Dynamic Hair Capture using Spacetime Optimization

Zexiang Xu[1]* Hsiang-Tao Wu† Lvdi Wang† Changxi Zheng‡ Xin Tong† Yue Qi*

*State Key Lab of Virtual Reality Technology and Systems, Beihang University †Microsoft Research ‡Columbia University

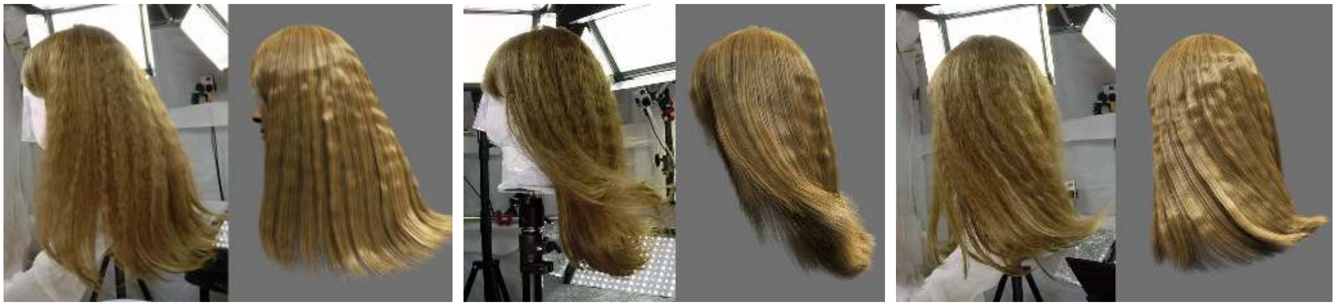


Figure 1: *Our dynamic hair capture system is able to recover different motions of wavy hair driven by wind (middle) and head shaking (right). The captured hair model (right of each pair) faithfully reconstructs the geometry and dynamics of video frames (left of each pair).*

Abstract

Dynamic hair strands have complex structures and experience intricate collisions and occlusion, posing significant challenges for high-quality reconstruction of their motions. We present a comprehensive dynamic hair capture system for reconstructing realistic hair motions from multiple synchronized video sequences. To recover hair strands' temporal correspondence, we propose a motion-path analysis algorithm that can robustly track local hair motions in input videos. To ensure the spatial and temporal coherence of the dynamic capture, we formulate the global hair reconstruction as a spacetime optimization problem solved iteratively. Demonstrated using a range of real-world hairstyles driven by different wind conditions and head motions, our approach is able to reconstruct complex hair dynamics matching closely with video recordings both in terms of geometry and motion details.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Geometric algorithms; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: image-based hair modeling, dynamic capture

1 Introduction

Along with a striking development of image sensing and computer vision techniques, the 3D dynamic capture has received considerable research in computer graphics, ranging from human body poses [de Aguiar et al. 2008; Vlasic et al. 2008], facial expression [Bickel et al. 2007; Bradley et al. 2010; Beeler et al. 2011], and cloth motions [Bradley et al. 2008] to natural phenomena such as liquids [Wang et al. 2009a; Gregson et al. 2012], flames [Ihrke and Magnor 2004], and even plant growth [Li et al. 2013]. 3D reconstruction of dynamic data enables runtime re-targeting and re-rendering, but also paves the way to understand specific dynamics and building various applications, such as the derivation of data-driven models and the validation of numerical simulations.

In this paper, we focus on the capture of 3D dynamic hair, a relatively underexplored dynamic capture problem. As one of the most recognizable parts of a human body, hair strands have diverse styles and exhibit rich spatiotemporal details. While many hair simulation techniques have been devised, these methods still face limitations in terms of fidelity, scalability, parameter selection, and contact resolution. 3D capture, on the other hand, provides a direct way to recover complex hair dynamics, and thus can be complementary to simulation methods.

Dynamic hair capture is challenging. At any time instance, the reconstructed hair needs to faithfully capture the overall hair shape as well as resolve spatial details. Addressing this problem has been a central subject in many recent *static* hair capture methods [Wei et al. 2005; Luo et al. 2013a; Hu et al. 2014]. Although fairly sophisticated, these methods, when applied for dynamic hair capture, lack the ability to establish strands' temporal correspondence across a whole sequence of static reconstructions. Further, the reconstructed hair distribution over a scalp can vary dramatically. Straightforward application of these methods fails to respect temporal coherence, resulting in unpleasant flickering artifacts.

Reconstruction of hair strands' temporal correspondence gives rise to many challenges. Dynamic hair constantly experiences mutual collisions, and some hair strands can move to the surface layer of a hair volume, occluding other stands in a time-varying way. Consequently, no matter how many cameras are placed around, the hair strands that can be tracked through video capture will be incomplete and temporally discontinuous. To make matters worse, the tiny and thin features of hair strands can cause severe motion blur in the captured video. Therefore, traditional feature-tracking algorithms such as optical-flow methods fail to track hair strands robustly. While there exist prior attempts of dynamic hair capture (e.g., [Yamaguchi et al. 2009; Luo et al. 2011; Zhang et al. 2012]), they still suffer from problems such as temporally incoherent hair fibers and topology, and lacking sufficient details.

Our approach to address the strand correspondence is based on a key observation: the motion of individual hair strands projected on a 2D image plane leaves clear trajectories on either horizontal or vertical slices of a spacetime video volume. Such trajectories, which we call *motion paths* (see Figure 5), are resilient to image noises, temporal aliasing and motion blur, and thus can serve as robust indicators of hair correspondence across frames. However, owing to the hair occlusions, motion paths that are recoverable from video sequences are likely to be sparse. To circumvent this problem,

[1] This work was done when Zexiang Xu was an intern at Microsoft.

we exploit a high-quality spatial hair reconstruction at each frame, and develop a spacetime optimization algorithm to infer the missing correspondence in a temporally coherent way.

Our proposed approach has the following major steps. We start by reconstructing static hair geometry at every frame. We build our approach upon the multi-sensor hair capture method [Luo et al. 2013a] to construct a set of *directed hair ribbons* (in § 3) representing a group of local hairs that have consistent shapes and directions. Meanwhile, we extract the *hair motion paths* from individual video sequences (in § 4). We then convert the per-frame hair ribbons into hair strands by solving a spacetime optimization problem (in § 5), which incorporates the extracted motion paths to ensure temporal coherence of hair motion as well as the hair length preservation and hair root distribution to achieve physically plausible hair geometries.

2 Related Work

Over the past decades, numerous methods have been devised to model, simulate, and render human hair. We refer the reader to a survey [Ward et al. 2007] of this prolific field. In this section, we mainly review work most closely related to ours: static and dynamic hair capture methods.

Static hair capture. Static hair reconstruction from multi-view images is pioneered by a series of work including [Kong et al. 1997; Grabli et al. 2002; Paris et al. 2004; Wei et al. 2005; Paris et al. 2008]. Several key ideas introduced therein, such as the estimation of a dense 3D vector field from 2D images and tracing hair strands through a vector field, are widely used in recent hair reconstruction work. Jakob et al. [2009] reconstructed detailed fiber-level hair geometry by sweeping the focal plane of a macro lens at different distances. This method generally takes a rather long time and is impractical for dynamic capture. Bonneel et al. [2009] estimated hair geometry and appearance from a single photograph. But their method aims for statistically plausible approximation rather than faithful reconstruction. Beeler et al. [2012] introduced a system to capture sparse facial hair and facial skin simultaneously using multi-view stereo matching. Using thermal imaging techniques, Herrera et al. [2012] sidestepped several common issues, such as anisotropic hair reflectance, hair-skin segmentation, and achieved high-quality reconstruction. But this method relies on the thermal emission from the head scalp, and is not suitable for long hair strands that grow far away from the head. Chai et al. [2012] proposed a single-view hair modeling algorithm that can generate a partial hair model suitable for image manipulation. Our dynamical hair capture is based on the static hair capture method [Luo et al. 2013a], which groups hair segments into ribbons, and then converts the ribbons into structured hair wisps. The reconstructed hair structures are further improved by a recent method [Hu et al. 2014] that exploits physical simulation methods to build a database of strand examples.

