



HAL
open science

Instant recovery of shape from spectrum via latent space connections

Riccardo Marin, Arianna Rampini, Umberto Castellani, Emanuele Rodolà,
Maks Ovsjanikov, Simone Melzi

► To cite this version:

Riccardo Marin, Arianna Rampini, Umberto Castellani, Emanuele Rodolà, Maks Ovsjanikov, et al.. Instant recovery of shape from spectrum via latent space connections. 2020 International Conference on 3D Vision (3DV 2020), Nov 2020, Fukuoka, Japan. pp.120-129, 10.1109/3DV50981.2020.00022 . hal-04479709

HAL Id: hal-04479709

<https://hal.science/hal-04479709>

Submitted on 27 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Instant recovery of shape from spectrum via latent space connections

Riccardo Marin
University of Verona

riccardo.marin.01@univr.it

Emanuele Rodolà
Sapienza University of Rome
rodola@di.uniroma1.it

Arianna Rampini
Sapienza University of Rome

rampini@di.uniroma1.it

Maks Ovsjanikov
Ecole Polytechnique, IP Paris
maks@lix.polytechnique.fr

Umberto Castellani
University of Verona

umberto.castellani@univr.it

Simone Melzi
Ecole Polytechnique, IP Paris
Sapienza University of Rome
melzi@di.uniroma1.it

Abstract

We introduce the first learning-based method for recovering shapes from Laplacian spectra. Our model consists of a cycle-consistent module that maps between learned latent vectors of an auto-encoder and sequences of eigenvalues. This module provides an efficient and effective linkage between Laplacian spectrum and geometry. Our data-driven approach replaces the need for ad-hoc regularizers required by prior methods, while providing more accurate results at a fraction of the computational cost. Our learning model applies without modifications across different dimensions (2D and 3D shapes alike), representations (meshes, contours and point clouds), as well as across different shape classes, and admits arbitrary resolution of the input spectrum without affecting complexity. The increased flexibility allows us to address notoriously difficult tasks in 3D vision and geometry processing within a unified framework, including shape generation from spectrum, mesh super-resolution, shape exploration, style transfer, spectrum estimation from point clouds, segmentation transfer and point-to-point matching.

1. Introduction

Constructing compact encodings of geometric shapes lies at the heart of 2D and 3D Computer Vision. While earlier approaches have concentrated on handcrafted representations, with the advent of geometric deep learning [12, 31], data-driven *learned* feature encodings have gained prominence. A desirable property in many applications, such as shape exploration and synthesis, is to be able to recover the shape from its (latent) encoding, and various auto-encoder architectures have been designed to solve this problem [2, 30, 32, 20]. Despite significant progress in this area, the structure of the latent vectors is arduous to control. For

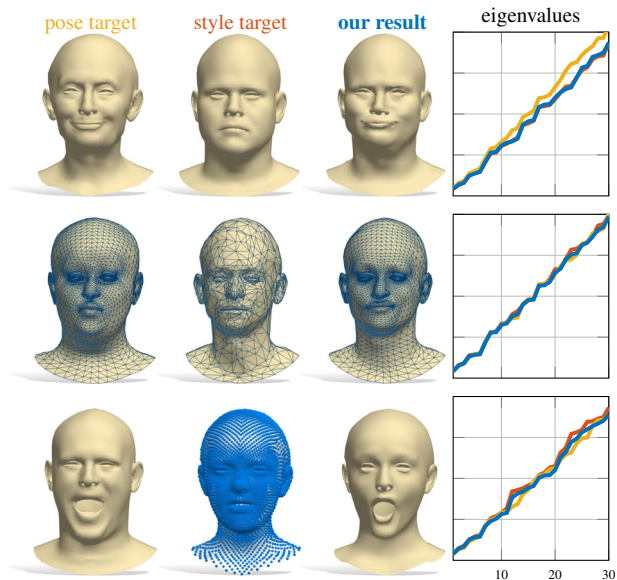


Figure 1. Our spectral reconstruction enables correspondence-free style transfer. Given pose and style “donors” (left and middle columns respectively), we synthesize a new shape with the pose of the former and the style of the latter. The generation is driven by a learning-based eigenvalues alignment (rightmost plots). Our approach handles different resolutions (middle row) and representations (bottom row; the surface underlying the point cloud is for visualization purposes only).

example, the dimensions of the latent vectors typically lack a canonical ordering, while invariance to various geometric deformations is often only learned by data augmentation or complex constraints on the intermediate features.

At the same time, a classical approach in spectral geometry is to encode a shape using the sequence of eigenvalues (spectrum) of its Laplacian operator. This representation is useful since: (1) it does not require any training, (2) it can be computed on various data representations, such as point

clouds or meshes, regardless of sampling density, (3) it enjoys well-known theoretical properties such as a natural ordering of its elements and invariance to isometries, and (4) as shown recently [19, 38], alignment of eigenvalues often promotes near-isometries, which is useful in multiple tasks such as non-rigid shape retrieval and matching problems.

Unfortunately, although encoding shapes via their Laplacian spectra can be straightforward (at least for meshes), the inverse problem of recovering the shape is very difficult. Indeed, it is well-known that certain pairs of non-isometric shapes can have the same spectrum, or in other words “one cannot hear the shape of a drum” [21]. At the same time, recent evidence suggests that such cases are pathological and that *in practice* it might be possible to recover a shape from its spectrum [19]. Nevertheless, existing approaches [19], while able to deform a shape into another with a given spectrum, can produce highly unrealistic shapes with strong artifacts failing in a large number of cases.

In this paper, we combine the strengths of data-driven auto-encoders with those of spectral methods. Our key idea is to construct a single architecture capable of synthesizing a shape from a learned latent code and from its Laplacian eigenvalues. We show that by explicitly training networks that aim to translate between the learned latent codes and the spectral encoding, we can both recover a shape from its eigenvalues and moreover endow the latent space with certain desirable properties. Remarkably, our shape-from-spectrum solution is extremely efficient since it requires a single pass through a trained network, unlike expensive iterative optimization methods with ad-hoc regularizers [19]. Furthermore, our trainable module acts as a proxy to differentiable eigendecomposition, while encouraging geometric consistency within the network. Overall, our key **contributions** can be summarized as follows:

- We propose the first learning-based model to robustly recover shape from Laplacian spectra *in a single pass*;
- For the first time, we provide a bidirectional linkage between learned 3D latent space and spectral geometric properties of 3D shapes;
- Our model is *general*, in that it applies with no modifications to different classes even across different geometric representations and dimensions and to data that does not belong to the datasets used at training time;
- We showcase our approach in multiple applications (e.g., Fig. 1), and show significant improvement over the state of the art; see Fig. 2 for an example.

