UTILISING UV-MAPPING FOR THE 3D-POINT CLOUD SEGMENTATION OF ARCHITECTURAL HERITAGE

FATIH UZUN[1], SEVGI ALTUN[2], SENA KAYASÜ[3], BERKAY ÖZTÜRK[4], YUSUF ŞAHIN[5], GÖZDE ÜNAL[6] and MINE ÖZKAR[7]
[1,3,4,5,6,7] *Istanbul Technical University*
[2] *The Chinese University of Hong Kong*
[1] *uzunfa17@itu.edu.tr, 0000-0002-8428-5482*
[2] *sevgialtun@cuhk.edu.hk, 0000-0002-5872-6985*
[3] *kayasu20@itu.edu.tr, 0000-0001-9281-6784*
[4] *ozturkber20@itu.edu.tr, 0000-0001-9343-8375*
[5] *sahinyu@itu.edu.tr, 0000-0002-4134-9423*
[6] *gozde.unal@itu.edu.tr, 0000-0001-5942-8966*
[7] *ozkar@itu.edu.tr, 0000-0003-0999-7401*

Abstract. This paper presents a procedural framework to process raw photogrammetric data intended for 3D point cloud segmentation of masonry structures. Raw data is segmented in order to obtain operable 3D models through processes that increasingly integrate computational methods, which can improve efficiency and accuracy. The idea is to improve the quality of the initial data, which can then be used to train a machine-learning system in identifying these materials more accurately. The approach incorporates a high-poly detailed mesh model generated through photogrammetry. The detailed model serves as a reference to extract colour information that we project onto a custom-created, low-poly representation of the dome architectural element, ensuring a precise fit with the target model. The model to utilise UV maps and height maps to preprocess data across various scales is a step towards facilitating the documentation and conservation of historic structures with an awareness of architectural knowledge.

Keywords. architectural heritage, unit-based, masonry documentation, 3D point cloud segmentation, UV mapping, point cloud processing

1. Introduction

Architectural conservation is reliant on documentation and full-scale surveys to record historical value as well as to inform any intervention. With contemporary tools, it is possible to create accurate, digital three-dimensional (3D) documentation of architectural objects for documentation required for surveys, analyses, and restoration. Point clouds are raw, 3D representations of real-life information that are collected through methods such as photogrammetry or laser scanning. They are useful in architectural heritage documentation because of their potential to create operable 3D

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 2, 313-322.

models. These models could facilitate the analysis, monitoring, and even reconstruction of historical structures. However, utilising point clouds for these purposes presents unique challenges.

Point clouds are widely used in 3D vision tasks across different fields. The structure offers that in an n-dimensional space with n being greater than zero, an n-tuple (x1, ..., xn) is referred to as a point. We can regard a point cloud as a collection of points that holds a meaningful arrangement. Given that we inhabit a 3D environment, we can describe a point within a point cloud as (x, y, z), where x, y, and z represent the point's coordinates relative to the x, y, and z axes.

Raw point clouds do not express meanings of architectural components. They must be segmented either manually or through computational means such as machine learning and image recognition for further analysis and accurate modelling. Different from the 2D images and videos, point clouds do not represent a structural form that can easily be processed through convolutional neural networks. Point cloud-specific machine learning models such as DGCNN (Wang et al., 2019), PointNet (Qi et al., 2017) and GACNet (Han et al., 2022) are produced for that reason. Computational methods are increasingly being utilised to automate the segmentation process in order to improve its speed and accuracy in architectural cases. In one such study, Francesca et al. (2020) offered a point cloud dataset for Italian cultural heritage sites Santo Stefano Street and the Church of St. Stefano and segmented it into walls, floor, arches, columns and other architectural elements. Battini et al. (2024) created a synthetic dataset to teach the machine learning methods to segment different types of vaults and lunettes and tested its results on the Ducal Palace of Urbino. Moyano et al. (2021) used a morphological segmentation algorithm to segment bricks, doors and other structures from a façade of Pilatos Palace. They also obtained the point clouds through laser scanning, which uses laser beams to take distance measurements and generate 3D models. Laser scanning is not very practical in many other cases due to high upfront cost and lack of portability because of its reliance on specific equipment.