All these methods focus on faithful reconstruction of static hair geometry. None of them were explicitly designed to ensure temporally coherent hair reconstruction. When provided with a sequence of input images, these methods can perform the reconstruction at every single time instance, but may lead to hair jumping or flickering artifacts.

Dynamic hair capture. Dynamic hair capture from videos is much less explored. A central challenge is to ensure temporal coherence while retaining the hair details. Ishikawa et al. [2007] attached small reflective markers to a few hair strands and used a motion capture system to reconstruct the motions of these "guide strands", followed by an interpolation scheme to generate a full dynamic hair model. They were limited by very coarse hair motions, as the details were smeared out by the interpolation of guide hairs. Yamaguchi et al. [2009] extended an earlier work [Wei et al. 2005]. The method, however, only works for straight hair with limited motion. Luo et al. [2011] computed multi-view correspondence from 2D orientation maps for dynamic hair capture, but still has artifacts of temporal incoherence in the results. Zhang et al. [2012] proposed a simulation-guided approach for dynamic hair capture. Simulated hair motions are used to ensure temporal coherence, but this method is limited by an over-smoothed hair geometry. Lastly, Chai et al. [2013] demonstrated a dynamic hair modeling technique based on single-view input, but their combination of optical flow with sparse features for tracking hair motion is limited by simple hairstyles and motions. Taking into account cross-view consistency of hair orientations, Luo et al. [2013b] proposed an algorithm to reconstruct a 3D hair surface of high accuracy. While their method also works on dynamic input, the output result is a surface mesh rather than a set of hair strands.

Different from all those methods, our approach exploits motion paths to ensure temporal coherence. Our extraction of motion paths is robust against image noises, temporal aliasing, and motion blur. We further incorporate the motion paths into an optimization framework to produce temporally coherent and detailed reconstruction.

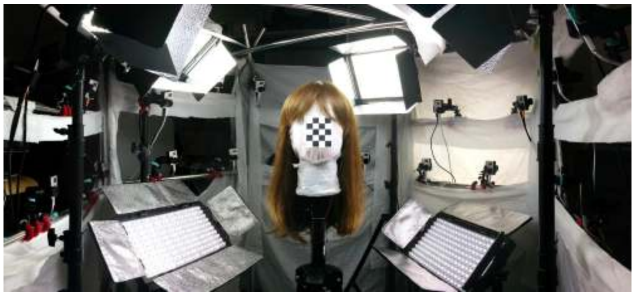


Figure 2: Our dynamic hair acquisition setup *consists of 21 GoPro cameras and six LED arrays.*

3 Data Capture and Preprocessing

Our system takes as input video frames of moving hair captured from multiple view angles. In the preprocessing step, the frames at every time instance form a multi-view stereo system. We preprocess video captured from every view individually, and reconstruct the hair geometric structures *statically*. In particular, we reconstruct at every time instance a set of *directed hair ribbons* following Luo et al.'s recent work [2013a]. These hair ribbons, although lacking temporal coherence, represent the hair's large-scale shapes and directions of growth, and will be used later as an initialization in our optimization framework (in § 5).

Acquisition setup. We capture moving hair using 21 GoPro HERO 3+ cameras mounted around the subject. They are able to record video in 720p at 120 fps, and cover an angular range of about 270° in longitude and 45° in latitude. We calibrate all the spatially mounted cameras using the Multicore Bundle Adjustment package [Wu et al. 2011], and synchronize their captures with a flashbulb. To reduce sharp shadows, we light the hair with six LED arrays (see Figure 2 for our lab setup).

Background removal. For each captured image frame, we remove non-hair background region automatically using a color classifier trained with a Gaussian Mixture Model [Chen et al. 2012]. For

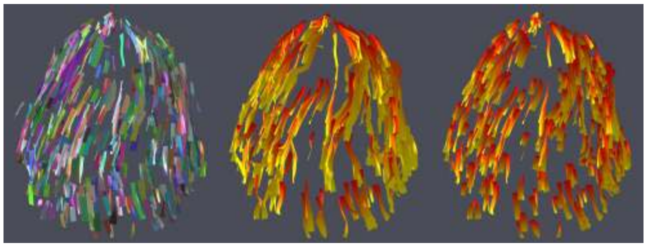


Figure 3: Hair ribbons and their directions: *We reconstruct hair ribbons of an input frame (left), connect them into hair wisps (middle), and build directed hair ribbons (right). The ribbon direction is color-coded using a color gradient from red to yellow.*

all datasets used in our examples, we manually paint $1 \sim 3$ binary masks, which are sufficient to learn a classifier of foreground and background pixels.

Head, scalp, and hair roots. We track the head's rigid motion by attaching a chessboard on the face of a dummy head (see Figure 2). We detect the chessboard using functionalities provided in OpenCV [Bradski 2000] and estimate the head's rigid motion at every time instance automatically. The estimated rigid transformations are used to transform a 3D head model initially provided and aligned by the user. We hypothesize that the hair roots are uniformly distributed over the entire scalp region of the head model. And the number of hair roots is a user-specified parameter, typically ranging from 30k to 50k in our examples.

Directed hair ribbons. We generate hair ribbons using the approach introduced in [Luo et al. 2013a]. Here we briefly review this method, while referring the reader to the original paper for more details. We start by generating a point cloud of the hair's outer surface using the Patch-based Multi-View Stereo (PMVS) algorithm [Furukawa and Ponce 2010]. Then we construct hair segments starting from each reconstructed 3D point. The segments are generated based on a 3D orientation field on the point cloud surface. We compute a 2D dense orientation field for every single view using a method similar to [Paris et al. 2004; Chai et al. 2012], and then combine them into a 3D orientation field [Wei et al. 2005]. Next, we group the segments into ribbons based on their spatial similarity in local regions. And lastly, we build a connection graph to represent possible ways of connecting hair ribbons. This graph is used to solve a graph-cut optimization problem for finalizing the ribbon connection into hair wisps. Different from [Luo et al. 2013a], we stop the pipeline at this step without moving on to the hair strand generation. Instead, we choose the ribbons that are connected into final wisps as our *directed hair ribbons*, as they have compatible growth directions (see Figure 3). The directed hair ribbons will be used as an initialization in § 5.

Remark. The rationale of keeping only directed hair ribbons is that they can serve as representatives of local geometric structures of hair at every time instance. In other words, they are "intrinsic" information that we can recover statically. Provided with this information, existing methods such as [Luo et al. 2013a] rely on heuristics to interpolate hair geometries that are invisible or missing. While hair interpolation is almost unavoidable because of hair occlusions, in the case of dynamic capture, it is prudent to design the interpolation scheme in a temporally coherent way. The static capture methods, however, were not designed to respect temporal coherence. Therefore, we postpone the computation of interpolated geometries such as invisible hair strands and hair attachment to the scalp. We will exploit motion paths introduced in the next section to bridge the detected local geometric structures with temporally coherent dynamics.

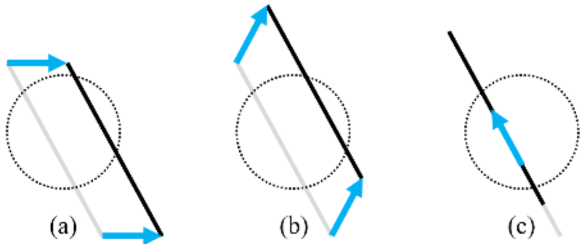


Figure 4: The ambiguity of hair motion: *A hair segment (black) moves in three different directions. Its movements are perceived in a local window (dashed circle). The two observed movements in (a) and (b) are indistinguishable. When the hair segment moves along its tangential direction (c), its movement can not even be observed.*

4 Motion Path Analysis

A key step of dynamic hair capture is the extraction of temporal information from input video sequences for tracking individual hair strands. Unfortunately, traditional optical-flow-based methods fail to track hair strands because of their thin structures. In fact, it is even practically impossible to track the exact motion of most hair segments, since the motion captured on an image can be ambiguous, a phenomenon known as "Barberpole illusion" (see Figure 4).

