varitations in vertex spacing, we define a novel smoothing term that is invariant to scene scale, and automatically adapts to different resolutions. We next derive a novel stereo matching cost function, which we call *equalized cross correlation*, that properly accounts for both camera shot noise and pixel sampling variance. This allows us to discriminate between reconstruction error caused by non-smooth features (i.e. hairs) versus that caused by noise. It also eliminates any ad-hoc weights for the smoothing term. We reconstruct a color texture map with the facial hair removed, to provide a plausible skin texture to underly the facial hair. First we compute a texture map by blending the input pixel values from the different camera views. Second, we measure the quality of the texture map by inspecting the stereo reconstruction error. We use the quality measure, combined with the steerable filter line detection, as an indicator of the presence of hair, enabling the hair to be painted out automatically, leaving cleanly shaved skin. We then smooth and inset the geometry that was painted over, to reduce the roughness of the hair blob and serve as underlying skin. We construct a cloud of oriented 3D hair particles by analyzing the detected 2D line segments and the facial geometry. The intersection of any pair of 2D line segments from two different camera views is a 3D line segment, and we may place an oriented 3D hair particle at its center. If we were to compute all such hair particles without further constraints, we would end up with many false positives that are not part of actual hair. Thus we define a *hair volume* that extends 1cm out from the skin, and construct only the hair particles that lie inside this volume. We score the hair particles based on how many nearby 2D line segments in any view may be explained by the hair particle. A hair particle may be discarded if its color is similar to the underlying skin texture in all views. Any pixel with a detected 2D line segment in any view retains only the highest scoring hair particle constructed from the 2D line segment. Finally, we revisit the facial texture to replace any pixels that were originally painted over, but turned out not to be occluded by the recovered hair. This produces a complete facial model, with hair particles to explain the fine parallax occuring in the facial hair, and a textured face mesh to explain the underlying color. We show results for a male subject with blonde / light brown hair, a male with short dark hair, and a female subject with dark hair in Figure 1. The facial geometry and hair particles are of high quality.

References

BEELER, T., BICKEL, B., BEARDSLEY, P., SUMNER, B., AND GROSS, M. 2010. High-quality single-shot capture of facial geometry. *ACM Trans. on Graphics (Proc. SIGGRAPH) 29*, 3, 40:1–40:9.

FREEMAN, W. T., AND ADELSON, E. H. 1991. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence 13*, 891–906.

FURUKAWA, Y., AND PONCE, J. 2010. Accurate, dense, and robust multiview stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell. 32*, 8, 1362–1376.

HERRERA, T. L., ZINKE, A., GMBH, G., WEBER, A., AND VETTER, T. Toward image-based facial hair modeling.

*e-mail: fyffe@ict.usc.edu

SIGGRAPH 2012, Los Angeles, California, August 5 – 9, 2012.
ISBN 978-1-4503-1435-0/12/0008