

Minimal Neural Atlas: Parameterizing Complex Surfaces with Minimal Charts and Distortion

Weng Fei Low[✉] and Gim Hee Lee[✉]

Institute of Data Science (IDS), National University of Singapore
NUS Graduate School’s Integrative Sciences and Engineering Programme (ISEP)
Department of Computer Science, National University of Singapore
{wengfei.low, gimhee.lee}@comp.nus.edu.sg
<https://github.com/low5545/minimal-neural-atlas>

Abstract. Explicit neural surface representations allow for exact and efficient extraction of the encoded surface at arbitrary precision, as well as analytic derivation of differential geometric properties such as surface normal and curvature. Such desirable properties, which are absent in its implicit counterpart, makes it ideal for various applications in computer vision, graphics and robotics. However, SOTA works are limited in terms of the topology it can effectively describe, distortion it introduces to reconstruct complex surfaces and model efficiency. In this work, we present *Minimal Neural Atlas*, a novel atlas-based explicit neural surface representation. At its core is a fully learnable parametric domain, given by an implicit probabilistic occupancy field defined on an open square of the parametric space. In contrast, prior works generally predefine the parametric domain. The added flexibility enables charts to admit arbitrary topology and boundary. Thus, our representation can learn a minimal atlas of 3 charts with distortion-minimal parameterization for surfaces of arbitrary topology, including closed and open surfaces with arbitrary connected components. Our experiments support the hypotheses and show that our reconstructions are more accurate in terms of the overall geometry, due to the separation of concerns on topology and geometry.

Keywords: Surface representation · 3D shape modeling

1 Introduction

An explicit neural surface representation that can faithfully describe surfaces of arbitrary topology at arbitrary precision is highly coveted for various downstream applications. This is attributed to some of its intrinsic properties that are absent in implicit neural surface representations.

Specifically, the explicit nature of such representations entail that the encoded surface can be sampled exactly and efficiently, irrespective of its scale and complexity. This is particularly useful for inference-time point cloud generation, mesh generation and rendering directly from the representation. In contrast, implicit representations rely on expensive and approximate isosurface extraction and ray casting. Furthermore, differential geometric properties of the surface can

also be derived analytically in an efficient manner [5]. Some notable examples of such properties include surface normal, surface area, mean curvature and Gaussian curvature. Implicit neural representations can at most infer such quantities at approximated surface points. Moreover, explicit representations are potentially more scalable since a surface is merely an embedded 2D submanifold of the 3D Euclidean space.

Despite the advantages of explicit representations, implicit neural surface representations have attracted most of the research attention in recent years. Nevertheless, this is not unwarranted, given its proven ability to describe general surfaces at high quality and aptitude for deep learning. This suggests that explicit representations still have a lot of potential yet to be discovered. In this work, we aim to tackle various shortcomings of existing explicit neural representations, in an effort to advance it towards the goal of a truly faithful surface representation.

State-of-the-art explicit neural surface representations [18,5,13,27] mainly consists of neural *atlas*-based representations, where each *chart* is given by a *parameterization* modeled with neural networks, as well as a predefined open square *parametric domain*. In other words, such representations describe a surface with a collection of neural network-deformed planar square patches.

In theory, these representations cannot describe surfaces of arbitrary topology, especially for surfaces with arbitrary connected components. This is clear from the fact that an atlas with 25 deformed square patches cannot represent a surface with 26 connected components. In practice, these works also cannot faithfully represent single-object or single-connected component surfaces of arbitrary topology, although it is theoretically capable given sufficient number of charts. Furthermore, these atlas-based representations generally admits a distortion-minimal surface parameterization at the expense of representation accuracy. For instance, distortion is inevitable to deform a square patch into a circular patch. Some of these works also require a large number of charts to accurately represent general surfaces, which leads to a representation with low model efficiency.

The root cause of all limitations mentioned above lies in predefining the parametric domain, which unnecessarily constrains its boundary and topology, and hence also that of the chart. While [27] has explored “tearing” an initial open square parametric domain at regions of high distortion, the limitation on distortion remains unaddressed. Our experiments also show that its reconstructions still incur a relatively high topological error on general single-object surfaces.

Contributions. We propose a novel representation, *Minimal Neural Atlas*, where the core idea is to model the parametric domain of each chart via an implicit *probabilistic occupancy field* [24] defined on the $(-1, 1)^2$ open square of the *parametric space*. As a result, each chart is free to admit any topology and boundary, as we only restrict the bounds of the parametric domain. This enables the learning of a distortion-minimal parameterization, which is important for high quality texture mapping and efficient uniform point cloud sampling. A separation of concerns can also be established between the occupancy field and parameterization, where the former focuses on topology and the latter on geometry and distortion. This enables the proposed representation to describe

surfaces of arbitrary topology, including closed and open surfaces with arbitrary connected components, using a *minimal atlas* of 3 charts. Our experiments on ShapeNet and CLOTH3D++ support this theoretical finding and show that our reconstructions are more accurate in terms of the overall geometry.

2 Related Work

Point clouds, meshes and voxels have long been the *de facto* standard for surface representation. Nonetheless, these *discrete surface representations* describe the surface only at sampled locations with limited precision. First explored in [18,35], *neural surface representations* exploits the universal approximation capabilities of neural networks to describe surfaces continuously at a low memory cost.

Explicit Neural Surface Representations. Such representations provide a closed form expression describing exact points on the surface. [18,35] first proposed to learn an atlas for a surface by modeling the chart parameterizations with a neural network and predefining the parametric domain of each chart to the open unit square. Building on [18], [5] introduced novel training losses to regularize for chart degeneracy, distortion and the amount of overlap between charts. [13] additionally optimizes for the quality of overlaps between charts. Such atlas-based representations have also been specialized for surface reconstruction [34,3,26]. However, these representations suffer from various limitations outlined in Sec. 1, as a consequence of predefining the parametric domain. Hence, [27] proposed to adapt to the target surface topology by “tearing” an initial unit square parametric domain. In addition to the drawbacks mentioned in Sec. 1, this single-chart atlas representation also theoretically cannot describe general single-object surfaces. Moreover, the optimal tearing hyperparameters are instance-dependent, as they are determined by the scale, sampling density and area of the surface, which cannot be easily normalized.

