

Reconstruction of Relatively Straight Medium to Long Hair Models using Kinect Sensors

Chao Sun, Zhongrui Li and Won-Sook Lee

School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON, Canada

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 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-

ages.

- 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.

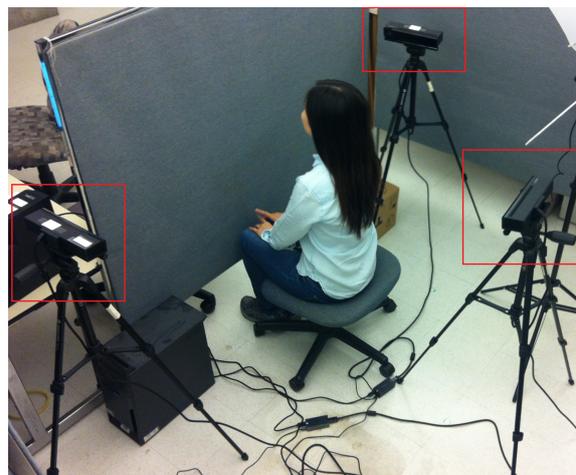


Figure 1: Hair capturing system.

This arrangement helps us to capture the images and the depth data of the hair. We apply a standard calibration (Macknoja et al., 2013) method to estimate the relative position and orientation between every pair of Kinect v2 sensors in order to obtain a composite 3D hair point clouds.

4 2D GUIDE 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 guide hair strands and add similar neighboring strands to obtain complete 3D hair models.

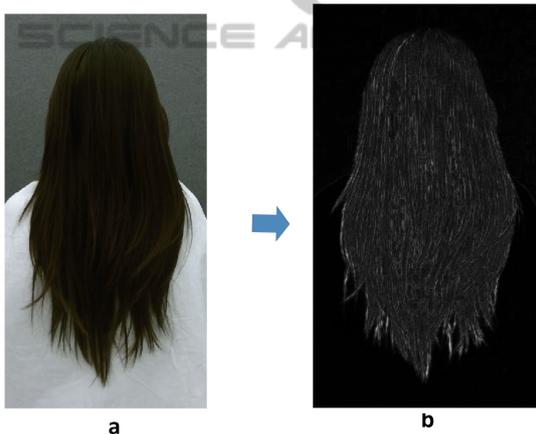


Figure 2: Gabor filtered Response. (a) the hair region image. (b) the Gabor filtered response.

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



Figure 3: Enhanced hair image.

(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 4. 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:

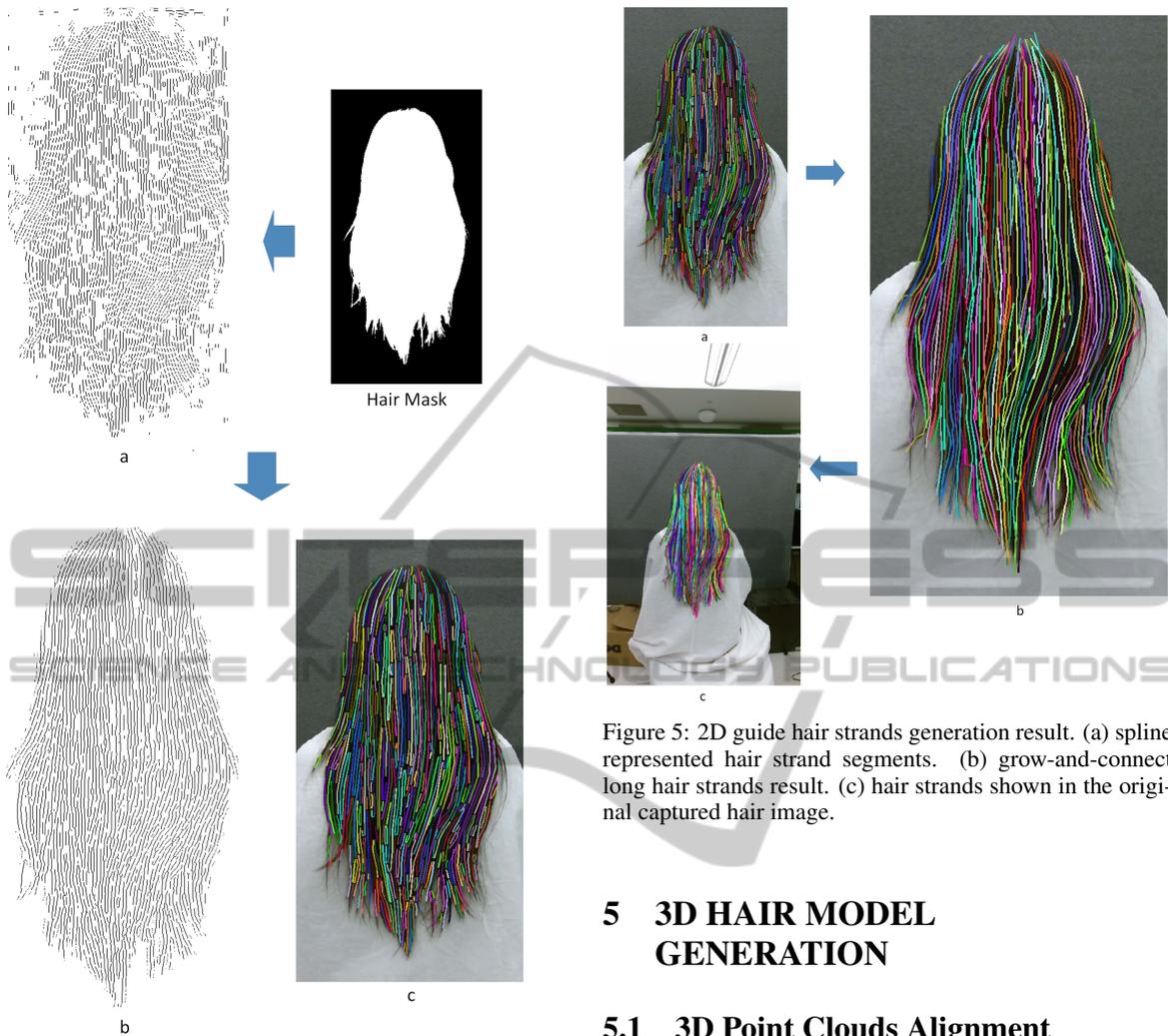


Figure 4: Hair strands segments extraction results. (a) eroded hair strand image. (b) eroded hair strand segments in hair region. (c) hair strand segments shown in different colors.

- 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 guide 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 GENERATION

5.1 3D Point Clouds Alignment

The point clouds alignment is performed with the Kinect calibration parameters. 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, the point cloud alignment results are 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 which are considered as plane curves in the 2D method do not take into account 3D information and it forces us to consider and analyze the strand segment connection in 3D space. We adopt a connection method focusing on short-distance connection and long-distance connection in different steps. First, we connect the strand

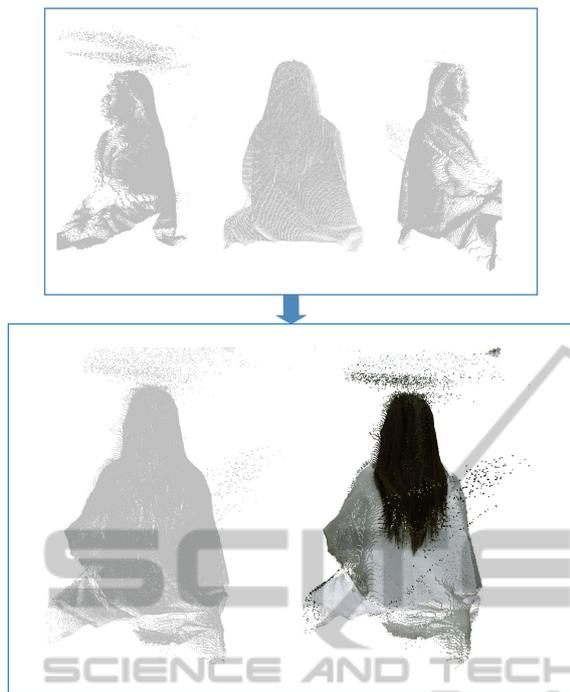


Figure 6: Hair point clouds alignment results. The upper row shows the three point clouds from the left, back and right point of views. The lower row shows the point clouds alignment result without/with texture.

segments which had relatively short distance between their end points. For each strand segment with the curve parametric equation:

$$p(x,y) = c(t) , t = [0, \dots, n]$$

The point $c(0)$ is defined as the head point, and the point $c(n)$ 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. We screen the tail point candidates and keep them only when the distance between a given strand's tail and current strand's head is shorter than d . We find the strand segment whose tail is the closest to the current strand's 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.