Inspired by the Epipolar-Plane Images (EPIs) of light fields [Bolles et al. 1987; Kim et al. 2013], we observe that the motion of hair strands often exhibits clear trajectories in the horizontal or vertical slices of a spatiotemporal video volume, as shown in Figure 5. This observation underpins our basic idea of hair strand tracking: we extract 2D hair trajectories robustly from each slice of the video volume. These trajectories, which we call *motion paths*, will then be used to track shape changes of 3D hair strands.

In this section, we first lay out in § 4.1 our algorithm of extracting motion paths represented as a set of polylines. While the motion paths are in 2D, when used in an optimization framework, they suffice to estimate the 3D shape changes of a hair strand, as elaborated in § 4.2.

4.1 Extraction of Motion Paths

Without loss of generality, consider a *horizontal* slice, which we call motion-path slice, at row y of a spatiotemporal video volume. For each pixel (x, t) on this slice, we wish to find (i) a *confidence* value indicating how well we can track hair's motion path at that pixel, and (ii) a *direction* pointing to a pixel $(x', t + 1)$ that are on the same motion path in the next time frame (see Figure 5(d)).

A hybrid direction detector. One way to estimate the moving direction at a pixel and the associated confidence value is using a Gabor filter [Paris et al. 2004; Jakob et al. 2009]: we convolve the image with a set of Gabor filters oriented at different angles, and for each pixel select the angle that maximizes the filter response as the detected pixel direction. The confidence value is computed based on the amplitude and variation of filter responses. This method is robust to image noise, and the resulting confidence map also corresponds to the ridge line features as demonstrated in [Chai et al. 2012]. This property helps us robustly re-center motion paths and avoid drifting when tracing the paths in temporal domain. However, simply applying oriented Gabor filters in our method is problematic. When a hair strand moves quickly (e.g., faster than the video frame

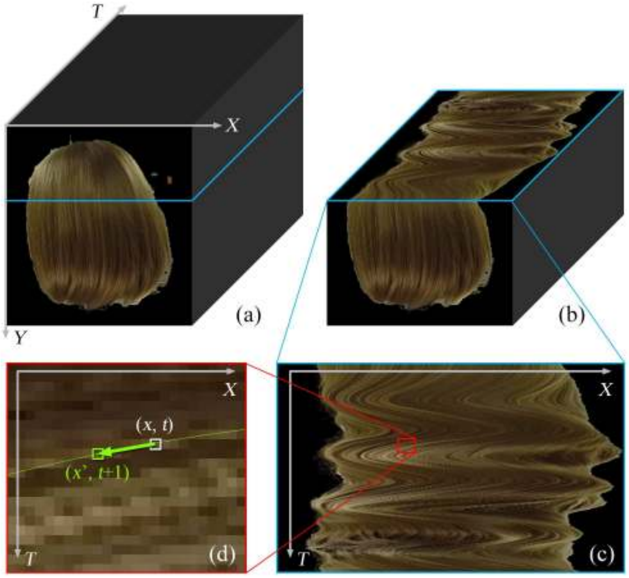


Figure 5: Motion paths: *In a spatiotemporal video volume (a), the motion of hair strands exhibits clear trajectories in a horizontal slice (b)(c). In (d), one of such trajectories, which we call a* motion path*, is highlighted in green.*

rate), its trajectories on a motion-path slice may become discontinuous. The resulting temporal aliasing will be picked up by Gabor filters as erroneous moving directions (see Figure 6(b)).

Another approach is to directly find in row $t+1$ a pixel $(x', t+1)$ whose 1D neighborhood within a radius r best matches that of (x,t). In particular, let $I_y(x,t)$ denote the pixel color at (x,t) of the motion path slice, we compute

$$x' = \arg\min_{x'} \sum_{j=-r}^{r} \left\| I_y(x+r,t) - I_y(x'+r,t+1) \right\|^2. \quad (1)$$

In this paper, we call this method *1D block matching*. Unlike Gabor-filter-based methods, this method is robust against temporal aliasing, but more prone to image noise. In addition, it is also hard to estimate a confidence map that corresponds to the ridges (see Figure 6(c)).

We propose a hybrid approach to leverage the advantages of both methods. Given a motion-path slice, we first estimate a direction field using the 1D block matching method. Let $\boldsymbol{d}_i$ denote the estimated motion-path direction at a pixel i. We then convolve at each pixel i with a Gabor kernel rotated exactly to the direction $\boldsymbol{d}_i$. The Gabor kernel we use has a form,

$$K_\theta(u,v) = \exp\left(-\frac{1}{2}\left[\frac{\tilde{u}^2}{\sigma_u^2} + \frac{\tilde{v}^2}{\sigma_v^2}\right]\right)\cos\left(\frac{2\pi\tilde{u}}{\lambda}\right), \quad (2)$$

where $\tilde{u} = u\cos\theta + v\sin\theta$ and $\tilde{v} = -u\sin\theta + v\cos\theta$ for a rotation angle θ. For all the examples, we use $\sigma_u = \sigma_v = 2, \lambda = 4$. The resulting filter responses are normalized over the entire image as the confidence map. Lastly, we apply one pass of the orientation refinement procedure as described in [Chai et al. 2012] to remove outliers. This hybrid method is robust against *both* temporal aliasing and image noise, as shown in Figure 6(d).

Tracing motion paths. Once we have obtained a direction and confidence field on every motion-path slice, we start to trace the motion paths along the estimated motion-path direction field. The confidence value at a pixel measures the quality of a local motion-path direction; a low confidence indicates that the direction might be contaminated by noises. Therefore, we are only interested in motion paths along which the confidence values are high.

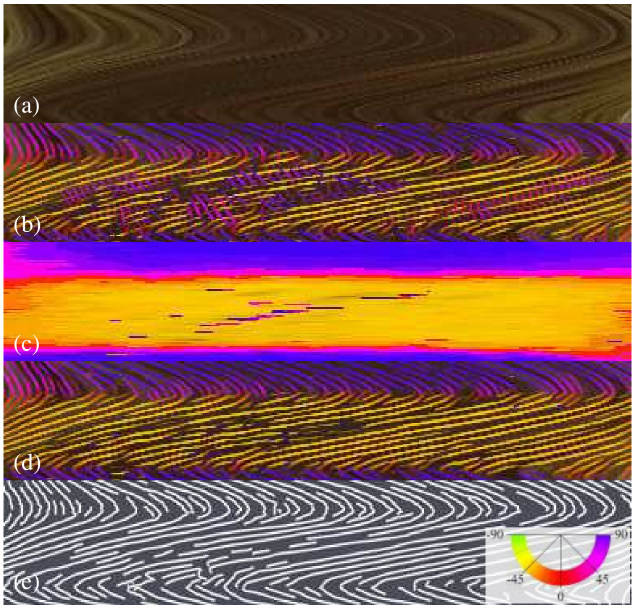


Figure 6: Motion path extraction: *(a) Given a motion path slice with temporal aliasing, (b) Gabor filters may detect incorrect directions due to aliasing. (c) An 1D block matching approach is robust against aliasing, but prone to image noise. (d) Our hybrid approach combines the advantages of (b) and (c), and produces a reliable sparse direction field from which (e) a set of polyline paths can be traced. The pseudo-color in (b-d) represents the detected direction as in the inset.*

In each motion-path slice, we sweep through every row t (corresponding to a frame t in time). For each untraced pixel (x,t) with a confidence value higher than a threshold σ, we start tracing a new motion path from that pixel. The tracing process repeatedly finds a new pixel $(x', t+1)$ along the motion-path direction at the current location (x,t) until it reaches a pixel whose confidence is below the threshold σ. For all our examples in this paper, we set $\sigma = 0.3$. At each tracing step, we also adopt a re-centering scheme similar to that in [Chai et al. 2012] to correct the new position to the local maximum of confidence. The pixels along a single trace form a polyline on a motion-path slice (see Figure 6(e)). At the end of this step, we obtain a set of polylines starting at different video frames in time. These polylines serve for two purposes in our system: (i) we use them to quickly check if a given pixel (x,y) at time frame t is on any motion-path; this kind of query will be frequently used when we optimize the hair geometry in § 5. And (ii) we use them to warp a hair strand from one time frame to adjacent time frames, as detailed in the next subsection.