Implicit Neural Surface Representations. These representations generally encode the surface as a level set of a scalar field defined on the 3D space, which is parameterized by a neural network. Some of the first implicit representations proposed include the *Probabilistic Occupancy Field* (POF) [24,9] and *Signed Distance Field* (SDF) [28]. These representations can theoretically describe closed surfaces of arbitrary topology and they yield accurate watertight reconstructions in practice. However, these works require ample access to watertight meshes for training, which might not always be possible. [7,17,1,2,4] proposed various approaches to learn such representations from unoriented point clouds. Nevertheless, POF and SDF are only restricted to representing closed surfaces. [10] proposed to model an *Unsigned Distance Field* (UDF) so that both open and closed surfaces can be represented as the zero level set. While this is true in theory, surface extraction is generally performed with respect to a small epsilon level set, which leads to a double or crusted surface, since there is no guarantee that the zero level set exists in practice. Consequently, UDF cannot truly represent general surfaces.

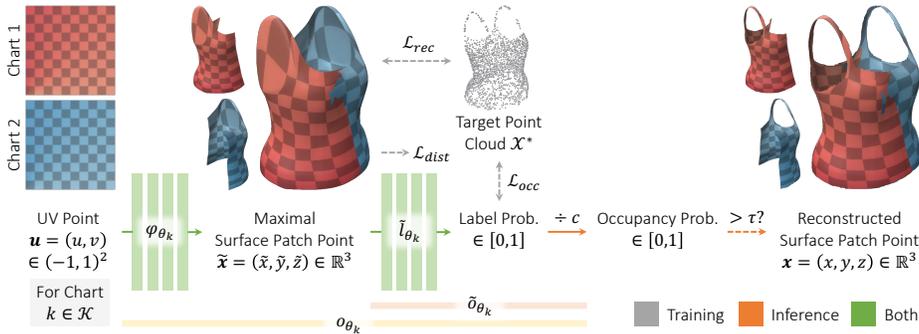


Fig. 1: Overview of our proposed method.

3 Our Method

We first present our proposed surface representation and its theoretical motivation (Sec. 3.1, 3.2). Next, we detail how to learn this representation (Sec. 3.3) and describe an approach for extracting point clouds and meshes of a specific size during inference (Sec. 3.4). An illustration of our method is given in Fig. 1.

3.1 Background

A *manifold* \mathcal{M} is a topological space that locally resembles an Euclidean space. A *surface* \mathcal{S} is merely a 2-dimensional manifold, or 2-manifold in short. In general, a manifold can be explicitly described using an *atlas*, which consists of *charts* that each describe different regions of the manifold. Formally, a chart on an n -manifold \mathcal{M} can be denoted by an ordered pair (U, φ) , whereby $U \subset \mathbb{R}^n$ is an open subset of the n -dimensional Euclidean space and $\varphi_k : U \mapsto \mathcal{M}$ is a *homeomorphism* or *parameterization* from U to an open subset of \mathcal{M} . An atlas for \mathcal{M} is given by an indexed family of charts $\{(U_k, \varphi_k) \mid k \in \mathcal{K}\}$ which forms an *open cover* of \mathcal{M} (i.e. $\bigcup_{k \in \mathcal{K}} \varphi_k(U_k) = \mathcal{M}$).

It is well-known that the *Lusternik-Schnirelmann category* [16,21,11] of a general n -manifold is at most $n + 1$. This implies that irrespective of its complexity, a general n -manifold always admits an atlas of $n + 1$ charts. Consequently, this defines the notion of a *minimal atlas* for a general n -manifold.

3.2 Surface Representation

Motivated by such a theoretical guarantee, we propose to represent general surfaces \mathcal{S} with a minimal atlas of 3 charts modeled using neural networks. Specifically, we model the surface parameterization of each chart k with a *Multi-Layer Perceptron* (MLP) parameterized by θ_k , which we denote as φ_{θ_k} . Furthermore, we employ a *probabilistic occupancy field* [24] defined on the \mathbb{R}^2 *parametric space* to implicitly model the *parametric domain* U_{θ_k} of each chart k .

More precisely, we model a probabilistic occupancy field o_{θ_k} with an MLP parameterized by θ_k on the $(-1, 1)^2$ open square of the parametric space. This allows us to implicitly represent the parametric domain U_{θ_k} as regions in the parametric space with occupancy probability larger than a specific threshold τ , or *occupied* regions in short. Our proposed *Minimal Neural Atlas* surface representation is formally given as:

$$\{(U_{\theta_k}, \varphi_{\theta_k}) \mid k \in \mathcal{K}\} , \quad (1)$$

where:

$$U_{\theta_k} = \{\mathbf{u} \in (-1, 1)^2 \mid o_{\theta_k}(\mathbf{u}) > \tau\} , \quad (2)$$

$$\varphi_{\theta_k} : U_{\theta_k} \mapsto \mathcal{S} , \quad (3)$$

$$o_{\theta_k} : (-1, 1)^2 \mapsto [0, 1] . \quad (4)$$

While we have formulated the proposed representation in the context of representing a single surface, conditioning the representation on a *latent code* $z \in \mathcal{Z}$ encoding any surface of interest facilitates the modeling of a family of surfaces. The latent code z can be inferred from various forms of inputs describing the associated surface, such as a point cloud or an image, via an appropriate encoder.

The key component that contrasts this atlas-based representation from the others is the flexibility of the parametric domain. In contrast to predefining the parametric domain, we only restrict its bounds. This eliminates redundant constraints on the boundaries and topology of the parametric domain and hence the chart. As a result, the proposed representation can learn a minimal atlas for general surfaces with arbitrary topology, including closed and open surfaces with arbitrary connected components. This also enables the learning of a distortion-minimal surface parameterization. A *separation of concerns* can thus be achieved, where o_{θ_k} mainly addresses the concern of discovering and representing the appropriate topology, and φ_{θ_k} addresses the concern of accurately representing the geometry with minimum distortion.

