Visual-Archaeo3D: Towards semi-autonomous graph-based videogrammetric reconstruction pipeline as web service

Marco Block-Berlitz, HTW Dresden, Germany
Dustin Heyer, HTW Dresden, Germany
Patrick Creutzburg, HTW Dresden, Germany
Lukas André Suthe, HTW Dresden, Germany
Paul Werner, HTW Dresden, Germany
Michael Bommhardt-Richter, HTW Dresden, Germany
Rohland Hendrik, HTW Dresden, Germany
Benjamin Ducke, DAI, Germany
Tim Kaiser, HTW Dresden, Germany
Christian Clemen, HTW Dresden, Germany

1. Motivation and introduction

Aerial 3D reconstructions by multicopter or underwater 3D reconstruction by mini submarines have become established for documentation in archaeological campaigns or for recording structures and artifacts[1]. The daily documentation work, the processing of image and video data in software packages with a focus on 3D reconstruction, requires experience and a variety of solution strategies when something doesn't lead to the desired outcome the first time. Here, from the user's point of view, there is often a need to make as few settings as possible to get satisfactory results. On the other hand, it is important to know that there is no predetermined reconstruction pipeline that can always deliver the best result for all input scenarios. Parameter adjustments, additional processing steps or rework are often necessary. The Archaeo3D software, published by Block et al. 2017a and developed in Java, combines various software packages that can be run console based. In this article we present Visual-Archaeo3D, a user-friendly web service that offers a graph-based user interface to Archaeo3D, the look and feel of which is comparable to packages such as AliceVision Meshroom. This allows more focus on data flow and more flexible parameter adjustment, while using an intuitive user-interface.

---

[1] http://www.archaeocopter.de

2. Theory and related work

Videogrammetry offers cost-effective and fault-tolerant data recording, allowing the UAV to capture data while moving, at relatively low resolution but with enormous numbers of strongly overlapping frames. This approach is inspired by real-time robotics and has the potential to be a first step towards real-time 3D processing and for providing at least reduced detail 3D models immediately after data capture. This will make it possible to verify on site, that the recorded material is adequate for reconstructing connected and gapless 3D site models. Higher quality and more detailed models can be obtained in the post-processing stage, in the office.

Videogrammetry (VG) can be simplified as an extension of photogrammetry (PG) by intelligent image selection (IS): VG = IS + PG. The approach of capturing and processing videos instead of photos has several advantages and disadvantages. In the case of photo recordings, geocoordinates can usually be stored as metainformation, which enables automatic georeferencing of the 3D models. The biggest disadvantage with videogrammetry is that the individual photos extracted from the videos are usually not geotagged and thus automatic georeferencing is not readily possible. However, this disadvantage can be solved satisfactorily in combination with photogrammetric images (see Fig. 1).

Fig. 1. Combination photo- and videogrammetry: As part of a 2018 documentation campaign in Kharabalgasun and Karakorum, the Erdene Zuu Buddhist Monastery was fully captured through a combination of traditional still photography and videogrammetry as shown in Block-Berlitz et al. 2021.

2.1 Reconstruction pipeline in Archaeo3D

In our experience, videogrammetry is the more error-tolerant, less expensive, and simpler, easier-to-use method for in-motion recording with multicopters or mini-submarines compared to pure photogrammetry. The JKeyFramer software, an automatic key frame selection tool, was one of the main results of the Archaeocopter project[2]. This tool uses videogrammetric methods for image selection and combines them according to the reconstruction objectives (fast calculation with reduced points to check coherence of data or complex calculation to maximize the level of detail). This tool was an important step towards fast 3d reconstruction at that time. In the meantime, the software pipeline has evolved to allow us to create fast preview models on site. Within the Archaeocopter project, the semi-automatic software Archaeo3D was developed to optimize and

---

[2] http://www.archaeocopter.de

control the whole reconstruction process. Videos and photos are automatically imported and processed. The software can rearrange or modify the pipeline modules and adjust the parameters, according to the current hardware and the image capturing situation.

A combination of VisualSFM[3], COLMAP (Schönberger and Frahm 2016), CMPMVS (Jancosek and Pajdla 2011) and Meshroom[4] formed the backbone of the processing chain in all Archaeocopter projects. The Archaeo3D reconstruction pipeline includes the following processing steps and software packages: 1. Data recording (GoPro Hero videos or photo sets), 2. Keyframe extraction (VLC5, MPlayer[5], ffmpeg[6], JKeyframer), 3. Image undistortion (OpenCV[7], JUndistortion), 4. Image enhancement (JResizer, JEnhancer), 5. Feature extraction (SiftGPU, JFeatureManager), 6. Sparse reconstruction (VisualSFM, COLMAP), 7. Dense reconstruction (CMVS+PMVS (Furukawa and Ponce 2010), OpenMVS[8]), 8. Compare or reduce point cloud (CloudCompare[9]), 9. SGM, Surface fitting (Poisson reconstruction (Kazhdan et al. 2006), CMPMVS, Meshroom, OpenMVS), 10. Producing orthoimages (CMPMVS), 11. Georeferencing, mesh cleaning (MeshLab (Cignoni et al. 2008)), 12. Integrate data into GIS (QGIS3). Additional software components such as JUndistortion for automatic camera calibration and JKeyFramer for automatic keyframe selection were developed and integrated. The pipeline automatically offloads processing to CPU or GPU depending on the hardware on which Archaeo3D is running. The number of parallel processing jobs is selected according to the available system memory. Our Archaeo3D pipeline allows us to create preview point clouds and quickly check them in the field for immediate validation of the results. The final reconstruction with Archaeo3D off-site, with more powerful computers, does provide more detailed results.

