# Capturing with Community: An Online Collaboration Platform for Cultural Heritage Practice

TAKEHIKO NAGAKURA, Massachusetts Institute of Technology, USA
WENZHE PENG, Massachusetts Institute of Technology, USA
RACHELLE VILLALON, Massachusetts Institute of Technology, USA
DIEGO CORNEJO BARRA, Massachusetts Institute of Technology, USA
DANIEL TSAI, Massachusetts Institute of Technology, USA

This paper introduces an online collaboration platform as means to facilitate a new type of cultural heritage practice. Over the past decade, capturing a 3D form and its texture has become easily attainable by anyone with an inexpensive device such as a smartphone in conjunction with photogrammetric modeling software. Consumer use now offers an opportunity for heritage documentation projects. Photogrammetric capture projects can collect a large amount of 3D data sampled on site through the help of untrained novices, rather than with skilled professionals utilizing special scanning or measurement equipment. However, unorganized 3D captures by novices are often fragmentary representations of parts of a site in varied qualities from different times. Taking advantage of the captured data requires placing them properly in the spatial and temporal context of the site and relating them to each other. MIT Design Heritage is a prototype online platform for posting, assembling and sharing 3D captures under development at MIT. Similar to 3D GIS software, it is a database and 3D visualization tool, with a toolset tailored to help individual participants share a communal project and integrate data collected by professionals and novices alike. This paper discusses the design requirements for this system and its deployment to various scenarios such as a private class, a public workshop, or a community event with open participation. Topics include handling of massive, heterogeneous 3D captures, optimization with multiple levels of detail, automation of assembly phases, a permission and moderation scheme for sharing and editing, communication among participants, and data extraction and visualization for story-telling. Examples are illustrated by recent cultural heritage projects including one in Machu Picchu. These projects demonstrate use cases that support the possibilities for forming a unique community of participatory practice around an effort to digitally preserve an important heritage sites.

Keywords:
Digital Heritage, Architecture, Collaboration Tool, Multimedia Database, Photogrammetric Modeling

CHNT Reference:
Takehiko Nagakura et al. 2019. Capturing with Community: An Online Collaboration Platform for Cultural Heritage Practice.

## INTRODUCTION

This paper illustrates a strategy to integrate a large amount of heterogeneous 3D captures collected by experts and novices alike through the use of an online platform. The design of this database platform, for conveniently posting and assembling 3D captures as well as distributing them for sharing and presentation is introduced with its toolset tailored to facilitate collaboration of individuals in a cultural heritage project. A communal approach to cultural heritage practice is discussed. A case of a Machu Picchu documentation project demonstrates use of a prototype online platform with database and visualization tools for collaborating participants.

Recent technology has brought about a variety of tools to capture a 3D model, including both its geometry and its surface texture, by scanning a cultural heritage site and objects on it. A convenient method is photogrammetric

Author's address: Takehiko Nagakura, Department of Architecture, Massachusetts Institute of Technology, 77 Massachusetts Ave. 7-304, Cambridge, MA 02139, USA; Email: takehiko@mit.edu

modeling that uses software and photographs of an object shot from multiple viewpoints and translates them into a 3D model. Using a smartphone camera and free cloud-based software, photographs can be sent through an app and processed into a 3D model. This is an extremely low-cost option that requires minimum operational knowledge for the user. A professional DSLR camera could be used to record high resolution photographs for better details, and high-end computers with fast processors and sophisticated applications could be setup to process photographs without relying on the internet and cloud. A similar method recently developed, typically requiring a laboratory setting, is photometric stereo modeling that uses photographs of the objects that are shot under multiple, controlled lighting conditions. [Esteban et al. 2008] Photometric stereo modeling uses variations of the shades produced on surfaces and calculates the geometry of the target object. Another option uses a special sensor device called RGB-D or Depth camera, to record a 2D image of a scene with the information of the distance from the camera to points in the scene. [Henry et al. 2010] Example devices include game consoles such as Kinect and specialized smartphones such as those compatible with the Google Tango platform. Software is available to produce a 3D model of the scene from data collected by such devices. In comparison, LiDAR, or Light Detection and Ranging, is a professional surveying method that uses an expensive and highly precise machine with pulsed laser light, and is often utilized on a construction site for measurement. [Rottensteiner and Briese 2002] LiDAR devices can be combined with a photographic camera to produce a 3D model with colored texture.

Each of these methods is capable of producing representations of cultural heritage sites as 3D models with photographic textures, but with different precision, cost, labor and level of expertise to operate. Therefore, different scenarios and purposes of the documentation projects are considered for deployment.

A traditional archaeological expedition with a large budget within a concentrated period can be made by a professional team with LiDAR equipment, using a small number of well trained staff to produce a precise recording of a large, complex site. Experts are usually required to move and operate expensive, sensitive devices on the site, and collect data while carefully avoiding missed targets caused by occlusions such as areas behind columns and walls. Scanning high surfaces like the top of a roof, if found on the site, is difficult and costly. Raw data collected through multi-point LiDAR scanning produces huge point-clouds. Carrying, compiling and editing this data into a representation of cohesive 3D captured models convenient for distribution is a challenging and laborious process.

In contrast, a very different scenario can be offered by the photogrammetric modeling method. Because it only requires a set of photographs of the target from multiple viewpoints, the documentation project can be performed by participants with lightweight equipment. The minimum equipment is a smart phone with a camera or a consumer digital camera, which a typical tourist has. For a dark interior space, a tripod or a DSLR camera with a bright lens would be helpful. For high surfaces and large areas, a photographic pole or a small drone with a camera and gimbal would be helpful. A special compact LiDAR device is becoming available for field use, but these simpler photographic tools are much more portable and inexpensive. Therefore, an alternative possibility has opened up to use many cameras and distribute the task of capturing a large site to many participants and many visits, instead of using one expensive machine operated by a small number of specialists. The participants' visit to the site can be casual and distributed rather than a concentrated expedition schedule.

A version of this alternative approach assumes each participant visits the site at a convenient time with any available recording camera and captures one or more portions of the site. The cultural heritage site is represented in the long run by sharing these 3D captures made over time and by many people, placed on the same coordinate system imposed on the target site, piece by piece. An idea to build a database of such fragmented 3D captures that places them properly in the spatial and temporal context of the site and lets them relate to each other has been previously proposed. [Nagakura et al. 2015] The entries in this database would be completely heterogeneous, potentially massive, and asynchronously collected. On the other hand, in order to achieve any meaningful representation of the cultural heritage site, a new challenge for the practice of cultural heritage projects is to find a means to enable productive collaborations among participants, each of whom, makes a spatially and temporally distributed contribution to the project. There are many issues relevant to such modeling collaborations and use of the database. For example, if a 3D model has a missing part or a hole, is there a good way to allow communication between the participants so that a new visitor to the site can efficiently patch up that part in the model made by a previous visitor? What is a good way to visualize overlapping 3D captures or stitch adjacent captures together? If the same portion of the site is captured by multiple participants at different times and in different quality, what interface would allow a user to conveniently search the database and visualize an appropriate one? What is a good strategy to moderate the assembly of heterogeneous parts when they include some disagreement due to errors produced by some participant or transformation of the site conditions due to transition of time?

