Capture of Hair Geometry from Multiple Images

Sylvain Paris   Hector M. Briceño   François X. Sillion

Artis GRAVIR/IMAG-INRIA*

Abstract

Hair is a major feature of digital characters. Unfortunately, it has a complex geometry which challenges standard modeling tools. Some dedicated techniques exist, but creating a realistic hairstyle still takes hours. Complementary to user-driven methods, we here propose an image-based approach to capture the geometry of hair.

The novelty of this work is that we draw information from the scattering properties of the hair that are normally considered a hindrance. To do so, we analyze image sequences from a fixed camera with a moving light source. We first introduce a novel method to compute the image orientation of the hairs from their anisotropic behavior. This method is proven to subsume and extend existing work while improving accuracy. This image orientation is then raised into a 3D orientation by analyzing the light reflected by the hair fibers. This part relies on minimal assumptions that have been proven correct in previous work.

Finally, we show how to use several such image sequences to reconstruct the complete hair geometry of a real person. Results are shown to illustrate the fidelity of the captured geometry to the original hair. This technique paves the way for a new approach to digital hair generation.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism; I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Shape I.5.2 [Pattern Recognition]: Design Methodology—Feature Evaluation and Selection

Keywords: Hair Capture, Hair Modelling, Reflectance Analysis, Shape from Shading, 2D Orientation Detection, Signal Processing

1 Introduction

Hair is an important part of the general aspect of a human character. It has a great impact on the look of the head and therefore requires special attention when creating digital characters. Today, hair manipulation relies strongly on user input. The process can be split into three parts: modeling, animating, and rendering. Great progress has been made for animation [Anjyo et al. 1992; Plante et al. 2001; Bertails et al. 2003] and rendering [Kajiya and Kay 1989; Kim and Neumann 2001; Marschner et al. 2003]. Unfortunately, hair modeling still requires a user-assisted process. Dedicated methods exist to drive the hair creation process and help the user: [Hadap and Magnenat-Thalmann 2000; Kim and Neumann 2002]. These clearly allow a fine control over the geometry but it becomes tedious if one wants to reproduce complex features like curls and waves of a real character.

We focus on this modeling step and propose an alternative solution to the creation of the hair geometry: using sequences of images

Figure 1: Sample result of our technique. Left: one of the original images; right: the reconstructed model. Our contribution is to capture the original hair geometry of a person; the rendering step is out of the scope of this paper and provided only for illustration purpose. From image sequences with fixed viewpoints and a moving light source, our technique automatically captures a dense set of hairs suitable for further processing.

of a real hair to produce a set of hair strands which capture most of the features of the original hair. This hair-from-images approach is complementary to the hair-from-user techniques. The produced set of strands can be used directly "as is" or may be the starting point of an user-driven approach. The advantage of the image-based approach is the reproduction of the hair of an existing person. The main application is virtual actors: more and more movies use digital clones of the human actors in their action scenes to shoot dangerous or technically impossible situations. Until now, the hair of these digital actors is either short or straight – often a poor match with the original hair. The presented technique is a first step toward duplicating a real hairstyle with complex features. Furthermore, one can also imagine using this approach in games (to have your own clone inside the game), styling software (to preview color changes of your hair), etc.

The goal of our method is to produce a set of hairs which can be rendered under different lighting conditions. An alternative method could produce a textured surface, but this would be hard to animate and render properly. We strive for a dense set of hair such that no holes are visible. We would, in the future, also like to capture the reflectance parameters and adapt the hair for animation. Our strategy is to overcome one of the major difficulties with hair: light interaction that results in large specularities, scattering, glints, etc. These properties hinder the use of 3D scanners and computer vision techniques. We turn this specific behavior into a source of information: we analyze the hair under various illuminations to extract their geometry. Our approach is to shoot image sequences of a still head with a fixed viewpoint under a moving light source. The correspondence between images is straightforward but gives no stereoscopic localization, so depth and shape information come from appearance variations. In that respect, it is related to *shape-from-shading* [Brooks and Horn 1989] and *shape-from-specularities* [Yang et al. 2003]. Unfortunately, these techniques only recover a surface and may have trouble with complex materials. Even example-based methods [Hertzmann and Seitz 2003] that overcome this last point, recover only a surface which is poor approximation of the hair geometry.

2 Related Work

There are many methods for rendering hair from synthetic data, but few for capturing it [Rushmeier 2001]. Traditional 3D capture

systems, such as laser scanners, have trouble with hair due to its complex reflection properties, and yield erroneous results.

Matusik *et al.* [2002] propose using image-based rendering for capturing hair. Their method may suffice for rendering, but without geometrical information, it would be difficult to animate or edit captured objects. For better editing and rendering, we need a geometrical model of hair. There are many packages for generating hair models [Daldegan et al. 1993; Hadap and Magnenat-Thalmann 2000; Kim and Neumann 2002], but they start from scratch and it can be time-consuming to generate complex hairstyles.

The work of Nakajima *et al.* [1998] considers modeling hair from pictures. It works by building a 3D hair volume from various viewpoints of the subject's hair. They point out the difficulty of extracting a complete hair strand from an input image, hence, they propose an heuristic to fill this volume with straight strands. There are no guarantees regarding the directionality of hair strands; thus, this method seems unlikely to handle complex hairstyles well.

Recent work in this area from Grabli *et al.* [2002], is the most relevant to ours. This system works by studying the subject's hair under various lighting conditions. By fixing the viewpoint, they can work with perfectly registered images. Their approach uses a single filter to determine the orientation of hair strands, therefore, many images and sequences may be required to achieve a dense sampling. They only consider one general viewpoint and thus do not reconstruct all of the hair. Our technique builds upon their approach and addresses these short-comings.

3   General Overview

Before exposing the main ideas of our method, we give a few definitions to clarify useful entities.

Definitions: We call a *fiber* a single and entire hair. A *strand* is a small group of fibers that are tightly grouped together along their whole length. This is the visible entity in images since a fiber width is smaller than a pixel. A *segment* is a small section (≈ 1 mm) of a strand, it is well approximated by a small line.

