Capturing History Bit by Bit

Architectural Database of Photogrammetric Model and Panoramic Video

Takehiko Nagakura¹, Daniel Tsai², Joshua Choi³
¹,²,³ MIT
¹,²,³ {takehiko|dtsai|joshuac}@mit.edu

Architecture changes in real time. It appears differently as the sun and weather shift. And over a long span, it naturally wears and decays or may be renovated. This paper discusses the use of two emerging low-cost technologies, photogrammetric modeling and panoramic video, for recording such transformations of buildings. These methods uniquely capture a moment in the existence of a building, and deliver its three dimensional appearance and the sense of traversing in it like no other conventional media. An approach with a database platform is proposed as a solution for storing recordings amassed from fieldwork and making useful heterogeneous representations out of these unique contents for studying architectural designs.

Keywords: Photogrammetry, Panoramic Video, Building Database, Digital Preservation, Architectural Design

INTRODUCTION: TWO NEW TOOLS

This paper describes how low cost, accessible emerging recording technologies can be deployed for architecture -- for its study, design, historical research, preservation, and renovation. Two technologies are highlighted: photogrammetry and panoramic video. Areas of use are described with example field study, characteristics found about these tools are discussed, and a new database platform is proposed.

Photogrammetry is the science of making measurements from photographs (Lowe,1999; Zisserman,2001). One application of this process is computation of a 3D representation from an appropriate set of photographs of an object. Contemporary photogrammetric modeling software uses photographs to create both 3D form and texture applied on its surfaces. (Fig. 6)

Panoramic video can record a 360 degree spherical or cylindrical view for each video frame. While the video is running or when paused, the viewer can freely look up, down, and completely around, as well as close-up and away. It can play interactively within a desktop viewer (Fig. 2) like YouTube, on a head mounted equipment such as the Oculus, or through a special panoramic projection system to get a more immersive experience [1].

Photogrammetric modeling and panoramic video are both becoming increasingly popular among consumers due in part to (1) low cost consumer equipment (e.g. smart phone, digital camera); (2) free or low cost software and cloud-services (e.g. YouTube, 123D); and (3) familiar foundations: making 3D models (new) from photographs (foundation) and more degrees of freedom (new) in video (foun-
As the technologies mature, their use will reduce the complexity and time from field observation to shareable, malleable, spatial and visual constructs.

**AREAS OF USE**

Over the past three years, these recording techniques have been tested in different areas of applications to help study architecture on site. This section is both descriptive of our studies and prescriptive.

**Architectural History and Preservation**

Unlike many fields of art, architects and architectural historians cannot take the subject of investigation back to the lab for study. Visiting a remote site is costly, and often there is only a moment to access an important space. It is also not possible to preserve and present architecture in museum galleries.

Photogrammetric models and panoramic videos are no replacement of the real building. However, their photographic ways of recording the physicality and experience of architecture provide unique means to store and convey the reality like no other media. A photogrammetric model captures a form with its surfaces covered with photographic textures of the material. Historians can virtually examine this photorealistic model not only from angles of the original photographs but also freely at other viewing positions impossible for photographers because of obstacles and accessibility around the target. (Fig. 1)

A panoramic video captures the surrounding environment in motion graphics format. On a constrained path of movement, a viewer is free to look at any desired direction as people would do in real spatial experience. Since this tool captures all viewable surfaces around it without specifically pointing at a target, the recording becomes a very convenient source to check appearance of any part of the building without revisiting it. (Fig. 2)

**Design Study**

The ability of these technologies to quickly and easily record the 3D form and appearance of a building with nominal equipment greatly facilitates onsite study and discourse of architecture. This sets them apart from more expensive, professional equipment such as laser scanning device. For instance, students with just small cameras can sample architecture at various distances, and study its site, building, components and details in respective scales.

