CROWD-SOURCED SURVEYING FOR BUILDING ARCHAEOLOGY: THE POTENTIAL OF STRUCTURE FROM MOTION (SFM) AND NEURAL RADIANCE FIELDS (NERF)

L. Vandenabeele [1*], M. Häcki[1], M. Pfister[1]

[1] Chair of Building Archaeology and Construction History, IDB, D-ARCH, ETH Zurich, Switzerland -
(vandenabeele, haecki, pfister)@arch.ethz.ch

KEY WORDS: Structure from Motion, NeRF, Crowd-sourced surveying, Remote surveying, Building Archaeology.

ABSTRACT:

This contribution presents a simple workflow for surveying historical buildings and sites using crowd-sourced images. The proposed approach involves collecting large datasets of images from the internet using free plugins, followed by automatic image analysis and filtering using AI-based tools. 3D reconstructions are then created with Structure from Motion (SfM) and neural radiance fields (NeRF). To assess the reliability of crowd-sourced surveys, the 3D reconstructions are compared to high-precision laser scans of large medieval churches. In addition, the paper demonstrates the potential of this workflow in the field of building archaeology through detailed geometrical analyses of several iconic domes such as Hagia Sofia. By enabling remote and 4D surveys, crowd-sourced reconstruction methods open up novel opportunities for rapid, affordable and borderless research on cultural heritage.

1. INTRODUCTION

Around the globe, historic buildings and sites are being photographed by crowds of visitors who share their images online. Since the upload of the first pictures on the web in the early 1990s, a crowd-sourced collection of billions of digital images has grown exponentially. These images offer a wealth of data that can be used for photogrammetry-based 3D reconstruction of buildings and sites, even after they have been demolished. The global availability of this data allows for documentation of cultural heritage from remote locations.

The use of crowd-sourced images for scientific research has been hindered by several challenges, including flaws like inadequate indexing, variable coverage, and varying quality. As a result, these images have mostly been used as a last resort to document lost heritage. Moreover, their heterogeneity poses a challenge for the creation of reliable 3D reconstructions. In addition, the lack of simple workflows covering each step of the process, from image collection to 3D analyses, limits their wider application.

This contribution aims to investigate the potential of crowd-sourced surveying as a research tool in building archaeology. To this end, we present a workflow for rapidly collecting and filtering crowd-sourced images using free web scrapers and an AI-based software. We also evaluate the precision of crowd-sourced 3D models through cloud-to-cloud comparisons with three large-scale terrestrial laser scanner (TLS) surveys. In addition, we compare renders and point clouds obtained through state-of-the-art Structure from Motion (SfM) and neural radiance fields (NeRF), a novel technique that uses deep neural networks to visualize 3D objects. Furthermore, we highlight promising applications of crowd-sourced surveying, including geometrical analyses of five significant domes and the surveying of Hagia Sofia in an earlier state.

This workflow employs mostly free user-friendly software that do not require in-depth IT knowledge, making it accessible to a wide audience. While the two popular SfM software used in this research do not allow exports in their free versions, the NeRF software is open-source. The AI-based software used to filter pictures offers a 2-week free trial.

2. STATE OF THE ART

2.1 Crowd-sourced Structure from Motion

The beginnings of crowd-sourced SfM date back more than 20 years. In 2001, a team from the Institute of Geodesy and Photogrammetry at ETH Zurich attempted to reconstruct the Bayon Tower at Angkor Wat in Cambodia. 13 "tourist-type" photographs taken on 35mm film, distributed evenly around the tower, resulted in a successful reconstruction containing 46850 points (Niederöst et al. 2001).
In 2002, the same team worked on a replica of the Buddha statues in the Bamiyan Valley in Afghanistan, destroyed one year earlier by the Taliban. Two sets of images were used for a reconstruction. Out of 15 images found on the internet, four with sufficiently high quality were selected. In addition, 3 metric images were taken with a photo-theodolite camera. The final reconstruction used for the replica was made using the metric images and manual measurements. However, an automatic reconstruction using their own software and the four internet images was successfully carried out. Only finer details, such as the folds of the robe, were missing (Grün et al. 2002).