We call *orientation*, the un-directed line containing a segment. To characterize one of the two corresponding directed lines, we need to provide a *direction*.

Global Approach: Our strategy relies on image sequences with fixed viewpoints and moving light source. On the one hand, since the camera does not move, a pixel always represents the same 3D location on the hair surface. On the other hand, this gives no stereoscopic information about this 3D location. Therefore the 3D information comes from the analysis of the image variations throughout the sequence. This analysis considers the segments individually and is split into two parts: a 2D analysis of the image properties and 3D analysis of the illumination variations. The former recovers the 2D orientation of a segment projected in the image plane and characterizes a plane containing the segment (Fig. 3). The latter gives a normal to the segment that results in a second plane. Intersecting both planes forms the 3D orientation of the segment. Linking these segments together builds strands.

To form a full geometric models of a person's hair, we propose in this paper a method to merge the information from several image sequences with different viewpoints.

Limitations: The method described in this paper is widely usable. However, there are few assumptions and some cases cannot be handled. Since we work from images, hidden hair strands are not captured. For instance, curls that form toward the camera are partially reproduced: only the visible half is captured. Furthermore, our technique relies on the assumption that hair strands are thin and that their orientation is visible in images. This implies that we cannot handle thick strands (like dreadlocks) or short hairs pointing toward the camera. Lastly, there are inputs where our system would not work well; for example, tangled hair might be unrecoverable
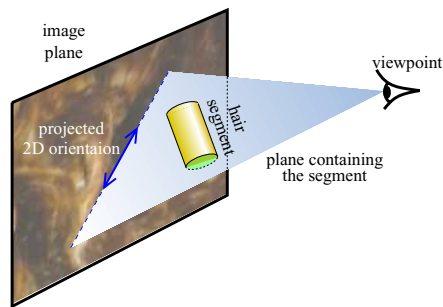

Figure 2: The viewpoint and the 2D orientation of a segment in image

because we work from the assumption that there is one orientation per pixel.

Contribution: This paper introduces the following new ideas:

- We outline a deep link between statistics and signal processing. This makes our 2D orientation computation robust, precise, and dense. This also allows us to evaluates the "quality" of the measurements, making it possible to compare different measurements.

- We show how to use only a minimal set of assumptions to recover the 3D aspect of the hair (in contrast to other methods which assume the whole reflectance model to be known, *e.g.*, [Grabli et al. 2002]).

- We explain how to merge several 3D parts of the same hair into a single consistent 3D model.

4   Orientation of the Segments

The first step in reconstructing complete hair strands is to retrieve the 3D orientation of each segment.

4.1   Capturing 2D Orientation

Our technique starts by measuring the 2D orientation of each segment projected in the images. Since a segment contains several fibers with the same orientation, it induces a local orientation in the image. This steps boils down to a classical signal processing issue: what is the local orientation of an image?

Many approaches exist to give an answer to this question. Ziou and Tabbone [1998] offer an overview of edge-detection technique, and the signal-processing literature suggests many approaches [Freeman and Adelson 1991; Granlund and Knutsson 1995; Donoho and Huo 2000]. In general, these methods are proven "optimal" under some theoretical hypotheses on the images. Unfortunately, we face a more complex problem. First, fibers are smaller than a pixel and introduce aliasing. Moreover, hair lighting properties (self-shadowing, light scattering, etc.) make it hopeless to predict any strong properties (*e.g.*, the size of oriented segments or the shapes of edges). Therefore, we only assume that there is one orientation in each pixel.

Our approach is not to propose a new filter but to choose among some existing filters the one that gives the most reliable results. Some existing methods [Granlund and Knutsson 1995; Felsberg and Sommer 2000; Meer and Georgescu 2001] evaluate their own reliability. This evaluation is filter-specific and cannot be compared across filters. Some case studies [Baker and Nayar 1999] compare filters on reference images regarding some known properties; *e.g.*, parallel edges. As discussed before, hair images have few such properties.

Our evaluation is based on *oriented* filters [Freeman and Adelson 1991]. A filter in this class is defined by its kernel K designed to detect an x-aligned orientation. To test an arbitrary orientation θ in a image I at pixel (x, y), K is rotated by θ (K_θ in short) and convoluted with I. It produces a score $F(\theta) = |K_\theta * I|(x, y)$. The

result of the filter is $\tilde{\theta} = \text{argmax}(F(\theta))$ – the orientation with the highest score.

We then observe the response curve of such a filter. We are seeking a peaked curve. If the curve is flat, the result is uncertain because other results can almost fulfill the same criterion. The peakiness of the curve is evaluated with its variance. Let $F(\theta)$ be the filter response for orientation $\theta \in]-\frac{\pi}{2}, +\frac{\pi}{2}]$, $\hat{F} = F/\int F$ be the normalized version of F and $d(\theta_1, \theta_2) = \min(|\theta_1 - \theta_2|, |\theta_1 - \theta_2 \pm \pi|)$ be the angular distance between θ_1 and θ_2. The measure is then the classical variance formula:

$$V(F) = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} d^2(\tilde{\theta}, \theta) \hat{F}(\theta) d\theta \qquad (1)$$

Comparison property: Since the variance only considers the normalized response curve of the oriented filter, it is independent of the image and of the oriented filter. Therefore, different filters for different pixels on different images can be rigorously compared together.

Moreover, this measure has other good points:

- The normalization makes it insensitive to the scale of the filter and of the image *e.g.* an intensity scale does not change our evaluation.
- When formula (1) is discretized, comparison can be done for any number of θ samples arbitrarily chosen as long as each sample is weighted with the measure function $d\theta$.

Ishikawa [2000] also introduces a statistical approach to chose among point-matching criteria. But since his method is based on entropy, it is sensitive to the number and positions of the samples and is invariant to a permutation of the response values, *i.e.*, it does not distinguish whether the significant scores are grouped or randomly spread. Yitzhaky and Peli [2003] use a statistical analysis to aggregate edge-detector results into a single edge map. But Forbes and Draper [2000] show that this evaluation is image dependent. Moreover, it requires numerous filters to properly work whereas our technique gives meaningful results with as few as two filters.