Figure 1
A photogrammetric model, 3D print and photo of Teatro Olimpico (A.Palladio and V. Scamozzi, 1585) (MIT Workshop 2013-5. D.Tsai, J. Choi and D. Rosenberg.)

Figure 2
A panoramic walkthrough (left) played from its equirectangular, spherical video (bottom) with its path shown in the plan of Villa Foscari (A. Palladio, 1563). (MIT Workshop 2014. D. Pinochet and T. Nagakura.)

Figure 3
Photogrammetric models of Brion Cemetery (C. Scarpa, 1978) captured from a drone (left) and on ground (right) at different times. (MIT Workshop 2015. T. Nagakura, J. Choi, D. Tsai, C. Cheng and Y. An.)
The same 3D forms can also be captured quickly at different times of the day, and its changing appearance due to lighting and weather can be compared. (Fig. 3) Panoramic videos captured in the same city piazza but with different events would inform how people inhabit the open space. (Fig. 11).

The ability to capture real world objects rapidly -- from details such as window frames, to facades, interior rooms, and whole buildings -- extends the traditional two dimensional photo collection to three dimensional representations. Objects can be studied in 3D, and used as components in new designs (Fig. 4) without de novo geometric construction. With careful editing, photogrammetric models can be made into physical scale models by 3D printing. (Fig. 1)

A detailed photogrammetric model of the original condition can be shared among the owner, architect, and contractors to review and discuss issues without being on the site. Subsequent models made before covering by the finish layer or painting can reveal the tectonics and materials used beneath the repaired surface. The model of the completed repair work can be compared with the original work to evaluate the effectiveness of the work. (Fig. 5)

In the long run, these recordings will also serve as valuable maintenance logs for facility management, especially if those models are accompanied by video recording showing the repairing process, contract documents, and other project information. (Fig. 7)

**FINDINGS: PHOTOGRAMMETRIC MODEL**

Photogrammetry and panoramic video tools have great potential for capturing architecture, but our field experiences suggest challenges and issues beyond just precision and image resolution that should be worked out as the technologies advance.

**Variability**

Photogrammetric models have varying quality. The process of turning photographs into 3D models is technical artistry. The reconstruction of forms and textures depends on the craft of taking appropriate photographs and managing their synthetic process.

Two prerequisites fundamental to photogrammetry computation are: (1) a set of photographs with partially overlapping views; (2) which are taken from different camera positions. (Fig. 6) The appropriateness of the photograph sets vary due to the photographic equipment used, the camera’s settings such as focus and exposure, the total number and angular variation of photographs, the sequencing of shots, and the amount of overlaps between shots.

The processing of a set of photographs into a 3D model is dependent on software and the underlying algorithms implemented. Uploading photographs to an online consumer service [2] and processing a model on a cloud server usually involves a simple selection among a few default settings, while advanced...
sifting and meshing software for 3D reconstruction [3, 4] requires tuning of many parameters with a wide range of possible results.

In an ideal situation where the camera can be placed anywhere around the object, an increased number of angles and resolution of photographs results in a model with superior details but at the cost of significant processing time and data size. It is desirable to prepare a model of quality that suits a particular purpose. For instance, preparing a single model with all details is different from representing an overall model with some separate detail models. (Fig. 3)

**Discreteness and boundary**
The makeup of a photogrammetric model is different from a 3D form made in conventional geometric modeling software. A photogrammetric model is formed of texture over meshed surfaces that derive from point clouds -- a set of easily recognizable marks observed in the original photograph set. The process of identifying these marks extracted from the set is called sifting and feature-matching. A pointed corner of an object as well as a small scratch on a concrete wall and tiny colored grain patterns on a marble floor indiscriminately qualify to make these point marks. Thus, seemingly discrete formal components appearing as a column, handrail, or floor in the model actually are a subset of a web of mesh that comprises the whole scene without any underlying articulation. (Fig. 8) The texture image imposed on this mesh gives the illusion of components as separate, but the underlying mesh is typically continuous. Converting it to a discrete component-based representation requires a laborious manual sorting operation or use of a special kind of emerging software that translates a captured mesh to a conventional geometric model assembled by discrete components [5].