2. Related work

Spectral quantities and in particular the eigenvalues of the Laplace-Beltrami operator provide an informative summary of the intrinsic geometry. For example, closed-form

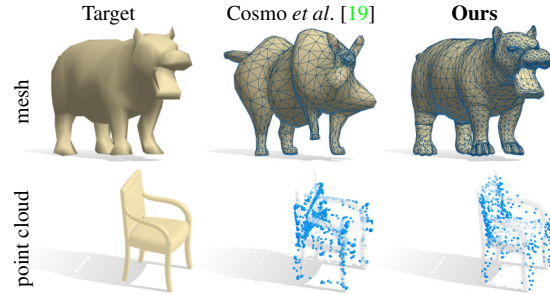


Figure 2. Comparison in estimating a shape from its Laplacian spectrum between the state-of-the-art method [19] (middle) and ours (right) for a mesh and a point cloud. The shapes recovered by our method are significantly closer to the target.

estimates and analytical bounds for surface area, genus and curvature in terms of the Laplacian eigenvalues have been obtained [13]. Given these properties, spectral shape analysis has been exploited in many computer vision and computer graphics tasks such as shape retrieval [41], description and matching [46, 3, 11, 33], mesh segmentation [40], sampling [34] and compression [26] among many others. Typically, the intrinsic properties of the shape are computed from its explicit representation and are used to encode compact geometric features invariant to isometric deformations.

Recently, several works have started to address the inverse problem: namely, recovering an extrinsic embedding from the intrinsic encoding [9, 19]. This is closely related to the fundamental theoretical question of “hearing the shape of the drum” [25, 21]. Although counterexamples have been proposed to show that in certain scenarios multiple shapes might have the same spectrum, there is recent work that proposes effective practical solutions to this problem. In [9] the shape-from-operator method was proposed, aiming at obtaining the extrinsic shape from a Laplacian matrix where the 3D reconstruction was recovered after the estimation of the Riemannian metric in terms of edge lengths. In [18] the intrinsic and extrinsic relations of geometric objects have been extensively defined and evaluated from both theoretical and practical aspects. The authors revised the framework of functional shape differences [43] to account of extrinsic structure extending the reconstruction task to non-isometric shapes and models obtained from physical simulation and animation. Several works have also been proposed to recover shapes purely from Laplacian *eigenvalues* [15, 1, 35] or with mild additional information such as excitation amplitude in the case of musical key design [8]. Most closely related to ours in this area is the recent *isoppectralization* approach introduced in [19], that aims directly to estimate the 3D shape from the spectrum. This approach works well in the vicinity of a good solution but is both computationally expensive and, as we show below, can quickly produce unrealistic instances, failing in a large number of cases in 3D, as shown in Fig. 2 for two examples.

In this paper we contribute to this line of work, and propose to replace the heuristics used in previous methods such as [19] with a purely data-driven approach for the first time. Our key idea is to design a deep neural network, that both constrains the space of solutions based on the set of shapes given at training, and at the same time, allows us to solve the isospectralization problem with a *single forward pass*, thus avoiding expensive and error-prone optimization.

We note that a related idea has been recently proposed in [24] via the so-called OperatorNet architecture. However, that work is based on shape difference operators [43] and as such requires a fixed source shape and functional maps to each shape in the dataset to properly synthesize a shape. Our approach is based on Laplacian eigenvalues alone and thus is completely correspondence-free.

Our approach also builds upon the recent work on learning generative shape models. A range of techniques have been proposed using the volumetric representations [47], point cloud auto-encoders [4, 2], generative models based on meshes and implicit functions [45, 22, 30, 28, 14], and part structures [29, 32, 20, 48], among many others.

Although generative models, and in particular auto-encoders, have shown impressive performance, the structure of the latent space is typically difficult to control or analyze directly. To address this problem, some methods proposed a disentanglement of the latent space [48, 4] to split it in more semantic regions. Perhaps most closely related to ours in this domain, is the work in [4], where the shape spectrum is used to promote disentanglement of the latent space into intrinsic and extrinsic components, that can be controlled separately. Nevertheless, the resulting network does not allow to synthesize shapes from their spectra.

Extending the studies of these approaches, our work provides the first way to connect the learned latent space to the spectral one, thus inheriting the benefits and providing the versatility of moving across the two representations. This allows our network to synthesize shapes from their spectra, and also to relate shapes with very different input structure (e.g., meshes and point clouds) across a vastness of sampling densities, enabling several novel applications.

3. Background

We model shapes as connected 2-dimensional Riemannian manifolds \mathcal{X} embedded in \mathbb{R}^3 , possibly with boundary $\partial\mathcal{X}$, equipped with the standard metric. On each shape \mathcal{X} we consider its positive semi-definite Laplace-Beltrami operator $\Delta_{\mathcal{X}}$, generalizing the classical notion of Laplacian from the Euclidean setting to curved surfaces.

Laplacian spectrum. $\Delta_{\mathcal{X}}$ admits an eigendecomposition

$$\Delta_{\mathcal{X}}\phi_i(x) = \lambda_i\phi_i(x) \quad x \in \text{int}(\mathcal{X}) \quad (1)$$

$$\langle \nabla\phi_i(x), \hat{n}(x) \rangle = 0 \quad x \in \partial\mathcal{X} \quad (2)$$

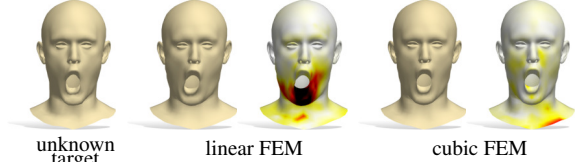


Figure 3. Reconstruction examples of our shape-from-spectrum pipeline. We show the results obtained with two different inputs: the eigenvalues of the Laplacian discretized with linear FEM, and those of the cubic FEM discretization. The heatmap encodes point-wise reconstruction error, growing from white to dark red.

into eigenvalues $\{\lambda_i\}$ and associated eigenfunctions $\{\phi_i\}$ ¹.