Practical Uses: For a given 2D location, several filters are tested with several parameters on images with different lighting. $V(\cdot)$ gives an objective rating to select the "best" option. Furthermore, as we will see later (§ 4.1.2), we can use these values to compare and compute the similarity between adjacent pixels to enhance our results.

4.1.1 Filters

We now detail the oriented filters that we use and their parameters. The first parameter is the scale at which the images are analyzed. Figure 3 shows that hair appearance significantly varies with the observed wavelength. Frequency selection is done with a band-pass filter (a difference-of-Gaussian filter in practice). Since we track hair strands which are very small, we only select the short wavelengths of the image with a Gaussian band-pass filter centered on $\lambda = 2$ to be as close as possible of the Nyquist sampling rate.

We test 64 orientations regularly spaced; this is a good trade-off between accuracy and computation time.

Gaussian first derivative (Canny)	$-x\,e^{-\frac{1}{2}(x^2+y^2)}$
Gaussian second derivative	$(x^2-1)\,e^{-\frac{1}{2}(x^2+y^2)}$
Canny Deriche	$x\,e^{-|x|-\frac{1}{2}y^2}$
Shen Castan	$-\frac{x}{|x|}\,e^{-|x|-\frac{1}{2}y^2}$
Gabor	$\cos(\omega x + \phi)\,e^{-\frac{1}{2}(x^2+y^2)}$

Table 1: Filters we use. Formulæ are given to detect a signal parallel to the *y* axis, rotation terms and scale factors are omitted for clarity.
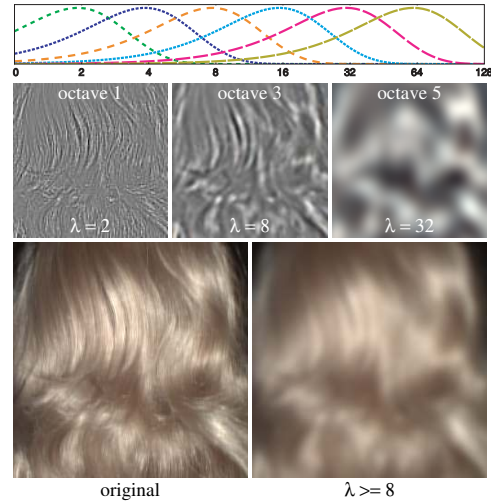

Figure 3: Frequency decomposition of an image. The top plot shows the frequency profiles of the band-pass filters. The middle row shows 3 sample frequency bands *a.k.a.* octaves (intensity is scaled to [0,1]). The bottom row shows the original image and a low-frequency version (octaves 1 and 2 are removed). The high-frequency image ($\lambda = 2$) mainly represents the micro-structure of the image due to the hair strands, while the low-frequency one ($\lambda \geq 8$) presents its macro-structure due to illumination. The second octave ($\lambda = 4$) is skipped because it still contains significant frequencies from the first one (observe the top plot).

We have chosen some oriented filters among the relevant existing ones (mainly edge and line detectors). Table 1 shows the kernels we use [Shen and Castan 1986; Deriche 1987; Feichtinger and Strohmer 2003; Canny 1983]. Canny [1983] shows that these filters can be decomposed into a *detector profile* that detects the signal variations and a *projection profile* orthogonal to it, that accounts for the neighborhood of the examined point.

Detector Profile: Depending on this profile, a filter detects the orientation of different features. This has an impact on the response curve (Fig. 4). Therefo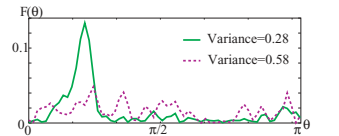re several profiles are used (plot in Figure 5). They are scaled so that their *pseudo-wavelength* (*i.e.*, the wavelength that best describes their largest variations) equals the high-pass wavelength (2 pixels in our case). For the Gabor filter, from our experiments, we have found that testing 4 values for $\phi \in \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\}$ sufficiently covers the possible phases. We set $\omega = 1$: smaller values do not add significant improvements and larger values detect wider features (several strands side by side) that do not correspond to what we track (the very local orientation of a segment) and degrade the results.

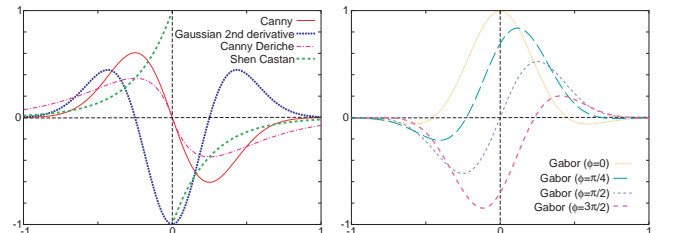Figure 4: Response curves of two different filters applied to the same pixel.

Figure 5: Detector profiles. These are the shapes of the filter kernels along the axis in which we are looking for a signal variation. The profiles are scaled to a unit pseudo-wavelength.
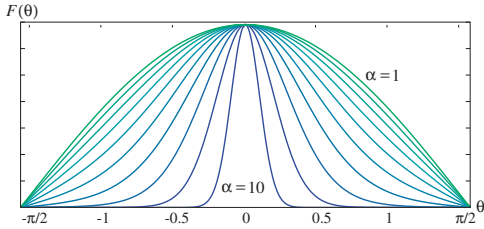
Figure 6: Influence of the projection profile. The plots show the responses $F(\theta)$ on a sine signal $(x,y) \mapsto \sin(x)$ of a Canny filter whose projection profile is stretched α times.

Projection Profile: This accounts for the local environment. A Gaussian shape is generally used for this profile. Canny [1983] shows that its extension improves efficiency; at the same time, he points out that orientation in real images is a local phenomenon and the extension cannot be too large. We have computed the theoretical response of Canny's filter on a perfect sine signal (Fig. 6). It proves that the variance estimator subsumes Canny's remark: the more extended the profile is, the lower is the variance (for an infinitely extended signal) – the better the filter is. In practice, the best results are reached with Gaussian profiles of standard deviations of 2, 4, and 8 pixels. Thus we test these 3 values.