To answer these questions, the authors' team conducted a series of fieldworks at various cultural heritage sites during the past several years in conjunction with developing and testing a prototype online platform for cultural heritage practices of collaborative capturing. The following sections present precedents of online collaboration and sharing platforms, the design specification of the proposed platform for cultural heritage practice, and an example project conducted using the prototype platform. This paper focuses on the 3D capturing practice through the photogrammetric modeling method, although the platform can accommodate 3D models made from any method.

## PRECEDENTS OF COLLABORATIVE DATA CREATION AND DISTRIBUTION TOOLS

Before developing our platform, several precedents were studied with regard to the collaboration of data creation and distribution. They include platforms that share some similar features with our platform, such as model sharing, photogrammetric scene reconstruction, regional 3D databases, collaborative modeling and moderation and review systems.

### A. Sketchfab

Sketchfab is a model sharing web-based platform built for distributing and discovering 3D Game/VR/AR contents similar to what YouTube does with videos. [Bilton 2013] Creators can upload geometric and photogrammetric models. Collections can be made by a user, who pays to download these models that suit his/her needs. For a given model, the platform provides various presentation tools such as view-sequence editing, and rating and commenting. However, Sketchfab only supports posting and sharing by each individual model. It does not support multi-model assembly, patching and editing.

### B. The Building Rome in a Day Project

Building Rome in a Day is a massive crowd-sourcing and photogrammetric modeling project. [Agarwal et al. 2011] A system was developed that matches and reconstructs 3D scenes from extremely large collections of photographs acquired through public posting sites on the internet. This reconstruction procedure was made possible by the contributions of an enormously and specifically collected dataset, and a special, giant modeling computation engine.

### C. Google Earth/Map

Google Earth is a 3D-GIS tool produced by the Google Team. It is fully integrated with the Google Map product. The Google Team maintains the whole modeling database regionally and periodically, including maps, satellite images, and photogrammetry models. In addition, Google allows some supplementary contributions from the user end, such as geo-tagged photos and panoramic photos.

### D. Minecraft

Minecraft is a game that is capable of modeling through user-collaboration. [Persson and Bergensten 2011] It does not support uploading existing models. The models in the Minecraft environment are represented by cubes, or Voxels. Each move has to be made by addition or subtraction of a Voxel.

### E. Wikipedia

Wikipedia is a famous knowledge sharing platform mainly based on text. It is created and edited by volunteers around the world and hosted by the Wikipedia Foundation. Wikipedia provides a moderation and reviewing system so that each post may have different editing restrictions for different user groups. Edits can be reviewed properly before being applied to the official presentation article.

## SPECIFICATION OF PLATFORM FUNCTIONS

The proposed online platform has features developed for the work flow of cultural heritage practice and tailored with consideration to collaborative settings. They are largely categorized in three functional areas: Catalog, Patchwork, and Narratives.

A typical process of compiling multiple 3D captures of a site into a cohesive integrated model includes a workflow, or "pipeline", of multiple software packages. For example, each mesh geometry and texture mapping image derived from photogrammetric modeling software is first imported into a mesh editing software such as ReCap or MeshLab, where unnecessary geometry is cleaned up, the mesh size is decimated, and texture image is down-sampled to sizes appropriate for editing and manageable for online distribution. Afterwards, multiple captures are imported to a modeling application such as Rhino or 3DS Max, where each object is scaled, rotated, trimmed, and moved to an appropriate position with reference to an overall site often represented by a plan or map. Many professional photogrammetric modeling software packages allow the use of GPS tags in the photographs or physical markers placed on the site in order to scale, rotate and position the model on the standard global coordinate system. However, heterogeneous models made by participants with different tools and levels of expertize often do not adhere to such a coordinate system and require post processing alignment for assembly, and this step usually is prone to inherent error or information loss. If desired, those component models are further edited to make presentations for various purposes. For example, an interactive content is made by importing the data into game engine software. Editing then allows the user to set predefined viewpoints or walk-through paths as well as selective visualization of components such as an interior section of a building showing a plan without a roof. (Fig. 4, left)

In the platform proposed, this working pipeline is built into the three distinctive sections. The "Catalog" section provides a MeshLab-like inspector and editor for individually contributed models; the "Patchwork " section works as a modeling application similar to Rhino where users can work collaboratively to import and "patch" multiple models into an assembly [Bidarra et al. 2002]; and the "Narratives" section allows for querying and discovering data, visualizing and presenting stories, and making community feedback about the assembly representing the heritage site. The specifications of the platform are described below.

## A. Catalog

### Viewing UI

The model viewer is built as a tool of 3D photogrammetry model visualization for general users, including architects and novices. Its functionalities are common with many 3D viewers. However, making the viewer UI as intuitive as possible with comprehensive user-friendly viewing modes, will help a wide spectrum of non-professional users. The viewing system allows angle changing and zooming through mouse actions (scrolling, clicking, and dragging) or touch-screen. The viewing system has preset viewing settings, such as "double-sided/single-sided texture", "high res./low-res./no texture", "wireframe", "black/white background", and ambient illumination strength levels as well as "auto-orbit viewing mode" and "auto-bow viewing mode".

*Fig. 1. Some viewing modes of the model viewer, from left to right: single-sided texture, double-sided texture, low illumination, high illumination, black background.*

### Level of Detail (LOD)/Titled Model Visualization

Due to internet bandwidth limitations and client side computation resource limitations, it can be challenging to stream and render the whole model at once. A typical solution would use an LOD loading/rendering mechanism [Luebke 2003], which optimizes the loading and rendering resource to the current viewing angle and zooming level. Alternatively, tiled model visualization [Shi and Wu 2010] is also widely used in regional mapping tools such as Google Map. Resource consumption is reduced by splitting the model into tiles, and only visualizing the tiles within a certain range. The two mechanisms may be combined to further optimize the user experience.

Another visualization approach [Koller and Levoy 2005] available today to reduce transmission involves video streaming of only the rendered frames to the client display at the cost of the server composing the view from the complex models. However, the level of detail (LOD) model visualization issue is not only a visual distribution issue, but also an off-line assembly and processing issue. A composite model is comprised of one or more existing components, and newly created components to be added to the existing ones. Assembly and processing requires "uploading" of the source components to the processing server even if it is local. Light weight field equipment will likely lack processing power for uploading a huge set of components. The use of LOD schemes with lighter intermediate model-constructs will be helpful.

## Online/Offline Database

In certain use scenarios such as remote field expeditions, the platform has to be used offline. This requires a database management mechanism. [Guturu et al. 2003] In a typical scenario, a participant downloads and caches the project on the platform from the server to a portable device prior to the fieldwork. While working in the field and offline without access to the server database, the website running on a portable, cached database can still be used to accommodate new data such as captures made on the site; As the offline session ends, the offline data should then be merged with the main database on the server. Another scenario is that a participant collects captures on the site without previously cached project datasets, and instead creates a new project database for temporary storage and adds new captures to it. These field samples still need to be integrated with the main project database once the fieldwork is completed. This online and offline settlement feature makes sure that the whole platform is valid and keeps track of all the modifications made offline.