The Laplacian eigenvalues of \mathcal{X} (its *spectrum*) form a discrete set, which is canonically ordered into a non-decreasing sequence

$$\text{Spec}(\mathcal{X}) := \{0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots\}. \quad (3)$$

In the special case where \mathcal{X} is an interval in \mathbb{R} , the eigenvalues λ_i correspond to the (squares of) oscillation frequencies of Fourier basis functions ϕ_i . This provides us with a connection to classical Fourier analysis, and with a natural notion of hierarchy induced by the ordering of the eigenvalues. In the light of this analogy, in practice, one is usually interested in a limited bandwidth consisting of the first $k > 1$ eigenvalues; typical values in geometry processing applications range from $k = 30$ to 100.

Furthermore, the spectrum is *isometry-invariant*, i.e., it does not change with deformations of the shape that preserve geodesic distances (e.g., changes in pose).

Discretization. In the discrete setting, we represent shapes as triangle meshes $X = (V, T)$ with n vertices V and m triangular faces T ; depending on the application, we will also consider unorganized point clouds. Vertex coordinates in both cases are represented by a matrix $\mathbf{X} \in \mathbb{R}^{n \times 3}$.

The Laplace-Beltrami operator $\Delta_{\mathcal{X}}$ is discretized as a $n \times n$ matrix via the finite element method (FEM) [16]. In the simplest setting (i.e., linear finite elements), this discretization corresponds to the cotangent Laplacian [36]; however, in this paper we use *cubic* FEM (see e.g. [40, Sec. 4.1] for a clear treatment), since it yields a more accurate discretization as shown in Fig. 3. Differently from [19, 38], this comes at virtually no additional cost for our pipeline, as we show in the sequel. On point clouds, $\Delta_{\mathcal{X}}$ can be discretized using the approach described in [17, 10].

4. Method

Our main contribution is a deep learning model for recovering shapes from Laplacian eigenvalues. Our model

¹Similarly to [19] we use homogeneous Neumann boundary conditions; see Eq. (2), where $\hat{n}(x)$ denotes the outward normal to the boundary.

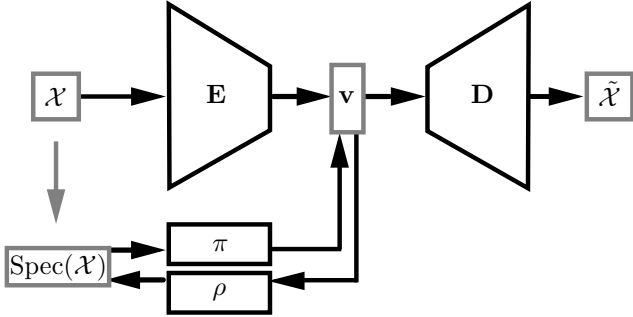


Figure 4. Our network model. The input shape \mathcal{X} and its Laplacian spectrum $\text{Spec}(\mathcal{X})$ are passed, respectively, through an AE enforcing $\mathcal{X} \approx \tilde{\mathcal{X}}$, and an invertible module (π, ρ) mapping the eigenvalue sequence to a latent vector \mathbf{v} . The two branches are trained simultaneously, forcing \mathbf{v} to be updated accordingly. The trained model allows to recover the shape purely from its eigenvalues via the composition $D(\pi(\text{Spec}(\mathcal{X}))) \approx \mathcal{X}$.

operates in an end-to-end fashion: given a spectrum as input, it directly yields a shape with a single forward pass, thus avoiding expensive test-time optimization.

Motivation. Our rationale lies in the observation that shape semantics can be learned from the data, rather than by relying upon the definition of ad-hoc regularizers [19], often resulting in unrealistic reconstructions. For example, a sheet of paper can be *isometrically* crumpled or folded into a plane (see inset figure). Since both embeddings have the same eigenvalues, the desirable reconstruction must be imposed as a prior. By taking a data-driven approach, we make our method aware of the “space of realistic shapes”, yielding both a dramatic improvement in accuracy and efficiency, and enabling new interactive applications.



Latent space connections. Our key idea is to construct an auto-encoder (AE) neural network architecture, augmented by explicitly modeling the connections between the latent space of the AE and the Laplacian spectrum of the input shape; see Fig. 4 for an illustration of our learning model.

Loosely speaking, our approach can be seen as implementing a coupling between two latent spaces: a learned one that operates on the shape embedding \mathcal{X} , and the one provided by the eigenvalues $\text{Spec}(\mathcal{X})$. In the former case, the *encoder* E is trainable, whereas the mapping $\mathcal{X} \rightarrow \text{Spec}(\mathcal{X})$ is provided via the eigen-decomposition and fixed a priori. Finally, we introduce the two coupling mappings π, ρ , trained with a bidirectional loss, to both enable communication across the latent spaces and to tune the learned space by endowing it with structure contained in $\text{Spec}(\mathcal{X})$.

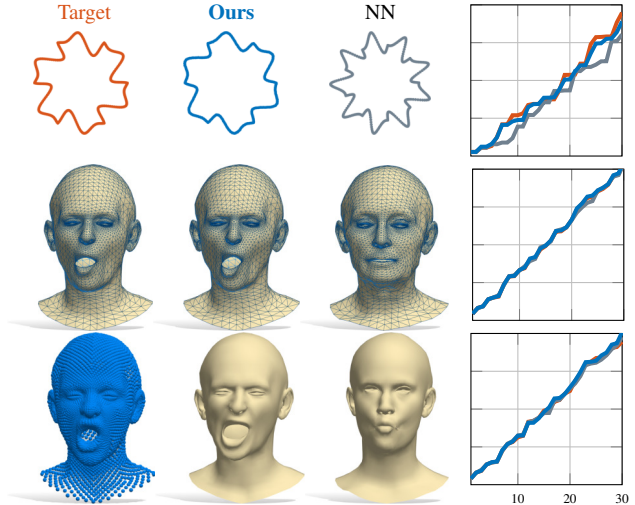


Figure 5. Shape reconstruction from eigenvalues using our approach on different representations (i.e. 2D contours, 3D meshes and point clouds). The eigenvalues of the shapes on the left are given to our network, which outputs the shapes in the middle. For each representation, the eigenvalues are computed on the appropriate Laplacian discretization as per Sec. 3. The NN column shows the nearest-neighbor solution sought in the training set.

