https://doi.org/10.2352/ISSN.2470-1173.2018.03.ERVR-452

Seeing the past: An augmented reality application for visualization the previous state of cultural heritage locations

Piotr Siekański[1], Eryk Bunsch[2], Robert Sitnik[1];
[1]Warsaw University of Technology, Warsaw, Poland; [2]Museum of King Jan III's Palace at Wilanów, Warsaw, Poland

Abstract

Augmented reality is very widely used to increase the attractiveness of sightseeing cultural heritage objects. Many locations are superimposed with additional computer generated data such as monuments, buildings, virtual characters, fauna and flora using dedicated systems that must be compatible with existing preservation policy. In this paper we present an example of an application that is used in King's Chinese Cabinet (Museum of King Jan III's Palace at Wilanów, Warsaw, Poland) that was scanned twice, before and after restoration works. This set of 3D data allows us to create an unique application that visualizes the real past state of preservation instead of computer generated artificial one. The interior is under a very strict protection and that increases significantly the requirements for an augmented reality system.

Introduction

Attractive and interactive sightseeing of the cultural heritage locations is a vivid topic. In recent years much effort has been done to get it to the whole new level – from static exhibitions into the interactive multimedia presentations. Augmented reality is very widely used to improve user's experience while sightseeing cultural heritage locations. Much effort has been made to superimpose existing cultural heritage locations with computer generated elements e.g. to virtually reconstruct monuments from ruins or to visualize and simulate everyday life of people living in the past in that places. This kind of approach is widely used in education and in entertainment but has a noticeable drawback: the computer generated data only simulates the possible previous state of the location.

The demand is growing and the advent of 3D scanning techniques has enabled the possibility to scan the locations before and after fundamental changes in its visual character (e.g. restoration) and then allow visitors to compare its actual state with the registered real state from the past. We exploit this approach to improve the attractivity of visiting King's Chinese Cabinet in Museum of King Jan III's Palace at Wilanów, Warsaw (Figure 1) which was scanned using structured light technique in 2011 prior to conservation and scanned again afterwards in 2015 [1](Figure 2).

***Figure 1**. Sample views from King's Chinese Cabinet*

***Figure 2**. Visual comparison between two registered point clouds from 2011 (left) and from 2015 (right)*

Related work

The use of 3D reconstruction in cultural heritage locations is very widely used due to popularity of image based model creation methods and laser scanning. [2] mentions many different advantages of 3D reconstruction in cultural heritage:

- documenting historic buildings and objects for reconstruction or restoration in case of fire, earthquake, flood, war, erosion, and so on;
- creating educational resources for history and culture students and researchers;
- reconstructing historic monuments that no longer or only partially exist;
- visualizing scenes from viewpoints impossible in the real world due to size or accessibility issues;
- interacting with objects without risk of damage;
- providing virtual tourism and virtual museum exhibits.

Based on a registered 3D cultural heritage locations, augmented reality applications are created to add 3D content to the real world images from the camera. Papagiannakis et al. [3]–[5] created an augmented reality application that enhance the users experience when visiting Pompeii by additional content such as virtual characters in ancient everyday activities or their unique fauna and flora. The users were required to wear a video-see-through HMD on which the additional content is rendered in real time to create an immersive experience.

Ruiz et al. [6] used augmented reality in a tropical forest of the Calakmul's Biosphere Reserve in Mexico, where an ancient Mayan metropolis is located. The main goal of this project is to overcome the difficulties in sightseeing due to the fact that it is located in a deep tropical forest and it takes about 5 hours to get there. The project is mainly focused on VR but AR is used in a physical reproduction of an ancient tomb originally located in the Mayan city and recreated in the San Miguel Museum in Campeche City. It allows the visitors of the Museum to see the physical reproduction of a funeral chamber as it was found by the archeologists and using augmented reality glasses the visualization of a funeral is possible to observe.

The PRISMA project developed by Fritz et al. [7] incorporates AR techniques into a tourist application. The camera is mounted on a tourist binoculars and its movement is tracked by the inertial sensors. Based on them additional simple 2D content is added exploiting the fact that the target locations are far away from the user and the correct perspective does not affect the users experience. Added information is a simple overlay.

The ARCHEOGUIDE project [8] held in Olympia, Greece is developed to superimpose existing ruins of ancient Olympia with artificially created models of the monuments. It offers a predefined paths for user to choose from. The system serves as a virtual guide as the user is moving along the path. The visualization is rendered on the special AR glasses and require user to carry a handy unit. Its main drawback is that it is limited by the very strict set of possible points of view.