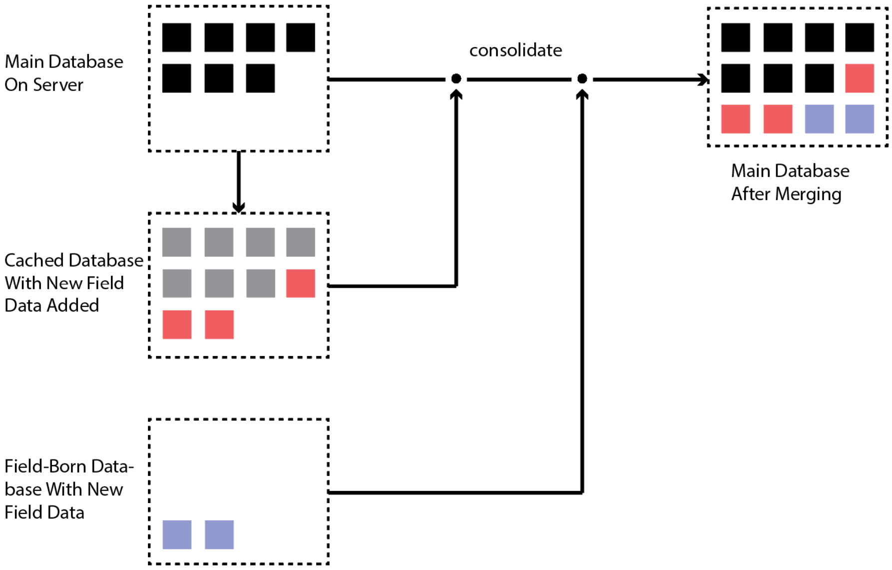


*Fig. 2. A database merging mechanism. Offline field capture data is stored in a cached database (red) or a new field-born database (blue), and reintegrated with the main database after the field work.*

## Source Validation

Photogrammetry models derived from photographs, and unfortunately, erroneous software processing often makes distorted models. Therefore, an option of including source photographs for validation is useful for the collaborating community members to identify problems. However, it is taxing to the database server to retain a large number of photographs used to generate each 3D model posted. By importing the photographs and computed corresponding camera positions from photogrammetric modeling software, participants can retrace and verify the model generation. Visualization of camera poses in the scene is also useful site condition information, because it often clearly indicates challenging positioning of cameras for a particular target. In addition, photographs from different models can be mixed and used for generating a new model and patching missing portions of existing models.

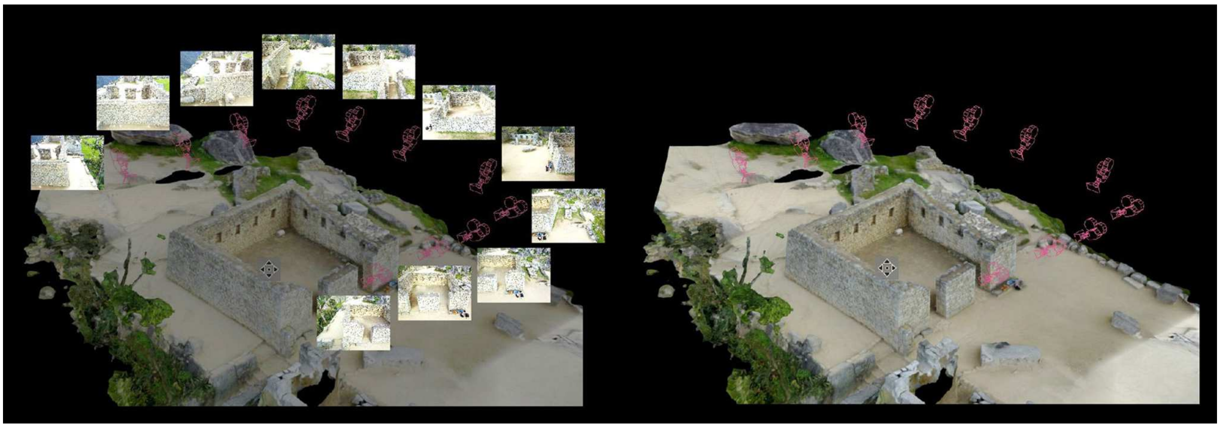

*Fig. 3. Validation photographs positioned in the scene together with the corresponding camera poses.*

## B. Patchwork

### Reference Model/Drawing

A user creating a new collaborative project initializes a Patchwork page with a small number of reference models or drawings, which can be any schematic presentation of the site. Incoming models are placed together as reference context, to have the correct scale, position, and orientation. This method allows a community project to proceed at multiple, disconnected locations in parallel. It also prevents accumulated positioning errors often produced by an alternative method of concatenating models one after another without a reference.

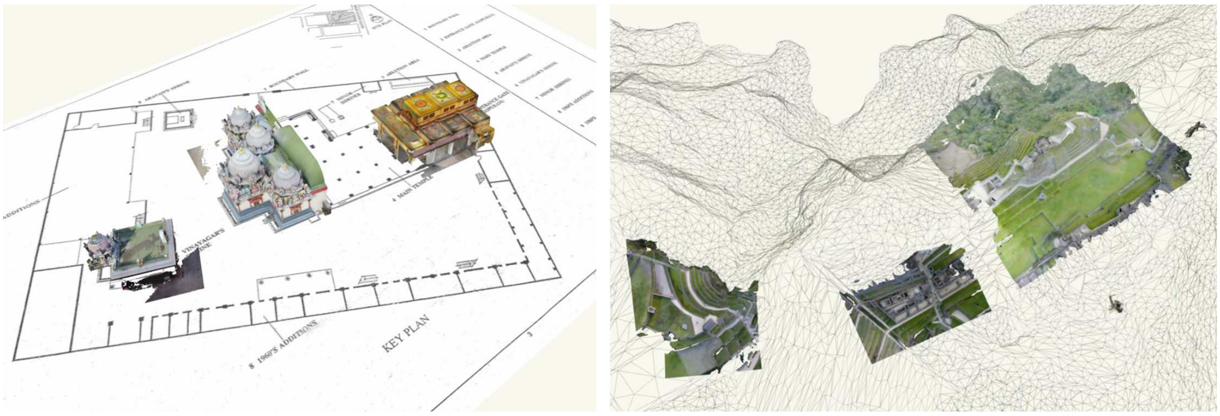


*Fig. 4. Reference drawing on the left, and reference model on the right.*

### Placement/Alignment of Models

The model placement tool is an important function of the Patchwork page. It makes the assembly possible by allowing a user to adjust the scale, rotation, and translation of a captured model to properly fit in the site. Eventually, the whole site is represented by numerous combined captures. Some convenient placement UI tools include (1) a widely used “gimbal” icon attached to the model that lets a user directly select, drag, rotate, and stretch it interactively by the mouse movement and eyeballing; (2) a “direction matching tool” that can be used to quickly orient the model, for example, by matching the tilted model’s floor area to that of the reference model or the plan, or by matching one model’s wall area to that of an adjacent model; (3) a “3-point alignment tool” that works by specifying 3 points from the target, 3 corresponding points from the reference model, and automatically matches the scale, rotation and position of the target with those of the reference. A reference model may ideally be a previously made approximate 3D model, a base map or drawing, or geo-coordinate based imagery (such as geo-tiff) underlay, if available. In a site with no such available information, the reference model starts as the first model, preferably as a light capture of the large context. The growing composite model serves as the reference for adding each new component thereafter.

