Recommendations for Image Acquisition for 3D Room Scanning

Adam Kaczmarek, Mariusz Szwoch, Dariusz Bartoszewski

Abstract—Proper image acquisition plays an extremely important role in 3D scanning of rooms, determining the completeness and correctness of the created model. Any mistakes made at this stage can cause a need to supplement the set of photographs or even to repeat the photo session, e.g. in case of changing the lighting conditions. This paper presents a set of recommendations minimizing the risk of making a mistake at the stage of acquiring photographs of interiors. These recommendations were based on numerous experiments carried out as part of the STERIO project which main goal is to develop a complete technology of faithful reconstruction of the real world interior sceneries in a virtual world of video games using stereo-photogrammetry methods.

Index Terms—video games, 3D scanning, image acquisition, stereo-photogrammetry

1 Introduction

3D scanning can be a useful tool for acquiring assets for video games using real object models, in particular those taking place in real-life locations. The recent development of photogrammetry algorithms allow now for effective, high-quality reconstruction of objects based on a complete set of photos covering the entire surface of the scanned object. Unfortunately, an important problem of these algorithms is the high computational complexity associated with the need to compare photos in pairs. In the case of small objects, this is not generally important, however, in the case of scanning larger interiors, there are significant problems associated with determining the proper method of their execution. So the problem arise how to ensure adequate and safe level of the object's coverage with as few as possible photographs. This issue is of particular importance in the case of scanning rooms with limited access, requiring troublesome travel and costly preparations. In this case, the improperly carried out photo session, that requires repetition, significantly extends the modeling process and increases its costs.

2 Related work

There are generally two main methods of creating 3D models from images, which are *photogrammetry* and *stereo-photogrammetry* [1]–[3]. In both approaches images are processed in pairs to reconstruct the third dimension (scene depth) but photogrammetry analysis all two-element subsets from the entire image set of single, uncorrelated photos, while stereo-photogrammetry analysis stereo-pairs with known parameters of the optical system. In both cases, special algorithms, like scale-invariant feature transform (SIFT) or speeded up robust features (SURF) are used to detect the characteristic points in those images. Then, matching algorithms, such as random sample consensus (RANSAC) are used to find the correspondence between images pair.

In the photogrammetry specialized algorithms, like structure from motion (SfM), are used first to reconstruct the exact location and orientation of camera for each single shot, and then to restore the depth of all landmarks in the scene, creating so called *point cloud*. This point cloud is then expanded by subsequent analysis of further images. Stereo-photogrammetry takes full advantage of stable and known relative positions between the pair of cameras, which allows for single pass, simpler and more accurate reconstruction of a point cloud. All received point clouds are then matched to create the final point cloud representing the whole object.

Although creation of a point cloud from an image pair is much simpler in stereo-photogrammetry, the process of matching different point clouds is of the same computational complexity as matching of subsequent image pairs in the photogrammetry. That is why, this second approach, which requires simpler optical system, and accepts photos from different sources, is more popular nowadays.

Numerous tools have been developed that allow reconstruction of three-dimensional objects and sceneries. Some of these tools did not go beyond the stage of academic projects, while some of them are the companies' proprietary tools used in their products, for example for the reconstruction of three-dimensional objects on Google Maps. Two groups of applications can be distinguished on the commercial market. The first group are applications dedicated to specific hardware solutions, such as Leica Geosystems HDS ™Cyclone. The remaining group includes systems that do not require dedicated equipment, accepting a set of photos taken with a regular camera. The most popular applications include Photometrix Australis and iWitness, PhotoModeller, and AgiSoft Photoscan [5]. These applications are used, among others for the reconstruction of monuments and buildings, including photographs taken from drone, archaeological sites, accidents and crime scenes and many more.