Decoupling Homeomorphic Ambiguity. Learning a minimal neural atlas in the present form possesses some difficulties. For a given *surface patch* described by a chart (U, φ) , there exists infinitely many other charts (U', φ') such that $U' = \phi(U)$ and $\varphi' = \varphi \circ \phi^{-1}$, where ϕ is a homeomorphism in the open square of the parametric space, that can describe the same surface patch. This statement is true because $\varphi'(U') = (\varphi \circ \phi^{-1})(\phi(U)) = \varphi(U)$. This coupled ambiguity of ϕ presents a great challenge during the learning of the two relatively independent components o_{θ_k} and φ_{θ_k} .

To decouple this homeomorphic ambiguity, we reformulate o_{θ_k} as:

$$o_{\theta_k} = \tilde{o}_{\theta_k} \circ \varphi_{\theta_k} , \quad (5)$$

where:

$$\tilde{o}_{\theta_k} : \varphi_{\theta_k}((-1, 1)^2) \mapsto [0, 1] \quad (6)$$

is an auxiliary probabilistic occupancy field defined on the *maximal surface patch* $\varphi_{\theta_k}((-1,1)^2) \subset \mathbb{R}^3$. This also requires us to extend the domain and codomain of φ_{θ_k} to the open square and \mathbb{R}^3 , respectively. Nonetheless, this is just a matter of notation since φ_{θ_k} is modeled using an MLP with a *natural domain* of \mathbb{R}^2 and codomain of \mathbb{R}^3 . Under this reformulation that conditions o_{θ_k} on φ_{θ_k} , \tilde{o}_{θ_k} can be learned such that it is invariant to ambiguities in ϕ . Particularly, since the same surface patch is described irrespective of the specific ϕ , it is sufficient for \tilde{o}_{θ_k} to be occupied only within that surface patch and *vacant* elsewhere (*i.e.* “trim away” arbitrary *surface patch excess*). This enables the learning of φ_{θ_k} with arbitrary ϕ that is independent of \tilde{o}_{θ_k} .

3.3 Training

To learn the minimal neural atlas of a target surface \mathcal{S}^* , we only assume that we are given its raw unoriented point cloud during training, which we denote as the set \mathcal{X}^* . For training, we uniformly sample a common fixed number of points or *UV samples* in the open square of each chart k to yield the set \mathcal{V}_k .

Due to the lack of minimal atlas annotations (*e.g.* target point cloud for each chart of a minimal atlas), a straightforward supervision of the surface parameterization and (auxiliary) probabilistic occupancy field for each chart is not possible. To mitigate this problem, we introduce the reconstruction loss \mathcal{L}_{rec} , occupancy loss \mathcal{L}_{occ} and metric distortion loss \mathcal{L}_{dist} . Without a loss of generality, the losses are presented similar to Sec. 3.2 in the context of fitting a single target surface. The total training loss is then given by their weighted sum:

$$\mathcal{L} = \lambda_{rec}\mathcal{L}_{rec} + \lambda_{occ}\mathcal{L}_{occ} + \lambda_{dist}\mathcal{L}_{dist} , \quad (7)$$

where λ_{rec} , λ_{occ} and λ_{dist} are the hyperparameters to balance the loss terms.

Reconstruction Loss. The concern of topology is decoupled from the parameterization in our proposed surface representation. As a result, we can ensure that the geometry of the target surface is accurately represented as long as the *maximal surface* $\tilde{\mathcal{S}}$, given by the collection of all maximal surface patches $\bigcup_{k \in \mathcal{K}} \varphi_{\theta_k}((-1,1)^2)$, forms a *cover* of the target surface \mathcal{S}^* . To this end, we regularize the surface parameterization of each chart with the unidirectional *Chamfer Distance* [15] that gives the mean squared distance of the target point cloud and its *maximal surface point cloud* $\bigcup_{k \in \mathcal{K}} \varphi_{\theta_k}(\mathcal{V}_k)$ nearest neighbor:

$$\mathcal{L}_{rec} = \frac{1}{|\mathcal{X}^*|} \sum_{\mathbf{x}^* \in \mathcal{X}^*} \min_{k \in \mathcal{K}} \min_{\mathbf{u} \in \mathcal{V}_k} \|\mathbf{x}^* - \varphi_{\theta_k}(\mathbf{u})\|_2^2 . \quad (8)$$

Occupancy Loss. To truly represent the target surface with the correct topology, it is necessary for the auxiliary probabilistic occupancy field of each chart \tilde{o}_{θ_k} to “trim away” only the *surface excess* given by $\tilde{\mathcal{S}} \setminus \mathcal{S}^*$.

Naïve Binary Classification Formulation. This is achieved by enforcing an occupancy of ‘1’ at the nearest neighbors of the target point cloud, and ‘0’

at other non-nearest neighbor maximal surface points, which effectively casts the learning of $\tilde{\theta}_k$ as a binary classification problem. However, this form of annotation incorrectly assigns an occupancy of ‘0’ at some maximal points which also form the target surface. Such mislabeling can be attributed to the difference in sampling density as well as distribution between the target and maximal surface, and effects of random sampling on the nearest neighbor operator.

Positive-Unlabeled Learning Formulation. Instead of the interpretation of mislabeling, we can take an alternative view of partial labeling. Specifically, a maximal point annotated with a label of ‘1’ is considered as a labeled positive (occupied) sample and a maximal point annotated with a label of ‘0’ is considered as an unlabeled sample instead of a labeled negative (vacant) sample. This interpretation allows us to cast the learning of $\tilde{\theta}_k$ as a *Positive and Unlabeled Learning* (PU Learning) problem [14,6] (also called *learning from positive and unlabeled examples*).