In addition to these tools, more sophisticated automatic placement tools can be considered. When a new model is dragged into the assembly scene, a smart tool can snap it into the assembly properly by identifying the matching portions between the new and the existing models, calculating the transformation matrix, and translating, rotating and scaling the new addition to fit into the assembly.

### Optimized and Selective Model Display

The database may include many models overlapping with each other, if many participants worked around a popular location on the site to produce individual captures, or a similar location is captured repeatedly to produce different versions with varied sampling quality, lighting, weather and other conditions. To view the assembly including some heavily overlapping models in the Patchwork page, a convenient tool is to automatically filter the models in the scene instead of rendering all the overlapping ones at the same time to produce an incomprehensible visualization. The filtering can be performed by considering the viewing angle, distance to the object and quality of the mesh and texture map, and selecting the best model for viewing. Or the user may specify criteria for selection such as a particular condition for a preferred model.

### Roles and Permissions

Different roles and permission options are provided in the platform to ensure that the collaborative practice works with a sufficient amount of security and accessibility for the participants. (Table 1)

When a project on the platform is accessed, each user is either a “creator”, an “editor”, a “registered user”, or a “guest”. The creator is the one who initiated the project and has all the permissions to make changes to the project database. The creator can manage the project, for example, by approving another user to become an editor, who can add and modify 3D captures to the assembly in the project Patchwork page. Also, each project can be designated as either a “private” project, or a “public” project. A private project serves only for participants approved as the editors of the project, and the project is completely invisible from other people. A public project is viewable by the public guests. In addition, users registered to the platform can participate in the project’s SNS (Social Networking Service) area, and set personal preferences for parameters related to viewing such as camera position and visual appearance of models.

*Table 1. Permission for different roles. “✓”: full privilege; “Δ”: limited privilege; “✕”: no privilege.*

| | User Roles | **Edit Camera** | **Upload/Edit Model** | **View Demo Page** | **Manage Project** | **SNS** |
|---|---|---|---|---|---|---|
| Public Project | Creator | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Editor | ✓ | ✓ | ✓ | ✕ | ✓ |
| | Registered | Δ | Δ | ✓ | ✕ | ✓ |
| | Guest | ✕ | ✕ | ✓ | ✕ | ✕ |
| Private Project | Creator | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Editor | ✓ | ✓ | ✓ | ✕ | ✓ |
| | Registered | ✕ | ✕ | ✕ | ✕ | ✕ |
| | Guest | ✕ | ✕ | ✕ | ✕ | ✕ |

### Masking Box/Slicing tool

As a complement to the automated “optimized and selective model display” function, there is a tool to apply masks to any part of a specified model, or slice a part of the model away. With this function, a user can customize the visible extent of each model individually, such that the preferred portion of a model can be kept visible, while the impaired or unnecessary part is hidden. In a case that two adjacent models have an overlapping area, the better version for that area can be kept, while the other can be masked, and the two models are stitched together with a clean joint.

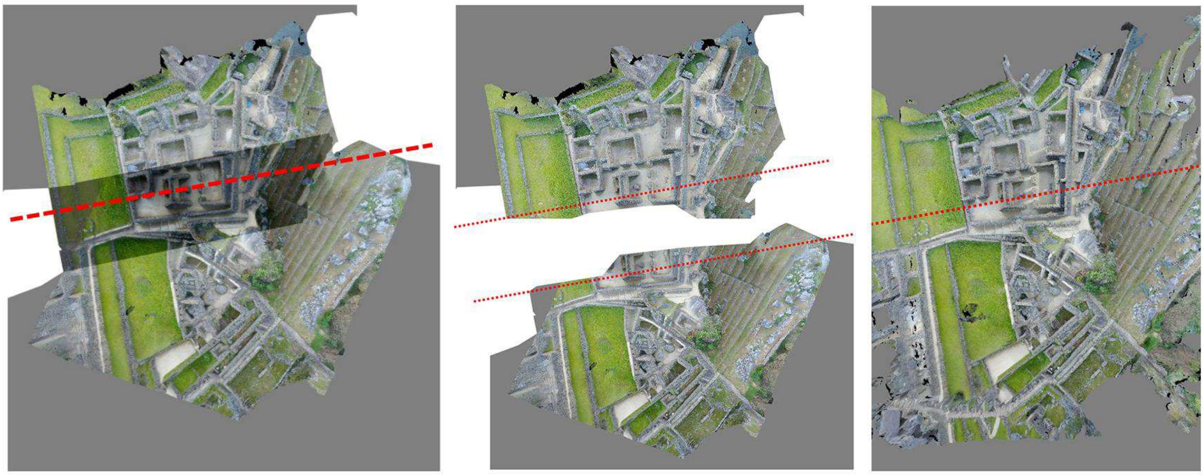

*Fig. 5. Masking/slicing a model. Left: before masking is applied; middle: masking is applied to each model at the same edge line (indicated in red dots) and different portions of the models are kept; right: model is combined with a clean joint.*

### Amend Mesh/Fill Holes

Several strategies are considered to remedy the problem of an incomplete portion or a hole in a captured model left by the photogrammetric modeling software. The masking box/slicing tool previously illustrated may be used to slice out a portion of another overlapping model and fit it into the hole. Alternatively, a more sophisticated tool can be developed so that the snapshots of this overlapping model around the area of the hole can be produced and used to complement the set of photographs originally used to make the model with the hole, and the new model can be generated without the hole. On the other hand, if the necessary photographs can be searched and found in the photographs submitted for validation of other models, they can be used to reprocess the model with the hole. And a quick and simple patching up of a hole is just interpolation of mesh geometry and texture from the area around the hole. An alternative approach to remedy the situation is to get assistance from people on or near the site to provide the necessary photographs to amend the model. An SNS function to help this is discussed in a following section.

## C. Narratives

### Quality Assessment

Each participant has interest and motivation to participate. 3D captures are made asynchronously under varying conditions of the site. The capture included in the assembly will be viewed and used by users with various intent. Therefore, it is not straightforward to estimate the contributive-worth of each captured model. At the same time, 3D captures made and placed by participants of different skill levels include ones that are heavily distorted, impaired or dislocated. Some captures made with deliberate sampling are likely to be useful and popular to more users than some other casual captures. Under these circumstances, two methods of quality assessment can be considered.

A moderation-based reviewing system uses a project moderator team or a public reviewing period for checking the uploaded capture before it is incorporated into the assembly publicly. This system approves and promotes the models that satisfy a shared quality standard much like Wikipedia system. The like/dislike voting system is a YouTube-like open rating system where each viewing user will be able to vote on the posted model, and the system will use the ratings to filter out the poorly-scored ones and optimize the search and viewing of captures to the high-scored ones.