Applications based on unrelated, single photos are good at reconstruction of objects with a large number of details taken in similar, good lighting conditions. Unfortunately,

- *Mariusz Szwoch, Gdansk University of Technology, Gdańsk, Poland, szwoch@eti.pg.edu.pl*
- *Adam Kaczmarek, Gdansk University of Technology, Gdańsk, Poland, adam.l.kaczmarek@eti.pg.edu.pl*
- *Dariusz Bartoszewski, Forever Entertainment S.A., Gdynia, Poland, dariusz.bartoszewski@forever-entertainment.com*

research has shown that these applications fail in more difficult cases, resulting in incomplete models. Another problem is the high computational and memory complexity of the algorithms used. One of the possibilities to solve these problems can be the use of stereo-photogrammetry methods, proposed in the STERIO project [6]. Because this technique assumes, in general, the use of a tripod for taking pictures, it allows to take a set of several photographs of more difficult fragments of the scene, using additional lighting or a graphic pattern. In addition, the ability to generate rare point clouds directly from the stereo-pairs significantly reduces the required scene processing time.

The common problem of all techniques based on image processing is the right way of such taking them that allows for suboptimal coverage of the entire surface of the scanned object. Too small number of photos does not cover with the the entire surface of the object, while too great number significantly increases the processing time. These problems become very important in a situation when access to the scanned room is limited. In such cases, it becomes particularly important to formalize the procedure for taking pictures. The proposed solution of this problem will be described in the next section.

3 Recommendations for taking images

In order to obtain a 3D scan of a building interior using photogrammetry or stereo-photogrammetry it is required to take a series of images of the scanned space. Such images have to meet two crucial requirements. Foremost, the entire area of a scanned interior needs to be visible or covered. If this requirement is not fulfilled blank areas will occur in a 3D scan obtained on the basis of this kind of photographs. Moreover, the same parts of a building interior need to be visible in many images. Otherwise, it will be problematic to match areas visible in different images in order to acquire a complete scans of an interior including all its parts. Overlapping areas makes it possible to merge data from these images resulting in obtaining a single 3D scan. This problem applies to all every part of a scanned area.

In general, greater number of images results in a more complete scan which more precisely reflects the shape of real spaces. However, preparing a great number of images has also disadvantages. The main one is the increase in the execution time of the algorithms which obtains a 3D scan on the basis of images. Moreover, also acquiring images is time consuming therefore it is advisable to reduce the number of images. However, images prepared for 3D scanning should be taken appropriately. A proper preparation of an image set concerns the selection of locations from which photographs are taken.

The problem of placing cameras during a photo-session was addressed in the manual of Agisoft PhotoScan which is a commercial software for obtaining 3D scans [5]. The document states that it is not correct to take images by aiming a camera in different directions from the same location. The camera should be relocated to a different position in order to take a subsequent photo. This requirement fully complies with our experiments concerning obtaining 3D scans on the basis of images.

In the STERIO project we have performed a series of experiments in which we acquired images of different building interiors and use them to reconstruct those rooms. We have developed a method of taking images which is based on placing cameras along straight lines parallel to walls. We assume that the process of acquiring images can be performed by an operator who is not familiar with the technology of photogrammetry and stereo-photogrammetry therefore instructions for taking images should be simple and easily understandable. The operator should take images by pointing a camera towards a wall of a building interior and moving from side to side in order to take a subsequent photo (Fig. 1).

Moreover, special concern should be given to all corners of the interior. For each corner additional series of images should be made from points of view located along a straight line placed opposite to a corner (Fig. 1). Making this series of images is required to properly combine parts of scans corresponding to adjacent walls. Apart from a series of images taken from locations which are close to a corner an additional images should be taken from points of view located along a straight line distant from the corner (Fig. 1). These images support providing completeness of a scan. The recommended location of cameras is presented in Fig. 1. The figure presents a sample building interior which is a corner of a parking lot.

