Open Science and Preserving the Digital Cultural Heritage

1st Maja Mihajlović
Codaxy d.o.o.
Banja Luka, Bosnia and Herzegovina
maja.mihajlovic9@gmail.com

2nd Andrea Baćo
Codaxy d.o.o.
Banja Luka, Bosnia and Herzegovina
andreabaco07@gmail.com

3rd Mihajlo Savić
Faculty of Electrical Engineering
University of Banja Luka
Banja Luka, Bosnia and Herzegovina
mihajlo.savic@etf.unibl.org
0000-0002-0289-3645

4th Miloš Ljubojević
Faculty of Electrical Engineering
University of Banja Luka
Banja Luka, Bosnia and Herzegovina
milos.ljubojevic@etf.unibl.org
0000-0002-7409-0972

***Abstract*—Cultural heritage is increasingly important in the modern world, not only as an important part of both collective and the individual identity, but also as a significant economic resource in the tourism industry. Enabling efficient and simple access to digital representations of relevant artefacts is among the most valuable missions of a modern museum or institution of culture. In this paper we present the system that enables the users to generate highly detailed digital reconstructions of real-world objects from a series of photographs through the use of Structure from Motion approach. The system is containerized and suitable for both standalone and hosted use. In order to increase the long term value of the system we use open science principles and open-source software.**

Keywords—open science, digital cultural heritage

I. Introduction

We live in turbulent times. While the advances in areas of information and communication technologies are rapid and transformative, there is an ever-present danger of alienation and isolation facing each individual. Loss of identity and lack of historical connections serving as a real-world anchors have only been exacerbated by the difficulties caused by the current pandemic, but both predate it, and will outlive it. Preservation of cultural heritage is of paramount importance in the modern fast-evolving digital world and digitizing it provides more efficient access and preservation means to us. One of the dangers connected to the process of digitization is the prospect of the Digital Dark Age due to outdated hardware, software or file formats [1].

One of the means of solving this issue is to use open and well documented formats, readable and writable by freely available and usable software – open-source software, while employing geographically distributed and redundant storage and compute resources through the use of modern cloud-based systems. This approach is based on the notion of FAIR data which if findable, accessible, interoperable and reusable.

The FAIR principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals and organizations [2]. If data is easy to find using unique identifiers and easy to integrate using a formal shared representation of knowledge, then they have a greater value. Such data is easy to share and reuse because machines have the means to understand where the data comes from and what it is about. It also speeds up research, facilitates reuse in the scientific research, and enhances collaboration.

Another issue is related to relatively complex software used to produce, manage and access the cultural heritage data. This issue arises from several factors, from the use of legacy software that evolved for a long time and accumulated a significant amount of very specific functionality that only needlessly complicates the use in a modern environment, through the need to create workflows consisting of several independent software packages with questionable compatibility, to the use of file formats that are of limited use in the modern environment.

In this paper, we present the system developed as a part of the project of digitization of cultural heritage collection of the Museum of Republic of Srpska as a part of VI-SEEM and NI4OS-Europe Horizon 2020 projects [3][4]. The wider system enables both the museum staff as well as the general public to easily access the relevant data, while the narrower system described in this paper enables the reconstruction and post-processing of three-dimensional representations of a real-world object from a series of photographs through the use of structure from motion approach.

II. Design of the System

Using photographs and similar visual aids to document physical artifacts collected by museum has a long history and it enables not only a simple way to reference the object, but also a relatively crude way to conduct some of the measurements without access to the object itself.

As photographs are two-dimensional representations of real-world objects, it is understandable that loss of information is inevitable and measurements that can be executed are limited in nature and prone to errors due to perspective distortion, different planes of reference and measurement, etc.

For example, measuring the items captured in pictures is difficult and imprecise. One option is to place a black and white strip with measurement markers that are clearly specified, next to the object. An example is given in Fig 1. By looking at it, visitors can at least determine the approximate size of the exposed object. Unfortunately, due to the perspective distortion, the measurement error increases with the height of the object rendering this method inadequate for anything with significant thickness.

Fig. 1. Example of a photo documentation of a Museum Artifact

Using the 3D model it is possible to determine significantly more information other than size, such as pose, volume, or shape of the chosen object.

As the technology progressed, the use of dedicated 3D scanners became more widespread. First, the scanners were based on manual scanning of individual points, then they evolved to employ line lasers and cameras to scan a line at a time, and finally to structured light scanners.