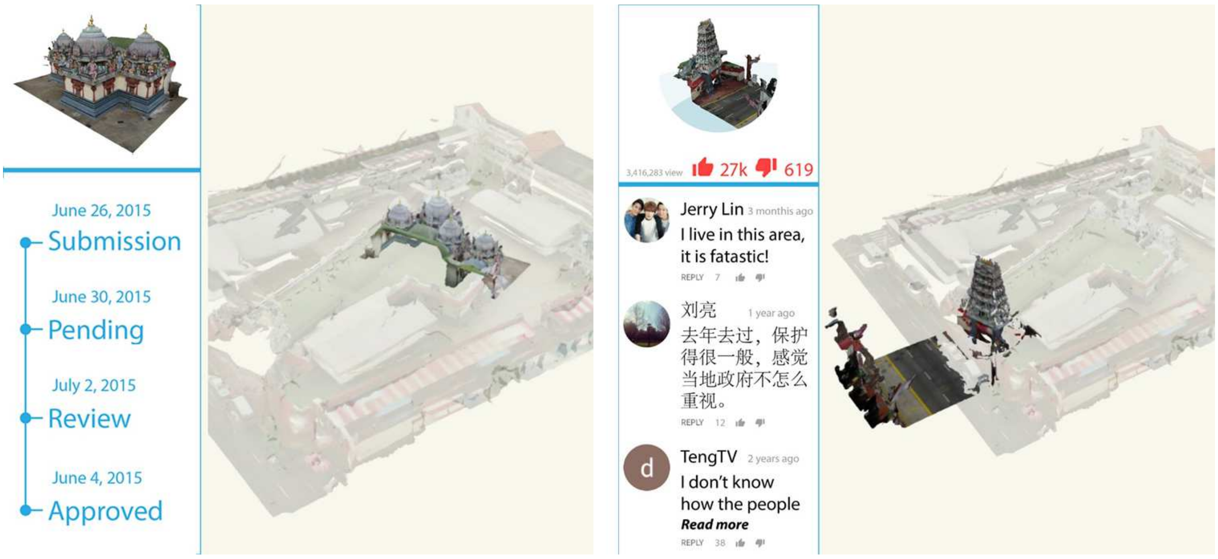


*Fig. 6. Two quality assessment options; left: moderation-based, right: like/dislike based.*

## SNS Comments/Conversation

The platform has a built-in SNS function for registered users to make comments associated with each capture model. The comments can be pinned to a specific part of the model for explicit reference, and can be used just as a production note, feedback made by other participants, or conversation to instigate collaboration among participants. For instance, a participant may produce a 3D capture that has holes due to missing photographs from the site. The model is posted on the platform with a mark-up comment asking others on or near the site to help take additional photographs necessary to amend the model.


*Fig. 7. Requesting missing information through an embedded SNS tool.*

## Query/Filtering Data

As many heterogeneous captures are accumulated in the database of the platform by collaborating participants, other users can search the database by querying and filtering the dataset by time, location, annotation, user rating and other criteria. For example, captures with high user ratings around a location of interest on the site can be selected and examined. The condition of the site during a certain time period can be studied by filtering the dataset to include only those made during that period.

### Storyline Creation

Visualization of project dataset in the Patchwork page can be used like an interactive version of a picture book or slide show for making stories about the cultural heritage site. The storyline tool can save a scene state that includes the viewpoint of the assembly and each captures' visual state such as visible/invisible/semitransparent as well as its transformation to indicate whether the model keeps in the position making the proper fit in the assembly or is in a temporarily moved/scaled/rotated situation by some specific amount. For instance, a user can move some captures out from the assembly and create an exploded view illustrating a composition of the site as seen in Figure 9 (right). Figure 9 (left-bottom) highlights a location of the site by making all adjacent models semi-transparent. Figure 10 (bottom) shows a series of orbiting views that can be set around an object of interest. The tool can keep a particular set of saved scene states together as storyline, and add transition effect from one state to another, as in the animated views that gradually moves the camera of one view onto that of that another. Figure 10 shows the presentation of a storyline that is made available to the public viewing in the Narrative section of the platform.

### Comparative Viewing

Users can specify two or more models and view them in a comparative viewing mode. In this mode, the control of the cameras is synchronized over all the models, and conveniently creates an interactive examination of similarity and difference of models from any desired view angle. For instance, 3D captures of the same location over time can be compared to study its transformation, or captures of the same location made by different participants can be accessed side by side for studying sampling qualities.

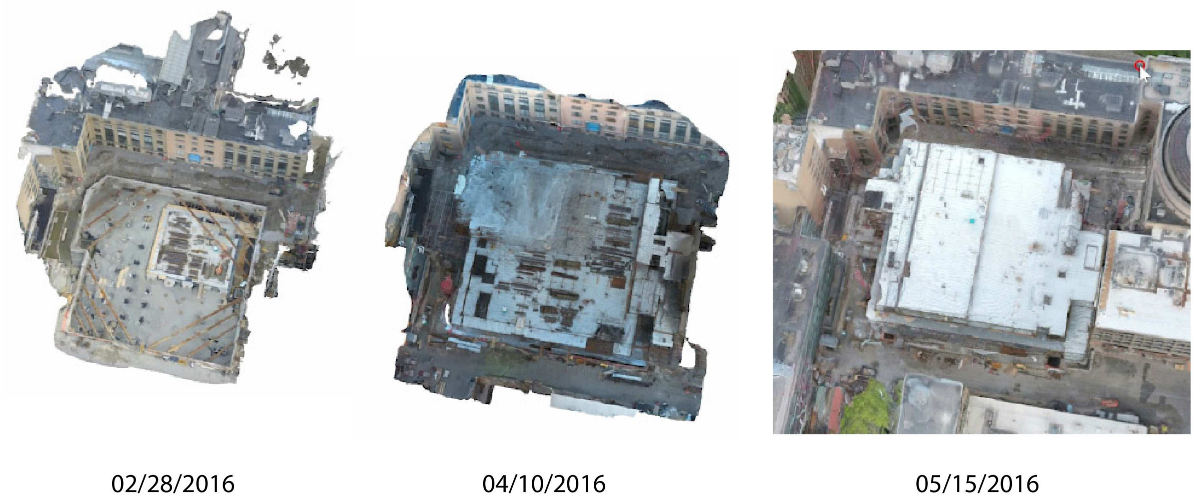


*Fig. 8. Comparative viewing of the MIT Nano building construction site captured in different months.*

## D. Multimedia Environment

### Media Other Than 3D Models

The supported media type includes but is not limited to audio, video (conventional/panorama), photograph (conventional/panorama), drawing and text. With documents of a variety of media mapped on the assembled model of the site, the platform can be a better all-around database for cultural heritage practice with a lot more flexibility as seen in Google Map. For instance, archival photographs geo-located on the site model assembly are significant for historians (Fig.10, stills 0 and 3).

### VR/AR Support

With the implementation of WebVR [Lv et al. 2011], both the Catalog model viewer pages and the Patchwork scenes can be directly rendered in VR format for a smartphone or HMD. Similarly to Sketchfab, the patching platform thereby also serves as an end-to-end VR platform to bring an immersive experience of cultural heritage sites without need to employ additional software.

Also, two AR application scenarios are considered: Onsite AR and offsite AR. Onsite AR can use objects in the site, like remains of the walls, as markers to match their 3D captures and bring up augmenting contents such as speculated reconstructions, audio guidance, and other annotations. For offsite AR, maps, photographs, drawings or scale models can be used as markers to bring up the corresponding 3D captures on a mobile screen, and let users interact with the platform contents. Similar attempts to use photogrammetry models combined with AR for cultural heritage presentations include Nagakura and Sung (2017).