4.1.2 Data Enhancement with Bilateral Filtering

Hair geometry has a local coherence: even if there are some discontinuities, in most cases, neighboring points are similar, thus interact analogously with the light. One consequence is the large extent of the highlights on the hair surface. Information can be extracted from this coherence. If a point has a poorly evaluated orientation, it can be "corrected" by the neighboring points that are more reliable. Furthermore, if these neighboring points have the same appearance in the images they are likely to have the same geometry.

So we locally spread the information according to the reliability of the measure (*i.e.*, its variance) and the appearance of the pixel in the sequence.

We apply a treatment inspired by the bilateral filter introduced by Tomasi and Manduchi [1998]. A diffusion process (*e.g.*, [Tschumperlé and Deriche 2002]) could also be used; our choice is motivated by recent results such as [Durand and Dorsey 2002] and by the theoretical properties from [Elad 2002].

The resulting orientation $\overline{\mathbf{o_p}}$ at point $\mathbf{p}$ is a weighted mean[1] over the neighborhood $\mathcal{N}_{\mathbf{p}}$:

$$\overline{\mathbf{o_p}} = \sum_{\mathbf{q} \in \mathcal{N}_{\mathbf{p}}} w_s(\mathbf{p},\mathbf{q}) w_v(\mathbf{p},\mathbf{q}) w_c(\mathbf{p},\mathbf{q}) \overline{\mathbf{o_q}}$$

The weight of $\overline{\mathbf{o_q}}$ is split: into a spatial term w_s that considers the location of $\mathbf{p}$ and $\mathbf{q}$; a variance term w_v that accounts for the reliability of the measure; and a color term w_c that accounts for the color similarity in the sequence. We use Gaussian functions for each of them. The first one uses the Euclidian distance: $w_s(\mathbf{p},\mathbf{q}) = \exp\left(-||\mathbf{p}-\mathbf{q}||^2/\sigma_d^2\right)$. For the variance term, we rely on the ratio $\rho(\mathbf{p},\mathbf{q}) = V_{\mathbf{q}}/V_{\mathbf{p}}$ to define $w_v = \exp(-\rho(\mathbf{p},\mathbf{q})/\sigma_\rho^2)$. The appearance similarity is evaluated from the maximum color difference Γ between $\mathbf{p}$ and $\mathbf{q}$ in the image sequence. This comparison must only account for illumination similarity disregarding other phenomena like glints that are related to the fiber structure and not to the strand geometry [Marschner et al. 2003]. Therefore, Γ is computed on the low frequencies of the images (Fig. 3). This leads to $w_c = \exp(-\Gamma^2(\mathbf{p},\mathbf{q})/\sigma_\Gamma^2)$. Figure 7 illustrates the significant improvement brought by this treatment.

In our experiments, the best results are achieved with: $\sigma_d = 3$, $\sigma_\rho = 1$ and $\sigma_\Gamma = 0.1$.

---

[1]To compute the weighted mean of several 2D orientations, the orientation $\phi \in [0, \pi[$ with weight w is mapped to the complex number $c = w \exp(2i\phi)$. The reverse mapping is then $\phi = \frac{1}{2}\arg(c)$. See [Watson 1983]

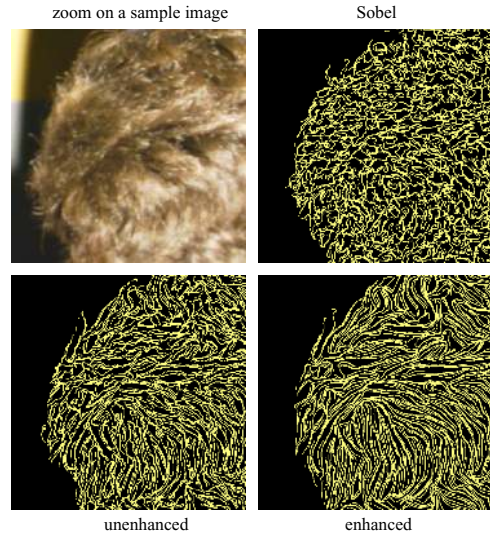zoom on a sample image   Sobel

unenhanced   enhanced

Figure 7: Flow lines from 2D orientation fields. Upper-left: A closeup image from the sequence. Upper-right: Results of Sobel filter. Lower-left: Our *unenhanced* results before the bilateral filter. Lower-right: Same results *enhanced* by the bilateral filter.

4.1.3 Comparison and Validation

Figure 7 compares the orientations computed by the Sobel technique used by Grabli *et al.* [2002] with our *unenhanced* results and with the same results *enhanced* by a bilateral filter. Sobel fails to provide any satisfying orientations whereas the enhanced results are convincing. Figure 8 show the importance of using various filters and bilateral filtering on the final result of the pipeline.

The method is further validated with a reference image of known orientations. The result of this experiment is shown in Figure 9. We have made a numerical evaluation of: the Canny filter in its classical use (the gradient is estimated from the x and y first derivatives of an isotropic Gaussian, the pseudo-wavelength is set to 2), the Sobel filter, our unenhanced and enhanced method; the mean errors for these filters are: 43°, 17°, 2.9°, and 2.3°, respectively. It outlines the precision of our method and shows that our bilateral filtering is a real enhancement (20% better) and not only visually pleasing.

4.2 Capturing the Normal Vector

From our 2D study, a given segment is constrained to lie in a plane (Fig. 3). We characterize its orientation inside that plane by analyzing its appearance under varying illumination. This relies on a minimal knowledge of the hair scattering model from Marschner *et al.* [2003]: The specular peak occurs in the standard reflection direction but the surface is slightly tilted toward the root because of the *cuticles*.

Therefore, from the intensity curves (intensity vs. light position) we can recover the light position corresponding to maximum reflection. As explained by Lu *et al.* [1999], this gives a normal to the hair fiber (Fig. 10).