Several projects extended the idea of crowd-sourced photogrammetry by assembling datasets of hundreds or even thousands of images. Projects such as *BigSfM* or *Photo Tourism* in 2008 aimed to reconstruct objects on a city scale (Snavely, 2012, Snavely et al. 2008). Other projects focusing on cultural heritage with datasets of hundreds of images were carried out in 2015 by *HeritageTogether* and the *Mosul project* (Griffiths et al. 2015, Vincent et al. 2015). In 2016, another ETH team was working on reconstructions from datasets of 15000 to 30000 images collected from *Flickr*. They took advantage of the few geotagged images to geo-reference the reconstructed point clouds (Hartmann et al. 2016). At the same time, a team from the University of Budapest collected three datasets of 6000-

---

7000 images, which they filtered down to 60-1600 usable images using a duplicate finder and hours of manual processing. A reconstructed statue was compared with point clouds from a TLS, SfM with DSLR images and SfM with smartphone images. However, no qualitative metrics are given, only that the crowd-sourced reconstructions do not provide accurate geometry and result in low point density (Somogyi et al. 2016).

In 2020, a dataset of 200000 Street View images and 700 other images was used to reconstruct facades and automatically estimate building height. The point clouds were scaled with 2D vector data (building outlines) from satellite imagery and compared to vehicle-based Lidar scans. Detailed analysis of point cloud accuracy was performed for 16 buildings with an average error of the closest point pairs of 0.391 m (Wu et al. 2020). The same year, 900 tourist images taken over 8 years were used to reconstruct the Temple of Bel in Palmyra, which was destroyed by ISIS in 2015 (McAvoy 2020). Models of the site before and after the destruction are available online.

Finally, three recent publications discuss the methods and efficacy of crowd-sourced image collection, focusing on how to increase public collaboration (Cheng and Ch'ng, 2022a, 2022b). One of them also suggests using a reference dataset of geo-referenced images taken for SfM to provide a base to which the other images can be aligned (Jaud et al. 2022).

2.2 Neural radiance fields

Neural radiance fields (NeRF) are a novel method from the field of deep-learning-based computer vision, which enable to synthesize new views of complex three-dimensional scenes with the help of a neural network. Starting from a set of images showing the same scene from different angles with the corresponding poses, the neural network is trained to represent this scene as a radiance field. In this field, the density in every point in space is determined, as well as the radiance in every direction from that point. Using volumetric rendering, the radiance field can be sampled to create a novel view from any given point. As the density is non-binary and the radiance is direction-dependent, accurate rendering of transparency and reflections is possible. The key difference to the SfM pipeline is therefore the use of deep-learning-based radiance fields instead of point clouds or meshes to represent a three-dimensional scene and volume rendering instead of more common rendering techniques for visualization. This approach offers a range of new possibilities in various fields such as image processing or surface reconstruction (Mildenhall et al. 2020, Gao et al. 2022).

Since their inception in 2020, NeRFs have undergone rapid and diverse development. While most of the research focuses on improvements in quality and speed, some developments can also be observed in other areas such as 3D reconstruction and *sparse views*. Consequently, NeRF becomes an interesting complementary or alternative process to SfM. Several attempts at 3D reconstruction have been made, including *Block-NeRF*, a city-scale reconstruction using 2.8 million street view images (Gao et al. 2022, Tancik et al. 2022). A comparison of NeRF and Multi view stereo algorithms (MVS) for 3D reconstruction has been carried out, yielding results with very limited quality (Condorelli et al. 2021).

As NeRFs are at their core a novel view synthesis method, and in most cases not designed to extract accurate point clouds, they still have to find their way into mainstream 3D reconstruction workflows. Moreover, the initial alignment of images is usually still achieved through SfM pipelines, although replacing these

with deep-learning-based methods is part of current research. Attempts to use NeRF's as a 3D reconstruction alternative to SfM have been successfully achieved for small objects (Rakotosaona et al. 2023).

3. COLLECTING CROWD-SOURCED IMAGES

3.1 Web scrapers

To collect crowd-sourced images, seven picture-sharing platforms or image search engines were selected based on the overall quantity of indexed images: Bing, Flickr, Google Images, Google Maps, Pinterest, Yahoo and Yandex. After typing the name of a specific building in these platforms in English and in the local language (e.g. "Baptistery of Florence, Battistero di Firenze"), all images resulting from the search were downloaded with free plugins: *Download All Images* (for Bing, Google Images, Pinterest, Yahoo and Yandex), *Image Downloader Continued* (for Google Maps) and *Flickr-scraper* (for Flickr) (Jocher, Ultralytics/Flickr_scraper, 2020). With these simple web scrapers, thousands of images could be downloaded in a few minutes. For tracking purposes, the images were renamed according to their source in order to count them after each step of the workflow.

