STAG: Smart Tools and Applications in Graphics (2024)
T. G. Dulecha, U. Castellani, and A. Giachetti (Editors)

*Poster*

# Semantic Stylization and Shading via Segmentation Atlas utilizing Deep Learning Approaches

S. N. Sinha[1], P. J. Kühn[1], R. Rojtberg[1], H. Graf[1], A. Kuijper[1,3] and M. Weinmann[2]

[1]Fraunhofer IGD, Germany
[2]Delft University of Technology, Netherlands
[3]TU Darmstadt, Germany

**Abstract**

*We present a novel hybrid approach for semantic stylization of surface materials of 3D models while preserving shading. Based on a hybrid approach that builds on directly applying style transfer on the object surface obtained by learning-based or traditional methods such as 3D scanners or structured light systems, thereby overcoming artifacts like halos, ghosting or lacking quality of the geometric representation produced by other 3D stylization methods. For this purpose, our methods involves (i) the initial generation of a segmentation map parameterized over the object surface inferred based on a deep-learning-based foundation model to guide the stylization and shading of different regions of the 3D model, and (ii) a subsequent 2D style transfer that allows the exchange or stylization of surface materials in high quality. By delivering high-quality semantic perceptive reconstructions in a shorter timeframe than current approaches using manual 3D segmentation and stylization, our approach holds significant potential for various application scenarios including creative design, architecture and cultural heritage.*

**CCS Concepts**

• ***Computing methodologies*** → ***Computer graphics; Reflectance modeling; Image manipulation; Scene understanding;***

## 1. Introduction

Advancements in deep learning-based image analysis and manipulation techniques, such as style transfer [GEB15] and segmentation [KMR*23], have improved stylization and segmentation of 2D images. Recent learning-based 3D scene representations that were built on volume rendering techniques, notably Neural Radiance Fields (NeRF) [MST*20] and 3D Gaussian Splatting (3DGS) [KKLD23], have shown great promise in photorealistic novel view synthesis applications. Examples of respective extensions support material asset generation [JTL*23], 3D style transfer [SGW24a, LZX*24], scene segmentation [QLZ*23], and spectral scene representation [SGW24b]. However, despite their potential, methods based on such representations often face limitations in terms of producing halos, ghosting, and inferior mesh quality, especially in polychromy reconstruction [Øs19]. Despite further challenges with specular and reflective surfaces, recent innovations like GaussianShader [JTL*23] and StyleGaussian [LZX*24] additionally improve neural rendering by extending these methods for reflectance estimation and style transfer. Overcoming the aforementioned challenges is crucial for enriching applications in digital media and cultural heritage preservation, potentially leading to more accurate and efficient reconstructions and immersive experiences. Our approach addresses these aforementioned challenges by first segmenting the scene mesh using a segmentation atlas and then semantically applying 2D style transfer and shading to accurate reconstructions obtained from high-quality 3D scanning devices. We propose a hybrid pipeline that leverages a real-time texturing approach [RK19] and atlas-based UV map approaches [WAA*17], with an aim to enhance mesh quality, material estimation and adaptability to new styles [NPLX22, WJC*23]. Our main contributions include a hybrid approach integrating deep learning with traditional methods of manual asset creation, and the development of an atlas-based segmentation methodology for 3D semantic stylization and shading. With this approach, we can project 2D image information onto 3D and thus map 2D segmentation masks onto 3D space.

## 2. Methodology

Our novel method (Fig. 1) integrates deep learning-based models like SAM [KMR*23] into traditional 3D mesh and material reconstruction pipeline by first generating a segmentation atlas [RK19] and then using it for semantic shading (applying different shading models and/or material assets to different segments of the 3D model) and stylization of meshes.

**Multi-view renderings and segmentation-map generation:** The 3D model is processed in BlenderProc [DWS*23], generating multi-view images through all-around camera animation to ensure comprehensive coverage for the segmentation atlas. We then gener-



