Computer-Aided Design 96 (2018) 31–41

Contents lists available at ScienceDirect

Computer-Aided Design

journal homepage: www.elsevier.com/locate/cad

Capture of hair geometry using white structured light[☆]

Yin Chen [a,c], Zhan Song [b], Shuai Lin [c], Ralph R. Martin [d], Zhi-Quan Cheng [b,e,*]

[a] School of Defense Engineering, Army Engineering University of PLA, China
[b] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
[c] School of Computer Science, National University of Defense Technology, China
[d] School of Computer Science & Informatics, Cardiff University, Wales, UK
[e] Avatar Science Company, Changsha, Hunan, China

ARTICLE INFO

Article history:
Received 21 December 2016
Accepted 26 October 2017

Keywords:
Structured light
Hair
Avatar
Strip-edge-based coding
Point cloud consolidation

ABSTRACT

Structured lighting (SL) scanning technology has been successfully applied to reconstruct personalized avatars, providing high-quality geometry for the face and body. However, previous white SL methods have typically been unable to capture and reconstruct hair geometry, preventing a complete, realistic avatar surface from being reconstructed. In this paper, we propose a novel hair capture system, producing a full-head manifold with complete hair, based on four white SL scanners made by ourselves. The key technical contribution is a robust strip-edge-based coding algorithm, using a projection pattern of 18 stripes; it allows geometric acquisition to an accuracy of 1 mm. Experiments are given to show that our system produces high-quality hair geometry, comparable to that captured by more expensive state-of-the-art multiview stereo systems. Our system is however easier to configure, and is more suited to real-world setups.

1. Introduction

Work on structured lighting (SL) scanning dates back over 40 years [1,2], and it is one of the most reliable techniques for capturing object surfaces. It has many intrinsic advantages, such as ease of configuration, and suitability for real-world applications outside a laboratory environment. Various SL methods [3–5] have been proposed, using binary or continuous light patterns sequentially projected onto the object of interest, while a digital camera captures images of it. Structured lighting has been widely applied to reconstruct personalized avatars with high-quality geometry for both the face and body. However, to the best of our knowledge, current white SL systems are unable to capture complete avatars, primarily due to the difficulty of capturing and reconstructing hair on the head.

The difficulty in hair reconstruction is mainly caused by the relatively low contrast in reflections, particularly for dark hair. A further problem is that light may penetrate through the hair if it is thin and sparse. 3D range data generated from such reflections is almost always unreliable, resulting in errors and artifacts in the captured geometry. This problem is made worse by the fine geometric detail of the hair: a state-of-the-art commercial SL system can achieve a resolution of about 1 mm, while the average diameter of a single hair fiber is about 0.08 mm. Therefore, it is widely believed that white SL is unsuitable for hair capture.

Fortunately, however, in many applications, the goal is to capture *overall* hair geometry, rather than precise detail of every strand. This is possible and meaningful as the hairs on the head do not form a random structure, but generally lie in groups of hairs forming tufts and curls. The general method of choice for high-quality hair geometry capture is to use expensive multiview stereo [6] with a complex, carefully configured camera array. However, as Echevarria notes [7], hair reconstruction from multiview stereo is very difficult to achieve outside a laboratory setting. Thus, as in [7], the goal of our technique is to reconstruct a high-quality *overall* hair surface, unlike other work which aims to reconstruct individual strands of hair [8–13].

In this work, we present a novel white SL hair capture system that produces high-quality overall hair geometry, with results which are comparable to those from state-of-the-art multiview stereo systems having that same goal. Our system can handle various hair-styles ranging from short to long, and from straight to curly. The captured hair geometry can be combined with the rest of the scanned head to build a full-head model (see Fig. 1).

The key technical contribution of the system is a robust strip-edge-based coding algorithm, using a projection pattern of 18 stripe images to provide a geometric acquisition accuracy of better than 1 mm. This algorithm is very reliable, and insensitive to hair appearance in terms of color, texture (e.g. curls) and reflectivity.

[☆] This paper has been recommended for acceptance by S. Schaefer.

[*] Corresponding author at: Avatar Science Company, Changsha, Hunan, China.
E-mail addresses: chris.leo.chan@gmail.com (Y. Chen), zhan.song@siat.ac.cn (Z. Song), shawn.shuai.lin@gmail.com (S. Lin), ralph@cs.cf.ac.uk (R.R. Martin), cheng.zhiquan@avatarscience.com (Z.-Q. Cheng).