The collected images (JPG and PNG) have sizes ranging from 3 KB to 30 MB. Most of the images have no metadata, except those from Flickr which often contain camera parameters. The largest files (above 15 MB) are usually obtained from Flickr, Google Maps and Yandex, whereas the smallest images come from Yahoo (rarely above 200 KB). A quick visual inspection immediately reveals that only a fraction of the images depicts the objects of interest. Some images are wrongly indexed (e.g. other building), some show objects in the vicinity (e.g. a statue), some are strongly edited and others are partly obstructed (e.g. selfies) (Figure 1). Hence, a great deal of selection work is required before attempting a 3D reconstruction.

Figure 1. Sample of internet images of the Basilica of St Anthony in Padua before filtering.

Three medieval churches previously surveyed with 3D laser scanner were selected to illustrate this workflow: the Basilica of St Anthony in Padua, the Church of St Philibert in Tournus and the Cathedral of St Peter in Lisieux. For each building, 6660, 3484 and 4251 pictures were collected, respectively. As these numbers are directly related to the popularity of the places, they show the limits of the approach: non-touristic places would hardly be reconstructed with crowd-sourced surveying.
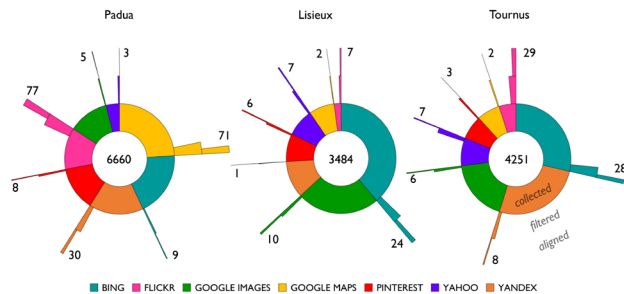

Figure 2. Evolution of datasets in a 3-step process (collection, filtering, alignment). The sizes of the initial and final datasets are indicated.

3.2 Filtering with AI

Faced with thousands of pictures for each case study, a manual selection was excluded in favour of automated image analysis. Therefore, the pictures had to be analysed and a fast content-based filtering had to be performed. The AI-based software *Excire Foto* was successfully applied following a 5-step approach: (1) import and analysis of all pictures, (2) selection of one picture showing the object/facade of interest, (3) content-aware similarity search, (4) removal of duplicates while keeping the largest files. Steps 2 and 3 were repeated to filter pictures from different sides of the buildings. In step 3, the AI tool to find similar photos based on content was used rather than the keywords attached automatically, which are too generic for a pertinent filtering (e.g. architecture, church, dome, religion, window). After filtering the downloaded images with AI, a small fraction varying between 2 and 5% of the initial datasets remained (Figure 2).

4. 3D RECONSTRUCTIONS

4.1 Structure from Motion

4.1.1 Alignment and reconstructed point clouds: Once groups of images consisting of different viewpoints of a defined case study were available, those were imported into the popular SfM software *Agisoft Metashape* and *Reality Capture*. Compared to a survey carried out by a single operator, crowd-sourced images are particularly varied in terms of camera, lighting condition and resolution. Between the two software, *Reality Capture* proved more adapted to this type of heterogeneous data. Indeed, groups of pictures that do not align perfectly are more appropriately divided into different components, whereas *Agisoft Metashape* tends to merge them. Moreover, the final texture of the point cloud is noticeably more uniform, which is valuable when generating orthophotos. In order to make this workflow as straightforward as possible, the second software was thus chosen for its suitability to the task in hand. In the three case studies, between 40 and 65% of the images previously filtered with AI were correctly aligned in *Reality Capture*. Although the main goal was to reconstruct the western facades of the three churches with images filtered accordingly, elements visible on the pictures' background were

reconstructed successfully (Figure 3). Using the Basilica of Saint Anthony as a case study, a second attempt was performed based on all pictures shot from different sides of the buildings, to test the possibility to construct complete models of large buildings in a single step.

Figure 3. SfM reconstruction of the Basilica of St Anthony from 203 crowd-sourced images of the western facade.

