



Revista Brasileira de Computação Aplicada, April, 2020

DOI: 10.5335/rbca.v12i1.10019 Vol. 12, No 1, pp. 84-94

Homepage: seer.upf.br/index.php/rbca/index

ORIGINAL PAPER

Improved point clouds from a heritage artifact depth low-cost acquisition

Pedro Oliveira Raimundo ^{10,1} and Karl Philips Apaza-Agüero ^{10,1}

¹Department of Computer Science, Federal University of Bahia (UFBA)

*pedro.o.raimundo@gmail.com; kaguero@ufba.br

Received: 2019-09-30. Revised: 2020-02-21. Accepted: 2020-02-23.

Abstract

Data acquired using low-cost depth cameras exhibits undesirable traits such as low-resolution and high amount of noise, yielding point clouds with insufficient information from an object. This limits the use of such devices for 3D reconstruction of heritage artifacts. This work aims to improve low-cost depth acquisition by using a new approach based on Super-Resolution techniques. The proposed approach has been applied to several artifacts of the Federal University of Bahia Museum of Archaeology and Ethnology (MAE/UFBA). The results show that our approach improves the quality of point clouds generated from several heritage artifacts. Our analysis concludes that whenever additional geometry is obtained via the proposed method there is actual reconstruction of detail, while any geometry that is removed is usually related to the removal of inconsistencies or noise from the input data, without loss of detail.

Keywords: Digital Heritage; Low-Cost Acquisition; Point Clouds; Super-Resolution.

Resumo

Dados adquiridos através de câmeras de profundidade de baixo custo possuem características indesejáveis como baixa resolução e forte presença de ruído, gerando nuvens de pontos com informação insuficiente a respeito dos objetos. Assim, a aplicabilidade dessas câmeras para reconstrução 3D de artefatos culturais é limitada. O presente trabalho visa melhorar as aquisições 3D de baixo custo através de uma proposta baseada em técnicas de Super-Resolução. O método proposto foi aplicado a diversos artefatos do Museu de Arqueologia e Etnologia da Universidade Federal da Bahia (MAE/UFBA). Os resultados obtidos indicam que o método proposto melhora a qualidade das nuvens de pontos obtidas a partir dos artefatos avaliados. A análise realizada indica que a geometria adicional obtida através do método proposto representa reconstrução de detalhe real do artefato, enquanto a eventual remoção de informação geométrica geralmente indica a eliminação de ruído ou dados inconsistentes, sem que ocorra a perda de detalhes.

Palavras-Chave: Aquisição de Baixo Custo; Nuvens de Pontos; Preservação Digital; Super-Resolução.

1 Introduction

Projects such as the Digital Michelangelo (Levoy et al., 2000) and the Great Buddha (Miyazaki et al., 2000) show examples of successful 3D reconstructions to digital heritage. However, as was the case with these two projects, digital heritage missions frequently employ highly specialized, and costly, equipment and personnel. Nevertheless, some recent

reconstruction pipelines focus on using low-cost acquisition technologies (Newcombe et al., 2011, Dias et al., 2006). Potential advantages of these low-cost approaches include the possibility of real-time reconstruction; GPU acceleration; and cheap, portable and lightweight equipment. The difference in price between low-cost and traditional 3D acquisition is of at least several orders of magnitude. An Intel Realsense low-cost camera costs in range of US\$100

and a Microsoft Kinect V1 (now discontinued) costs around US\$50, whereas traditional laser scanners can cost from US\$10,000 to more than US\$100,000.

The Kinect V1 is the first version of a cheap motiontracking device based on a infrared emitter/receiver pair. The Kinect works by projecting a pattern of structured infrared light (Pavlidis et al., 2007) upon the scene, much like older low-cost reconstruction pipelines (Rocchini et al., 2001). The infrared receiver detects the distortions caused by the reconstructed scene on the pattern; based on these distortions and the known parameters of the acquisition hardware (distance between the sensors, etc.), it is possible to obtain the distance of the detected points to the infrared emitter and obtain depth images of the scene at a rate of 30 frames per second. Moreover, the Kinect V1 has an RGB camera which is used to capture color information at the same rate of the depth images.

