Relatively Straight Medium to Long Hair Reconstruction using Kinect Sensors

Chao Sun, Zhongrui Li and Won-Sook Lee
School of Electrical Engineering and Computer Science, University of Ottawa, ON, Canada
{csun014, zl109, wslee}@uottawa.ca

Keywords: Hair Modeling, Kinect v2 Sensors

Abstract: Most existing hair capturing methods reconstruct 3D hair models from multi-view stereo based on complex capturing systems composed of many digital cameras and light sources. In this paper, we introduce a novel hair capturing system using consumer RGB-D (Kinect sensors). Our capture system, consisting of three Kinect v2 sensors, is much simpler than previous hair capturing systems. We directly use the 3D point clouds captured by Kinect v2 sensors as the hair volume. Then we adopt a fast and robust image enhancement algorithm to adaptively improve the clarity of the hair strands geometry based on the estimated local strands orientation and frequency from the hair images captured by the Kinect colour sensors. In addition, we introduced a hair strand grow-and-connect algorithm to generate relatively complete hair strands. Furthermore, by projecting the 2D hair strands onto the 3D point clouds, we can obtain the corresponding 3D hair strands. The experimental results indicate that our method can generate plausible 3D models for long, relatively straight hair.

1 INTRODUCTION

Recent advances in consumer RGB-D sensors have facilitated real-world object capturing and modeling. Consumer RGB-D sensors can now provide relatively accurate images and 3D point clouds with an easy, efficient, and inexpensive capturing procedure. For example, the Kinect v2 depth sensor provides improved ability of 3D capturing and 3D visualization than the Kinect for Xbox 360. In addition, the Kinect v2 sensor has a 1080p color camera which can capture clear images and videos. Researchers have previously used Kinect depth sensors to capture 3D human body models, (Tong et al., 2012) (Wang et al., 2012) (Li et al., 2013) (Shapiro et al., 2014), however, using Kinect sensors to capture real human hair models has not been explored.

Hair modeling remains one of the most challenging tasks due to the characteristics of hair, such as omnipresent occlusion, specular appearance and complex discontinuities (Ward et al., 2007). For hair modeling, most existing methods utilize multi-view 2D hair images to obtain 3D hair geometry information (Paris et al., 2008) (Luo et al., 2012) (Luo et al., 2013b) (Luo et al., 2013a). However, such methods usually require complex capture systems composed of many digital cameras and light sources and produce a large number of hair images. The large amount of the input hair images makes the reconstruction a time-consuming procedure. In addition, user assistance is needed to a certain extent (for example: to clear the outliers from the reconstruction results). In our proposed method, our capture system is much simpler. We use three Kinect v2 sensors to obtain both the 2D images and 3D depth data from different view angles of real hair or a wig. Since the 3D hair point clouds are directly captured by Kinect v2 depth sensors rather than reconstructed from images, our method is computationally efficient and can effectively reduce the cost. Based on the hair images captured by the color sensors, we apply a fast and robust image enhancement algorithm to abstract hair strand segments from hair images. Then we apply a grow-and-connect algorithm to obtain relatively complete 2D hair strands represented by a predefined quantity of control points. Finally, we project the 2D hair control points on the 3D point clouds to obtain the 3D hair strands.

Our contributions are:

- Consumer level RGB-D sensors (Kinect v2 sensors) can be used to perform easy and inexpensive real straight hair capturing.
- Our image enhancement based 2D hair strands extraction method is computational efficient and robust with respect to the quality of input hair im-


Our grow-and-connect algorithm can provide relatively complete hair strands based on the local feature of hair strand segments.

