Harvesting Dynamic 3D Worlds from Commodity Sensor Clouds

EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2016)
C. E. Catalano and L. De Luca (Editors)

HARVEST4D: Harvesting Dynamic 3D Worlds from Commodity Sensor Clouds

T. Boubekeur[1], P. Cignoni[2], E. Eisemann[3], M. Goesele[4], R. Klein [5], and M. Wimmer[6]

[1]Telecom ParisTech, Paris, France
[2]CNR-ISTI, Pisa, Italy
[3]Technical University Delft, Delft, The Netherlands
[4]Technische Universitat Darmstadt, Darmstadt, Germany
[5]Universitat Bonn, Bonn, Germany
[6]Technische Universitet Wien, Wien, Austria

Abstract

The FP7 FET Open project "Harvest4D: Harvesting Dynamic 3D Worlds from Commodity Sensor Clouds" deals with the acquisition, processing, and display of dynamic 3D data. Technological progress is offering us a the wide spread availability of sensing devices that deliver different data streams, which can be easily deployed in the real world and produce streams of sampled data with increased density and speed, allowing us to iterate the sampling process in time. These data need to be processed and displayed in a new way. The Harvest4D project proposes a radical change in acquisition and processing technology: instead of a goal-driven acquisition that determines the devices and sensors, Harvest4D method lets the sensors and resulting available data determine the acquisition process. A variety of challenging problems need to be solved: huge amount of data, different modalities, varying scales, dynamic, noisy and colourful data. Moreover, the project focus is also in developing new approaches for the interactive visualization of complex sampled data.
This short contribution presents a selection over the many scientific results produced by HARVEST4D. We will focus on those results which could bring a major impact to the Cultural Heritage domain, namely facilitating the acquisition and analysis of the sampled data or providing better visual analysis capabilities.

Categories and Subject Descriptors (according to ACM CCS): Computer Graphics [Computing methodologies]: Shape Modeling—Visualization [Human-centered computing]: Visualization systems and tools—

1. Introduction

The FP7 FET Open project Harvest4D deals with the acquisition, processing, and display of dynamic 3D data. The consolidated acquisition process for visual models of 3D worlds requests:

- planning a specific scanning campaign,
- carefully selecting the (often costly) acquisition devices,
- performing the on-site acquisition at the required resolution,
- post-processing the acquired data to produce a beautified triangulated and textured model.

Technological progress is offering us a the wide spread availability of sensing devices that deliver different data streams, which can be easily deployed in the real world and produce streams of sampled data with increased density and easier iteration of the sampling process. These data need to be processed and displayed in a new way.

The Harvest4D project proposes a radical change in acquisition and processing technology: instead of a goal-driven acquisition that determines the devices and sensors, Harvest4D method lets the sensors and resulting available data determine the acquisition process. A variety of challenging problems need to be solved: huge amount of data, different modalities, varying scales, dynamic, noisy and colourful data.

Harvest4D involves six partners from five EU member states: TUWien (AT), Technische Universitat Darmstadt, and Universitat Bonn (DE), CNR-ISTI (IT), Telecom ParisTech (FR) and Technical University Delft (NL).

Harvest4D produced excellent results, scoring as one of the best among recent FET projects, with many papers published on first rank venues (i.e. more than 6 ACM Siggraph/TOG and 4 CGF/Eurographics publication). The full list of scientific results is available on the project web at http://harvest4d.org

This short contribution will present a selection over the many scientific results produced by Harvest4D, with a major focus on those results which could bring a major impact to the Cultural Heritage (CH) domain. In this framework, an Harvest4D application scenario is 3D sampling/documentation of a large evolving infrastructure; an immediate example of the former is the management of a major archaeological site, which encompasses all phases of the digitization and digital access pipeline. In this specific domain, we need new technologies for sampling evolving scenarios (e.g. the evolving progress of an archaeological excavation) and for processing the huge quantity of 3D samples produced. Once massive high-resolution 3D models are in place, the subsequent issue is how to provide ubiquitous and easy interactive access to those data. A work package in the Harvest4D project deals with new visualization and interaction paradigms, for example, visualization of large 3D data over the web. In the following, we describe our first results in the mentioned research directions.