We phrase our overall training loss as follows:

$$\ell = \ell_{\mathcal{X}} + \alpha \ell_{\lambda}, \quad \text{with} \quad (4)$$

$$\ell_{\mathcal{X}} = \frac{1}{n} \|D(E(\mathbf{X})) - \mathbf{X}\|_F^2 \quad (5)$$

$$\ell_{\lambda} = \frac{1}{k} (\|\pi(\boldsymbol{\lambda}) - E(\mathbf{X})\|_2^2 + \|\rho(E(\mathbf{X})) - \boldsymbol{\lambda}\|_2^2) \quad (6)$$

where $\boldsymbol{\lambda}$ is a vector containing the first k eigenvalues in $\text{Spec}(\mathcal{X})$, \mathbf{X} is the matrix of point coordinates, E is the encoder, D is the decoder (Fig. 4), $\|\cdot\|_F$ denotes the Frobenius norm, and $\alpha = 10^{-4}$ controls the relative strengths of the reconstruction loss $\ell_{\mathcal{X}}$ and the spectral term ℓ_{λ} . The blocks D , E , π , and ρ are learnable and parametrized by a neural network (see the supplementary material for implementation details). Eq. (6) enforces $\rho \approx \pi^{-1}$; in other words, π and ρ form a translation block between the latent vector and the spectral encoding of the shape.

At test time, we recover a shape from the spectrum Spec simply via the composition $D(\pi(\text{Spec}))$ (Section 5). For additional applications we refer to Section 6.

Shape representation. We consider two different settings: triangle meshes in point-to-point correspondence *at training time* (typical in graphics and geometry processing), and unorganized point clouds *without* a consistent vertex labeling (typical in 3D computer vision).

Autoencoder architecture. Our model can be built with potentially any autoencoder. In our applications we chose relatively simple ones to deal with meshes and unorganized

	full res	1000	500	200
Ours	1.61	1.62	1.71	2.13
Ours without ρ	1.89	1.82	2.06	2.42
NN	4.45	4.63	4.01	2.65
Cosmo <i>et al.</i> [19]	—	16.4	7.11	4.08

Table 1. Shape-from-spectrum reconstruction comparisons with NN (nearest neighbors between spectra) and the state of the art [19]; we report average error over 100 shapes of an unseen subject from COMA [39]. Best results are obtained with our full pipeline. ‘—’ denotes out of memory; all errors must be rescaled by 10^{-5} .

point clouds, although more powerful generative methods would be equally possible. The latent space dimension is fixed to 30 (the same as k). We refer to the supplementary material for details about the architecture, both in the case of meshes and point clouds.

Remark. Our architecture takes $\text{Spec}(\mathcal{X})$ as an input, *i.e.*, the eigenvalues are *not* computed at training time. By learning an *invertible* mapping to the latent space, we avoid expensive backpropagation steps through the spectral decomposition of the Laplacian $\Delta_{\mathcal{X}}$. In this sense, the mapping ρ acts as an efficient proxy to differentiable eigendecomposition, which we exploit in several applications below.

Since eigenvalue computation is only incurred as an offline cost, it can be performed with arbitrary accuracy (we use cubic FEM, see Fig. 3) without sacrificing efficiency.

5. Results

In this section we report the results on our core application of shape from spectrum recovery.

To evaluate our method, we trained our model on 1,853 3D shapes from the COMA dataset [39] of human faces; 100 shapes of an unseen subject are used for the test set. We repeated this test at four different mesh resolutions: $\sim 4\text{K}$ (full resolution), 1K, 500 and 200 vertices respectively. For each resolution, we independently compute the Laplacian spectrum and use these spectra to recover the shape.

Comparison. We compared our method in terms of reconstruction accuracy to the state-of-the-art isospectralization method of Cosmo *et al.* [19], as well as to a nearest-neighbors baseline, consisting in picking the shape of the training set with the closest spectrum to the target one. In addition, we trained two separate architectures (with and without the ρ block) and compared them. The test without this network component is an ablation study we carry out to validate the importance of the *invertible* module connecting the spectral encoding to the learned latent codes.

The quantitative results are reported in Table 1 as the mean squared error between the reconstructed shape and the ground-truth. Figures 2 and 5 further show qualitative comparisons with the different baselines involving differ-

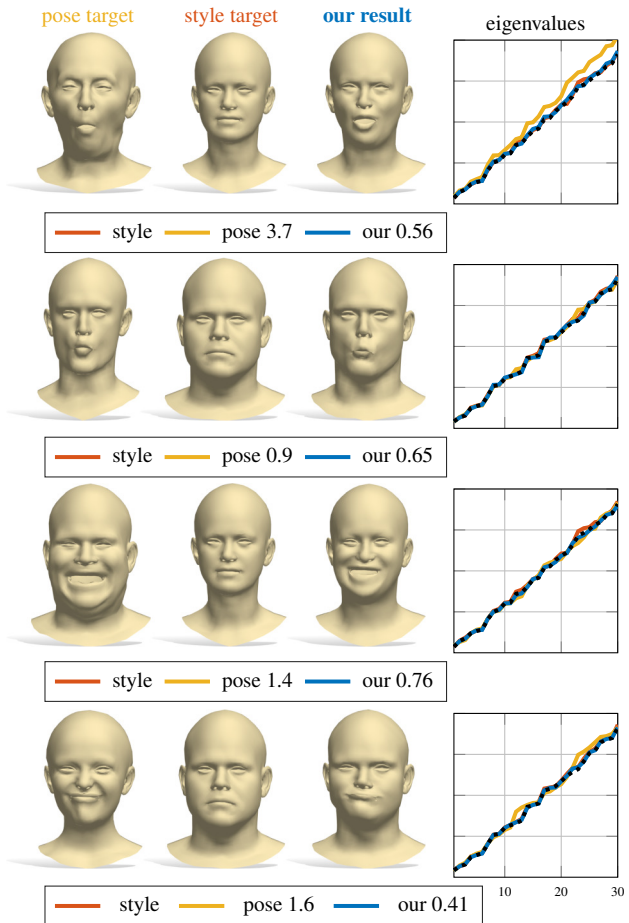


Figure 6. Examples of style transfer. The target style (middle) is applied to the target pose (left) by solving problem (7) and then decoding the resulting latent vector (right). For each example we also report the corresponding eigenvalue alignment (rightmost plots). The black dotted line is the image of ρ . The numbers in the legend denote the distance from the target “style” spectrum to the source pose and to our generated shape; a small number suggests near-isometry between the generated shape and the style target.

ent shape representations. In Fig. 5, for the sake of illustration, similarly to [19, 38], we also include 2D contours, discretized as regular cycle graphs.

As the results suggest, the ρ block both contributes to reduce the reconstruction error, and to enable novel applications (see in Sec. 6). Note that our method achieves a significant improvement over nearest neighbors in terms of accuracy, and an order of magnitude improvement over isospectralization. Also, the latter approach consists in an expensive optimization which requires hours to run, while our method is instantaneous at test time.

