Medial Spectral Coordinates for 3D Shape Analysis

Morteza Rezanejad[1], Mohammad Khodadad[2], Hamidreza Mahyar[3], Herve Lombaert[4], Michael Gruninger[1], Dirk Walther[1], Kaleem Siddiqi[5]

[1] University of Toronto, Toronto, Canada, [2]Sharif University of Technology, Tehran, Iran, [3] McMaster University, Hamilton, Canada
[4]ÉTS, Montréal, Canada, [5]McGill University, Montréal, Canada

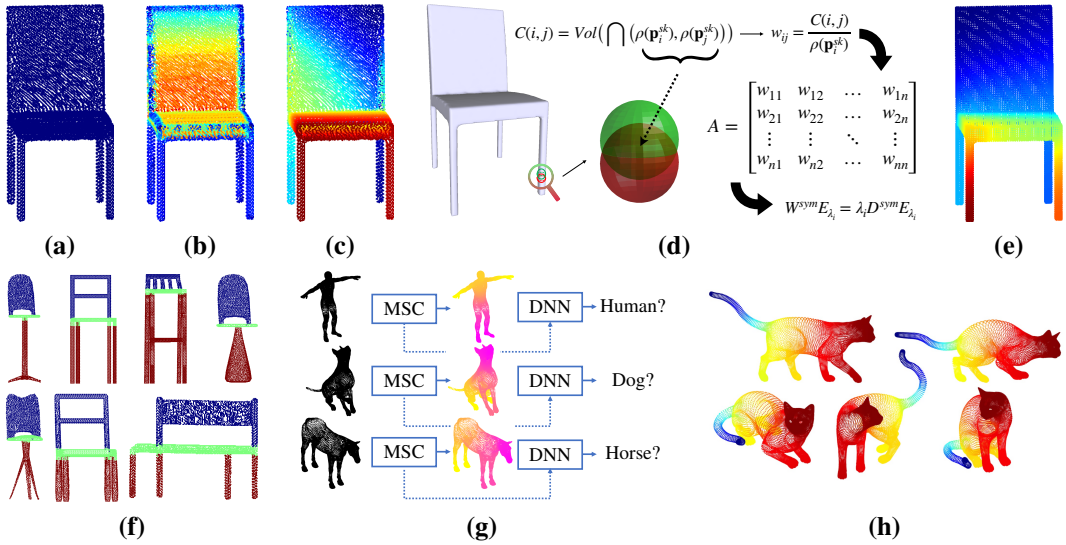


Figure 1. An overview of our main results. **First row:** **(a)** We begin with a 3D object and compute its medial surface. **(b)** The radius associated with the medial surface is mapped to its boundary. **(c)** Boundary points that are associated with the same medial surface point are coupled and shown with the same color. **(d)** Entries of the adjacency matrix are computed using the intersection volume of the associated medial spheres to drive adjacency weights, for subsequent spectral analysis (see Section 3 for a detailed explanation). **(e)** The first non-trivial eigenvector is mapped onto the original shape. **Second row:** The use of our medially driven spectral coordinates for segmentation **(f)**, object classification **(g)**, and shape correspondence **(h)**.

Abstract

In recent years there has been a resurgence of interest in our community in the shape analysis of 3D objects represented by surface meshes, their voxelized interiors, or surface point clouds. In part, this interest has been stimulated by the increased availability of RGBD cameras, and by applications of computer vision to autonomous driving, medical imaging, and robotics. In these settings, spectral coordinates have shown promise for shape representation due to their ability to incorporate both local and global shape properties in a manner that is qualitatively invariant to isometric transformations. Yet, surprisingly, such coordinates have thus far typically considered only local surface positional or derivative information. In the present article, we propose to equip spectral coordinates with medial (object width) information, so as to enrich them. The key idea is to couple surface points that share a medial ball, via the weights of the adjacency matrix. We develop a spectral feature using this idea, and the algorithms to compute it. The incorporation of object width and medial coupling has direct benefits, as illustrated by our experiments on object classification, object part segmentation, and surface point correspondence.

1. Introduction

Advances in manufacturing sensors that can perceive 3D information, including RGBD cameras, are beginning to revolutionize 3D perception using computer vision methods. In this context, 3D object shape analysis has received much attention due to its relevance to a variety of applications in robotics, medical imaging, autonomous driving, surveillance, and other domains. The technologies that are

being developed for these applications must grapple with fundamental problems including 3D shape classification, part segmentation, and shape matching.

The problem of finding correspondences between 3D objects has a long history [4, 7, 19, 20, 30, 31, 40]. Moving beyond object matching, the segmentation of a 3D object into its constituent parts is another classic computer vision problem for which research is very much still active [18, 21, 29, 32, 47, 54]. In addition to part segmentation, 3D shape classification remains a topic of active research. On this front, deep learning methods, including PointNet [37], PointNet++ [38], DeepSet [58], ShapeContextNet [53], PointGrid [26], DynamicGCN [56], and SampleNet [24], have shown promise. As such models grow and undergo further development, they also become more data-hungry and require more resources in terms of computing time. Deep neural network models are typically limited by their design and they cannot easily learn features that are not describable by the parts from which they are made. In the context of 3D shape analysis, it is also important for a neural network based model to be able to learn appropriate latent features when the point cloud is sparse or has been geometrically transformed [55].

To facilitate these tasks, a descriptive and versatile representation for 3D object shape is key. To this end, in the present article we propose a novel 3D object shape representation method that computes spectral coordinates using the associated 3D medial surface. Medial representations have been used for 2D shape analysis in the literature for decades [3, 33, 39, 44, 46, 52]. Applications of medial representations also exist in 3D [43, 45], but challenges exist around computing them reliably and robustly. In addition, the computer vision community has shifted away from the use of hand-crafted features to embrace the use of deep learning based descriptors, learned from geometric data. Taken together, these methods have allowed for better performance in general, in most tasks that the models have been curated for.

In the present article, we aim to equip present models with an enriched representation of 3D object shape, one that leverages the local width of the object to construct a spectral feature at each of its boundary points. We make three main contributions.

1. We develop a robust, and reliable implementation of an algorithm for computing 3D medial surfaces based on an average outward flux measure.

2. We introduce a novel approach to computing spectral coordinates, exploiting the duality between the medial surface and the boundary it represents. This approach brings inferred local and global geometrical information to the boundary and allows for an enriched spectral representation of that boundary. In particular, surface points that are bi-tangent to medial spheres that have a common volume are coupled to capture local object part symmetry.