4.2 Hair Prediction using Motion Paths

With a set of motion paths ready, we now present our algorithm to warp a hair strand from its shape in time frame t to an adjacent time frame t'. The algorithm introduced here will be used thereafter in § 5. We will refer back to this subsection later when presenting the hair geometry optimization algorithm.

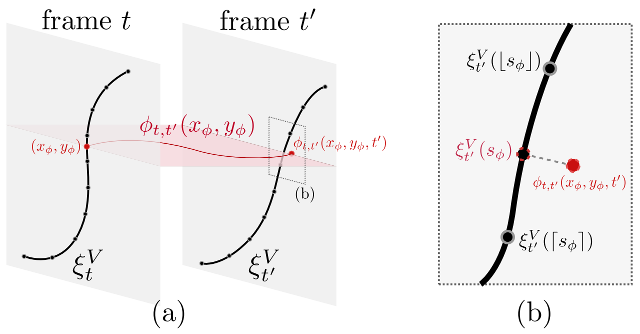

frame t
frame t'
(x_ϕ, y_ϕ)
$\phi_{t,t'}(x_\phi, y_\phi)$
$\phi_{t,t'}(x_\phi, y_\phi, t')$
(b)
ξ_t^V
$\xi_{t'}^V$
(a)
$\xi_{t'}^V(\lfloor s_\phi \rfloor)$
$\xi_{t'}^V(s_\phi)$
$\phi_{t,t'}(x_\phi, y_\phi, t')$
$\xi_{t'}^V(\lceil s_\phi \rceil)$
(b)

Figure 7: Motion path prediction in a single view: *(a) A motion path (red curve) intersecting with projected strand ξ_t^V at (x_ϕ, y_ϕ) traces to position $\phi_{t,t'}(x_\phi, y_\phi, t')$ at frame t'. (b) In a zoomed-in view, we predict the shape of the projected strand $\xi_{t'}^V$ by minimizing the distance between $\phi_{t,t'}(x_\phi, y_\phi, t')$ and its closest strand point $\xi_{t'}^V(s_\phi)$. The latter can be represented as a linear combination of the two adjacent projected strand particles.*

Notations. We denote the shape of a hair strand at a time frame t as ξ_t, and let ξ_t^V be the 2D projection of ξ_t on a camera view V. A hair strand is represented as a polyline consisting of N particles. We use $\xi_t(i), i = 1...N$ to denote the position of the i-th particle on the strand and use $\xi_t^V(i)$ to indicate the corresponding projected position on the view V. Additionally, we use $\phi_{t,t'}(x, y)$ to denote a motion path between frame t and t' that passes through pixel (x, y) at frame t. Similarly, $\phi_{t,t'}(x, y, t_c)$ denotes the pixel position this motion path passes through at frame t_c, so that $\phi_{t,t'}(x, y, t) = (x, y)$.

Single-view prediction. Given a 2D projection of strand ξ in view V and frame t (i.e., ξ_t^V) and the extracted motion paths, our goal here is to predict the hair strand shape at time frame t' by estimating the projected positions of hair particles $\xi_{t'}^V(i), i = 1...N$. We first collect a set Φ^V of motion paths that pass through the hair projection ξ_t^V from time frame t to t', namely,

$$\Phi^V = \left\{ \phi_{t,t'}(x, y) \mid (x, y) \in \xi_t^V \right\}.$$

In other words, the hair strand deforms along all the paths in Φ^V, so the projected shape of ξ in frame t' should pass through $\phi_{t,t'}(x, y, t')$ for all the paths $\phi_{t,t'} \in \Phi^V$. Because of the possibility of "Barberpole illusion", we can not simply interpret a motion path $\phi_{t,t'}$ as a point on ξ_t^V moving *horizontally* from (x, y) to $\phi_{t,t'}(x, y, t')$. To estimate the hair strand shape $\xi_{t'}^V$ at time frame t', we propose to solve an optimization problem in which the point-to-curve distance between $\phi_{t,t'}(x, y, t')$ and $\xi_{t'}^V$ is minimized for all $\phi_{t,t'} \in \Phi^V$ in a least-squares sense (see Figure 7).

Concretely, consider a motion path $\phi \in \Phi^V$ that intersects with the projected hair strand ξ_t^V. Let (x_ϕ, y_ϕ) denote the intersection point. Following along the motion path $\phi_{t,t'}$, its predicted position at t' is $\phi_{t,t'}(x_\phi, y_\phi, t')$. Denoting the unknown projected hair shape at t' as $\xi_{t'}^V$, we define an energy function

$$E_\Phi = \sum_{\phi \in \Phi^V} \left\| \xi_{t'}^V(s_\phi) - \phi_{t,t'}(x_\phi, y_\phi, t') \right\|^2, \qquad (3)$$

where $\xi_{t'}^V(s_\phi)$ denotes the *closest point* on $\xi_{t'}^V$ from the motion path-predicted point $\phi_{t,t'}(x_\phi, y_\phi, t')$. If s_ϕ is a real-valued parameterization of the polyline $\xi_{t'}^V$, we can rewrite the position of $\xi_{t'}^V(s_\phi)$ as a piece-wise linear interpolation of the hair particles, namely,

$$\xi_{t'}^V(s_\phi) = (\lceil s_\phi \rceil - s_\phi)\xi_{t'}^V(\lfloor s_\phi \rfloor) + (s_\phi - \lfloor s_\phi \rfloor)\xi_{t'}^V(\lceil s_\phi \rceil),$$

where $\lceil s_\phi \rceil$ and $\lfloor s_\phi \rfloor$ are respectively the subsequent and preceding hair particles of hair strand ξ_t^V at the parameter s_ϕ.

In addition, since the motion paths are typically sparse in comparison to strand vertices, we introduce two regularization terms. The first one is to preserve strand length,

$$E_{\text{len}} = \sum_i^N \left\| \xi_{t'}^V(i+1) - \xi_{t'}^V(i) \right\|^2,$$

and the second is to preserve local shape details represented by Laplacian coordinates [Sorkine et al. 2004]:

$$E_\Delta = \sum_i \left\| 2\xi_{t'}^V(i) - \xi_{t'}^V(i-1) - \xi_{t'}^V(i+1) \right\|^2.$$

Finally, the total energy is defined as

$$E = \alpha_\Phi E_\Phi + \alpha_{\text{len}} E_{\text{len}} + \alpha_\Delta E_\Delta. \qquad (4)$$

We set $\alpha_\Phi = 300$ and $\alpha_{\text{len}} = \alpha_\Delta = 1$ in all our examples.

Now we need to solve for the unknown hair particle positions $\xi_{t'}^V(i), i = 1..N$ by minimizing the energy function (4). The evaluation of (4) involves computing the parameterization s_ϕ of closet points, as used in (3). If we know s_ϕ *a priori*, minimizing the energy function (4) amounts to solving a sparse linear system of projected vertex positions of $\xi_{t'}^V$ (i.e., $\xi_{t'}^V(i), i = 1..N$). We therefore solve the minimization problem of (4) in an iterative way similar to the Iterative Closest Point (ICP) algorithm: in each iteration, we first find for each $\phi_{t,t'}(x, y, t')$ the closest point on the current estimation of $\xi_{t'}^V$, and then solve the resulting linear system (using CHOLMOD [Davis and Hager 2009; Chen et al. 2009]) to compute updated vertex positions of $\xi_{t'}^V$.

Multi-view prediction. When we have input from multiple views in which ξ_t is visible, the 2D projection of $\xi_{t'}$ in each view needs to match as closely as possible the hair strand prediction by that view's motion paths. These 2D motion path predictions collectively constrain the 3D shape of $\xi_{t'}$. The formulation of the optimization is similar to the single-view case. The only difference is that the penalty term (3) for the mismatch of motion paths now includes all views where ξ_t is visible.

