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

Surface Reconstruction from Silhouette and Laser Scanners as a Positive-Unlabeled Learning Problem

M. Gottardo, M. Pistellato, and F. Bergamasco

DAIS, Università Ca'Foscari di Venezia, 155, Via Torino, Venice, Italy

Abstract

Typical 3D reconstruction pipelines employ a combination of line-laser scanners and robotic actuators to produce a point cloud and then proceed with surface reconstruction. In this work we propose a new technique to learn an Implicit Neural Representation (INR) of a 3D shape $\mathcal{S}$ without directly observing points on its surface. We just assume being able to determine whether a 3D point is exterior to $\mathcal{S}$ (e.g. observing if the projection falls outside the silhouette or detecting on which side of the laser line the point is). In this setting, we cast the reconstruction process as a Positive-Unlabelled learning problem where sparse 3D points, sampled according to a distribution depending on the INR's local gradient, have to be classified as being interior or exterior to $\mathcal{S}$. These points, are used to train the INR in an iterative way so that its zero-crossing converges to the boundary of the shape. Preliminary experiments performed on a synthetic dataset demonstrates the advantages of the approach.

CCS Concepts

• ***Computing methodologies*** $\rightarrow$ ***Shape modeling; Shape representations; Reconstruction;***

1. Introduction

Surface reconstruction is a central topic in Computer Vision and Graphics. This process has wide-ranging applications across various fields such as cultural heritage [PTB20], 3D shape segmentation [PBAT19], and industrial quality control [Hai17]. The task can be approached through a variety of methods, with both traditional scanning techniques [PBC*18] and Machine Learning-based models [HWW*24, SS23]. Traditional methods typically involve extracting a point cloud, followed by mesh generation to obtain the surface [KBH06, LLX*23]. Modern approaches often leverage machine learning models, particularly Multi-Layer Perceptrons (MLPs). These models, however, tend to prioritize realistic rendering over precise surface reconstruction. One of the most representative class of models in this category is Neural Radiance Fields (NeRF) [MST*21, XXP*22], which represents a significant advancement in realistic object rendering. Our method is tailored for any machine vision scenario where an accurate 3D reconstruction is needed. Most sensing devices, like laser scanners, give a sampling of the object surface in terms of a (reasonably dense) point cloud in space. For basic tasks the sparse point cloud is sufficient, but for others, a volumetric description of the object occupancy is equally important. However, creating voxel representations or triangulated surfaces from a point cloud are complex tasks per se, sometimes representing the most significant source of error in the final result. This is particularly true when the point clouds are noisy due to non-Lambertian reflecting materials [YWLG23] or if some regions can not be directly observed because of self-occlusions [LLW*23].

2. The Proposed Method

Our method is based on directly constructing an implicit representation of the object to reconstruct. The shape is described as an occupancy function $\mathcal{S} : \Omega \subset R^3 \rightarrow \{-1 \dots 1\}$ where a point (x,y,z) is inside the object iff $\mathcal{S}(x,y,z) \leq 0$ and outside when positive. The given representation is implicit because the surface is not directly encoded, but obtainable as the zero level-set of the function. To represent $\mathcal{S}$, we rely on a coordinate-based neural representation as a highly effective alternative over discrete representations like voxel grids. Our choice is not driven solely by the popularity of such models, but is justified by some properties that we can exploit for the reconstruction process. In particular: (i) Implicit Neural Representations (INR) are continuous functions (mathematically speaking, not accounting for the floating point approximations) so there is no need to discretize the space, and (ii) INRs are analytically derivable with respect to the input. Therefore, a neural occupancy network can convey information about the surface location (zero level-set) and the surface normal (∇S) without numerical differentiation.

2.1. Occupancy estimation as a PU learning problem

A straightforward way to train $\mathcal{S}$ is to provide a dataset of 3D points labelled as 1 if outside the object and -1 if inside. This labelling is reasonable in computer graphics applications where a triangular mesh is already known, but an implicit representation is needed for rendering purposes. On the other hand, sensing devices used for 3D reconstruction can reliably determine if a point is exterior to an object (i.e. should be labelled as 1) but not if it is interior.