The main difference between using a low-cost sensor and high-end acquisition hardware lies in the quality of the acquired data. The data provided by the Kinect scanner is noisy, low-resolution (both the depth maps and color images are captured at 640x480 resolution) and inconsistent (pieces of information appear and disappear even in successive frames). Cheaper acquisition hardware also indirectly limits the scale of the reconstruction targets, which is constrained by factors such as scanning resolution and operational range.

In this context, the central contribution of the present research is a new approach, based on Super-Resolution (SR) techniques (Nasrollahi and Moeslund, 2014), to enhance the 3D data obtained from a low-cost device. Our approach generates a high-resolution (HR) depth image from low-resolution (LR) depth images of the object. The associated 3D reconstruction pipeline processes this HR data and generates a point cloud with more detail and less noise compared to clouds generated from the original LR observations. The approach has been applied in a practical context of digital heritage, capturing several cultural objects of the UFBA Museum of Archaeology and Ethnology.

An important observation is that even though a Microsoft Kinect V1 was used during the acquisition phase of this work, the method is not tied to a specific depth sensor. Therefore, even better results could be possible with newer and more accurate low-cost acquisition devices such as the Microsoft Kinect V2 and the Intel Realsense depth cameras. Moreover, it should be noted that while several studies work with 3D meshes and point clouds directly, our method treats the acquired data as a grayscale 16-bit 2D image during the acquisition stage. This allows the usage of 2D image processing techniques instead of 3D geometry processing algorithms for tasks such as noise removal and registration, which ultimately yielded good results. Extra care was taken to never quantize the input data using less than 16 bits, thus preserving its original precision.

The rest of this work is structured as follows: Section 2 shows the previous work focused on improving data quality from low-cost acquisitions;

Section 3 details the proposed approach based on Super-Resolution; Section 4 presents experimental results, showing some cultural heritage artifacts used in this work; finally, Section 5 presents the conclusions of this work and some directions for future research.

Related Work

The acquisition phase of a 3D reconstruction pipeline is concerned with using one or more acquisition devices to capture the data of an object (geometry, color, etc.) that will be used throughout the other phases of a 3D reconstruction process (Bernardini and Rushmeier, 2002). For the acquisition of geometry, there is a variety of hardware with considerable differences. When selecting which device to use, the acquisition devices have to be compared regarding characteristics such as precision, flexibility, reconstruction speed, portability, and operational scale (Gomes et al., 2014). Regarding low-cost acquisition technologies, there are both active (which project some sort of light upon the scene) and passive (which use image data captured by RGB images) capture approaches, this work focuses on the former. Some arguments for the usage of active low-cost 3D acquisition techniques are: a) active sensors provide much faster captures because the depth information is calculated from one or more physical measurements instead of relying on image processing and feature matching techniques; b) active sensors are more robust regarding external factors such as changes in lighting and focal length; and c) active sensors perform better in reconstructing textureless and featureless surfaces.

A survey on low-cost 3D reconstruction of cultural heritage artifacts has already been proposed (Raimundo Nevertheless, due to limitations et al., 2018). of low-cost depth cameras, the raw data that they provide is usually noisy, low-resolution and inaccurate (Silva et al., 2013, Cui et al., 2013). A way to improve the quality of sensor data shown in several studies (Zollhöfer et al., 2015, Silva et al., 2013) is the utilization of SR techniques (Park et al., 2003, Richardt et al., 2012). A comprehensive survey and a thorough taxonomy of this area have been proposed (Nasrollahi and Moeslund, 2014). Super-Resolution is the process of obtaining HR images from one or more LR observations of the same object, where one or several parameters of the imaging model (position, focal length, noise model, etc.) vary between the LR images (Nasrollahi and Moeslund, 2014). In this context, resolution can refer to one or more image characteristics such as spatial resolution or temporal resolution. With that said, like most SR approaches (Nasrollahi and Moeslund, 2014, Huang and Yang, 2010), the proposed SR method aims to improve the spatial resolution of the images, increasing the amount of high-frequency information (i.e. object detail), by varying the position of the depth camera slightly between the captures. This is different from simple image interpolation because the latter only increases the amount of pixels of the input image,

without seeking to reconstruct detail.

