

Single View Facial Hair 3D Reconstruction

Gemma Rotger¹, Francesc Moreno-Noguer², Felipe Lumbreras¹, and Antonio Agudo²

¹ Computer Vision Center and Departament Ciències de la Computació, UAB, Spain

² Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain

Abstract. In this work, we introduce a novel energy-based framework that addresses the challenging problem of 3D reconstruction of facial hair from a single RGB image. To this end, we identify hair pixels over the image via texture analysis and then determine individual hair fibers that are modeled by means of a parametric hair model based on 3D helices. We propose to minimize an energy composed of several terms, in order to adapt the hair parameters that better fit the image detections. The final hairs respond to the resulting fibers after a post-processing step where we encourage further realism. The resulting approach generates realistic facial hair fibers from solely an RGB image without assuming any training data nor user interaction. We provide an experimental evaluation on real-world pictures where several facial hair styles and image conditions are observed, showing consistent results and establishing a comparison with respect to competing approaches.

Keywords: 3D Vision, Shape Reconstruction, Facial Hair Modeling.

1 Introduction

In the last years, computer-generated characters are provided of an increasing level of realism, heightening the standards of the film and video game industry. Hair plays an essential role to build convincing digital humans and fulfill them with an identity. Pessig *et al.* [17] advised that the eyebrows alone may be the most relevant feature for facial recognition. Unfortunately, it is well known that retrieve hair geometry solely from images is a challenging task due to its structural complexity. Although several image-based methods are capable to create high-quality reconstructions [4, 13–16, 18], they require specialized and expensive hardware. Furthermore, except [4] none of them proved any accomplishment over facial hair.

Recently, the earliest models handling hair reconstruction from a single RGB image have emerged. The first approaches [5, 7] required several user interactions. Particularly, in [5] was required depth information and a user to annotate the hair, as well as head-region annotations in the image. In [7] was expected the user to define sparse strokes in order to determine the local direction ambiguity. It dropped to new data-driven solutions [11], which reduced the required volume of user interaction but not omitted. Later, remarkable results have been achieved with the use of structured light patterns [8] and electro-luminescent wires [9]. Unfortunately, no achievement was reported over facial hair. Similarly, deep-learning approaches contributed to the field as hair growing direction estimation [6] and hair-style parametrization [21] amongst others. However, these

approaches tended to be highly data demanding, and active methods required expensive hardware setups.

In this paper, we present a novel optimization framework that uses the texture information in an RGB image to predict the facial hair fibers geometry in 3D. We take advantage of orientation analysis texture methods [15] to detect the fibers. Later, we track across this detection the individual fibers at pixel-level, splitting connections when their orientation significantly varies. Similar to [4], we allow hair crossings to be detected by connecting fibers which are close in the space and have a similar overall orientation. Afterward, we parametrize our hair generation model by minimizing a set of four different energies that take into account different 2D detection properties. To improve the computation efficiency, we arranged the detected hair fibers in different groups according to their 2D properties (these are position, length, and orientation), obtaining a different parametrization per group. More specifically, eyelashes and eyebrows require to be studied separately due to the evident differences regarding beard and mustache fibers. Finally, we enhanced more realism by adding hair density and small random noise.

Our main contribution is to propose a model for growing 3D hair fibers together with a 3D face model from a single RGB image, providing hair fibers with high resolution. To this end, no further information, user interaction, or training data is required. We demonstrate the effectiveness of our method over a wide variety of facial hair styles and geometries. Our approach is extensively evaluated on real high-quality RGB images from the Internet, proving the suitability of our framework to reconstruct plausible 3D facial hairs directly from pictures.

2 Related Work

The 3D reconstruction of realistic faces has prevailed a topic of interest [1, 2, 12]. Concerning the recovering of human hair fibers as an essential step to realism and verisimilitude, several methods addressed the problem from different perspectives [4–6, 8, 9, 13, 15, 18, 21]. While there exist a wide range of methods reconstructing a hairstyle as a parametric model [10, 11, 14, 19, 20], unfortunately, the area of research studying hair reconstruction fiber-by-fiber from a single view is shorter.

In particular, fiber-by-fiber was tackled from the perspective of multi-view stereo. The work of [15] presented a fiber-by-fiber reconstruction approach by growing hair within the restrictions established by the visual hull. In a similar practice, in [18] was synthesized the hair fibers from local image orientations. In [13] was presented a new setup to capture hair arrangement fiber-by-fiber, and in [4] was proposed the first approach to reconstruct facial hair fibers. Lately, active-light methods established their solutions in the field. In [8] was proposed a novel robust strip-edge-based coding method, based on a projection pattern. In [9] was introduced a novel method for braid acquisition and 3D guide hair reconstruction based on electro-luminescent wires.