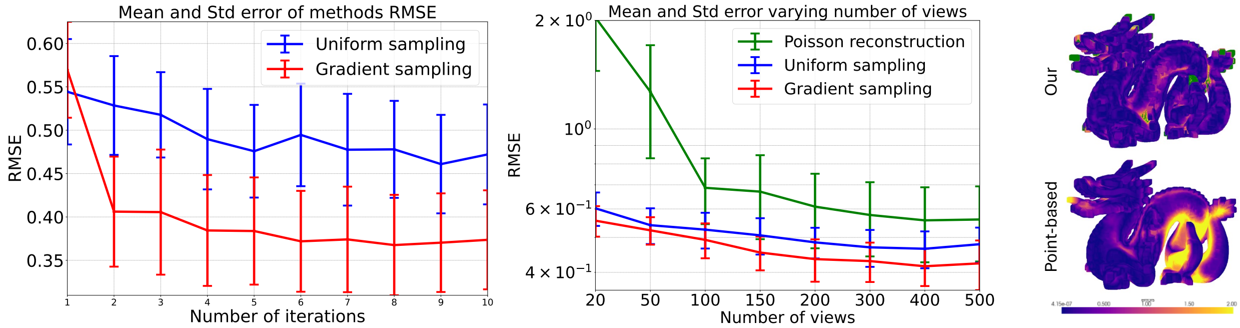

Figure 1: *Left: average reconstruction error increasing the iterations for uniform and gradient-based sampling. Center: surface error increasing the number of views for different approaches. Right: per-vertex error for our approach and point-based surface reconstruction.*

Suppose a simple space-carving application where we observe an object from multiple points of view. By extracting the silhouette (with any suitable foreground-background segmentation) we can reliably state if any 3D point is exterior to the object simply by checking if the projection falls outside the silhouette. Indeed, seeing the background implies that no solid matter is occluding the light ray exiting the camera. Note however that, from a single image, if the projection falls inside the silhouette we cannot state certainly that it is an interior point. The same applies to a laser scanner, where a 3D point is surely exterior if its projection falls on the side of the projected laser line where the epipole of the laser projector is located. When applied to a neural-based occupancy function, this asymmetry in determining if a point is interior or exterior can be cast as a Positive-Unlabelled (PU) learning problem [JS19, BD20].

2.2. Reconstruction pipeline

Assume to observe an object from different points of view with instruments capable of determining if a 3D point $\mathbf{x} \in R^3$ is surely exterior to the object. That is, if $\mathbf{x}$ is exterior the label will be 1 (with no error). Otherwise, the "unknown label" 0 is given, meaning it is impossible to determine if the point is in or out. All the 3D points $\mathbf{x}_1, \mathbf{x}_2, \ldots$ are realizations of a continuous (3-dimensional) random variable $\mathbf{X}$ generating points on the object's surface according to a probability density function f. Moreover, we assume to know the extent of a bounding box β containing the object. The proposed reconstruction pipeline then proceeds as follows:

0. initialization Let f be a continuous uniform distribution in β. This reflects the initial uncertainty of the surface, equally likely in β. Let $\mathcal{S}$ be the INR initialized with random weights.

1. sampling Sample M 3D points $\mathbf{x}_1, \ldots \mathbf{x}_M$ from $\mathbf{X}$ and use the data acquired from the optical sensors to assign a label (1 or 0) to each point. This results in a sequence $(\mathbf{x}_1, l_1), \ldots, (\mathbf{x}_M, l_M)$ of Positive-Unlabelled data. This operation is performed with Gibbs' sampling using the (discretized) CDF of f.

2. PU learning Solve a PU learning problem to assign the unknown points $(\mathbf{x}_i, 0)$ to either the exterior $(\mathbf{x}, 1)$ or interior $(\mathbf{x}_i, -1)$ classes. We use a K-NN classifier as discussed in [JS19].

3. train Use the labelled data to train $\mathcal{S}$ until convergence. In our setup, each $\mathbf{x}_i$ and l_i are the input and expected values to learn, respectively. Binary cross-entropy is used as the loss function.

4. f update Use $\|\nabla \mathcal{S}\|$ as a proxy to compute a new empirical PDF f. Since the surface is the zero-level set of $\mathcal{S}$, and it is constant inside or outside the object, the magnitude of $\nabla \mathcal{S}$ is related to the likelihood of a point to lie on the object's surface. Therefore, we compute $\|\nabla \mathcal{S}\|$ on a uniform discrete grid in β and normalize so that it sums to one to make it a distribution.

5. loop Return to step 1 and use the updated f to sample a new set of points. The process is repeated until a specified amount of iterations is reached.

The idea is to train the INR iteratively: in each iteration, a set of points are randomly sampled from a distribution exhibiting higher density closer to the object surface. In this way, we reserve more data where the function abruptly changes, and therefore more information is needed.