5 Spacetime Hair Optimization

So far we have reconstructed a sequence of directed static hair ribbons at all time instances (in § 3) and extracted motion paths in motion-path slices of every view's spatiotemporal video volume (in § 4). In this section, we combine them together to generate a final animation of hair strands in a spatiotemporally coherent way.

We start by estimating a per-frame 3D direction field from the directed hair ribbons (§ 5.1). Since the ambiguity of hair's direction of growth has been resolved by direction analysis in our data preprocessing step [Luo et al. 2013a], we can trace physically plausible hair strands using the direction field (§ 5.2). However, these initial strands resemble independent geometric shapes of hair in each time frame, neglecting their temporal coherence. Consequently, the resulting hair strands have flickering artifacts. To address this problem, we propose an iterative optimization method that effectively re-matches each strand's external part with its root, producing a final dynamic hair model that is both temporally smooth and detail-preserving (§ 5.3).

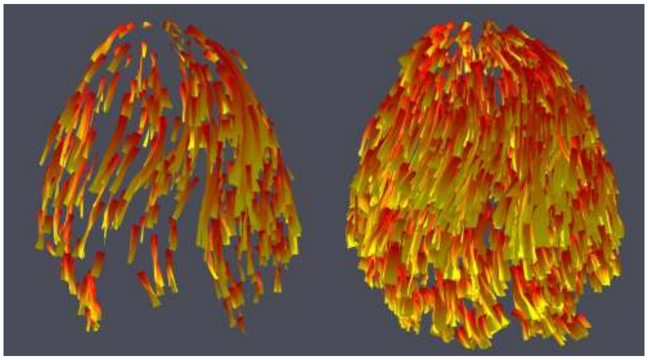


Figure 8: Ribbon warping: *(left) Directed hair ribbons reconstructed at a single frame are typically sparse. (right) By adding the warped ribbons from adjacent (± 2 in this case) frames, the resulting ribbons are much denser. They retain hair details and temporally coherence.*

5.1 Per-frame 3D Direction Fields

The directed hair ribbons produced in the preprocessing step (§ 3) represents the most salient static hair structures at each frame, but they are likely to be sparse (see Figure 8). Our first attempt of inferring the missing data was computing a dense 3D direction field using the sparse known directions as constraints, followed by tracing hair strands using the resulting direction field. This approach produces smooth vector field, mainly because the directions and shapes represented by directed hair ribbons are consistent throughout the entire hair volume. However, since the directed hair ribbons are sparse, the resulting direction field tends to be overly smoothed, losing hair structure details. Moreover, we observed that the sparsity of hair ribbons can change dramatically from frame to frame, leading to temporally incoherent direction field. On the other hand, it suggests that we can fill in missing ribbons in one frame by warping ribbons from adjacent frames using motion paths.

Ribbon warping using motion paths. In each time frame t, we warp ribbons from an adjacent time frame $t + \Delta$ toward the frame t and add them into the set of hair ribbons at frame t (see Figure 8). According to [Luo et al. 2013a], every hair ribbon resembles a set of isocurves. Warping a ribbon from frame $t + \Delta$ to t amounts to a multi-view motion-path prediction (introduced in § 4.2) of each isocurve. In practice, we use $\Delta \in \{-2, -1, 1, 2\}$, and only warp a ribbon if all its isocurves have at least 5 intersecting motion paths. Recall that in § 4.1 we extract motion paths as polylines. We check the intersection between a polyline and an isocurve by comparing the closest distance between them against a small threshold value of 2 pixels.

Hair root directions. Hair roots are almost always invisible, yet their direction can significantly affect overall hair shape as well as the reproduction of hair features such as parting lines or vortices. Our system allows the user to draw directly a hair parting line on the scalp. For each hair root, we then estimate an *incline direction* using the gradient of a geodesic distance field defined on the parameterized scalp surface [Wang et al. 2009b]. The hair root direction is set by tilting the scalp surface normal to the incline direction by a user-adjustable amount. We use a 90 degree incline angle for all our examples.

Direction field diffusion. Lastly, we generate a complete direction field at a time frame by diffusing the directions defined on all directed ribbons, including both the original and the warped ones. The direction filed is solved in a bounding box of hair volume. We discretize the bounding box into voxels, among which the ones containing either part of a ribbon or a hair root are considered as *known voxels* whose directions are the ribbon's direction or root direction. We use the known voxels as Dirichlet boundary condition to solve a Laplace equation. This diffusion procedure is similar to [Fu et al. 2007], and the voxel size used in our examples is 3 mm.

5.2 Generating initial strands

Interior distance field. To confine hair strands to captured hair volume, we also compute a scalar distance field, denoted as D, similar to the one used by Luo et al. [2013a]. We reuse the 3D grid used for the direction field diffusion to solve and store D there. A voxel that is occupied by a directed hair ribbon or a head model is considered as a *hair boundary voxel*, in which we store a zero distance-field value. The distance values of the boundary voxels of the bounding box volume is set to be the smallest Euclidean distance between that voxel and any hair boundary voxel. The values of all other voxels are computed by solving a diffusion equation.

Strand growing. Given a time frame t, after computing the 3D direction field $\mathbf{V}$ and the interior distance field D, we now generate an initial hair model by growing individual strands from the hair roots following the 3D direction $\mathbf{V}$,

$$\xi(i) = \xi(i-1) + \epsilon \mathbf{V}(\xi(i-1)), \tag{5}$$

The growing of a hair strand terminates when any of the following conditions is met: (i) the strand has grown out of the hair volume where $D(\xi(i)) > d_{\text{bound}}$; (ii) the growing direction has a sharp change $\mathbf{V}(\xi(i)) \cdot \mathbf{V}(\xi(i-1)) < 0$; (iii) the strand has reached a predefined maximum length. We set d_{bound} to be 10 mm in practice.

Strand length correction. Following the process above, we can generate an initial dynamic hair model $\mathcal{S}_0$ from all the input frames. Each dynamic strand in $\mathcal{S}_0$ has a fixed root, but its length may vary significantly through the sequence. This is because it is very difficult to detect hair *tips* from input video frames. Previous image-based methods tend to generate *as-long-as-possible* hair strands, causing hair tips to concentrate unnaturally in a narrow region (usually near the bottom volume boundary in the case of a draping hairstyle).

To remedy this issue, for each dynamic strand ξ, we calculate its *average length* through the entire sequence as a preferred length, and trim ξ so that its length in any frame is no longer than the preferred length. Figure 9 compares the initial hair strands before and after length correction. This correction step greatly reduces strand length variation while keeping the overall hair shape nearly unchanged.

5.3 Optimizing Final Strands

Being traced directly from the 3D direction field, the initial dynamic hair model $\mathcal{S}_0$ in general matches the large-scale geometry of the original hair quite well. However, the connections between exterior strand segments and hair roots are temporally inconsistent and prone to small differences from the input. As a result, a strand may largely jitter even though its root is fixed on the scalp, and the entire hair model will appear flickering.

EM-like optimization. We therefore seek hair strand motion that can both robustly grow to the exterior and match the spatial geometric details in each frame. This problem is similar to the case in Expectation-Maximization (EM) [McLachlan and Krishnan 1997], where the *variables* (the strand shape) and the *parameters* of the

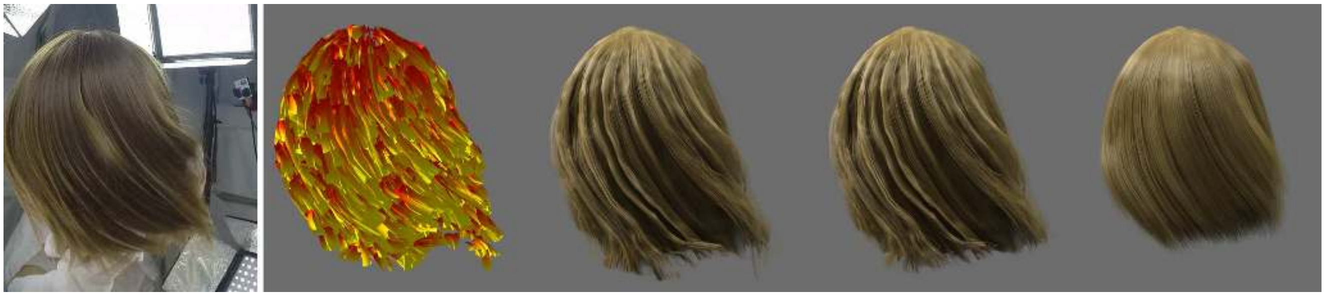


Figure 9: *From left to right: input video frame, directed hair ribbons, initial strands without length preserving, initial strands with length preserving, and the final optimization result.*

energy function (the spatial constraints this strand should satisfy) are both unknown. We thus introduce an EM-like algorithm for optimizing the final shape of each strand. It alternates between the E and the M steps:

- In the *E-step*, for each strand ξ_t in frame t, we use multi-view motion path prediction (§ 4.2) to warp the corresponding strand $\xi_{t+\Delta}$ from an adjacent frame $t+\Delta$ to the current frame, with an additional hard constraint that the root of the warped strand must coincide with $\xi_t(0)$.

