High-Quality Hair Modeling from A Single Portrait Photo

Menglei Chai* Linjie Luo† Kalyan Sunkavalli† Nathan Carr† Sunil Hadap† Kun Zhou*

*State Key Lab of CAD&CG, Zhejiang University †Adobe Research

(a) Input photo & strokes **(b)** 3D hair model **(c)** Side view **(d)** Relighting (ours) **(e)** [Chai et al. 2013]

Figure 1: *From a single photo and a few user strokes (a), our system reconstructs a 3D model that captures the intricate details of the hair (b) & (c). This model can be used for high-quality portrait relighting with more realistic shadowing effects (d) than the previous method (e). Original image courtesy of Ross Heale-Whittle.*

Abstract

We propose a novel system to reconstruct a high-quality hair depth map from a single portrait photo with minimal user input. We achieve this by combining depth cues such as occlusions, silhouettes, and shading, with a novel 3D helical structural prior for hair reconstruction. We fit a parametric morphable face model to the input photo and construct a base shape in the face, hair and body regions using occlusion and silhouette constraints. We then estimate the normals in the hair region via a Shape-from-Shading-based optimization that uses the lighting inferred from the face model and enforces an adaptive albedo prior that models the typical color and occlusion variations of hair. We introduce a 3D helical hair prior that captures the geometric structure of hair, and show that it can be robustly recovered from the input photo in an automatic manner. Our system combines the base shape, the normals estimated by Shape from Shading, and the 3D helical hair prior to reconstruct high-quality 3D hair models. Our single-image reconstruction closely matches the results of a state-of-the-art multi-view stereo applied on a multi-view stereo dataset. Our technique can reconstruct a wide variety of hairstyles ranging from short to long and from straight to messy, and we demonstrate the use of our 3D hair models for high-quality portrait relighting, novel view synthesis and 3D-printed portrait reliefs.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Geometric algorithms;

Keywords: Face reconstruction, hair capture, model-based reconstruction, Shape from Shading

1 Introduction

The growing popularity of digital entertainment, 3D printing, and virtual reality applications has led to a parallel interest in efficient ways to create 3D content, especially personalized 3D face and body models. Despite the recent progress in high-quality 3D acquisition via scanning, most 3D scanning systems require expensive equipment [Beeler et al. 2010; Beeler et al. 2012] or lengthy scanning processes [Paris et al. 2008]. Photos, on the other hand, are easy to capture, and the ability to reconstruct 3D human models directly from single photos can enable 3D content creation for everyone. Hair plays a particularly important role in the way people perceive each other, and modeling the structures of a hairstyle from a single photo is an important and challenging task for 3D content creation.

While human faces can be reconstructed well from a single photo using parametric 3D models [Blanz and Vetter 1999], such parametric models for hair are difficult to find. This is because faces tend to vary in very constrained ways making it possible to represent them using low-dimensional models. Hair on the other hand can exhibit extreme variability and geometric complexity, and current single-view hair reconstruction methods [Chai et al. 2012; Chai et al. 2013] rely on local cues such as hair occlusion and strand smoothness to reconstruct approximate hair models. While these hair models are adequate for image-based rendering and editing tasks, they lack the geometric accuracy for more demanding applications such as relighting (see Fig. 1).

Shape from Shading (SFS) methods can recover detailed geometry (in the form of surface normals) for general objects from a single photo given lighting and material (albedo) estimates [Johnson and Adelson 2011]. Previous techniques have used SFS to refine the geometric details of a rough 3D model by using the lighting inferred from the rough model [Valgaerts et al. 2012; Garrido et al. 2013; Suwajanakorn et al. 2014]. These techniques typically assume that the surface albedo is piece-wise constant – an inaccurate assumption for hair because most hairstyles have smoothly varying hair color. Another challenge in applying SFS-based techniques to hair is that they typically assume a diffuse shading model. Hair, on the other hand, exhibits specular appearance and high-frequency shadowing, leading to reconstruction errors.

The goal of our work is to reconstruct a high-quality hair depth map

from a single portrait photo with minimal user input. Our technique achieves this by combining three components: a coarse depth map reconstructed using occlusions and silhouettes constraints, detailed surface normals inferred from shading, and a novel 3D helical prior that enforces the geometric structures that are typical of hair. We initialize our 3D model by fitting a parametric morphable face model to the input portrait photo and constructing a base shape in the face, hair and body regions using the boundary constraints between these regions based on occlusion and silhouette cues. Next, we estimate the detailed normals by performing SFS computation using the lighting inferred from the face model as well as an *adaptive albedo model* that accounts for the typical color and occlusion variations of hair. Finally, we introduce a 3D *helical hair prior* that reinforces the hair geometric structure, and can be automatically estimated from the input photo. While applying any of these cues individually will lead to poor reconstructions, by combining them, our system is able to produce robust, accurate 3D hair models.

We show that our system can reconstruct a wide variety of hair models ranging from short to long and from straight to messy. Furthermore, it can handle portrait photos captured in the wild, with varying pose and unknown illumination; all the results in this paper have been created from photos downloaded from the Internet. Our method produces significantly better reconstructions than previous single-image methods. It also closely matches a state-of-the-art multi-view stereo method [Fuhrmann et al. 2014] on a multi-view stereo dataset. The accuracy of our reconstruction enables applications like high-quality portrait relighting with realistic shadowing effects. We also show that our hair models can be used to create 3D portrait models that can be 3D printed as high-relief sculptures with compelling geometric details.

Contributions. In summary, our main contributions are:

- A complete system to reconstruct accurate 3D hair models from single photos and sparse input by combining cues based on shading, occlusions, silhouettes and a helical hair prior.
- A novel method to fit helices to a hair photo and a 3D helical hair prior that captures the characteristic geometry of hair structures.
- An adaptive albedo model for effective SFS-based estimation of detailed normals on hair.

2 Related Work

Most single-view reconstruction techniques can be categorized into two main classes: model-based methods that assume certain prior knowledge about the object being reconstructed, and *Shape from X* methods that infer depth from various image cues (as represented by *X*), such as shading, contours and occlusions. Here we only review those techniques that are most relevant to our goal of reconstructing 3D hair models from single portrait photos.

Model-based reconstruction. The variation in the shape of human faces is constrained and can be well characterized by a low dimensional space. This has lead to a plethora of parametric morphable face models that can capture the identities and expressions of the human subjects from single-view inputs [Blanz and Vetter 1999; Vlasic et al. 2005; Cao et al. 2014].

