INTRODUCING AN AUTOMATED PIPELINE FOR A BROWSER-BASED, CITY-SCALE MOBILE 4D VR APPLICATION BASED ON HISTORICAL IMAGES

Sander Münster[a], Ferdinand Maiwald[b], Mathias Hofmann[b], Walpola Perera[a]

[a] *Chair for Digital Humanities (Image/Objects), Friedrich Schiller University Jena, Germany, sander.muenster@uni-jena.de*
[b] *Media Center, Technische Universität Dresden, Germany*

Abstract. The process for automatically creating 3D city models from contemporary photographs and visualizing them on mobile devices is now well established, but historical 4D city models are more challenging. The fourth dimension here is time. This contribution describes an automated VR pipeline based on historical photographs and resulting in an interactive browser-based device-rendered 4D visualization and information system for mobile devices. Since the pipeline shown is currently still under development, initial results for stages of the process will be shown and assessed for feasibility.

Keywords: urban history, photogrammetry, 3D modelling pipeline, browser based viewer

INTRODUCTION

A major interest of the project UrbanHistory4D (www.urbanhistory4d.org) is to develop scalable technological workflows to create 4D city models from historical images in order to enable virtual cultural heritage tourism, as well as to support information and scholarly research on urban history (Münster et al., 2017). Because currently the creation of virtual models of historical buildings and cities requires a lot of interpretation and individual modelling and is therefore labor-intensive (Pfarr-Harfst, 2016, Münster et al., 2018), we are aiming to automate workflows for creating models from historical images. Against this background, this article describes an automated VR pipeline based on historical photographs and resulting in an interactive browser-based device-rendered 4D visualization and information system for mobile devices. The described workflow ('pipeline') is currently still under development. However, prototypes of the individual components are already functional and have been tested on small samples. Results of the individual steps in the pipeline will be presented in this poster.

PIPELINE

The pipeline shown in Figure 1 consists of four steps. Step 1 and 2 recognize similar buildings and views as well as positions and orientation of photos. In step 3, building ground plots are vectorized. Finally, in step 4, a 3D model is created and visualized in a browser-based mobile virtual reality application.

- **Step 1: Historical Buildings Recognition:** The first step is to identify the same objects and views, creating the basis for reconstructing the image orientation. Since benchmark datasets for the recognition of buildings currently use large-scale sets of contemporary images, specific challenges in our case are the detection of non-landmark buildings which are hard to distinguish because many of them are highly similar (cf. Gominski et al., 2019), as well as the general sparsity of image data available. Against this background, we have created a test data set containing historical photographs of different objects – 220 time-varying images of a landmark building, 10 images of another landmark building, 10 images of a non-landmark building, 30 random historical images. First tests with the pre-trained network VGG16 (Simonyan and Zisserman, 2014) led to promising results for the recognition of the landmark structure. Regarding small samples, we expect algorithmic approaches to be promising. Therefore, we currently implement contrasting pipelines based on strong feature points as detected by Scale-Invariant Feature Transform (SIFT) algorithm (Lowe, 2004) and following outliers and inliers which are identified using the Random Sample Consensus (RANSAC) algorithm (Fischler and Bolles, 1981). RANSAC algorithm is being used in the pipeline for excluding the outliers and find a linear model which only uses inliers in its calculation. As a contrasting approach., another pipeline design currently under testing is based on DELF as combination of AI- and feature-based detection (Noh et al., 2018).
- **Step 2: Retrieving Image Orientation:** The second step requires retrieving the camera orientation automatically in order to map the historical images on a 3D city model. This step is still a critical issue and thus, is often performed manually (Schindler and Dellaert, 2012, Zawieska and Markiewicz, 2016, Bevilacqua et al., 2019). Consequently, for an automatic structure-from-motion workflow, a complete