By fixing the viewpoint (and the subject) we can observe the same segment under varying lighting conditions. Consider a light describing a circular motion around a segment modeled as a cylinder. If the light motion is not perpendicular to the segment axis, we are guaranteed to have a specular highlight. If the light motion is perpendicular to it, there is no intensity peak because we are always in the specular region. To avoid this, we capture images with the light moving in two orthogonal planes (from left-to-right and bottom-to-top). Thus, we are sure to have a useful sequence. We select the one whose plane has the lowest angular deviation with the computed 2D orientation. For example, if the segment is vertically aligned, we choose the bottom-to-top path so that the trajectory is never orthogonal to the segment. Figure 11 shows sample intensity

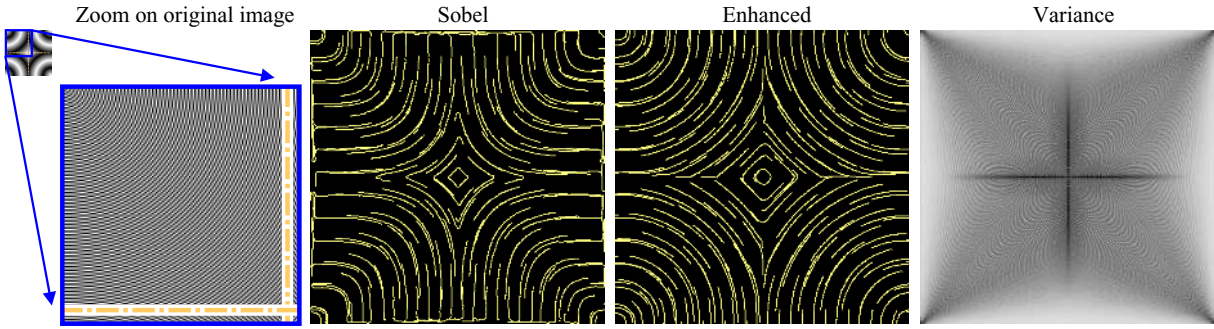| Zoom on original image | Sobel | Enhanced | Variance |

Figure 9: Validation of the orientation measure. The original image is composed of 4 radial sine signals centered on the 4 corners with wavelength $\lambda = 2$ (dotted lines are the symmetry axes of the image). The image is aliased by itself because of the short wavelength; there may be additional artifacts due to your printer or screen. Orientation measures are actually dense (1 per pixel) but only some flow lines are shown for clarity. The Sobel filter produces too curved results whereas our *enhanced* (using variance-based filter selection and bilateral filtering) data retrieve correct values. The right image shows the reliability (variance: white for 0, black over 1) of the filter selected by our method. As expected, it distinguishes the discontinuities and circle centers that make the measure less reliable.
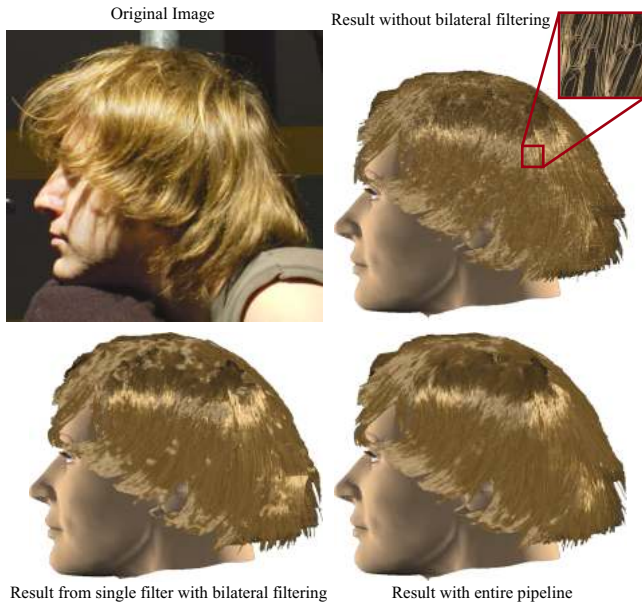

Figure 8: Performance. Upper-left: Original Image from sequence. Upper-right: Result from selecting among many filters (different pixels can use different filters) but no bilateral filtering; note the noisy aspect of the image. Lower-left: Result from using single Gabor filter on a single image with bilateral filtering; note the large error on the right and wrong highlights. Lower-right: Result from selecting among many filters with bilateral filtering.

profiles that are obtained from these light paths.

This technique still holds for an elliptical cylinder hair model without modification, but it would need to compensate for the cuticle angle by slightly rotating the computed segment toward its root. However, we have found that the results are satisfactory without this offset and in practice, on some complex hairstyles, the root direction may be unknown.

One important caveat is that the light source does not describe a full circle. Our measurements would be impaired. Forward scattering would occur if the segment were between the light and the camera. We therefore limit the light-view angle to lie between $\pm\frac{\pi}{2}$ (For the vertical light path the bottom angle is limited by the floor to $\approx -20^{\circ}$). This bounds the detectable normal between $\pm\frac{\pi}{4}$. Fortunately, this addresses the case of a great majority of the visible segments. Moreover, the segments which have normals outside this interval generally lie near the silhouettes and will be better captured from another point of view.

5 Practical Implementation

In the previous section we have shown how to build a 3D orientation field (sometimes named *line field* in literature) from a given image sequence from a given viewpoint. To build a full model of a person's hair we need multiple viewpoints. For our initial implementation, we use four sequences: right, left, back, and top. We register these four sequences to build a 3D orientation field that covers the whole head. The final part of the algorithm is to grow strands following the 3D orientations.

5.1 Viewpoint Registration

To generate hair throughout a head, we need a 3D orientation field that covers the whole head. We use a simple setup to capture image sequences using a single camera. Hence, we need to register the different viewpoints on a common coordinate system (the intrinsic parameters of the camera are known and constant). A more sophisticated setup using multiple cameras could obviate this step.

Our registration is done manually: An ellipsoid is fitted to bound the hair volume. In practice, a bounding ellipse (the projection of the ellipsoid into the viewing plane) is fitted for each viewpoint. Using the camera parameters and the ellipses, the ellipsoid is characterized and the cameras are located relatively to it (details omitted because of space limitation). The camera-to-image precision is in the order of 1mm (previous steps rely only on this one) whereas the registration between the different viewpoints is in the order of 10mm impacting only the thin overlapping regions (Fig. 12).