The KinectFusion system (Izadi et al., 2011, Newcombe et al., 2011) performs real-time 3D reconstruction using the Microsoft Kinect V1 as its acquisition hardware. KinectFusion showed a way to improve the quality of raw data by applying a bilateral filter (Tomasi and Manduchi, 1998) to remove noise from the input data while preserving its edges and fine detail, providing cleaner input to the other steps of the reconstruction process. Nevertheless, even after applying a bilateral filter to the depth image, holes or missing detail could be generated because to the temporal inconsistency of Kinect. Also, given that real-time reconstruction reduces the viability of our approach and is not a requirement for heritage artifact reconstruction, this constraint is removed in the present work to allow the usage of computationally intensive Super-Resolution techniques to improve the quality of the data of captured cultural artifacts as much as possible.

3 Proposed Approach

For each captured artifact an acquisition protocol consisting of the number of depth captures, the angular displacement between captures, and the boundaries of the capture volume is established. Considering the strict quality requirements associated with digital heritage, it is also necessary to work directly with the sensor data instead of processed or filtered data provided by existing capture tools. Thus, an interactive tool (Fig. 1) which captures depth and color images and generates point clouds was developed to acquire the raw data from the depth sensor. This application can also be configured to capture a user-defined number of images in a single burst.

In an incremental fashion, two different techniques to solve the problems of low-resolution and noise present in data from low-cost 3D scanners were developed: Smooth Accumulation and Super-Resolution, both of which leverage the burst-capture functionality of the developed tool. While Smooth Accumulation was eventually replaced by a custom SR approach, it was used in several case studies and some of its ideas were reused, such as the utilization of more than one depth frame to improve the acquired depth image.

3.1 Smooth Accumulation

Initial assessments of the depth data showed that even after applying a bilateral filter (Tomasi and Manduchi, 1998) to the depth map, several holes were present and part of the geometric information was missing because of the temporal inconsistency of the depth stream. To tackle this issue, data from multiple depth frames was accumulated, maintaining information that was temporarily absent due to fluctuations in the sensor measurements. Eq. (1) formalizes this accumulation technique for two grayscale input images A and B to obtain a third image (C); as a convention, right

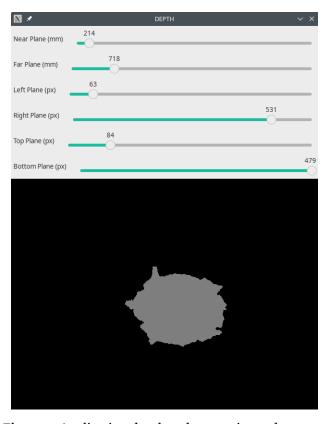


Figure 1: Application developed to acquire and process depth information. The slider controls at the top are used to define minimum and maximum thresholds (in millimeters) of the capture and the clipping limits (in pixels) of the image plane, effectively determining the capture volume.

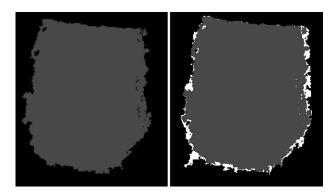


Figure 2: Smooth accumulation: an initial attempt to solve the problems of noise and lack of detail in the acquired data. The leftmost depth image shows the captured data without accumulation, and the rightmost one depicts the data accumulated from multiple depth frames, with the added data presented in white. The depth information appears as a single shade of gray here because of the leveling necessary to make the reconstructed object visible.

subscription is used to refer to the image element at row i and column j of an image:

$$accum(A,B) = C \text{ where } \begin{cases} C_{ij} = A_{ij} & \text{if } A_{ij} \neq 0 \\ C_{ij} = B_{ij} & \text{if } A_{ij} = 0 \end{cases}$$
 (1)

In practice, the resulting image C can be accumulated with the next captured frame and so on. Fig. 2 depicts the result of this process. The combination of this data accumulation technique and bilateral filtering is referred to as Smooth Accumulation.

Super-Resolution

The theoretical framework employed by most SR approaches assumes that each LR image is a warped, blurred, decimated and noisy version of the original HR image (Nasrollahi and Moeslund, 2014). Therefore, the SR problem consists of finding the correct set of transformations that turn each LR image back to its HR version and fuses their information in some way. Fig. 3 illustrates the procedure of going from several LR images to one HR image, through sequential noise removal, upsampling, deblurring, and image fusion operations.

