3D Optical Reconstruction of Building Interiors for Game Development*

Mariusz Szwoch[1][0000−0002−4683−4727],
Dariusz Bartoszewski[2][0000−0003−1231−3322]

[1] Gdańsk University of Technology, Gdańsk, Poland
szwoch@eti.pg.edu.pl
[2] Forever Entertainment s.a., Gdynia, Poland
dariusz.bartoszewski@forever-entertainment.com

Abstract. This paper presents current possibilities of using stereophotogrammetry to scan interiors for video games. This technology offers great possibilities for effective room scanning based on popular cameras, however, it also has some limitations in the implementation of a fully automatic image processing. As the result of the STERIO project, the Sterio Reconstruction tool has been created that offers semi-automatic 3D reconstruction of building interiors. The proposed technology can be used, among others, for modeling rooms for video games which action takes place in locations reflecting real interiors. The original, large-scale experiment has been carried out in order to validate the created technology in real conditions. The goal of the experiment was to develop a prototype 3D video game by a small development studio using created solutions. The experiment showed great possibilities of the developed tool and the technology used.

Keywords: image processing · stereo-photogrammetry · 3D scanning · video games · interiors reconstruction.

1 Introduction

Three-dimensional vision plays an extremely important role in the everyday life of a human being, allowing, among others, for orientation in the field, avoiding obstacles, efficient movement and precise handling. For many years, it has also found numerous applications in digital systems. Among the many fields of applications, one can mention here simultaneous localization and mapping (SLAM) systems for robotics and autonomous vehicles, vehicle support systems (active cruise control, obstacle detection systems, and road analysis), optical quality control, 3D scanning of static and animated objects.

The latter solutions can be used for the development of modern video games that require hard use of high quality 3D assets for both game characters and the

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

scenery. In the traditional approach, all such assets are created by 3D artists or bought from a store. The first approach is a very labor-intensive task, while the second one does not guarantee uniqueness of bought assets, what is easily recognized by players. In case when game concept requires that the action takes place in a location that represents a real world place, traditional, handmade modeling seems to be non-creative, yet time and labor consuming task that could be replaced by automatic or at least semi-automatic software solutions. Our studies and large-scale experiments have shown that automation of room reconstruction significantly reduces the time and cost of modeling the 3D scene [1]. Moreover, it also supports rapid prototyping giving preliminary object models, which is a very important element of agile development process and often determines the further direction of the game development. For those reasons, it is very important to accelerate the process of modeling 3D scenes that are meant to reflect real objects. Three-dimensional scanning has already been successfully used in the development of several commercial games, such as The Vanishing of Ethan Carter by The Astronauts in 2014, Star Wars Battlefront by EA DICE in 2015, and Get Even by The Farm 51 in 2017.

There are available several technologies allowing for reconstruction of 3D objects, including photogrammetry [2], [3], laser based devices, e.g. light detection and ranging (LIDAR) [4] and time-of-flight (ToF) technologies, and finally depth sensors [5],[6]. Unfortunately, all of the mentioned solutions have some drawbacks that limit their usage in video game industry, especially by small and medium developer teams. Laser based scanners are very precise as to geometry scanning but are very expensive and do not support textured models. ToF sensors are cheaper but have limited resolution and also do not provide texture information. In turn, RGB-D depth sensors, based mostly on the structured infrared light patterns, like Microsoft Kinect or Intel RealSense, can provide complete 3D models of usually enough quality but are limited to small distances of a few meters. Finally, photogrammetry methods can provide high quality textured models but require numerous input data and high computational and memory resources. Moreover, it often fails to fully reconstruct objects with flat, uniform surfaces and poorly illuminated ones.

In this paper, the main concept of the STERIO project is presented, which aims to overcome problems of above approaches by using photogrammetry and stereo-photogrammetry methods for 3D scanning of building interiors for game development purposes at reasonable financial costs. The performed experiments proved that the proposed approach is faster than traditional one and allows to provide good quality models that can be used in video games.