2.2 Graph-based user interfaces and usability engineering

Since a command line interface requires a lot of training for those unfamiliar with the concept, a more intuitive, user-friendly way to use Archaeo3D was required. First evaluations and mock-ups for the use of graph-based UI elements for the reconstruction pipeline Archaeo3D were conducted in the context of Patrick Creutzburg's bachelor thesis in 2021. The promising results of the user-oriented user tests for this UI concept and the first mock-ups based on current findings of usability engineering were the basis of the VideoBIM project. Graph-based UIs are elementary concepts of Blender[10] or Meshroom, for example, and often offer users an intuitive way to control complex workflows.

3. Visual-Archaeo3D as web service

With Visual-Archaeo3D we will offer a web-accessible platform, where individual 3D reconstruction pipelines can be combined and procedures for automated parameter adjustment can be used (see

---

[3] http://ccwu.me/vsfm/index.html

[4] https://alicevision.org/#meshroom

[5] https://www.mplayerhq.hu

[6] https://www.ffmpeg.org/

[7] https://opencv.org

[8] https://cdcseacave.github.io/openMVS/

[9] https://www.danielgm.net/cc/

[10] https://www.blender.org/

Fig. 2). Most of the time in Archaeo3D is spent watching videos from flights or dives in an external player and noting the partial sequences with start and end seconds that are relevant for processing. This has the advantage of not overloading the processing with unnecessary computational overhead. Unfortunately, the transfer of the data (video name and time sequences) is a source of error and must be done carefully. Therefore, a very important development is the built-in video sequencer, which can be used to play videos and mark the sequences very easily.
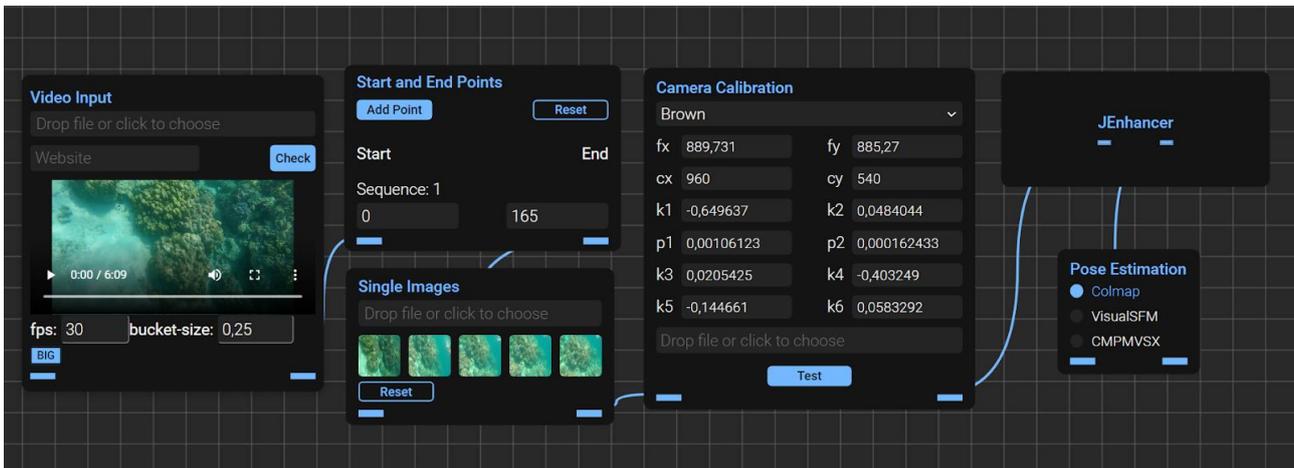

Fig. 2. Here we see a sample configuration for video processing: On the left, a video can be dragged and dropped, then played back and the desired sequences marked for processing. Furthermore, photos can also be loaded. From the selected sequences a new processing node is created, which extracts the images from the sequences. Afterwards, the JKeyframer is used to make a quick or detailed intelligent image selection, as desired. The images are then corrected by Camera Calibration, enhanced by JEnhancer and then the pose estimation is calculated in COLMAP.