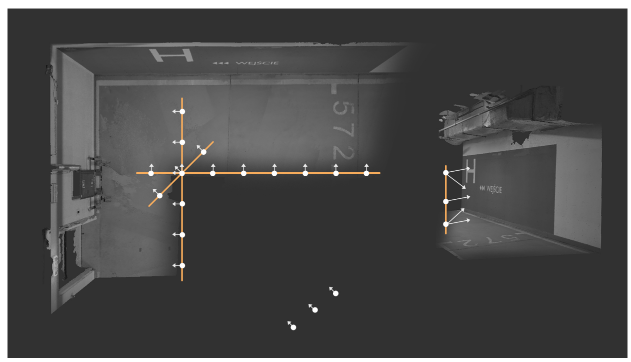

Fig. 1. Recommenced locations from which images should be taken for 3D scanning two adjacent walls and their common corner

Due to significant height of most rooms it is also advisable that images are taken from few different heights for each point in order to cover the entire area of a wall. This recommendations is visually presented in the rightmost part in Fig. 1.

We have also experimented with several other configurations of taking images. For example, in the case of small rooms very similar results can be obtained if cameras are placed around a room and aimed at the opposite wall (Fig.2). Such approach is much simpler and gives smaller sets of images.

Two other approaches failed, giving worse results or causing technical problems during the photo session. The procedure of moving camera along the arc paths (Fig. 3) does not improve the quality and completeness of a resulting 3D scan. Additional, high scale differences of subsequent images cause serious problems with their matching. Also approach in which pictures are taken form several locations around 360 degrees from each (Fig.4) gives poor

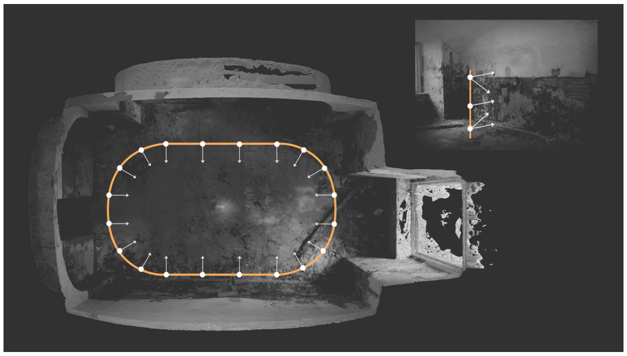

Fig. 2. A method of taking images recommended for small spaces

results as there are serious problems with finding images covering the same area.

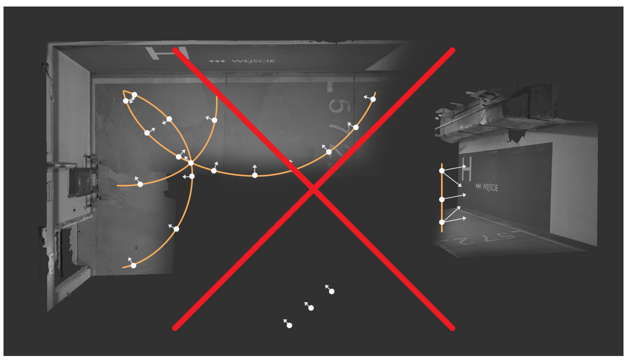

Fig. 3. A method of taking images which does not improve results

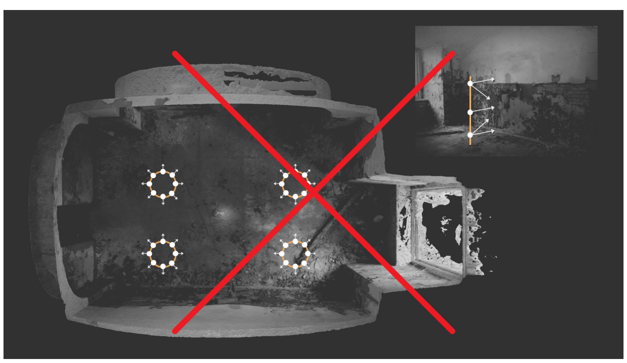

Fig. 4. An incorrect method of selecting locations for making images

4 RESULTS

Within the STERIO project an appropriate technology and a software tool is developed the effective reconstruction of 3D interior models using photogrammetry and stereo-photogrammetry tools. The result of the project will allow for faithful recreation of real world scenery in the virtual world of video games. Applying the proposed technology will allow for more effective and less costly location generation for games in real places. Such approach will have a big advantage over the indirectly competitive 3D laser-scanning technology as it produces a similar effect with fewer resources and shorter production time.