https://doi.org/10.1016/j.cad.2017.10.006

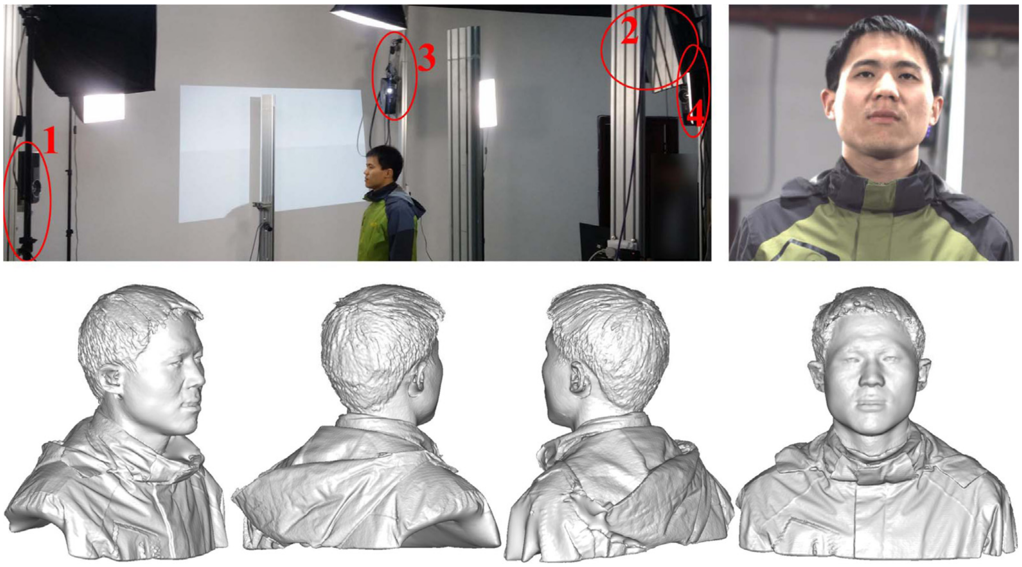

Fig. 1. Our system can acquire high-quality hair geometry using white structured light. Top left: our system uses four scanners (shown in red) around the subject. Top right: subject, from the front. Bottom: reconstructed result, from various directions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Acquisition of hair geometry data using a single white SL scanner we have built around this algorithm takes just 0.3 s. Our overall system uses four such scanners in a laboratory; they capture data sequentially to avoid inference between projected light patterns. Thus, scanning overall takes about 1.2 s, which is a short enough time for a subject to hold steady, with little movement between different scans.

After the raw hair data is captured by our system, erroneous points in regions with sparse and short hair (e.g. the temples) are re-positioned using a point cloud consolidation process, which is our second technical contribution. In such regions, the hair position may appear offset because of light penetration through the sparse or short hair to the scalp. Point cloud consolidation is achieved by a variant of a bilateral filtering scheme which moves the offset points to their expected positions. This allows our system to achieve visually pleasing reconstruction results even in the presence of sparse and short hair.

In the rest of this paper, we briefly review related work in Section 2. We then give technical details of data acquisition in Section 3 and point cloud consolidation in Section 4. Section 5 presents experiments and analyzes the results. Conclusions are discussed in Section 6.

2. Related work

Hair capture is an active research topic in computer graphics. A recent survey [14] has summarized and categorized existing methods from different perspectives. As previously noted, we focus on the smooth surface reconstruction of the *overall* hair style [7], rather than reconstructing highly-detailed individual wisps or hair strands as is done in [8,9,11,12].

State-of-the-art approaches for overall hair capture are often based on multi-view stereo reconstruction systems, using multiple digital cameras [6]. A typical example can be found in [7], which reconstructs a personal hair-style, with fine-scale geometric detail suitable for e.g. 3D printing. However, as noted by [7], such multi-view systems are costly, and have complex capture setups which are more suited to the laboratory than the real world.

Commodity RGB-D cameras (e.g. Kinect [15], Primesense [16], Xtion Pro Live [17]) provide low-cost scanning hardware, which can be used as a basis for capturing 3D personalized avatars with hair. The pioneering KinectFusion [18,19] is a GPU-based capture system using an RGB-D camera for both tracking and rigid surface reconstruction, allowing users to incrementally capture geometrically accurate 3D models. It has motivated many follow-up algorithms, using differing strategies to improve tracking speed and to expand spatial mapping capabilities to larger scenes or deformable avatars. In particular, the recent DynamicFusion algorithm [20] demonstrates robust performance when reconstructing a non-rigidly deforming human body in real-time. However, such low-cost hardware can only provide low-quality results, with an acquisition accuracy limited to several mm, and further suffers from long capture time due to redundant data collection techniques.

Structured lighting (SL) scanning is the most successfully commercialized technique, for a variety of commercial and technical reasons, including scanning quality, price, reliability, and capture speed. Since the invention of SL scanning [1,2], numerous SL systems have been devised for surface acquisition. Several survey papers [3–5] demonstrate that they have been widely applied to build personalized avatars with high-quality geometry for human faces and bodies. However, the hair on the head of the reconstructed avatar is always missing in previous work, due to the difficulty in distinguishing structured light patterns on dark, highly detailed hair surfaces.

Using expensive near-infrared lights is one approach to providing sufficient contrast in dark areas, allowing the recovery of dark surfaces under general illumination [21]. However, it requires the subjects being scanned to be small and close to the infrared scanner (about 20 mm from it): the limited power of near-infrared light projectors greatly limits the size of the working volume. This limitation restricts its suitability for human hair capture.

Instead, we revisit the question of whether (dark) hair can be effectively captured by using white structured light. The key question is whether the projected signal, in the form of a light pattern, can be coded in a way that can be accurately recovered. As Geng [5] notes, existing SL techniques can be classified into multiple-shot (sequential) and single-shot methods, according to the number of projected patterns. Multiple-shot methods often generate more reliable and accurate results, provided that the 3D target is static. We use a robust, multiple-shot, strip-edge-based coding algorithm to acquire hair geometry. Compared to using raw