Modeling Repetitive Motions in Real-World 3D Scenes

Yi Xu  Daniel G. Aliaga
Purdue University

1. Introduction

Obtaining models of dynamic 3D objects is an important part of content generation for computer graphics. If the states or poses of the dynamic object repeat often during a sequence (but not necessarily periodically), we call such a repetitive motion. There are many objects, such as toys, machines, and humans, undergoing repetitive motions. Our key observation is that for repetitive motions we can use one fixed camera to perform robust motion analysis and a second capture-device to provide 3D information of each motion state. After the motion sequence, we group temporally disjoint observations of the same motion state and produce a smooth space-time reconstruction of the scene. Effectively, the dynamic scene modeling problem is converted to a series of static scene reconstructions, which are much easier to tackle. The second device can be either a passive camera or an active-light projector, resulting in two different modeling techniques. Based on this observation, we present an efficient passive multi-viewpoint acquisition and a robust structured-light acquisition of repetitive motions.

2. Multiple-Viewpoint Modeling

To passively obtain a multiple-viewpoint model of repetitive motions, we combine the motion analysis benefits provided by a static camera with the object reconstruction ability of one or more moving cameras observing a static scene, together yielding an efficient capture system using as few as two cameras. Traditional multiple-camera acquisition methods can successfully capture 3D objects by using a static installation of cameras. However, these methods typically need a large number of calibrated and synchronized cameras. The fixed installation often precludes portability and limits the flexibility to adapt to the number of images and viewpoints needed by the reconstruction algorithm. In sharp contrast, our approach allows us to model dynamic objects without the costs and limitations of large and static multi-camera acquisition systems and without having to tackle correspondence establishment of a moving camera seeing a moving scene.

Our approach proceeds in two phases. In a brief preprocessing phase, the static camera observes the dynamic object and determines a sequence of \( M \) states of the repetitive motion. Subsequently, during a real-time capture phase, the images from the static camera are classified to one of the previously determined \( M \) states and the moving camera, synchronized to the static camera, simultaneously captures the state from different viewpoints. Once each state is sampled from a sufficient number of viewpoints, acquisition is complete. We then perform a silhouette-based volumetric reconstruction of each state of the acquired object and use a space-time optimization to align images over space and to stabilize the motion over time. The motion state sequence can be rearranged to produce new motion observable from novel viewpoints.

3. Structured-Light Acquisition

To actively obtain a dense geometry and color model of repetitive motions, we replace the second capture-device with a digital projector. Passive methods, such as the one in the previous section, are unobtrusive and simultaneously obtain color and depth information, but they are not as robust as active methods, which add energy into the scene. Active structured light methods that require “one-frame” are easily suitable for ranging dynamic scenes. However, only limited reconstruction density can be achieved by using a single pattern. If temporally disjoint images capturing the same motion state but under different structured light illumination patterns can be corresponded together, time-multiplexed codes can be used to acquire high density depth samples. Furthermore, if the state of the moving scene can also be matched against fully-illuminated images, the color and texture of the moving scene can be recovered as well.

Our approach uses a geometrically and spectrally calibrated camera-projector pair to capture a scene containing repetitive motions. For acquisition, an all white image and a set of two-color Gray-code patterns are sequentially projected onto the scene. Image analysis is performed to find a set of motion-state images under white illumination that generate a smoothly changing and repeating image sequence and that can be well-matched against the two-color structured-light illuminated images. Matching is done using an image differencing operator that is calibrated to work with different colored light sources. Since the observed motion tends to repeat, many desired motion states are eventually sampled by all patterns, yielding the ability to reconstruct motion states individually. The collection of reconstructed motion states can then be used to re-create a scene similar to the original or to produce new motion sequences.

4. Contributions

Our contributions can be summarized as: a methodology for modeling repetitive motions in real-world 3D scenes, an efficient passive multiple-viewpoint modeling method using as few as two cameras, and an active structured-light method for acquiring dense color and depth samples (Figure 1). More visual details are in the accompanying video.

![Figure 1. Reconstruction Results. a) Multi-viewpoint acquisition results. b) Structured light acquisition results. The top row shows a moving observer seeing a static scene. The bottom row shows a static viewer seeing the motion.](image-url)