Spectral bandwidth has a direct effect on reconstruction accuracy, since increasing this number brings more high-frequency detail into the representation. Following [19, 38, 42], in all our experiments we use $k = 30$. In the supplementary material we report results for different k .

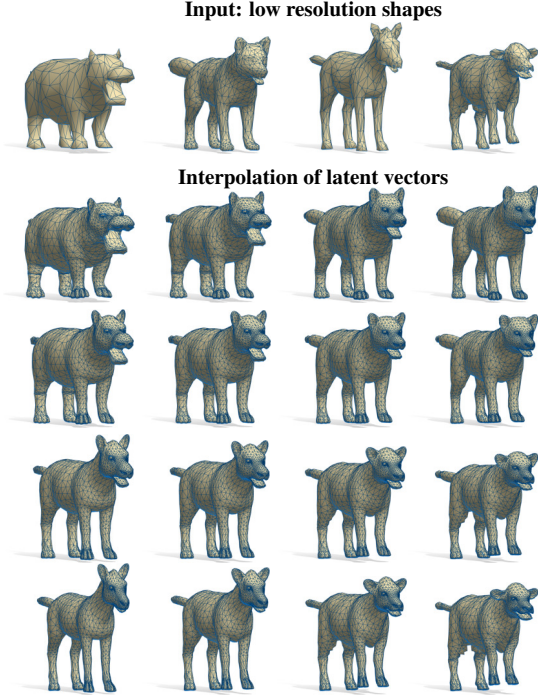


Figure 7. Latent space interpolation of four low-resolution shapes with different connectivity (top row, unseen at training). The spectra of the input shapes are mapped via π to the latent space, where they are bilinearly interpolated and then decoded to \mathbb{R}^3 . The reconstructions of the input are depicted at the corners of the grid.

6. Additional applications

Our general model enables several additional applications, by exploiting the connection between spectral properties and shape generation. Due to the limited space, we collect in the supplementary materials the details of the training and test sets and the parameters used in our experiments.

6.1. Style transfer

As shown in Fig. 1, we can use our trained network to transfer the style of a shape $\mathcal{X}_{\text{style}}$ to another shape $\mathcal{X}_{\text{pose}}$ having both a different style and pose. This is done by a search in the latent space, phrased as:

$$\min_{\mathbf{v}} \|\text{Spec}(\mathcal{X}_{\text{style}}) - \rho(\mathbf{v})\|_2^2 + w \|\mathbf{v} - E(\mathcal{X}_{\text{pose}})\|_2^2 \quad (7)$$

Here, the first term seeks a latent vector whose associated spectrum aligns with the eigenvalues of $\mathcal{X}_{\text{style}}$; in other words, we regard style as an intrinsic property of the shape, and exploit the fact that the Laplacian spectrum is invariant to pose deformations. The second term keeps the latent vector close to that of the input pose (we initialize with $\mathbf{v}_{\text{init}} = E(\mathcal{X}_{\text{pose}})$). We solve the optimization problem by back-propagating the gradient of the cost function of Eq. (7) with respect to \mathbf{v} through ρ . The sought shape is then given by a forward pass on the resulting minimizer. In Fig. 6, we

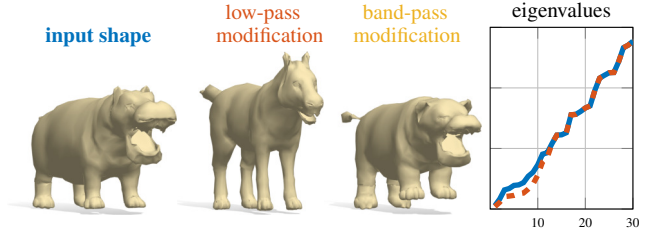


Figure 8. Exploring the space of shapes in real time via manipulation of the spectrum. The low-pass modification (middle) decreases the first 12 eigenvalues of the input shape; the band-pass modification (right) amplifies the last 12 eigenvalues. The damping of low eigenvalues leads to more pronounced geometric features (e.g. longer legs and snout), while amplification of mid-range eigenvalues affects the high-frequency details (e.g. the ears and fingers); see the supplementary video for a wall-clock demo.

show four examples (others can be found in the supplementary material). We emphasize here that the style is purely encoded in the input eigenvalues, therefore it does not rely on the test shapes being in point-to-point correspondence with the training set. This leads to the following:

Property 1 *Our method can be used in a **correspondence-free** scenario. By taking eigenvalues as input, it enables applications that traditionally require a correspondence, but side-steps this requirement.*

This observation was also mentioned in other spectrum-based approaches [19, 38]. However, the data-driven nature of our method makes it more robust, efficient and accurate, therefore greatly improving its practical utility.

6.2. Shape exploration

The results of Sec. 6.1 suggest that eigenvalues can be used to drive the exploration of the AE’s latent space toward a desired direction. Another possibility is to regard *the eigenvalues themselves* as a parametric model for isometry classes, and explore the “space of spectra” as is typically done with latent spaces. Our bi-directional coupling between spectra and latent codes makes this exploration feasible, as remarked by the following property:

Property 2 *Latent space connections provide both a means for **controlling** the latent space, and vice-versa, enable **exploration** of the space of Laplacian spectra.*

Since eigenvalues change continuously with the manifold metric [5], a small variation in the spectrum will give rise to a small change in the geometry. We can visualize such variations in shape directly, by first deforming a given spectrum (e.g., by a simple linear interpolation between two spectra) to obtain the new eigenvalue sequence μ , and then directly computing $D(\pi(\mu))$.

In Fig. 7 we show a related experiment. Here we train the network on 4,430 animal meshes generated with

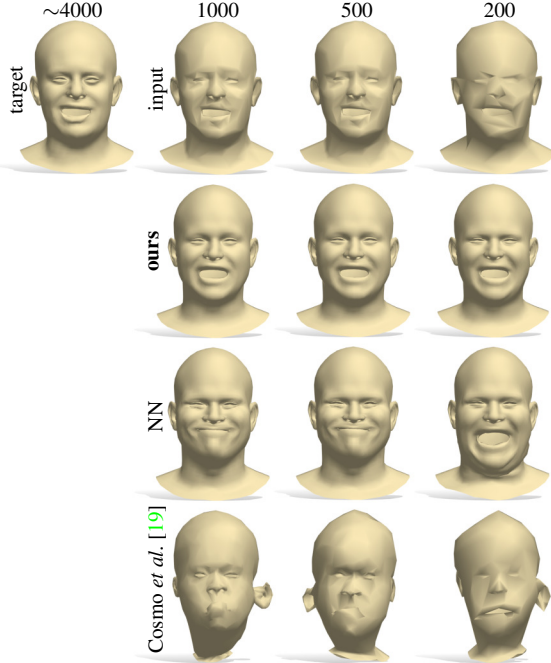


Figure 9. Mesh super-resolution for inputs at decreasing resolution (top row, left to right). Our method fits closely the original input shapes (top left), while other approaches either predict the wrong pose (NN baseline) or generate an unrealistic shape (Cosmo *et al.*).

