https://doi.org/10.2352/issn.2168-3204.2020.1.0.55

Automated 3D mass digitization for the GLAM sector

Santos, Pedro[1]; Tausch, Reimar[1]; Domajnko, Matevz[1]; Ritz, Martin[1]; Knuth, Martin[1] and Fellner, Dieter[1,2]; [1]Fraunhofer Institute for Computer Graphics Research and TU-Darmstadt, Germany; [2]Graz University of Technology, Institute of Computer Graphics and Knowledge Visualization, Austria

Abstract

The European Cultural Heritage Strategy for the 21st century has led to an increased demand for fast, efficient and faithful 3D digitization technologies for cultural heritage artefacts. Yet, unlike the digital acquisition of cultural goods in 2D which is widely used and automated today, 3D digitization often still requires significant manual intervention, time and money. To overcome this, the authors have developed CultLab3D, the world's first fully automatic 3D mass digitization technology for collections of three-dimensional objects. 3D scanning robots such as the CultArm3D-P are specifically designed to automate the entire 3D digitization process thus allowing to capture and archive objects on a large-scale and produce highly accurate photo-realistic representations.

Introduction

In this paper we present a workflow, from photogrammetric acquisition in challenging environments to a representation of the acquired 3D models in different ways, usable for online visualization and color 3D printed replicas.

Central to our workflow is the CultArm3D-P, a fully automated color-faithful 3D-scanner able to scan arbitrary objects at reproducible high resolution at the push of a button.

To this end the CultArm3D-P robotic scanning system features autonomous view-planning with depth and pose estimation making sure the optimal number of images are taken, with an optimal amount of overlap, in a way, as to gather sharp information from every region of the surface of the artefact the image sensor can physically see. For faithful color reproduction ICC profiles are calculated.

Our work lays the foundation for a color-faithful end-to-end reproduction of artifacts, which has successfully been applied to several digitization projects of which we present a few results, such as a project with the Reiss-Engelhorn Museums in Mannheim. We present results of what is to our knowledge the first ever autonomous photogrammetry-based 3D digitization of very small golden artifacts, in this case from the island of JAVA in Indonesia, at very high resolution using a polarization approach combined with focus stacking.

We conclude with an outlook on the end-to-end reproduction workflow leading to virtual replicas (online 3D visualization, virtual and augmented reality) and physical replicas (3D printed objects using Fraunhofer IGD's photopolymer printer driver cuttlefish).

CultArm3D-P

In this chapter we describe the CultArm3D-P, a fully automated color-calibrated 3D-scanner able to scan arbitrary objects at reproducible high resolution at the push of a button.

The motivation for this scanning station is to relieve the human operator from tedious tasks, such as repositioning a camera around the object and keeping track of the scanning progress to eventually ensure a complete surface coverage and stable quality. Those are challenging tasks even for expert scanning operators, especially for high resolution scanning where the scanner measurement volume (defined by the camera optics) is normally a lot smaller than the object itself, and thus the resulting high resolution 3D model is comprised of many single scans or images. The impact of this scanning station is twofold. Firstly, the overall scanning time is effectively reduced by a high data acquisition rate with automated and parallel processing. Secondly, the automated and adaptive data acquisition enables the economical use of focused camera optics, such as macro lenses, to the scanning task even for larger objects, and thus effectively increases resulting surface resolution and 3D model quality.

The components of the CultArm3D scanning station can be classified into to capturing and positioning devices, which are synchronized and controlled by a standard PC. For capturing a high-resolution photo camera is combined with a customized mounted ring light and an optional background light. The camera must feature a PC control interface for triggering and transferring the images, such as the Canon 5DS R (50MP) or the PhaseOne iXG (100MP). For positioning a light-weight robot arm holding the camera is combined with a turntable for the object. Thus, the object can be captured from all sides while movements of the camera can be restricted to one side of the turntable and the robot arm is not required to reach over the object for capturing it from the other side, resulting in a safer workspace and enabling the use of a static photo background. The positioning devices must also feature a near real-time PC interface, such as the collaborative robot arm series from Universal Robots.

For the CultArm3D scanning station there are currently two versions available, one light-weight compact desktop version (see Fig. 1) and one heavier out-of-box version that comes with a centerless glass turntable. While the first version features a space-saving flat turntable, second version enables capturing the object even from below through a glass plate in one scanning pass without having to reposition it. This further reduces the manual interaction with, e.g. fragile objects that cannot simply be repositioned upside-down or sideways.