5.2 Hair Growth

For each view, we define a hair region (mask) which we use to compute the *visual hull* [Laurentini 1994] of the hair volume. We limit the influence of a viewpoint to half of the volume because the left viewpoint would interfere with the right one – for instance, the boundaries around the ears may not match. The visual hull is only
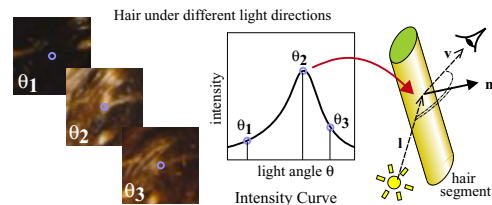

Figure 10: Determination of a vector normal to a hair segment. For each pixel, the maximum intensity in the image sequence characterizes the peak reflection configuration for the underlying segment.
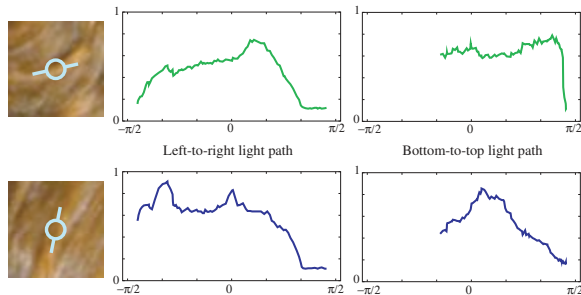
Figure 11: Intensity plots for two segments with different orientations. The light paths orthogonal to the segments' axes are flat (with potentially the 2 caustic lobes predicted by Marschner *et al.* [2003]) and give no information about the normal. The other paths have a "diffuse+specular" response that characterizes the normal. The choice between both paths is made from the 2D orientation of the segment.
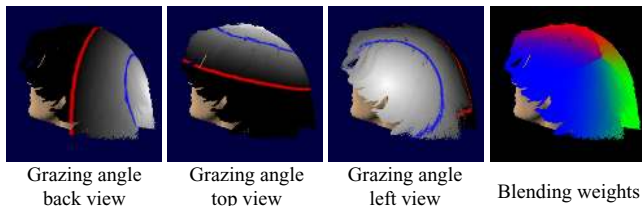

Figure 12: The grazing angles relatively to three views; from white (0°) to black (over 90°). The blue and red lines correspond to 45° and 90°. The last image shows the blending weights deduced for the three views (blue: left, green: back, red: top).

used to ensure that the synthetic hair lies inside the original hair volume.

The bounding ellipsoid we used for the viewpoint registration (§ 5.1) is then shrunk to fit inside the hull to approximate of the *skull*. More precisely, this will be the surface of the starting points of the visible strands. When the hair is thick, this surface will be slightly bigger than the real skull. Alternatively, another method for acquiring or approximating a person's skull could be used, *e.g.*, medical X-rays.

To form a strand, a point is picked on this skull. From it, we iteratively chain the 3D segments (Fig. 13). The 3D segment at point **p** is computed by projecting **p** into the visible image planes. The 3D orientation at the projected point is computed in one of two ways:

If only one camera "sees" the point. The 3D segment is straightforwardly computed with a length that corresponds to the back-projection of the image pixel ($\approx$1mm); its direction is the one that is most similar to the last segment added to this strand (no sharp angles). The initial direction of the strand is chosen to match the general hair direction e.g. up-to-down for most long hairs. If no such direction exists, the direction pointing outside the skull is used.

Alternatively, if more than one camera sees the point, we limit the use of grazing lines of sight, because corresponding normals may not be accurate as previously discussed. We select the two views with the lowest grazing angles ψ (Fig. 12). We then average the corresponding orientations[2] weighted by $\cos^2(\psi)$. Figure 12 shows the influence of each view and confirms that 4 views (top, back, left and right) correctly cover the hair volume. In practice, only in limited overlapping areas is the 3D information for pixels in different views blended.

A strand is ended if it touches the visual hull boundary or if it reaches a certain length; the latter being more common. To re-

---

[2]In 3D, the sum of more than 2 orientations is not defined. But it is defined for 2 orientations by using the 2D case in their common plane.

duce unnecessary computation, we end the hair strand if more than a number of them pass over a pixel (10 segments/pixel in practice). Depending on the model, we generate from 30 000 to 70 000 strands.

Though it is a robust process, some isolated segments may still be wrongly oriented. Even if their number is limited, their visual effects can be noticeable. Therefore, in a post-process we reduce the highest curvatures of a strand with a diffusion process on the vertices **p** for which the curvature κ is higher than a threshold κ_0. The strand evolves according to $\partial \mathbf{p}/\partial t = \partial^2 \mathbf{p}/\partial^2 s$ if $\kappa > \kappa_0$; 0 otherwise (with s the curvilinear abscissa) until it stabilizes. In practice, we use κ_0 to bound the curvature radius over $\approx$ 10mm.
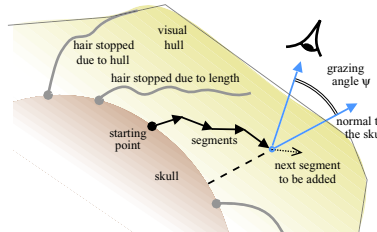

Figure 13: Hair strand growth: the starting point is picked on the skull, segments are then added one by one. The influences of the cameras are weighted by their grazing angles.

6   Captured Hair Results

Setup: To acquire the image sequences, we have a setup in which the subject whose hair is to be captured is able to keep his or her head fixed for several seconds. A fixed video camera captures images at 7.5 frames per second with a 1024×768 resolution (in practice a region of $\approx 550 \times 550$ for the hair because the mirror balls must visible). A point light source aimed at the subject's hair is moved while its distance to the head is constrained ($\approx$1 m) and its angular position is known thanks to 2 mirror balls. Figure 14 shows a picture of our setup. To acquire several viewpoints, the person turns his head.