Deep-learning methods have contributed to the topic providing solutions from a single RGB image. For instance, in [5] was introduced a 3D helical hair prior that captures the geometrical structure of the hair from a single image, but it required manual hair segmentation and direction guidance. In [6] user interactions were overcome with a novel hierarchical deep-neural network for automatic hair segmentation and hair growth direction estimation. Additionally, in [21] was presented a convolutional neu-

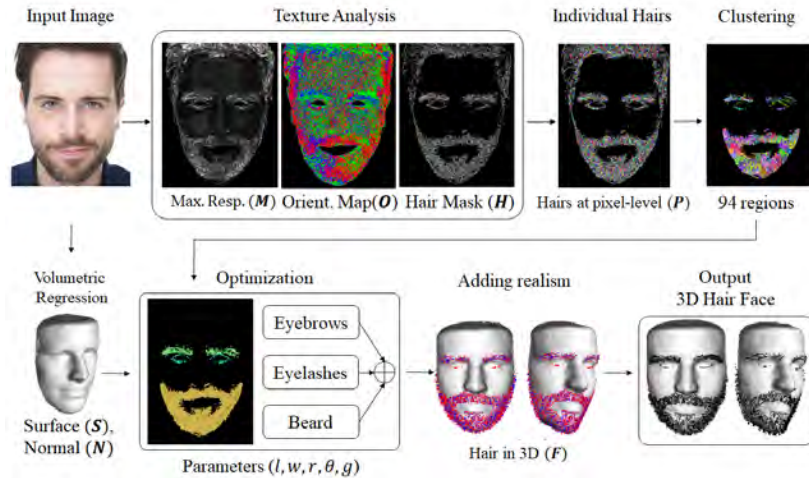


Fig. 1. An overview of our pipeline. From a single image, our approach first retrieves a 3D facial model (coded by N and S) by applying a volumetric regression CNN approach [12]. Later, a hair map is detected over the image via Gabor texture analysis, where some attributes are obtained: the maximum response and the maximum orientation response (every orientation is represented by a different color) are denoted as M and O , respectively, obtaining the final hair detection in the binary matrix H . Next, we trace individual hair fibers via pixel-connectivity, orientation differences, and endpoints distance in P . As different areas (beard and mustache, eyebrows, and eyelashes), need a different parametrization due to their variability, hair fibers are grouped in 94 different regions according to their location, orientation, and 2D length, which are included in one of the previous macro-classes. For these macro-classes, model parameters are estimated by optimization. Eventually, we append the results and add small random variations in the hair length and orientations as well as density to increase the realism. Red and blue lines represent the estimated and computed hairs, respectively.

ral network that adopts the 2D orientation fields of the image as input and generates a strand feature. The resulting parametrizations allowed interpolating between several hairstyles. However, deep-learning approaches do not handle fiber-by-fiber retrieving, recalling these methods ineffective for accurate facial hair reconstruction. Further, they are demanding in data terms. On the other side of multi-view and active-light based methods, they require expensive or specific hardware and controlled illumination, reducing its applicability in real-world scenarios. Besides, most of the previous works had not demonstrated its suitability to reconstruct facial hair fibers but head hair fibers or hairstyles. We find in [4] an interesting exception. It consists of a coupled hair and skin multi-view stereo method based on images acquired in a controlled studio. In this paper, we present a similar concept but using a single-view image under uncontrolled lighting. Our method can handle hair fibers estimation from a good quality RGB image without requiring training data or other setups.

3 Our Approach

This section describes the computation of the 3D hair strands from a single RGB image, and the final post-processing step to add further realism to the reconstruction.

First, our approach detects hair fibers in the RGB image via Gabor texture analysis. After that, based on an orientation analysis, the full facial hair is outlined from its root to

the tip at pixel-level and group them in a cluster of hairs with similar properties to speed up the optimization process. In combination with the estimated 3D facial model [12], we solve a set of optimizations to find the parameters that better model each cluster of hair. Finally, we append all the computed hairs and provide further realism by adding small random variations in length and root orientation. Figure 1 depicts a schematic of the overall approach, and how every part is connected. In the remaining of this section, each step is illustrated in detail.