2 Background

Three-dimensional video games are the most spectacular part of the market as they allow to present reliable game worlds that provide deep immersion experience. 3D games dominate primarily on the platforms of personal computers and video consoles, but they become an essential and growing part of games for mobile and browser platforms. Because the action of many 3D games takes place in worlds that resemble, reflect or even simulate our world, the creation of realistic models and scenes has become a very important task of 3D graphics, animators, and level designers. In general, scenes can be divided into outdoor and indoor ones with different characteristics, requirements and limitations. External scenarios usually include a wide area with diversified terrain, vegetation cover and numerous objects, e.g. buildings, vehicles, animals, opponents, etc. It requires, among other things, paying special attention to the scene rendering performance, which in some cases may require the use of simpler objects with low number of vertices (low-poly). In turn, indoor scenes have a very limited horizon, which allows the player to register the details of rooms, which usually enforces the use of more accurate models with more vertices (mid- and high-poly).

In the case of manual modeling of objects, high-poly models are usually created, which can be easily converted to simpler ones tailored to the needs of the designer. However, manual approach is time-consuming and requires high imaginary, design, and modelling skills. That is why different methods, algorithms, and tools supporting the 3D scanning have been developed recently. In general, they can be divided into *active* methods that use some additional light source for measurements and *passive* ones, that are based on the image analysis of the existing, non-enriched scene. Active methods may use the measurement of flight time or phase difference (ToF) of a laser beam reflecting from the object's surface (e.g. Microsoft Kinect v.2) or use triangulation method based on lasers (LIDAR) or a structural light patterns (e.g. Microsoft Kinect v.1). Precise active scanners are usually very expensive and often scan only geometry of the models. In turn, passive methods generally use triangulation on a set of images taken by calibrated stereo-camera set (stereo-photogrammetry) or by different, independent and not calibrated cameras (multi-view stereo, MVS). The first approach demands more effort to properly take pictures but provides higher precision and can deal with many problematic issues.

3D scanning technology has been developed for the last decades and used in different fields, such as preserving objects significant for a cultural heritage [7], reconstruction of buildings for 3D maps, character modeling for 3D animations etc. Initially, most efforts focused on scanning small and medium sized 3D objects using relatively cheap cameras, 3D scanners, and depth sensors based on structural-light technology [8]. Unfortunately, due to their limitations, such 3D scanners cannot be used for scanning bigger objects such as buildings and their interiors.

Modeling of larger objects and whole 3D locations could be carried out mainly using expensive laser scanners which are available to larger development studios [9],[10]. Laser based technics do not provide color information, unless hybrid laser-camera systems are used. With a smaller game budget, the modeling of larger assets remained the domain of 3D graphic designers. It is only recently that methods allowing the use of cameras for scanning buildings, their interiors, and even vegetation have attracted a lot of interest [5],[7],[8],[9].

3D scans can be also obtained with use of standard, optical cameras which are very accessible and relatively cheap. In general, there are two kinds of methods for acquiring 3D scans with cameras, which are *photogrammetry* and *stereo-photogrammetry*. The photogrammetric approach bases on a series of usually unrelated photos taken from around a scanned object. Special algorithms, such as scale-invariant feature transform (SIFT) or speeded up robust features (SURF) are used to detect the characteristic points in those images [11]. Matching algorithms, such as random sample consensus (RANSAC), are then used to find the correspondence between images pair. On this base, algorithms, such as structure from motion (SfM), allow to reconstruct a 3D point cloud representing the scanned object. This approach is most popular nowadays as it allows for the reconstruction of buildings, plants, and other objects based on uncorrelated images possible taken using different cameras at different time, e.g. as is the case in social media services. This approach is also used in a professional tools, such as Agisoft Photoscan [12]. The greatest disadvantage of this approach is the computational power needed to process big sets of images.

Another technique which allows to retrieve a 3D view of objects is stereo-photogrammetry that belongs to the same class of technologies as stereo vision [3]. In this method a rigid set of two cameras is used called a stereo-camera. As the relative position of cameras is fixed, it is possible to reconstruct partial point cloud of the scanned object directly from a stereo-image, simplifying the calculations time. This can be an essential advantage over SfM in some application area such as rapid modeling and 3D scenes prototyping in video games. Although, this approach requires using a tripod, it can become another advantage in some situations, such as insufficient illumination or hard-to-process surfaces, when it is possible to take several photos from exactly the same location in different lighting conditions [13].