During the alignment process, different components were generated for different facades of the building but those could not be merged automatically. Indeed, the vast majority of pictures are taken from a few locations (e.g. main square, cloister open to public, street behind the church), leaving poorly documented parts in the survey that impede the alignment of all pictures in a single component. These limitations might be overcome through the manual identification of common features (i.e. markers) where pictures poorly overlap or by cloud-to-cloud alignment. This shows that in the case of large unbounded scenes, crowd-sourced surveying is strongly limited by the access of the public to different viewpoints. As shown in section 5.1, the approach performs very well in bounded spaces that are easy to access and document comprehensively.

4.2 Neural radiance fields

4.2.1 Alignment and reconstructed point clouds: To test the current potential of NeRF for crowd-sourced surveying in building archaeology, the software *Nerfstudio* was selected (Tancik et al. 2023). The open-source project offers a simple pipeline to create NeRFs from images. *Nerfstudio* is based on a Python framework that consolidates various NeRF techniques into modular components to ensure a balance between training speed and quality. It allows the training and rendering of NeRF scenes as well as the export of renders, point clouds and meshes via its web viewer. However, the initial alignment of images is not based on machine learning but rather on existing SfM tools like *Colmap*, *Agisoft Metashape* or *Reality Capture* (Tancik et al. 2023).

Nerfstudio has developed a model called *Nerfacto* that combines and optimises various methods from recent research such as *NeRF--* (Wang et al. 2022), *instant-NGP* (Müller et al. 2022), *NeRF-W* (Brualla et al. 2021) and *Ref-NeRF* (Verbin et al. 2021). The resulting model performs well for unbounded and real-world environments (Tancik et al. 2023), which is why the *Nerfacto* model was selected in the present workflow. Two datasets were used to investigate the potential of NeRF: crowd-sourced images from the western facade of the Basilica of St Anthony in Padua and the dome of the Baptistery of St

Giovanni in Florence. To compare the results with those of SfM tools, 3D point clouds and renders of both objects were generated. First trials revealed that pixel resolution and sharpness of input images are elementary for the quality of a NeRF. Therefore, the datasets had to be further filtered based on picture size and resolution to extract only the most detailed images. Doing so, 23 high-quality images were used for the western facade of the basilica and 27 for the dome of the baptistery. After aligning these images in *Reality Capture*, the camera poses were imported as a CSV file into *Nerfstudio*. The pictures and the camera poses were then used to train the neural network in *Nerfstudio*.

Although NeRF developments have mostly aimed at the creation of videos and renders, point clouds can already be exported from *Nerfstudio*. Additionally, from the model trained in the web viewer, various rendered views could be exported as images. At this stage, orthophotos cannot yet be generated.

4.3 Comparison of results

4.3.1 Point clouds precision: A qualitative assessment of crowd-sourced surveys requires dependable reference surveys. Hence, three TLS surveys (Leica RTC 360) of the churches of Padua, Lisieux and Tournus were used as a basis for comparison in the software *CloudCompare*. For consistency in the evaluation process, only one frontal setup position for each case was used to eliminate registration errors. Once scaled and finely aligned on the TLS surveys, the point clouds generated with SfM software showed a close match with the reference surveys (Figure 4). In the cases of Padua and Tournus, respectively 56 and 46% of the crowd-sourced surveyed points fall within a distance of less than 1.25 cm from their closest neighbours on the TLS survey (Figure 5). The more intricate geometry of Lisieux's church and the impossibility to capture the facade from its left side could be the reasons this value drops to 21%. As the models were scaled equally in all directions, these results also demonstrate that the general proportions obtained with crowd-sourced images are accurate.
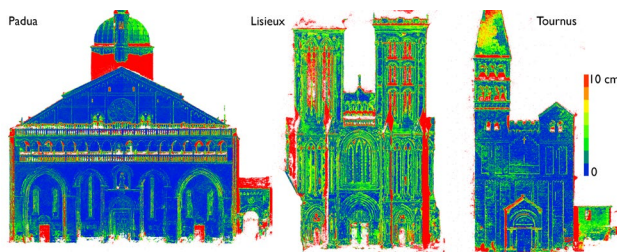

Figure 4. Cloud-to-cloud comparisons between reference TLS surveys and crowd-sourced SfM surveys of 3 churches.
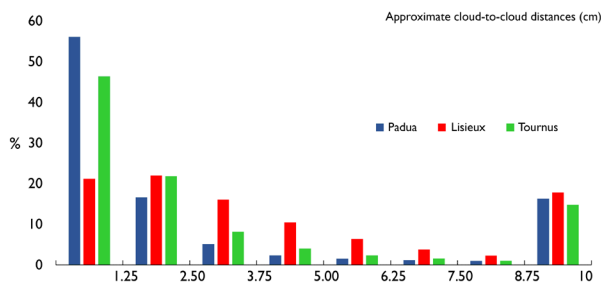

Figure 5. Deviations between reference TLS surveys and crowd-sourced SfM surveys.