However, it is very difficult, if not impossible, to find a "morphable hair model" for single-view hair reconstruction due to vast variations in the appearance and geometric complexity of hair. Most existing methods rely on low-level primitives, boundary cues, occlusion and orientation cues to reconstruct hair on the strand, or wisp level. Chai et al. [Chai et al. 2012; Chai et al. 2013] demonstrated systems to create a 3D hair model from a single view using hair orientation and face-hair occlusion relationships for creative manipulation on portrait images and videos. From a 2D sketch of hair, Wither et al. [2007] inferred simulation parameters based on the super-helices model [Bertails et al. 2006] to generate a 3D hair model that matches an input sketch. Bonneel et al. [2009] proposed a method to infer the parameters of a hair appearance model to generate hair renderings matching a single photograph. By analyzing hair geometry and orientation with geometric primitives, simulated and procedural examples, Luo et al. [2013] and Hu et al. [2014a; 2014b] demonstrated 3D hair capture systems that preserve plausible hair structures. Another thread of work focus on fine-scale geometric detail refinement based on mesoscopic image features for face models [Beeler et al. 2010] and hair models [Echevarria et al. 2014].

Our approach is inspired by the super-helices hair model [Bertails et al. 2006] that approximates hair structures as piece-wise helices from the input photo. These found helices can then be used to regularize the reconstruction and reinforce the hair structures. This approach is related to previous work on helix curve-fitting from sketches [Cherin et al. 2014] which optimizes for piecewise helical parameters to fit single 2D sketch curves. However, our technique differs from this method in a number of aspects. First, in our case, the 2D helix projections on an input photo are not precisely defined, and we need to first reliably detect local curve tangents from the photo. Second, capturing long-range hair structures requires us to enforce a global 3D helical prior, making local curve information insufficient by itself. We account for this by proposing a novel solution to merge sparse local cues into global clusters that correspond to long helices. Finally, the inherent ambiguities in fitting 3D helices to noisy 2D projections can lead to over-fitting and many false connections. To avoid this, we reduce the optimization space by using a restricted homogeneous helix model and robustly fit as-long-as-possible helix curves with regularization.

In concurrent work, Hu et al. [2015] demonstrate an interactive system to model 3D hair model from a single photo by effectively combining examples from a database of hairstyles. In comparison, our work focuses on improving reconstruction accuracy by relying on image-based cues (such as shading, silhouettes, and gradients). A validation on a multi-view stereo dataset shows that our technique compares favorably with a state-of-the-art multi-view stereo algorithm [Fuhrmann et al. 2014].

Shape from X methods. In the single-lighting case concerned in this work, Shape from Shading (SFS) algorithms analyze the shading in photos to recover high-resolution geometry in the form of surface normals. This is done by expressing observed image intensities as a function of scene properties (reflectance, illumination, geometry) and invert this function to estimate geometry. Since this is a highly under-constrained problem, most SFS techniques make strong assumptions about the scene, like the commonly known Lambertian albedo and directional lighting [Zhang et al. 1999]. SFS has been shown to be better constrained under natural illumination [Johnson and Adelson 2011]. Barron et al. [2012] design complex priors on reflectance, illumination, and geometry to further constrain the optimization, but they can only recover coarse geometry. Additional information (such as coarse geometry [Han et al. 2013]) often helps to improve the robustness and quality of SFS reconstruction. As a result, SFS methods are often used to refine the coarse geometry captured in 3D acquisition systems [Wu et al. 2011; Valgaerts et al. 2012; Garrido et al. 2013; Suwajanakorn et al. 2014]. These systems typically reconstruct coarse geometry, use it to estimate illumination, which is then used for SFS computation. To handle albedo variations, these systems assume piece-wise constant albedo or a sparse set of albedo values. These assumptions are especially inaccurate for hair because most hairstyles have smoothly transitioned hair color. Instead, we introduce an adaptive albedo model for normal estimation using SFS

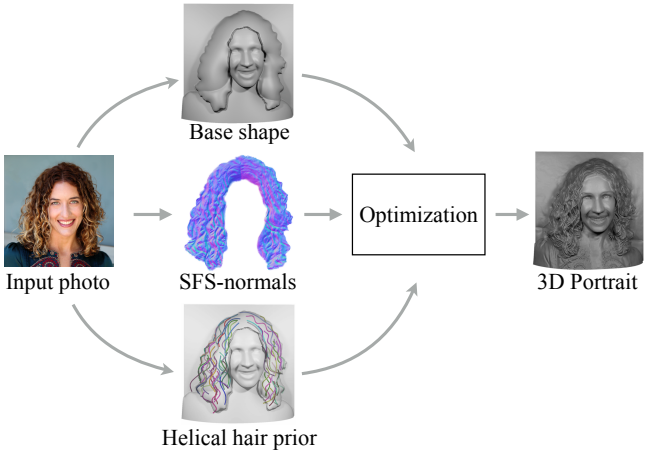


Figure 2: **Overview of our method.** *Given a single input photo, our method uses a single optimization framework to create the final 3D portrait that incorporates the base shape obtained from the boundary constraints, the normal map estimated by SFS (SFS-normals), and the helical hair prior of fitted helices to the input photo. Original image courtesy of vgm8383.*

that can handle albedo variations and ambient occlusion effects on hair.

Besides shading, contours are informative cues to infer 3D shape. Karsch et al. [2013] studied how well different contour cues (e.g. folds, silhouettes and self occlusions) inform shape reconstruction. Sykora et al. [2014] generate global illumination renderings of hand-drawn characters using only the contours and the layering relationships of the components. We augment our hair reconstruction using silhouette constraints.

3 Overview

Given a single portrait photo, our goal is to build a 3D model. In particular, we use a depth map representation for the geometry; this allows us to effectively combine the detailed normals from SFS and various depth cues derived from the input photo in a single optimization framework. To aid us in the reconstruction, we rely on the user to roughly annotate the hair and the body regions in the photo.

To initialize our reconstruction with a rough shape estimate, we construct a *base shape* (Sec. 4) for the face, hair and body regions. We start by fitting a bilinear 3D face model to the detected facial landmarks in the photograph [Blanz and Vetter 1999; Cao et al. 2014]. This gives us the base shape of face with the right pose, identity and expression of the subject. Projecting this model on to the image gives us the face region, Ω_f. We then compute the base shape for body and hair regions specified by the user – Ω_b and Ω_h respectively – using the face model, and occlusion and silhouette constraints. The body region is then constructed based on the body silhouettes and the hair region is built from hair silhouettes, depth constraints of the face model and face-hair occlusion relationships. This step gives us a base shape depth map, that we denote as d^b.

To use SFS to estimate detailed surface normals on hair region (referred to as SFS-normals from now on), we first need to estimate the environment lighting. We use the (low-resolution) surface normals of the face model to infer the environment lighting from the observed images intensities (Sec. 5.1). We then use SFS to estimate per-pixel face and hair surface normals from the observed image intensities based on this inferred lighting. We regularize the albedo in this optimization using a novel adaptive albedo model that accounts for the albedo variations on the hair (Sec. 5.2). We denote the surface normals reconstructed using this step as $\mathbf{n}^{SFS}$.

