3D Reconstruction Techniques for Virtual Worlds

Christopher Stanwyck

Comp 5470 Computer Graphics II

Oct 10, 2017

Abstract

Accurate and efficient reconstruction of virtual worlds from real life sources provides an intriguing way create virtual realities without requiring intensive synthetic modeling and artistry. As the speed of graphics hardware and software continues to increase, larger and more detailed worlds are becoming possible. Filling these worlds with higher detailed models, an increasing number of models, larger landscapes, greater numbers of structures, and all the things that make up these virtual scenes also increases in complexity. The ability to create these larger and more detailed worlds must keep up with the capability of rendering them. While it is important to be able to synthetically create virtual elements to maintain full artistic latitude, the ability to draw from a much larger pool of real-life art, objects, and landscapes can be used to quickly fill in the remainder of a virtual world.

Introduction

The process of 3D reconstruction includes both methods for acquiring real world data as well as methods for processing that data. Typical acquisition methods may include "…optical laser-based range scanners, structured light scanners, and LiDAR scanners, as well as passive methods such as multi-view stereo" [Berger et al.]. More recently, RGB-D cameras such as the Microsoft Kinect are also being used for acquisition. These cameras deliver image based solutions with an additional depth map provided by an infrared scanning device. More recently, other novel techniques such as the volumetric displacement of water have been used [Aberman et al.].

Once the data has been collected, several methods exist for analyzing it and creating the desired virtual element. Since the data almost always has some imperfections, the use of priors, or expectations based on similar models, is used to fill in the gaps. Otherwise, reconstructing the element is an ill posed problem with infinitely many possible solutions [Berger et al.]. In particular, when generating texture data from images which contain noisy or incomplete geometries, patch-based approaches may be used to process the data into a plausible result [Bi et al.]. Sometimes data acquisition may be very incomplete and strict bounds must be placed to obtain a detailed representation. For instance, reconstruction of wire art is very difficult for devices such as the Kinect due to the small and fine features. But, the features of wire art can be defined in a limited way to still provide detailed reconstruction [Liu et al.]. Other recent advances include the introduction of Convolutional Neural Networks (CNNs) to recreate textures for objects based on novel lighting conditions [Li et al.]. At present, there are a wide variety of methods for processing the data based on the desired representation and element class.

Surface appearance

One of the challenges in the reconstruction of 3D surfaces is the ability to regenerate that surface with novel lighting conditions. In order to prevent having to take many pictures at many angles under a substantial amount of lighting conditions, the properties of the material being imaged and current lighting conditions must be understood. Artists have the ability to look at images of surfaces and create plausible reflectance maps and similar surfaces. However, this is a time consuming process. Recent advancements in image processing have been put to use in the reconstruction of surfaces. Convolutional Neural Networks which have been trained with the appropriate data may be used to create reflectance maps of planar surfaces so that scanned images may be used as textures in virtual world under new lighting conditions with plausible results [Li et al.]. Li et al. have also created an improved self-augmentation approach for the CNN which requires less training data. Images may be synthesized to fill in gaps in the target manifold due to the fact that the forward rendering algorithm provides an exact inverse representation. This approach still requires training data, i.e. photographs of similar surfaces, and may have trouble extrapolating to more novel surfaces. Nevertheless, it may prove extremely useful for synthesizing similar surfaces.

Obtaining surface appearance from non-planar surfaces poses a different challenge. Creating an appropriate texture map requires knowledge of the geometry of the object being scanned. Commodity depth sensing cameras provide adequate but incomplete results. The inaccuracies in obtaining geometry may be transferred to creating a texture map. Bi et al. have devised a method which uses multiple images of an object and patch-based synthesis to create a plausible result. Multiple images are blended together which requires exact alignment or else ghosting effects can take place. Synthesized target images are created from each source image to provide an aligned image. Then, an energy function is used to measure both the completeness and coherence between the images. This ensures that the synthesized image is both true to the original image and free of artifacts due to incomplete or inconsistent geometry data. In some cases where the source image does not match the geometry, this can result in objects disappearing. However, this is more desirable than incomplete and ghosted images appearing in the resulting texture map.

Geometry

The other half of reconstructing an object is faithfully or plausibly re-creating the geometry. Most scans will produce a variety of imperfections which complicate the reconstruction process and make it ill-posed unless the use of priors is made [Berger et al.]. Liu et al. demonstrate that even with significant data imperfections a reliably accurate geometry can be captured. In their work, they show that wire art can be reconstructed when bounded by the features of typical hand-made wire art. The mapping process uses a multiple traveling salesman problem to find smooth paths. The algorithm favors curves with the shortest number of sections. However, these assumptions limit the usefulness of the technique. A new technique or set of rules must be made for each set of objects scanned.