## EXAMPLE OF USE IN THE FIELD STUDY AND EVALUATION

The authors' team has built a prototype, MIT Design Heritage Platform, in which the functionalities described in the previous section have been selectively implemented, while testing the platform under development in several different cultural heritage projects. The platform is based on the standard LAMP (Linux, Apache, MySQL and PHP) framework, with the WebGL-based Three.js package for 3D visualization and D3.js for UI designs.

An example use case of this prototype made by a small MIT research team is an ongoing project for Machu Picchu in collaboration with a university at Cusco and the regional government bureau of Peru. Initially, a team of five students led by a professor spent six days on this site situated on the mountain ridge 2430 meters above sea level, and then a follow-up expedition was made by a team of three students. As for the level of experience, one student has worked on multiple photogrammetric capturing projects with an extensive record of piloting drones, and other students have taken one digital heritage class that included an exercise of photogrammetric captures. The main area of Machu Picchu encompassing all the major archaeological settlement remains of granite stone structures measured approximately 200 meters by 500 meters. This area was recorded with ground-based photographic equipment for photogrammetric modeling as well as with panoramic video cameras for general documentation. Additionally, a larger area including the nearby Huayna Picchu Mountain and the mountain sides of Machu Picchu was recorded mainly by drone-based cameras. Approximately 26000 photographs were taken in the first trip and processed to create dozens of 3D models of various sections of the site as well as overall aerial models (Fig. 9, left-top). The second expedition captured approximately14000 photographs, and the 3D capture production is currently under way.

Prior to the expedition, the research team was not familiar with the Machu Picchu geography and had little idea about navigating around it. All arrangements for permissions and navigation was made by the local partners, a local government director and professors in a university in Cusco. The project started by preparing a very schematic 3D model of the site from various online information such as satellite views. A new project entry was created in the prototype platform with a Patchwork page and this schematic model was included as its first entry. On this page, the model was to serve as a template for placing fragments of 3D captures to be made on the site later. Sharing this before the field trip among the research team helped the members study the location through mapping published photographs and YouTube videos to estimated locations.

During the initial expedition, every evening, the team members processed photographs taken on the site into preliminary 3D captures. On the platform, each participant was registered as an editor member of the Machu Picchu project, and got an individual Catalog page on which captured models were accumulated with information of dates, creators and other notes for sharing. As the number of 3D captures grew, some were placed in the platform's Patchwork page, initially with reference to the template model and later to a schematic model produced from drone shots. The visualization facilitated an efficient planning vehicle as team members shared each individual's progress and placed each production in the context of the site and project goal.

Some sets of photographs failed to become meaningful 3D models due to a variety of issues such as inadequate exposure and focus quality, equipment problems, and interference from tourists on the site. Retake strategies for those were discussed around the in-context visualization to help plan on-site activities for the following day. Because access to power, internet and operational space in the field was limited, processing photographs into 3D models in situ and in real time was not feasible. Afterwards, in the hotel, the participants used individual laptops and processing software installed locally to make these preliminary models, since sending the large amount of photographs to a cloud-based modeling service on the internet was still unrealistic. Therefore, real-time feedback was not possible in the field, but the platform worked efficiently enough by keeping the try-error-retry cycle moving with adequate communication between participants without delayed feedback.

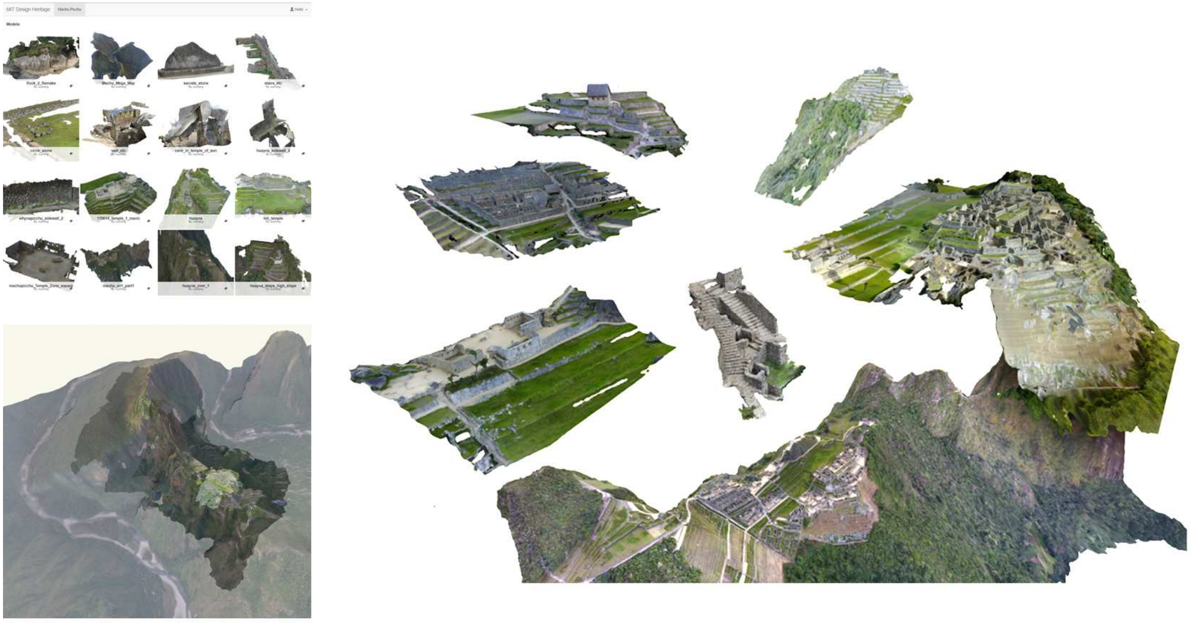

*Fig. 9. Use case: Machu Picchu; Left-top: a Catalog page showing the individual models processed and uploaded; Left-bottom: the Patchwork page showing the assembled Machu Picchu site; Right: Exploded-model-view built in the Narratives section of the prototype platform.*

The lack of adequate power and internet also affected on-site activities. It is possible to carry a limited number of small captured models on a smart phone or tablet and check the problems in the previous captures during the retake activity. However, the prototype platform lacked an implemented data caching function for carrying the project on a portable devise without internet, and so reviewing these models on site was not possible. Also, thanks to the local collaborators, the research team members were give permission to access nearly all parts of the site complex, including hard to reach locations and unmapped areas beyond the usual tourist paths. Communication between the sub-groups was often difficult because there were no street addresses and visibility was very low with narrow winding paths, steep topography and occasional cloud enveloping the site. The 3D viewing of the captured topographic model would have been very useful as a navigation and communication tool on the site just like a Google Map in an urban environment.

The first and second expeditions were six months apart, and more 3D captures have been produced during that period in the lab at MIT on desktop computers with high processing power and fast internet. These captures were placed in the context model on the platform, and many were carefully marked and commented as targets of follow-up recording. For instance, some captures still had holes and distortions, and some important areas between captures were missing. The participants of the second expedition carried those mark-ups to the site, and performed efficient recordings to amend the problematic portions of the models and remedy the entire site documentation.