After evaluating several third-party methods (Mitzel et al., 2009, Farsiu et al., 2004), a new SR approach was developed specifically to improve the quality of depth maps acquired using low-cost devices. The main motivation for developing a novel SR approach was that the results obtained through the evaluated approaches, geared towards images and videos of real-world scenes, suggested that these approaches do not correctly handle some traits of depth data (such as temporal inconsistency and gaps in the data, which are not common in real-world scenes), introducing artifacts in the final geometry.

The stages of the proposed Super-Resolution approach are established as follows: pre-processing stage, registration stage, upsampling stage, warping stage, and reconstruction stage. Although the names of these stages are somewhat novel, this approach still lies within theoretical SR frameworks outlined by related work (Park et al., 2003, Nasrollahi and Moeslund, 2014). Notwithstanding, this naming scheme was devised to allow the reader to further differentiate the proposed approach from existing ones and drill-down on the specifics of stages which are often bundled together in other studies.

Pre-processing: given the importance of using a priori information to improve the quality in other lowcost 3D reconstruction studies (Raimundo et al., 2018). This stage was included in the proposed SR approach to calculate the bounding box of the 3D volume corresponding to the HR image. The information calculated during this stage of the SR pipeline is used in latter steps of the SR pipeline to avoid introducing extreme values (i.e. values outside of the observed volume) and remove non-linear noise that might be introduced when performing the subsequent image transformations.

Registration: due to the acquisition protocol and nature of the captured artifacts, only global motion between the LR frames must be compensated in this study. Again, drawing from an image processing background, a 2D registration technique (ECC image alignment (Evangelidis and Psarakis, 2008)) was used to acquire a sub-pixel registration of the LR frames to a template (usually the first frame of the sequence). Using ECC, an affine model of the translation and rotation between the frames is obtained, which can then be applied to align the images (Fig. 4 - top). The key here is that the displacement between the subsequent LR frames must be small enough to be compensated accurately via rigid 2D alignment. The current approach yields good results with an angular displacement between 1° and 3° and linear displacement of about 1 cm. Despite that, the method is robust with regards to larger displacements, as invalid depth information generated by non-overlapping areas of the LR frames is filtered out in the next stages of the SR pipeline.

Upsampling: to take advantage of the subpixel alignment obtained during registration on the following stages of the SR pipeline, the LR images must be upsampled. In the proposed approach, nearestneighbor scaling is applied to the LR frames to avoid introducing invalid depth data at this stage of the SR pipeline. It has been determined that a scaling factor of 4 works well for 16 LR frames of input; nevertheless, this constant likely has to be adjusted for other acquisition protocols and devices. higher spatial resolution obtained from upsampling the images allows the non-redundant information present therein to be interwoven during the alignment and reconstruction stages. Fig. 4 (middle) shows how the registration obtained previously is still valid for the upsampled images, also due to the usage of nearest-

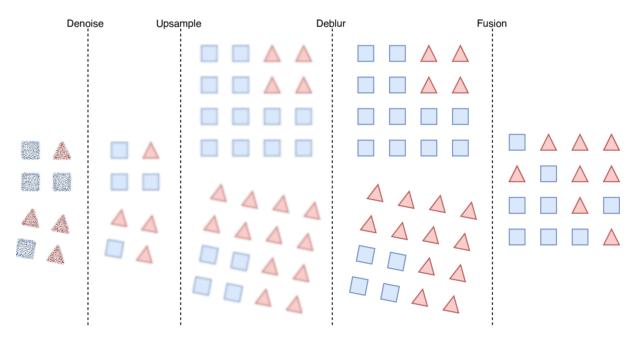


Figure 3: Illustrated summary of SR. The figure shows how multiple LR images can be mapped back to a single HR image via a set of successive image transformations, with the triangles and squares denoting different features in the images.

neighbor scaling, which preserves the features of the input data.

Warping: given the upsampled LR frames and the registration obtained previously, it is possible to perform an affine warp that corresponds the matching parts of the upsampled images (Fig. 4 - middle). Due to the higher pixel resolution obtained in the previous step, the sub-pixel displacements obtained via ECC in the registration phase now correspond to "whole" pixels. Thus it is possible to align these images with greater precision than before, without losing information because of aliasing. After applying the warp operations, the data from the LR frames is ready to be fused in the next step.