  Let $\xi_{t+\Delta,t}$ denote a strand warped from frame $t+\Delta$ to t, for each ξ_t we calculate a weighted average of all warped strands:

$$\xi_t^* = \sum_{\Delta=-3}^{3} w_\Delta \xi_{t+\Delta,t}, \tag{6}$$

  where w_Δ is normalized Gaussian weight, and $\xi_{t,t} = \xi_t$. The warped strands are uniformly re-sampled before averaging.

- In the *M-step*, we align each average strand ξ_t^* computed in the E-step with the local spatial constraints. We select the aligned strand ξ_t' by iteratively minimizing the following energy function:

$$E_{\mathrm{v}} = \sum_i \Big(\left\| \xi_t'(i+1) - \xi_t'(i) - \epsilon \mathbf{V}_t(\xi_t(i)) \right\|^2 + \left\| \xi_t'(i) - \xi_t^*(i) \right\|^2 \Big), \tag{7}$$

  where $\mathbf{V}_t$ is the 3D direction field in frame t, and $\epsilon = 3$ mm the step length for initial hair tracing.

  This process deforms ξ_t^* locally so that it better matches the geometric details encoded in the direction field and in the mean time preserve the temporally consistent large-scale shape, effectively performing a *re-connection* between exterior hair segments and hair roots on the scalp.

In practice we find that 10∼25 EM iterations are sufficient to converge to a good result in most cases.

6 Results

We have captured three different hairstyles including a short straight one, a long wavy one, and a long straight one, each with several different types of motions. As shown in Figure 13, our results can faithfully reproduce the prominent shapes of the original hair strands even in the case where a group of hair wisps move in front of other hairs. When blown by wind (Figure 1 and 13), the stochastic strand motion can also be reconstructed by our method. For the long wavy hairstyle, the reconstructed hair model preserves the original characteristic waves while maintaining temporal consistency. We refer the readers to the supplemental video that includes all the dynamic results.

Algorithm Step	*Time cost (hours)*
Image background removal	0.2
PMVS point cloud	4
Directed hair ribbons	1
Motion path analysis	0.5
Align ribbons	4
Initial strands	2
Final strands	3
Total	14.7

Table 1: Timings of our algorithm *on a typical 1200-frame, 21-view dataset, measured on a shared cluster with about 200 effective processing cores.*

In Figure 10, hair strands dynamically interact with a ball, mimicking the cases where real hair collides with obstacles such as one's shoulders. The reconstructed hair strands move around the virtual ball in a convincing way, even though some internal strands may penetrate into the ball, as our current algorithm does not resolve collisions explicitly.

Our input video sequences have a length of 3∼10 seconds (120 frames per second), and the reconstructed hair models in our examples consist of 30K∼50K strands. Note that the hair's level of details largely depends on the input video quality and the resolution of reconstructed vector field. In our case, 30K∼50K strands are sufficient for capturing the geometric details while keeping a low storage and I/O cost. For rendering purpose, however, a denser set of hair can be interpolated.

6.1 Implementation details

Our algorithm pipeline is mostly automatic. Given a set of multi-view video captures, the only needed manual input is to paint 1-3 hair masks for training the foreground/background classifier, and aligning the head model in a single frame. It takes about 15 min for manually setting a dataset.

We process the captured data on a shared computer cluster whose available cores vary from several dozens to a few hundreds. For a typical 1200-frame dataset, the average processing time is about 15 hours. Table 1 shows the timings for individual algorithm steps. The input image frames and the output of every steps take about 500GB storage in total for each dataset.

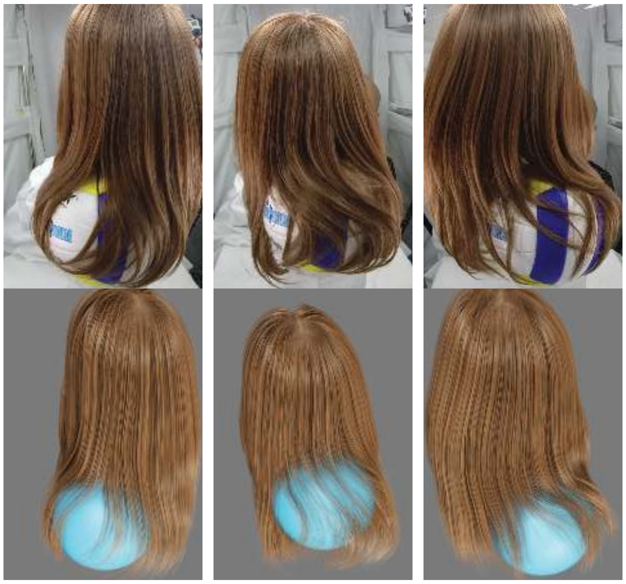


Figure 10: Hair interacting with a ball: *Three input frames are shown in the top row, and the corresponding reconstructed hair models are in the bottom row.*

6.2 Discussion

Comparison with [Zhang et al. 2012]. Our captured hairstyles are similar to those in [Zhang et al. 2012]. However, the results from both approaches have significant differences. Spatially, our results capture the hair's global shape. In comparison with [Zhang et al. 2012], our results preserve more local details. Notably, we resolve the non-smooth wisp structures and the gaps in-between (e.g., the valley-shaped hair region in the fourth group of Figure 13 - Dataset #1). Temporally, Zhang et al.'s results only demonstrate with gentle and relatively slow hair motions, as shown in their video. In contrast, our method can handle rapid head shaking as well as wind blows under which nearby strands may move quite differently. This robustness is due to our robust motion path tracking and the space-time optimization that takes both spatial and temporal constraints into account in each iteration.

Motion path slice direction. In all our examples, we use *horizontal* motion-path slices for motion-path extraction, because the hair strands are more likely to intersect with nearly perpendicular slices. If the 2D projection of a strand intersects with slices nearly in parallel, the resulting motion paths may become ambiguous and thus difficult to detect, as shown in Figure 11. We note that neither our motion path extraction nor the motion prediction algorithm depends on a specific slice direction. When necessary, one can choose different slice directions according to the strands' local orientations.

Limitations. The types of dynamic hair that our method can robustly reconstruct are limited by the geometric and motion inconsistency of the hair subject. In our pipeline, a number of steps such as point cloud reconstruction, 2D orientation fields, directed ribbons, and 3D direction field all assumes that the hair geometry have locally limited shape variance. While this assumption is valid for a wide range of real-world hairstyles (e.g., the wavy one shown in Figure 1 and 13), certain hairstyles under large-scale head motions or external forces such as in a strong wind can result in severely large local variance of hair geometry. For those cases, our method may produce overly smoothed results. One failure case in our experiments is shown in Figure 12, where the hair strands are blown by a strong electronic fan, producing a rather chaotic hair motion. In this case, the PMVS fails to reconstruct a reliable point cloud, and the directed ribbons are sparse even after ribbon augmentation. What makes it worse, resolving the ambiguity of hair's direction of growth in a temporally coherent way also becomes quite challenging. While our method still produces a complete and temporally smooth dynamic model, the rich motion details in the input are largely missing in the reconstruction.