***Figure 1**. CultArm3D-P front and back light;*

The custom designed ring-light has a D50 spectrum suited for color calibration with attachable polarization filters for capturing shiny or other challenging objects without specularities. The ring-light is used at close distance to illuminate and dissolve even cavities while the automated surface-adaptive robot motion guarantees a steady distance to the area in focus and thus a uniform light intensity throughout the whole scanning task. An active backlight can optionally be installed to support object segmentation from the background. Figure 1 shows both lights in operation, the ring-light at close focus distance and the back light for capturing the object's silhouettes from far distance.

Cyberphysical Aequivalence and View Planning

All relevant hardware components are spatially modelled and integrated into one virtual 3D monitoring and planning environment (see Fig. 2). The viewer shows the actual robot pose and the resulting camera angle by processing essential robot sensor readouts in near real-time and combining them to a human conceivable virtual 3D representation of the present reality. Furthermore, future scanning actions are planned and visualized (as a green transparent overlay). Intermediate scanning results are displayed within the cylindrical object safety volume (orange transparent overlay) and updated as soon as they are reconstructed and become available to provide a visual preview and indication of the scanning progress.

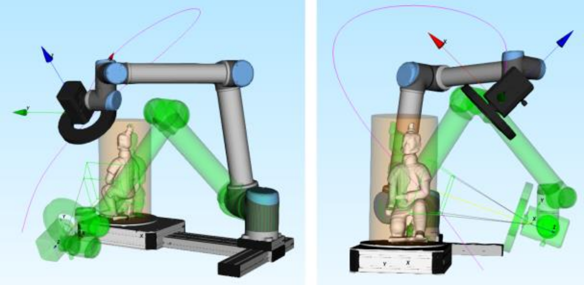

***Figure 2**. CultArm3D-P Autonomous View Planning environment*

In particular, techniques of analytical forward and inverse kinematics are applied to plan transitions between consecutive camera views and the implied robot trajectories in a safe and time-efficient way. For this reason all vital robot parts are also augmented with colliders, i.e. surrounding geometric primitive, for rapid prediction and avoidance of collisions (see Fig. 2). In this way, camera views that would compromise the object's safety are either automatically rejected or rearranged [1].

Initial system calibration

To further increase the precision of the planning and scanning results, the previously defined and modelled spatial transformations from the virtual 3D representation between different parts (such as turntable, robot and camera) are refined and corrected by an additional geometric calibration phase.

This scanning station is mobile deployable and quickly installed at new scanning locations with different camera lens combinations. However, this usually requires to temporarily unmount the camera using the quick release adapter, or to detach the robot arm from the turntable for compact packing, and this will result in a slight change in the spatial transformations and camera intrinsics when reassembling parts of the station. Therefore, the CultArm3D scanning station can compensate for this change by an initial system self-calibration that automatically carries out the following four three steps in order:

1. Camera intrinsics calibration
2. Hand-eye calibration
3. Turntable calibration

First the camera intrinsics are retrieved based on [2] defining the actual field of view and compensating the lens distortion. Next the hand-eye transformation between the camera's optical center (eye) and the robot's tool frame (hand), where the camera is mounted at, is retrieved based on [3]. Using the robot arm limp calibration, which is usually provided manufacturer, the optical center of the camera can now be positioned with respect to the robot arm base. In order to position the camera with respect to the object on the turntable, a final step, the turntable calibration is carried out to find the surface and the rotation axis of the scanning volume. All essential calibration data is automatically acquired after placing a calibrated ring-board target on the turntable.

After the geometric calibration is complete, the camera is characterized for color by replacing the calibration target with a known color board (such as X-Rite ColorChecker SG for normal setup or Rez Checker Target.for macro setup) and capturing it within the scanning volume at the desired fixed focus distance with the ring-light illumination.