3.1 2D Hair Detection

Texture Analysis. Our first observation is that human hair may vary in a considerable range of tones and shapes, making the problem harder to address. Therefore, we notice that the RGB space is not suitable to achieve a proper texture analysis. For this reason, we convert the initial image to the HSV color space and only consider the saturation and value channels to exploit the greater uniformity of the hair fibers, together with a significant difference with regards to the skin pixels. The hue channel is employed to identify the inner part of the eyes and the mouth where the texture is not analyzed. It is well known that these regions do not hold hair roots. However, they can contain other features that can be confused with hair fibers in the texture analysis such as veins, or strong lip textures. For this, and similar to [4], we consider convenient not to analyze the texture on these specific areas.

As it can be analyzed in [4, 15], orientation responses are suitable to estimate hair in images. Formally, they use a filter kernel K_θ for different θ orientations, at every 10 degrees, and keep the orientation that produces the largest score in the function $F(x, y) = |K_\theta * V|_{(x,y)} + |K_\theta * S|_{(x,y)}$, at a pixel (x, y) , for the value V and saturation S channels. In a similar manner, we apply the real part of a Gabor filter bank consisting of 5 different wavelengths $\lambda = \{2, 2.5, 3, 3.5, 4\}$ and 18 orientations θ (from 0 to 170, at each 10 degrees). For each pixel, we hold the maximum response of the filter $M(x, y)$ and the orientation of the maximum response $O(x, y)$ which later will allow us to detect individual hair fibers. Both are defined as:

$$M(x, y) = \max(F(x, y)), \quad (1)$$

$$O(x, y) = \theta_{\max(F(x, y))}. \quad (2)$$

Next, we binarize the maximum response by applying a simple threshold τ to exclude low-confidence responses. To this end, we define a binary hair mask H as:

$$H(x, y) = M(x, y) > \tau. \quad (3)$$

Individual hair trace. Let us define a hair at a pixel level as P^h , and the set of all the pixels in a hair as $\mathbf{p}_i^h \in P^h$ with $\mathbf{p}_i^h = (x_i^h, y_i^h)$. The goal in this stage is to transform the different blobs in H into an ordered set of pixels $\mathbf{p}_i^h \in P^h$ where \mathbf{p}_0^h is the hair root and \mathbf{p}_L^h the hair tip, where L denotes the length of the fiber.

We manage the hair map H regard to 8-connected region analysis techniques to find connected pixels. However, due to hair constitution, crossings, and strong shading, different samples can be grouped as a single detection. To determine a single hair fiber trace, we study the orientation map O and ensure all the pixels in the same connected region has an orientation difference with respect to the following pixel is smaller than

10 degrees. When this does happen, i.e., $|\mathbf{O}(\mathbf{p}_i^h) - \mathbf{O}(\mathbf{p}_{i+1}^h)| > 10$, we detach the connection.

The previous step results in small regions equivalent to visible hair fibers in the image. However, due to shadows and intersections, hair can split into several detections. In [4] was pointed that hair fibers can be re-joined in 3D if they satisfy three conditions: 1) their endpoints are unconnected, 2) they are close in the space (or overlapped), and 3) the orientation variation is lower than 20 degrees. We adopt a similar setup in 2D, limiting the angle to 10 degrees to be consistent with our previous step. A maximum distance limit for re-connect the endpoints equal to three pixels. For all combinations of hairs, we accept those satisfying the previous conditions. We additionally allow more than two hairs to join in the same section if the overall of the segments fulfills the orientation restriction and their endpoints are allowed to connect in a two-by-two association. The relevance of this step is imperative since the estimation of the resulting 3D hair properties is directly related to the measures extracted from visible 2D hair segments.

Endpoint labeling. In this step, we determine if a hair fiber has its endpoints (root and tip) in the proper order. This step is essential since the hair model presented in the following sections is implemented over the root position and grows unto the tip location. We employ certain facial landmarks extracted with [3]. These landmarks include the nose tip for beard, mustache, and eyebrows, and the averages of the eye landmarks as the central eye-points. We define the growth direction determining the root as the endpoint with the minimum Euclidean distance value with respect to the corresponding landmark, and the tip as the endpoint with the maximum value, such as:

$$\begin{aligned} \mathbf{p}_{root}^h &= \min(d(\mathbf{p}_i^h, (x_j^l, y_j^l))), \\ \mathbf{p}_{tip}^h &= \max(d(\mathbf{p}_i^h, (x_j^l, y_j^l))), \end{aligned} \quad (4)$$

where (x_j^l, y_j^l) denotes the landmarks. In case the endpoints are reversed, we shift the entire \mathbf{P}^h values to satisfy the new endpoint labeling.