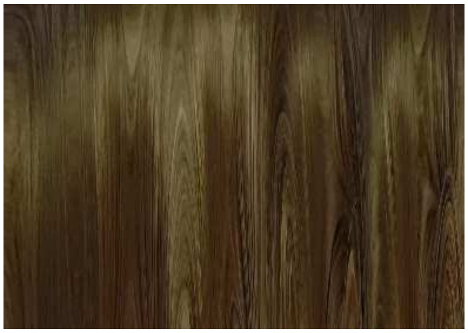


Figure 11: *A vertical motion path slice of the video volume in Figure 5(a).*

Another limitation of our method is that the silhouette (visual hull) of the reconstructed hair may not closely match the input, especially near the tips of draping hairs. By using a more sophisticated video matting algorithm, we might be able to cut hair tips near the silhouette based on soft hair masks, and thereby further improve the direction estimation and strand tracing. Another alternative is to replace the PMVS-based point cloud reconstruction by [Luo et al. 2013b], which can generate a high-quality hair surface using a wide-baseline setup similar to ours.

Real hair often has a small portion of stray strands. Their shapes and motion can differ from their nearby hair strands noticeably. This kind of spatially local variance poses difficulties for many existing reconstruction methods including ours. Missing those spatial details can degrade the visual realism of final results. One possible approach to circumvent this problem is to add some procedurally generated or physically simulated stray hairs to the reconstruction results in a post-processing step.

When hair strands interact with an object (see Figure 10), our current algorithm reconstructs hair strands that might intersect with the object mainly because of two reasons. First, when the object is partially occluded by hair, the object surface behind the hair becomes invisible from cameras. Thus it is difficult to recover that part of 3D shape of the object unless some prior knowledge about the object's geometry is assumed. Second, even with a known object's geometry, our current framework does not resolve the collisions explicitly. However, it is possible to correct the hair strand positions in a postprocessing step to resolve collisions using methods such as position-based dynamics [Müller et al. 2007; Cai et al. 2014]. In this paper, we leave the dynamic capture of both objects and hair strands as a future work, and focus on the capture of hair strands only.

Capturing the hair of real human subjects requires a robust head tracking algorithm (e.g., [Bradley et al. 2010; Beeler et al. 2011]). Our algorithm has no assumption about the material or physical properties of hair fibers, so it should be able to reconstruct most hair strands of a human subject. But similar to the case involving hair-object interactions, one needs to take extra care when processing

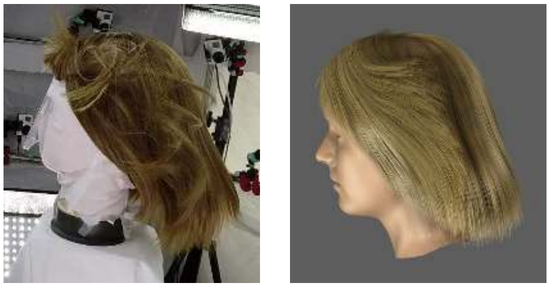


Figure 12: Failure Case: *When the hair motion is highly chaotic, our method will produce over-smoothed result.*

hair strands near the boundaries. Lastly, integrating our hair capture technique into a full-body human dynamic capture system with other specialized techniques is an interesting, albeit challenging, direction for future research.

7 Conclusion

We have presented a comprehensive system for reconstructing a dynamic hair model from synchronized multi-view video sequences. Our approach exploits hair motion paths extracted from horizontal slices in a spatiotemporal video volume. These motion paths enable us to warp hair ribbons and constrain hair motions, and thereby to formulate the dynamic hair reconstruction as a spacetime optimization problem. When reconstructing hair motion details, we avoid favoring either the spatial or the temporal smoothness over the other. Instead, we postpone the generation of a complete hair model until we have analyzed directed ribbons and motion paths separately. We then reconstruct the hair model in an optimization framework that considers both spatial and temporal constraints together. Demonstrated with a number of hairstyles, our reconstructed hair motion faithfully capture the real hair dynamics in a temporally coherent way while retaining plausible motion details.

Acknowledgment

We would like to thank the anonymous reviewers for their constructive feedback, Bosheng Zhou for the help on device setup, Menglei Chai for his assistance on hair rendering, and Linjie Luo for kindly providing testing data. Yue Qi is supported by National Natural Science Foundation of China (No. 61272348, 61202235), Ph.D. Program Foundation of Ministry of Education of China (No. 20111102110018), National Key Technology Research & Development Program of China (No. 2014BAK18B01).

References

BEELER, T., HAHN, F., BRADLEY, D., BICKEL, B., BEARDSLEY, P., GOTSMAN, C., SUMNER, R. W., AND GROSS, M. 2011. High-quality passive facial performance capture using anchor frames. *ACM Trans. Graph. 30* (Aug.), 75:1–75:10.

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, 117:1–10.

BICKEL, B., BOTSCH, M., ANGST, R., MATUSIK, W., OTADUY, M., PFISTER, H., AND GROSS, M. 2007. Multi-scale capture of facial geometry and motion. *ACM Trans. Graph. 26*, 3 (July).

BOLLES, R. C., BAKER, H. H., AND MARIMONT, D. H. 1987. Epipolar-plane image analysis: An approach to determining structure from motion. *International Journal of Computer Vision 1*, 7–55.

BONNEEL, N., PARIS, S., PANNE, M. V. D., DURAND, F., AND DRETTAKIS, G. 2009. Single photo estimation of hair appearance. *Computer Graphics Forum 28*, 1171–1180.

BRADLEY, D., POPA, T., SHEFFER, A., HEIDRICH, W., AND BOUBEKEUR, T. 2008. Markerless garment capture. *ACM Trans. Graph. 27*, 3 (Aug.), 99:1–99:9.

BRADLEY, D., HEIDRICH, W., POPA, T., AND SHEFFER, A. 2010. High resolution passive facial performance capture. *ACM Trans. on Graphics (Proc. SIGGRAPH) 29*, 3.

BRADSKI, G. 2000. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

CAI, M., ZHENG, C., AND ZHOU, K. 2014. A reduced model for interactive hairs. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2014) 33*, 4 (Aug.).

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

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

CHEN, Y., DAVIS, T. A., HAGER, W. W., AND RAJAMANICKAM, S. 2009. Algorithm 887: Cholmod, supernodal sparse cholesky factorization and update/downdate. *ACM Trans. Math. Softw.*.

CHEN, Q., LI, D., AND TANG, C.-K. 2012. KNN matting. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012*, 869–876.

DAVIS, T. A., AND HAGER, W. W. 2009. Dynamic supernodes in sparse cholesky update/downdate and triangular solves. *ACM Trans. Math. Softw. 35*, 4.

DE AGUIAR, E., STOLL, C., THEOBALT, C., AHMED, N., SEIDEL, H.-P., AND THRUN, S. 2008. Performance capture from sparse multi-view video. *ACM Trans. Graph. 27*, 3 (Aug.), 98:1–98:10.

FU, H., WEI, Y., TAI, C.-L., AND QUAN, L. 2007. Sketching hairstyles. In *Proc. EUROGRAPHICS Workshop on Sketch-Based Interfaces and Modeling*, 31–36.

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

GRABLI, S., SILLION, F., MARSCHNER, S., AND LENGYEL, J. E. 2002. Image-based hair capture by inverse lighting. In *Proc. Graphics Interface*.

GREGSON, J., KRIMERMAN, M., HULLIN, M. B., AND HEIDRICH, W. 2012. Stochastic tomography and its applications in 3D imaging of mixing fluids. *ACM Trans. Graph. 31*, 4 (July), 52:1–52:10.

HERRERA, T. L., ZINKE, A., AND WEBER, A. 2012. Lighting hair from the inside: A thermal approach to hair reconstruction. *ACM Trans. Graph. 31*, 6, 146:1–9.

HU, L., MA, C., LUO, L., AND LI, H. 2014. Robust hair capture using simulated examples. *ACM Trans. Graph.*.

IHRKE, I., AND MAGNOR, M. 2004. Image-based tomographic reconstruction of flames. In *Proceedings of the 2004 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, SCA '04, 365–373.