Reconstruction: following the lead of most other SR approaches, the usage of mean and median filters was evaluated for the reconstruction of the HR image. These are some effective and computationally efficient ways to fuse information from multiple images, which, in some cases, accurately reproduce the results obtained via complex analytical approaches (Nasrollahi and Moeslund, 2014). However, unlike what happens in general SR methods, which are geared towards regular color or grayscale images, a simple mean fusion did not yield good results for depth images due to the discontinuity of the data (Fig. 5 – top-left). As unreliable depth data is registered as a 0 or other invalid value by the depth sensor, simply averaging the values introduces additional noise on the depth map, a median fusion improved over these results but still left some invalid information (Fig. 5 - top-right), with the added disadvantage of introducing depth plateaus on the point cloud. In the final reconstruction approach, these problems were solved by eliminating zeros from

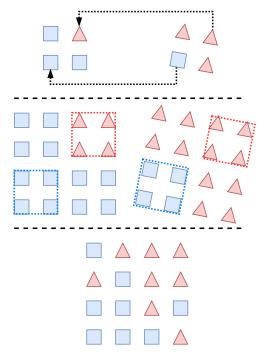


Figure 4: Representation of the registration of two hypothetical LR images and their combination into a single HR image. From top to bottom the registration, upsampling and reconstruction stages of the SR pipeline are illustrated.

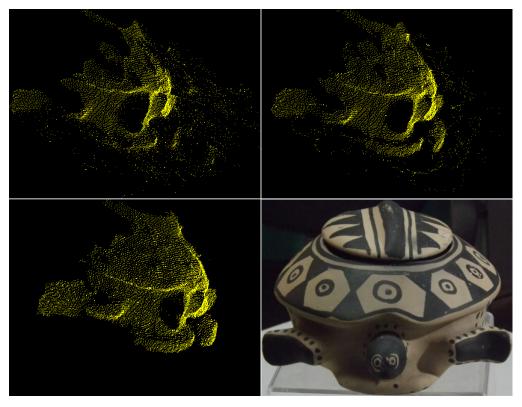


Figure 5: Visual comparison of the generated super-resolved point cloud using the mean (top-left), median (top-right) and zero-eliminating mean (bottom-left) operations during the reconstruction phase of the SR pipeline. For reference, the bottom-right shows a picture of the artifact.

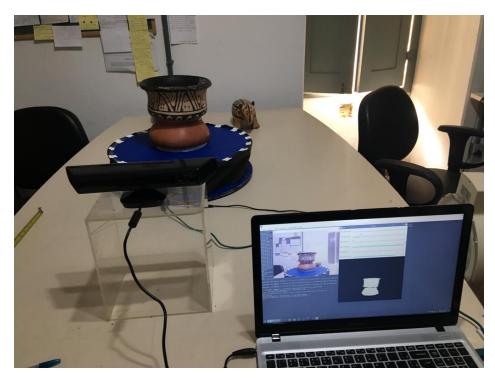


Figure 6: Photo of the experimental setup used throughout this evaluation. The setup consists of a laptop, a turntable, a Microsoft Kinect V1, and the heritage artifacts themselves, in addition to the transparent support used to position the Kinect at an adequate height.

the mean calculations during the image fusion, which ultimately yielded good results (Fig. 5 - bottom), with very little remaining noise (an expected side-effect of the mean operation) and smooth, credible, geometry.

Results

One of the main challenges faced by low-cost 3D reconstruction pipelines is the low-resolution, lack of detail, and high amount of noise of the depth data provided by the hardware (Cui et al., 2013). Therefore, considerable effort was dedicated to improve the results of the acquisition phase because if better geometric data is passed on to the next stages of a low-cost 3D reconstruction pipeline, better 3D models are produced. Henceforth, the results of applying the developed SR method to captured heritage artifacts are presented and discussed.

4.1 Experimental Setup

The main execution environment of the software developed within this work was a laptop computer with an 2.5 GHz Intel Core i7 6500-U processor, 8 GB of RAM and a discrete Geforce 940MX graphics card. The acquisition devices were a Microsoft Kinect V1 scanner and a turntable. An overview of the experimental setup is provided in Fig. 6.