Our labeling mechanism satisfies the *single-training-set scenario* [14,6] since the maximal points are *independent and identically distributed* (i.i.d.) on the maximal surface and are either labeled positive (occupied) or unlabeled to form the “training set”. Following [14], we assume that our labeling mechanism satisfies the *Selected Completely at Random* (SCAR) assumption, which entails that the labeled maximal points are i.i.d. to, or selected completely at random from, the maximal points on the target surface. Under such an assumption, the auxiliary probabilistic occupancy field defined on can be factorized as:

$$\tilde{\theta}_k(\tilde{\mathbf{x}}) = \frac{\tilde{l}_{\theta_k}(\tilde{\mathbf{x}})}{c}, \quad (9)$$

where:

$$\tilde{l}_{\theta_k} : \varphi_{\theta_k}((-1, 1)^2) \mapsto [0, 1] \quad (10)$$

returns the probability that the given maximal point $\tilde{\mathbf{x}}$ is labeled, and c is the constant probability that a maximal point on the target surface is labeled. In the PU learning literature, \tilde{l}_{θ_k} and c are referred to as a *non-traditional classifier* and *label frequency* respectively. Note that c is proportional to the relative sampling density between the target and maximal point cloud.

\tilde{l}_{θ_k} can now be learned in the standard supervised binary classification setting with the *Binary Cross Entropy* (BCE) loss as follows:

$$\mathcal{L}_{occ} = -\frac{1}{\sum_{k \in \mathcal{K}} |\mathcal{V}_k^*|} \sum_{k \in \mathcal{K}} \sum_{\mathbf{u} \in \mathcal{V}_k} \text{BCE}(\mathbb{1}_{\mathcal{V}_k^*}(\mathbf{u}), \tilde{l}_{\theta_k} \circ \varphi_{\theta_k}(\mathbf{u})), \quad (11)$$

where $\mathcal{V}_k^* \subseteq \mathcal{V}_k$ is the set of UV samples corresponding to the target point cloud nearest neighbors. As a result of the reformulation of the probabilistic occupancy field, it can be observed that the surface parameterization of each chart directly contributes to the occupancy loss. In practice, we prevent the backpropagation of the occupancy loss gradients to the surface parameterizations. This enables the parameterizations to converge to a lower reconstruction loss since they are

now decoupled from the minimization of the occupancy loss. Furthermore, this also facilitates the separation of concerns between o_{θ_k} and φ_{θ_k} .

Metric Distortion Loss. To learn a minimal neural atlas with distortion-minimal surface parameterization, we explicitly regularize the parameterization of each chart to preserve the *metric* of the parametric domain, up to a common scale. We briefly introduce some underlying concepts before going into the details of the loss function.

Let $J_k(\mathbf{u}) = [\partial\varphi_{\theta_k}/\partial u \ \partial\varphi_{\theta_k}/\partial v]$, where $\mathbf{u} = [u \ v]^\top$, be the Jacobian of the surface parameterization of chart k . It describes the tangent space of the surface at the point $\varphi_{\theta_k}(\mathbf{u})$. The *metric tensor* or *first fundamental form* $g_k(\mathbf{u}) = J_k(\mathbf{u})^\top J_k(\mathbf{u})$ enables the computation of various differential geometric properties, such as length, area, normal, curvature and distortion.

To quantify metric distortion up to a specific common scale of L , we adopt a scaled variant of the *Symmetric Dirichlet Energy* (SDE) [30,31,29], which is an isometric distortion energy, given by:

$$\frac{1}{\sum_{k \in \mathcal{K}} |\mathcal{W}_k|} \sum_{k \in \mathcal{K}} \sum_{\mathbf{u} \in \mathcal{W}_k} \frac{1}{L^2} \text{trace}(g_k(\mathbf{u})) + L^2 \text{trace}(g_k(\mathbf{u})^{-1}), \quad (12)$$

where the distortion is quantified with respect to the set of UV samples denoted as \mathcal{W}_k . We refer this metric distortion energy as the *Scaled Symmetric Dirichlet Energy* (SSDE). As the SSDE reduces to the SDE when $L = 1$, the SSDE can be alternatively interpreted as the SDE of the derivative surface parameterization, given by post-scaling the parameterization of interest by a factor of $1/L$.

Since we are interested in enforcing metric preservation up to an arbitrary common scale, it is necessary to deduce the optimal scale L^* of the SSDE, for any given atlas (hence given g_k). To this end, we determine the L^* by finding the L that minimizes the SSDE. As the SSDE is a convex function of L , its unique global minimum can be analytically derived. Finally, the metric distortion loss used to learn a minimal neural atlas with distortion-minimal parameterization is simply given by:

$$\mathcal{L}_{dist} = 2\sqrt{\text{mean}_{\mathcal{V}^*}(\text{trace} \circ g_k) \ \text{mean}_{\mathcal{V}^*}(\text{trace} \circ g_k^{-1})}, \quad (13)$$

where:

$$\text{mean}_{\mathcal{W}}(f) = \frac{1}{\sum_{k \in \mathcal{K}} |\mathcal{W}_k|} \sum_{k \in \mathcal{K}} \sum_{\mathbf{u} \in \mathcal{W}_k} f(\mathbf{u}). \quad (14)$$

Note that we only regularize UV samples corresponding to nearest neighbors of the target point cloud, which are labeled as occupied. This provides more flexibility to the parameterization outside of the parametric domain. While [5] has proposed a novel loss to minimize metric distortion, our metric distortion loss is derived from the well-established SDE, which quantifies distortion based on relevant fundamental properties of the metric tensor, rather than its raw structure. We also observe better numerical stability as \mathcal{L}_{dist} is given by the geometric mean of two values roughly inversely proportional to each other.

3.4 Inference

Label Frequency Estimation. The label frequency c can be estimated during inference with the *positive subdomain* assumption [6]. This requires the existence of a subset of the target surface that is uniquely covered by a chart, which we assume to be true. We refer such regions as *chart interiors* since they are generally far from the chart boundary, where overlapping between charts occur.