Once a decent volume of data is accumulated, the platform is a database for anyone to browse, search and query 3D captures and tailor-make a representation appropriate for a specific purpose. In parallel to the expedition efforts, some research team members are using the dataset accumulated on the platform for making representations suitable for a variety of contexts including classroom material of a preservationist, historical research of Machu Picchu, and presentations for funding agencies. For example, one participant scholar has a project to make a narrative of the early history of Machu Picchu along with the archival diary and photographs produced by its first western explorer, Hiram Bingham. Intiwatana is the notable ritual stone at the top of a hill on the site, where Bingham's team took a well-known photograph in 1911 (Fig. 10, still 0). Using the platform function to make a storyline from the 3D capture assembly, a presentation (Fig. 10) was made of slides including the sight of the stone that matches the sight of Bingham's photograph, a view of the hilltop including the stone, an aerial view highlighting the location of the hilltop in Machu Picchu, and orbital close-up views showing the details of the stone. The slide presentation shows animation between the views to conveniently illustrate their spatial relations.

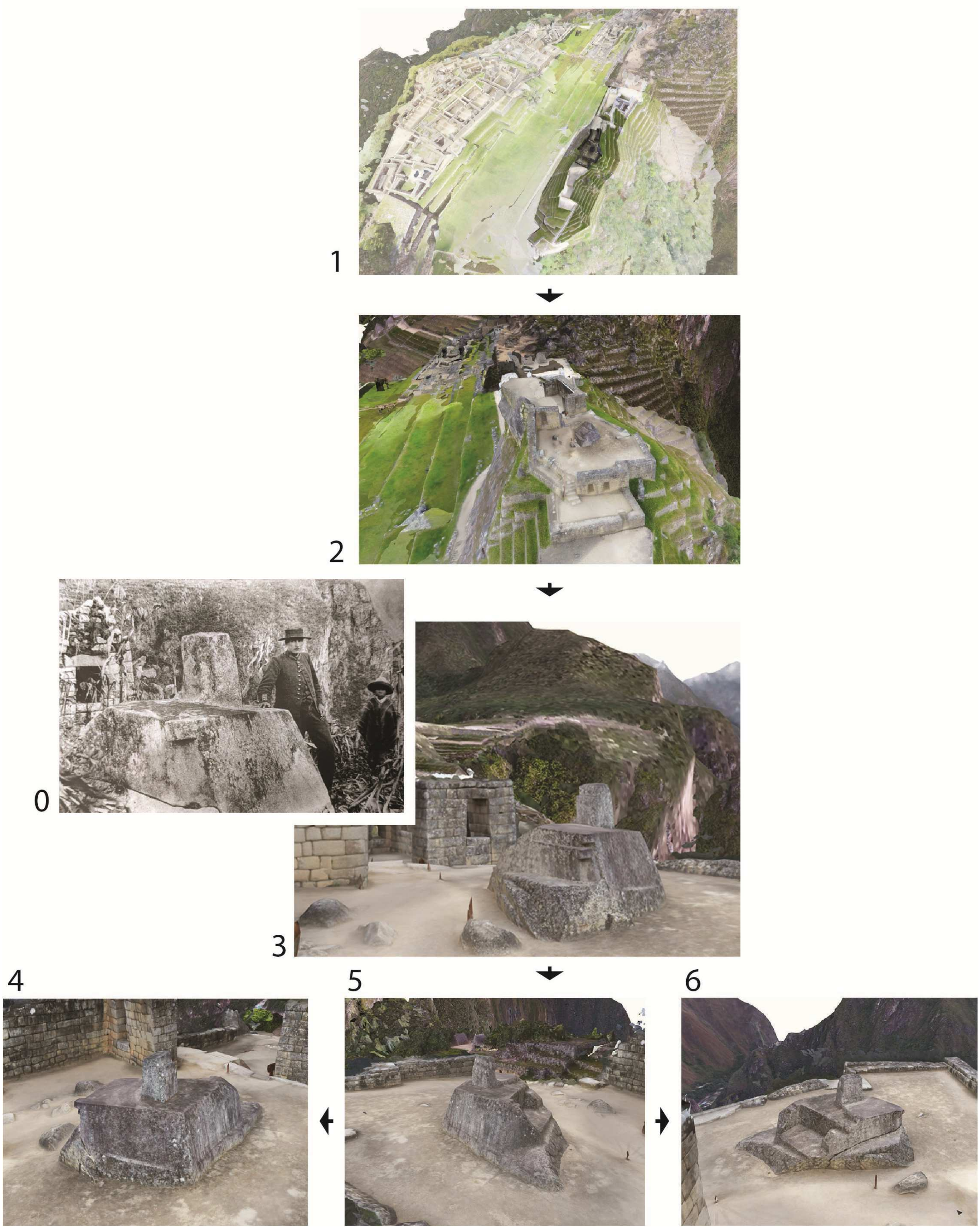

*Fig. 10. One storyline animation made using the platform's Narratives tool; 0: historic photograph taken in the 1911 Peruvian expeditions of Yale University led by Hiram Bingham III, showing the "Intiwatana", the top stone; 1: Intiwatana bird eye view on the site, the context has a opacity of 0.5; 2: closer look view of the hilltop; 3: view from the same vantage point as the historic photograph; 4,5,6: detailed view around the top stone.*

This online platform aims to support applications for collaborative cultural heritage practices for different kinds of communities. The Machu Picchu project is a simple but very common example situation in which members in a small research team are familiar with each other and the role of each towards a shared goal is well understood. In a classroom or workshop setting with somewhat larger participation, new community members may need to be assembled and configured quickly. As the independence of the participants increases in the community, an instructor can prepare and structure the project, and make active, continuous interventions to oversee and moderate members working on different parts. The most challenging case is a project with open participation, where a large number of participants with varied motivations and interests may join with little oversight by the initial project creator. Communication between participants may require self-initiated and self-managed configurations with little supervisory intervention.

In a number of cultural heritage locations, the prototype under development was tested in settings similar to the Machu Picchu project, and also in workshop settings of one to two dozen students. Based on these experiences, it is possible to estimate the scalability of the platform to much larger and more public community projects. For participant communications and management of access control, the specification of the functions for the prototype provides all the basic tools. Some of these are yet to be implemented in the current prototype. But it is a reasonable assumption that similar functionalities seen in widely used SNS platforms such as Wikipedia, Sketchfab and Google Map solves the issues typically seen in a large community participation, including evaluations by rating, conflict settlement through moderation, and enforcement of ethical conduct by a small group of people playing administrative roles.

A likely challenge for scalability, however, is the computational resource necessary for each participant to deal with the massive amount of 3D captured model data. In the current server-client model of the prototype platform, the first bottleneck is the internet bandwidth as a large amount of geometric models and texture images needs to be transferred over the internet to a participant’s device for viewing the large assembly. The second limitation is the graphic performance of the participant’s device which determines how much 3D capture data can be displayed at once with reasonable interactivity. The prototype platform does not implement any optimization or compression method. For distributing a captured model with a 100k mesh and 4k resolution texture via an ASCI OBJ file and JPG format with decent quality settings, the transfer requires about 15 megabytes of data. In a 4G connection over a carrier data network of 10 Mbps speed, this will take around 10 seconds. If there are 100 captures making up an assembly, downloading would take an impractical 15 minutes. However, with WIFI or gigabit LAN, this can be done in a minute or less. As for the performance of the display device, the prototype platform uses a WebGL-based software solution, and the graphic hardware system to execute it is another concern. With an ordinary office computer or a laptop computer without a dedicated high-end graphic card, loading one 100k mesh is not a problem. However, after loading a few dozen meshes, many computers encounter a situation known as "snag", which halts a browser due to lack of graphic memory capacity. A consumer smartphone is likely to hit this problem after several captures are loaded. A computer with a gaming GPU can hold more 3D models to display, but loading hundreds of mesh models at a time is problematic.