Because of this makeup, the natural assumption about a building or its digital model as a construction of sensible parts does not apply to a raw photogrammetric model. For instance, a component cannot be easily selected unlike discrete shapes such as boxes, cylinders, and spheres made by conventional geo-

Therefore, returning to the same building site a year or even a day later encounters totally different conditions, and produces different photogrammetric models. (Figs. 3, 5, and 7) Over time, multiple models of the same space and objects can be created as 3D snapshots. These snapshots represent a life of the building in many versions, over time, and in varying conditions of season, weather, and lighting. This multiplicity is an essential feature of photogrammetric method that captures reality at a moment, and should be considered as the nature of the transient subject rather than just a limitation of the tool.

**Multiplicity**
Environmental conditions shift day by day, minute by minute, with the sun position and the weather changing illumination and shadows of the building surfaces. This affects the recording of their photographic texture. Buildings also have moving parts such as doors, windows and blinds. Over a long period of time, the form and the texture transforms through decay and weathering, and also with repairs and renovations.

Figure 6 A façade model of San Giorgio Maggiore (A. Palladio, 1611) processed in a cloud-based photogrammetric modeling tool, Autodesk 123D. Eight photos were taken from multiple locations and with overlapping views. (MIT Workshop 2013. J. Choi.)

Figure 7 Photogrammetric models showing a wall before and after repair, and a still from video recording of restoration process. (John Andrew Mansion Survey 2014-5. T. Nagakura, D. Tsai, J. Choi, and W. Sung.)
metric modeling software.

Figure 8
A photogrammetric model shown with and without texture: the boundary is irregular, and the underlining mesh has no discrete components articulated. (MIT Workshop 2014. T. Nagakura.)

Such a makeup also entails irregular jagged boundaries. As we see walls, floors, columns, and windows around us, few architectural parts exist physically separated from other parts. Photographs on sites are taken with these components constructed together, and the boundary of the resulting model is arbitrarily generated without being aligned with that of any building components.

With patience, editing the model can trim or split its mesh into one or more meshes close to the shapes of desired components, but their boundary edges remain jagged. Modifying an edge further to a cleaner, more linear one makes it difficult to reassemble the split meshes together without causing a visible seam.

Figure 9
A drone capturing the photographs of Villa Pisani at Bagnolo (A. Palladio, 1545). (MIT Workshop 2014. V. Leung.)

Completeness and overlap
Making a complete photogrammetric model out of a single session of taking photographs on site is impractical or not feasible for many kinds of building components, let alone the whole building. The process presents some challenges that are not typical of the capture process for small objects such as a sculpture or vase that may be placed and lit in an ideal condition for taking necessary photographs.

- Access to photograph a target from all sufficient angles often is problematic. A tree, waterway, and other built structures as well as the presence of people and cars may prevent sufficient circling shots without interference. The street may be too narrow for capturing the exterior of a building. Scaffolding or a drone with a camera is likely needed to capture the details of a roof, column capital, and other high locations. (Fig. 9)
- A classical column, for instance, has details that require close-up shots of the base and capital, while its long shaft with a monotonous surface is better captured with commanding distant shots. Ample space around the object and good planning for systematically taking all necessary photographs in sequence and with overlaps are needed.
- A table top, for example, has a well illuminated upper surface and a darker underside. Taking good photographs of both these sides requires steep adjustment of exposure around the edge, but, photogrammetric process of sifting and feature-matching prefers exposure levels that do not rapidly change.
- For an interior space, taking photographs of each surrounding surface of the room from reasonably varied camera positions is often difficult due to the tightness of the space and obstructing furniture on the floor. Capturing the floor requires high camera angles. Also, illumination is often low inside and with high contrast around windows.
- Architectural designs and their tectonics often have characteristics photogrammetric modeling applications have difficulty in processing to various degrees. They include materials that are transparent or highly reflective; monotonous surfaces like uniformly painted white walls; and heavily repetitive patterns such as an elevation of a high-rise building...
captured through a set of photographs of its portions. These features tend to confuse the sifting and feature-matching process of photogrammetric software.