Similar to training, we uniformly sample the open square to infer a set of maximal surface points. Given that l_{θ_k} is well-calibrated [14], maximal points on a chart interior have a \tilde{l}_{θ_k} value of c under the positive subdomain assumption. In practice, we identify such points by assuming at least η percent of maximal points lie on a chart interior, and these points correspond to the maximal points with the highest confidence in \tilde{l}_{θ_k} . We refer η as the *minimum interior rate*. c can then be estimated by the mean \tilde{l}_{θ_k} of the interior maximal points [14,6]. As \tilde{l}_{θ_k} is not explicitly calibrated and the SCAR assumption does not strictly hold in practice, we adopt the median estimator instead for improved robustness.

Point Cloud and Mesh Extraction. After the label frequency has been estimated, the *reconstructed minimal neural atlas* $\{(U_{\theta_k}, \varphi_{\theta_k}) \mid k \in \mathcal{K}\}$, and hence the *reconstructed surface* $\mathcal{S} = \bigcup_{k \in \mathcal{K}} \varphi_k(U_{\theta_k})$, are then well-defined. As a result, we can extract the *reconstructed surface point cloud* $\mathcal{X} = \bigcup_{k \in \mathcal{K}} \varphi_{\theta_k}(\mathcal{V}_k \cap U_{\theta_k})$. Furthermore, we can also extract a mesh from the reconstructed minimal neural atlas, similar to [18]. We refer this as the *reconstructed mesh*. This can be done by first defining a regular mesh in the open square of each chart and then discarding triangles with vertices outside of the reconstructed parametric domain. The mesh is then transferred to the reconstructed surface via the parameterization of each chart.

Nevertheless, it is often useful to extract a point cloud or mesh with a specific number of vertices. We achieve this in an approximate but efficient manner by adopting a two-step batch *rejection sampling* strategy. Firstly, we employ a small batch of UV samples to estimate the *occupancy rate*, which quantifies the extent to which the open square of all charts are occupied. Given such an estimate, we then deduce the number of additional UV samples required to eventually yield a point cloud or mesh with approximately the target size.

4 Experiments

We conduct two standard experiments: surface reconstruction (Sec. 4.1) and single-view reconstruction (Sec. 4.2) to verify that our representation can effectively learn a minimal atlas with distortion-minimal parameterization for surfaces of arbitrary topology. The first experiment considers the basic task of reconstructing the target surface given its point cloud, while the second is concerned with the complex task of surface reconstruction from a single image of the target. In addition to the benchmark experiments, we also perform ablation studies to investigate the significance of various components in our representation (Sec. 4.3).

Datasets. We perform all experiments on the widely used *ShapeNet* dataset [8], which is a large-scale dataset of 3D models of common objects. Specifically, we adopt the dataset preprocessed by ONet [24]. Instead of the default unit cube normalization on the point clouds, we follow existing atlas-based representations on a unit ball normalization. The ShapeNet dataset serves as a strong benchmark on representing general single-object closed surfaces.

Additionally, we also perform the surface reconstruction experiment on the CLOTH3D++ [23] dataset, which contains approximately 13,000 3D models of garments across 6 categories. Following ONet, we preprocess the dataset by uniformly sampling 100,000 points on the mesh of each garment. The point clouds are then similarly normalized to a unit ball. With this dataset, we are able to evaluate the representation power on general single-object open surfaces.

Metrics. We adopt a consistent set of metrics to assess the performance of a surface representation on all experiments. To quantify the accuracy of surface reconstruction, we employ the standard bidirectional Chamfer Distance (CD) [15] as well as the *F-score* at the default distance threshold of 1% (F@1%) [22,32], which has been shown to be a more representative metric than CD [32]. Following prior works on atlas-based representations, we report these two metrics on the reconstructed surface point cloud, given by regularly sampling the parametric domain of each chart. Furthermore, we also report the metrics on the *reconstructed mesh point cloud*, given by uniformly sampling the reconstructed mesh. This is similarly done in [19], as well as in implicit representation works. We refer to the first set of metrics as Point Cloud CD and F@1%, while the second as Mesh CD and F@1%. The reported reconstruction metrics are computed with a point cloud size of 25,000 for both the reconstruction and the target.

As pointed out by [28], topological errors in the reconstructed surface are better accounted for when evaluating on the reconstructed mesh point cloud. This is due to the non-uniform distribution of the reconstructed surface point cloud, especially at regions of high distortion where sampled points are sparse. Nevertheless, we still report the point cloud metrics to assess the reconstruction accuracy in the related task or setting of point cloud reconstruction.

We employ a set of metrics to quantify the distortion of the chart parameterizations. In particular, we use the SSDE at the optimal scale L^* (Eq. 13) to measure the metric distortion up to a common scale. We also quantify the area distortion up to a common scale using a distortion energy, which is derived in a similar manner from the equi-area distortion energy introduced in [12]. Lastly, we measure the *conformal* distortion, or distortion of local angles, using the MIPS energy [20]. The reported distortion metrics are computed with respect to the UV samples associated with the reconstructed surface point cloud. We offset the distortion metrics such that a value of zero implies no distortion.

Baselines. We benchmark minimal neural atlas against state-of-the-art explicit neural representations that can be learned given raw unoriented target point clouds for training. Specifically, we compare with AtlasNet [18], DSP [5] and

Table 1: Surface Reconstruction on CLOTH3D++.