3.2 A simple hair growing parametric model

When an incipient hair shaft grows, it follows the normal surface direction. However, after reaching a determined length, its trajectory is affected by other factors, for instance, the shaft weight, the gravity effect, and the follicle cross-section, which determines the shaft thickness and the curliness. The length of the shaft also affects the curvature in a two-dimensional plane, since the longer the hair becomes, the more burden the tip of the hair shaft supports, so the greatest is the gravity effect over the fiber. However, the curliness cannot be represented as deformation in a two-dimensional plane but as a 3D local helix.

To overcome all these possible variations, we have determined a hair model that ensures local coherence and smoothness. Further, it represents both the curliness as a 3D local helix and a gravity-like effect. Our model has five different parameters that control the hair growing conditions and provide hair fiber-like results. These parameters define the size (length l , and width w), the curliness (radius r , angle θ) and a gravity-like effect (g) that avoids shafts to lie suspended in the air. s is the last parameter of the model, which determines the resolution of the generated hairs. It is defined in advance by the

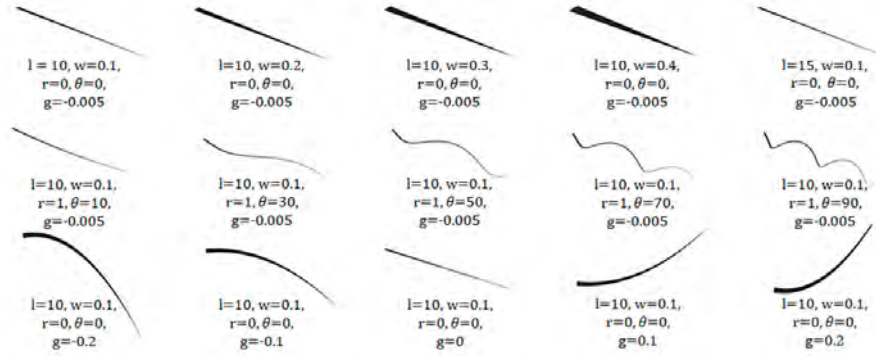


Fig. 2. Parametric hair model. Our parametric model depends on five parameters: length l , width w , the curliness parameters r and θ , and the gravity coefficient g . As it can be seen in the figure, thanks to our model we can obtain a wide variety of fibers. **Top:** Some instances varying l and w . **Middle:** Some instances as a function of the curliness parameters. **Bottom:** Modifying the gravity-like parameter.

user according to the requirements of the desired solution. The greater the value, the larger the resolution, and the slower to perform all the computations and visualizations. In our experiments, we adopted a value of $s = 25$ for all the estimations.

Let us consider $F^h(\mathbf{p}^h, \mathbf{n}^h, l, w, r, \theta, g)$ the parametrization of a hair fiber, where \mathbf{p}^h denotes the 3D position and \mathbf{n}^h the 3D growing direction. In Fig. 2, we show several examples of how our model acts as a function of the parameters. Since the image is aligned with the initial 3D, the detections over the image are likewise aligned. For each detected root \mathbf{p}_0^h we can interpolate the three closets mesh vertices and obtain \mathbf{p}^h as the average of the positions and \mathbf{n}^h as the average of the vertex normals. With these considerations and the parameters explained above we can define our facial hair model as:

$$F_i^h(\mathbf{p}^h, \mathbf{n}^h, l, w, r, \theta, g) = \mathbf{p}_0^h + \mathbf{R}(\mathbf{n}^h) \frac{(i-1) \cdot l}{s} Hx(r, \theta, i) - Gy(g, i), \quad (5)$$

where i denotes the i -th point in the hair fiber, $\mathbf{R}(\mathbf{n}^h)$ represents the rotation matrix, which adjusts the 3D helix's direction with the corresponding normal vector at the given position.

The 3D helix is defined over the x -axis as follows:

$$Hx(r, \theta, i) = (i-1, r \cdot \sin(\theta \cdot (i-1)), r \cdot \cos(\theta \cdot (i-1))), \quad (6)$$

where $Hx(r, \theta, i)$ represents the function producing the 3D coordinates of the 3D helix at the i -th point. r is the radius and θ the angle between two consecutive points.

The gravity-like effect is computed as follows:

$$Gy(g, i) = \mathbf{v} \cdot (i-1) \cdot (\cos(\alpha), -\frac{g(i-1)^2}{2}, \sin(\alpha)), \quad (7)$$

where $Gy(g, i)$ represents the effect of a gravity-like function at the i -th point of the helix. After several tests, we define $\mathbf{v} = [\frac{1}{\alpha}, 1, \frac{1}{\alpha}]$, being $\alpha = 10$ degrees as are the values that adapt better the hair fibers behavior.