3. We demonstrate that by using such spectral coordinates that are associated with the medial surface, there are performance benefits for popular 3D shape analysis tasks, including finding correspondences between surface points, 3D object part segmentation, and 3D object shape classification.

Our article is organized as follows. In Section 2, we review average outward medial surfaces and their computation. In Section 3, we introduce our novel medial spectral coordinates. In Section 4 we develop a number of applications that take advantage of our these spectral coordinates. Finally, we conclude with a summary of contributions and a discussion of possible future research directions, in Section 5.

2. Medial Representation of 3D Shapes

Introduced by Harry Blum [5], medial axes or skeletons have been used in the vision and computational geometry literature for years [2, 12, 16, 27, 43, 44, 48]. Formally, the medial axis is defined as the locus of centers of maximal inscribed spheres within a watertight object. In 3D this locus can include both 3D curves and 3D medial manifolds, that together comprise the 3D medial surface of an object. A promise of this representation is that it captures the duality between both boundary and interior object properties, the latter by considering the radius associated with the inscribed sphere to capture local object width. In addition, it is possible to reconstruct the original object from its medial surface, so in this sense the representation is complete. Unfortunately, in practice, computing a medial surface robustly turns out to be a challenging computational problem. We focus here on the medial surface computation approach of [42], which is based on a notion of average outward flux (AOF), and which was later used for hippocampal shape analysis [6] as well as other applications [43].

2.1. Average Outward Flux Medial Surfaces

To provide the intuition behind this algorithm consider the object shown in Figure 2, together with a local portion of its medial surface (light green) and the associated boundary surface patches (purple). The medial surface is homotopic to the original object and thus reflects its topology while making local reflective symmetries explicit. Let $\mathcal{S}$ represent the closed surface of a 3D object and $D_E(\mathbf{p})$ be the Euclidean distance to the boundary of this object at each point $\mathbf{p}$ in its interior. Now consider the gradient vector field $\dot{q}$ of $D_E(\mathbf{p})$. Where there is just one closest point $\mathbf{b} = (b_x, b_y, b_z)$ on the boundary, to a particular interior-point $\mathbf{a} = (a_x, a_y, a_z)$, the gradient vector at $\mathbf{a}$ is computed

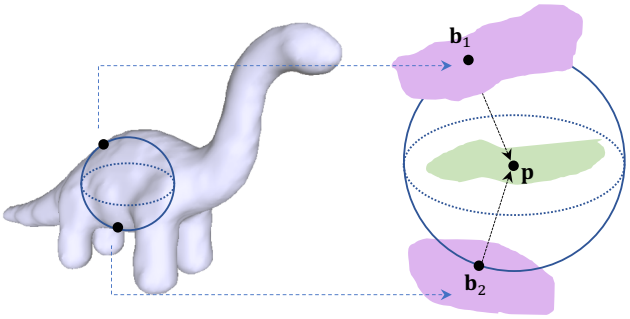


Figure 2. A schematic of a portion of a medial surface together with a selected maximal inscribed sphere, and the associated bi-tangent points ($\mathbf{b}_1$, $\mathbf{b}_2$) where the sphere touches the boundary. The medial surface patch is shown in light green and the associated local boundary patches are in light purple. Each point $\mathbf{p}$ on the medial manifold is associated with two distinct bi-tangent points on the object's surface, to which it is closest in the sense of Euclidean distance (the figure is adapted from [6]).

as:

$$\dot{q}(\mathbf{a}) = \frac{\mathbf{a} - \mathbf{b}}{||\mathbf{a} - \mathbf{b}||} \tag{1}$$

However, the gradient vector field is multi-valued at locations on the medial surface. Exploiting this property, a measure was proposed in [42], based on the outward flux of $\dot{q}$ through a shrinking sphere averaged over the surface area of that sphere:

$$AOF(\mathbf{p}) = \frac{\int_{\partial R} \langle \dot{q}, \mathcal{N}_0 \rangle \partial \mathcal{S}}{Area(\partial R)}. \tag{2}$$

Here ∂R represents the surface area of the shrinking sphere and $\partial \mathcal{S}$ is the surface area element. $\mathcal{N}_0$ is the outward normal at each point on the sphere. Earlier work on medial axis computations has shown that as the size of this sphere shrinks to zero, this measure gives non-zero values on the medial surface, but zero-values off the medial locus, and hence provides an effective means for localizing medial surface points [14, 42]. In the present article, we use a robust implementation of this algorithm. Figure 3 illustrates the sequence of steps for 3D medial surface generation.

2.2. Preserving object topology

Whereas in a continuous setting one can associate medial surface points with locations where the AOF is non-zero, in a discrete setting care has to be taken to ensure that the resultant set of voxels is homotopic to the original object. We follow past work [6], which rests on a definition of endpoints and simple points on a digital 3D lattice:

1. A *simple point* is a point that cannot be removed without changing the topology of the object. Its removal will either disconnect the object or create a hole or a cavity.

Algorithm 1 Medial Surface Generation Procedure

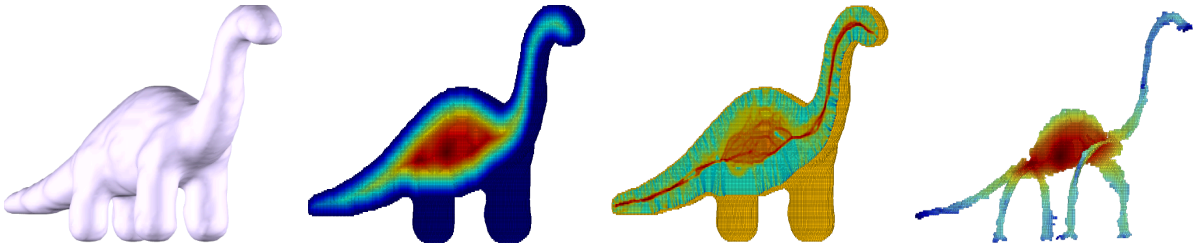
procedure AVERAGE OUTWARD FLUX COMPUTATION
  Compute the Euclidean Distance Function $D_E(\mathbf{p})$
  Compute the gradient vector field ∇D_E
  Compute the average outward flux of ∇D_E using 2
procedure AOF TOPOLOGY PRESERVING THINNING
  Heap $\rightarrow H$, Threshold $\rightarrow \tau$
  for each point $\mathbf{p}$ on the boundary of the object **do**
    if $\mathbf{p}$ is simple **then**
      insert point $\mathbf{p}$ into H with key value $AOF(\mathbf{p})$
  while H is not empty **do**
    $\mathbf{p}$ = HeapExtractMax(H)
    if $\mathbf{p}$ is simple **then**
      if $\mathbf{p}$ is end point and $AOF(\mathbf{p}) > \tau$ **then**
        Label $\mathbf{p}$ as a medial surface point
      else
        Remove $\mathbf{p}$
    for all neighbors $\mathbf{q}$ of $\mathbf{p}$ **do**
      if if $\mathbf{q}$ is simple **then** inset point $\mathbf{q}$ into H