Combining the base shape and SFS-normals recovers only incomplete and blurred hair structures (Fig. 6); this is due to the approximations that these models make about hair appearance. To better capture the hair structures in the reconstruction, we introduce a *helical hair prior* inspired by the super-helices model [Bertails et al. 2006] that hair strands can be modeled as piece-wise helices. To formulate the prior, we discover helical hair structures from the input photo using a novel RANSAC-based approach. We cluster the pixels of the hair into super-pixels based on hair orientation and proximity using k-way graph cuts (Sec. 6.1). Each super-pixel is then fit with the best 2D projection of a 3D helix on a set of rotated axes. Adjacent super-pixels that can be fit with the same helix are iteratively combined to construct long 2D helix projections. Finally, we recover the true 3D helix parameters for these 2D helix projections, and use the 3D helices as the helical hair prior to constrain the optimization to match these hair structures (Sec. 6.2). We denote the depth of the 3D helical structures that we recover by d^h.

Finally, we reconstruct the final depth map, d, by minimizing an energy function that combines the base shape, SFS-normals, and 3D helical hair prior:

$$\begin{aligned} E(d) &= \lambda_b E_b + \lambda_n E_n + \lambda_h E_h \\ &= \sum_p \lambda_b \left\| d_p - d_p^b \right\|^2 + \lambda_n \left\| \nabla d_p - \mathbf{n}_p^{SFS} \right\|^2 + \lambda_h \left\| d_p - d_p^h \right\|^2, \quad (1) \end{aligned}$$

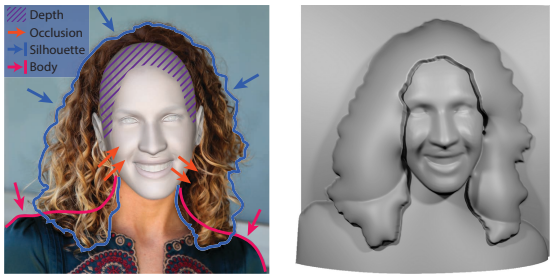
where E_b, E_n and E_h are the energies for the base shape, SFS-normals and the helical hair prior respectively. E_b constrains the depth map to lie close to the base shape (and is enforced more strongly in the face and body regions), E_n requires the gradients of the depth map to explain the SFS surface normals, and finally, E_h enforces the 3D helical hair prior estimated in the hair region. Fig. 2 shows an overview of our system.

4 Base Shape

While SFS reconstruction can capture fine detail in surface reconstructions, they are prone to gross low-frequency errors because the small errors in per-pixel normal estimates accumulate when being integrated. We avoid these errors by computing a base shape using the user strokes specifying the body and hair regions. This base shape serves as a large-scale regularizer on the subsequent SFS-based geometric refinement. We fit a morphable face model to the detected facial landmarks in the photograph. Then we construct different regions of the base shape in a back-to-front occlusion order, i.e., background, face, body and hair, by enforcing boundary constraints based on occlusion and silhouette relationships.

4.1 Face Fitting

We use a morphable face model to estimate the 3D face model from the input photograph. Morphable face models represent face geometry as a linear combination of low-dimensional basis vectors, which are computed using principle component analysis on 3D face geometry data and are designed to capture the variation in face geometry over different identities and expressions. Given a set of detected facial landmarks on an input photograph, we recover the rigid pose and the coefficients of the morphable face model that minimize the distance between the projected landmarks and the detected ones. The recovered rigid pose and basis coefficients define the full face model. In our work, we use a morphable face model trained on FaceWarehouse [Cao et al. 2014], and estimate the identity and expression coefficients using an iterative optimization. For more details, please see [Cao et al. 2014].



(a) Boundary constraints **(b)** Base shape

Figure 3: **Base shape construction.** *We use a set of boundary constraints (a) to construct the base shape (b): the depth of the face model, the occlusion relationships between face and hair, the hair silhouettes and the body silhouettes.*

(a) Fit face model **(b)** Lighting estimate area **(c)**

Figure 4: **Face fitting and lighting estimation.** *We fit a face model to the input photo (a), and then use the computed pure skin region Ω_s (b) to estimate the environment lighting (c).*

4.2 Boundary Constraints

Face and background. We directly assign the depth from the fit face model as the base shape in the face, Ω_f, and set the background depth to the depth of the center-line of the head.

Body. To construct the base shape in the body region, Ω_b, we use the body silhouettes specified by the user strokes along the body boundary $\partial\Omega_b$ (Fig. 3). The body base shape can then be obtained by solving:

$$\min_{d_p} \sum_{\partial\Omega_b} (\|d_p - d_p^*\|^2 + \|\mathbf{n}_p - \nabla\Omega_b\|^2) + \sum_{(p,q)\in\Omega_b} \|\mathbf{n}_p - \mathbf{n}_q\|^2, \quad (2)$$

where d_p^* denote the background depth values, $\nabla\Omega_b$ denotes the gradient of the body silhouette, $\mathbf{n}$ denotes the normals corresponding to the base shape. The first term sets the body base shape depth along the silhouette to the background. The second term enforces that the normals $\mathbf{n}_p$ along the silhouette lie in the same direction as $\nabla\Omega_b$, or in other words, orthogonal to the viewing direction. Finally, the third term encourages the normals to be smooth.

Hair. The base shape in the hair region, Ω_h, can be constructed using a modification of Eqn. 2. The term d_p^* denotes the depths of the face and body that have been constructed previously and ensures that the hair that touches the face and the body stays on it. Finally, the normal constraints on $\partial\Omega_h$ apply to the normals of the hair silhouettes.

Combining the depths estimated in the face, body, and hair regions give us the base shape, d^b that is used to constrain the final reconstruction following Eqn. 1. Fig. 3 shows the hair and body boundary constraints indicated by the user input and the corresponding base shape constructed from them.

5 Shape from Shading

The base shape reconstructed in the previous section (Fig. 3) captures the gross shape of the hair, but does not capture any of the detail in the geometry. In this section, we should how we infer per-pixel surface normals in the hair region to augment the base shape. To do this, we estimate the lighting in the photograph using the morphable face model, and use this lighting in a SFS-based scheme to recover surface normals while enforcing an adaptive albedo model that accounts for variations in hair color.

5.1 Lighting Estimation

Fig. 4 shows the morphable face model fit to the photograph gives us very coarse geometry that is restricted to the face region. We use this face geometry to estimate the lighting in the image and use this lighting to perform SFS computation to recover per-pixel SFS-normal $\mathbf{n}_p^{SFS}$ in the hair region.

We project the 3D face model back to the image plane to estimate a per-pixel depth, d_p^f, and normal, $\mathbf{n}_p^f$, in the face region, Ω_f. We average the the pixel intensities in the region to estimate the average skin color c_s. To remove the regions with different albedo color (eyes, mouth, facial hair) or shadows, we shrink Ω_f to estimate the pure skin regions, Ω_s, by clustering the chrominance values of the pixels in the face region. Fig. 4 shows an example of our face model fitting and skin area segmentation.