3 The STERIO project

The goal of the STERIO project was to develop a low-cost technology for making 3D scans of building interiors based on images from standard photo cameras [13]. The results of the project developed by Forever Entertainment s.a. allow for faithful recreation of real world scenery in the virtual world of video games. Applying the proposed technology also allows for more effective and less costly location generation for games which action takes place in real locations. 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. The technology is intended for use in small- and medium-sized game development studios.

3.1 The STERIO data set

During the project, a comprehensive collection of over 10,000 photos of 26 various interiors were taken. These photos, single and stereo pairs, were used to develop, test and verify the algorithms and tools being developed. For each location, a different number of data subsets were prepared, consisting of images taken under different lighting conditions with different exposition parameters, camera location and orientation points, additional visual markers displayed, and other experiment setting. Some datasets are accompanied by additional images of a special chessboard pattern for cameras calibration. Sample images from the STERIO dataset are presented in Fig.1.

a) b)

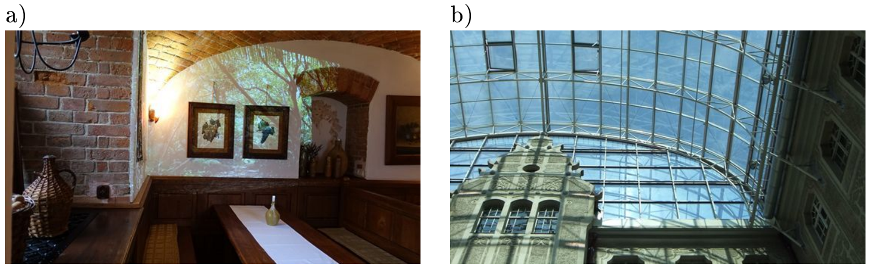

Fig. 1. Sample images from the STERIO dataset: the restaurant's cellar (a), the university hall (b)

3.2 Sterio Reconstruction tool

The main effect of the project is the Sterio Reconstruction tool that allows designer to perform a full process of processing the set of input images of the object into its 3D model. The tool uses image processing algorithms based on open source OpenMVG and openMVS libraries [14]. This tool offers a user-friendly interface that allows full control over the reconstruction process, starting from selection of input photos, through configuration of parameters of individual processing stages, saving of intermediate data, and presentation of results (Fig.2). SR offers also some manual and semi-automatic tools allowing for modifying intermediate point clouds, flattening regions etc.

The whole processing pipeline consist of several steps including (Fig.3):

- collecting photos, initial processing, and determining characteristic points (landmarks);
- searching for pairs of pictures covering a common fragment of a room;
- calculating relative and global camera positions and orientations for each analyzed pair;
- building local sparse point clouds and fitting them into the global sparse point cloud;
- building global dense point cloud;
- building object's mesh;
- building textures;
- decimating and retopologizing the model.

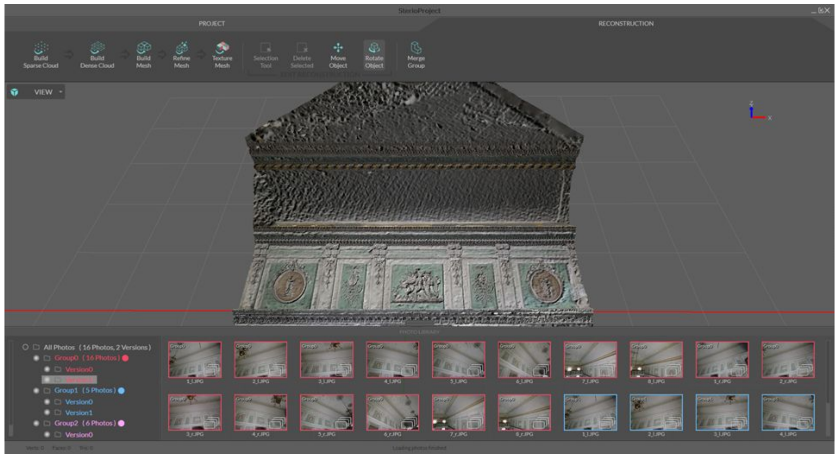

Fig. 2. Example screen of the Sterio Reconstruction tool

Fig. 3. Visualization of main steps of image processing in the Sterio Reconstruction tool: calculating camera positions and orientations (a), building global sparse (b) and dense (c) point cloud, building object's mesh (d), decimating the mesh (e), retopologizing the model with textures (f)