Scanners based on line lasers can be usable outdoors in varying lighting conditions due to the fact that lasers can be relatively strong light sources and can easily overpower even the direct sunlight. The problem with their use is the slowness of the scanning process as they scan a line at a time and have to provide for a very precise movement of the scanning line as the difference between two adjacent lines defines the resolution in one direction.

Structured light scanners project a pattern onto the object while recording the deformed pattern with a camera module offset to the projector [5]. Based on the deformations, the scanner software can calculate the shape of the object. This process can be seen as a massively parallel line scanning process, but it is also much more robust as the light is not limited to simple lines and can resolve features that are difficult to deal with by using only line-based patterns. An example of such scanner is given in Fig. 2.

The issue with structured light scanners lies in the fact that they need a fairly well controlled lighting environment and, as such, are mostly suitable for indoor use. They are also aimed at being used for objects of modest dimensions, from several centimeters to few meters, and are rarely used for outdoor scanning of large structures. As such structures are very common in cultural heritage (monuments, sacral objects, archeological excavations, etc.) there is a clear need to provide a suitable solution for this problem.

Fig. 2. Zeiss COMET L3D 5M structured light scanner system

Due to rapid advancements in digital photography, modern cameras are practical, affordable and high-quality imaging devices usable in a wide variety of conditions and lighting environments. By obtaining a sufficient number of photographs of the object from different viewpoints it is possible to first calculate the relative positions of camera and then to reconstruct the shape of the surface of the object. As the photographs contain texture data as well, the produced model can be further enhanced by incorporating this data. This approach of digital reconstruction of a real-world object is called structure from motion (SfM) [6][7].

Motion from structure is most suitable for a multitude of images with a high degree of overlap, which encompasses the three-dimensional structure of an object viewed from a wide range of positions. Based on photographs from different angles, the software can reconstruct the 3D shape and the texture of the object.

There are several commercial desktop software packages and cloud services that can produce satisfactory reconstructions, but, keeping in mind the benefits of using open science and open source software, we have developed a containerized solution that is usable in both standalone manner (running on a user computer) and as a web based hosted solution requiring no additional software or high performance hardware from the user as the system uses the resources of ETFBL CC01 compute cluster for allowed users. The system utilizes several open-source projects such as VisualSFM [8], PVMS/CVMS [9], Poisson [10], etc.

A wide variety of imaging sensors can be used for SfM, from video stills, through low-grade compact digital cameras to professionally created photographs with carefully selected equipment. The primary requirement is well-exposed photographs of the feature(s) of interest. Whereas image quality and resolution are improved by using increasingly expensive modern digital camera models, images captured at the highest resolutions (e.g. over 12 megapixels) will almost inevitably need to be re-sized (with the consequent loss of image detail) to avoid lengthy processing times. [11]

Although no special hardware is required it takes a lot of patience and precision. Unlike structured light and laser scanning, this method of scanning is also suitable for extremely large objects such as buildings, monuments, etc.

When taking photographs, it is necessary to take into account the lighting of the subject. If the object is in the exterior, it is ideal to take photos when it is cloudy or heavily overcast, in order to use even and flat illumination and avoid shadows in the pictures. This is important for a number of reasons. First, since the camera has a finite dynamic range capability, it prevents the parts of the image to be crushed to black (too dark to distinguish the features) or blown-out (too bright to preserve any features) which is very important as the system uses surface features for the critical part of the processing pipeline, second, it provides a fairly uniform lighting needed for texture reconstruction as the shadows and bright spots would otherwise be permanently baked into the textures thus providing false data, a finally, it prevents the camera operator and equipment from introducing the false data into the scene by projecting their own shadows.

III. Results

Since our intention was to create a system that will be available for both hosted and standalone use and to be accessible from different environments, we decided to use Docker as a form of application packaging and operating system level virtualisation solution.

Docker provides the ability to package and run an application in a loosely isolated environment called a container [12].

Containers allow us to bundle an application with all of its required components, including libraries and other dependencies, and deploy it as a single file. Thanks to the use of container, we can be confident that the program will operate on any other Linux machine, independent of any specific settings the machine may have that differ from the machine used to write and test the code. Like all other tools used in this system, Docker is also an open-source solution.

The workflow is very efficient and simple as it consists of creating a project, uploading the photographs and starting the reconstruction process.