ISHIKAWA, T., KAZAMA, Y., SUGISAKI, E., AND MORISHIMA, S. 2007. Hair motion reconstruction using motion capture system. *ACM SIGGRAPH 2007 posters*.

JAKOB, W., MOON, J. T., AND MARSCHNER, S. 2009. Capturing hair assemblies fiber by fiber. *ACM Trans. Graph. 28*, 5, 164:1–9.

KIM, C., ZIMMER, H., PRITCH, Y., SORKINE-HORNUNG, A., AND GROSS, M. 2013. Scene reconstruction from high spatio-angular resolution light fields. *ACM Trans. Graph. 32*, 4, 73:1–12.

KONG, W., TAKAHASHI, H., AND NAKAJIMA, M. 1997. Generation of 3D hair model from multiple pictures. In *Proc. Multimedia Modeling*.

LI, Y., FAN, X., MITRA, N. J., CHAMOVITZ, D., COHEN-OR, D., AND CHEN, B. 2013. Analyzing growing plants from 4D point cloud data. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2013) 32*, 6.

LUO, L., LI, H., WEISE, T., PARIS, S., PAULY, M., AND RUSINKIEWICZ, S. 2011. Dynamic hair capture. Tech. rep., Princeton University.

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

LUO, L., ZHANG, C., ZHANG, Z., AND RUSINKIEWICZ, S. 2013. Wide-baseline hair capture using strand-based refinement. *Computer Vision and Pattern Recognition*.

MCLACHLAN, G. J., AND KRISHNAN, T. 1997. The EM algorithm and extensions. In *Wiley series in probability and statistics*. John Wiley & Sons.

MÜLLER, M., HEIDELBERGER, B., HENNIX, M., AND RATCLIFF, J. 2007. Position based dynamics. *Journal of Visual Communication and Image Representation 18*, 2, 109–118.

PARIS, S., BRICEÑO, H., AND SILLION, F. 2004. Capture of hair geometry from multiple images. *ACM Trans. Graph. 23*, 3, 712–719.

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, 30:1–9.

SORKINE, O., COHEN-OR, D., LIPMAN, Y., ALEXA, M., ROESSL, C., AND SEIDEL, H.-P. 2004. Laplacian surface editing. In *Symposium on Geometry Processing*.

VLASIC, D., BARAN, I., MATUSIK, W., AND POPOVIĆ, J. 2008. Articulated mesh animation from multi-view silhouettes. *ACM Trans. Graph. 27*, 3 (Aug.), 97:1–97:9.

WANG, H., LIAO, M., ZHANG, Q., YANG, R., AND TURK, G. 2009. Physically guided liquid surface modeling from videos. *ACM Trans. Graph. 28*, 3 (July), 90:1–90:11.

WANG, L., YU, Y., ZHOU, K., AND GUO, B. 2009. Example-based hair geometry synthesis. *ACM Trans. Graph. 28*, 3, 56:1–9.

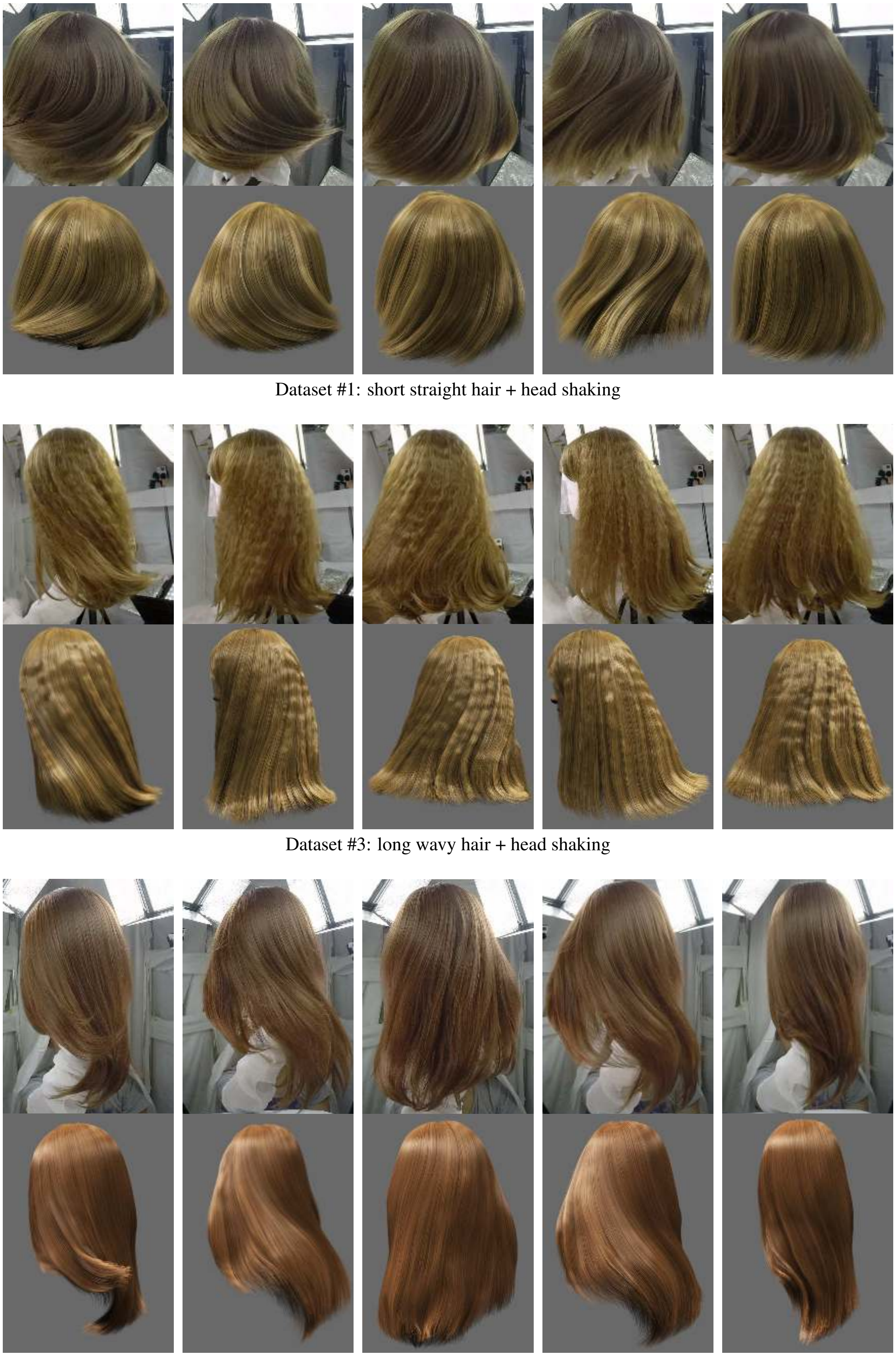
WARD, K., BERTAILS, F., KIM, T.-Y., MARSCHNER, S. R., CANI, M.-P., AND LIN, M. C. 2007. A survey on hair modeling: styling, simulation, and rendering. *IEEE Transactions on Visualization and Computer Graphics 13*, 2, 213–234.

WEI, Y., OFEK, E., QUAN, L., AND SHUM, H.-Y. 2005. Modeling hair from multiple views. *ACM Trans. Graph. 24*, 3, 816–820.

WU, C., AGARWAL, S., CURLESS, B., AND SEITZ, S. M. 2011. Multicore bundle adjustment. In *Proc. CVPR*.

YAMAGUCHI, T., WILBURN, B., AND OFEK, E. 2009. Video-based modeling of dynamic hair. In *the 3^{rd} Pacific Rim Symposium on Advances in Image and Video Technology*, 585–596.

ZHANG, Q., TONG, J., WANG, H., PAN, Z., AND YANG, R. 2012. Simulation guided hair dynamics modeling from video. *Computer Graphics Forum 31*, 7, 2003–2010.



Dataset #1: short straight hair + head shaking

Dataset #3: long wavy hair + head shaking

Dataset #5: long straight hair + head shaking

Figure 13: Dynamic hair capture results: *Please see the accompanying video for the complete motion results.*