Photogrammetry is an effective method of obtaining point clouds and 3D models of architectural objects through a simple camera. It is easily applicable and highly accessible due to its low upfront cost and non-invasive nature (Lanzara et al. 2022). It can also be a highly precise method, depending on the camera, lighting conditions, and environmental set up. Randazzo et al. (2020) showed that photogrammetric examination provides a higher resolution image than laser scanners as they determined the stone material deterioration of a historical church. Similarly, Karataş et al. (2022) combined photogrammetric methods with laser scans to obtain a better visualisation of the damage in a historical railway station. Photogrammetry can be used to model and document architectural elements that are physically difficult to reach via the integration of drone photography. This is particularly important for elements such as domes, which may be difficult to reach when using traditional methods of documentation.

A challenge for point clouds generated through photogrammetry is that they often lack depth on specific surfaces, as they are based on two-dimensional photographs. Taking photographs of architectural elements from different angles can mitigate this issue in certain cases, but it is not always possible. For instance, the details on a curved brick surface may appear flush in a point cloud even though they may not be in reality. This makes analyses and high-precision documentation of intricate surfaces difficult.

Existing applications in point cloud segmentation have primarily focused on a larger scale, such as building-scale stylistic segmentation and architectural elements, but there is still need for research on the smaller scale.

The change in scale introduces some difficulties. Smaller elements require higher levels of detail for successful segmentation. This creates complications, especially in cases where the documented element is physically out of reach, and image resolution is low. Although there are some solutions to improve the level of detail during initial documentation, these considerably increase the cost. There is a need for applications to balance cost, accuracy, and level of detail in order to maximise the potential of photogrammetry in architectural conservation.

Presented here is a procedural framework that utilises UV mapping and a post-processing workflow to provide approximate depth to visual and point cloud data and facilitate the training of deep learning systems to address the challenges mentioned above. The framework aims to enhance segmentation efficiency in photogrammetric documentation, specifically focusing on the segmentation of brick types and mortar in the dome of a historical building in Central Anatolia. The framework employs a high-poly detailed mesh model generated through photogrammetry as a reference, enabling the precise subtraction of colour information. Projection of colour information onto a custom-created, low-poly representation of the dome ensures an accurate fit with the target model. The transfer of intricate details from the high-poly model to the low-poly model is accomplished using the technique of normal mapping, which maps the surface orientations of the high-poly model onto the low-poly model's surface. This pre-processing step prepares the raw data for subsequent three-dimensional point cloud segmentation workflows, resulting in improved accuracy and efficiency in the semantic segmentation process. Consequently, synthetic data sets can be created for supervised machine-learning algorithms that operate on colour data. Even though post processing may introduce some errors, these are negligible when segmenting a dome into its visual patterns and obtain the details. Hence, enhancing the photogrammetric process through UV mapping could allow a better representation and more accurate segmentation.

2. Materials and Methods

The historic building utilised as a case study was Sahip Ata Hanikah, abbreviated as SAH from now onwards. SAH is a 13th-century building within a Seljuk period complex located in Konya, Türkiye. The glazed bricks constituting the main dome were mostly reconstructed based on previous documentation in the late 2000s. These were deemed sufficient for this study as they retain the visual and geometric properties of similar historic domes in the region (Kayın 2017). As part of the study, we filtered the brick pattern forming the inner surface of the main dome of SAH at the brick-unit level. The photogrammetric model, the low-poly representative model, and the UV mapping on to this model served as tools for the filtering process. Subsequently, we prepared a dataset for machine learning using the displacement map created for 3D point cloud segmentation. This model follows the workflow shown in Figure 1.