the SMAL parametric model following the official protocol [50]. Given four *low-resolution* shapes \mathcal{X}_i as input, we first compute their spectra $\text{Spec}(\mathcal{X}_i)$, map these to the latent space via $\pi(\text{Spec}(\mathcal{X}_i))$, perform a bilinear interpolation of the resulting latent vectors, and finally reconstruct the corresponding shapes. Finally, in Fig. 8 we show an example of interactive spectrum-driven shape exploration. Given a shape and its Laplacian eigenvalues as input, we navigate the space of shapes by directly modifying different frequency bands with the aid of a simple user interface. The modified spectra are then decoded by our network in *real time*. The interactive nature of this application is enabled by the efficiency of our shape from spectrum recovery (obtained in a single forward pass) and would not be possible with previous methods [19] that rely on costly test-time optimization. We refer to the accompanying video and the supplementary materials for additional illustrations.

6.3. Super-resolution

A key feature that emerges from the experiment in Fig. 7 is the perfect reconstruction of the low-resolution shapes once their eigenvalues are mapped to the latent space via π . This brings us to a fundamental property of our approach:

Property 3 *Since eigenvalues are largely insensitive to mesh resolution and sampling, so is our trained network.*

This fact is especially evident when using cubic FEM discretization, as we do in all our tests, since it more closely

approximates the continuous setting and is thus much less affected by the surface discretization.

Remark. It is worth mentioning that existing methods can employ cubic FEM as well; however, this soon becomes prohibitively expensive due to the differentiation of spectral decomposition required by their optimizations [19, 38].

These properties allow us to use our network for the task of mesh super-resolution. Given a low-resolution mesh as input, our aim is to recover a higher resolution counterpart of it. Furthermore, while the input mesh has *arbitrary* resolution and is unknown to the network (and a correspondence with the training models is *not* given), an additional desideratum is for the new shape to be in dense point-to-point correspondence with models from the training set. We do so in a single shot, by predicting the decoded shape as:

$$\mathcal{X}_{\text{hires}} = D(\pi(\text{Spec}(\mathcal{X}_{\text{lowres}}))). \quad (8)$$

This simple approach exploits the resolution-independent geometric information encoded in the spectrum along with the power of a data-driven generative model.

In Fig. 9 we show a comparison with nearest-neighbors between eigenvalues (among shapes in the training set), and the isospectralization method of Cosmo *et al.* [19]. Our solution closely reproduces the high-resolution target. Isospectralization correctly aligns the eigenvalues, but it recovers unrealistic shapes due to ineffective regularization. This phenomenon highlights the following

Property 4 *Our data-driven approach replaces ad-hoc regularizers, that are difficult to model axiomatically, with realistic priors learned from examples.*

This is especially important for deformable objects; shapes falling into the same isometry class are often hard to disambiguate without using geometric priors.

6.4. Estimating point cloud spectra

As an additional experiment, we show how our network can directly predict Laplacian eigenvalues for unorganized point clouds. This task is particularly challenging due to the lack of a structure in the point set, and existing approaches such as [17, 6] often fail at approximating the eigenvalues of the underlying surface accurately. The difficulty is even more pronounced when the point sets are irregularly sampled, as we empirically show here. In our case, estimation of the spectrum boils down to the single forward pass:

$$\widetilde{\text{Spec}}(\mathcal{X}) = \rho(E(\mathcal{X})). \quad (9)$$

To address this task we train our network by feeding unorganized point clouds as input, together with the spectra computed from the corresponding meshes (which are available at training time). As described in the supplementary materials, for this setting we use a PointNet [37] encoder

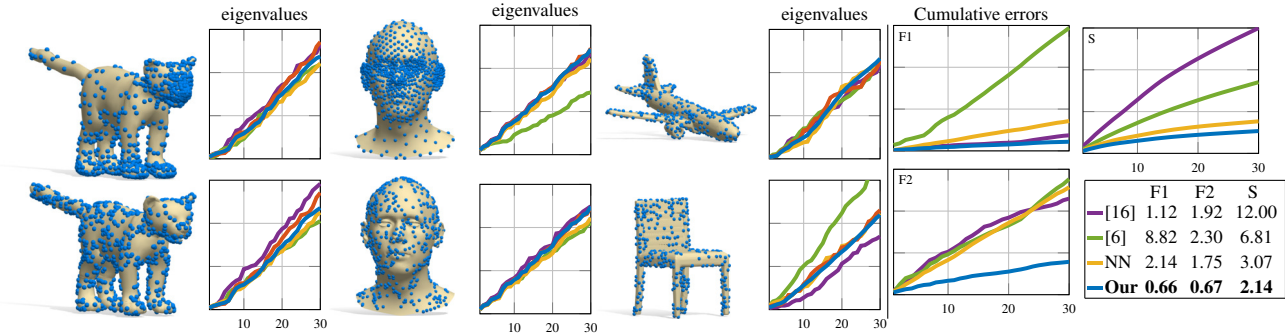


Figure 10. Qualitative and quantitative evaluation of point cloud spectra estimation. On the left we show the qualitative comparison for different samplings on three classes (animals, human faces and objects). We show the eigenvalues estimations alongside the input point cloud (depicted as surface samplings), and the ground truth spectrum (in red). On the last two columns, we report the average cumulative error curves evaluated on the FLAME dataset for the two different distributions (F1 and F2) and on ShapeNet (S).

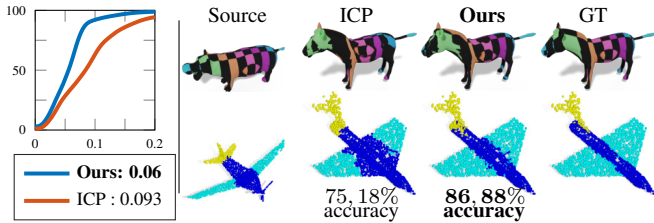


Figure 11. On the left, quantitative evaluation of matching [27] between 100 pairs of animals. On the right, the qualitative comparison on texture and segmentation transfer.

and a fully connected decoder, and we replace the reconstruction loss of Eq. (5) with the Chamfer distance. This application highlights the generality of our model, which can accommodate different representations of geometric data.