DOI: 10.2312/stag.20241352

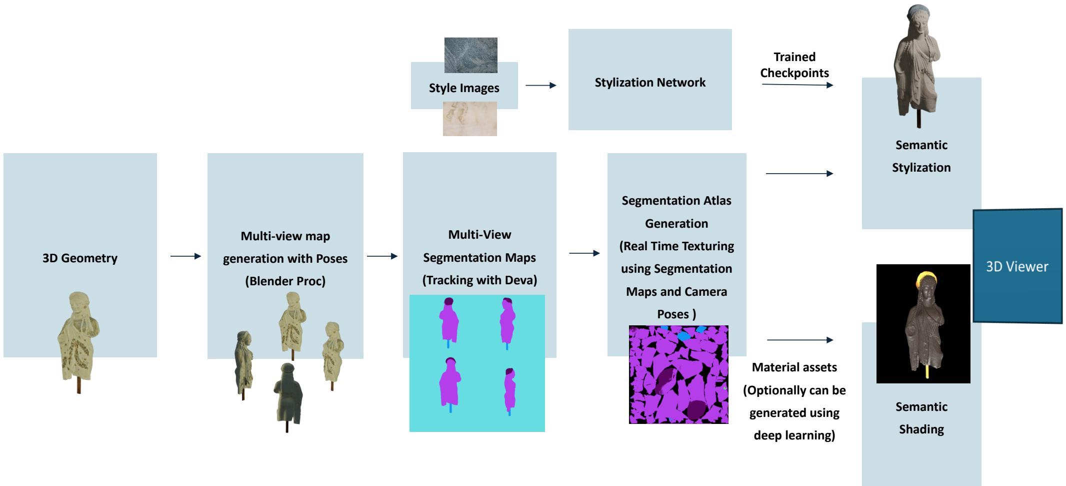


**Figure 1:** *Pipeline of our hybrid approach: The 3D model was rendered from various views with BlenderProc. Multi-view segmentation maps were subsequently generated using learning-based methods. The resulting segmentation atlas was then employed for semantic stylization or shading of the 3D model.*

ate multi-view segmentation maps (i.e. segmented maps of an object from multiple viewpoints) using the Tracking-Anything-with-Deva technique [COP*23] and SAM [KMR*23].

**Segmentation atlas generation:** Since CAD data typically lacks texture atlas coordinates, we generate the texture atlas using Blender's 'Smart UV Project' algorithm [Ble24], and then generate the segmentation atlas by projecting the model per view and mapping the texture into a 2D atlas [RK19].

| Segment | Average IoU Score |
|---|---|
| | 0.856 |
| | 0.798 |
| | 0.999 |
| | 0.955 |
| Average | 0.902 |

**Table 1:** *Average IoU for each segment (colors represent each segment as in Fig. 2) is calculated between the ground truth multi-view maps and the view generated by the segmented atlas.*

**Semantic shading and stylization:** We leverage a semantic atlas, deep stylization models and various shading models for region-specific stylization and shading, enhancing the visual appeal of 3D models in a hybrid pipeline, implemented in Three.js which uses WebGL for rendering [Bur12, CT82, Bli77, Xr, SGGW24].

## 3. Evaluation

We conducted our experiments on the Core dataset [KPR*13] using a machine with an Nvidia RTX 3090 graphics card.

**Quantitative and Qualitative Analysis:** We validated our segmentation atlas by calculating the Intersection over Union (IoU) [Cho21] with ground truth segmentation maps produced by SAM [KMR*23], revealing an overall average of 0.902 and indicating the potential for further enhancements through improved UV parametrization and atlas generation methods. For qualitative analysis we show that we can stylize different parts of the statue with different stones as shown in Fig. 2

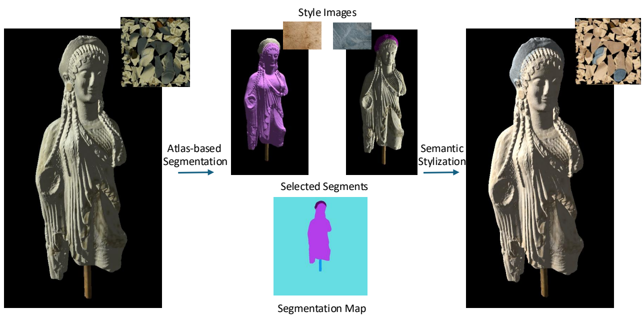


**Figure 2:** *Qualitative analysis: Different segments of the statue rendered using different stones using our pipeline*

## 4. Conclusion and Future work

We presented a hybrid pipeline for stylizing and shading distinct regions of a 3D model, which offers great potential for creative content creation in various applications including design, cultural heritage and architecture. The quality of the atlas generated depends on the segmentation maps derived over multiple frames. Future work may aim to optimize this process through AI-powered tools and the further integration of state-of-the-art atlas generation techniques, thereby further enhancing the pipeline's ability to meet industry needs for high-quality 3D model stylization and shading.

**Acknowledgement** The work in this paper was partially funded by the European Commission for the PERCEIVE project (grant agreement 101061157).

## References


[Ble24] BLENDER: Smart uv project, 2024. Accessed: Oct. 25, 2024. URL: https://docs.blender.org/manual/nb/2.79/editors/uv_image/uv/editing/unwrapping/mapping_types.html. 2

[Bli77] BLINN J. F.: Models of light reflection for computer synthesized pictures. In *Proceedings of the 4th Annual Conference on Computer Graphics and Interactive Techniques* (1977), Association for Computing Machinery. doi:10.1145/563858.563893. 2

[Bur12] BURLEY B.: Physically-based shading at disney. URL: https://api.semanticscholar.org/CorpusID:7260137. 2

[Cho21] CHO Y.: Weighted intersection over union (wiou): A new evaluation metric for image segmentation. *CoRR abs/2107.09858* (2021). URL: https://arxiv.org/abs/2107.09858, arXiv:2107.09858. 2

[COP*23] CHENG H. K., OH S. W., PRICE B., SCHWING A., LEE J.-Y.: Tracking anything with decoupled video segmentation. In *ICCV* (2023). 2