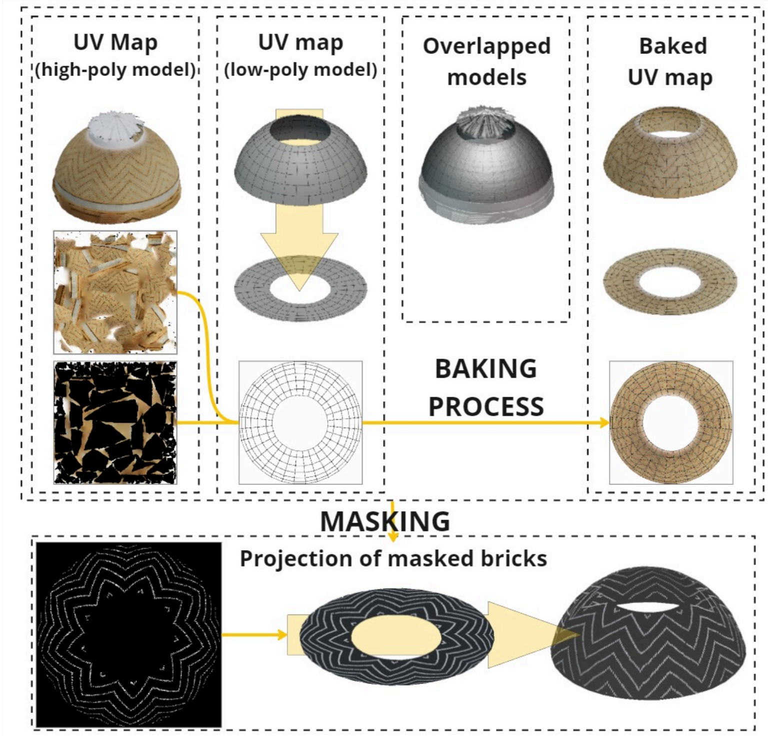

Figure 1. Pipeline for preprocessing data with UV Projection

2.1. THE PREPARATION OF POINT CLOUDS

The photogrammetry-based mesh model for the central dome is characterised by a high polygon count and intricate detailing. Additionally, the UV map of the relevant model is automatically generated by photogrammetry software, leading to an irregular UV map unsuitable to be used as a filter. Consequently, we identified the need for a representative low-poly model. This model boasts a reduced polygon count and a UV map generated through a projection method. Subsequently, we adjusted the representative model in a Blender 3D environment to achieve optimal alignment with the photogrammetry-based high-poly model.

To prepare the low-poly model that is seamlessly aligned with UV mapping and baking stages, we imported the high-poly model into Blender 3D in .fbx format and adjusted its orientation of the model. Subsequently, we represented the dome with a half sphere in the low-poly model. We utilised the "Shrinkwrap" command to achieve a perfect alignment between the photogrammetric model and the representation model. The primary objective of this process is to reduce the distance between the surface normals of the two models.

UV maps are commonly employed in computer graphics to project regularly sampled 2D data, including colours, normals, or displacements, onto surfaces within a

3D environment (Poranne et al., 2017). We found that the UV map of the prepared low-poly representative model needs to be unfolded. At this stage, we transferred the colour data on the dome surface obtained through Photogrammetry along the Cartesian plane's Z-axis. Therefore, we open the UV map of the Low-poly model with an orthographic view from the -Z direction. This allows the diffuse map texture images present in the high-poly model to align seamlessly with the UV map of the low-poly model during the baking process.

Finally, to generate the diffuse map, we created a blank image with a resolution of 2048 x 2048 px. Subsequently, we incorporated the colour data from the high-poly model into the UV map of the low-poly model. Following these steps, we exported the created diffuse map in (.png format), thus concluding the modelling process. Consequently, the diffuse map has been prepared and masked to attain a suitable format for generating the height map. The UV maps of the photogrammetry model and the low-poly representation model can be observed in Figure 2.

Figure 2. High-poly and low-poly models' UV maps of SAH

2.1.1. The 2D Filtering of Images

In the orthographic unwrapping and baking of the UV map, bricks become smaller and more bricks get distorted as we move outwards from the centre of the dome. However, this variation does not pose an issue for the filtering process and the preparation of the resulting data. To perform the desired filtering, we selected bricks in Photoshop according to their colour information and generated a black-and-white mask. Since we could address the entire surface in one go, executing this process once sufficed.