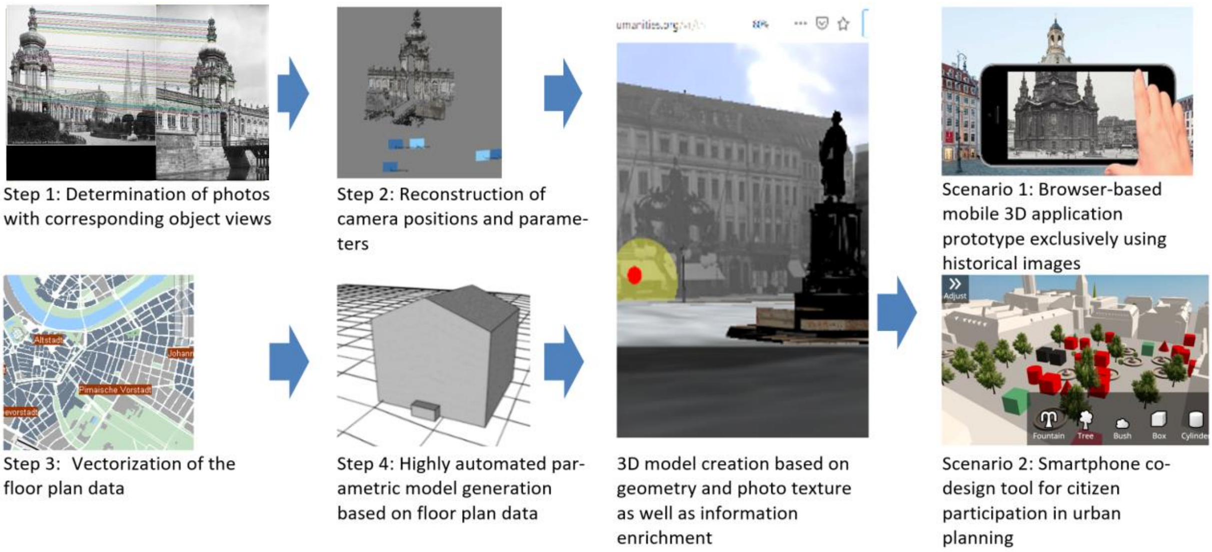

Figure 1 - Pipeline

pipeline has to be built—from selection of appropriate images via image orientation through to sparse model generation. Promising results could be achieved using additional recent images of the same object (Maiwald et al., 2017) in combination with specialized feature matching tools appropriate for historical images (Maiwald et al., 2019).

- **Step 3: Ground Plots from Images:** After the orientation and mapping of historical photographs introduced in the previous steps, the next task is to create basic geometric shapes of the buildings to enable facade texturing based on these photographs. As a basis for the 3D geometry generation shown in Step 4, we used cadaster plans of the historical city of Dresden. Outlines of building footprints were manually redrawn as polylines using ArcGIS. As a result, we obtained WGS 84-coded point data of the corners of the buildings. Since polylines currently require manual processing by clicking on the corners of each building, we started to automate building outline detection using machine learning (Oliveira et al., 2017).

  Step 4: Mobile VR Visualization: Recommendations for cultural heritage applications include allowing users to use their own mobile devices and avoiding the need for installing any additional application. Therefore, we decided on a browser-based application capable of running on both iOS/WebKit and Android/Chrome browsers. Geometry information is either stored as point information of the ground plot outline for simple level of detail (LOD) 1.5 (cf. Special Interest Group 3D, 2007)—block model and schematic roof—geometries, or as Collada files for complex geometries. In the case of point information, these are linked to a path and extruded to the unified height of the building. The central functionality is texture projection in order to ensure the dynamic compilation of textured buildings. This is realized by a dynamic shader, calculating a projection from the virtual camera frustum. Image metadata about the orientation and position of the photos are created within step 2 and loaded from a SQL database.

TEST

Since the individual components of the pipeline are already working at this stage, we are conducting several tests to validate and optimize their results. Step 1 has been tested for the VGG 16 based pipeline with the already mentioned test set of 220 landmark images. We currently run tests with the contrasting pipelines as well as for the sparse and non-landmark image sets. Step 2 is conducted with a set of currently 1070 historical photographs of Dresden obtained from the Deutsche Fotothek (deutschefotothek.de). A smaller subset, currently of 22 images, serves as test data for the texture projection as shown in Step 4. For the VR browser application in this step, a model of the inner city of Dresden was used. This model contains 30 building models with LOD 2.5—facade tiling and roof roughly modeled—as Collada data derived from the model of the municipality of Dresden (Stadt Dresden, 2018) for extant or rebuilt buildings. Another 105 buildings that are no longer extant were automatically modeled as LOD 1.5 from historical cadaster data.

USAGE SCENARIOS

As investigated empirically, it is not necessary for models to be very detailed to enable humans to identify specific buildings and urban structures or assess their properties (Münster, 2018). Therefore, despite the fact that models created by our pipeline are low in details, one scenario addressed by our work is virtual tourism (Scenario 1), in particular tour guiding. The main usage scenario is to support tourists by providing a visual impression of a historic situation at their current location (Niebling et al., 2018), as well as to provide access to textual information about single objects on demand (Mauro Ceconello, 2015). Our work also represents a useful technological basis for public participation and citizen science projects (Scenario 2). Building on the technological foundations described in step 4, we have developed a prototype for a smartphone application allowing citizens to participate in collaborative design processes in the context of urban planning projects. Participants are able to position, relocate, and resize abstract shapes (boxes, cylinders, cones), which they can assign to functions (e.g., office, living, traffic) and specific objects (e.g., trees, shrubs, fountains). This tool also facilitates co-creation by allowing numerous individuals to work collaboratively and simultaneously on a joint project, responding to and building on others' proposals.