[CT82] COOK R. L., TORRANCE K. E.: A reflectance model for computer graphics. *Association for Computing Machinery* (1982). URL: https://doi.org/10.1145/357290.357293. 2

[DWS*23] DENNINGER M., WINKELBAUER D., SUNDERMEYER M., BOERDIJK W., KNAUER M., STROBL K. H., HUMT M., TRIEBEL R.: Blenderproc2: A procedural pipeline for photorealistic rendering. *Journal of Open Source Software* (2023). URL: https://doi.org/10.21105/joss.04901, doi:10.21105/joss.04901. 1

[GEB15] GATYS L. A., ECKER A. S., BETHGE M.: A neural algorithm of artistic style, 2015. 1

[JTL*23] JIANG Y., TU J., LIU Y., GAO X., LONG X., WANG W., MA Y.: Gaussianshader: 3d gaussian splatting with shading functions for reflective surfaces. *arXiv* (2023). 1

[KKLD23] KERBL B., KOPANAS G., LEIMKÜHLER T., DRETTAKIS G.: 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics 42*, 4 (July 2023). URL: https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/. 1

[KMR*23] KIRILLOV A., MINTUN E., RAVI N., MAO H., ROLLAND C., GUSTAFSON L., XIAO T., WHITEHEAD S., BERG A. C., LO W.-Y., DOLLÁR P., GIRSHICK R.: Segment anything. *arXiv:2304.02643* (2023). 1, 2

[KPR*13] KEIL J., PUJOL L., ROUSSOU M., ENGELKE T., SCHMITT M., BOCKHOLT U., ELEFTHERATOU S.: A digital look at physical museum exhibits: Designing personalized stories with handheld augmented reality in museums. In *2013 Digital Heritage International Congress (DigitalHeritage)* (2013), vol. 2, pp. 685–688. doi:10.1109/DigitalHeritage.2013.6744836. 2

[LZX*24] LIU K., ZHAN F., XU M., THEOBALT C., SHAO L., LU S.: Stylegaussian: Instant 3d style transfer with gaussian splatting, 2024. 1

[MST*20] MILDENHALL B., SRINIVASAN P. P., TANCIK M., BARRON J. T., RAMAMOORTHI R., NG R.: Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV* (2020). 1

[NPLX22] NGUYEN-PHUOC T., LIU F., XIAO L.: Snerf: Stylized neural implicit representations for 3d scenes. *arXiv* (2022). 1

[QLZ*23] QIN M., LI W., ZHOU J., WANG H., PFISTER H.: Langsplat: 3d language gaussian splatting. *arXiv* (2023). 1

[RK19] ROJTBERG P., KUIJPER A.: Real-time texturing for 6d object instance detection from rgb images. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (2019). doi:10.1109/ISMAR-Adjunct.2019.00-25. 1, 2

[SGGW24] SINHA S. N., GORSCHLUETER F., GRAF H., WEINMANN M.: Deepmaterialinsights: A web-based framework harnessing deep learning for estimation, visualization, and export of material assets from images. In *Proceedings of the 29th International ACM Conference on 3D Web Technology* (New York, NY, USA, 2024), Association for Computing Machinery. URL: https://doi.org/10.1145/3665318.3677152. 2

[SGW24a] SINHA S. N., GRAF H., WEINMANN M.: Semanticsplatstylization: Semantic scene stylization based on 3d gaussian splatting and class-based style transfer. In *Eurographics Workshop on Graphics and Cultural Heritage* (2024), The Eurographics Association. doi:10.2312/gch.20241256. 1

[SGW24b] SINHA S. N., GRAF H., WEINMANN M.: Spectralgaussians: Semantic, spectral 3d gaussian splatting for multi-spectral scene representation, visualization and analysis, 2024. URL: https://arxiv.org/abs/2408.06975, arXiv:2408.06975. 1

[WAA*17] WILSON D., ANGLIN C., AMBELLAN F., GREWE C. M., TACK A., LAMECKER H., DUNBAR M., ZACHOW S.: Validation of three-dimensional models of the distal femur created from surgical navigation point cloud data for intraoperative and postoperative analysis of total knee arthroplasty. *International Journal of Computer Assisted Radiology and Surgery 12*, 12 (2017), 2097 – 2105. URL: https://link.springer.com/content/pdf/10.1007%2Fs11548-017-1630-5.pdf, doi:10.1007/s11548-017-1630-5. 1

[WJC*23] WANG C., JIANG R., CHAI M., HE M., CHEN D., LIAO J.: Nerf-art: Text-driven neural radiance fields stylization. *IEEE Transactions on Visualization and Computer Graphics* (2023). 1

[Xr] X-RITE: Total appearance capture ecosystem. Accessed: 2024-09-22. URL: https://www.xrite.com/categories/appearance/total-appearance-capture-ecosystem. 2

[Øs19] ØSTERGAARD J. S.: "reconstruction" of the polychromy of ancient sculpture: A necessary evil? *Technè* (2019). 1



Proceedings published by Eurographics - The European Association for Computer Graphics.