The number of generated images depends on the variety of the ornaments. In SAH, there are three different types of bricks which are turquoise, black and natural coloured ones and they are all filtered separately into three images. The number of generated images could be changed according to the focused analysis. A UV-unwrapped, single-texture image is imported into Photoshop and made colour sampling from different areas in the image because of the colour differences on the surface. The number of samples changes between 10 to 20 depending on the reflections or the deformations on the brick surfaces. In the next step of precisely defining the edges with the “Select and Mask” tool, two main settings, namely edge detection and contrast, are changed. For a 2048 x 2048 px sized texture, a 2 px radius size is enough to determine the edges of brick units. Utilising edge detection allows small discrepancies to be mended and separation between bricks to be clearly defined. Additionally, increasing the contrast level to 100% helps to create a complete black-and-white image with sharp edges. This process is repeated for all different colours or patterns on the surface of the studied building. The final black-and-white image is exported in .jpeg file format for use as a displacement map.

2.1.2. Displacement Map

The displacement corresponds to a scalar adjustment applied along the normal of the mesh at the specified point, and the map is depicted through a grayscale image that encodes the magnitude of displacement to be implemented at that specific point (Guidi and Angheleddu, 2016). Displacement maps are created using gradients between black and white colour tones, where black represents the lowest point, and white signifies the highest point. This allows for the preparation of data suitable for machine learning algorithms. With the displacement map crafted as in the section above, we can impart volume to the surface of the previously prepared low-poly model. In Blender3D software, the “Subdivision” command under the modify tab is applied to obtain a higher-resolution model from the low-poly model prior to bringing in the displacement map. The operations at this stage do not affect the UV map.

In the subsequent step, the “Displace” command under the modify tab defines a new and empty image to the low-poly model, selecting UV from the Coordinates option. The previously prepared displacement map is imported by utilising the Open option in the Texture Properties menu. Through this method, the height level of the surface texture of the bricks on the model becomes parametrically controllable.

2.2. THE EXPERIMENTAL SETUP

All experiments were carried out using an NVIDIA A5000 RTX GPU and an Intel(R) Core (TM) i7-7700K CPU @ 4.20 GHz, operating within the PyTorch framework. For the assessment of segmentation performance, the SAH's point cloud was reduced to nearly 2 million points via distance-based Poisson Disk sampling, thus obtaining a homogenous distribution. Following this, a slice representing 10% of the dome's hemisphere, containing around 200,000 points, was chosen for the training phase (as in Figure 3). Given that both DGCNN and CurveNet feature point aggregation operators, which combine each point’s features with those derived from its dynamic or graph-walk neighbours, and have demonstrated notable efficacy in various segmentation tasks, we employed these architectures in our training. The dome was segmented into four categories: mortar, naked brick, turquoise-glazed brick, and black-glazed brick.

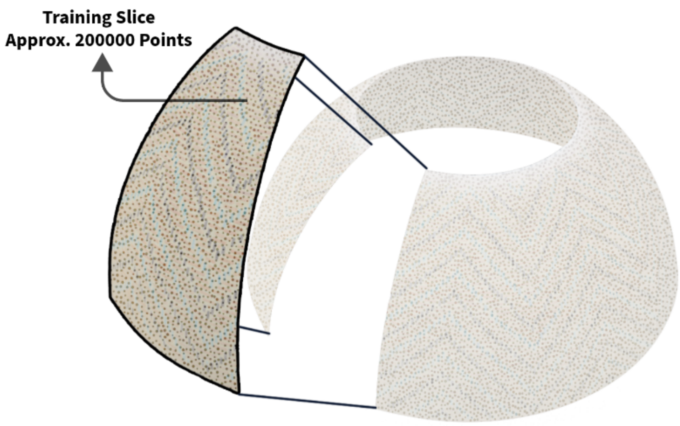

Figure 3. Division of the test and training data

In the training procedure, random windows containing 10,000 points are cropped from the training slice and fed to the networks. To augment the training data and make the networks more generalizable, random rotations of 15 degrees in XYZ-directions are applied. Also, on point colours, a jittering augmentation is applied to change the saturation slightly. Networks are trained using Adam optimizer with a learning rate of 1e-4 and a batch size of 32. During the testing phase, the networks are fed with patches, each consisting of a point and its nearest 10,000 neighbours. Then, for the intersecting points between multiple patches a majority voting is applied (as in Figure 4).