3. Experiments and Conclusion

Preliminary experiments were performed with a new synthetic dataset created using the state-of-the-art raytracer Mitsuba [JSR*22] to simulate line laser scanners observing an object rotating on a turntable. We included 6 different meshes from the Stanford Scanning Repository. Figure 1 displays some results in term of average surface accuracy (point-based RMSE) with respect to the ground truth. In the leftmost plot we show the average error during the training process for our model using a simple uniform sampling and the proposed gradient-based sampling updated at each iteration. As the number of iteration increases, the gradient-based sampling offers a better accuracy in representing the surface. The central plot shows the average error increasing the number of acquired views: we report results for uniform sampling, gradient sampling (proposed) as well as a standard point-based triangulation followed by surface reconstruction as proposed in [KBH06]. Compared to point-based reconstructions, our method can better represent the object surfaces even with a smaller number of views. Finally, we show a qualitative plot of the object "dragon" where per-vertex error is displayed for our method and the point-based one: we can observe a relevant improvement in some areas where the acquired point cloud is sparse due to self-occlusions. As a future work, we plan to include a more in-depth analysis of the method capabilities, involving experiments on the PU classification approach, a dataset extension and real-world scenarios.

References

[BD20] BEKKER J., DAVIS J.: Learning from positive and unlabeled data: A survey. *Machine Learning 109*, 4 (2020), 719–760. 2

[Hai17] HAIBO Y.: Industrial design applications of surface reconstruction algorithm based on three dimensional point cloud data. In *2017 International Conference on Robots & Intelligent System (ICRIS)* (2017), IEEE, pp. 178–181. 1

[HWW*24] HUANG Z., WEN Y., WANG Z., REN J., JIA K.: Surface reconstruction from point clouds: A survey and a benchmark. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2024). 1

[JS19] JASKIE K., SPANIAS A.: Positive and unlabeled learning algorithms and applications: A survey. In *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)* (2019), IEEE, pp. 1–8. 2

[JSR*22] JAKOB W., SPEIERER S., ROUSSEL N., NIMIER-DAVID M., VICINI D., ZELTNER T., NICOLET B., CRESPO M., LEROY V., ZHANG Z.: Mitsuba 3 renderer, 2022. https://mitsuba-renderer.org. 2

[KBH06] KAZHDAN M., BOLITHO M., HOPPE H.: Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing* (2006), vol. 7. 1, 2

[LLW*23] LI Z., LONG X., WANG Y., CAO T., WANG W., LUO F., XIAO C.: Neto: neural reconstruction of transparent objects with self-occlusion aware refraction-tracing. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2023), pp. 18547–18557. 1

[LLX*23] LEI N., LI Z., XU Z., LI Y., GU X.: What's the situation with intelligent mesh generation: A survey and perspectives. *IEEE transactions on visualization and computer graphics* (2023). 1

[MST*21] MILDENHALL B., SRINIVASAN P. P., TANCIK M., BARRON J. T., RAMAMOORTHI R., NG R.: Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM 65*, 1 (2021), 99–106. 1

[PBAT19] PISTELLATO M., BERGAMASCO F., ALBARELLI A., TORSELLO A.: Robust cylinder estimation in point clouds from pairwise axes similarities. p. 640 – 647. doi:10.5220/0007401706400647. 1

[PBC*18] PISTELLATO M., BERGAMASCO F., COSMO L., GASPARETTO A., RESSI D., ALBARELLI A.: Neighborhood-based recovery of phase unwrapping faults. vol. 2018-August, p. 2462 – 2467. doi:10.1109/ICPR.2018.8546052. 1

[PTB20] PISTELLATO M., TRAVIGLIA A., BERGAMASCO F.: Geolocating time: Digitisation and reverse engineering of a roman sundial. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12536 LNCS* (2020), 143 – 158. doi:10.1007/978-3-030-66096-3_11. 1

[SS23] SAMAVATI T., SORYANI M.: Deep learning-based 3d reconstruction: a survey. *Artificial Intelligence Review 56*, 9 (2023), 9175–9219. 1

[XXP*22] XU Q., XU Z., PHILIP J., BI S., SHU Z., SUNKAVALLI K., NEUMANN U.: Point-nerf: Point-based neural radiance fields. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2022), pp. 5438–5448. 1

[YWLG23] YANG R., WANG Y., LIAO S., GUO P.: Dpps: A deep-learning based point-light photometric stereo method for 3d reconstruction of metallic surfaces. *Measurement 210* (2023), 112543. 1

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