In a few seconds, hundreds of images are taken. Thus we can assume that the hair is not moving throughout the sequence. For each viewpoint we currently segment the hair region manually. Four viewpoints are used: top, right, left, and back. This is a minimal set for reconstructing the whole model of hair. A more sophisticated setup with more cameras would do this in one pass.

Rendering: We use the rendering model of Marschner *et al.* [2003] for our visual comparisons. The parameters are determined to roughly match the original color of the hair. This model is restricted to a single fiber and does not account for hair-to-hair shadowing, scattering, etc. A neutral head is placed under the hair to provide a consistent image. Each strand is rendered by anti-aliased GL lines; the color is computed at each vertex. This rendering step is only provided for illustration purpose: highlights convey useful visual cues about the hair geometry. We do not claim any contribution about hair rendering.
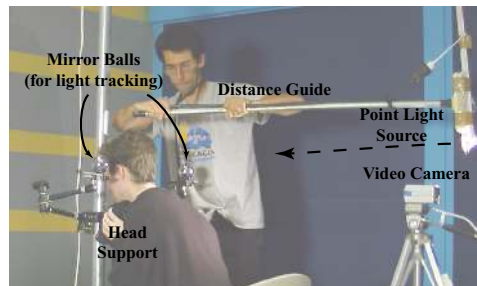

Figure 14: A simple setup to capture a subject under a moving light.

Figure 15: Capture of a black tangled hair. Left: comparison with original view. Right: a view different from the input sequences.

Overall Accuracy: Ideally we would like to compare our captured model with the ground truth which is obviously unknown. A useful consistency check however involves verifying that a rendered image sequence using the light path from the capture setup on our hair model creates similar reflection patterns. So, for each pixel we find the angle in 3D space of the light direction corresponding to the specular peak in the synthesized sequence and compare it to the angle in the original sequence (Fig. 16; we compare an image sequence like Fig. 1). The actual corresponding normal error is half this angular error (from the classical mirror reflection formula). For example, a 5° error at a pixel means that the highlight in the synthetic image occurs at a light position that is 5° different from the real sequence. We find a mean error of 13° and median error of 6.4^o. This difference shows that there are a few large errors. Those are mainly near the silhouette because of hair-to-hair forward scattering which is not rendered. This is confirmed by only considering the front facing region (inside the blue line in Figure 12): the mean and median errors drop down to 7.6^o and 5.0^o, respectively.

This error seems reasonable since it sums the errors from the whole process. Moreover a visual comparison conveys a convincing similarity. Figure 1 shows one such result, notice the highlights at the back and near the top of the head are aligned.
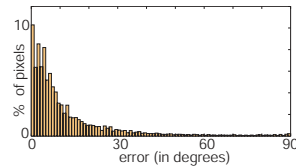

Figure 16: Angular Specular Error Distribution.

Difficulties: For some persons, the skull is poorly approximated by an ellipsoid (it would require a more sophisticated shape such as a super-quadric); for long-haired person, it may also be hard to guess it under the hair. Figure 15 shows an example of this problem; the ellipsoid cannot perfectly match both side and back silhouettes. We have chosen to better approximate the side profile, thus the back profile does not match; however, the hairstyle is still correct.

Acquiring data for long hair requires special care: Hair strands are likely to move when the subject turns for a new viewpoint – especially for the sequence from the top. In this view, the hair silhouette is no longer consistent with the other views; we ignore it for the hull computation. To minimize the potential perturbation, the blending weights of the top view are halved. These slight changes overcome these difficulties with a limited impact on the overall quality. A complete system with several cameras would eliminate this issue.

The hair falling from the top of the head to the top of the face are captured with a lower precision because we do not have a front view. An additional camera should ameliorate this point.

Timing: The acquisition of the 4 sequences take about 5 minutes. Segmenting the hair areas takes 5 to 10 minutes. Running the filters and selecting the lowest variance lasts ≈ 1 hour per view point (the convolutions involve fast Fourier transforms on large domains); we found that using nine images representative of each sequence yields good results. All the other steps (light tracking, segment chaining, etc.) take only a few minutes. (We use an Intel-Xeon 2.4 GHz.)

Types of Hair: Our technique is able to work on a wide range of hairstyles and colors as illustrated by Figures 1, 15 and 17. Large and small curls and waves are accurately captured. Black hair is challenging because the strands are only visible in the highlights. Nonetheless, our method is robust enough to handle it, even with complex small features.

Figure 17: Capture of a long wavy hair. Left: comparison with original view. Right: a view different of the input sequences.

7 Summary and Future Work

We have presented a technique to capture the real geometry of a person's hair from multiple images. This system uses the complex reflectance properties of hair to retrieve its 3D geometry. To take advantage of this light interaction, we introduce a way to compare results from different filters and parameters. The presented variance method links signal processing and statistics to reach precise and robust measures. The theoretical foundation of this approach is studied and shown to be related to other classical methods. We also expose how to exploit the light reflection of hair to extract valuable 3D information.

As a proof-of-concept, we propose a simple practical setup that exploits different image sequences of a real person and show how to blend their results to generate a full hair geometry. The satisfying results reached by this experiment justify the conception of a more sophisticated system. Several cameras can be used at the same time with a light source on a robotic gantry. Such a setup would reach a high level of precision that could open new research directions. For instance, it could be possible to densely measure the appearance properties of the hair to retrieve the parameters of a scattering model. We believe that it would also be possible to acquire such a model from a single sequence. Because this valuable knowledge would permit to work with less information, *e.g.*, less images and light positions, one can conceive the motion capture of hair.

It would also be interesting to adapt this technique to the other methods that manipulate hair. Our poly-lines cannot be directly used for efficient animation and editing. Creating specific data structures from images for these tasks like wisp hierarchy [Kim and Neumann 2002; Bertails et al. 2003] or fluid flow [Hadap and Magnenat-Thalmann 2000] seems a promising direction to explore.

From a theoretical point of view, the study of filters also requires further attention. We have presented and validated some parameters that we use and vary in our filter selection. Future work can look at other parameters and their application to other fields.