2. Consider a plane that passes through a point $\mathbf{p}$ such that its intersection with the 3D object results in an open curve. If this curve ends at $\mathbf{p}$, then $\mathbf{p}$ is an *endpoint* of that 3D curve. Additional examples of endpoints include when $\mathbf{p}$ is on the rim or corner of a 3D surface.

On a 3D digital lattice we can classify both simple points and endpoints by considering the 26 neighborhoods of a particular voxel (the cubic lattice of $3 \times 3 \times 3$). We implement these two notions of simple points and endpoints according to the characterization of [36], which is based upon different sets of neighbors that share either a point, an edge, or a face with the considered voxel. Algorithm 1 describes the resultant medial surface generation method, which combines the AOF measure with a topology preserving thinning approach, using the above characterization of simple points and endpoints. Figure 3 provides several examples for a variety of 3D objects. At the end of this process the surface of each 3D object $\mathcal{S}$ can be represented by a set of skeletal points where:

$$\mathcal{S} = \bigcup_{i=1}^{n} \mathbf{p}_i^{sk} = \bigcup_{i=1}^{n} (x_i, y_i, z_i, r_i, \lambda_i). \tag{3}$$

Here the triplet (x_i, y_i, z_i) represents the location of the surface point $\mathbf{p}_i$, r_i is the radius of the maximal inscribed sphere touching the object at $\mathbf{p}_i$, and λ_i is the average outward flux value at the medial surface point at the center of that sphere. As shown in [43] the AOF value also reveals the object angle, and thus provides additional useful object and symmetry information. In the next subsection, we will



(a) 3D Shape **(b)** $D_E(\mathbf{p})$ **(c)** $\frac{\int_{\partial R}\langle \dot{q},\mathcal{N}_0\rangle \partial \mathcal{S}}{Area(\partial R)}$ **(d)** MAT

Figure 3. Using the divergence theorem, medial axis voxels can be identified by considering the behavior of the average outward flux (AOF) of the gradient of the Euclidean distance function to the boundary of a 3D object, through a shrinking sphere [43]. In particular, the limiting AOF value of all points not located on the skeleton is equal to zero. Starting with the boundary of a 3D object (**a**), we first compute the distance map to the boundary (**b**), and then compute the 3D AOF map (**c**) from this distance. Finally, by retaining the non-zero AOF voxels, we obtain the 3D medial axis transform (**d**). Since at each medial voxel the maximal inscribed sphere (Fig. 2) touches the boundary at least at two distinct points, it is possible to reconstruct the boundary purely from the medial locus.

demonstrate that this representation can be used for object reconstruction.

2.3. Boundary Reconstruction Using Medial Points

Perhaps the most obvious criterion for the suitability of an object shape representation method is whether it can be used to reconstruct the original object. This criterion, alongside other criteria such as completeness, hierarchy, invariance, stability, and similarity has been used to judge shape representation choices. Whereas medial representations satisfy several of these criteria [43], algorithms implemented on a discrete lattice may not allow for faithful object reconstruction. Here, for purposes of reconstruction, we compare the AOF skeletonization process against a modern centerline extraction method [6] which has an open source implementation in Itk (https://bit.ly/3rZPY57). Since the medial surface consists of the locus of maximally inscribed spheres, the original object can be reconstructed by considering the envelope of the maximally inscribed spheres at all the medial surface points. A quantitative comparison via a mean intersection over union (mIoU) measure with original objects is shown in Table 1, over the entire datasets of ModelNet10 [51] and COSEG [50]. For all object categories in each dataset the medial surface based reconstructions are more complete than those obtained using medial curves.

COSEG			ModelNet10		
Model	Ours	Curve	Model	Ours	Curve
Candelabra	96.99	73.25	Table	99.68	88.84
Chairs	96.38	81.64	Bathtub	93.25	86.57
Fourleg	90.33	73.24	Bed	95.31	73.39
Goblets	90.68	82.24	Chair	96.10	86.08
Guitars	93.19	76.80	Desk	97.78	81.80
Lamps	95.30	68.19	Dresser	94.23	91.81
Vases	96.54	70.99	Monitor	90.90	73.60
Irons	94.07	80.71	Night stand	92.66	83.10
Tele-aliens	98.19	76.62	Sofa	91.53	81.52
Chairs (L)	97.18	80.26	Toilet	91.05	83.29
Mean	94.89	76.39	Mean	94.25	83.00

Table 1. The mean Intersection-Over-Union (mIoU) reconstruction measures, for the two datasets of COSEG and ModelNet10, comparing the use of flux driven centrelines (https://bit.ly/3rZPY57) and AOF medial surfaces [42].

3. Spectral Coordinates using Medial Manifolds

We now propose to use the medial manifold to equip the object's surface with a new spectral signature, one that explicitly considers local object width. Whereas spectral approaches have been popular for shape correspondence [9, 10, 20, 31] and spectral methods have also been successfully used in combination with graph convolution to learn features from point sets and meshes [15, 49, 57], the explicit consideration of object width and the associated symmetry properties and coupling of surface points which share a medial ball is new.