Finally, we compose a cylinder over each segment of \mathbf{F}^h . Where the initial width in \mathbf{F}_0^h is defined beside the parameter w , and it decreases consistently unto the last point \mathbf{F}_l^h where the cylinder width is $w = 0$.

3.3 3D growing via energy minimization

Hair geometry is locally consistent, yet it has different properties according to face regions. For instance, eyebrows are not similar to beard fibers, but, each is similar within the same class. To consider these different behaviors, we have defined three different optimization groups. The first corresponds to beard and mustache and clusters a total of 70 groups with similar length, position, and orientation via K -means. The second group distinguishes the hairs belonging to the subject eyelashes, in total four clusters corresponding to each of the eye lines upper and lower of both eyes. The third group corresponds to the eyebrows, and it has ten clusters per eyebrow, likewise arranged with K -means. We optimize each of the 94 clusters separately to satisfy the following energy minimization:

$$\mathcal{E}_{total} = \mathcal{E}_{len} + \mathcal{E}_{ori} + \mathcal{E}_{tip} + E_{cur}. \quad (8)$$

Length Term \mathcal{E}_{len} . The length energy term is a direct segment longitude comparison. We use the sum of Euclidean distances from the root pixel detection \mathbf{p}_0^h to the tip pixel detection \mathbf{p}_l^h and compare it upon the same distance in the xy -plane for the estimated hair root \mathbf{f}_0^h and tip \mathbf{f}_l^h .

$$\mathcal{E}_{len} = \sum_h \left\| (\mathbf{p}_l^h - \mathbf{p}_0^h) - (\mathbf{f}_l^h - \mathbf{f}_0^h) \right\|_2^2. \quad (9)$$

Orientation Term \mathcal{E}_{ori} . It encourages hair fibers to have similar orientations than the detected in the given image. In practice, the global 3D orientation is determined by the surface normal, though, the gravity-like parameter can force the fiber to grow in a different orientation.

$$\mathcal{E}_{ori} = \sum_h \left\| \tan^{-1} \left(\frac{p_{ly}^h - p_{0y}^h}{p_{lx}^h - p_{0x}^h} \right) - \tan^{-1} \left(\frac{f_{ly}^h - f_{0y}^h}{f_{lx}^h - f_{0x}^h} \right) \right\|_2^2. \quad (10)$$

Tip-to-tip Term \mathcal{E}_{tip} . Tip-to-tip cost encourages hairs fibers to have the tip projection on the 2D plane close to the tip detection in the image.

$$\mathcal{E}_{tip} = \sum_h \left\| (\mathbf{p}_l^h - \mathbf{f}_l^h) \right\|_2^2. \quad (11)$$

Curviness Term \mathcal{E}_{cur} . It limits the hair to coincide barely with root and tip but fail in the remaining pixels. It computes the perpendicular distance from each fiber point to the closest point in the root-tip segment and compares the equivalent procedure with the detected hair on the image.

$$\mathcal{E}_{cur} = \sum_h \sum_i \left\| \frac{(\mathbf{p}_l^h - \mathbf{p}_0^h) - (\mathbf{p}_0^h - \mathbf{p}_i^h)}{(\mathbf{p}_l^h - \mathbf{p}_0^h)} - \frac{(\mathbf{f}_l^h - \mathbf{f}_0^h) - (\mathbf{f}_0^h - \mathbf{f}_i^h)}{(\mathbf{f}_l^h - \mathbf{f}_0^h)} \right\|_2^2. \quad (12)$$

Optimization. We solve Eq. (8) over the different groups in parallel, by using the non-linear least squares. Toward the eyelash optimization, the procedure is equivalent though we force the detected roots to move along the spline formed by the eyelid landmarks. While the eyebrows optimization differs from the rest since we force the hair fibers to grow in the tangent direction instead of the normal one.

3.4 Adding density and further realism

We found that estimating the hair parameters with large clusters even when they have related properties we lose significant realism. To address this issue, we propose a post-processing step which involves two actions: adding density to the retrieved result and adding further realism by appending random small variations to the final 3D hairs.

Adding density. When we explore the individual hair fibers in 2D, we may discard hairs with severe occlusions and reject sections with poor connections. In this post-processing step, we aim to estimate the missed detections as a combination of the closest visible hairs. We compare the binary mask generated by the orthographic projection of the computed hair elements $\pi(\mathbf{F})$, with the initial hair map \mathbf{H} . To avoid false positives in the estimation, we expand the mask generated by $\pi(\mathbf{F})$ with morphological operators and estimate new hair only on the resulting region:

$$G = \sum_i \sum_j (\mathbf{H}(i, j) - \pi(\mathbf{F})(i, j)) \cdot (\pi(\mathbf{F}) \oplus \mathbf{D})(i, j), \quad (13)$$

where G represents the number of available pixels to grow a new hair, \mathbf{D} is a binary 5×5 dilation mask, and \oplus represents the binary dilation operator. G is a positive integer if the further density is required, and a negative integer if we added more than necessary hair fibers. If G is equal to zero, it implies that there is the exact amount of pixels in the hair map than projections of \mathbf{F} .