2 RELATED WORKS

Image-based hair capturing have been explored recently. (Paris et al., 2004) proposed a system that captured 2D hair image sequences of a stationary head with a fixed viewpoint under a moving light source. They estimated the 2D hair orientation of the highlight and used the light information to obtain the 3D normal vector for hair modeling. (Wei et al., 2005) used a hand-held camera or video camera to capture 2D hair images under natural lighting conditions. They detected the local orientation of every pixel in each hair image and represented every hair fiber using a sequence of chained line segments. They applied triangulation for each fiber segment using image orientations of multiple views to reconstruct the 3D hair fibers and generated a visual hull to constrain the synthesis of the hair fibers. (Paris et al., 2008) presented an active acquisition system called Hair Photobooth. The system is composed of 16 cameras, 150 LED light sources and 3 projectors. They acquire hair images under different lighting directions from a number of cameras to recover the hair reflectance field. They also used triangulation to retrieve the location of the hair strands. (Jakob et al., 2009) proposed an approach that can obtain accurate individual hair strands by using focal sweeps with a robotic-controlled camera equipped with a macro-lens. (Beeler et al., 2012) used an algorithm to reconstruct facial hair strand geometry using a high resolution dense camera array. They developed an algorithm to refine the facial hair strand connections and remove outliers. (Luo et al., 2012) presented a hair modeling method based on orientation fields with structure-aware aggregation. This method can reconstruct detailed hair structures for a number of different hair styles. (Luo et al., 2013b) developed a hair modeling method based on an 8-camera wide-baseline capture system. They applied strand-based refinement to reconstruct an approximate hair surface and evaluated their reconstruction method on a set of synthetic hair models, resulting in an average reconstruction error of about 3 mm. (Luo et al., 2013a) proposed a structure-aware hair capture system. They had two systems to capture the wig and real hairstyle: A camera held by a robotic arm takes 50 images from different viewpoints for each wig and the real hairstyle capture system consisted of 30 cameras. The system reconstructed 3D point clouds from multi-view images. They also calculate the 3D orientation field based on the 2D orientation fields in each image. Then complete hair models were generated using a procedure that started from strand segmentations to ribbons and finally to complete wisps. However, the performance of this method depended on a good initial point cloud from multi-view stereo capture. Moreover, a careful and time-consuming manual clean-up procedure is needed. (Hu et al., 2014) introduced a hair capture system using simulated examples. They used the Super-Helices model to simulate static hair strands and generated 18 hair model databases. They applied a strand-fitting algorithm to fit cover strands they reconstructed from multi-view hair images onto the generated models in order to obtain structural plausible hair models. By introducing the simulated examples, they avoid the procedure of manually cleaning up the outliers in 3D reconstruction. Determining which hair model should be used is a key step in their method. However, the strand-fitting algorithm may need to go through all available databases to determine the fitness and the corresponding procedure is time-consuming.

3 CAPTURING SYSTEM

We have adopted the Kinect v2 sensor which was released by Microsoft in August, 2014. The new generation Kinect has a higher definition camera with a resolution of 1920 by 1080 and is equipped with a new depth sensor which employs time-of-flight (ToF) technology. Our hair capture system consists of three Kinect v2 sensors. The three Kinect v2 sensors are placed at the back side, right side, and left side of the model, as shown in Figure 1.

Figure 1: Hair Capturing System.
This arrangement helps us to capture the images and the depth data of the hair. We apply a standard calibration (Macknojia et al., 2013) and registration method to build the relationship between multiple Kinect v2 sensors in order to obtain a composite 3D hair point clouds.

4 2D KEY HAIR STRANDS GENERATION

A human head typically consists of a large volume of small-diameter hair strands (Ward et al., 2007), thus it is very difficult to abstract each single hair strand from hair images. Since our system is designed to capture 3D models of long straight hair, and noting that adjacent hair strands tend to be alike, it is possible to generate key hair strands and add similar neighboring strands to obtain complete 3D hair models.

From previous hair capturing methods, we discovered that the most significant information of hair images is the orientation of the hair strands. Gabor filters are well suited to estimating the local orientation of hair strands (AK. and F., 1990). The intensity hair images are convolved with Gabor filters of varying filter kernels for different orientations (we use 10 equidistant orientations covering a range of 0° to 180°. At each pixel position, the orientation that produces the highest Gabor response is stored in the orientation map and the maximum response is saved as the Gabor response image, as shown in Figure 2.