4 Experiments

The goal of the experiment was to validate the usefulness of the developed technology and tools in the video game industry. A large-scale experiment was carried out, using them in the development of a prototype video game, which takes place in locations based on real places. Within the experiment, complete design work was carried out that allowed to create a prototype survival horror game "The Mansion", containing one complete level. Particular attention was paid to the design of the game world, especially the layout of the title mansion, allowing for the correct placement of chambers and rooms created using the developed solution (Fig.4a). As part of the task, a complete scenography was created, consisting of a reconstructed mansion, additional rooms and an open area with small architecture. A set of several simple logic puzzles allows players to explore the area and assess the quality of its components.

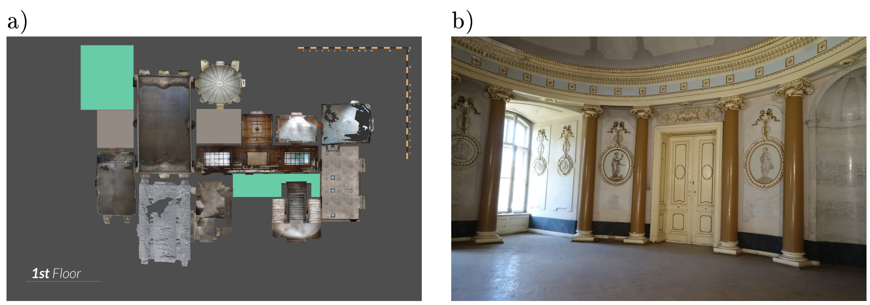

Fig. 4. The designed 1st floorplan for The Mansion game (a), example photography of the Bożków Mansion (The Pillar Room)

In order to create interiors corresponding to the actual location, a complete photographic session was made in accordance with the developed recommendations for a historic mansion from the 16th century located in the village of Bożków in Silesia region (Fig.4b). Searching for the target location revealed a big threat to this type of projects, which is the lack of consent or exorbitant financial expectations of property owners. All photos were taken in accordance to the previously developed recommendations for image acquisition for 3D room scanning [15]. These recommendations advise how to cover the room with photographs, properly balancing the completeness of the cover guaranteeing the completeness of the final model and the number of photographs deciding on the processing time and the required memory. Additional criteria set out the rules for exposure of photographs, as well as cases in which you can use individual photos to speed up the work, and when you need to use a stereophoto set to get better pictures in dark rooms or flat surfaces with no landmarks. On such surfaces special graphical pattern was displayed that gives the best results according to previous tests and experiments [1].

According to the designed plan of the rooms, several chambers were selected from the mansion in Bożków for full reconstruction and use in the prototype game. For each room an initial photo set was chosen for the Sterio Reconstruction tool according to the experience of the level designer. After a full reconstruction of the room, the model was inspected for its completeness. In the case of significant space shortages to the initial set of photographs, additional photos of a given part of the room were added.

Each fully reconstructed room underwent further processing aimed at simplifying the mesh of the model. At this stage, the built-in Sterio Reconstruction tool was used to flatten the selected surfaces, such as walls, ceilings, floors or their fragments. This operation allows for a radical reduction in the number of mesh tops. However, it causes some problems with the continuity of the grid at the border of the simplified area, which makes it necessary to leave a certain margin allowing its smooth connection with the environment. After each change of the mesh geometry one should also remember about the need to convert the texture mapping (retopologization).