For each new hair, we grow a new hair fiber \mathbf{F}^k as a combination of the three closest hairs. This is achieved by averaging the parameters of the hair strand or the equivalent, averaging all the j -th points in $\mathbf{F}_j^k = \frac{1}{3} \sum_{i=1}^3 \mathbf{F}_j^i$ where i denotes each of the three nearest neighbors. The process is iterated until G is equal or slightly lower to zero.

Adding small random variations. The estimation of a group of hairs with similar parametrization leads to homogeneous hair in small areas and consequently lack of realism. To overcome this situation, we add small random noise to the resultant length $\lambda_l \in [-0.05, 0.05]$ and a small random rotation amongst all the axes $\lambda_r \in [-1, 1]$, where $\dim(\lambda_r) = 3$. Both variations λ_r and λ_l are summed to the fiber estimated parameters to generate a hair shaft such as:

$$\mathbf{F}^h(\mathbf{p}^h, \mathbf{n}^h + \lambda_r, l + \lambda_l, w, r, \theta, g), \quad (14)$$

where λ_r adjusts the orientation given by \mathbf{n}^h and λ_l the fiber length given by l .

4 Experimental evaluation

We now present our experimental results for different types of pictures, including several hairstyles for both genders, obtained from Pexels platform. Additionally, we also present a qualitative comparison with respect to [4].

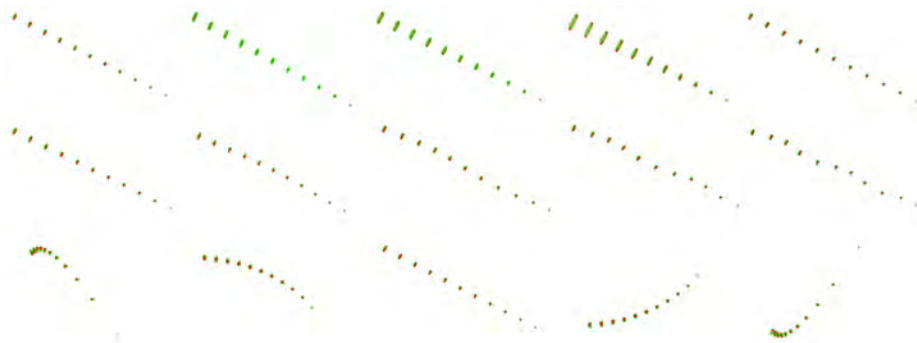


Fig. 3. Qualitative evaluation of several hair fibers. We depict the average amongst all the hairs in a single hair. Green circles represent the ground truth, and red dots our estimation. Best viewed in color.

(row,col)	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	Average
error 3D	0.027	0.004	0.006	0.014	0.017	0.014	0.013	0.017	0.015	0.013	0.002	0.002	0.006	0.003	0.008	0.011
time (s)	40.443	40.936	50.613	51.977	40.996	52.167	52.873	62.767	51.463	51.853	40.443	59.267	61.289	61.311	64.152	52.170

Table 1. Quantitative evaluation of several hair fibers and time budget. 3D errors of the hair fibers depicted in Fig. 3, and the corresponding computation time in seconds. The last column represents the average error for all the points in the 150 hairs.

First, we had evaluated our method quantitatively with synthetic samples. To this end, we have generated a total of 150 synthetic hairs over half sphere and recovered the parameters from their 2D coordinates. Our goal is to reconstruct the 3D hairs with those parameters and obtain the 3D error between points in the initial and the reconstructed hairs. For completeness, we have used the configurations that were described in Fig. 2. Table 1 reports the 3D errors as an average of the 3D error amongst all the points. Figure 3 depicts the average hair, displaying where these errors are located. As it can be observed, in both cases our method presents very competing results and provides accurate hair reconstructions.

To show the effectiveness of our approach on real scenarios, we first consider from short to full beards and mustaches. Some results are displayed on Fig. 4, where we can observe how our method produces realistic solutions on challenging scenarios with short and long beards, mustache, eyebrows, and eyelashes. Further, it works satisfactorily with partial occlusions, as self-occluding beards (see subjects (2,1) and (3,1) on the previous figure). We also report some numbers regarding these experiments in Table 2, where the number of hair fibers is included. As it can be observed, our approach is available to recover a large number of hair fibers in different areas.