Image acquisition and 3D reconstruction

The CultArm3D scanning station reconstructs 3D models using the established technique of photogrammetry. Therefore, the acquired raw data consists of high resolution photos of the object, covering each surface part several times, in order to use structure-from-motion [4][5] and multi-view-stereo [6][7] techniques to recognize features and triangulate 3D information. In general, given a set of photos with complete surface coverage, then the higher the camera image resolution, the higher the resolution of the resulting reconstructed 3D model. Hence, for highest resolution focused camera macro optics can be applied even for objects much larger than the actual camera measurement volume (defined by field of view and depth of field) and capturing only a small part of the object surface per image but in high resolution. The drawback is that because of the small measurement much more photos are necessary to cover the complete surface. This trade-off between time and quality is addressed by this scanning station.

At the current state the CultArm3D scanning system captures a high resolution image in average every 4 seconds (approx. 900 images per hour). The images are directly transferred via USB3 connection and store in a project folder for further processing. For small camera movements between views, e.g. during focus stacking, the system's acquisition speed becomes mostly limited by transfer speed of the camera. To achieve this high acquisition rate a state machine is modelled in software that synchronizes the different hardware and software components, such as camera, lights, robot arm and turntable, reconstruction, view and trajectory planning modules. Images are captured and processed in groups, allowing for parallelization, e.g. capturing the next group while processing the previous group.

In comparison to other 3D reconstruction techniques, such as structured-light or space-time analysis for laser triangulation, photogrammetry is computationally expensive, normally resulting in long processing times of several hours for the final full resolution model. Therefore, the final 3D model in highest quality based on the full resolution images is usually calculated offline after the data acquisition process with the automated scanning station finished. In order to predict an adequate reconstruction quality for the final model and ensure complete surface coverage, intermediate low quality 3D models are reconstructed in clusters already during the scanning task and serve as a quality indicator and decision base for further scanning actions.

View planning

Especially in the cultural heritage domain, objects are very unique and vary in size and shape. Therefore, the scanning strategy and the camera views cannot be simply predefined but have to be carefully adapted to the individual object to achieve best results.

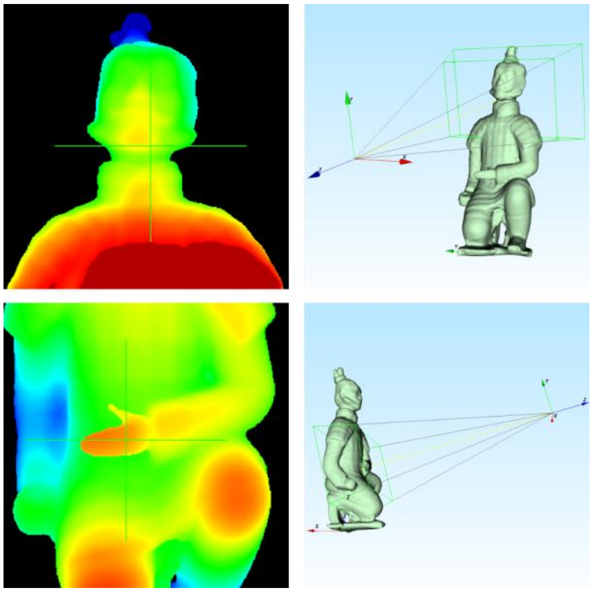

***Figure 3**. CultArm3D-P - Depth of field estimation of two camera view candidate (green: optimally focused, blue: far plane, red: near plane)*

In this scope, view planning describes the process of computing a sufficient set to camera views in position and orientation that captures the object of interest as complete as possible resulting in a 3D model with the desired quality [8]. This can be even an incremental process involving a feedback loop of planning, capturing and reconstructing, where the intermediate incremental reconstruction serves as the input for the next planning phase. It can be also regarded as an optimization problem, with the objective function of maximizing an overall model quality estimate while satisfying the safety constraints. The challenge however lies in the definition of a proper quality estimate that can be frequently evaluated during the scanning process and predicts / correlates with the desired quality of the final 3d model.

The CultArm3D features a hybrid approach that relies on little initial user input of setting a bounding and safety volume by defining a cylinder in height and diameter around the object on the turntable. Then a first set of approximated views can be quickly calculated and carried out mainly based on the volume size, the camera field of view and focus distance. Figure 3 shows the intermediate reconstruction result from an initial quick scan with 40 images reduced to low resolution. The point cloud density is automatically evaluated and low density areas and holes are identified (and highlighted with red color) and can be distinguished from areas with sufficient density (blue color).