Although, the software tool is still under development the results received so far are very encouraging. The STERIO tool is able to reconstruct the dense point cloud from a set of images and convert it to a mesh model that can be exported into OBJ file format accepted by most game engines. Sample results of the reconstruction process carried out according to presented recommendations are presented in Fig.5-7.

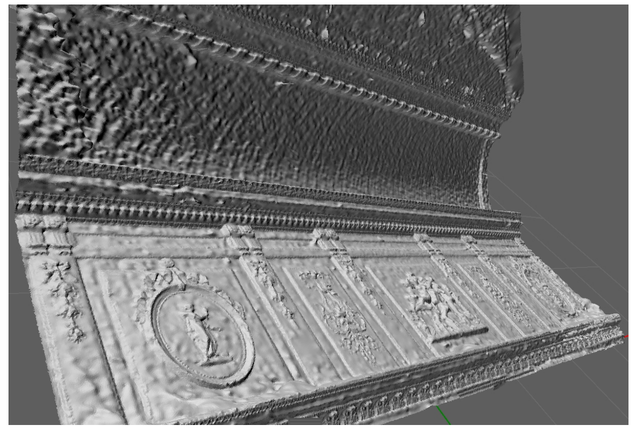

Fig. 5. Reconstructed geometry of a sample wall fragment with bas-reliefs

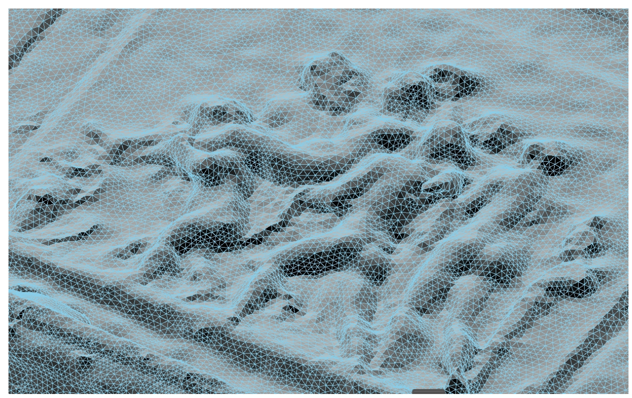

Fig. 6. Detailed mesh of a fragment of the bas-relief wall

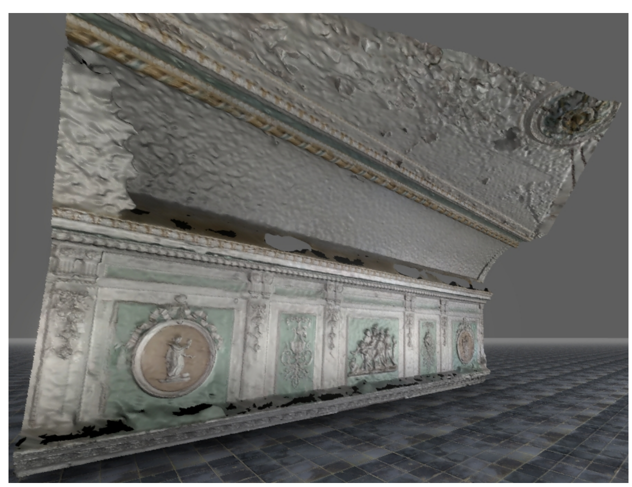

Fig. 7. Reconstructed sample wall, fully textured and lit in a STERIO viewer tool

At the moment, there are two significant problems with the developed technology. The first is the great computational complexity that is common for all photogrammetry based tools. More complex rooms require hundreds photos to cover all their surface, significantly extending the processing time. Moreover, matching algorithms use lot of memory often exceeding the amount of available physical memory in the system that additionally slower the process. The experiments with cloud computing based on Microsoft solution did not solve the problem due to the limits of memory and transfer time, so the general processing GPU approach will be examined as the next step.