In addition, we also consider some cases over subjects with eyebrows and eyelashes, including a challenging case concerning a subject with eyeglasses. Particularly, the subject (5,2) represents a challenging scenario due to the poor texture produced by makeup. Fortunately, our approach is able to detect small pieces of eyebrow hair instead of full hairs owing to the large image resolution and quality. Another example, it is the subject (6,1), Frida Kahlo, where our approach is evaluated for a low-resolution picture. As in the previous case, our approach also produces a visually realistic solution. Again,



Fig. 4. Face reconstruction with different facial hair styles. In both sides, we represent the same information. **First column:** Input image. **Second and third columns:** Frontal and side views of our 3D hair+face reconstruction over a textured face. The hair fibers are represented by red lines. **Fourth and fifth columns:** Frontal and side views of our estimated geometry, without considering any texture. **Sixth and seventh columns:** Just observing our hair estimation.

(row,col)	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
im res.	849×1273	1024×768	1200×825	1600×2400	4000×6000	393×588	1265×1920	845×650	1760×2640	1750×1168	3999×3999	5075×5760
hairs up	766(61)	62(2)	72(6)	181(19)	804(6)	168(13)	551(63)	609(55)	66(9)	334(67)	7132(238)	162(7)
hairs lo	10251(510)	9575(155)	17874(408)	4647(41)	0(0)	0(0)	6250(212)	7332(746)	7733(440)	9384(361)	0(0)	0(0)

Table 2. Number of reconstructed hair fibers for pictures on Fig. 4. We report for every picture on Fig. 4, its resolution, and the number of retrieved hairs on the upper/lower parts of the face, showing in parenthesis the added hairs in the post-processing step. To this end, we consider the location of every picture on the figure, indicating its row and column position.

some numbers about these experiments are reported in Table 2. In Fig. 5 we show some detailed close-ups, where it can be observed the realism we achieve with our approach.

In all cases, we use un-optimized Matlab code on an Intel(R) Xenon(R) CPU ES-1620 v3 at 3.506GHz. The full pipeline run-time depends on the amount and complexity of the hairs to recover. It takes from 15 minutes to recover examples with upper hair only to 10 hours to recover the subject (3,1) in Fig. 4, where the hair density is extremely large.

Finally, we also establish a qualitative comparison with respect to [4]. In this case, it is worth mentioning that while our approach only needs an RGB image under general and uncontrolled lighting conditions as input, [4] requires a calibrated multi-camera system. Despite this disadvantage in terms of hardware resources, our approach obtains competing results (see Fig. 6 for a qualitative comparison). As it can be observed, our method is effective in locating the hair fibers and optimizing their parametrization.

Discussion: Regardless of the surpassing performance in terms of single image facial hair fiber capture, our method works better in the presence of short and scattered hairs on high-resolution pictures. Similarly, we find helpful clear and noiseless textures with high contrast between skin and face. Although our method can reconstruct like-



Fig. 5. Close-up results. Some close-ups of detailed instances are displayed. **First and second column:** eyelashes and a piece of beard around the mouth are represented for the subject (1,1) on Fig. 4. **Third column:** unveils the thick eyebrow of Frida Kahlo, subject (6,1). **Fourth column:** represents a man’s mustache, picture (4,1). In all cases, we can observe how the hair fibers are successfully recovered, and they are visually coherent.



Fig. 6. Qualitative comparison on 3D hair+face reconstruction. **First column:** input RGB image for our approach. It is worth noting that the solution in [4] requires 14 cameras along with 4 flashes, i.e., a very constrained calibration is demanded. **Second and third column:** frontal and side views using [4]. **Fourth and fifth column:** our solution.

wise facial hairs when the previous scenarios are not favorable, we find the hair fibers with big orientation changes are difficult to be recovered (see Dali’s example in Fig. 4), since our approach is not able to fully trace individual hairs with angles larger than 10 degrees. Moreover, the hair should be evident at pixel-level in order to be detected.

5 Conclusion

In this paper, we have proposed a framework that successfully recovers 3D facial hair from a single RGB image without any training data. To this end, we proposed a facial hair parametric model based on 3D helixes and a set of energies which rely on 2D hair detections over different face areas to estimate the parameters directly from the image. Furthermore, it does not require any training data, user interaction, or any specific setup. We have extensively validated our approach over a collection of several images with uncontrolled illumination and show consistent and realistic results even in challenging cases as thick beards, eyeglasses, low-resolution pictures, and eyebrow makeup. Further, we compare our approach with the current state of the art, where our procedure retrieves competing results despite the clear disadvantage in terms of hardware and single-image versus multiview setups. Facial hair is an essential step in the reconstruction of realistic faces. For this reason, future research lines are to join different procedures to retrieve various aspects of detail in human faces without delimiting the face area but the full head structure including the neck, which is a plausible spot for men to have hair.