While classic SFS from images captured under frontal white point light source is an ill-posed problem [Zhang et al. 1999; Durou et al. 2008], recent work has shown that shape estimation under natural lighting is better constrained and far more accurate [Johnson and Adelson 2011; Oxholm and Nishino 2012; Barron and Malik 2012]. Following the work of Johnson and Adelson [2011], we represent the scene illumination using a quadratic lighting model, $\mathcal{L}(\mathbf{A}, \mathbf{b}, c)$. The shading induced by this lighting model at every pixel in the scene is given by:

$$\mathbf{I}_p = \mathcal{L}(\mathbf{A}, \mathbf{b}, c) * n_p = \mathbf{n}_p^T \mathbf{A} \mathbf{n}_p + \mathbf{b}^T \mathbf{n}_p + c, \quad (3)$$

where $\mathbf{I}_p$ and $\mathbf{n}_p$ are the observed color and surface normal at pixel p respectively, and $\mathbf{A}$, $\mathbf{b}$, and c are the parameters of the lighting model. This quadratic lighting model captures a wide range of lighting (both low-frequency and directional components) and also allows some deviations from Lambertian reflectance. But note that the albedo at pixel p is not accounted for in this model; for uniform albedo regions it gets rolled into the lighting parameters, and later we show how we can account for variations in the albedo during normal estimation.

To estimate the parameters of this lighting model, previous methods have used calibration devices [Johnson and Adelson 2011], or initial shape estimates [Han et al. 2013; Haque et al. 2014]. In our case, we use the coarse face geometry reconstructed using the morphable face model, $\mathbf{n}_p^f$, to estimate the lighting parameters. We estimate the lighting coefficients by minimizing the following linear least squares system:

$$\min_{\mathbf{A},\mathbf{b},c} \sum_{p\in\Omega_s} ||\mathcal{L}(\mathbf{A}, \mathbf{b}, c) * \mathbf{n}_p^f - \mathbf{I}_p||^2. \quad (4)$$

In practice, we regularize this optimization using $\lambda\|\mathbf{A}\|^2 + \lambda\|b\|^2 + \lambda c^2, \lambda = 0.01$ for our results. We solve for these three parameters in every color channel independently, thus we have $\mathbf{A}^l, \mathbf{b}^l, c^l, l \in \{R, G, B\}$. Because the albedo is not accounted for in this model, we restrict the error function to the detected facial skin region Ω_s to ensure a roughly uniform albedo. Fig. 4 shows an example of the lighting estimated using this method.

(a) (b) (c) (d)

***Figure 5:* Normal estimation.** *We use an adaptive albedo model (a) to account for the albedo variations on face and hair that defines a relative compensation r_p (b) to handle both global albedo differences between face and hair (blue boxes) and local shading variation (yellow boxes) within the hair region (green dashed lines). The estimated normal map is shown in (c). (d) shows the comparative result assuming constant hair albedo.*

5.2 Normal Estimation

Given the pre-computed lighting, SFS-normals $\mathbf{n}_p^{SFS}$ can be recovered by minimizing the following data term:

$$E_p(\mathbf{n}_p) = \sum_{l\in\{R,G,B\}} \left\| \mathcal{L}(\mathbf{A}^l, \mathbf{b}^l, c^l) * \mathbf{n}_p - \mathbf{I}_p^l \right\|^2. \quad (5)$$

There are two issues with the normals estimated using this energy term. Firstly, optimizing for per-pixel normals independently will lead to noisy estimates. We resolve this by adding two pair-wise smoothness terms:

$$E_s(\mathbf{n}_p) = \sum_{q\in\mathcal{N}(p)} \|\mathbf{n}_p - \mathbf{n}_q\|^2 \quad (6)$$

$$E_i(\mathbf{n}_p) = \|\nabla \times \mathbf{n}_p\|^2. \quad (7)$$

The first constraint enforces normal similarity between neighboring pixels, and the second term enforces that the estimated normals are integrable by penalizing the curl of the normal vector field around each pixel.

Secondly, because we did not account for the albedo in the lighting model (Eqn. 3), deviations from the mean skin color (which is subsumed into the lighting model) will lead to errors in the normal estimates. This is especially problematic in the hair regions because the albedo of hair is often dramatically different from that of face, and hair regions typically have a lot of albedo variation, shadowing, and ambient occlusion that are not handled in the shading model. To account for these effects, we introduce an *adaptive albedo model* that defines a *relative* compensation r_p at each pixel in order to handle both the global albedo differences and local shading variation:

$$r_p = r_0 r'_p, \quad r'_p \in [r_{min}, r_{max}], \quad (8)$$

where r'_p indicates the local per-pixel albedo variation at each pixel, and r_0 accounts for global deviations in the albedo from the mean skin color. The value r_0 in the hair region is set to the ratio of the average color of hair over face, i.e., $r_0 = (\sum_{p\in\Omega_h} \mathbf{I}_p/|\Omega_h|)/(\sum_{p\in\Omega_s} \mathbf{I}_p/|\Omega_s|)$. r'_p is a per-pixel grayscale compensation term bounded by r_{min} and r_{max} that accounts for local shading and shadowing effects.

We then modify Eqn. 5 to account for r_p:

$$E_p(\mathbf{n}_p) = \sum_{l\in\{R,G,B\}} \left\| r_p \mathcal{L}(\mathbf{A}^l, \mathbf{b}^l, c^l) * \mathbf{n}_p - \mathbf{I}_p^l \right\|^2. \quad (9)$$

To ensure that this optimization is well-constrained, we impose a smoothness constraint for r_p:

$$E_r(r_p) = \sum_{q\in\mathcal{N}(p)} \|r_p - r_q\|^2. \quad (10)$$

The final energy combines E_p, E_s, E_i and E_r and we solve this constrained nonlinear least-squares system iteratively on a patch basis. Within each patch, Levenberg-Marquardt method is adopted with the unit length constraint on the of normal vectors using the method in Johnson and Adelson [2011]. During each iteration, patches are solved in a sweep-line order, allowing updated information to be propagated across overlapping patches to ensure proper global constraints.

Fig. 5 shows an example of our normal estimation method. As this figure shows, our relative compensation term, r_p, accounts for the significant differences between the albedo of the face and hair regions, as well as smooth albedo variations along the hair using the relative compensation, giving us accurate surface normals that are not corrupted by these variations. On the other hand, assuming a constant albedo, as is done in previous SFS-based techniques, produces very poor surface normals (Fig. 5(d)).

5.3 Shape Integration

We merge the SFS-normals and the base shape to reconstruct a depth map with the global shape of the base shape and geometric details in the SFS-normals. Combining depth and normal information can be done in a similar way to Nehab et al. [2005], and corresponds to the first two terms in Eqn. 1, where the parameter λ_n and λ_b control how strongly the SFS normal or the original base shape is to be preserved. Visually, we found that enhancing the details on the face too much tends to lead to artifacts. In contrast, the hair region requires more enhancement. Therefore, we use $\lambda_n = 0.6, \lambda_b = 0.4$ and $\lambda_n = 0.1, \lambda_b = 0.9$ for the hair and face regions respectively. This integrated shape is used in Sec. 6.2 to disambiguate the parameters when a 2D helix is projected back to 3D.