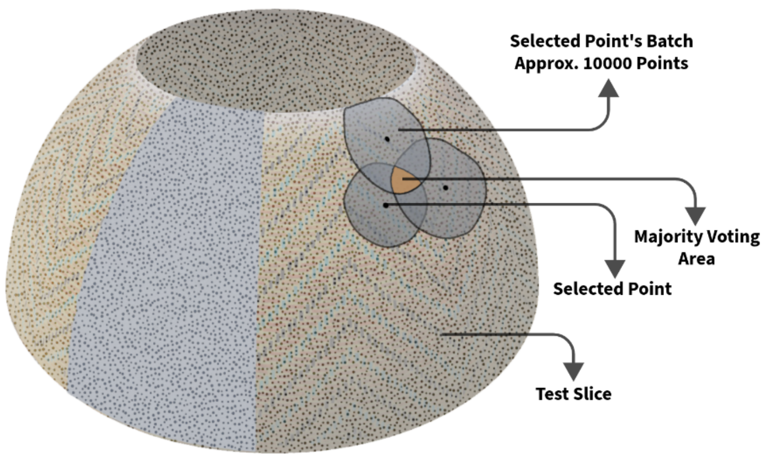

Figure 4. Patch selection during testing

3. Results & Discussion

We trained the DGCNN and CurveNet machine-learning models using photogrammetric data with and without UV mapping. In order to compare the efficacy of our proposed method, we compared accuracy using two different metrics: mIoU (mean Intersection over Union) and point accuracy. mIoU measures the accuracy of a segmentation model by calculating the intersection over union for each class and then taking the mean across all classes. We obtained the following scores.

Table 1. Accuracy comparison of UV mapping technique

	With UV Mapping		Without UV Mapping	
	mIoU	Point Accuracy	mIoU	Point Accuracy
DGCNN	72.49	91.62	71.22	90.56
CurveNet	68.21	77.46	67.64	75.65

The results indicate that the UV-mapping method provides incremental improvements when training a machine-learning system for automated segmentation.

As seen in the table, there is approximately a one percent increase in segmentation accuracy achieved through the UV-mapping method. Though this increase may seem minor numerically, it represents a quantitative evidence of our UV-mapping alongside the visual enhancement of the structure. Such a contribution can be attained with minimal effort invested via the UV-mapping method.

The quality of photogrammetric data is heavily dependent on-site conditions during the survey process. During the photogrammetric survey of SAH, the survey permit allowed for photographs to be taken from the ground-floor level with existing lighting conditions. As a result of the limitations in the conditions of our survey, it was difficult to gather more information about the surface details such as the depth difference between bricks and mortar through photogrammetry or manual measurements. In such situations where site conditions are suboptimal, enhancing the obtained data becomes important to get more precise results and to achieve a higher success rate.

It must be noted that the proposed method was used to create synthetic data, and that its congruence with the physical characteristics of the dome are undetermined because the study did not involve finer measurements of depth between bricks and mortar.

The use of synthetic data when investigating historical buildings is controversial, with ethical practices when applying machine learning to architectural surveys still being determined. However, synthetic data enhanced through such methods as UV mapping may be useful in the development of computer-vision and machine-learning workflows, which can then be tweaked and corrected as new and higher-quality data is collected.

The proposed methods contribute to the segmentation process by enhancing the level of detail for low-quality photogrammetric data and improving the accuracy of outcomes. However, it is important to note that the efficiency of the method is directly proportional to the quality of the data obtained. In other words, as the quality of the photogrammetric data improves, the effectiveness of the proposed method in segmentation results also increase.

4. Conclusion

If there is little or no distinct difference in colour between multiple unit elements of a historical building, manually masking them in the digital data becomes an error-prone, tedious and time-consuming process, turning into a bottleneck. Methods for preprocessing and preparing data play a critical role in advancing this field, allowing for the documentation and conservation of historic structures while preserving architectural knowledge.