Acknowledgments: This work has been partially supported by the Spanish Ministry of Science and Innovation under projects FireDMMI TIN2014-56919-C3-2-R, BOSS

TIN2017-89723-P, and HuMoUR TIN2017-90086-R; by the Spanish State Research Agency through the María de Maeztu Seal of Excellence to IRI MDM-2016-0656, and by the CSIC project R3OBJ 201850I099.

References

1. A. Agudo and F. Moreno-Noguer. Combining local-physical and global-statistical models for sequential deformable shape from motion. *IJCV*, 122(2):371–387, 2017.
2. A. Agudo and F. Moreno-Noguer. Force-based representation for non-rigid shape and elastic model estimation. *TPAMI*, 40(9):2137–2150, 2018.
3. B. Amos, B. Ludwiczuk, M. Satyanarayanan, et al. Openface: A general-purpose face recognition library with mobile applications. *CMUSCS 2016*.
4. T. Beeler, B. Bickel, G. Noris, P. Beardsley, S. Marschner, R. W. Sumner, and M. Gross. Coupled 3D reconstruction of sparse facial hair and skin. *TOG*, 31(4):117, 2012.
5. M. Chai, L. Luo, K. Sunkavalli, N. Carr, S. Hadap, and K. Zhou. High-quality hair modeling from a single portrait photo. *TOG*, 34(6):204, 2015.
6. M. Chai, T. Shao, H. Wu, Y. Weng, and K. Zhou. AutoHair: fully automatic hair modeling from a single image. *TOG*, 35(4), 2016.
7. M. Chai, L. Wang, Y. Weng, X. Jin, and K. Zhou. Dynamic hair manipulation in images and videos. *TOG*, 32(4):75, 2013.
8. Y. Chen, Z. Song, S. Lin, R. R. Martin, and Z.-Q. Cheng. Capture of hair geometry using white structured light. *Computer-Aided Design*, 96:31–41, 2018.
9. H. Hachmann, M. Awiszus, and B. Rosenhahn. 3D braid guide hair reconstruction using electroluminescent wires. *The Visual Computer*, 34:793–804, 2018.
10. L. Hu, C. Ma, L. Luo, and H. Li. Robust hair capture using simulated examples. *TOG*, 33(4):126, 2014.
11. L. Hu, C. Ma, L. Luo, and H. Li. Single-view hair modeling using a hairstyle database. *TOG*, 34(4):125, 2015.
12. A. S. Jackson, A. Bulat, V. Argyriou, and G. Tzimiropoulos. Large pose 3D face reconstruction from a single image via direct volumetric CNN regression. In *ICCV*, 2017.
13. W. Jakob, J. T. Moon, and S. Marschner. Capturing hair assemblies fiber by fiber. *TOG*, 28(5):164, 2009.
14. L. Luo, H. Li, and S. Rusinkiewicz. Structure-aware hair capture. *TOG*, 32(4):76, 2013.
15. S. Paris, H. M. Briceño, and F. X. Sillion. Capture of hair geometry from multiple images. *TOG*, 23(3):712–719, 2004.
16. S. Paris, W. Chang, O. I. Kozhushnyan, W. Jarosz, W. Matusik, M. Zwicker, and F. Durand. Hair photobooth: geometric and photometric acquisition of real hairstyles. *TOG*, 27(3):30, 2008.
17. J. Peissig, T. Goode, and P. Smith. The role of eyebrows in face recognition: With, without, and different. *Journal of Vision*, 9(8):554–554, 2009.
18. Y. Wei, E. Ofek, L. Quan, and H.-Y. Shum. Modeling hair from multiple views. *TOG*, 24(3):816–820, 2005.
19. X. Yu, Z. Yu, X. Chen, and J. Yu. A hybrid image-CAD based system for modeling realistic hairstyles. In *SIGGRAPH*, 2014.
20. M. Zhang, M. Chai, H. Wu, H. Yang, and K. Zhou. A datadriven approach to four-view image-based hair modeling. *TOG*, 36(4):156, 2017.
21. Y. Zhou, L. Hu, J. Xing, W. Chen, H.-W. Kung, X. Tong, and H. Li. HairNet: Single-view hair reconstruction using convolutional neural networks. In *ECCV*, 2018.