In the next stage, the model is automatically simplified (decimation) to the extent that can be used in the game. Unfortunately, modern decimation algorithms still do not work optimally. If the level of reduction is too high, it often causes different artifacts, such as lack of edge and corner retention, highly irregular surface division, disappearance of details in important places, etc. Therefore, each reconstructed model was subjected to final processing by 3D graphics, whose task was to eliminate any imperfections of the model and its optimization and adaptation to the needs of the game. An exemplary appearance of the models of the reconstructed rooms is shown in Fig.5.

a) b)

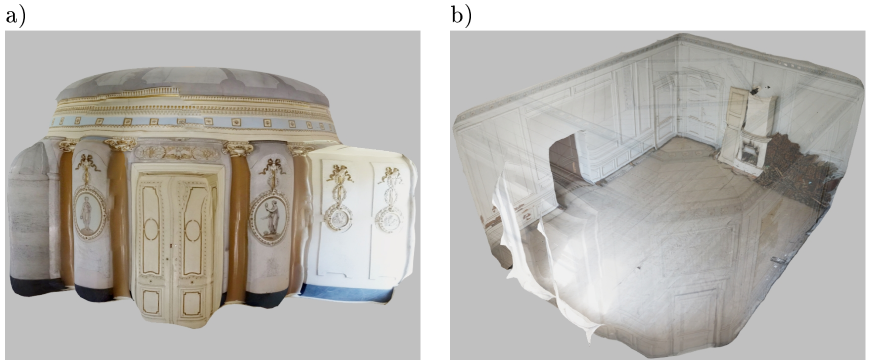

Fig. 5. An exemplary appearance of the models of the reconstructed rooms: The Pillar Room a), The White Room (b)

Finally, optimized models were imported into Unity 3D engine and put together, creating a plan of the first floor of the title mansion. In the level design process the models were combined with other room models, hand-modeled due to their nature, preventing them from 3D scanning (the orangery). All rooms have been adequately lit, enriched with detail models, interaction elements and additional special effects, e.g. floating dust, moon glow, mirror reflections, etc. An example of the appearance of rooms in the game is shown in Fig. 6.

Fig. 6. An exemplary appearance of the rooms during the gameplay

The entire process of reconstruction of the room was assessed by the level designer and 3D graphic designer in terms of suitability in the process of designing the rooms in the actual course of production. The overall rating was positive. The level designer emphasized the advantages of quickly receiving room prototypes, omitting the detailed concept sketches required in the traditional production process. In turn, 3D graphic designer emphasized the generally good quality of reconstructed models and their compatibility with 3D modelling tools like 3ds Max, Blender 3D and ZBrush. Although models still required a relatively large amount of work to eliminate all shortcomings, the process of creating models was faster than in the traditional approach.

5 Conclusion and future work

The experiment carried out with the use of the Sterio Reconstruction tool showed the high suitability of the photogrammetric approach for room reconstruction. Tools of this type can be valuable elements of the game design and development process, whose locations are based on real places.

The solutions proposed within the project result in greater efficiency in the reconstruction of rooms [15]. You can mention here about the set of recommendations regarding taking of photographs extremely important due to the high costs of potential repetition and photo session. Comprehensive collection of photographs of diverse rooms is a great base for testing and developing algorithms [13]. The proposed method for displaying a special graphic patterns with an LCD projector in rooms with poor lighting or flat surfaces allows to effectively scan these rooms using stereo-photogrammetry [1].

The conducted research also indicated some limitations of the photogrammetric approach. One of the most important limitations is very large computing and memory power requirements for image matching algorithms, creation and joining of point clouds. This significantly limits the possibility of using typical workstations, even with usage of the power of graphics cards (GPGPU). On such devices one should use a smaller set of pictures, which does not guarantee full coverage of the reconstructed room and thus completeness of the obtained model mesh that must be corrected manually. The solution to the problem may be the use of expensive, specialized workstations or the use of a cloud computing model.

Another problem is the correct decimation of the model grid. It is a well-known problem of combining simple and complex walls in 3D modeling, for which a completely satisfying solution has not been developed so far. This limits cooperation in some way with the proposed semi-automatic options for inviting flat surfaces, enforcing the need to leave a certain margin of unprocessed vertices. One solution may be development of decimation methods that keep edges and corners.

Another group of problems is related to the availability and condition of scanned rooms, which should also be taken into account in this type of projects. During the project there were unexpected problems with finding a manor house that could be formally used to implement the game. In many situations, some obstacles were: legal problems (lack of owner), lack of consent for temporary removal of furniture or absurdly high costs of granting permission for use in the production of the game.

