Augmented Dynamic Shape for Live High Quality Rendering

Tony Tung*
Rakuten, Inc.

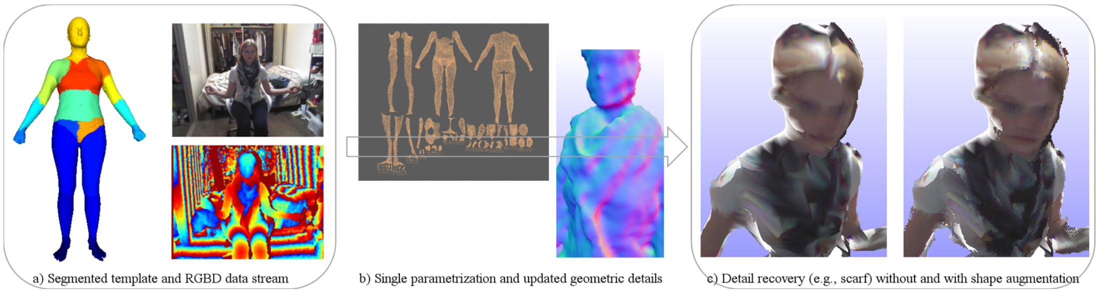

Figure 1: a) Input data, b) Normal map updated across time, c) complete shape recovery without and with augmentation information.

1 Introduction

Consumer RGBD sensors are becoming ubiquitous and can be found in many devices such as laptops (e.g., Intel's RealSense) or tablets (e.g., Google Tango, Structure, etc.). They have become popular in graphics, vision, and HCI communities as they enable numerous applications such as 3D capture, gesture recognition, virtual fitting, etc. Nowadays, common sensors can deliver a stream of color images and depth maps in VGA resolution at 30 fps. While the color image is usually of sufficient quality for visualization, depth information (represented as a point cloud) is usually too sparse and noisy for readable rendering.

Moreover, as a fixed camera only captures a scene from a single viewpoint, observed objects are usually incomplete. Hence it has been uncommon to use this technology for online 3D object visualization. In this work, we present a novel approach to achieve live high quality rendering of 3D objects in motion using a single consumer RGBD camera (e.g., Microsoft Kinect 1). The current method is particularly designed to work with human body shape, and is optimized for rendering using standard graphics pipeline and GPU. We believe it can be utilized for many applications such as 3D video chat, 3D games, special effects, e-shopping, etc.

2 Our Approach

The idea is to deform a low resolution 3D template shape in real time, while augmenting it with geometric details provided by higher resolution depth information. The proposed method consists of two parts: 1) shape modeling and 2) live rendering of details, and is applied to dynamic shape reconstruction of human in motion.

First, for shape modeling, we employ a deformable (parametric) template learned from real-world 3D scans of humans in various positions by PCA. The parameters represent coordinates in a subspace of human body shape and can be tuned to match any body shape [Anguelov et al. 2005]. The template is also divided into body parts which are associated to a skeleton. Hence, by applying a state-of-the-art body part recognition on observed depth maps (see Kinect SDK), we can recover bone transformations and manipulate the template by inverse kinematic (like a virtual avatar). Hence, we obtain a sequence of 3D meshes reproducing body motions captured by a RGBD sensor. Importantly, the meshes have consistent data structure over time (i.e., same number of vertices, and same mesh connectivity). Moreover, as the relative position of the observed shape is given by the detector, we can dynamically align and fuse depth measurements on the deformed template over time by average weighting.

Second, to obtain geometric details on a discrete surface, one could displace template vertices to the closest depth map point. While this works, a high resolution mesh would be required to obtain high quality rendering. In that case, computation performance could decrease if the mesh is too big. On the other hand, as meshes have consistent data structure over time, a unique parametrization (i.e., UV map) can be used for the whole sequence. Hence, we propose to compute and update normal maps obtained by projecting observed points (from depth maps) onto the template mesh surface. For the implementation, we actually go through the template mesh triangles (25000), check for visibility with respect to sensor's viewpoint position (at $(0, 0, -1)$ by default), and project the visible vertex onto the depth image plane (see Fig. 1 c)). Displacement information is contained at the pixel coordinate found in the image plane. As the parametrization is pre-computed only once, the complexity is log-linear. Hence, the pipeline is optimal for live high quality rendering of dynamic shape. Note that the sequence can also be encoded in geometry video for efficient geometry compression. We captured subjects with a single RGBD camera for validation.

To our knowledge, no similar method with body shape estimation, online template-based completion and rendering of augmented shape using normal maps has been proposed in the literature.

*e-mail:tony.tung@rakuten.com

SIGGRAPH 2015 Posters, August 09 – 13, 2015, Los Angeles, CA.
ACM 978-1-4503-3632-1/15/08.
http://dx.doi.org/10.1145/2787626.2787643

References

ANGUELOV, D., SRINIVASAN, P., KOLLER, D., THRUN, S., RODGERS, J., AND DAVIS, J. 2005. Scape: Shape completion and animation of people. *SIGGRAPH*.