Software tools were developed from third-party libraries for various purposes. OpenCV (Bradski and Kaehler, 2000) and PCL (Rusu and Cousins, 2011) were used to handle and process images and point clouds, OpenGL (Woo et al., 1999) was the graphics library used for the visualization module, and the libfreenect ¹ driver was used in the depth acquisition application to access and retrieve data from the Microsoft Kinect V1.

Two heritage artifacts were captured: a pot with fish-like carvings (heretofore denominated "Fish Pot", Fig. 7) and a turtle-shaped clay pan ("Turtle Pan", Fig. 8). Both pieces belong to the MAE/UFBA collection and are part of the material culture of the Waujá indigenous tribe and have been chosen for this analysis due to their differences in geometry, finishing and motifs.

4.2 Experimental results and discussion

The objective of the proposed SR pipeline was to improve over the smooth accumulation approach whilst still mitigating the two main traits of depth maps acquired using low-cost RGB-D hardware: heavy noise and low amount of detail. Fig. 9 illustrates the results of using the proposed SR pipeline on the Turtle Pan. With comparison to the raw depth map acquired from a Microsoft Kinect V1, more data is present in the capture (\approx 12,000 vs. \approx 10,000 captured points), several holes were filled, and the geometry is overall smoother.

Complementing what was already shown in Fig. 9,

¹https://github.com/OpenKinect/libfreenect



Figure 7: Fish Pot – Decorated clay pot crafted by members of the Waujá tribe. Part of the MAE/UFBA collection.



Figure 8: Turtle Pan - Turtle-shaped clay pan made by members of the Waujá tribe. Also part of the MAE/UFBA collection.

Fig. 10 shows further results of this approach depicting the geometry acquired from the Turtle Pan. The smoothing capabilities of the SR technique are evident in the image, as the plateaus in the data are much less noticeable; however, it is also important that details have not been lost in the operation and actually became more distinguishable. While in Fig. 10 – Left the shape of the object appears somewhat flattened, on Fig. 10 - Right approximates the surface of the object more accurately, and even presents a piece of the object which had not been captured previously.

Fig. 11 shows the SR results for the Fish Pot. In this case, the smoothing properties of the technique are still present, but the acquisition of additional geometry and reconstruction of details is more apparent. The overall shape and features of the scanned artifact are at the same time smoother and more well-defined. The original point cloud (without SR) also presents some distortion, which flattens the geometry of the object in the same way that happened to the Turtle Pan, while the SR version of the point cloud more closely approximates the round shape of the object.

Table 1 quantitatively presents the results of using the proposed SR technique on other point clouds of tested heritage pieces, the increase in captured vertices ranges from 2.97% to 38.44%. These results confirm that more vertices are obtained using SR than without. Together with a qualitative evaluation of the final reconstructions, this indicates that reconstruction of detail is obtained via the proposed method, while a

reduction in number of vertices, if any, would indicate the removal of noise or invalid geometry.

Conclusions

The depth data obtained from low-cost sensors is usually lacking in detail and consistency, which affects the quality of the models obtained in 3D reconstruction. Through the proposed Super-Resolution approach, it has been possible to enhance this data and reconstruct detail beyond what the sensor initially provides, as indicated by our experimental results. Moreover, the proposed method also improves the overall quality of the data in terms of smoothness and presence of holes.

Future research should focus on an extended evaluation of the current approach through both the reconstruction of more heritage artifacts and a quantitative analysis of the final 3D reconstructions, which entails the existence of some baseline ground truth. Another interesting goal for future works is the deployment of a complete 3D reconstruction pipeline for the preservation of heritage artifacts by museum staff. Such a pipeline should take the limitations of low-cost hardware and existing heritage practice into consideration to be as streamlined as possible.

Acknowledgments

The authors would like to thank the National Council for Scientific and Technological Development (CNPq) for funding this research via the project 406719/2016-0. The authors would also like to thank the entire MAE/UFBA staff for all their support.

References

Bernardini, F. and Rushmeier, H. (2002). The 3d model acquisition pipeline, Computer graphics forum 21(2): 149-172. https://doi.org/10.1111/1467-8659. 00574.

Bradski, G. and Kaehler, A. (2000). Opency, Dr. Dobb's journal of software tools 3.