Then we provide three connection methods 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. In this case, we connect the new head to the new tail.
- For the missed segments, we delete three points from the tail of the target strand, and connect the new tail to the head of the reference strand.

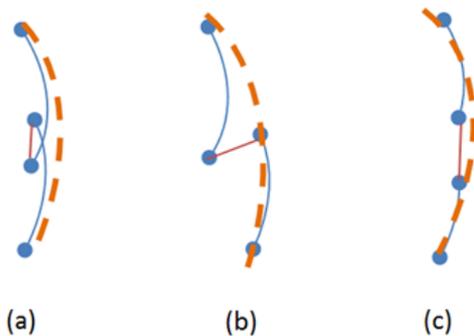


Figure 7: Short distance connection, (a) overlapped segments, (b) missed segments, (c) regular separated segments.

- For the regular separated segments, we connect the head and tail directly.

In addition, 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 cone's base. Segments with their tail points 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 \vec{AB} , and the direction vector of the tangent line at the head point of the reference segment is defined as \vec{V} . If the vector angle between \vec{V} and \vec{AB} is smaller than 15° , the segment is considered as a target segment. The vector angle of \vec{AB} and \vec{V} is defined by equation:

$$\theta = \cos^{-1} \left(\frac{\vec{AB} \cdot \vec{V}}{|\vec{AB}| \cdot |\vec{V}|} \right)$$

In Figure 8, the vector angle θ is greater than 15° , thus the segment C is not a candidate. For all target segments, we first calculate a connection weight value between the tails and the reference's head. Then we connect the reference segment and target segment with the highest connection weight. The connection weight 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 and the slope

For a reference curve C_r , the slope at the head is defined as S_{rh} and the curvature at the head is defined as K_{rh} . For a target curve C_t , the slope at the tail is defined as S_{tt} and the curvature at the tail is defined as K_{tt} . The distance between the reference's head point

and the target's tail point is defined as d . A rough connection-weight w can be defined by the equation:

$$w = \alpha/|S_{rh} - S_{tt}| + \frac{\beta}{d} + \gamma/|K_{rh} - K_{tt}|$$

where α , β and γ are the 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

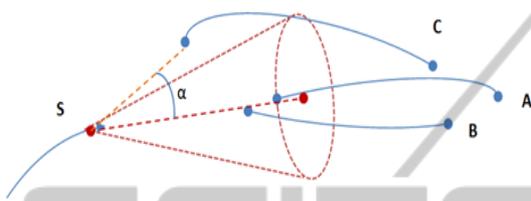


Figure 8: Long distance connection.

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

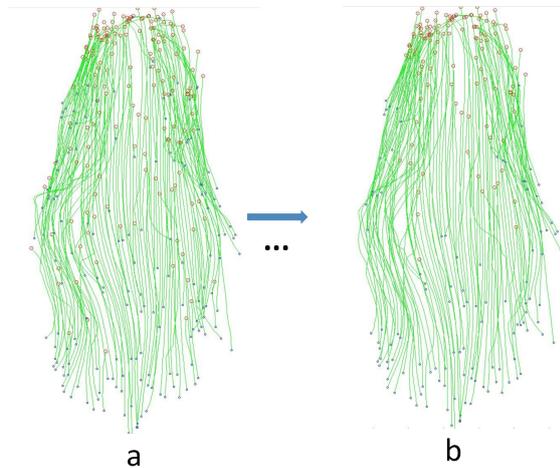


Figure 9: 3D guide hair strands. The red points are the head points, the blue points are the tail points. (a) the original 3D guide hair strands. (b) the 3D guide hair strands after connection.

With the curve parametric function, we select a serial of continuous 3D control points on every spline curve. The guide hair strands generated with the curve function represent the primary hair style and a particle system can be used to create a number of child

hair strands surrounding each guide hair strand. A complete hair model can be generated this way.

With the control points, the child hair strands' density, style, and length can be determined.