Hair strands, especially long hair strands, are difficult to directly extract from Gabor filter results using traditional edge-detection algorithms. Thus, we apply the image enhancement algorithm developed by (Hong et al., 1998) to enhance the hair strands geometry in hair images based on the estimated local orientation and frequency. We perform normalize the image, then both the orientation and frequency are estimated. Furthermore, we generate the region masks and filter the image again. The enhanced hair image is shown in Figure 3. This image enhancement algorithm is computationally efficient and robust with respect to the quality of input hair image.

We then erode the enhanced image to give the hair strands the one-pixel-width presentation which is easy to track by using a standard line-tracing algorithm. However, the eroded image contains some individual points and bifurcation points need to be removed. We also remove the segments which are shorter than the predefined length threshold. In addition, we apply the hair region mask on the hair image to obtain only the strand segments in hair region, as shown in Figure. We use 10 control points to represent each hair strand segment (1 head point, 1 tail point and 8 body points), as shown in Figure 5 (a).

In order to connect the hair strand segments into long hair strands, we apply a grow-and-connect algorithm. The grow and connect result is shown in Figure 5. The procedure of the algorithm is:
Figure 4: Hair strands segments extraction results. (a) Eroded hair strand image. (b) Eroded hair strand segments in hair region. (c) Tracked hair strand segments shown in different colors.

- Step 1: Current hair strand segment grows in two directions with a predefined increment threshold;
- Step 2: Calculate the distance between the head/tail point of current hair strand segment and other segments;
- Step 3: Choose the pair of possible connection points with the minimum distance.
- Step 4: If the minimum distance is smaller than the predefined distance threshold then connect the segments.
- Step 5: After all possible connections have been made, repeat step 1 to step 4 a predefined number of times.

Figure 5: 2D hair strands generation result. (a) Spline represented hair strand segments. (b) grow-and-connect long hair strands result. (c) hair strands shown in the original captured hair image.

5 3D HAIR MODEL

5.1 3D Point Clouds Alignment

With the calibration parameters we obtained, the point clouds can be transformed and aligned together. And we adopted point-to-plane based ICP (Besl and McKay, 1992) algorithm which utilized iterative method to minimize the distance between the points of two point clouds to improve the alignment process, as shown in Figure 6.

5.2 3D Connection Analysis

With the 2D hair strand connection method, we can partially solve the hair occlusion and missing data problem. However, the hair segments are considered as plane curves in the 2D method, and the 3D information is not taken into account, forcing us to consider and analyze the strand segment connections in 3D space. We adopt a three-step method to connect the strand segments in 3D focusing on short-distance connection and long-distance connection in different phases. First, we merge the strand segments which had relatively short distance between their end points.
For each strand segment with the curve parametric equation:

\[ c = f(t) \]

where \( t \) ranges from 0 to \( n \).

The end point \( f(0) \) is defined as the head point, and the end point \( f(n) \) is defined as the tail point. For each strand segment’s head point, we search for the nearest tail point attempting to find the proper connection candidates. A distance threshold value \( d \) is used to control the search algorithm and limit the distance between the head and tail of two segments. We screen the tail point candidates and keep them only when the distance between a given strand’s tail and current strands head is shorter than \( d \), and find the strand segment whose tail is the closest to the current strands head. We define the current strand as reference strand and candidate strand as target strand. As shown in Figure 7, there are three basic conditions for the possible connection in the first step.

We provide three merging solutions for these conditions. For the overlapped segments, we delete three points from the head of the reference strand and the tail of the target strand. For the missed segments, we delete three points from only the tail of the target strand, and connect the new tail to the head of the reference strand. For the regular separated segments, we connect the head and tail directly.