Fig. 6(f) shows the result of integrating only the SFS normals and the base shape. As can be seen, adding the SFS normals to the base shape brings out a significant amount of detail in the hair, face, and body regions. In some cases, the details on the face and the body may not correspond to true geometric structures (for e.g., in the eyes), but we found that they added to the visual quality of the result. This is similar to sculptural techniques that use geometry to depict texture detail (like the iris in sculpted busts).

6 Helical Hair Prior

Combining the base shape with SFS-normals gives us reconstructions with nice visual detail. However, as can be seen in Fig. 6, the reconstruction may not capture the rich structural detail in the hair region. Hair has complex BRDF and local lighting effects that violate our shading model. Our patch-based reconstruction is robust to this but at the cost of blurring out some of the hair detail. In this section, we discuss how we use a geometric prior for hair to capture intricate hair structures. In particular, we rely on the observation that hair can be approximated well by piece-wise 3D helices [Bertails et al. 2006]. We infer these structures from the input photo by clustering pixels with consistent hair orientation and color and fitting 2D projected helical models to the clusters. We then use the integrated depth (Sec. 5.3) to recover the true 3D helices. Finally, we enforce depth continuity along these inferred 3D helices as the energy term E_h in Eqn. 1.

6.1 Super-Pixel Clustering

We first compute a robust orientation map of the photo using a bank of oriented filters that are uniformly sampled in $[0, \pi)$. By analyzing the convolution response at each angle, we choose the orientation θ_p with maximum response and calculate the corresponding confi-

(a) Clustering **(b)** Single-cluster fit **(c)** Multi-cluster fit **(d)** Projected helices **(e)** With the prior **(f)** Without the prior

Figure 6: **Helix fitting and the helical hair prior.** *We first perform super-pixel clustering based on color and orientation similarity (a). Single-cluster fitting (b) fits the best helix for each cluster followed by multi-cluster fitting (c) that merges neighboring compatible helices into longer ones. These helices are then projected back to the model to resolve ambiguity (d). Using these 3D helices as the helical hair prior, our optimization can recover intricate hair structures (e), compared to the incomplete and blurred ones without this prior (f).*

dence value c_p by measuring how distinct it is compared to other angles, as in Chai et al. [2012].

We then sample a set of cluster seed pixels that satisfies the following conditions: its confidence is a) no less than a threshold (i.e., $c_p \geq c_{min}$), and b) locally maximal in a neighboring window of size 5. We use these samples as cluster centers, and apply k-way graph cuts [Karypis and Kumar 1998] to segment the image into super-pixels, $\mathcal{C}$, using both color and orientation distance:

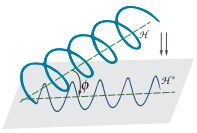
$$w_{(p,q)} = w_c \left| c_p - c_q \right| + w_\theta \left| \theta_p - \theta_q \right|. \tag{11}$$

6.2 Helix Fitting

Single-Cluster Helix Fitting. A single 3D helix $\mathcal{H}$ can be parametrized within a local frame (with its main axis aligned with the z-direction) in terms of a coordinate t.

$$\mathcal{H}(t) = \begin{cases} x(t) &= a\ \cos(t) + m_x, \\ y(t) &= b\ \sin(t) + m_y, \\ z(t) &= c\ t + d_z. \end{cases} \tag{12}$$

Projecting it to a 2D plane (while denoting rotation about y-axis with ϕ) gives us a generalized trochoid curve $\mathcal{H}^*$ (We assume an orthogonal projection to simplify the formulation):



$$\mathcal{H}^*(t) = \begin{cases} x(t) &= a\ \sin(\phi)\ \cos(t) + c\ \cos(\phi)\ t + m_x, \\ y(t) &= b\ \sin(t) + m_y, \end{cases} \tag{13}$$

with the tangent given by:

$$\dot{\mathcal{H}}^*(t) = \frac{b\ \cos(t)}{-a\ \sin(\phi)\ \sin(t) + c\ cos(\phi)}. \tag{14}$$

We fit the 2D helix model to each super-pixel cluster, $\mathcal{C}_i$, estimated in the previous step. Given the super-pixel points with 2D position $\mathbf{p}$ and orientation tangent $\dot{\mathbf{p}}$, we can fit a single helix to them by solving the following minimization problem:

$$\min_{t_\mathbf{p}, \mathcal{H}^*} \sum_\mathbf{p} \left\| \mathbf{p} - \mathcal{H}^*(t_\mathbf{p}) \right\|^2 + \left\| \dot{\mathbf{p}} - \dot{\mathcal{H}}^*(t_\mathbf{p}) \right\|^2 + w_r \left\| t_\mathbf{p} \right\|^2, \tag{15}$$

where the last term penalizes over-fitting by minimizing the parameter coordinate of every point.

The fitting process also needs to determine the local orientation of the 2D frame in which the helix is defined. Incorporating it in the energy term above will complicate the optimization significantly. Instead, we uniformly sample a set of helix axes, and estimate the remaining parameters by solving Eqn. 15 for each axis. The axis with the minimal fitting error is chosen. We found that uniformly sampling 32 orientations was sufficient for our experiments.

Multiple-Cluster Helix Fitting. The 2D projected helices fit to single super-pixel clusters are often too short and ambiguous to generate 3D helices that are long and accurate enough to extract large-scale hair structures (see Fig. 6(b)). Therefore, we construct long helices by incrementally merging compatible neighboring single-cluster segments into multi-cluster helices.

In each iteration of this process, we start with a single-cluster helix $\mathcal{H}_i^*$, collect all its neighboring clusters $\mathcal{C}_{\mathcal{H}^*}$, and for each pair of clusters, re-fit a new helix, $\mathcal{H}_{i+1}^*$. We measure the fitting score as the average fitting error across all the merged clusters:

$$e(\mathcal{H}^*) = \frac{\sum_{\mathbf{p} \in \mathcal{C}_{\mathcal{H}^*}} \left\| \mathcal{H}^*(t_\mathbf{p}) - \mathbf{p} \right\|^2}{\left| \mathcal{C}_{\mathcal{H}^*} \right|}. \tag{16}$$

If the fitting score is below a set threshold $e \geq e_{min}$ (set to about 25), the clusters are merged and used for further extension. When this iteration is terminated for every helix, we further remove redundant helices that belong to an identical set of initial clusters (only keep the one with minimal fitting error), and all helices with length less than a threshold l_{min}, which is set to 50 pixels.

To improve the fitting performance, we make the assumption that the new helix, $\mathcal{H}_{i+1}^*$, shares the same axis as the previous helix $\mathcal{H}_i^*$, so that we don't need to sample axes again as in the single-cluster fitting. Fig. 6(c) shows the multi-cluster 2D helices that we are able to detect using this method.