React[11] forms the web frontend as a Javascript FrameWork and is used to read in the data and adjust the free parameters for the reconstruction pipeline. The data is first assembled into a LogFile and then loaded as a midend by MinIO[12]. MinIO is an open-source cloud storage provided through Docker. This is where all exchange data is stored for upload and download. After the data is uploaded, Java works as a backend and prepares the delivered data into a folder structure. There, the steps recorded in the LogFile are read and executed. After processing, partial results are uploaded back to MinIO and made available via the web interface.

4. Local software versus web service

Alicevision's Meshroom software represents a milestone in the field of freely available software for 3D reconstruction. The possible customization of the pipeline and the graph-based user interface make it possible to realize complex workflows in simple steps and provide good usability. Visual-Archaeo3D has set itself the goal of taking these positive features, embedding further functionalities by supporting easy video processing (including enhancement of scenarios under water), and offering them as a web service. A User-friendly pipeline and parallel processing are the focus of the processing, so that several users can have data packages processed on the server at the same time.

---

[11] https://reactjs.org/
[12] https://min.io/

When using 3D reconstruction software for a documentation campaign, its availability as a web service can be an important basis for decision-making. For example, during the documentation campaign in Mongolia in 2018 presented by Block-Berlitz et al. 2021, in which large areas of more than 50 km2 were captured with simple multicopters in a few days, the data was uploaded directly from the yurt after the flights to the web service of DroneDeploy[13] and the first results were examined within a few hours. This has the advantage of not having to carry large equipment for on-site computation. The approach of using 3D reconstruction software as a web service and to work without own hardware is becoming more and more popular, as it reduces the cost for mobile high-end hardware and at the same time eliminates the risk of damage or loss of equipment in the field. On the other hand, the processing pipeline of DroneDeploy's web service is not known, which does not allow the user to independently select algorithms or adjust parameters. Another obvious disadvantage of a web service is the stable internet connection required, especially for uploading large amounts of data, which can be critical in some remote areas but can be solved by satellite internet. Visual-Archaeo3D aims to combine the benefits of graph-based user interfaces of an individually configurable 3D reconstruction pipeline with the advantages of a web service, relieving the users from the necessity to maintain the hard- and software stack.

5. Conclusions and future work

The graph-based videogrammetric web service 'Visual-Archaeo3D' presented in this article is a part of the VideoBIM project funded by the SMWK as presented in Kaiser et al. 2022. In VideoBIM, the results of videogrammetric processing after pose estimation are investigated for the co-registration of image blocks within the context of Building Information Modeling (BIM). This allows point clouds to be embedded in BIM and, for example, to carry out building inspections. So far, web access to Visual-Archaeo3D and VideoBIM are still limited but can be unlocked upon request. We are working with the German Archaeological Institute's IT Department to provide a freely accessible server solution and look forward to presenting Visual-Archaeo3D and VideoBIM as an open-source software soon.

Bibliography

Block-Berlitz M., Ducke B., Rohland H., Franken C., Suchowska P., Batbayar T., Erdenebat U. (2021) Area-Optimized, Rapid UAV-Borne Recording of Medieval Heritage in Central Asia, Journal of Field Archaeology, Taylor & Francis, 2021, pp. 1-15, DOI 10.1080/00934690.2021.2007661

Block M., Dworsky C., Löw C., Seidl da Fonseca H., Gehmlich B., Wittchen D., Görsch N., Suchowska P., and Ducke B. (2017a). Underwater Videogrammetry with Adaptive Feature Detection at "See am Mondsee", Austria. *Studies in Digital Heritage*, *1*(2), pp. 547-565. https://doi.org/10.14434/sdh.v1i2.23202

Block M., Gehmlich B., Hettmanczyk D. (2017b): Automatic Underwater Image Enhancement using Improved Dark Channel Prior, In: Journal of Studies in Digital Heritage, Indiana University, Vol. 1, No. 2, pp. 566-589

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., Ranzuglia, G. (2008). Meshlab: an open-source mesh processing tool. Eurographics Italian chapter conference, 2008, Salerno, Italy, pp. 129-136

Creutzburg P. (2021) Graphbasierte UI-Konzepte zur intuitiven Anpassung einer 3D-Rekonstruktionspipeline, Bachelor thesis, HTW Dresden, Germany

---

[13] https://www.dronedeploy.com/

Furukawa, Y., Ponce, J. (2010). Accurate, Dense, and Robust Multiview Stereopsis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(8), pp. 1362-1376.

Jancosek, M., Pajdla, T. (2011). Removing hallucinations from 3D reconstructions. Technical Report CMP CTU.

Kaiser T., Clemen C., Block-Berlitz M. (2022) Co-Registration of Videogrammetric Pointclouds with BIM – First conceptual results, Conference on Measurement, Visualisation and Processing in BIM for Design and Construction Management II (MVPBIM2022), Prague

Schönberger, J. L., Frahm, J.-M. (2016). Structure-from-Motion Revisited. Conference on Computer Vision and Pattern Recognition (CVPR), DOI 10.1109/CVPR.2016.445