Seo et al. [9] created an augmented reality based tour guide in Gyeongbokgung Palace, Korea in which camera position is estimated by tracking simple geometric primitives and special wooden labels. The view is superimposed by virtual characters and additional information synchronized with projection. Their system also supports personalized tour guides based on users location and preferences.

All of the solutions presented above are based on artificially generated content that superimpose the view of a location. On the other hand the availability of scans of King's Chinese Cabinet before and after conservation makes our solution unique due to the fact that what we visualize is the state of our location prior to conservation without any artificial content created by computer graphics designers [10] (Figure 3).

***Figure 3**. Input camera frame (left) and rendered point cloud (right)*

Application requirements

The availability of precise and very dense 3D data allows us to create an application that tracks the camera position relative to the interior and displays the view on the model registered before restoration resulting in time-travel like visual experience (Figure 4). Our core requirements were imposed by the interior and can be summarized as follows:

- The Cabinet is covered by a very fragile paintings so it was forbidden to attach any additional markers. This forced us to prefer markerless solutions.
- The paintings are also very sensitive to the light so the Museum allows using only the visible light spectrum. That rejects RGBD sensors like Kinect that use infrared laser to estimate the depth of an observed scene.

The solution should work on a mobile devices such as tablets or smartphones that usually have only monocular camera so we decided to limit our application to only one camera as our input device and develop an algorithm to estimate its position in real time keeping in mind that the registered point clouds were scanned using structural light technique instead of typical state-of-the-art methods: SfM or laser scanners.

***Figure 4**. Application in use*

Hardware

We use a single Logitech B910 camera with resolution reduced to 640x480 px and calibrated offline. We connect it to Lenovo Yoga convertible laptop with Intel Core i5-5200U processor and Nvidia 840M graphics card. This card is able to handle complex visualization as well as camera pose estimation algorithm. To achieve the best performance we implement remote processing option on a dedicated workstation that communicates over wi-fi with our laptop. Our server is an ordinary PC computer with Intel i7 960 3,2Ghz processor and Nvidia GeForce GTX 590 graphics card. The remote processing is not necessary but it decreases the processing time because of higher performance of a server and separation between two main GPU tasks – point cloud rendering and camera pose estimation algorithm.

***Figure 5**. East wall of the Cabinet - sample ortho render*

Algorithm

Our solution works in a two stages mode: Firstly, in an offline stage the 3D scanned model is rendered using ortho projection (Figure 5). We render one image per Cabinet wall that results in 4 renders to cover the whole Cabinet. After that we compute the keypoints and descriptors on them to create our train set. We also store the 3D coordinates of each keypoint.

Our runtime loop is similar to the state-of-the-art (Figure 6). For every image taken by the camera, we extract keypoints and descriptors, then match them to our train set and filter them according to the ratio test proposed by Lowe [11]. After that we estimate the camera pose based on 2D-3D correspondences using ePnP algorithm [12] with RANSAC procedure to reject outliers. Due to the fact that we estimate camera pose in each frame separately the resulting camera path is not stable and the rendered view is shaking. To achieve pleasant visual experience we developed a Gaussian weighted averaging on camera positions that eliminates jitter in camera movement (Figure 7).

Finally, filtered camera position is transferred to the visualization module and the frame of the previous state is rendered according to the estimated camera position.

The most time consuming task is keypoints and descriptors extraction process and matching ant thus we decided to implement it on GPU using Nvidia CUDA technology.

Camera pose averaging

We assume that the error in camera poses is normally distributed and therefore we introduce two parameters: the number N of past frames that influence the current camera position and σ that is used as a smoothing parameter in our Gaussian averaging.

For each frame i we solve the PnP problem to estimate camera pose C_i and save the result as 4x4 matrix using homogeneous coordinates (1) (X_i, Y_i, Z_i, T_i denote 3-element column vectors). We also save the timestamp t_i of a frame:

$$C_i = \begin{bmatrix} X_i & Y_i & Z_i & T_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (1)$$

To calculate averaged camera position we save last N frames and for each past frame i we calculate time difference Δt between current camera frame timestamp and t_i.