Let the object's surface $\mathcal{S}$ be represented by its set of boundary points $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$. We voxelize the object and compute the medial surface using Algorithm 1 in Section 2. Let the set of resulting medial surface voxels be $\{\mathbf{p}_1^{sk}, \mathbf{p}_2^{sk}, \ldots, \mathbf{p}_m^{sk}\}$. For each medial surface point $\mathbf{p}_j^{sk}$ there is a direct relationship between the object angle θ (2) and the AOF value at it (as discussed in [43]). Using this relationship, for each $\mathbf{p}_j^{sk}$, we generate a set of candidate boundary points by rotation of unit tangent vectors by θ, followed by scaling by the medial radius, in planes that contain the medial surface normal. Of these candidates, we retain those that are within a distance of 1 voxel to the object's surface. We refer to the set of such candidate boundary

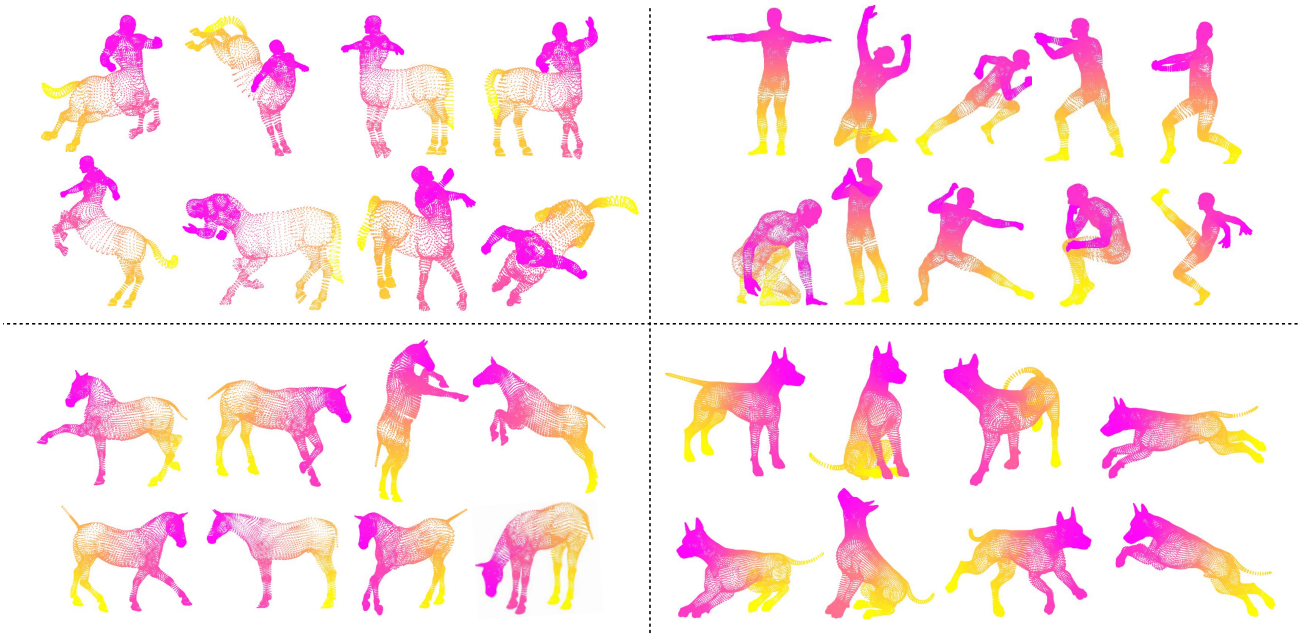


Figure 4. Qualitative results showing correspondences obtained using our medial spectral coordinates across different poses of the same models in the TOSCA [8] dataset.

points as $\mathcal{S}_{\text{Rec}}$. Then, for each vertex on the original object's surface, we apply a k-nearest neighbor search with $k = 1$, to find the closest corresponding boundary point in $\mathcal{S}_{\text{Rec}}$. In this fashion, each boundary vertex point $\mathbf{b}_i$ is associated with a unique medial surface point $\mathbf{p}_j^{sk}$ and is given a unique medial radius value. Now, assuming that the boundary is tessellated so that the boundary points are the vertices (or set of nodes) in a connected graph, eigenmaps of these nodes are computed as follows. We first compute an adjacency matrix (as shown in Figure 3 (**d**)) where a particular metric is used to compute the adjacency weights in the matrix W, as follows. Consider two boundary points (nodes) $\mathbf{b}_i$ and $\mathbf{b}_j$. Let these boundary points be associated with skeletal points $\mathbf{p}_i$ and $\mathbf{p}_j$, respectively. Now, let the maximal inscribed sphere at skeletal point $\mathbf{p}_i$ be given by $\rho(\mathbf{p}_i^{sk})$. We consider the degree of overlap between the inscribed spheres touching the object's surface at $\mathbf{b}_i, \mathbf{b}_j$ to construct the adjacency weight between the associated nodes

$$w_{ij} = \frac{C(i,j)}{\rho(\mathbf{p}_i^{sk})}, \tag{4}$$

where $C(i,j)$ is the volume of the intersection between the spheres $\rho(\mathbf{p}_i^{sk})$ and $\rho(\mathbf{p}_j^{sk})$. We then construct the general Laplacian operator as $L = D - W$, where W is the weighted adjacency matrix of the graph with affinity weights (see [17]), and the degree matrix, D, is a diagonal matrix, where $D_{ii} = \Sigma_j w_{ij}$. We then compute the eigenvectors of the Laplacian matrix:

$$LE_{\lambda_i} = \lambda_i E_{\lambda_i}, i = 1, \ldots, n. \tag{5}$$

Here, since matrix W is not symmetric (leading to a possibility of imaginary numbers numerically), we create two symmetric matrices W^{sym} and D^{sym} as follows. W^{sym} has entries given by

$$w_{ij}^{sym} = C(i,j), \tag{6}$$

and D^{sym} which is a diagonal matrix, with entries $d_{ii}^{sym} = \rho(\mathbf{p}_i^{sk})$. Now, by using the Arnoldi method of ARPACK [28], we can obtain eigenvectors of the following equation:

$$W^{sym}E_{\lambda_i} = \lambda_i D^{sym}E_{\lambda_i}, i = 1, \ldots, n \tag{7}$$

where the obtained eigenvalues are real and non-negative. Now that we have our eigenvectors, we can obtain spectral coordinates by mapping the coordinates from the boundary of the 3D shape. For each boundary point $\mathbf{b}_i$ element (e.g. x, y, z or its corresponding sphere radius value r from the medial surface), we can compute the mapped information on the i^{th} basis eigenvectors as:

$$\mathbf{m}_i = \mathbf{b}_i D^{sym} E_{\lambda_i}. \tag{8}$$

Referring back to the Dirichlet energy equation introduced by Bronstein *et al.* [10] that measures change over the boundary $\mathcal{S}$, we can introduce our spectral coordinates as follows:

$$S_C(\mathcal{S}) = \Big(\bigcup_{i=1}^{n} \lambda_i \sum_{j=1}^{k} (\mathbf{m}_{ij}^2) \Big). \tag{9}$$