Afterward, we try to find the long-distance connections. We define a cone whose apex is the reference segment’s head point, and whose cone angle is equal to 30°. The tangent line at the head point of the reference segment points to the center point of the cones base. Segments with their tail point within the area of the cone are the candidate segments (target segments). To determine whether a segment is a candidate or not, we define the head point of the reference segment as A, and the tail point of a segment as B. The vector from A to B is defined as \( AB \), and the direction vector of the tangent line at the head point of the reference segment is defined as \( V \). If the vector angle between \( V \) and \( AB \) is smaller than 15°, the segment is considered a target segment. The vector angle of \( AB \) and \( V \) is defined by equation:

\[
\cos^{-1}\left(\frac{\langle AB, V \rangle}{|AB||V|}\right)
\]

In Figure 8, the \( \alpha \) is greater than 15°, thus the segment \( C \) is not a candidate. For all target segments, we calculate a connection possibility between the head and tail and connect the reference and target segment with the highest possibility. The connection possibility was introduced into the determination procedure as a value that can measure the viability of a connection between two strand segments. Our algorithm considers the straight-line distance between the tail and head of two curves, the end point curvature in 3D, the slope, and the plane curve curvature in 2D.

For a reference curve \( C_r \), the slope at the tail is defined as \( S_{rt} \) and the curvature at the tail is defined as \( K_{rt} \). For a piece curve \( C_p \), the slope at the head is defined as \( S_{ph} \) and the curvature at the head is defined as \( K_{ph} \). The distance between the reference tail point and piece head point is defined as \( d \). A rough connection
possibility can be calculated by the equation:

$$\text{possibility} = \frac{\alpha}{|S_r - S_{ph}|} + \frac{\beta}{|K_r - K_{ph}|} + \frac{\gamma}{|S_r - S_{ph}| + |K_r - K_{ph}|}$$

where $\alpha$, $\beta$, and $\gamma$ are variable connection coefficients. By adjusting the coefficients, the relevance of each factor in the check can be altered to match a given scenario. In our experiment, we chose connection coefficients $\alpha = 5$, $\beta = 40$ and $\gamma = 0.5$. The connection results are shown in Figure 8.

![Figure 8: Long distance connection.](image)

The mapping relationship of the 2D hair image and the 3D point clouds can be obtained through Kinect coordinate mapping functions. Based on this mapping relationship, we can project 2D key hair strands onto 3D point clouds and obtain 3D key hair strands. By combining 3D key hair strands from different views, we can obtain the composite 3D key hair strands, as shown in Figure 9.

![Figure 9: 3D key hair strands. The red points are the head points and blue points are the tail points of the 3D hair strands. (a) the original 3D key hair strands. (b) the 3D key hair strands after connection in 3D](image)

With the curve parametric function from the data analysis stage, we computed the position of serial continuous 3D control points. The key hair strands generated with the curve function represent the primary hair style and a particle system can be used to create a large number of child hair strands surrounding each key hair, with child hairs being sub-particles inheriting their properties from their parent, a key hair. A complete hair model can be generated this way.

With the control point information, the child hair strand density, style, and length can be determined. Child hair strands are sub-particles, making it possible to work primarily with a relatively low number of parent particles for which the physics are calculated. They carry the same material as their key hair and are colored according to the exact place from which they are emitted. The number of children per key hair affects the overall density of the hair and informs its look. Experimental results are shown in Figure 10.

6 CONCLUSIONS AND FUTURE WORK

In this paper, we present a novel system for capturing straight or relatively straight hair of medium or long length based on Kinect v2 sensors. We take advantages of the depth and color sensors of the Kinect v2 to obtain reliable 3D depth data and 2D hair images. Based on the 2D hair images, we introduce an image enhancement algorithm to abstract 2D hair strand segments followed by a grow-and-connect algorithm to generate long 2D key hair strands. By projecting the 2D key hair strands onto 3D point clouds, we can obtain key 3D hair strands and generate the surrounding child strands. Since the 3D hair strands of our long straight hair models are presented using control points, our hair model can be easily adapted for use in rendering and animation.

The modeling of relatively straight hair of medium to long length is the first step in our hair modeling method based on Kinect v2 sensors. In the future, we will apply our hair modeling system to more complicated hairstyles, such as curly hairstyles.
REFERENCES