And then we calculate local Gaussian weight w_i of a frame based on previously computed Δt (2):

$$w_i = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\Delta t}{\sigma}\right)^2} \qquad (2)$$

Then we sum up all of the local Gaussian weights to calculate global weight factor w (3):

$$w = \sum_N w_i \qquad (3)$$

This factor is further used to calculate filtered camera translation T_f based on local weights w_i and local translation T_i (4):

$$T_f = \sum_N \frac{w_i \cdot T_i}{w} \qquad (4)$$

Because we want to keep the views in adjacent frames as close as possible we also have to maintain the direction that the camera is looking. To minimalize differences in z axis (camera axis) we create new filtered z axis Z_f (5) and then normalize it.

$$Z_f = \sum_N w_i \cdot Z_i \qquad (5)$$

And then we create new rotation axes X_f, Y_f (6) based on Z_f ($\times$ denotes vector multiplication and Z denotes unit vector along z axis (0,0,1)):

$$X_f = Z \times Z_f, \qquad Y_f = X_f \times Z_f \qquad (6)$$

Finally we apply new filtered transformation matrix for camera C_f:

$$C_f = \begin{bmatrix} X_f & Y_f & Z_f & T_f \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (7)$$

We experimentally set the neighborhood N to 15 poses and σ to 1 as it is the optimal tradeoff between smooth camera movement and application responsiveness.

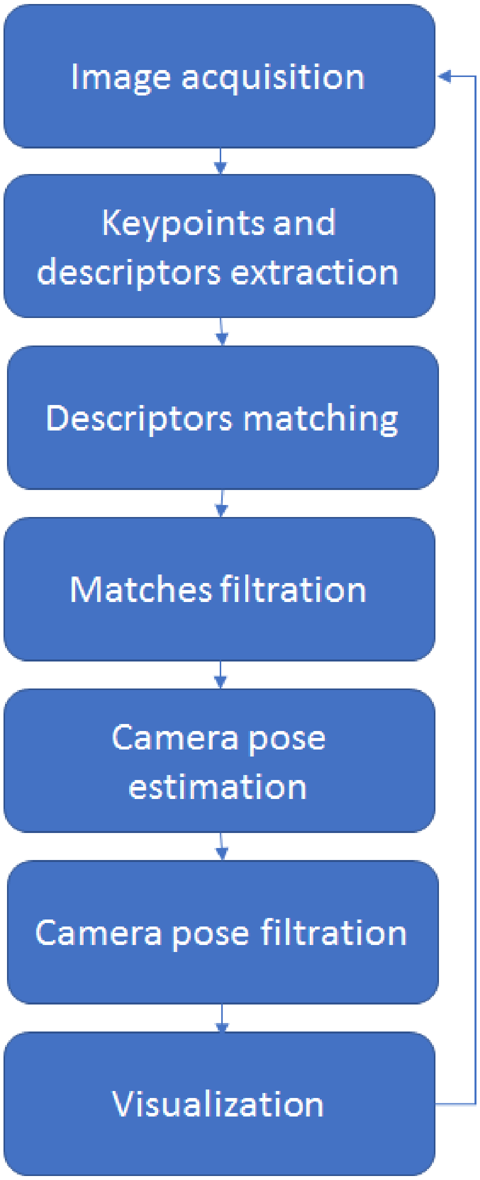

***Figure 6**. Runtime loop*

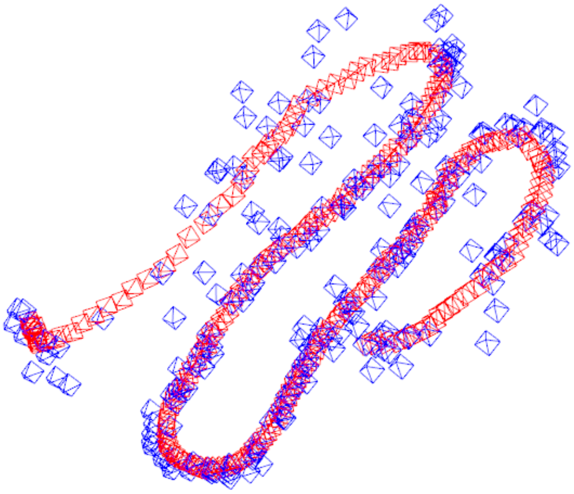

***Figure 7**. Estimated camera path. Blue cameras represent camera path estimated by our solution and red ones represent camera path with our Gaussian averaging*

Conclusions and future works

The potential of augmented reality in cultural heritage presentation is enormous. While many solutions rely on artificially generated content to superimpose real world image, we instead show to the users a real past state of the King's Chinese Cabinet in Museum of King Jan III's Palace at Wilanów, Warsaw, Poland prior to a massive restoration while observing its current state. We implement and test our method and provide a real time solution with a complex visualization. The interior is under a very strict protection that restricts significantly the set of possible augmented reality techniques to use inside.

In the future we plan to reduce the computational complexity and port our solution to mobile devices such as tablets and smartphones and test it in different cultural heritage locations.