Cui, Y., Schuon, S., Thrun, S., Stricker, D. and Theobalt, C. (2013). Algorithms for 3d shape scanning with a depth camera, IEEE Transactions on Pattern Analysis and Machine Intelligence 35(5): 1039-1050. https:// doi.org/10.1109/TPAMI.2012.190.

Dias, P., Matos, M. and Santos, V. (2006). reconstruction of real world scenes using a low-cost 3d range scanner, Comp.-Aided Civil and Infrastruct. Engineering 21: 486-497. https://doi.org/10.1111/j. 1467-8667.2006.00453.x.

Evangelidis, G. D. and Psarakis, E. Z. (2008). Parametric image alignment using enhanced correlation coefficient maximization, Transactions on Pattern Analysis and Machine Intelligence 30(10): 1858-1865. https://doi.org/10.1109/TPAMI. 2008.113.

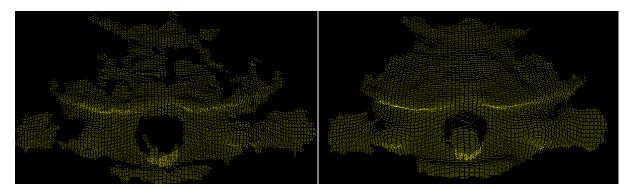


Figure 9: Point cloud captured using a low-cost 3D scanner (Microsoft Kinect V1) before (left) and after using (right) the proposed SR approach on 16 LR depth frames. The hole-filling and smoothing properties of the technique are evident, and the features of the artifact are more recognizable in general.



Figure 10: Acquired point cloud before (left) and after (right) SR processing - Turtle Pan. Upon closer inspection, it is possible to see that the plateaus in the data are thoroughly smoothed by the SR approach, without sacrificing detail.

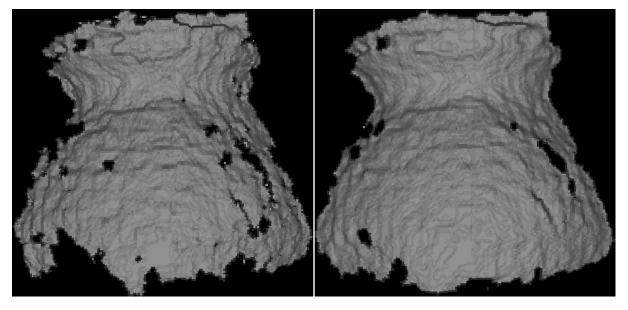


Figure 11: Acquired point cloud before (left) and after (right) SR processing - Fish Pot. The smoothing properties of the SR approach are once again visible, but in this case the reconstruction of additional geometry is clearer.

Artifact (View)	No. of Points (raw)	No. of Points (SR)	Geometry Delta (%)
Turtle Pan (0°)	10,165	13,041	+28.29
Turtle Pan (80°)	11,357	11,694	+2.97
Turtle Pan (180°)	7,524	10,416	+38.44
Turtle Pan (260°)	12,922	14,590	+12.91
Fish Pot (0°)	19,514	20,638	+5.8
Fish Pot (80°)	19,585	20,228	+3.3
Fish Pot (180°)	19,506	20,441	+4.8
Fish Pot (260°)	21,287	21,876	+2.8

Table 1: No. of Vertices of several point clouds with regards to the utilization of SR