As a result of these issues, trying to capture the entirety of a built work or a complete component is often impossible on site, or is not a practical approach to implement without extraordinary planning and special arrangements. And field experiences indicate an expedition of photogrammetric recording of a building usually produces many incomplete models of architectural parts, with some missing portions. On the other hand, a casual, intuitive, but educated expedition can be very productive if the capture-target is divided into reasonable parts prior to its recording, and each part comes with a task of capturing the area around it.

The incomplete 3D models derived from field work have overlaps in many cases. (Fig. 10) Divided recording tasks over sessions and individuals are planned to generate overlapping models with consideration of later assembly. Uncoordinated recordings would produce models of similar portions as common areas of interest and good accessibility.

FINDINGS: PANORAMIC VIDEO
Although the media for panoramic video is different from that of photogrammetric models, similar challenges are identified for recording and processing.

Variability
Panoramic video is delivered in a video file format including the sequence of frames just like conventional video. But in each frame, the environment around the camera is projected onto a flat image. The spherical equirectangular format is a widely used method that this paper discusses and captures the full spherical environment around the camera. There are other popular formats including cylindrical and cubic projections with different ways of encoding the scene onto an image frame. (Fig. 2)

Panoramic video will come in varying quality as much as photogrammetric models. The quality depends on all crafting factors regarding conventional video recording such as exposure, focus, shutter speed, shakiness, and speed of movement. Exposure and focus are particularly difficult since a frame includes all objects around the camera, not just those found within the finder of conventional video. For instance, in an interior space, a camera can be placed near one wall of the room but distant from another, and with a well-lit floor and a dark ceiling.

The equipment for panoramic video is rapidly evolving. Full spherical capture uses multiple video cameras fixed to a rig, and typically, processing software is required later to stitch the videos from all cameras into a single video format [6], although some systems include a stitching circuit on the rig.

Alternatively, a single camera or cellphone can be equipped with a special lens or mirror and capture panoramic video that covers a limited part of the surrounding environment. One disadvantage of the single camera method is image resolution, especially for capturing a very wide angle of environment. However, it advantageously does not require the process of stitching and the file size of clips with lower resolution is relatively small.

The stitching process to make panoramic video with multiple cameras involves technical artistry when dealing with the seams between video clips. After time syncing, each clip may have its own de-
sired exposure level appropriate for subjects in each image, and stitching needs either using an averaged level or some blending. Moreover, objects near the shared boundary in the frames of adjacent cameras often do not match exactly, because the cameras are positioned slightly away from each other, without physically being able to share a single point of conversion. (Fig. 12) This issue is known as parallax. Avoiding seams between clips often requires delicate and laborious adjustment in the stitching software.

Multiplicity
Just as photogrammetric models, panoramic video captures a moment of reality. Recordings on the same site made at different times will likely deliver different appearances of the building due to changing environmental condition and state of the building as well as events on the site. The presence of resting people and cars can obstruct recording of panoramic video, but they often become helpful additions if they are moving and making events that describe states of inhabited space and inform its use. (Fig. 11)

Discreteness and boundary
A panoramic video clip’s temporal boundary is as sharp and clear as that of a conventional video. It starts at a certain time, and finishes at another, while playing at a set speed in between. Therefore, a regular video editing software can be used for montaging panoramic clips, and can make up a complex panoramic story presentation on timeline. A clip can be trimmed in duration, concatenated with another back to back or with transition effects such as dissolving. Multiple clips can be spatially composed within the spherical image plane. (Fig. 13) The playing speed can be adjusted as well.