Child hair strands are sub-particles which makes it possible to work primarily with a relatively low number of parent particles. They carry the same physical properties and materials as their guide hair. And they are colored according to the exact location where they



Figure 10: 3D Hair model. The upper row shows the input hair images. The first row shows the input hair image. The second row shows the corresponding 3D hair model from the right, back and left views. The third row shows the 3D hair model from other point of views.

are emitted. The number of children hair affects the overall density of the hair. 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 guide hair strands. By projecting the 2D guide hair strands onto 3D point clouds, we can obtain guide 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

- AK., J. and F., F. (1990). Unsupervised texture segmentation using gabor filters. In *Systems, Man and Cybernetics, 1990. Conference Proceedings., IEEE International Conference on*, pages 14–19.
- Beeler, T., Bickel, B., Noris, G., Beardsley, P., Marschner, S., Sumner, R. W., and Gross, M. (2012). Coupled 3d reconstruction of sparse facial hair and skin. *ACM Trans. Graph.*, 31(4):117:1–117:10.
- Besl, P. and McKay, N. D. (1992). A method for registration of 3-d shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(2):239–256.
- Hong, L., Wan, Y., and Jain, A. (1998). Fingerprint image enhancement: Algorithm and performance evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:777–789.
- Hu, L., Ma, C., Luo, L., and Li, H. (2014). Robust hair capture using simulated examples. *ACM Transactions on Graphics (Proceedings SIGGRAPH 2014)*, 33(4).
- Jakob, W., Moon, J. T., and Marschner, S. (2009). Capturing hair assemblies fiber by fiber. In *ACM SIGGRAPH Asia 2009 Papers, SIGGRAPH Asia '09*, pages 164:1–164:9, New York, NY, USA. ACM.
- Li, H., Vouga, E., Gudym, A., Luo, L., Barron, J. T., and Gusev, G. (2013). 3d self-portraits. *ACM Transactions on Graphics (Proceedings SIGGRAPH Asia 2013)*, 32(6).
- Luo, L., Li, H., Paris, S., Weise, T., Pauly, M., and Rusinkiewicz, S. (2012). Multi-view hair capture using orientation fields. In *Computer Vision and Pattern Recognition (CVPR)*.
- Luo, L., Li, H., and Rusinkiewicz, S. (2013a). Structure-aware hair capture. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 32(4).
- Luo, L., Zhang, C., Zhang, Z., and Rusinkiewicz, S. (2013b). Wide-baseline hair capture using strand-based refinement. In *Computer Vision and Pattern Recognition (CVPR)*.
- Macknojjia, R., Chavez-Aragon, A., Payeur, P., and Laganier, R. (2013). Calibration of a network of kinect sensors for robotic inspection over a large workspace. In *Robot Vision (WORV), 2013 IEEE Workshop on*, pages 184–190.
- Paris, S., Briceño, H., and Sillion, F. (2004). Capture of hair geometry from multiple images.
- Paris, S., Chang, W., Kozhushnyan, O. I., Jarosz, W., Matusik, W., Zwicker, M., and Durand, F. (2008). Hair photobooth: Geometric and photometric acquisition of real hairstyles. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)*, 27(3):30:1–30:9.
- Shapiro, A., Feng, A., Wang, R., Li, H., Bolas, M., Medioni, G., and Suma, E. (2014). Rapid avatar capture and simulation using commodity depth sensors. *Computer Animation and Virtual Worlds*.
- Tong, J., Zhou, J., Liu, L., Pan, Z., and Yan, H. (2012). Scanning 3d full human bodies using kinects. *Visualization and Computer Graphics, IEEE Transactions on*, 18(4):643–650.
- Wang, R., Choi, J., and Medioni, G. (2012). Accurate full body scanning from a single fixed 3d camera. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*, pages 432–439.
- Ward, K., Bertails, F., Kim, T.-Y., Marschner, S. R., Cani, M.-P., and Lin, M. C. (2007). A survey on hair modelling: styling, simulation, and rendering. *IEEE Trans. on Visualization and Computer Graphics*, 13(2):213–234.
- Wei, Y., Ofek, E., Quan, L., and Shum, H.-Y. (2005). Modeling hair from multiple views. *ACM Trans. Graph.*, 24(3):816–820.