- Farsiu, S., Robinson, M. D., Elad, M. and Milanfar, P. (2004). Fast and robust multiframe super resolution, IEEE transactions on image processing 13(10): 1327–1344. https://doi.org/10.1109/TIP.2004.834669.
- Gomes, L., Bellon, O. R. P. and Silva, L. (2014). 3d reconstruction methods for digital preservation of cultural heritage: A survey, Pattern Recognition Letters **50**: 3-14. https://doi.org/10.1016/j.patrec.2014.03.
- Huang, T. and Yang, J. (2010). Image super-resolution: Historical overview and future challenges, Superresolution imaging, CRC Press, pp. 19-52. https: //doi.org/10.1201/9781439819319.
- Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A. and Fitzgibbon, A. (2011). Kinectfusion: Real-time 3d reconstruction and interaction using a moving depth camera, Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, UIST '11, ACM, New York, NY, USA, pp. 559-568. http://doi.acm.org/10.1145/ 2047196.2047270.
- Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J. et al. (2000). digital michelangelo project: 3d scanning of large statues, Proceedings of the 27th annual conference on Computer graphics and interactive techniques, ACM Press/Addison-Wesley Publishing Co., pp. 131-144. https://doi.org/10.1145/344779.344849.
- Mitzel, D., Pock, T., Schoenemann, T. and Cremers, D. (2009). Video super resolution using duality based tv-l1 optical flow, in J. Denzler, G. Notni and H. Süße (eds), Pattern Recognition, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 432-441. https://doi.org/10. 1007/978-3-642-03798-6_44.
- Miyazaki, D., Ooishi, T., Nishikawa, T., Sagawa, R., Nishino, K., Tomomatsu, T., Takase, Y. and Ikeuchi, K. (2000). The great buddha project: Modelling cultural heritage through observation, VSMM2000 (6th international conference on virtual systems and multimedia), pp. 138-145.
- Nasrollahi, K. and Moeslund, T. B. (2014). Superresolution: a comprehensive survey, Machine vision and applications 25(6): 1423-1468. https://doi.org/ 10.1007/s00138-014-0623-4.

- Newcombe, R. A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A. J., Kohi, P., Shotton, J., Hodges, S. and Fitzgibbon, A. (2011). Kinectfusion: Realtime dense surface mapping and tracking, Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on, IEEE, pp. 127-136. https://doi.org/ 10.1109/ISMAR.2011.6092378.
- Park, S. C., Park, M. K. and Kang, M. G. (2003). Super-resolution image reconstruction: a technical overview, IEEE signal processing magazine 20(3): 21–36. https://doi.org/10.1109/MSP.2003.1203207.
- Pavlidis, G., Koutsoudis, A., Arnaoutoglou, F., Tsioukas, V. and Chamzas, C. (2007). Methods for 3d digitization of cultural heritage, Journal of cultural heritage 8(1): 93-98. https://doi.org/10.1016/j. culher.2006.10.007.
- Raimundo, P., Apaza-Agüero, K. and Apolinário, A. L. (2018). Low-cost 3d reconstruction of cultural heritage artifacts, Revista Brasileira de Computação Aplicada 10(1): 66-75. https://doi.org/10.5335/rbca. v10i1.7791.
- Richardt, C., Stoll, C., Dodgson, N. A., Seidel, H.-P. and Theobalt, C. (2012). Coherent spatiotemporal filtering, upsampling and rendering of rgbz videos, Computer Graphics Forum 31(2pt1): 247-256. https: //doi.org/10.1111/j.1467-8659.2012.03003.x.
- Rocchini, C., Cignoni, P., Montani, C., Pingi, P. and Scopigno, R. (2001). A low cost 3d scanner based on structured light, Computer Graphics Forum 20(3): 299-308. https://doi.org/10.1111/1467-8659.00522.
- Rusu, R. B. and Cousins, S. (2011). 3D is here: Point Cloud Library (PCL), IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China. https://doi.org/10.1109/ICRA.2011.5980567.
- Silva, J. W., Gomes, L., Agüero, K. A., Bellon, O. R. P. and Silva, L. (2013). Real-time acquisition and superresolution techniques on 3d reconstruction, Image Processing (ICIP), 2013 20th IEEE International Conference on, IEEE, pp. 2135-2139. https://doi.org/10.1109/ ICIP.2013.6738440.
- Tomasi, C. and Manduchi, R. (1998). Bilateral filtering for gray and color images, Computer Vision, 1998. Sixth International Conference on, IEEE, pp. 839-846. https: //doi.org/10.1109/ICCV.1998.710815.

- Woo, M., Neider, J., Davis, T. and Shreiner, D. (1999). OpenGL Programming Guide: The Official Guide to Learning OpenGL, Version 1.2, 3rd edn, Addison-Wesley Longman Publishing Co., Inc., USA.
- Zollhöfer, M., Siegl, C., Vetter, M., Dreyer, B., Stamminger, M., Aybek, S. and Bauer, F. (2015). Low-cost real-time 3d reconstruction of large-scale excavation sites, *J. Comput. Cult. Herit.* 9(1). https://doi.org/10.1145/2770877.