Based on the intermediate 3D reconstruction a second more detailed set of views is planned utilizing rendering technologies and the calibrated camera intrinsics to simulate the effect of each view candidate. Figure 3 shows how the camera depth of field is visualized on the object. The view candidates are selected in such a way that the area in focus, which correlates with the local density gain, is maximized. Special attention is given to the previously identified low density areas that are often caused by occlusions. Additional views candidates are generated targeting those areas to eventually reach a sufficient density. With this new set of selected view candidates the next scan phase can be carried out by the robot and the intermediate 3D reconstruction is updated. The process repeats until the desired surface density.

A strong correlation was found between the density estimate on an intermediate point cloud and the surface quality of a resulting final 3D reconstructed with images in full resolution. Furthermore, careful focus planning is essential to the quality gain of intermediate and final the 3D model, especially because the use of lens autofocus focus functionality is not recommended for photogrammetry because it changes camera intrinsics for each view resulting in incomparable camera images. Therefore the lens focus stays fixed throughout the whole scanning task while the robot automatically adapts the distance to the surface with respect to the selected camera view candidate.

RESULTS

Gold is extremely difficult to digitize due to its highly reflective material properties. In a project with the Reiss-Engelhorn Museums in Mannheim, Germany, we used our CultLab3D-P to capture 20 very small gold artifacts from the island of Java in Indonesia (see Fig. 4 and 5). We autonomously scanned artefacts ranging from rings, earrings to bracelets in extremely high resolution with faithful color reproduction, so they would be displayed on 65” autostereoscopic displays in the exhibition showing their creators’ craftsmanship and enormous attention to detail.

***Figure 4**. Reiss Engelhorn Museums – JAVA Gold Exhibition: Earring with mythological Daemon, Java 14.- 15. AD © CES / 3D-Model: Fraunhofer IGD*

The reason why highly reflective surfaces are so hard to digitize, a problem equally found in structured light systems and the image-based approach of photogrammetry used by CultArm3D-P, is that both systems draw their knowledge of geometric depth from triangulation that is based on active or passive coding of the target surface.

In the first case, structured light is overlaid over the surface to allow spatial decoding, and thus a correlation of points on the encoded surface and the corresponding pixels in the camera image. In the second case, fine structural patterns already existing on the object surface found in one image are identified on the same exact surface location, observed from several different camera perspectives, again serving as basis for correlation of surface points and camera pixels for triangulation. In the case of reflective surfaces, the signal cast on the object surface, be it structured light or simply diffuse light required for photogrammetry, is reflected back to the camera, leading to overexposure of the sensor. Reducing exposure time, closing the aperture or reducing the light source intensity still leaves bright sensory readout in the regions of high reflectivity, depending on the camera and light angle in relation to the surface normal, while other regions are underexposed, leading to no usable correlation information.

***Figure 5**. Reiss Engelhorn Museums – JAVA Gold Exhibition: Earring with mythological Daemon, Java 14.- 15. AD © CES / 3D-Model: Fraunhofer IGD*

Instead of trying to find a suboptimal tradeoff between the components involved, we simply separate the received light feedback into the desired (diffuse) component, and the undesired (specular) channel. We can then use the purely diffuse information to extract geometric information, while entirely shutting out the specular compound. We achieve this by exploiting the physical effect of circular polarization, which is also used in photography. The effect only exists for mostly parallel light and observer direction (camera), which additionally implies that the light source must be as close as possible to the camera. We satisfy this requirement using diffuse ring lights of our own design which surround the lens, providing a narrow band of diffuse light around the lens aperture. Both the light source and the camera lens are equipped with circular polarizers that are tuned such that they let the diffuse light component pass while blocking the specular component.

The effect is based on the fact that light originating from the light source and being directly reflected off the target surface is phase-shifted during the reflection, and cannot pass the polarizer in front of the camera (analyzer) which is tuned accordingly.

Fig 6: Java Gold Exhibition at Reiss-Engelhorn-Museen. © CES / 3D-model: Fraunhofer IGD / United Screens GmbH.

All other light contributions that are reflected by the surface but find their way under different angles and numbers of reflections, pass the analyzer, and contribute to a well-lit diffuse image.