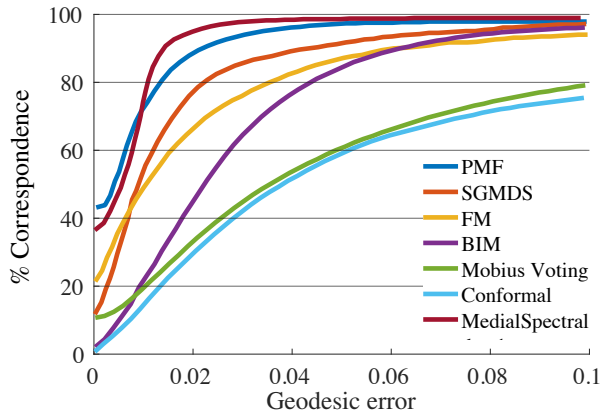


Figure 5. Correspondence accuracy on the TOSCA dataset [8] for different methods. The methods we have compared include PMF [23], SGMDS [1], FM [35], BIM [22], Mobius Voting [30], Conformal [22], and the use of our new MedialSpectral signature.

We are careful to always remove the trivial eigenvalue of 0 (the smallest eigenvalue) from our set of spectral coordinates.

4. Applications

We now evaluate the use of our medial spectral coordinates for three popular applications in 3D shape analysis: a) finding correspondence between surface points, b) 3D object part segmentation of 3D objects and c) 3D object classification. For each case, we demonstrate that the explicit use of local object width, as reflected in the medial radius, and the coupling of surface points that share a common medial sphere, provides direct benefits and improved performance.

4.1. Shape Correspondence

To obtain correspondences between surface points on two distinct objects, one needs a matching algorithm that matches one against the other. In the present setting, we have spectral coordinates alongside eigenvalues that are obtained from a Laplacian operator. When spectra are computed, two situations are possible that make the direct comparison of spectral coordinates challenging. First, eigenvector computation may generate a sign ambiguity. Second, it is possible that when eigenvectors are being computed for the same data but for two different shapes, they might be computed in opposite orders due to the fact that the ordering of the lowest eigenvector may change. [31] suggests a method for mitigating the effects of this flipping problem by favoring three factors: 1) pairs of eigenvectors that are most likely to match based on the similarity between their eigenvalues 2) histograms 3) the spatial distributions of their spectral coordinate value. We use these techniques to reorder and align spectra during the correspondence finding process. The process of reordering is sped up by downsampling all eigenvectors. After reordering and aligning the spectra, two points that are closest in the embedded representations can be treated as corresponding points across the two objects. This is achieved by using the Coherent Point Drift (CPD) method [34]. We test our correspondence finding algorithm on the TOSCA dataset [8], which includes a total of 80 objects, including four limb animals such as cats and dogs, female and male figures. We compare our methods against several other existing methods and report the results in Figure 5. The results show the advantages of using medially driven spectral coordinates when compared to the best conformal model that we could find. We also achieve results that are comparable to those produced by the state of the art deep neural network based models, such as PMF [23]. We show qualitative results of correspondences obtained on four different models from the TOSCA dataset in Figure 4.

4.2. Part Segmentation

We now show that the addition of medially driven spectral coordinates can be beneficial to the task of part segmentation. To achieve this aim, we propose the following task to carry out unsupervised 3D shape decomposition. We collect all the spectral coordinates computed and add them as additional features to the boundary representation $\mathbf{b}_i = (x_i, y_i, z_i, r_i)$. We then use the high-dimensional data spectral clustering of [11] to cluster all object models within a category together. The algorithm of [11] provides a spectral clustering approach based on subspace randomization and graph fusion for high-dimensional data, which enables us to carry out both segmentation and co-segmentation tasks for a single shape or family of shape models (co-segmentation). To examine whether these added features to the original shape help with the segmentation task, we tested our off-the-shelf method on the Princeton Segmentation Benchmark [13]. We report our results in Table 2 and show typical qualitative part segmentation results in Figure 6.

Our results show that in several of these object categories, there is a boost in performance when we use spectral coordinates that use medial surface information over the use of raw data (with the spectral clustering algorithm). These results are reflected in the second (SpectralMedial) and third (Spectral) columns of Table 2. We also observe that the use of spectral coordinates outperforms several of the other existing approaches including the deep model-based approach of [25]. Finally, the mean performance and the rand index measure of our approach is the best overall on the entire Princeton dataset [13].

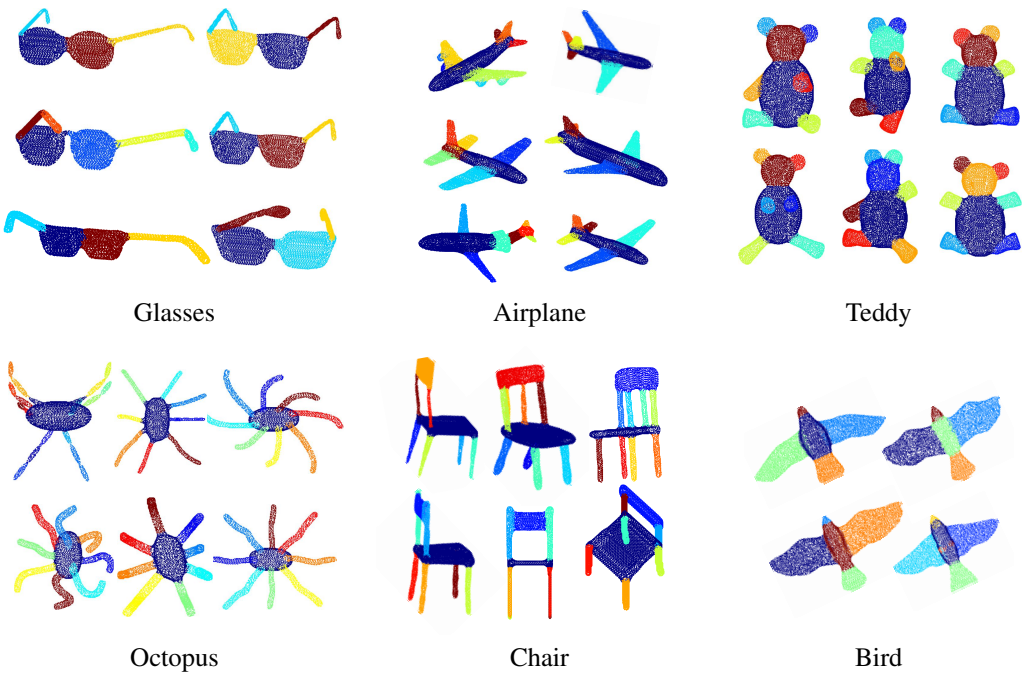


Figure 6. Qualitative part segmentation results obtained using medially driven spectral signatures on the Princeton segmentation benchmark (see text for details).