Depth-Guided 3D Helix Estimation. Up to this point, we have inferred a set of sparsely distributed 2D helix projections. We now recover their corresponding 3D structures by making use of the estimated depths using the base shape and SFS normals (Sec. 5.3), and enforce this 3D structure in the portrait reconstruction process. In order to recover a 3D helix from the 2D projections we have inferred, we still need to estimate the rotation angle relative to the projection plane, ϕ, and the displacement vector d_z along the projection axis (see Eqn. 12). The value ϕ encodes the convex/concave ambiguity when a 3D helix is projected on to a 2D plane, and plays a critical role in resolving the 3D structure.

The unknown depth component of the projected 3D helix is then:

$$d\left(\mathcal{H}^*(t)\right) = \cos(\phi)\ (a\ \cos(t) + b\ \sin(t)) + c\ \sin(\phi)\ t + m_z. \tag{17}$$

In order to estimate it, we rely on the model depth, d_p reconstructed using only the base shape and SFS-normals in Sec. 5.3. We project the 2D helices on to this model, and sample the depth at pixels (t_p, d_p) along the project 2D helices. We solve for optimal values

Our method [Chai et al. 2013]

Figure 7: **Relighting results.** *Our method enables realistic hair-face shadowing and hair self-shadowing effects compared to [Chai et al. 2013]. See the accompanied video for more results.*

of ϕ and d_z that best fit this depth map for complete 3D helices (still within the local frame that rotates around the z-axis):

$$\arg\min_{\phi,d_z} \sum_p \|d\left(\mathcal{H}^*(t_p)\right) - d_p\|^2. \tag{18}$$

The recovered parameters are substituted in Eqn. 17 to recover the 3D helix depth, d_p^h. This 3D helix depth is used in Eqn. 1 to improve the reconstruction of the final result. Fig. 6(e) shows the reconstructed geometry with our helical hair prior. In contrast to the result without this prior, Fig. 6(f), our result faithfully recovers the intricate geometry in the hair region. Please see more comparisons in the accompanied video.

7 Results

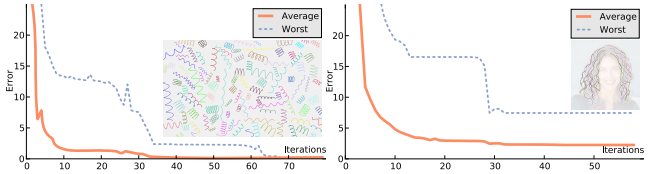
To demonstrate the robustness of our method, we apply our method to a variety of portrait photos captured in the wild, with a variety of hairstyles ranging from short to long and from straight to messy. The reconstructed hair models are shown in Figs. 1 and 13. For each example in Fig. 13, we show the input photo, the reconstructed 3D hair model as well as two side views of the model with and without the texture. In the accompanied video, we also show our results in continuous rotation for better visualization. As shown in these results, our method can faithfully recover both long and continuous wisp structures thanks to our helical hair prior as well as fine-scale hair details thanks to the SFS normal estimation and integration.

Our high quality hair model can significantly improve portrait relighting applications. Fig. 7 shows two relighting examples under changing global illumination with a comparison to the results using the portrait model generated by [Chai et al. 2013]. In our implementation, we not only take into account the surface geometry, but also grow individual hair strands occupying the hair volume bounded by the surface as in [Chai et al. 2013] and use a realistic hair appearance model [Marschner et al. 2003] with self-shadowing to render the relit hair. Thanks to our high quality portrait models,

Figure 8: *3D printed high-relief portraits from different views. From left to right, original images courtesy of vgm8383, Qsimple and Denise Mahoney.*

our relighting results show better face-hair shadowing and hair self-shadowing effects as well as realistic moving highlights on the hair as the lighting changes. In contrast, the relighting results by [Chai et al. 2013] look flat and unrealistic as if the entire hair region was on a smooth surface. Please refer to the accompanied video for full relighting results.

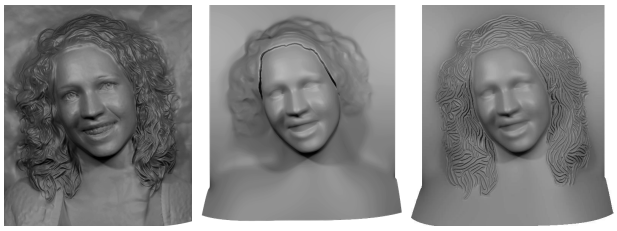
We can create high quality 3D portrait models with our hair models and produce high-relief portrait sculptures using 3D printing. A few printed portraits are shown in Fig. 8 from different static views. These models are shown under dynamic views in the accompanied video; the added viewing dimensionality becomes quite striking when compared to the original photograph due to the compelling geometric details produced by our method.



(a) Synthetic data (b) Real data

Figure 9: *Convergence curves of our helix fitting solver on synthetic curves (left), and a real photo (Fig. 6) (right) with average and worst (the helix with largest error in that iteration) curves shown respectively. Error is defined as average pixel distance to the pixel clusters per-iteration.*

Analysis of Helix Fitting Solver. To evaluate the convergence and robustness of our helix fitting solver, we test it on both synthetically generated curves as well as helix clusters from real photos. For synthetic testing, we generate 1000 2D helix curves with parameter values randomly chosen within large ranges, and rasterize them into pixel clusters (similar to Sec. 6.1). As the fitting error convergence curves show (Fig. 9), our solver finds correct solutions for most synthetic samples with close-to-zero errors, and converges to reasonably good results on real photos. Clusters with non-uniform helical shapes (varying period or circle radius) can introduce large errors, due to our uniform helix model. A more generalized model with varying parameters may be helpful but will make the optimization more difficult. We chose this uniform model since we do not seek to recover super long strands but consistent segments with moderate length to capture local hair structures.



Our method [Chai et al. 2013] [Echevarria et al. 2014]

Figure 10: **Comparison with previous methods.** *Our reconstruction result reveals both gross geometric structures and realistic fine-scale details compared to [Chai et al. 2013] and [Echevarria et al. 2014] (applied to our base shape).*

Evaluation and Comparison. We first compare our method with [Chai et al. 2013] and [Echevarria et al. 2014] on reconstruction quality in Fig. 10. Since Echavarria et al. [2014] is a stylized refinement method on coarse hair geometry, we apply their method on the base shape of hair for fair comparison. As shown in the comparison, although the method of [Echevarria et al. 2014] creates stylized hair features suitable for 3D fabrication, it lacks reconstruction accuracy in both aggregate wisp structures and fine-scale hair details compared to our method. The method of [Chai et al. 2013] can reconstruct approximate hair shape based on the boundary constraints, hair occlusion and strand smoothness cues, but it fails to recover any fine geometric details of the hair as seen in our result.