References

ANJYO, K., USAMI, Y., AND KURIHARA, T. 1992. A simple method for extracting the natural beauty of hair. In *Proc. of SIGGRAPH*, ACM.

BAKER, S., AND NAYAR, S. 1999. Global measures of coherence for edge detector evaluation. In *Conference on Computer Vision and Pattern Recognition*, vol. 2.

BERTAILS, F., KIM, T.-Y., CANI, M.-P., AND NEUMANN, U. 2003. Adaptive wisp tree. In *Proc. of Symposium on Computer Animation*.

BROOKS, M. J., AND HORN, B. K. P. 1989. Shape and source from shading. In *Shape from Shading*, B. K. P. Horn and M. J. Brooks, Eds. MIT Press.

CANNY, J. 1983. *Finding Edges and Lines in Images*. Master's thesis, MIT.

DALDEGAN, A., MAGNENAT-THALMANN, N., KURIHARA, T., AND THALMANN, D. 1993. An integrated system for modeling, animating and rendering hair. *Computer Graphics Forum 12*, 3, 211–221.

DERICHE, R. 1987. Using Canny's criteria to derive a recursively implemented optimal edge detector. *The International Journal of Computer Vision 1* (May).

DONOHO, D., AND HUO, X. 2000. Beamlet pyramids. In *Proc. of SPIE conference*.

DURAND, F., AND DORSEY, J. 2002. Fast bilateral filtering for the display of high-dynamic-range images. In *Proc. of SIGGRAPH*, ACM.

ELAD, M. 2002. On the bilateral filter and ways to improve it. *IEEE Trans. on Image Processing 11*, 10 (October), 1141–1151.

FEICHTINGER, H. G., AND STROHMER, T., Eds. 2003. *Advances in Gabor Analysis*. Birkhauser.

FELSBERG, M., AND SOMMER, G. 2000. A new extension of linear signal processing for estimating local properties and detecting features. In *Proc. of DAGM Symposium Mustererkennung*.

FORBES, L. A., AND DRAPER, B. A. 2000. Inconsistencies in edge detector evaluation. In *Conference on Computer Vision and Pattern Recognition*.

FREEMAN, W. T., AND ADELSON, E. H. 1991. The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence 13*, 9, 891–906.

GRABLI, S., SILLION, F., MARSCHNER, S. R., AND LENGYEL, J. E. 2002. Image-based hair capture by inverse lighting. In *Proc. Graphics Interface*, 51–58.

GRANLUND, G. H., AND KNUTSSON, H. 1995. *Signal Processing for Computer Vision*. Kluwer Academic Publishers.

HADAP, S., AND MAGNENAT-THALMANN, N. 2000. Interactive hair styler based on fluid flow. In *Proc. of Workshop on Computer Animation and Simulation*, Eurographics.

HERTZMANN, A., AND SEITZ, S. 2003. Shape and materials by example: A photometric stereo approach. In *Proc. of Conference on Computer Vision and Pattern Recognition*, 576–584.

ISHIKAWA, H. 2000. *Global Optimization Using Embedded Graphs*. PhD thesis, New York University.

KAJIYA, J. T., AND KAY, T. L. 1989. Rendering fur with three dimensional textures. In *Proc. of SIGGRAPH*, ACM.

KIM, T.-Y., AND NEUMANN, U. 2001. Opacity shadow maps. In *Proc. of Rendering Techniques conf.*, Springer, 177–182.

KIM, T.-Y., AND NEUMANN, U. 2002. Interactive multi-resolution hair modeling and editing. In *Proc. of SIGGRAPH conference*, ACM.

LAURENTINI, A. 1994. The visual hull concept for silhouette-based image understanding. *IEEE Trans. on Pattern Analysis and Machine Intelligence 16*, 2.

LU, R., KOENDERINK, J. J., AND KAPPERS, A. M. 1999. Specularities on surfaces with tangential hairs or grooves. In *Proc. of the International Conference on Computer Vision*, IEEE, 839–846.

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

MATUSIK, W., PFISTER, H., ZIEGLER, R., NGAN, A., AND MCMILLAN, L. 2002. Acquisition and rendering of transparent and refractive objects. In *Proc. of the Eurographics Workshop on Rendering*.

MEER, P., AND GEORGESCU, B. 2001. Edge detection with embedded confidence. *IEEE Trans. on Pattern Analysis and Machine Intelligence 23*, 12.

NAKAJIMA, M., MING, K. W., AND TAKASHI, H. 1998. Generation of 3d hair model from multiple pictures. In *IEEE Computer Graphics & Applications (12) 1998*, 183–169.

PLANTE, E., CANI, M.-P., AND POULIN, P. 2001. A layered wisp model for simulating interactions inside long hair. In *Proc. of Computer Animation and Simulation*, Eurographics.

RUSHMEIER, H. E. 2001. 3D capture for computer graphics. In *Third International Conference on 3D Digital Imaging and Modeling*.

SHEN, J., AND CASTAN, S. 1986. An optimal linear operator for edge detection. In *Proc. of Conference on Computer Vision and Pattern Recognition*, IEEE.

TOMASI, C., AND MANDUCHI, R. 1998. Bilateral filtering for gray and color images. In *Proc. of the International Conference on Computer Vision*, IEEE, 839–846.

TSCHUMPERLÉ, D., AND DERICHE, R. 2002. Orthonormal vector sets regularization with pde's and applications. *International Journal on Computer Vision 50* (12), 237–252.

WATSON, G. S. 1983. *Statistics on spheres*. John Wiley and Sons.

YANG, R., M., P., AND WELCH, G. 2003. Dealing with textureless regions and specular highlights. In *Proc. of the International Conference on Computer Vision*, IEEE.

YITZHAKY, Y., AND PELI, E. 2003. A method for objective edge detection, evaluation and detector parameter selection. *IEEE Trans. on Pattern Analysis and Machine Intelligence 25*, 8.

ZIOU, D., AND TABBONE, S. 1998. Edge detection techniques - an overview. *International Journal of Pattern Recognition and Image Analysis 8*, 537–559.