We consider two types of point clouds: (1) with similar point density and regularity as in the training set (shown in the supplementary materials), and (2) with randomized non-uniform sampling. We compare the spectrum estimated via $\rho(E(\mathcal{X}))$ to axiomatic methods [17, 6], and to the NN baseline (applied in the latent space); see Fig. 10. The qualitative results are obtained by training on SMAL [50] (left), COMA [39] (middle) and ShapeNet watertight [23] (right). To highlight its generalization capability, the network trained on COMA is tested on point clouds from the FLAME dataset, while on ShapeNet we consider 4 different classes (airplanes, boats, screens and chairs). We compute the cumulative error curves of the distance between the eigenvalues from the meshes corresponding to the test point clouds. The mean error across all test sets is also reported in the legend. Our method leads to a significant improvement over the closest state-of-the-art baseline [6].

6.5. Matching from spectrum

Finally, we compute dense correspondences between shape pairs using only their spectra. These are fed into our network; since the output points are naturally ordered by

the decoder, we exploit this to establish a sparse correspondence. In the case of meshes, we extend it to a dense one by using the functional maps framework [33]. In the case of point clouds, we can propagate a semantic segmentation using nearest neighbors. We perform a quantitative evaluation on SMAL [50], testing on 100 non-isometric pairs of animals from different classes. Two applications that benefit from our approach are texture and segmentation transfer; we tested them respectively on animals and segmented ShapeNet [49]. The comparison baseline consists of 100 iterations of ICP [7] to rigidly align the two shapes followed by nearest-neighbor assignment as correspondence. See Fig. 11 and the supplementary for further details.

7. Conclusions

We introduced the first data-driven method for shape generation from Laplacian spectra. Our approach consists in enriching a standard AE with a pair of cycle-consistent maps, associating ordered sequences of eigenvalues to latent codes and vice-versa. This explicit coupling brings forth key advantages of spectral methods to generative models, enabling novel applications and a significant improvement over existing approaches. Our limitations are shared with other spectral methods in the computation of a robust Laplacian discretization. Adopting the recent approach [44] for such borderline cases is a promising possibility. Further, while the Laplacian is a classical choice due to its Fourier-like properties, spectra of other operators with different properties may lead to other promising applications.

Acknowledgements We gratefully acknowledge Luca Moschella and Silvia Casola for the technical support, Nicholas Sharp for the useful suggestions about pointcloud spectra. Parts of this work were supported by the KAUST OSR Award No. CRG-2017-3426, the ERC Starting Grant No. 758800 (EXPROTEA), the ERC Starting Grant No. 802554 (SPECGEO), the ANR AI Chair AI-GRETTE, and the MIUR under grant “Dipartimenti di eccellenza 2018-2022” of the Department of Computer Science of Sapienza University and University of Verona.

References

- [1] D. Aasen, T. Bhamre, and A. Kempf. Shape from sound: toward new tools for quantum gravity. *Physical review letters*, 110(12):121301, 2013. [2](#)
- [2] P. Achlioptas, O. Diamanti, I. Mitliagkas, and L. Guibas. Learning representations and generative models for 3d point clouds. In *International Conference on Machine Learning*, pages 40–49, 2018. [1](#), [3](#)
- [3] M. Aubry, U. Schlickewei, and D. Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pages 1626–1633. IEEE, 2011. [2](#)
- [4] T. Aumentado-Armstrong, S. Tsogkas, A. Jepson, and S. Dickinson. Geometric disentanglement for generative latent shape models. In *International Conference on Computer Vision (ICCV)*, 2019. [3](#)
- [5] S. Bando and H. Urakawa. Generic properties of the eigenvalue of the laplacian for compact riemannian manifolds. *Tohoku Mathematical Journal, Second Series*, 35(2):155–172, 1983. [6](#)
- [6] M. Belkin, J. Sun, and Y. Wang. Constructing laplace operator from point clouds in rd. In *Proceedings of the twentieth annual ACM-SIAM symposium on Discrete algorithms*, pages 1031–1040. Society for Industrial and Applied Mathematics, 2009. [7](#), [8](#)
- [7] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, Feb 1992. [8](#)
- [8] G. Bharaj, D. I. Levin, J. Tompkin, Y. Fei, H. Pfister, W. Matusik, and C. Zheng. Computational design of metallophone contact sounds. *ACM Transactions on Graphics (TOG)*, 34(6):223, 2015. [2](#)
- [9] D. Boscaini, D. Eynard, D. Kourounis, and M. M. Bronstein. Shape-from-operator: recovering shapes from intrinsic operators. *Computer Graphics Forum*, 34(2):265–274, 2015. [2](#)
- [10] D. Boscaini, J. Masci, E. Rodolà, M. M. Bronstein, and D. Cremers. Anisotropic diffusion descriptors. *Computer Graphics Forum*, 35(2):431–441, 2016. [3](#)
- [11] A. M. Bronstein, M. M. Bronstein, L. J. Guibas, and M. Ovsjanikov. Shape google: Geometric words and expressions for invariant shape retrieval. *ACM Transactions on Graphics (TOG)*, 30(1):1, 2011. [2](#)
- [12] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017. [1](#)
- [13] I. Chavel. *Eigenvalues in Riemannian Geometry*. Academic Press, 1984. [2](#)
- [14] Z. Chen and H. Zhang. Learning implicit fields for generative shape modeling. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5939–5948, 2019. [3](#)
- [15] M. Chu, G. Golub, and G. H. Golub. *Inverse eigenvalue problems: theory, algorithms, and applications*, volume 13. Oxford University Press, 2005. [2](#)
- [16] P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40. Siam, 2002. [3](#)
- [17] U. Clarenz, M. Rumpf, and A. Telea. Finite elements on point based surfaces. In *Proceedings of the First Eurographics conference on Point-Based Graphics*, pages 201–211. Eurographics Association, 2004. [3](#), [7](#), [8](#)
- [18] E. Corman, J. Solomon, M. Ben-Chen, L. Guibas, and M. Ovsjanikov. Functional Characterization of Intrinsic and Extrinsic Geometry. *ACM Transactions on Graphics*, 17, 2017. [2](#)
- [19] L. Cosmo, M. Panine, A. Rampini, M. Ovsjanikov, M. M. Bronstein, and E. Rodolà. Isospectralization, or how to hear shape, style, and correspondence. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7529–7538, 2019. [2](#), [3](#), [4](#), [5](#), [6](#), [7](#)
- [20] L. Gao, J. Yang, T. Wu, Y.-J. Yuan, H. Fu, Y.-K. Lai, and H. Zhang. Sdm-net: Deep generative network for structured deformable mesh. *arXiv preprint arXiv:1908.04520*, 2019. [1](#), [3](#)
- [21] C. Gordon, D. L. Webb, and S. Wolpert. One cannot hear the shape of a drum. *Bulletin of the American Mathematical Society*, 27(1):134–138, 1992. [2](#)
- [22] T. Groueix, M. Fisher, V. G. Kim, B. Russell, and M. Aubry. AtlasNet: A Papier-Mâché Approach to Learning 3D Surface Generation. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2018. [3](#)
- [23] J. Huang, H. Su, and L. Guibas. Robust watertight manifold surface generation method for shapenet models. *arXiv preprint arXiv:1802.01698*, 2018. [8](#)
- [24] R. Huang, M.-J. Rakotosaona, P. Achlioptas, L. Guibas, and M. Ovsjanikov. Operatornet: Recovering 3d shapes from difference operators. In *ICCV*, 2019. [3](#)
- [25] M. Kac. Can one hear the shape of a drum? *The american mathematical monthly*, 73(4P2):1–23, 1966. [2](#)
- [26] Z. Karni and C. Gotsman. Spectral compression of mesh geometry. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00*, pages 279–286. ACM Press/Addison-Wesley Publishing Co., 2000. [2](#)
- [27] V. G. Kim, Y. Lipman, and T. Funkhouser. Blended intrinsic maps. In *ACM Transactions on Graphics (TOG)*, volume 30, page 79. ACM, 2011. [8](#)
- [28] I. Kostrikov, Z. Jiang, D. Panozzo, D. Zorin, and J. Bruna. Surface networks. In *Proc. CVPR*, 2018. [3](#)
- [29] J. Li, K. Xu, S. Chaudhuri, E. Yumer, H. Zhang, and L. Guibas. Grass: Generative recursive autoencoders for shape structures. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2017)*, 36(4):52–56, 2017. [3](#)
- [30] O. Litany, A. Bronstein, M. Bronstein, and A. Makadia. Deformable shape completion with graph convolutional autoencoders. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1886–1895, 2018. [1](#), [3](#)
- [31] J. Masci, E. Rodolà, D. Boscaini, M. M. Bronstein, and H. Li. Geometric deep learning. In *SIGGRAPH ASIA 2016 Courses*, page 1. ACM, 2016. [1](#)