No. of Charts	Surface Representation	Point Cloud		Mesh		Distortion		
		CD, 10^{-4} ↓	F@1% ↑	CD, 10^{-4} ↓	F@1% ↑	Metric ↓	Conformal ↓	Area ↓
1	AtlasNet	4.074	88.56	18.99	82.40	15.54	3.933	0.8428
	AtlasNet++	4.296	87.88	4.937	86.20	13.22	3.368	0.6767
	DSP	7.222	82.41	21.86	78.93	1.427	0.5746	0.1032
	TearingNet	6.872	84.99	8.321	83.40	17.40	3.407	1.149
	Ours w/o \mathcal{L}_{dist}	4.206	88.38	4.373	87.85	3.263	1.328	0.1516
	Ours	4.296	88.00	4.476	87.36	1.600	0.5688	0.1652
2	AtlasNet	3.856	89.78	7.075	86.71	6.411	1.931	0.4031
	AtlasNet++	4.106	88.62	4.734	86.93	20.65	5.095	1.116
	DSP	4.710	87.12	5.536	85.91	0.2160	0.0771	0.0283
	Ours w/o \mathcal{L}_{dist}	3.603	90.78	3.846	89.91	3.654	1.328	0.2227
	Ours	3.775	90.08	3.982	89.50	0.9227	0.3582	0.0847
	25	AtlasNet	3.396	91.47	6.269	88.74	8.028	2.485
AtlasNet++		3.368	91.62	3.652	90.67	12.89	3.615	1.073
DSP		3.227	92.06	3.501	91.26	0.4284	0.1252	0.0439
Ours w/o \mathcal{L}_{dist}		3.300	91.87	3.684	90.72	4.603	1.654	0.3047
Ours		3.299	91.90	3.554	91.07	0.5637	0.1770	0.0940

Table 2: Surface Reconstruction on ShapeNet.

No. of Charts	Surface Representation	Point Cloud		Mesh		Distortion		
		CD, 10^{-4} ↓	F@1% ↑	CD, 10^{-4} ↓	F@1% ↑	Metric ↓	Conformal ↓	Area ↓
1	AtlasNet	8.131	79.74	13.37	74.60	21.23	4.151	1.574
	AtlasNet++	8.467	78.46	10.82	75.69	30.40	5.687	2.017
	DSP	14.22	70.03	16.29	68.58	0.4684	0.1618	0.0580
	TearingNet	11.64	75.96	20.01	70.86	21.96	5.092	1.882
	Ours w/o \mathcal{L}_{dist}	6.684	83.05	7.133	81.76	8.264	2.654	0.4130
	Ours	7.559	80.45	7.959	79.39	2.546	0.8929	0.2246
3	AtlasNet	7.071	81.98	10.96	77.68	16.27	4.037	1.041
	AtlasNet++	7.516	80.59	9.280	78.17	32.39	6.252	2.626
	DSP	10.79	76.39	11.98	74.85	0.4130	0.1571	0.0380
	Ours w/o \mathcal{L}_{dist}	6.266	84.04	6.875	82.22	10.23	3.303	0.5155
	Ours	6.311	83.63	6.761	82.23	2.189	0.7094	0.2521
	25	AtlasNet	6.285	83.98	7.855	81.25	14.48	4.522
AtlasNet++		6.451	83.50	7.333	81.87	20.34	5.643	1.981
DSP		7.995	81.54	8.609	80.08	0.9477	0.2900	0.1040
Ours w/o \mathcal{L}_{dist}		5.844	85.11	6.646	83.52	7.639	2.595	0.4464
Ours		5.780	85.28	6.726	83.86	1.178	0.3760	0.1576

TearingNet [27]. Furthermore, a variant of AtlasNet that is trained with the Mesh CD and SSDE at the optimal scale, in addition to the original Point Cloud CD loss, is also adopted as an additional baseline, which we refer to as AtlasNet++. It serves as a strong baseline since topological errors in the reconstructions are explicitly regularized with the Mesh CD, unlike other baselines. The remaining losses help to minimize the excessive distortion caused by optimizing the Mesh CD, as similarly mentioned in [19].

4.1 Surface Reconstruction

In this experiment, we consider the specific setting of reconstructing the target surface given an input point cloud of size 2,500. The input point cloud also

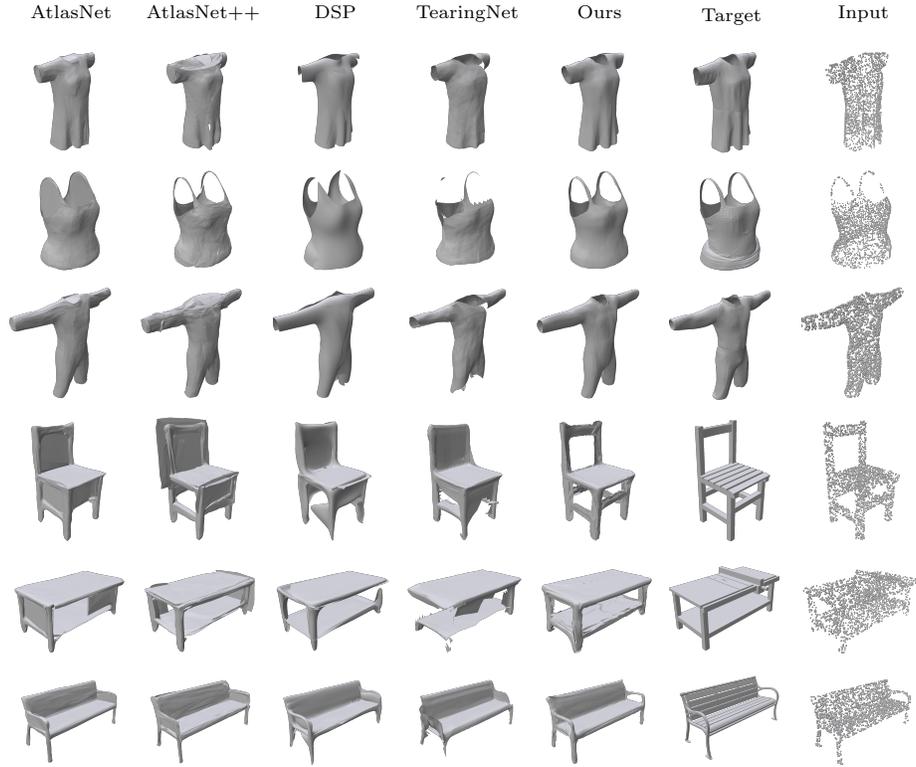


Fig. 2: Surface Reconstruction on CLOTH3D++ and ShapeNet.

serves as the target point cloud for training. This is consistent with previous works such as [18,5,13,19]. Nevertheless, we adopt a larger UV sample size of 5,000 for training in all works.