However, point clouds obtained with NeRF are much less reliable. For example, the point cloud obtained for the Basilica

of St Anthony in Padua displays only 7% of points within a range of accuracy of 1.25 cm compared to a TLS survey. Indoor tests provided better overall results, as shown by the comparison of SfM and NeRF point clouds (Figure 6).
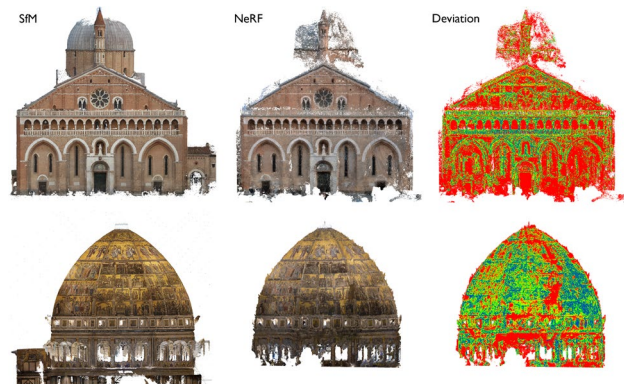

Figure 6. Comparison of SfM and NeRF point clouds generated with crowd-sourced images of the Basilica of St Anthony in Padua and the Baptistery of Florence.

Figure 7. Orthophotos of the churches of Lisieux and Tournus generated with crowd-sourced SfM.
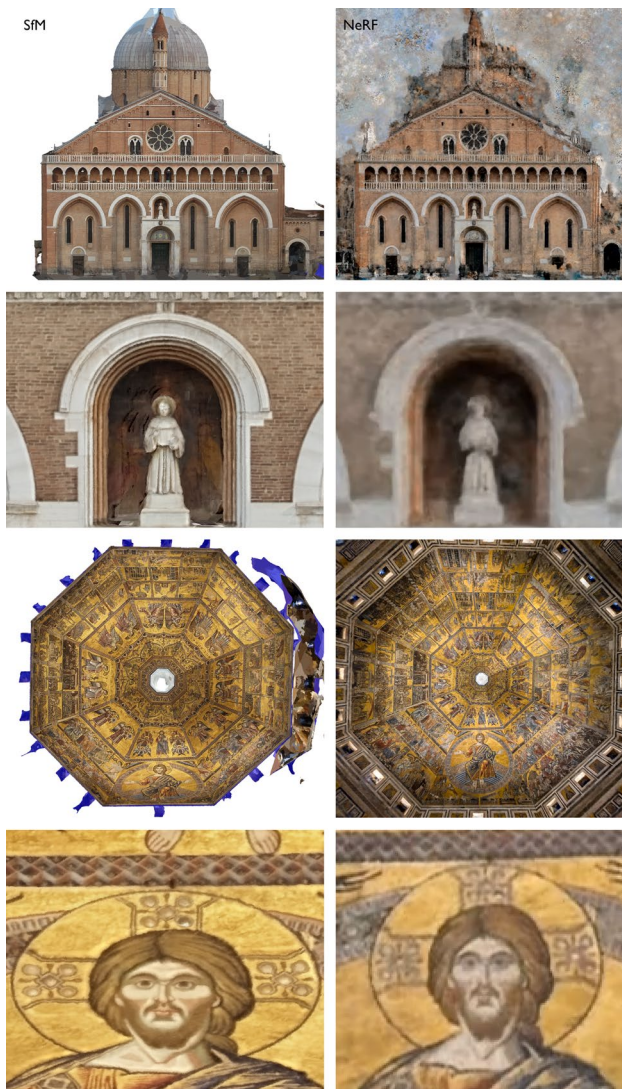
The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

Figure 8. Comparison of renders generated with SfM (orthophotos) and NeRF (perspective views) with magnifications.