Structure from Motion component is based on:

- VisualSFM,
- SIFT - by using VLFeat [13],
- Yasutaka Furukawa's PMVS/CMVS,
- Screened Poisson Surface Reconstruction,
- Texrecon from mvs-texturing [14],
- MeshLab [15], and
- Nexus by Visual Computing Laboratory - ISTI - CNR [16].

VisualSFM is used for initial project setup, feature detection and matching as well as the first pass 3D reconstruction.

First, VisualSFM imports all the provided images and performs feature detection using the VLFeat software, which is open-source library that implements popular computer vision algorithms specializing in image understanding and local features extraction and matching. The main algorhitm is SIFT (Scale Invariant Feature Transformation) and can be executed on general purpose CPU or on GPGPU cards if they are present in the system. GPGPU provides for significantly faster execution but limits the availability of the solution as it requires a presence of a card that supports NVIDIA CUDA technology.

Then, it determines the pairs of overlapping images and forms sets of similar images, based on which it estimates the position of cameras (and other camera parameters, such as focal length and angle of view), and after that, based on the position of the feature and the relative movement of the camera, it does a sparse reconstruction of the object's points.

Based on the sparse reconstruction and photographs, it is possible to perform a dense point cloud reconstruction by following not only the feature but all the elements in the images and filling the space between the points of the sparse reconstruction. For the dense reconstruction, our system integrates the execution of Yasutaka Furukawa's PMVS/CMVS toolchain.

The output is a dense point cloud of the object. To obtain the surface it is necessary to perform a surface reconstruction. There are many algorithms, but we decided to use Screened Poisson Surface Reconstruction because, according to our experience, it gives good results in a wide range of cases.

This process produces a monochromatic mesh surface, which is acceptable for certain uses, but since we want a better quality and texture is very important in cultural heritage documentation, it is necessary to approach generating textures with the reconstructed object.

This is done by the Texrecon tool from the mvs-texturing package by projecting a photo onto the reconstructed object from a camera perspective and calculating its texture. Result of that is an output which contains a number of files (one OBJ for the object, one MLT for the texture definition, and several image files that represent compound textures).

The resulting object is very detailed, but it contains many artificial parts that may be removed by running MeshLab in a headless mode called meshlabserver, which executes the script we created and performs final mesh processing and refining.

The script starts by removing small isolated pieces of a specific size, then moves on to isolated pieces with a limited number of faces/vertices, and finally determines the color information of each face based on the texture (to get a model that is colored even when no texture is used in a separate file because textures can be extremely large). The reconstructed object is now in PLY format which is widely supported by a large number of different programs.

After the process finishes, the user is presented with the full output of the SfM pipeline as seen in Fig. 3 and Fig. 4. The user is then free to further process the generated object through the web based version of Meshlab software by choosing the "Edit" option, providing users with the most important and most frequently used tools in an easy to use package.

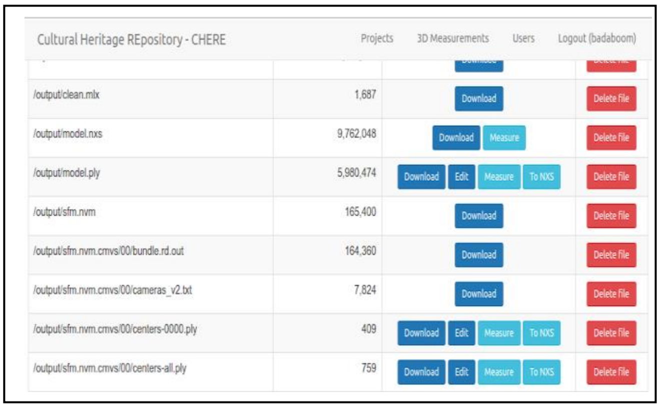

Fig. 3. Output of the SfM pipeline

The application also supports transforming PLY input files to NXS format, using the Nexus package. NXS format represents the mesh model in a streaming form, with first step containing the rough surface model of the object, and each successive step being more detailed.