2. Technologies for advanced sampling

3D reconstruction from images is bringing a revolution in 3D acquisition of CH, due to the economy and simplicity of the process (we do not need anymore to displace a costly scanner in the acquisition context).

MVE, the Multi-View Environment, is an end-to-end multi-view geometry reconstruction software [**?**]. In contrast to most image-based geometry reconstruction approaches, our system is focused on reconstruction of multi-scale scenes, an important aspect in many areas such as CH (Figure **??**) . It allows to reconstruct large datasets containing some detailed regions with much higher resolution than the rest of the scene. The MVE system provides a graphical user interface for structure-from-motion reconstruction, visual inspection of images, depth maps, and rendering of scenes and meshes. Another feature of MVE is to avoid to fill holes in regions with insufficient data for a reliable reconstruction. This may leave gaps in models but does not introduce artificial geometry, common to many global reconstruction approaches.

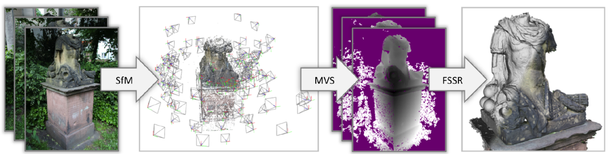

Figure 1: *The MVE system, providing enhanced multi-scale surface reconstruction from set of 2D photographs.*

The MVE system is based on an innovative surface reconstruction algorithm, Floating Scale Surface Reconstruction [**?**], that takes into account the inherent scale associated to each 3D sample. The method works with oriented, scale-enabled sample points and operates on large, redundant and potentially noisy point sets. The approach draws upon a simple yet efficient mathematical formulation to construct an implicit function as the sum of compactly supported basis functions. The implicit function has spatially continuous floating scale and can be readily evaluated without any preprocessing. One of the key properties of the approach is that it is virtually parameter-free even for complex, mixed-scale datasets. In addition, our method is easy to implement, scalable and does not require any global operations.

3. Technologies for advanced processing of sampled data

Processing huge dataset or multiple set of data sampling the same context at different time is not a simple task. Harvest4D results include a solution for accurate real-time simplification. Adaptive geometric simplification is still a complex task on non-trivial datasets. The highly efficient solution proposed in [**?**] is based on a new concept: Morton Integrals. By summing up quadric error metric matrices along Morton-ordered surface samples, we can extract concurrently the nodes of an adaptive cut in the so-defined implicit hierarchy, and optimize all simplified vertices in parallel. This approach is inspired by integral images and exploits recent advances in high performance spatial hierarchy construction and traversal. As a result, the GPU implementation can down-sample a mesh made of several millions of polygons at interactive rates, while providing better quality than uniform simplification and preserving important salient features (Figure **??**).

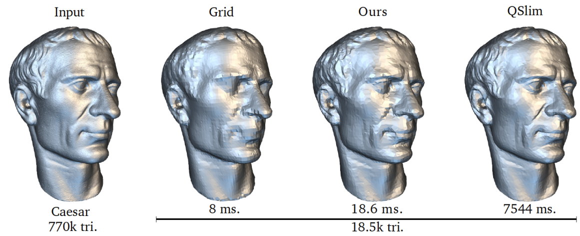

Figure 2: *An example of the results obtained in real-time using the geometric simplification algorithm proposed in [?].*