To evaluate whether a method can faithfully learn a minimal representation, we benchmark all surface representations at 3 different number of charts, except for TearingNet. In particular, we evaluate at 1, 2 and 25 charts on CLOTH3D++ and 1, 3 and 25 charts on ShapeNet. The inconsistency of 2 and 3 charts between both datasets is attributed to the fact that these baselines theoretically admit a minimal atlas of 2 and 3 charts for general single-object open and closed surfaces respectively. Benchmarking at 1 and 25 charts, which is the default for the baselines, also allows us to assess the limiting performance of a surface representation as the number of charts decreases or increases, respectively. Furthermore, we also evaluate at 1 chart because our proposed representation admits a minimal atlas of 1 chart for general single-object open surfaces.

The quantitative results for surface reconstruction on CLOTH3D++ and ShapeNet are reported in Table 1 and 2, respectively. In general, our surface representation achieves higher point cloud reconstruction performance at any given number of charts, especially on the more complex ShapeNet dataset. This

Table 3: Single-View Reconstruction on ShapeNet.

No. of Charts	Surface Representation	Point Cloud		Mesh		Distortion		
		CD, 10^{-3} ↓	F@1% ↑	CD, 10^{-3} ↓	F@1% ↑	Metric ↓	Conformal ↓	Area ↓
1	AtlasNet	3.100	57.78	4.512	52.85	34.60	4.771	3.189
	AtlasNet++	3.254	55.74	4.049	54.13	33.48	6.579	3.657
	DSP	5.210	46.28	6.048	44.55	1.113	0.3718	0.1294
	TearingNet	3.633	55.27	5.250	51.65	30.52	5.610	2.778
	Ours w/o \mathcal{L}_{dist}	3.754	61.63	3.891	60.38	8.148	2.510	0.4778
	Ours	3.840	59.90	4.002	58.71	2.674	0.9069	0.2420
3	AtlasNet	2.992	59.08	4.125	54.92	25.00	4.759	2.059
	AtlasNet++	3.077	57.84	3.706	56.26	38.07	6.263	3.419
	DSP	4.447	50.12	5.096	48.11	0.5316	0.1505	0.0901
	Ours w/o \mathcal{L}_{dist}	3.582	62.11	3.704	60.68	9.972	3.177	0.5906
	Ours	3.621	61.90	3.744	60.56	1.776	0.5537	0.2304
	25	AtlasNet	2.883	60.68	3.655	57.21	18.77	4.853
AtlasNet++		2.961	59.40	3.469	57.88	25.76	6.201	2.522
DSP		3.582	55.60	4.336	52.94	1.492	0.3712	0.2254
Ours w/o \mathcal{L}_{dist}		3.413	62.71	3.464	61.42	6.620	2.206	0.4532
Ours		3.437	63.05	3.514	61.93	1.037	0.3321	0.1497

indicates that the overall surface geometry is more accurately reconstructed by our representation, which can be attributed to the separation of concerns between o_{θ_k} and φ_{θ_k} . Furthermore, minimal neural atlas significantly outperforms the baselines in terms of the mesh reconstruction accuracy at any given number of charts, which is particularly true for lower number of charts and on ShapeNet. Together with the observation that our point cloud and mesh reconstruction metrics are relatively on par with each other, this suggests that minimal neural atlas can also reconstruct the topology of the target surface more accurately. These conclusions are also supported qualitatively in Fig. 2, where we show the reconstructed meshes of all surface representations at 2 and 3 charts, except for TearingNet, on CLOTH3D++ and ShapeNet.

The reconstruction metrics of our representation are also substantially more consistent across a wide range of charts. It is also worth noting that despite using fewer charts, minimal neural atlas often outperforms the baselines in terms of reconstruction accuracy. This affirms the ability of our representation to learn a minimal atlas for general surfaces. Moreover, the chart parameterizations of minimal neural atlas exhibit inherently lower distortion as it achieves lower metric values compared to AtlasNet, TearingNet and even AtlasNet++ without explicit regularization of distortion. Unlike DSP, the reported results also indicate that our representation is able to significantly reduce distortion without sacrificing reconstruction accuracy by additionally minimizing the metric distortion loss.

4.2 Single-View Reconstruction

This experiment adopts the exact same setting as the surface reconstruction experiment, except the input is an image of the target surface. The quantitative results on ShapeNet reported in Table 3 remains largely similar to the previous experiment. While our representation incurs a relatively higher point cloud CD,

Table 4: Ablation Study of Minimal Neural Atlas with \mathcal{L}_{dist} .

Variant	Point Cloud		Mesh		Metric Distortion ↓	Occupancy Rate ↑
	CD, 10^{-4} ↓	F@1% ↑	CD, 10^{-4} ↓	F@1% ↑		
No o_{θ_k} reformulation	15.60	69.12	16.59	67.31	2.348	80.49
No \tilde{o}_{θ_k} factorization	327.8	29.51	386.3	27.59	4.951	6.762
No \tilde{l}_{θ_k} pos. encoding	7.419	82.04	8.050	80.23	1.848	83.45
Full Model	6.311	83.63	6.761	82.23	2.189	77.48

it consistently achieves a higher reconstruction performance on the more representative F@1% metric [32]. We can thus reach to the same conclusions, as per the previous experiment.

4.3 Ablation Studies

The ablation studies are conducted in the same setting as surface reconstruction on ShapeNet using 3 charts. The results reported in Table 4 verifies the immense importance of decoupling the homeomorphic ambiguity by reformulating o_{θ_k} with Eq. 5, as well as casting the learning of \tilde{o}_{θ_k} as a PU Learning problem, which can be easily solved given the factorization of \tilde{o}_{θ_k} in Eq. 9, instead of a naïve binary classification problem. It also shows that it is crucial to apply positional embedding [25,33] on the input maximal point coordinates of \tilde{l}_{θ_k} to learn a more detailed occupancy field for better reconstructions, albeit at a minor cost of distortion.