4.3.2 Quality of renders: The comparison of renders obtained with SfM and NeRF tilts in favour of the first approach. Especially in the case of outdoor scenes, the results are significantly better with SfM software like *Reality Capture*. However, as shown by the renders of the Baptistery of Florence, in bounded scenes the NeRF outputs are almost as good as those of SfM (Figure 8). Unfortunately, corrected orthophotos cannot be generated at this stage in *Nerfstudio*.

Whereas the quality of SfM-generated point clouds will likely stay unmatched as long as NeRF software rely on the same SfM tools to align pictures, NeRF-generated renders and orthophotos will certainly be able to compete in a near future.

5. POTENTIAL APPLICATIONS

5.1 Remote surveying

Using the workflow presented in the previous paragraphs, we surveyed remotely five iconic domes in Italy and Turkey: the Baptistery of Ravenna, Hagia Sofia in Istanbul, the Baptistery of Florence, the Great Mosque (Ulu Camii) of Malatya and the Baptistery of Padua. Depending on the popularity of the building, between 70 and 400 relevant pictures were imported

for each dome in *Reality Capture*. The five resulting point clouds reproduce remarkably the detailed textures of the domes (mosaics or bricks) (Figure 9). Furthermore, due to the high overlap of pictures, the geometry of each dome is captured with great precision. Based on the previous comparisons with TLS surveys, one can estimate that the accuracy is in the order of one or two centimetres.

The possibility to accurately survey remote or hardly accessible buildings and sites based on online images could greatly facilitate comparative studies on buildings' typologies and geometries. For example, in just a few minutes, the design principles and deformations of the surveyed domes could be analysed and compared (Figure 10).
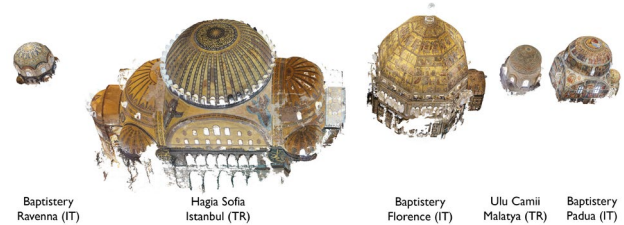

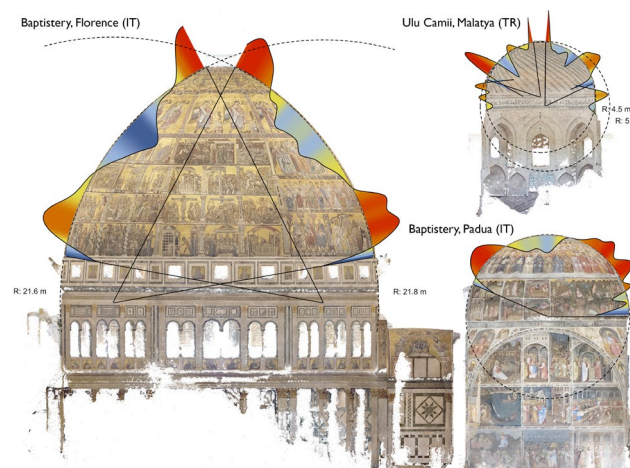**Figure 9.** Remote surveys of five iconic domes in Italy and Turkey.

Figure 10. Geometrical analyses of three crowd-surveyed domes showing their design principles and deformations.

5.2 4D surveying

Pictures available on the internet span over two decades at least. Consequently, crowd-sourced surveying offers the chance to reconstruct earlier states of buildings and sites. Depending on the number of pictures that can be collected and the identification (or interpretation) of capture dates, the temporal evolution of specific places can be traced through 3D reconstructions.

To illustrate this prospect, we choose the dome of Hagia Sofia in Istanbul, which has been almost permanently hidden by scaffolding in the last decade. Based on the combination of different TLS surveys from various international research teams, a deformation analysis of the central dome was performed recently. However, about a fourth of the geometry could not be studied due to dense scaffolding (Bianchini, 2020). To complete the survey, the AI-based filtering in *Excire Foto* was carried out in two steps. Firstly, a group of images of the dome of the building was created; then those showing a scaffolding were filtered out. The remaining 401 pictures were