Conversely, panoramic video has no spatial boundary. A fully spherical recording does not require pointing a camera at a subject of interest. The continuous physical world around the rig is traced without consciously framing a discrete view of selected subjects. When the video is played, a viewer can see any desired direction by scrolling the image up, down, left or right; or turning the head with head-mounted display or inside a spherically projected room. The point of viewing interest is selected by the viewer, not by the person recording it.
Andre Bazin, a prominent film critic, wrote that the edges of the film screen are "the edges of a piece of masking that shows only a portion of reality." (Bazan, 1967) Accordingly, film-makers use this important tool to point the attention of the audience to intended subjects in the space. But architects do not assume a visitor’s view is artificially constrained within the space they design. Panoramic video not only captures and shows more surfaces than conventional video but also lets the viewer study and appreciate architecture in a way closer to what people do in the actual physical space.

When a viewer does not actively participate in changing the viewing direction, the view of the default direction is seen, and this direction can be set by the stitching software. Consistently keeping this direction to the same compass orientation or automatically rotating the view according to the position of the camera rig on a path or events happening around it needs to be carefully considered. (Fig. 14)

Completeness and overlap
Completeness of panoramic video recording on a building site needs discussed with recognition of two kinds of essential information values delivered by this medium: the recording (1) provides the appearance of space around the position of the camera rig; and (2) conveys to the viewer a sense of traversing in and around architecture. By visiting all possible locations around the exterior, and through each floor and every room, panoramic video recording of a building can thoroughly cover visible surfaces from a visitor. This would make a densely sampled version of Google Map's Street View on a building site.

The following are some challenges: Access may be blocked to some areas. The camera cannot go to high positions without a long pole or large drone. Also, if a moving rig is placed above the head of the person carrying it, the panoramic video includes a higher view than the eyes of a person.

And while all important locations may well be visited with the panoramic rig, recording all possible traversing sequences for circulation often is impractical and redundant. (Fig. 15, left)

Selecting and recording only major circulation routes would be a typical scenario for a common architectural representation project. The recording is incomplete and may include overlapping footages. On the other hand, the panoramic video clips can be edited conveniently just like regular video clips. Therefore, another scenario is to split the recordings of systematically corrected traversing sequences to a set of unitary paths linking each space to the next, and then reassemble appropriate ones back in order to create a desired traversing route. (Fig. 15, right)

Conventional video editing software can prepare this assembled clip manually, or a custom player can be developed to process the linking dynamically.

BUILDING DATABASE: CUMULATIVE APPROACH
In photogrammetric models, spatial forms and texture are captured. In panoramic video, the appearance of place and the sense of traversing it is recorded with viewing freedom and events. Any casual, low-cost, field-based work of recording architectural design can quickly deploy them complemen-
The processes and outcomes of these two technologies possess similar pragmatic and conceptual characteristics. Both media capture the physical environment at a moment in time, and multiple recordings on the same site can capture a variety of transient states of architecture. Both processes tend to amass recordings crafted at various levels of quality, and are convenient for rapidly producing valuable recordings by portion - albeit incompletely.

**Platform for collection and distribution**

How do we take advantage of the recordings of these unique characteristics? Creating a singular snapshot of entire architecture could be approached, but it is often difficult to achieve in one recording session or expedition, and without extraordinary planning, equipment, and arrangement.

One alternative approach is to develop a platform that stores and organizes surveyed information gathered over time by various people into a dynamic building database. Such a database includes photogrammetric models and panoramic videos sampled at various distance, interiors and exteriors, context and site, with and without furniture, and in various weather and lighting conditions. As new versions of recordings are created, the database grows.

The platform can be used with a browser that retrieves related recordings from the database and combines them to make a larger 3D assembly or concatenated video sequence. Alternatively, it can visualize selected recordings in a form suitable for comparative studies.

This approach accumulates historical data "bit by bit", and generates a composite, heterogeneous representation of a building, instead of a holistic single snapshot often targeted by a traditional surveying expedition that can be hit or miss.