Model	SpectralMedial	Spectral	MV-RNN [25]	Shu [41]	WcSeg	RandCuts	ShapeDiam	NormCuts	CoreExtra	RandWalks	FitPrim	KMeans
Human	**0.0557**	0.1240	0.106	0.116	0.128	0.131	0.179	0.152	0.225	0.219	0.153	0.163
Cup	0.1413	0.1262	0.100	**0.096**	0.171	0.219	0.358	0.244	0.307	0.358	0.413	0.459
Glasses	0.0802	0.1304	**0.066**	0.173	0.173	0.101	0.204	0.141	0.301	0.311	0.235	0.188
Airplane	**0.0728**	0.0891	0.085	0.150	0.089	0.122	0.092	0.186	0.256	0.248	0.166	0.211
Ant	0.0279	0.0900	0.021	**0.001**	0.021	0.025	0.022	0.047	0.065	0.068	0.086	0.131
Chair	0.0483	0.0518	0.051	**0.040**	0.103	0.184	0.111	0.088	0.187	0.156	0.212	0.213
Octopus	0.0255	0.0835	**0.022**	0.036	0.029	0.063	0.045	0.061	0.051	0.067	0.101	0.101
Table	0.0630	0.1158	0.072	**0.040**	0.091	0.383	0.184	0.093	0.244	0.131	0.181	0.369
Teddy	0.0521	0.1105	0.035	**0.024**	0.056	0.045	0.057	0.121	0.114	0.128	0.132	0.182
Hand	**0.0670**	0.1443	0.076	0.135	0.116	0.090	0.202	0.155	0.155	0.189	0.202	0.154
Plier	0.0970	0.1433	**0.054**	0.151	0.087	0.109	0.375	0.183	0.093	0.230	0.169	0.263
Fish	**0.1030**	0.2339	0.146	0.288	0.203	0.297	0.248	0.394	0.273	0.388	0.424	0.413
Bird	**0.0049**	0.0850	0.059	0.171	0.101	0.107	0.115	0.184	0.124	0.250	0.196	0.190
Armadillo	**0.0178**	0.1265	0.060	0.073	0.081	0.092	0.090	0.116	0.141	0.115	0.091	0.117
Bust	**0.1081**	0.1729	0.162	0.275	0.266	0.232	0.298	0.316	0.315	0.298	0.300	0.334
Mech	0.1018	0.1133	0.121	**0.073**	0.182	0.277	0.238	0.159	0.387	0.211	0.306	0.425
Bearing	0.0660	0.1516	0.080	**0.056**	0.122	0.124	0.119	0.183	0.398	0.246	0.188	0.280
Vase	**0.0772**	0.1504	0.106	0.212	0.161	0.133	0.239	0.236	0.226	0.246	0.257	0.387
FourLeg	**0.0921**	0.1772	0.135	0.140	0.152	0.174	0.161	0.208	0.191	0.218	0.185	0.193
Average	**0.0696**	0.1273	0.082	0.118	0.123	0.153	0.176	0.172	0.211	0.215	0.210	0.251

Table 2. The Rand Index (error) segmentation scores, for object categories in the Princeton segmentation benchmark, comparing across different methods. In this table, a lower number indicates better performance. The use of medially drive spectral coordinates (column two) consistently outperforms the use of spectral coordinates without the medial (radius) component (column three). The approach is also competitive against many other approaches for several of the object categories.

MNet10		bathtub	bed	chair	desk	dresser	monitor	night stand	sofa	table	toilet
	Vanila	83.87	92.85	94.73	77.77	83.87	96.48	84.02	91.70	80.78	95.87
	Ours	85.71	94.28	92.85	82.80	85.54	96.51	84.33	96.44	83.40	97.02
MNet40		airplane	bathtub	bed	bench	bookshelf	bottle	bowl	car	chair	cone
	Vanila	100.00	79.86	93.88	74.93	92.81	93.98	100.00	97.77	95.83	100.00
	Ours	100.00	83.50	98.19	76.38	96.57	94.67	100.00	98.58	98.88	100.00
		cup	curtain	desk	door	dresser	flower pot	glass box	guitar	keyboard	lamp
	Vanila	69.85	89.83	78.94	94.89	64.90	29.89	93.93	100.00	100.00	89.85
	Ours	73.95	92.46	81.39	99.29	66.83	32.20	98.71	100.00	100.00	90.93
		laptop	mantel	monitor	night stand	person	piano	plant	radio	range hood	sink
	Vanila	100.00	95.90	94.86	82.41	84.93	88.63	72.85	69.81	90.99	79.93
	Ours	100.00	98.87	98.30	84.37	88.03	92.77	73.09	70.41	95.77	83.19
		sofa	stairs	stool	table	tent	toilet	tv stand	vase	wardrobe	xbox
	Vanila	95.87	84.94	89.95	87.86	94.88	98.88	86.87	78.79	59.93	69.91
	Ours	97.13	86.97	90.60	89.31	96.26	100.00	87.74	80.50	60.59	74.32

Table 3. The difference in object classification accuracy when using vanilla features and our medial spectral features with the PointNet model [37], using the ModelNet10 and ModelNet40 datasets [51].

4.3. Object Classification

The last set of experiments we present in this paper relate to the problem of object classification. For this task, we examine whether the inclusion of medial spectral coordinates leads to an improvement in the performance of a vanilla neural network model. To achieve this goal, we carried out the following classification task. Inspired by the idea presented in [55], we created a new feature vector for each point in a 3D point cloud as a Geometry Similarity Connection, capturing the local behavior of each point in the eigenvector space. First, we find the k nearest neighbors to each point in the eigenvector space. Then, we add the mean and standard deviation of those k points to the x,y,z coordinates of the considered point. Finally, we add these additional features to the point's features, creating a vector of 9 features for each point in the point cloud. We tested the added features on the popular PointNet model [37]. The added features result in an improvement in the overall accuracy of the model from 88.76% to 90.41% on ModelNet-10 and from 86.23% to 88.26% on ModelNet-40. We present a detailed account of the results for each category for this experiment in Table 3.