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

imported into *Reality Capture* and 246 images were successfully aligned. The 3D reconstruction was scaled using the field surveys carried out by MIT researcher Robert L. Van Nice from the 1930s to the 1980s (Robert L. Van Nice fieldwork records and papers) (Figure 11). An ideal sphere was then fitted onto the resulting geometry and the deformations were highlighted using a *Grasshopper* script in *Rhinoceros*. The deformation map leads to the same observations as those obtained with TLS (Figure 12). On top of that, the entire dome is documented.
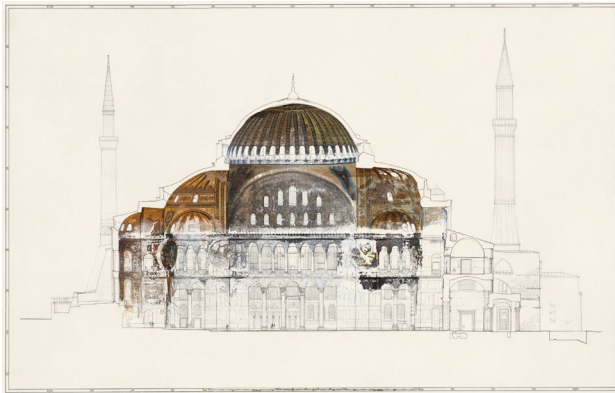

Figure 11. Scaling the crowd-sourced SfM reconstruction of Hagia Sofia's dome using a longitudinal section by Robert L. Van Nice (1930s-1980s).
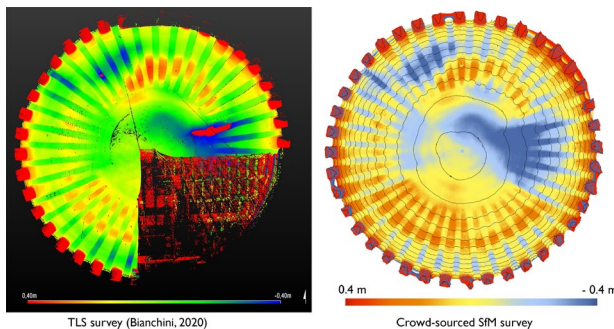

Figure 12. Comparison of a TLS survey and the crowd-sourced SfM survey of the dome of Hagia Sofia, showing its deviations from an ideal sphere.

6. CONCLUSIONS

The workflow presented in this contribution offers a simple way to document historical buildings and sites from crowd-sourced images. Leveraging on the maturity of SfM software, detailed surveys could be reconstructed based on heterogeneous datasets filtered with AI-based tools. The accuracy of the results was then assessed based on TLS reference measurements, which confirmed the reliability of the method. The most detailed and accurate surveys were obtained with state-of-the-art SfM software. The application of NeRF showed promising perspectives in terms of visualisation of historical buildings, especially indoors. Although at this early stage NeRF geometries do not provide usable models, AI-based tools clearly open opportunities to gain new insights into cultural heritage from unconventional sources.

ACKNOWLEDGEMENTS

REFERENCES

Bianchini, C., 2020. A Methodological Approach for the Study of Domes. *Nexus Network Journal*, 22, 983-1013. doi.org/10.1007/s00004-020-00526-9.

Cheng, D., Ch'ng, E., 2022a. Harnessing Collective Differences in Crowdsourcing Behaviour for Mass Photogrammetry of 3D Cultural Heritage. *Journal on Computing and Cultural Heritage*, 16, 1-23. doi.org/10.1145/3569090.

Cheng, D., Ch'ng, E., 2022b. Facilitating Situated Crowdsourcing of 3D Cultural Heritage via Asynchronous Virtual Collaboration. *Computer Supported Cooperative Work (CSCW)*. Preprint in review. doi.org/10.21203/rs.3.rs-2245108/v1.

Condorelli, F., Rinaudo, F., Salvadore, F., Tagliaventi, S., 2021a. A Comparison between 3d Reconstruction Using Nerf Neural Networks and Mvs Algorithms on Cultural Heritage Images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2021, 565-70. doi.org/10.5194/isprs-archives-XLIII-B2-2021-565-2021.

Gao, K., Gao, Y., He, H., Lu, D., Xu, L., Li, J., 2022. NeRF: Neural Radiance Field in 3D Vision, A Comprehensive Review. *arXiv*. arxiv.org/abs/2210.00379.