**Prototype**

A prototype platform has been under development and tested in a surveying project to record the John Andrew Mansion designed by McKim, Mead and White, a building currently used as a fraternity and undergoing restoration. Through numerous short visits to the site near the MIT campus, approximately 100 recordings have been made in photogrammetric model and panoramic video format and stored in the database. (Fig. 16) The following summarizes this platform's design in progress.

- The database stores, organizes, and distributes photogrammetric models and panoramic videos, as well as geometric models, conventional videos, drawings, photos, notes, audio and other materials important in creating a representation of architecture.
- Each piece of data is attached with annotations and information tags to support searching and filtering the database for finding relevant spatial forms and visual contents.

- The platform provides an interactive graphic environment to allow examining the data in intuitive layouts. For instance, changed states such as before and after a restoration can be compared side by side in a 3D browser (Figs. 5 and 7), and a panoramic walk-through can be accompanied with a plan showing its path. (Figs. 2) Using a common game environment, this environment ports to multiple ecosystems and mobile platforms.
• The platform supports dynamic spatial assembly of 3D models (Fig. 16) as well as dynamic temporal and spatial editing of panoramic videos. (Figs. 13 and 15) This platform creates their composite representations in that the whole becomes greater than the sum of its parts.

• The platform promotes openness, participation and sharing by extending the access levels from personal, to groups, to the public. People can collaborate on representing a building or place, with an ultimate option for crowd-sourcing 3D captures and panoramic videos.

CONCLUSIONS

Photogrammetric modeling and panoramic video are emerging techniques that bring new forms of digital record of the real world around us. For architectural study, design, historical research, preservation, and renovation, these quick and inexpensive techniques enable the individual on site to capture the transient state of a building at a moment in its existence.

Field experiences reveal their unique characteristics and challenges for use. The very nature of a photogrammetric model as sampled mesh and texture stands in contrast to a geometric model assembled by discrete parts. Likewise, the very nature of panoramic video causes the discrete framing of subjects in a conventional video recording to disappear. Use of both technologies will produce abundant recordings with variations in quality, sensitivity to environmental conditions, and peculiar degrees of completeness.

The recordings collected for an architectural site can be stored within a database. It can store a growing number of recordings mixed with other conventional models and visual footages. The data can be sorted, combined, and provide useful heterogeneous representations for studying architecture.

Through such a platform, photogrammetric models and panoramic videos bring about a new understanding and novel virtual experience derived from real world physical objects and space. They will empower architectural representations in unique ways that conventional media alone were never able to deliver.

ACKNOWLEDGEMENTS

The examples come from the projects executed at MIT including academic workshops Palladio Digitale 2013-2014, Digital Heritage 2015, and a UAV research project in 2014 with funding from Autodesk, Singapore University of Technology and Design, Chi Phi Fraternity System, and MIT Department of Architecture. Projects on the architecture of Palladio and Scarpa were developed in collaboration with Professor Howard Burns, President of the Scientific Committee, CISA Palladio. Special acknowledgment to Professor Antonio Foscari, Director Guido Beltramini, CISA Palladio, Manuela Bedeschi, owner of the Villa Pisani at Bagnolo, Dr. Damiana Paterno, and Tom Holtey, Treasurer of Chi Phi’s Fraternity board overseeing John Andrew Mansion renovation project. Thanks to all students who participated in these projects with excellence. In memory of W. J. Mitchell.

REFERENCES

Bazin, A 1967, What is Cinema?, University of California Press, Berkeley

Choi, J 2014, Democratic Play: crowd-sourcing through games for architectural design, Master’s Thesis, MIT


Lowe, DG 1999, 'Object Recognition from Local Scale-Invariant Features,' Proceedings of the Seventh IEEE International Conference on Computer Vision, 2, pp. 1150-1157:

Palladio, A 1738, The Four Books of Architecture, Isaac Ware, London