To quantitatively evaluate our reconstruction accuracy, we also compare our method to a state-of-the-art multi-view stereo method [Fuhrmann et al. 2014] on a multi-view dataset used for high quality hair reconstruction in Hu et al. [2014a]. We choose to evaluate our method on a real dataset instead of synthetic ones because real datasets reflect the complex nature of real-world hairstyles and lighting conditions. While we use all of the 56 multi-view images for the multi-view stereo reconstruction, we only use one selected frontal photo to reconstruct our model. The difference map is computed as the depth difference after optimal rigid alignment is applied to the model as shown in Fig. 11. Our average reconstruction difference compared to the multi-view reconstruction result is about 2cm. Despite the overall better reconstruction accuracy of multi-view stereo methods, the result is noisy and discontinuous which reaffirms the difficult nature of hair as a reconstruction target. Visually, our result looks more coherent and smooth thanks to our helical hair prior for reinforcing the hair structures. In addition, our result reveals very fine-scale hair detail as a result of our novel helix prior and SFS pipeline; this is not recovered by the multi-view stereo method.

Input Resolution. In our experiments, higher resolution input photos can produce visually better results with finer details and clearer hair structures, but the computation cost will also increase proportionally. As a practical choice, all results in this paper are generated from input photos of about 800×600 pixels, which can produce good results within acceptable time.

User Interaction. The only user interaction needed is to guide image segmentation with simple strokes. It takes less then a minute for an untrained average user to finish so. And in practice, since this interaction is quite straightforward, different user inputs led to almost identical results and had little influence on following steps. Please see the accompanying video for an example user interaction session.

Parameters. We use the default set of parameters for all our examples except the multiple-cluster merge threshold e_{min} and the regularization weight w_r in Eqn. 15 when denser and curlier helix fitting are desired for some highly curly hairstyles. e_{min} can be decreased to allow longer and denser helices; and w_r can be relaxed to allow more curly helices. We summarize important parameters in Table. 1 for easy look-up.

λ_b	λ_n	λ_h	r_{min}	r_{max}	w_c	w_θ	c_{min}	w_r
0.4	0.6	0.1	0.25	4	1	0.1	0.3	0.1
Eq. 1			Eq. 8		Eq. 11			Eq. 15

Table 1: *Parameter values used in our experiments.*

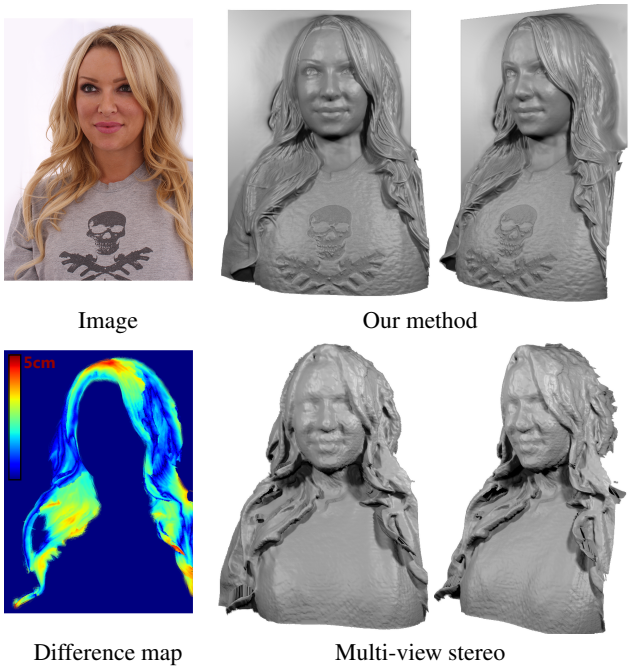


Image Our method

Difference map Multi-view stereo

Figure 11: **Comparison with multi-view stereo.** *Our single-view reconstruction result closely matches the result of the state-of-the-art multi-view stereo method [Fuhrmann et al. 2014] on a multi-view dataset (with an average difference of 2cm) and preserves more coherent and detailed hair structures.*

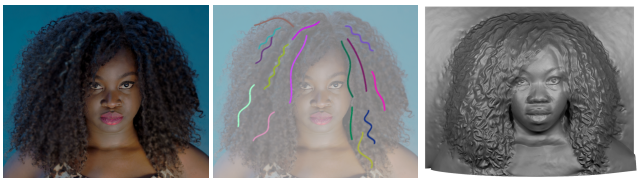


Figure 12: **Limitations.** *For frizzy hairstyles our method may fail to generate a reliable helical hair prior resulting in loss of geometric details. Original image courtesy of Peter Grifoni.*

Timings. For a typical input photo of 800×600 pixels, our prototype system takes about 10 minutes to process all subsequent steps, including about 5 minutes for normal estimation, about 5 minutes for helix fitting and less than 1 minute for final shape integration.

Limitations. While our technique can handle a wide range of hairstyles, it has problems with frizzy ones (Fig. 12). The silhouettes for such examples are not well defined and the hair strands are too wavy to be resolved by our helix fitting method. And our technique may not be able to generate accurate structures for some styled hairstyles such as braids. Another limitation is that our morphable face model [Cao et al. 2014] cannot fit profile portraits since landmark detection may fail. For relighting, the highlights may

Figure 13: **More results.** *For each example, we show the input photo, the 3D hair model and its two side views with and without texture. From top to bottom, original images courtesy of Qsimple, Denise Mahoney, Qsimple, Chris Zerbes and Alex Masters.*

look static because they are partially baked into the albedo layer. A robust method for single image specularity decomposition would help to resolve this issue but is still a very challenging problem.

8 Conclusion

We have demonstrated a system for high quality 3D hair reconstruction from a single photo with minimal user input. By effectively combining single-view depth cues such as shading, silhouettes and occlusions as well as a novel helical hair prior in an optimization framework, we are able to achieve reconstruction accuracy closely matched to a state-of-the-art multi-view stereo algorithm on a multi-view dataset. Our method can handle a variety of hairstyles in the wild, from short to long and from straight to messy. Our accurate hair models enable high-quality portrait relighting with realistic shadowing effects on the complex hair structures. We can create high quality 3D portrait models with our hair models and produce high-relief portrait sculptures using 3D printing.

In the future we would like to further enhance our system to cover a broader range of content. By applying stronger shape priors, such as a hairstyle database, we can extend our method to build a full 3D hair model beyond a depth map. By extending our shape templates to include the torso and body, we can reconstruct portraits that capture more of the human pose and expression. Finally, our system is tuned to generate compelling results from frontal portraits, additional exploration is needed to handle both profile and extreme views which may have more challenging occlusions.

Acknowledgment

We would like to thank Hao Li for sharing the multi-view stereo data, the Flickr users for letting us use their work under the Creative Commons License, and the anonymous reviewers for their constructive comments. This work is partially done when Menglei

Chai was an intern at Adobe Research. Kun Zhou was supported in part by NSFC (No. 61272305), and National Program for Special Support of Eminent Professionals of China.

References

BARRON, J. T., AND MALIK, J. 2012. Color constancy, intrinsic images, and shape estimation. In *ECCV*, Springer, 57–70.

BEELER, T., BICKEL, B., BEARDSLEY, P., SUMNER, B., AND GROSS, M. 2010. High-quality single-shot capture of facial geometry. *ACM Trans. Graph. 29*, 4, 40:1–40:9.