Another important issue is detecting geometric changes between different sampling of the same scene, performed by multiple acquisitions occurring at different times. This is a critical operation for all systems requiring a precise segmentation between change and no-change regions. Unfortunately, typical 3D scanning setups cannot provide any one-to-one mapping between measured samples in static regions: in particular, both extrinsic and intrinsic sensor parameters may vary over time while sensor noise and outliers additionally corrupt the data. A multi-scale approach was adopted in [**?**] to robustly tackle these issues. Starting from two point clouds, first outliers are removed using a probabilistic operator. Then, the actual change is detected using the implicit surface defined by the point clouds under a Growing Least Square reconstruction that, compared to the classical proximity measure, offers a more robust change/no-change characterization near the temporal intersection of the scans and in the areas exhibiting different sampling density and direction. The resulting classification is enhanced with a spatial reasoning step to solve critical geometric configurations that are common in man-made environments.

Finally, compression is another important issue with the usual huge datasets produced in CH applications, to support efficient storage and transmission. A compression approach that is efficient in storage requirements as well as in computational cost, as it can compress and decompress point cloud data in real-time, was proposed in [**?**]. Furthermore, it is capable of compressing incrementally acquired data, local decompression and of decompressing a subsampled representation of the original data. The method is

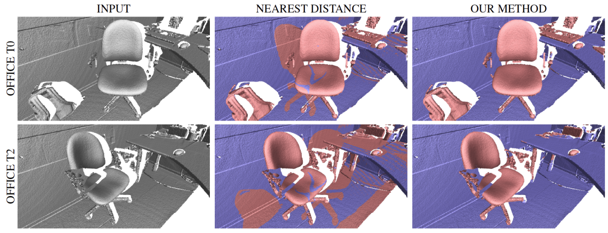

Figure 3: *An example of the results obtained with the change detection algorithm proposed in [?] .*

based on local 2D parameterizations of surface point cloud data, for which we describe an efficient approach.

4. Technologies for advanced visualization

Providing interactive visualization of the huge representation produced i the CH domain is not an easy task. We have either digitizations technologies able to produce tend of million samples/triangles even in the case of small or medium scale artworks; or, in other application cases, we could be interested in sampling very large scale subjects, up to the size of an archaeological site or an entire historical city. The issues related to the interactive visualization (either local or remote) and the need of effective approach to interact and navigate complex spaces are a key focus in Harvest4D.

A novel scene representation for the visualization of large-scale point clouds accompanied by a set of high-resolution photographs has been proposed in [**?**]. Many real-world applications deal with very densely sampled point-cloud data, which are augmented with photographs that often reveal lighting variations and inaccuracies in registration. Consequently, the high-quality representation of the captured data, i.e., both point clouds and photographs together, is a challenging and time-consuming task. A two-phase approach is proposed in [**?**]: first, a (preprocessing) phase generates multiple overlapping surface patches and handles the problem of seamless texture generation locally for each patch; second, these patches are stitched at render-time to produce a high-quality visualization of the data. This approach results to be one order of magnitude faster than equivalent mesh-based texturing techniques and provides improved flexibility when dealing with growing data sets.

With the enormous advances of the acquisition technology (increased sampling speed and resolution), converting sampled points into good-quality surfaces is becoming more and more complex. The consolidated approach (range maps registration, global reconstruction and texturing) could become impractical with increasing data sizes. The progress of GPU technology and of raytracing solutions allows to skip these pre-processing requirements, by introducing an efficient raytracing of multiple depth maps [**?**]. In a preprocessing phase, we first generate high-resolution textured depth maps by rendering the input points from image cameras and then perform a graph-cut based optimization to assign a small subset of these points to the images. At runtime, we use the resulting point-to-image assignments (1) to identify for each view ray which depth map contains the closest ray-surface intersection and (2) to efficiently compute this intersection point. The resulting algorithm accelerates both the texturing and the rendering of the depth maps by an order of magnitude (see Figure **??**).