The proposed method can be applied to cases that require depth information to enhance the features of architectural elements in a digital model. While it is not essential to regulate the UV map for geometries with flat surfaces, the method can contribute by enhancing the difference in depth between units. For models with complex and curved surfaces, unwrapping the UV-map with the projection method provides further ease of operation. For its adaptation to different cases, it is essential to understand and preserve the geometrical characteristics of the unit element in that architectural case and the overall surface while defining the initial geometry of the low-poly model.

The proposed method also serves the preparation of parametric layers in the Historic Building Information Model (HBIM) through segmentation with the help of reduced mesh geometry for the BIM environment. It helps to reduce the computational cost by using representative models but preserving the essential information.

Acknowledgements

This research is supported by TÜBİTAK (The Scientific and Technological Research Council of Türkiye) under project number 119K896. Many thanks to Mustafa Cem Güneş for taking photographs of the dome as well as Alican Mertan and Oğuzhan Karabacak for their contribution to the implementation of semantic segmentation.

References

Battini, C., Ferretti, U., De Angelis, G., Pierdicca, R., Paolanti, M., & Quattrini, R. (2024). Automatic generation of synthetic heritage point clouds: Analysis and segmentation based on shape grammar for historical vaults. Journal of Cultural Heritage, 66, 37-47. https://doi.org/10.1016/j.culher.2023.10.003

Guidi, G., & Angheleddu, D. (2016). Displacement mapping as a metric tool for optimizing mesh models originated by 3D digitization. Journal on Computing and Cultural Heritage, 9(2), 1-23. https://doi.org/10.1145/2843947

Han, B., Zhang, X., & Ren, S. (2022). PU-gacnet: Graph attention convolution network for point cloud Upsampling. *Image and Vision Computing*, 118, 104371. https://doi.org/10.1016/j.imavis.2021.104371

Karataş, L., Alptekin, A., & Yakar, M. (2023). Investigating the material deteriorations on the facades of stone structures by terrestrial laser scanning method: Case study of Mardin mansion. *Afyon Kocatepe University Journal of Sciences and Engineering*, 23(3), 700-711. https://doi.org/10.35414/akufemubid.1197891

Kayin, A. (2017). Konya Sâhip Ata Hanikahı Çinileri. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (37), 206-221.

Lanzara, E., Catello, D., & Casciello, M. (2022). Scan to HBIM for complex reflective metal artefacts. 3d digitisation and restoration. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-2/W1-2022, 121-128. https://doi.org/10.5194/isprs-archives-xlviii-2-w1-2022-121-2022

Matrone, F., Lingua, A., Pierdicca, R., Malinverni, E. S., Paolanti, M., Grilli, E., Remondino, F., Murtiyoso, A., & Landes, T. (2020). A benchmark for large-scale heritage point cloud semantic segmentation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2020, 1419-1426. https://doi.org/10.5194/isprs-archives-xliii-b2-2020-1419-2020

Moyano, J., León, J., Nieto-Julián, J. E., & Bruno, S. (2021). Semantic interpretation of architectural and archaeological geometries: Point cloud segmentation for HBIM parameterisation. *Automation in Construction*, 130, 103856. https://doi.org/10.1016/j.autcon.2021.103856

Poranne, R., Tarini, M., Huber, S., Panozzo, D., & Sorkine-Hornung, O. (2017). Autocuts. *ACM Transactions on Graphics*, 36(6), 1-11. https://doi.org/10.1145/3130800.3130845

Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). Pointnet: Deep learning on point sets for 3d classification and segmentation. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 652-660).

Randazzo, L., Collina, M., Ricca, M., Barbieri, L., Bruno, F., Arcudi, A., & La Russa, M. F. (2020). Damage indices and photogrammetry for decay assessment of stone-built cultural heritage: The case study of the San Domenico church main entrance Portal (South Calabria, Italy). *Sustainability*, 12(12), 5198. https://doi.org/10.3390/su12125198

Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., & Solomon, J. M. (2019). Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 38(5), 1-12. https://doi.org/10.1145/3326362