- [32] K. Mo, P. Guerrero, L. Yi, H. Su, P. Wonka, N. Mitra, and L. J. Guibas. StructureNet: Hierarchical graph networks for 3d shape generation. *arXiv preprint arXiv:1908.00575*, 2019. 1, 3
- [33] M. Ovsjanikov, M. Ben-Chen, J. Solomon, A. Butscher, and L. Guibas. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (TOG)*, 31(4):30:1–30:11, 2012. 2, 8
- [34] C. Öztireli, M. Alexa, and M. Gross. Spectral sampling of manifolds. *ACM Transactions on Graphics (TOG)*, 29(6):168, 2010. 2
- [35] M. Panine and A. Kempf. Towards spectral geometric methods for euclidean quantum gravity. *Physical Review D*, 93(8):084033, 2016. 2
- [36] U. Pinkall and K. Polthier. Computing discrete minimal surfaces and their conjugates. *Experimental mathematics*, 2(1):15–36, 1993. 3
- [37] C. R. Qi, H. Su, K. Mo, and L. 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. 7
- [38] A. Rampini, I. Tallini, M. Ovsjanikov, A. M. Bronstein, and E. Rodolà. Correspondence-free region localization for partial shape similarity via hamiltonian spectrum alignment. In *International Conference on 3D Vision (3DV)*, 2019. 2, 3, 5, 6, 7
- [39] A. Ranjan, T. Bolkart, S. Sanyal, and M. J. Black. Generating 3D faces using convolutional mesh autoencoders. In *European Conference on Computer Vision (ECCV)*, 2018. 5, 8
- [40] M. Reuter. Hierarchical shape segmentation and registration via topological features of laplace-beltrami eigenfunctions. *International Journal of Computer Vision*, 89(2-3):287–308, 2010. 2, 3
- [41] M. Reuter, F.-E. Wolter, and N. Peinecke. Laplace-spectra as fingerprints for shape matching. In *Proceedings of the 2005 ACM symposium on Solid and physical modeling*, pages 101–106. ACM, 2005. 2
- [42] J.-M. Roufosse, A. Sharma, and M. Ovsjanikov. Unsupervised deep learning for structured shape matching. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1617–1627, 2019. 5
- [43] R. M. Rustamov, M. Ovsjanikov, O. Azencot, M. Ben-Chen, F. Chazal, and L. Guibas. Map-based exploration of intrinsic shape differences and variability. *ACM Transactions on Graphics (TOG)*, 32(4), 2013. 2, 3
- [44] N. Sharp, Y. Soliman, and K. Crane. Navigating intrinsic triangulations. *ACM Trans. Graph.*, 38(4):55:1–55:16, July 2019. 8
- [45] A. Sinha, A. Unmesh, Q.-X. Huang, and K. Ramani. SurfNet: Generating 3d shape surfaces using deep residual networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 791–800, 2017. 3
- [46] J. Sun, M. Ovsjanikov, and L. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum*, 28(5):1383–1392, 2009. 2
- [47] J. Wu, C. Zhang, T. Xue, W. T. Freeman, and J. B. Tenenbaum. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. In *Advances in Neural Information Processing Systems*, pages 82–90, 2016. 3
- [48] Z. Wu, X. Wang, D. Lin, D. Lischinski, D. Cohen-Or, and H. Huang. Sagnet: Structure-aware generative network for 3d-shape modeling. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2019)*, 38(4):91:1–91:14, 2019. 3
- [49] L. Yi, L. Shao, M. Savva, H. Huang, Y. Zhou, Q. Wang, B. Graham, M. Engelcke, R. Klokov, V. Lempitsky, et al. Large-scale 3d shape reconstruction and segmentation from shapenet core55. *arXiv preprint arXiv:1710.06104*, 2017. 8
- [50] S. Zuffi, A. Kanazawa, D. Jacobs, and M. J. Black. 3D menagerie: Modeling the 3D shape and pose of animals. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, July 2017. 7, 8