Acknowledgements

This work has been partially supported by the European Union within the Infrastructure and Environment Operational Programme and Statutory Work of Warsaw University of Technology within the project "Revitalization and digitalization Wilanów the only Baroque royal residence in Poland".

References

[1] R. Sitnik, E. Bunsch, G. Mączkowski, and W. Załuski, "Towards automated, high resolution 3D scanning of large surfaces for cultural heritage documentation," in *IS&T/SPIE Symposium on Electronic Imaging: Science &Technology, IS&T*, 2016.

[2] S. F. El-Hakim, J.-A. Beraldin, M. Picard, and G. Godin, "Detailed 3D reconstruction of large-scale heritage sites with integrated techniques," *IEEE Comput. Graph. Appl.*, vol. 24, no. 3, pp. 21–29, May 2004.

[3] G. Papagiannakis *et al.*, "Mixing virtual and real scenes in the site of ancient Pompeii," *Comput. Animat. Virtual Worlds*, vol. 16, no. 1, pp. 11–24, 2005.

[4] G. Papagiannakis and N. Magnenat-Thalmann, "Mobile Augmented Heritage: Enabling Human Life in Ancient Pompeii," *Int. J. Archit. Comput.*, vol. 5, no. 2, pp. 395–415, 2007.

[5] G. Papagiannakis *et al.*, "LIFEPLUS: revival of life in ancient Pompeii, virtual systems and multimedia," in *Proceedings of the 8th International Conference on Virtual Systems and Multimedia*, 2002, pp. 25–27.

[6] R. Ruiz, S. Weghorst, J. Savage, P. Oppenheimer, T. F. Iii, and Y. Dozal, "Virtual Reality for Archeological Maya Cities," in *UNESCO World Heritage Centre Conference*, 2002.

[7] F. Fritz, A. Susperregui, and M. T. Linaza, "Enhancing cultural tourism experiences with augmented reality technologies," *6th Int. Symp. Virtual Real. Archaeol. Cult. Herit. VAST*, pp. 20–21, 2005.

[8] V. Vlahakis *et al.*, "Archeoguide: An augmented reality guide for archaeological sites," *IEEE Computer Graphics and Applications*, vol. 22, no. 5. pp. 52–60, 2002.

[9] K. Kim, B.-K. Seo, K. Kim, and J.-I. Park, "Augmented Reality-Based On-Site Tour Guide : A Study in Gyeongbokgung," in *Asian Conference on Computer Vision ACCV 2010: Computer Vision – ACCV 2010 Workshops*, 2010, no. November.

[10] E. Bunsch, A. Guzowska, and R. Sitnik, "3D scanning documentation of two different objects - The King's Chinese Cabinet in Wilanow Palace Museum and a Roman gravestone from archeological excavations in Moesia Inferior as a part of multidisciplinary research," in *Proceedings of the 2012 18th International Conference on Virtual Systems and Multimedia, VSMM 2012: Virtual Systems in the Information Society*, 2012, pp. 633–636.

[11] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[12] F. Moreno-Noguer, V. Lepetit, and P. Fua, "Accurate non-iterative o (n) solution to the pnp problem," *Comput. Vision, 2007. ICCV 2007. IEEE 11th Int. Conf.*, pp. 1–8, 2007.

Author Biography

Piotr Siekański received his BSc Eng in Mechatronics with specialization in Multimedia Techniques from the Warsaw University of Technology (2014) and MSc Eng within the same field and University in 2016. He is currently a PhD student at Warsaw University of Technology. His work and research is focused on the real time multimedia systems and markerless camera pose estimation algorithms.

Eryk Bunsch, M.Sc., dipl. conservator of art. Collaborates with National Institute for Museums and Public Collections and National Audiovisual Institute since 2011 as an expert in the field of modern 3D documentation of cultural heritage. Head of Laboratory for 3D Documentation at Museum of King Jan III's Palace at Wilanów. His current scientific interest considers developing technology of precise 3D scanning of cultural heritage objects. He has managed many research projects related to 3D documentation. His articles are combining different spheres of his scholar activity: conservation of art, 3D precise documentation and archaeology.

Robert Sitnik (Member of SPIE) received his MSc Eng (1999), PhD (2002) in applied optics from the Warsaw University of Technology. He has authored and co-authored more than hundred scientific papers. His interests are structured light shape measurement (3D/4D), triangulation methods, digital image processing, computer graphics, animation software development and virtual reality techniques. He has been a leader of projects from various fields like 3D optical metrology, virtual and augmented reality and supporting medical diagnosis by opto-numerical solutions. He is head of Virtual Reality Techniques Division at WUT.