5. Conclusion

We have proposed a novel spectral coordinate for 3D shape analysis applications, one that includes the local object width associated with the medial surface radius function. To our knowledge, this is the first attempt to include medial width in a spectral feature, and our experiments demonstrate the considerable benefits of doing this. Our model ties surface points on a 3D object to their associated medial surface points and uses the duality between the medial surface of an object and its boundary to extract spectral coordinates with explicit consideration of local object width and object part symmetry. We have introduced a novel way to compute the adjacency weights between boundary nodes, based on the volume of intersection of their associated inscribed spheres. The use of this medially driven spectral coordinate leads to improved correspondences when compared against the use of conformal methods. When applied to 3D part segmentation the use of the medial surface to group together points on the boundary that are not necessarily close to one another in geodesic terms, but are tied to each other geometrically in terms of local object symmetry, has advantages. Finally, our object classification experiments using the medial spectral coordinates in addition to the raw data demonstrate that a vanilla deep neural network model can benefit from the incorporation of these additional spectral features.

We intend to release all our implementations on GitHub, to allow our results to be reproduced. In future work, we aim to examine whether it is possible to compute medial spectral coordinates implicitly, via a deep neural network model. We also hope to examine the potential of medial spectral coordinates for graph neural network models. Our method suffers from a potential limitation, which is that the complexity of the skeletonization algorithm is a linear function of the number of voxels. Hence, voxelization could impose a computational burden when compared to some alternate methods [45]. However, in practice, refined voxelization at increased computatiional cost does provide greater robustness to boundary perturbations.

Acknowledgments: We are grateful to the University of Toronto and NSERC for research support.

References

[1] Yonathan Aflalo, Anastasia Dubrovina, and Ron Kimmel. Spectral generalized multi-dimensional scaling. *International Journal of Computer Vision*, 118(3):380–392, 2016. 6

[2] Nina Amenta, Sunghee Choi, and Ravi Krishna Kolluri. The power crust, unions of balls, and the medial axis transform. *Computational Geometry*, 19(2-3):127–153, 2001. 2

[3] Xiang Bai, Xingwei Yang, Deguang Yu, and Longin Jan Latecki. Skeleton-based shape classification using path similarity. *International Journal of Pattern Recognition and Artificial Intelligence*, 22(04):733–746, 2008. 2

[4] Silvia Biasotti, Simone Marini, Michela Spagnuolo, and Bianca Falcidieno. Sub-part correspondence by structural descriptors of 3d shapes. *Computer-Aided Design*, 38(9):1002–1019, 2006. 2

[5] Harry Blum et al. *A transformation for extracting new descriptors of shape*, volume 43. MIT press Cambridge, MA, 1967. 2

[6] Sylvain Bouix, Jens C Pruessner, D Louis Collins, and Kaleem Siddiqi. Hippocampal shape analysis using medial surfaces. *Neuroimage*, 25(4):1077–1089, 2005. 2, 3, 4

[7] Rune Brincker, Anders Skafte, Manuel López-Aenlle, Aldo Sestieri, Walter D'Ambrogio, and Alfonso Canteli. A local correspondence principle for mode shapes in structural dynamics. *Mechanical Systems and Signal Processing*, 45(1):91–104, 2014. 2

[8] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. *Numerical geometry of non-rigid shapes*. Springer Science & Business Media, 2008. 5, 6

[9] Michael M Bronstein and Alexander M Bronstein. Shape recognition with spectral distances. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):1065–1071, 2010. 4

[10] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017. 4, 5

[11] Xiaosha Cai, Dong Huang, Chang-Dong Wang, and Chee-Keong Kwoh. Spectral clustering by subspace randomization and graph fusion for high-dimensional data. *Advances in Knowledge Discovery and Data Mining*, 12084:330, 2020. 6

[12] Charles-Olivier Dufresne Camaro, Morteza Rezanejad, Stavros Tsogkas, Kaleem Siddiqi, and Sven Dickinson. Appearance shock grammar for fast medial axis extraction from real images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14382–14391, 2020. 2

[13] Xiaobai Chen, Aleksey Golovinskiy, and Thomas Funkhouser. A benchmark for 3D mesh segmentation. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 28(3), Aug. 2009. 6

[14] Pavel Dimitrov, James N Damon, and Kaleem Siddiqi. Flux invariants for shape. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, volume 1, pages I–I. IEEE, 2003. 3

[15] Nicolas Donati, Abhishek Sharma, and Maks Ovsjanikov. Deep geometric functional maps: Robust feature learning for shape correspondence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8592–8601, 2020. 4

[16] Jacob Feldman and Manish Singh. Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences*, 103(47):18014–18019, 2006. 2

[17] Leo J Grady and Jonathan R Polimeni. *Discrete calculus: Applied analysis on graphs for computational science*, volume 3. Springer, 2010. 5

[18] Sindhu Hegde and Shankar Gangisetty. Pig-net: Inception based deep learning architecture for 3d point cloud segmentation. *Computers & Graphics*, 95:13–22, 2021. 2

[19] Masaki Hilaga, Yoshihisa Shinagawa, Taku Kohmura, and Tosiyasu L Kunii. Topology matching for fully automatic similarity estimation of 3d shapes. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 203–212, 2001. 2

[20] Varun Jain and Hao Zhang. Robust 3d shape correspondence in the spectral domain. In *IEEE International Conference on Shape Modeling and Applications 2006 (SMI'06)*, pages 19–19. IEEE, 2006. 2, 4

[21] Evangelos Kalogerakis, Melinos Averkiou, Subhransu Maji, and Siddhartha Chaudhuri. 3d shape segmentation with projective convolutional networks. In *proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3779–3788, 2017. 2

[22] Vladimir G Kim, Yaron Lipman, and Thomas Funkhouser. Blended intrinsic maps. *ACM transactions on graphics (TOG)*, 30(4):1–12, 2011. 6

[23] Zorah Lähner, Matthias Vestner, Amit Boyarski, Or Litany, Ron Slossberg, Tal Remez, Emanuele Rodola, Alex Bronstein, Michael Bronstein, Ron Kimmel, et al. Efficient deformable shape correspondence via kernel matching. *arXiv preprint arXiv:1707.08991*, 2017. 6