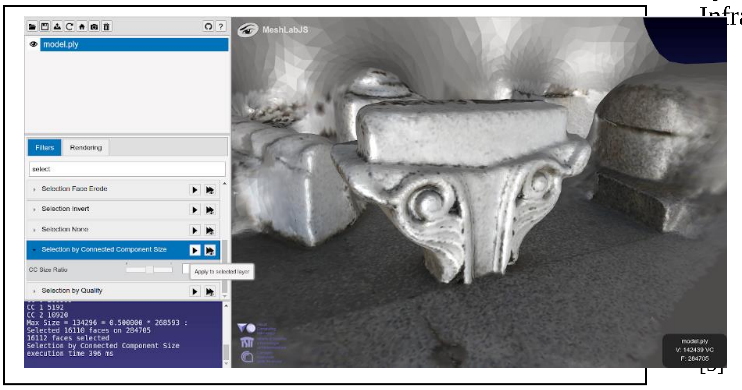

Fig. 4. Editing generated mesh in web version of Meshlab

When the user is satisfied, it is possible to measure various elements of the reconstructed object directly in a web environment that is based on 3D HOP viewer and enables users to rotate, move and scale the objects, to alter the lighting in order to better observe features, to create cross-sections, etc. This component is presented in Fig. 5.

Fig. 5. Measuring the reconstructed object in web based environment

The core of the system was adapted and used for measurements of dental plaster study models and compared with manual and photogrammetric measurements with all three methods producing measurements in agreement [17].

IV. Conclusions

Digitization and the possibility of online access to cultural material can facilitate and enrich participation in the cultural life of a large part of the population. The possibility of online access to library collections, for example, significantly facilitates the availability of books, while visiting museum websites encourages users to visit and participate in cultural life. The old goals of cultural policies can be achieved and enhanced using the new technologies. Therefore, work on the digitization of cultural heritage is extremely important. The development of information technologies in the world provides countless opportunities to preserve and increase the public availability of cultural heritage.

Acknowledgment

The research leading to these results has been co-funded by the European Commission under the H2020 Research Infrastructures contract no. 857645 (project NI4OS-Europe).

References

Saving European Archeology from the Digital Dark Age [Online]. Available: https://www.seadda.eu/. [Accessed: 19-May-2021].

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. "The FAIR Guiding Principles for scientific data management and stewardship", Sci Data 3, 160018 ,2016.

VI SEEM [Online]. Available: https://vi-seem.eu/. [Accessed: 19-May-2021].

NI4OS-Europe [Online]. Available: https://ni4os.eu/. [Accessed: 19-May-2021].

S. McPherron, T. Gernat and J. Hublin, "Structured light scanning for high-resolution documentation of in situ archaeological finds", Journal of Archaeological Science, vol. 36, no. 1, pp. 19-24, 2009.

[6] C. Wu, S. Agarwal, B. Curless, and S. M. Seitz, "Multicore bundle adjustment," Cvpr 2011, 2011.

[7] J. E. Nunes Masson and M. R. Petry, "Comparison of Algorithms for 3D Reconstruction," 2019 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), 2019.

[8] C. Wu, "Towards Linear-Time Incremental Structure from Motion", 2013 International Conference on 3D Vision, 2013.

[9] Q. Shan, R. Adams, B. Curless, Y. Furukawa, and S. M. Seitz, "The Visual Turing Test for Scene Reconstruction," 2013 International Conference on 3D Vision, 2013.

[10] M. Kazhdan, M. Chuang, S. Rusinkiewicz, and H. Hoppe, "Poisson Surface Reconstruction with Envelope Constraints," Computer Graphics Forum, vol. 39, no. 5, pp. 173â€"182, 2020.

[11] M.J.Westoby, J.Brasington, N.F.Glassera, M.J.Hambreya, J.M.Reynoldsc (2012). "Structure-from-Motion photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, vol. 179, Elsevier, 300-314, 2012.

[12] Docker Documentation. 2021. Docker. [Online] Available at: https://docs.docker.com/ [Accessed 6 June 2021].

[13] Vlfeat.org.VLFeat - Home. [online] Available at: https://www.vlfeat.org/ [Accessed 8 June 2021].

[14] GitHub. nmoehrle/mvs-texturing. [online] Available at: https://github.com/nmoehrle/mvs-texturing [Accessed 8 June 2021].

[15] Meshlab.net. MeshLab. [online] Available at: https://www.meshlab.net/ [Accessed 8 June 2021].

[16] Ponchio, F., [online] Vcg.isti.cnr.it. Available at: http://vcg.isti.cnr.it/nexus [Accessed 8 June 2021].

[17] M. Arapović-Savić, et al, "Photogrammetry based space analysis measurements in orthodontic diagnosis", Stomatoloski glasnik Srbije, vol. 65, no. 2, pp. 78-88, 2018.