Griffiths, S., Edwards, B., Wilson, A., Karl, R., Labrosse, F., LaTrobe-Bateman, E., Miles, H., Moeller, K., Roberts, J., Tiddeman, B., 2015. Small Works, Big Stories. Methodological Approaches to Photogrammetry through Crowd-Sourcing Experiences. *Internet Archaeology*, 40. doi.org/10.11141/ia.40.7.2.

Grün, A., Remondino, F., Zhang, L., 2002. Reconstruction of the Great Buddha of Bamiyan, Afghanistan. *Photogrammetry and Remote Sensing*, 58(3), 225-234. doi.org/10.1016/S0031-8663(02)00019-5.

Hartmann, W., Havlena, M., Schindler, K., 2016. Towards complete, geo-referenced 3D models from crowd-sourced amateur images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-3, 51-58. doi.org/10.5194/isprs-annals-III-3-51-2016.

Jaud, M., Dantec, N., Parker, K., Lemon, K., Lendre, S., Delacourt, C., Gomes, R., 2022. How to include crowd-sourced photogrammetry in a geohazard observatory - Case study of the Giant's Causeway coastal cliffs. *Remote Sensing*, 14, 3243. doi.org/10.3390/rs14143243.

Jocher, G., 2020. Ultralytics/Flickr_scraper: Release V1. Zenodo. doi.org/10.5281/ZENODO.3678807 (30 April 2023).

McAvoy, S., 2020. Temple of Bel - 3D Tourist Photo Reconstruction. *Digital Media Lab. UC San Diego Library Digital Collections*. doi.org/10.6075/J0SQ8XX4 (30 April 2023).

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., Ng, R., 2022. NeRF: Representing scenes as neural radiance fields for view synthesis. *arXiv*. http://arXiv:2003.08934.

Niederöst, J., Zhang, L., Grün, A., 2001. Generating a 3D model of a Bayon tower using non-metric imagery. In *ISPRS WG V/6*

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

Workshop "Recreating the Past - Visualization and Animation of Cultural Heritage", 123-128. doi.org/10.3929/ethz-a-004331583.

Rakotosaona, M.-J., Manhardt, F., Arroyo, D. M., Niemeyer, M., Kundu, A., Tombari, F., 2023. NeRFMeshing: Distilling Neural Radiance Fields into Geometrically-Accurate 3D Meshes. *arXiv*. http://arxiv.org/abs/2303.09431.

Robert L. Van Nice fieldwork records and papers, Dumbarton Oaks, Trustees for Harvard University, Washington DC. https://hollisarchives.lib.harvard.edu/repositories/16/resources/7 13 (30 April 2023).

Snavely, N., Seitz, S. M., Szeliski, R., 2008. Modeling the World from Internet Photo Collections. *International Journal of Computer Vision*, 80(2), 189–210. doi.org/10.1007/s11263-007-0107-3

Snavely, N., 2012, BigSFM: Reconstructing the World from Internet Photos, Cornell University. www.cs.cornell.edu/projects/bigsfm (30 April 2023).

Somogyi, A., Barsi, A., Molnár, B., Lovas, T., 2016. Crowdsourcing based 3D modeling. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B5, 587–590. doi.org/10.5194/isprs-archives-XLI-B5-587-2016.

Tancik, M., Casser, V., Yan, X., Pradhan, S., Mildenhall, B. P., Srinivasan, P., Barron, J. T., Kretzschmar, H., 2022. Block-NeRF: Scalable Large Scene Neural View Synthesis. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8248-8258. doi.org/10.1109/CVPR52688.2022.00807

Tancik, M., Weber, E., Ng, E., Li, R., Yi, B., Kerr, J., Wang, T., et al., 2023. Nerfstudio - A Modular Framework for Neural Radiance Field Development. *arXiv*. https://arxiv.org/abs/2302.04264.

Vincent, M. L., Gutierrez, M. F., Coughenour, C., Manuel, V. L.-M. B., Remondino, F., Fritsch, D., 2015. Crowd-sourcing the 3D digital reconstructions of lost cultural heritage. *Digital Heritage*, 1, 171–172. https://doi.org/10.1109/DigitalHeritage.2015.7413863.

Wu, J., Mao, J., Chen, S., Zhuoma, G., Cheng, L., Zhang, R., 2020. Building Facade Reconstruction Using Crowd-Sourced Photos and Two-Dimensional Maps. *Photogrammetric Engineering & Remote Sensing*, 86(11), 677-694. https://doi.org/10.14358/PERS.86.11.677.