Another promising technique has been demonstrated by Aberman et al. They demonstrate that full geometry, even self occluding, can be captured by using the volumetric displacement of water during a sequence of dips at varying angles. They name this the 3D dip transform. It is obtained by taking volume measurements at fixed intervals to create volume slices of an object. These slices can be viewed as sinograms, similar to computed tomography, with a sinogram for each dipping profile which is taken along a specific orientation of the object. Reconstruction is treated as a least-squares optimization problem. For each rotation, the sum over the rows of the rotated matrix is computed and an M-length vector is created. The end result is a voxelized binary solution.

Analysis and Limitations of Current Work

In generating the intermediate data for surface appearance, researchers have benefited from using synthesized data. Many problems can arise from alignment issues, coherency issues, and other conflicting or incomplete data. The intermediate synthesized step is a way to massage the data into a plausible result. This can compromise on creating an exact representation, but a plausible result is better than a malformed one.

Techniques for analyzing surface appearance have demonstrated great results, but with limitations. The requirements for creating reflectance maps for surface appearance were only demonstrated on planar surfaces. Adding geometry data to non-planar images would likely expand the training data for CNNs greatly and extend processing time to where it may not be feasible.

The use of priors is currently required for many reconstruction techniques. This rises due to errors on acquiring the data for processing. Improvements in current hardware, or new scanning techniques may be able to reduce the need for priors and help create a truly universal scanning and reconstruction mechanism. A current scanner may be set up for a general scene and be unable to reconstruct wire art, for example, unless manual user interaction is added to specify a particular region of the scene as needing select processing techniques.

Using water as a sensor is an interesting and novel technique. Obvious limitations include the acquisition time and the scale of acquisitions. In addition, porous materials are ill suited to displacing an accurate amount of water. This technique appears more useful for reconstructing a select range of art or hard objects of small stature.

Conclusions

While there are many different techniques and algorithms for reconstructing 3D objects and scenes from real-life objects, there currently is not a one size fits all approach. Great progress is made when the correct priors are known, but even then semantic information can be lost resulting in a representation which isn't completely faithful to the original [Bi et al.].

Despite the limitations, 3D reconstruction can be of great use in 3D and virtual reality applications. When synthesizing new worlds, using a real-life object as a reference and automatically generating a model is a great head start on populating the world. Faithful representation is not needed if the object has plausible appearance.

In addition to creating a virtual world based on real life-objects, mixed reality applications can benefit from immediate object reconstruction. These principles can be used to acquire information from the world, disseminate an object, and provide meaningful data to augment the world.

A future improvement in scanner cost and fidelity alongside a more comprehensive list of priors would enable reconstruction of vast landscapes, a capability that could assist with creating virtual tours, virtual commuting, or any such virtual representation of our world.

Bibliography

Aberman, Kfir and Katzir, Oren and Zhou, Qiang and Luo, Zegang and Sharf, Andrei and Greif, Chen and Chen, Baoquan and Cohen-Or, Daniel 2017. Dip Transform for 3D Shape Reconstruction. *ACM Transactions on Graphics 36, 4 (2017).*

Matthew Berger, Andrea Tagliasacchi, Lee M. Seversky, Pierre Alliez, Joshua A. Levine, Andrei Sharf, and Claudio T. Silva. 2014. State of the Art in Surface Reconstruction from Point Clouds. In Eurographics 2014 - State of the Art Reports, Sylvain Lefebvre and Michela Spagnuolo (Eds.). □e Eurographics Association. DOI:h□ps://doi.org/10.2312/egst.20141040

Liu, Lingjie and Ceylan, Duygu and Lin, Cheng and Wang, Wenping and Mitra, Niloy J. 2017. Image-based Reconstruction of Wire Art. *ACM Transactions on Graphics, 36, 4 (2017).*

Li, Xiao and Dong, Yue and Peers, Pieter and Tong, Xin 2017. Modeling Surface Appearance from a Single Photograph Using Self-augmented Convolutional Neural Networks. *ACM Transactions on Graphiccs, 36, 4, (2017).*

Bi, Sai and Kalantari, Nima Khademi and Ramamoorthi, Ravi 2017. Patch-based Optimization for Image-based Texture Mapping. *ACM Transactions on Graphics 36, 4 (2017).*