CONCLUSION

Since the prototype pipeline for photographic image orientation is generally functional, an ongoing task is to further assess the steps involved in object recognition and image orientation reconstruction for larger image sets. The sketched pipeline is relevant for the Digital Humanities conference because it offers a highly automated creation of 4D city models. Those models are already a widely used basis in Digital Humanities, as they fertilize research e.g. of viewer effects, historical urban development and transition as well as enable intuitive 4D interfaces to access textual information (cf. Kuroczyński et al., 2019).

REFERENCES

BEVILACQUA, M. G., CAROTI, G., PIEMONTE, A. & ULIVIERI, D. 2019. Reconstruction of lost Architectural Volumes by Integration of Photogrammetry from Archive Imagery with 3-D Models of the Status Quo. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* XLII-2/W9**,** 119-125.

FISCHLER, M. A. & BOLLES, R. C. 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM,* 24**,** 381-395.

GOMINSKI, D., POREBA, M., GOUET-BRUNET, V. & CHEN, L. Challenging deep image descriptors for retrieval in heterogeneous iconographic collections. ACM Multimedia, 2019 Nice.

KUROCZYŃSKI, P., MÜNSTER, S. & PFARR-HARFST, M. (eds.) 2019. *Der Modelle Tugend 2.0: Digitale 3D-Rekonstruktion als virtueller Raum der architekturhistorischen Forschung,* Heidelberg: Heidelberg University Press.

LOWE, D. G. 2004. Distinctive image features from scale-invariant keypoints. *International journal of computer vision,* 60**,** 91–110.

MAIWALD, F., BRUSCHKE, J., LEHMANN, C. & NIEBLING, F. 2019. A 4D information system for the exploration of multitemporal images and maps using photogrammetry, web technologies and VR/AR. *Virtual Archaeology Review,* 10**,** 1-13.

MAIWALD, F., VIETZE, T., SCHNEIDER, D., HENZE, F., MÜNSTER, S. & NIEBLING, F. 2017. Photogrammetric analysis of historical image repositories for virtual reconstruction in the field of digital humanities. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,* 42**,** 447.

MAURO CECONELLO, A. S., DAVIDE SPALLAZZO, UMBERTO TOLINO 2015. Playing Design- Mobile Serious Games to Valorize Design Culture in the Urban Space. *Digital Heritage 2015.* Granada, Spain.

MÜNSTER, S. 2018. Cultural Heritage at a Glance. Four case studies about the perception of digital architectural 3D models. *In:* ALONSO, F. (ed.) *2018 3rd Digital Heritage International Congress (DigitalHERITAGE) held jointly with 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018).* San Francisco: IEEE.

MÜNSTER, S., KRÖBER, C., FRIEDRICHS, K., BRUSCHKE, J., MAIWALD, F. & NIEBLING, F. 2017. Urban History in 4 Dimensions - Supporting Research and Education. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* XLII-2/W5, 26th International CIPA Symposium 2017, 28 August–01 September 2017**,** 525-530.

MÜNSTER, S., KUROCZYŃSKI, P. & MESSEMER, H. Digital 3D reconstruction projects and activities in the German-speaking countries. EUROMED 2018, 2018.

NIEBLING, F., MAIWALD, F., MÜNSTER, S., BRUSCHKE, J. & HENZE, F. Accessing Urban History by Historical Photographs. 2018 3rd Digital Heritage International Congress (DigitalHERITAGE) held jointly with 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018), 2018 San Francisco. 1-8.

NOH, H., ARAUJO, A., SIM, J., WEYAND, T. & HAN, B. 2018. *Large-Scale Image Retrieval with Attentive Deep Local Features*.

OLIVEIRA, S. A., LENARDO, I. D. & KAPLAN, F. Machine Vision Algorithms on Cadaster Plans. Conference of the International Alliance of Digital Humanities Organizations (DH 2017) , Montreal, Canada , August 8-11, 2017, 2017.

PFARR-HARFST, M. 2016. Typical Workflows, Documentation Approaches and Principles of

3D Digital Reconstruction of Cultural Heritage. *In:* MÜNSTER, S., PFARR-HARFST, M., KUROCZYŃSKI, P. & IOANNIDES, M. (eds.) *3D Research Challenges in Cultural Heritage II*. Cham: Springer LNCS.

SCHINDLER, G. & DELLAERT, F. 2012. 4D Cities: Analyzing, Visualizing, and Interacting with Historical Urban Photo Collections. *Journal of Multimedia*, 7**,** 124-131.

SIMONYAN, K. & ZISSERMAN, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

SPECIAL INTEREST GROUP 3D (ed.) 2007. *CityGML Specification*.

STADT DRESDEN 2018. Virtuelles 3D-Stadtmodell.

ZAWIESKA, D. & MARKIEWICZ, J. 2016. Development of Photogrammetric Documentation of the Borough at Biskupin Based on Archival Photographs - First Results. 10059**,** 3-9.