BEELER, T., BICKEL, B., NORIS, G., BEARDSLEY, P., MARSCHNER, S., SUMNER, R. W., AND GROSS, M. 2012. Coupled 3D reconstruction of sparse facial hair and skin. *ACM Trans. Graph. 31*, 4, 117:1–117:10.

BERTAILS, F., AUDOLY, B., CANI, M.-P., QUERLEUX, B., LEROY, F., AND LÉVÊQUE, J.-L. 2006. Super-helices for predicting the dynamics of natural hair. *ACM Trans. Graph. 25*, 3, 1180–1187.

BLANZ, V., AND VETTER, T. 1999. A morphable model for the synthesis of 3D faces. In *Proceedings of SIGGRAPH*, ACM, 187–194.

BONNEEL, N., PARIS, S., VAN DE PANNE, M., DURAND, F., AND DRETTAKIS, G. 2009. Single photo estimation of hair appearance. In *EGSR*, Eurographics Association, 1171–1180.

CAO, C., WENG, Y., ZHOU, S., TONG, Y., AND ZHOU, K. 2014. FaceWarehouse: A 3D facial expression database for visual computing. *IEEE Trans. TVCG 20*, 3, 413–425.

CHAI, M., WANG, L., WENG, Y., YU, Y., GUO, B., AND ZHOU, K. 2012. Single-view hair modeling for portrait manipulation. *ACM Trans. Graph. 31*, 4, 116:1–116:8.

CHAI, M., WANG, L., WENG, Y., JIN, X., AND ZHOU, K. 2013. Dynamic hair manipulation in images and videos. *ACM Trans. Graph. 32*, 4, 75:1–75:8.

CHERIN, N., CORDIER, F., AND MELKEMI, M. 2014. Modeling piecewise helix curves from 2D sketches. *Comput. Aided Des. 46*, 258–262.

DUROU, J.-D., FALCONE, M., AND SAGONA, M. 2008. Numerical methods for shape-from-shading: A new survey with benchmarks. *Comput. Vis. Image Underst. 109*, 1, 22–43.

ECHEVARRIA, J. I., BRADLEY, D., GUTIERREZ, D., AND BEELER, T. 2014. Capturing and stylizing hair for 3D fabrication. *ACM Trans. Graph. 33*, 4, 125:1–125:11.

FUHRMANN, S., LANGGUTH, F., AND GOESELE, M. 2014. MVE - a multi-view reconstruction environment. In *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 11–18.

GARRIDO, P., VALGAERT, L., WU, C., AND THEOBALT, C. 2013. Reconstructing detailed dynamic face geometry from monocular video. *ACM Trans. Graph. 32*, 6, 158:1–158:10.

HAN, Y., LEE, J.-Y., AND KWEON, I. S. 2013. High quality shape from a single RGB-D image under uncalibrated natural illumination. In *ICCV*, IEEE, 1617–1624.

HAQUE, S. M., CHATTERJEE, A., AND GOVINDU, V. M. 2014. High quality photometric reconstruction using a depth camera. In *CVPR*, IEEE, 2283–2290.

HU, L., MA, C., LUO, L., AND LI, H. 2014. Robust hair capture using simulated examples. *ACM Trans. Graph. 33*, 4, 126:1–126:10.

HU, L., MA, C., LUO, L., WEI, L.-Y., AND LI, H. 2014. Capturing braided hairstyles. *ACM Trans. Graph. 33*, 6, 225:1–225:9.

HU, L., MA, C., LUO, L., AND LI, H. 2015. Single-view hair modeling using a hairstyle database. *ACM Trans. Graph. 34*, 4, 125:1–125:9.

JOHNSON, M. K., AND ADELSON, E. H. 2011. Shape estimation in natural illumination. In *CVPR*, IEEE, 2553–2560.

KARSCH, K., LIAO, Z., ROCK, J., BARRON, J. T., AND HOIEM, D. 2013. Boundary cues for 3D object shape recovery. In *CVPR*, IEEE, 2163–2170.

KARYPIS, G., AND KUMAR, V. 1998. Multilevel K-way partitioning scheme for irregular graphs. *J. Parallel Distrib. Comput. 48*, 1, 96–129.

LUO, L., LI, H., AND RUSINKIEWICZ, S. 2013. Structure-aware hair capture. *ACM Trans. Graph. 32*, 4, 76:1–76:12.

MARSCHNER, S. R., JENSEN, H. W., CAMMARANO, M., WORLEY, S., AND HANRAHAN, P. 2003. Light scattering from human hair fibers. *ACM Trans. Graph. 22*, 3, 780–791.

NEHAB, D., RUSINKIEWICZ, S., DAVIS, J., AND RAMAMOORTHI, R. 2005. Efficiently combining positions and normals for precise 3D geometry. *ACM Trans. Graph. 24*, 3, 536–543.

OXHOLM, G., AND NISHINO, K. 2012. Shape and reflectance from natural illumination. In *ECCV*, Springer, 528–541.

PARIS, S., CHANG, W., KOZHUSHNYAN, O. I., JAROSZ, W., MATUSIK, W., ZWICKER, M., AND DURAND, F. 2008. Hair photobooth: Geometric and photometric acquisition of real hairstyles. *ACM Trans. Graph. 27*, 3, 30:1–30:9.

SUWAJANAKORN, S., KEMELMACHER-SHLIZERMAN, I., AND SEITZ, S. M. 2014. Total moving face reconstruction. In *ECCV*, Springer, 796–812.

SÝKORA, D., KAVAN, L., ČADÍK, M., JAMRIŠKA, O., JACOBSON, A., WHITED, B., SIMMONS, M., AND SORKINE-HORNUNG, O. 2014. Ink-and-ray: Bas-relief meshes for adding global illumination effects to hand-drawn characters. *ACM Trans. Graph. 33*, 2, 16:1–16:15.

VALGAERTS, L., WU, C., BRUHN, A., SEIDEL, H.-P., AND THEOBALT, C. 2012. Lightweight binocular facial performance capture under uncontrolled lighting. *ACM Trans. Graph. 31*, 6, 187:1–187:11.

VLASIC, D., BRAND, M., PFISTER, H., AND POPOVIĆ, J. 2005. Face transfer with multilinear models. *ACM Trans. Graph. 24*, 3, 426–433.

WITHER, J., BERTAILS, F., AND CANI, M.-P. 2007. Realistic hair from a sketch. In *SMI*, IEEE, 33–42.

WU, C., VARANASI, K., LIU, Y., SEIDEL, H.-P., AND THEOBALT, C. 2011. Shading-based dynamic shape refinement from multi-view video under general illumination. In *ICCV*, IEEE, 1108–1115.

ZHANG, R., TSAI, P.-S., CRYER, J. E., AND SHAH, M. 1999. Shape from shading: A survey. *IEEE Trans. PAMI 21*, 8, 690–706.