The specification of the platform discussed earlier indicated possible remedies for these problems such as using level-of-detail display methods and optimized and selective model display. There are also other approaches proposed for sharing of large 3D models such as the use of compressed or binary format for 3D capture files traveling over the internet, and the rendering of the view on the server side and distributing the view image over the internet. [Ma et al. 2013] Implementation of any one of these methods is a substantial contribution to the scalability of the platform.

Finally, two important issues are worth revisiting for improving the operation of the prototype platform. One is the portability of the platform in the field without internet and with limited computation power. As discussed in the case of the Machu Picchu expedition, it would be very convenient if a part of the project dataset on the server can be (i) cached onto a smart phone, (ii) used to browse the existing data and then (iii) used to add new information on the site. This loop would act to re-integrate the revised contents when the participants come back from the field. This scenario is similar to caching a map in a smart phone application. The second area of important improvement concerns the process of adding a captured 3D model to an existing assembly. Novices participating in an open project without knowledge of 3D modeling software will benefit from automatic positioning and scaling of captured 3D models fitting into the existing assemblies, and automatic or assisted stitching of the overlapping model. Such functionality has been previously discussed as part of the platform specification.

## CONCLUSION AND FUTURE DEVELOPMENT

This paper described a new type of online platform for cultural heritage practice, which aims to assist the process of collection and assembly of 3D captures in a database, especially through collaboration of multiple participants. The database platform functions include those to help collaborating participants comprehend and assemble complex 3D shapes captured, resolve discrepancies and remedy incompletion, work online and offline, and efficiently deal with a large heterogeneous dataset and its asynchronous production. In addition, the platform is designed to help users to extract appropriate data and compose different representations and narratives of the heritage site. The 3D capture database can also accommodate video, image, audio and other media, and let users make virtual or augmented reality output for enhanced interactivity and immersive experience.

A prototype platform was built in limited form. The benefit of its use in a relatively small research community setting was tested and elucidated by its field application in a Machu Picchu documentation project. This field test indicated that the scalability of the platform is a major challenge due to the limited computational resources available under ordinary project contexts. To cope with the limitation of internet communication bandwidth and graphic card display capability to deliver and visualize a large number of 3D captures, the proposed key features are required to be fully implemented with some further algorithmic development, before realizing the exciting prospect of use in a truly massive and openly collaborative practice of cultural heritage.

## ACKNOWLEDGEMENT

The prototype online platform discussed in this paper was created through funding support of International Design Center at MIT and Singapore University of Design and Technology. A part of examples included in this paper derived from the Cusco and Machu Picchu expeditions funded by MISTI MIT-Peru program and conducted in collaboration with UNSAAC and Cusco Culture Directorate of Peru. Authors thank Prof. Rene Navarro, Prof. Wilbert Salazar and Director Cesar Medina for all the support provided during the fieldwork sessions as well as Paloma Gonzalez for leading the expedition team and Eytan Man, Nikolaus Vlavianos, Xu Zhang, Takuro Kikuchi and Chan Liu for project operations. Another part of examples was produced through two of the Digital Heritage workshops held at and funded by Singapore University of Design and Technology. We thank Prof. Kang Shua Yeo, Prof. Shaohui Foong, Prof Saikit Yeung and all the students and teaching assistants of MIT and SUTD who participated in these workshops. We also thank Joanne Tsai and Evelynn Doone for additional and continuing supports of digital heritage research at MIT.

## REFERENCE

S. Agarwal, Y. Furukawa, N. Snavely, I. Simon, B. Curless, S. Seitz and S. Szeliski. 2011. “Building Rome in a Day”, Communications of the ACM, Vol. 54, No. 10, Pages 105-112.

R. Bidarra, N. Kranendonk, A. Noort, and W. F. Bronsvoort. 2002. "A collaborative framework for integrated part and assembly modeling." Proceedings of the seventh ACM symposium on Solid modeling and applications. ACM.

T. Bilton.2013. “Sketchfab raises $2M to make itself the 'YouTube for 3D design'”. https://venturebeat.com/2013/12/05/sketchfab-raises-2m-to-make-itself-the-youtube-for-3d-design/

C. H. Esteban, G. Vogiatzis and R. Cipolla.2008. "Multiview photometric stereo." IEEE Transactions on Pattern Analysis and Machine Intelligence 30.3, 548-554.

P. Guturu, J. Pal, T. E. Heaven, Z. Zhu, W. J. Jordan, H. Beffert and R: McNamara. 2003. "System and method for database synchronization." U.S. Patent No. 6,581,075. 17 Jun. 2003.

P. Henry, M. Krainin, E. Herbst, X. Ren and D. Fox. 2010. "RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments." In the 12th International Symposium on Experimental Robotics (ISER).

D. Koller and M. Levoy. 2003. “Protecting 3d graphics content,” Commun. ACM, vol. 48, no. 6, p. 74, Jun. 2005

Luebke D. P. “Level of detail for 3D graphics”. Morgan Kaufmann.

Z. Lv, T. Yin, Y. Han, Y. Chen and G. Chen. 2011. "WebVR——web virtual reality engine based on P2P network." Journal of Networks 6.7, Pages 990-998.

T. Nagakura, D. Tsai and J. Choi.2015. “Capturing History Bit by Bit - Architectural Database of Photogrammetric Model and Panoramic Video”, eCAADe.

T. Nagakura and W. Sung. 2017. “AR Mail from Harbin” In. Proceedings of SIGGRAPH '17 VR Village, Los Angeles, CA, USA, 2017. DOI: https://dx.doi.org/10.1145/3089269.3089274.

K. Ma, L. Fang and C. Li.2013. "Pipeline for network based server-side 3D image rendering." U.S. Patent No. 8,386,560. 26 Feb. 2013.
M. Persson and J. Bergensten. 2011. "Minecraft." [Computer software]. Stockholm, Sweden: Mojang AB. Retrieved from http://minecraft.net.
F. Rottensteiner and C: Briese. 2002. "A new method for building extraction in urban areas from high-resolution LIDAR data." International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences 34.3/A, 295-301.
N. X. Shi and X. Wu. 2010. "Method of client side map rendering with tiled vector data." U.S. Patent No. 7,734,412. 8 Jun. 2010.

*Imprint:*
*Proceedings of the 22nd International Conference on Cultural Heritage and New Technologies 2017.*
*CHNT 22, 2017 (Vienna 2019). http://www.chnt.at/proceedings-chnt-22/*
*ISBN 978-3-200-06160-6*

*Editor/Publisher: Museen der Stadt Wien – Stadtarchäologie*
*Editorial Team: Wolfgang Börner, Susanne Uhlirz*
*The editor's office is not responsible for the linguistic correctness of the manuscripts.*
*Authors are responsible for the contents and copyrights of the illustrations/photographs.*