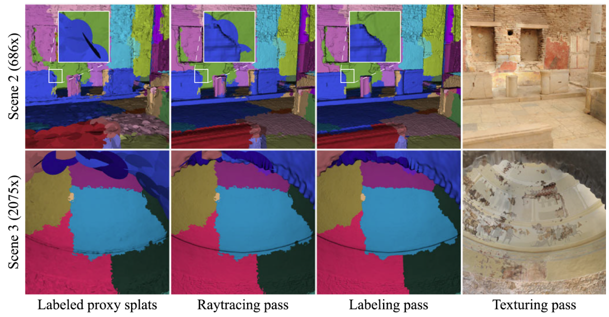

Figure 4: *An example of the results obtained with the raytracing solution proposed in [?].*

Another alternative approach is based on the conversion from sampled points or surfaces to a compressed voxel dataset. Voxel representations, coupled with raytracing rendering approaches have been used since two decades. Complexity resides in the need of providing very high-resolution and, usually, also uneven densities (as in most sampled datasets).

Recently, directed acyclic graphs (DAGs) were successfully used for compressing sparse voxel scenes as well, but they are restricted to a single bit of (geometry) information per voxel. A method to compress arbitrary data, such as colours, normals, or reflectance information, has been presented in [**?**]. By decoupling geometry and voxel data via a novel mapping scheme, it was possible to apply the DAG principle to encode the topology, while using a palette-based compression for the voxel attributes, leading to a drastic memory reduction. This method outperforms existing state-of-the-art techniques and is well-suited for GPU architectures, achieving real-time performance on commodity hardware for coloured scenes with up to 17 hierarchical levels ($128K^3$ voxel resolution), which are stored fully in core (see Figure **??**).

Figure 5: *Compressed voxelized scene at different levels of detail, rendered in real time using raytracing only [?].*

The remote visualization and navigation of 3D data directly inside a web browser is becoming feasible due to recent efforts in standardising the components for 3D rendering.

A method for the easy remote navigation of complex archeological 3D environments [**?**], represented by multi-resolution triangle meshes, has been implemented on top of the 3DHOP platform [**?**]. This system supports two intuitive navigation modes (see Figure **??**): the user can explore the model from the top (bird's eye mode, left image) or can walk inside the environment in a walk-through fashion (first-person mode, right image).

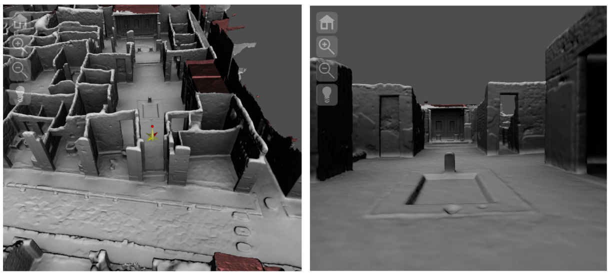

Figure 6: *Web-based visual navigation over Pompei's houses: switching from birds eye to first-person mode preserves the user position and orientation, as it is visible in these images.*

5. Conclusions

The Harvest4D project has significantly contributed to promote a technological progress by proposing a radical change in acquisition and processing pipelines focusing on on the efficient management of huge amount of 3D acquired data with different modalities, in varying scales, and considering dynamic, noisy and appearance rich enviroments data.

The six partners of the project contributed with several technical papers published on first rank venues like Siggraph/ACM TOG, Eurographics/CGF and IEEE TVCG making this project as one of the top scoring project in the CS area.

The contribution given by this project to the Cultural Heritage (CH) domain has been significant. For example the proposed technologies can be of paramount importance in 3D sampling/documentation of a large evolving environments like the management of a major archaeological site, a problem which encompasses all phases of the digitization and digital access pipeline. In this specific domain, the project provided new technologies for sampling evolving scenarios (e.g. the evolving progress of an archaeological excavation) and for processing the huge quantity of 3D samples produced. Moreover for the massive high-resolution 3D models generated the project has provided new visualization and interaction paradigms and tools for visualization of large 3D data over the web.