Future work includes further development of photogrammetric image processing algorithms and the Sterio Reconstruction tool to overcome all mentioned problems and to maximize the automation level of the entire process. In particular, the efficiency of cloud computing and high-performance workstations will be examined. Advanced algorithms for flattening the surface and decimation will also be developed.

References

1. Szwoch, M., Kaczmarek, A.L., Bartoszewski, D.: Using Stereo-photogrammetry for Interiors Reconstruction in 3D Game Development. In: Choraś M., Choraś R. (eds) Image Processing and Communications Challenges 10. IP&C 2018. Advances in Intelligent Systems and Computing, vol 892. Springer, Cham (2019).
2. Mikhail, E.M., Bethel, J.S., McGlone, J.C.: Introduction to Modern Photogrammetry. Wiley, New Jersey (2011)

3. Kaczmarek, A.L.: Stereo vision with Equal Baseline Multiple Camera Set (EBMCS) for obtaining depth maps of plants. Computers and Electronics in Agriculture 135:23–37 (2017)
4. Musialski, P., Wonka, P., Aliaga, D.G., Wimmer, M., Van Gool, L., Purgathofer, W.: A survey of urban reconstruction. Computer Graphics Forum, 32 (6), 146–177 (2013)
5. Escorcia, V., Dávila, M.A., Golparvar-Fard, M., Niebles, J.C.: Automated Vision-Based Recognition of Construction Worker Actions for Building Interior Construction Operations Using RGBD Cameras. In: Construction Research Congress (2012)
6. Szwoch, M., Pieniążek, P.: Detection of Face Position and Orientation Using Depth Data. In: The 8th Intern. Conf. on Image Processing and Communications; Image Processing And Communications Challenges 7 Book Series: Advances in Intelligent Systems and Computing. Springer International Publishing 389:239–251 (2016).
7. Balsa-Barreiro, J., Fritsch, D.: Generation of visually aesthetic and detailed 3D models of historical cities by using laser scanning and digital photogrammetry. Digital Applications in Archaeology and Cultural Heritage (2017)
8. Ha, H., Oh, T.H., Kweon, I.S.: A multi-view structured-light system for highly accurate 3d modeling In: Intern. Conf. on 3D Vision (3DV). IEEE 118–126 (2015).
9. Kazmi, W., Foix, S., Alenya, G.: Plant leaf imaging using time of flight camera under sunlight, shadow and room conditions. In: Intern. Symp. on Robotic and Sensors Environments (ROSE). IEEE 192–197 (2012)
10. Tse R.O., Gold C., Kidner D. (2008) 3D City Modelling from LIDAR Data. In: van Oosterom P., Zlatanova S., Penninga F., Fendel E.M. (eds) Advances in 3D Geoinformation Systems. Lecture Notes in Geoinformation and Cartography. Springer, Berlin, Heidelberg (2008)
11. Bay H., Tuytelaars T., Van Gool L. (2006) SURF: Speeded Up Robust Features. In: Leonardis A., Bischof H., Pinz A. (eds) Computer Vision – ECCV 2006. ECCV 2006. Lecture Notes in Computer Science, vol 3951. Springer, Berlin, Heidelberg (2006)
12. Agisoft PhotoScan user Manual (2014), https://www.agisoft.com/pdf/photoscan-pro_1_4_en.pdf. Last accessed 10 July 2019
13. 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) Łódź University of Technology, Łódź, Poland (2017)
14. Kaczmarek, A.L., Szwoch, M., Bartoszewski, D.: Using Open Source Libraries for Obtaining 3D Scans of Building Interiors. In: Borzemski L., Świątek J., Wilimowska Z. (eds) Information Systems Architecture and Technology: Proceedings of 39th International Conference on Information Systems Architecture and Technology – ISAT 2018. ISAT 2018. Advances in Intelligent Systems and Computing, vol 852. Springer, Cham (2019)
15. Szwoch, M., Kaczmarek, A.L., Bartoszewski, D.: Recommendations for Image Acquisition for 3D Room Scanning. Journal of Applied Computer Science, Vol. 26 No. 2, pp. 45–55 (2018)