5 Conclusion

In this paper, we propose *Minimal Neural Atlas*, a novel explicit neural surface representation that can effectively learn a minimal atlas with distortion-minimal parameterization for general surfaces of arbitrary topology, which is enabled by a fully learnable parametric domain. Despite its achievements, our representation remains prone to artifacts common in atlas-based representations, such as intersections and seams between charts. Severe violation of the SCAR assumption, due to imperfect modeling of the target surface, non-matching sampling distribution *etc.*, also leads to unintended holes on the reconstructed surface, which we leave for future work. Although we motivated this work in the context of representing surfaces, our representation naturally extends to general n -manifolds. It would thus be interesting to explore its applications in other domains.

Acknowledgements. This research/project is supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG2-RP-2021-024), and the Tier 2 grant MOE-T2EP20120-0011 from the Singapore Ministry of Education.

References

1. Atzmon, M., Lipman, Y.: SAL: Sign agnostic learning of shapes from raw data. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2020)
2. Atzmon, M., Lipman, Y.: SALD: Sign agnostic learning with derivatives. In: International Conference on Learning Representations (2021)
3. Badki, A., Gallo, O., Kautz, J., Sen, P.: Meshlet priors for 3d mesh reconstruction. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
4. Baorui, M., Zhizhong, H., Yu-shen, L., Matthias, Z.: Neural-Pull: Learning Signed Distance Functions from Point Clouds by Learning to Pull Space onto Surfaces. In: International Conference on Machine Learning (ICML) (2021)
5. Bednarik, J., Parashar, S., Gundogdu, E., Salzmann, M., Fua, P.: Shape Reconstruction by Learning Differentiable Surface Representations. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2020)
6. Bekker, J., Davis, J.: Learning from Positive and Unlabeled Data: A Survey. Machine Learning (2020)
7. Boulch, A., Langlois, P., Puy, G., Marlet, R.: Needrop: Self-supervised shape representation from sparse point clouds using needle dropping. In: 2021 International Conference on 3D Vision (3DV) (2021)
8. Chang, A.X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., Xiao, J., Yi, L., Yu, F.: ShapeNet: An Information-Rich 3D Model Repository. Tech. Rep. arXiv:1512.03012 [cs.GR], Stanford University — Princeton University — Toyota Technological Institute at Chicago (2015)
9. Chen, Z., Zhang, H.: Learning implicit fields for generative shape modeling. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
10. Chibane, J., Mir, A., Pons-Moll, G.: Neural Unsigned Distance Fields for Implicit Function Learning. In: Advances in Neural Information Processing Systems (2020)
11. Cornea, O., Lupton, G., Oprea, J., Tanré, D.: Lusternik-Schnirelmann category. American Mathematical Soc. (2003)
12. Degener, P., Meseth, J., Klein, R.: An adaptable surface parameterization method. In: IMR (2003)
13. Deng, Z., Bednarik, J., Salzmann, M., Fua, P.: Better Patch Stitching for Parametric Surface Reconstruction. In: Proceedings - 2020 International Conference on 3D Vision, 3DV 2020 (2020)
14. Elkan, C., Noto, K.: Learning classifiers from only positive and unlabeled data. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2008)
15. Fan, H., Su, H., Guibas, L.: A point set generation network for 3d object reconstruction from a single image. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017)
16. Fox, R.H.: On the lusternik-schnirelmann category. Annals of Mathematics (1941)
17. Gropp, A., Yariv, L., Haim, N., Atzmon, M., Lipman, Y.: Implicit Geometric Regularization for Learning Shapes. In: International Conference on Machine Learning (2020)

18. Groueix, T., Fisher, M., Kim, V.G., Russell, B.C., Aubry, M.: A Papier-Mâché Approach to Learning 3D Surface Generation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2018)
19. Gupta, K., Chandraker, M.: Neural Mesh Flow: 3D Manifold Mesh Generation via Diffeomorphic Flows. In: Advances in Neural Information Processing Systems (2020)
20. Hormann, K., Greiner, G.: Mips: An efficient global parametrization method. Tech. rep., Erlangen-Nuernberg Univ (Germany) Computer Graphics Group (2000)
21. James, I.: On category, in the sense of lusternik-schnirelmann. *Topology* (1978)
22. Knapitsch, A., Park, J., Zhou, Q.Y., Koltun, V.: Tanks and temples: Benchmarking large-scale scene reconstruction. *ACM Transactions on Graphics* (2017)
23. Madadi, M., Bertiche, H., Bouzouita, W., Guyon, I., Escalera, S.: Learning cloth dynamics: 3d + texture garment reconstruction benchmark. In: Proceedings of the NeurIPS 2020 Competition and Demonstration Track, PMLR (2021)
24. Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., Geiger, A.: Occupancy networks: Learning 3d reconstruction in function space. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
25. 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: European Conference on Computer Vision (ECCV) 2020 (2020)
26. Morreale, L., Aigerman, N., Kim, V., Mitra, N.J.: Neural surface maps. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021)
27. Pang, J., Li, D., Tian, D.: TearingNet: Point Cloud Autoencoder to Learn Topology-Friendly Representations. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021)
28. Park, J.J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S.: DeepSDF: Learning continuous signed distance functions for shape representation. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
29. Rabinovich, M., Poranne, R., Panozzo, D., Sorkine-Hornung, O.: Scalable locally injective mappings. *ACM Transactions on Graphics* (2017)
30. Schreiner, J., Asirvatham, A., Praun, E., Hoppe, H.: Inter-surface mapping. *ACM Transactions on Graphics* (2004)
31. Smith, J., Schaefer, S.: Bijective parameterization with free boundaries. *ACM Trans. Graph.* (2015)
32. Tatarchenko, M., Richter, S.R., Ranftl, R., Li, Z., Koltun, V., Brox, T.: What Do Single-View 3D Reconstruction Networks Learn? In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
33. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is All you Need. In: Advances in Neural Information Processing Systems (2017)
34. Williams, F., Schneider, T., Silva, C., Zorin, D., Bruna, J., Panozzo, D.: Deep Geometric Prior for Surface Reconstruction. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2019)
35. Yang, Y., Feng, C., Shen, Y., Tian, D.: Foldingnet: Point cloud auto-encoder via deep grid deformation. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2018)