The second problem concerns development of the methods to automatically reduce geometry complexity of modeled rooms. So far received models contain too many vertices and triangles to be directly used in a video game. Many different algorithms and software tools were tried but gave no satisfactory results. The main problem is not preserving edges and plane surfaces which is not acceptable. The solution is to enable this process to be done semi-automatically, offering the user a handy and useful tool to interact with.

5 Conclusion and future work

The method of making and selection of photos for the needs of photogrammetry and stereo-photogrammetry systems can significantly influence the effectiveness and time of reconstruction of 3D objects. Because it is not always possible to supplement photos or to take them again, it is extremely important to conduct a proper scan session. This paper presents the most important recommendations in this field, which significantly improve the quality of the models obtained for the tested rooms.

Experiments with data acquisition also concerned testing results of setting different camera parameters including exposure time, ISO, aperture value and focal length. In the experiments lens with different focal length were used. Conclusions drawn from these tests are such that there is no perfect camera setup which is always best one for 3D data retrieval. Therefore, cameras should be set appropriately to light and other environment conditions in which images are taken. However, the selection of camera lens is significant, because lens having low field of view cause the necessity to take a great number if images in order to cover the entire space which is under scanning. Thus, it is recommended to use wide angle lens.

Further work within the STERIO project includes the development of the application for reconstruction of rooms using both available methods of photo analysis. This hybrid approach will allow to take advantage of both methods, that is, the easiness of taking and processing individual photos in photogrammetry, and much greater capabilities to analyze difficult surfaces in stereo-photogrammetry. The developed application will also contain a whole set of tools for model processing at each stage of the room reconstruction process.

Acknowledgment

This work was supported by the Sectoral Programme GAMEINN within the Operational Programme Smart Growth 2014-2020 under the contract no POIR.01.02.00-00-0140/16.

References

[1] Scharstein, D., Szeliski, R.: *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms*. International Journal of Computer Vision, 47(1/2/3), 2002, pp. 7-42.

[2] Geddes, K.O., Czapor, S.R., Labahn, G.: *A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms*. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 2006, pp. 519-528.

[3] Kaczmarek, A.L.: *Stereo vision with equal baseline multiple camera set (EBMCS) for obtaining depth maps of plants*. Computers and Electronics in Agriculture 135, 2017, pp. 23-37 .

[4] Geiger, A., Lenz, P., Urtasun, R.: *Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite*. Conference on Computer Vision and Pattern Recognition (CVPR), IEEE 2012, pp. 3354-3361.

[5] *Agisoft PhotoScan User Manual Professional Edition, Version 1.3.* http://www.agisoft.com/pdf/photoscan_1_3_en.pdf Agisoft LLC, 2017.

[6] Szwoch, M., Kaczmarek, A.L., Bartoszewski, D.: *STERIO - Reconstruction of 3D Scenery for Video Games Using Stereo-photogrammetry*, Computer Game Innovations in Monograph of the Lodz University of Technology In. Napieralski P, Wojciechowski A (eds.), Lodz University of Technology, Lodz, Poland, 2017

Mariusz Szwoch received his Ph.D. from the Faculty of Electronics, Telecommunications and Informatics of Gdask University of Technology in 2002. His research interests include image processing, pattern recognition and computer vision, interactive multimedia systems, computer graphics, and video games. He is the author of more than 60 peer reviewed publications.

Adam L. Kaczmarek received the Ph.D. degree in informatics in 2012 from Gdansk University of Technology, Gdansk, Poland. He also received the M. Sc., Eng. degrees in informatics from this university. Mr. Kaczmarek worked as an Intern at Intel Technology Poland from 2005 to 2007. Since 2007 he is working at Gdansk University of Technology. Currently, he is an Assistant Professor. His research interests include information retrieval, stereo vision and 3D vision.

Dariusz Bartoszewski received his M.Sc. degree in economics from Gdansk University, Gdansk, Poland. He is a game developer with over 20 years of experience. Mr. Bartoszewski worked at Young Digital Poland and Aidem Media. Currently, he is a production manager at Forever Entertainment company.