CONCLUSION

We presented our CultArm3D-P, an autonomous, color-calibrated 3D scanning robot yielding reproducible high quality 3D models (see Fig. 6) as an example of how to put automated 3D mass digitization for the digital preservation of entire collections to practice [9]. In line with the overall objective of the European Commission's Digital Agenda for Europe [10][11], our approach provides a solid foundation for future research and development of 3D technologies in the realm of cultural heritage. The system is highly flexible and can come in multiple configurations. It serves as a platform for future improvements and the inclusion of advanced technologies (e.g volumetric measurement sensors, ultrasound or others) towards 3D consolidated data models and therefore represents an important contribution to leveraging innovative digital technologies for the cultural heritage sector. The work conducted marks an important milestone in the journey of cultural heritage research towards a connected and digital future. It fosters new applications using high-quality 3D models which range from better visitor experiences in museums by innovative exhibition concepts based around 3D replicas, to new business models such as 'virtual loans' of 3D models or educational applications. Our CultLab3D developments offer the chance for cultural institutions to tap into additional revenue streams and secure their funding with complementary business models and enable digital heritage preservation and research for future generations to come.

ACKNOWLEDGMENT

This work was supported in part by German Ministry of Economy and Energy Affairs grant 01MT12022E "CultLab3D – fast economic digitization of cultural heritage in high quality".

References

[1] Antonio De Stefano, Reimar Tausch, Pedro Santos, Arjan Kuijper, Giuseppe Di Gironimo, Dieter W. Fellner, Bruno Siciliano, Modeling a virtual robotic system for automated 3D digitization of cultural heritage artifacts, Journal of Cultural Heritage, Volume 19, 2016,Pages 531-537, ISSN 1296-2074, https://doi.org/10.1016/j.culher.2015.11.008.

[2] Zhang, Z. (2000): A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), pp. 1330-1334.

[3] R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," in IEEE Transactions on Robotics and Automation, vol. 5, no. 3, pp. 345-358, June 1989. doi: 10.1109/70.34770

[4] Gao Yi, Luo Jianxin, Qiu Hangping and Wu Bo, "Survey of structure from motion," Proceedings of 2014 International Conference on Cloud Computing and Internet of Things, Changchun, 2014, pp. 72-76. doi: 10.1109/CCIOT.2014.7062508

[5] Özyeşil, O., Voroninski, V., Basri, R., & Singer, A. (2017). A survey of structure from motion. Acta Numerica, 26, 305-364. doi:10.1017/S096249291700006X

[6] Goesele M., Klowsky R., Ackermann J., Ritz M., Fuhrmann S., Langguth F., Maeijcke P.: Scene reconstruction from community photo collections. Computer 43 (06 2010), 48–53. 3, 9

[7] Michael Goesele u. a. „Multi-View Stereo for Community Photo Collections“. In: Proceedings of the

[8] 11th International Conference on Computer Vision (ICCV 2007). Rio de Janeiro, Brazil: IEEE, 2007, S. 265–270 (siehe S. 6).

[9] Scott, W. R., Roth, G., & Rivest, J.-F. (2003). View planning for automated three-dimensional object reconstruction and inspection. ACM Computing Surveys, 35(1), 64-96. http://dx.doi.org/10.1145/641865.641868

[10] Tausch, R.; Domajnko, M.; Ritz, M.; Knuth, M.; Santos, P.; Fellner, D. W., "Towards 3D Digitization in the GLAM (Galleries, Libraries, Archives, and Museums) Sector – Lessons Learned and Future Outlook", IPSI BgD Transactions on Internet Research (TIR), Special issue - "Digital Heritage and Related Tools", January 2020, Vol 16, No 1, ISSN 1820 – 4503, 2020

[11] European Commission: Commission recommendation of 27 October 2011 on the digitisation and online accessibility of cultural material and digital preservation (2011/711/EU) [online], available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:283:0039:0045:EN:PDF (accessed November 11, 2019).

[12] European Commission: (2011/711/EU) [s. above] and European Commission: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A Digital Agenda for Europe (COM(2010)245 final) [online], available from: http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52010DC0245&from=en (accessed November 11, 2019).

Author Biography

Pedro Santos is Head of the Competence Center for Cultural Heritage Digitization and with Fraunhofer Institute for Computer Graphics Research since 2002. His department develops the world's first approach for fast, economic, and automated 3D digitization of cultural heritage, recognized with the EU Prize for Cultural Heritage / Europa Nostra Award 2018 in the category of research. Pedro Santos is author/co-author of over 80 publications as well as reviewer for Eurographics, ACM and IIIE.