[24] Itai Lang, Asaf Manor, and Shai Avidan. Samplenet: Differentiable point cloud sampling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7578–7588, 2020. 2

[25] Truc Le, Giang Bui, and Ye Duan. A multi-view recurrent neural network for 3d mesh segmentation. *Computers Graphics*, 66:103–112, 2017. Shape Modeling International 2017. 6, 7

[26] Truc Le and Ye Duan. Pointgrid: A deep network for 3d shape understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9204–9214, 2018. 2

[27] Der-Tsai Lee. Medial axis transformation of a planar shape. *IEEE Transactions on pattern analysis and machine intelligence*, (4):363–369, 1982. 2

[28] Richard B Lehoucq, Danny C Sorensen, and Chao Yang. *ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*. SIAM, 1998. 5

[29] Cheng Lin, Lingjie Liu, Changjian Li, Leif Kobbelt, Bin Wang, Shiqing Xin, and Wenping Wang. Seg-mat: 3d shape

segmentation using medial axis transform. *IEEE Transactions on Visualization and Computer Graphics*, page 1–1, 2020. 2

[30] Yaron Lipman and Thomas Funkhouser. Möbius voting for surface correspondence. *ACM Transactions on Graphics (TOG)*, 28(3):1–12, 2009. 2, 6

[31] Herve Lombaert, Leo Grady, Jonathan R Polimeni, and Farida Cheriet. Focusr: feature oriented correspondence using spectral regularization–a method for precise surface matching. *IEEE transactions on pattern analysis and machine intelligence*, 35(9):2143–2160, 2012. 2, 4, 6

[32] Yecheng Lyu, Xinming Huang, and Ziming Zhang. Learning to segment 3d point clouds in 2d image space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12255–12264, 2020. 2

[33] Diego Macrini, Sven Dickinson, David Fleet, and Kaleem Siddiqi. Bone graphs: Medial shape parsing and abstraction. *Computer Vision and Image Understanding*, 115(7):1044–1061, 2011. 2

[34] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*, 32(12):2262–2275, 2010. 6

[35] Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (TOG)*, 31(4):1–11, 2012. 6

[36] Chris Pudney. Distance-ordered homotopic thinning: a skeletonization algorithm for 3d digital images. *Computer vision and image understanding*, 72(3):404–413, 1998. 3

[37] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017. 2, 8

[38] Charles R Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv preprint arXiv:1706.02413*, 2017. 2

[39] Morteza Rezanejad. *Medial measures for recognition, mapping and categorization*. McGill University, 2020. 2

[40] Dinggang Shen, Wai-him Wong, and Horace HS Ip. Affine-invariant image retrieval by correspondence matching of shapes. *Image and Vision Computing*, 17(7):489–499, 1999. 2

[41] Zhenyu Shu, Chengwu Qi, Shiqing Xin, Chao Hu, Li Wang, Yu Zhang, and Ligang Liu. Unsupervised 3d shape segmentation and co-segmentation via deep learning. *Computer Aided Geometric Design*, 43:39–52, 2016. Geometric Modeling and Processing 2016. 7

[42] Kaleem Siddiqi, Sylvain Bouix, Allen Tannenbaum, and Steven W Zucker. Hamilton-jacobi skeletons. *International Journal of Computer Vision*, 48(3):215–231, 2002. 2, 3

[43] Kaleem Siddiqi and Stephen Pizer. *Medial representations: mathematics, algorithms and applications*, volume 37. Springer Science & Business Media, 2008. 2, 3, 4

[44] Kaleem Siddiqi, Ali Shokoufandeh, Sven J Dickinson, and Steven W Zucker. Shock graphs and shape matching. *International Journal of Computer Vision*, 35(1):13–32, 1999. 2

[45] Andrea Tagliasacchi, Thomas Delame, Michela Spagnuolo, Nina Amenta, and Alexandru Telea. 3d skeletons: A state-of-the-art report. *Computer Graphics Forum*, 35(2):573–597, 2016. 2

[46] Roger Tam and Wolfgang Heidrich. Shape simplification based on the medial axis transform. In *IEEE Visualization, 2003. VIS 2003.*, pages 481–488. IEEE, 2003. 2

[47] Hugues Thomas, Charles R Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, François Goulette, and Leonidas J Guibas. Kpconv: Flexible and deformable convolution for point clouds. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6411–6420, 2019. 2

[48] Stavros Tsogkas and Sven Dickinson. Amat: Medial axis transform for natural images. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2708–2717, 2017. 2

[49] Chu Wang, Babak Samari, and Kaleem Siddiqi. Local spectral graph convolution for point set feature learning. In *Proceedings of the European conference on computer vision (ECCV)*, pages 52–66, 2018. 4

[50] Yunhai Wang, Shmulik Asafi, Oliver Van Kaick, Hao Zhang, Daniel Cohen-Or, and Baoquan Chen. Active co-analysis of a set of shapes. *ACM Transactions on Graphics (TOG)*, 31(6):1–10, 2012. 4

[51] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1912–1920, 2015. 4, 8

[52] Jun Xie, Pheng-Ann Heng, and Mubarak Shah. Shape matching and modeling using skeletal context. *Pattern Recognition*, 41(5):1756–1767, 2008. 2

[53] Saining Xie, Sainan Liu, Zeyu Chen, and Zhuowen Tu. Attentional shapecontextnet for point cloud recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4606–4615, 2018. 2

[54] Mutian Xu, Junhao Zhang, Zhipeng Zhou, Mingye Xu, Xiaojuan Qi, and Yu Qiao. Learning geometry-disentangled representation for complementary understanding of 3d object point cloud. *arXiv preprint arXiv:2012.10921*, 2020. 2

[55] Mingye Xu, Zhipeng Zhou, and Yu Qiao. Geometry sharing network for 3d point cloud classification and segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 12500–12507, 2020. 2, 8

[56] Fanfan Ye, Shiliang Pu, Qiaoyong Zhong, Chao Li, Di Xie, and Huiming Tang. Dynamic gcn: Context-enriched topology learning for skeleton-based action recognition. In *Proceedings of the 28th ACM International Conference on Multimedia*, pages 55–63, 2020. 2

[57] Li Yi, Hao Su, Xingwen Guo, and Leonidas J Guibas. Syncspeccnn: Synchronized spectral cnn for 3d shape segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2282–2290, 2017. 4

[58] Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan Salakhutdinov, and Alexander Smola. Deep sets. *arXiv preprint arXiv:1703.06114*, 2017. 2