Spectral Analysis Towards Geometric Auto-Encoding of Subcortical Structures

Loic Le Folgoc[1], Bishesh Khanal[2], Amir Alansary[1],
Konstantinos Kamnitsas[1], Ian Walker[1], and Ben Glocker[1]

[1]Imperial College London, UK , email: l.le-folgoc@imperial.ac.uk
[2]King's College London, UK.

Abstract

Statistical shape analysis can benefit from algorithms that are intrinsic to the shape; multi-scale and hierarchical; robust to perturbations, yet sensitive to fine-grained content. For this purpose we investigate Deep Spectral Kernels (DSKs), trainable and hierarchical similarity functions based on spectral analysis. DSKs are shape encoders. They are multi-layer and compositional architectures that map semi-structured, high dimensional objects such as meshes to representations fit for machine learning. The encoding generates a sequence of increasingly high-level geometries, augmented with functional maps that summarize relevant details from finer scales. At the core of the procedure, the Spectral Wavelet Transform allow for the encoding of structural and functional data to be done in a shape-intrinsic manner. We experiment with unsupervised clustering of subcortical structures.

1 Introduction

Most approaches to statistical shape analysis fundamentally rely on registration, from landmark based representations and active shape models,to medial representations and Principal Geodesic Analysis,to deformable registration and diffeomorphometry.In contrast, relevant information is often invariant to the object pose. Enforcing such invariances within statistical representations should help algorithms learn and generalise well from little data. Spectral shape descriptors (8; 7), built from the spectrum and eigenfunctions of the Laplace(-Beltrami) operator, have achieved popularity in that context.Still it remains an open question how to best structure the framework of spectral analysis into powerful statistical learning tools. Truncating the spectrum (8) yields a finite-dimensional shape descriptor, albeit with a loss of fine-grained information. Low-dimensional embeddings can be derived instead from metrics based on the full spectrum (10; 5) or on histograms of local spectral descriptors (1). As an alternative to finite-dimensional representations, (6) designs spectral kernels (*i.e.* non-linear similarity metrics between shapes) that make generic kernel methods available for the purpose of shape analysis. Several limitations remain to be addressed for spectral methods to be more discriminative and interpretable, including the loss of fine-grained information in non-hierarchical approaches; and the lack of specificity to the spatial configuration of salient features in bag-of-feature based methods. This extended abstract discusses Deep Spectral Kernels (DSKs) as a family of trainable, multi-scale, hierarchical shape kernels. DSKs process shapes sequentially in fine-to-coarse manner, to derive a shape representation that is: (1) sensitive to both *local* fine-grained structure and *global* characteristics; (2) *invariant* to pose (registration-free), *robust* to perturbations and *smoothly* varying under smooth deformation of the shape; and (3) *compositional*: it uses learned template "shape patches" and information about their joint spatial layout to describe the higher-level geometry. DSKs can be interpreted as shape auto-encoders. This analogy suggests a simple scheme to learn DSK parameters, similar to the pre-training proposed by Hinton et al. (4) for image auto-encoders.

1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

2 Method

The study of deep architectures on non-Euclidean geometries has recently emerged as Geometric Deep Learning (2). Much of the literature has focused on extending the widely successful Convolutional Neural Network architecture from regular grids (*e.g.* for time series and images) to general graphs. In contrast to images, convolution on non-Euclidean meshes does not define a shift-invariant operation. Thus one of the appeals of the CNN architecture is lost, namely its interpretation as compositional template matching. How strongly the spectral filtering operation is tied to the underlying geometry is an open problem. For this reason, we aim instead to extend *template matching*. A layer records the similarity of each point's local neighborhood on the mesh to a bank of learned prototypes, via an intermediate shape-intrinsic representation described below.

Multiscale mesh representation. We consider the case of orientable surface meshes in $3D$. Meshes are smoothed and coarsened at multiple scales (the lth layer of the DSK operates on such a smoothed mesh at scale σ_l). The mesh at scale σ_l is obtained by evolving the coordinates following an incompressible mean curvature flow $\partial_t X = \Delta X + 2h_{\text{avg}}/A_t \cdot \vec{n}(X)$, starting from the initial mesh, from time 0 to $T = \sigma_l^2$ (with h_{avg} the average mean curvature over the surface, A_t the surface area, $\vec{n}(X)$ the surface normal).

Spectral Wavelet Transform (SWT). The SWT is a (multidimensional) signature for functions over shapes, obtained by recording the response with a bank of spectral filters. We restrict ourselves to filters for which $\langle h|\phi_n\rangle \coloneqq \bar{h}(\lambda_n)$ is a function of the spectrum only. These filters can be defined independently of the object geometry by how they modulate the signal in various spectral bands (*e.g.*, low- or band-pass filters), and by abuse of notation we write $h(\lambda_n)$. We define it analogously to (3), but with spatial pooling. Given $\boldsymbol{h} = (h_1 \cdots h_K)$ a bank of spectral filters and $f : \Omega \to \mathbb{R}$ a function on a shape Ω, let $\mathcal{S}_{\boldsymbol{h}}[f] \triangleq (\langle f * h_i\rangle_\Omega)_{i=1\cdots K} \in \mathbb{R}_+^K$. It is defined entirely in terms of statistics of f within various spectral bands. Hence it is suitable for similarities between functions on *distinct* shapes. $\mathcal{S}_{\boldsymbol{h}}$ offers *frequency* (spectral) localisation. *Spatial* localisation is recovered by pooling over a local neighborhood instead of the whole shape. **Shape signature.** The SWT is a (local) signature for a function over a shape. It can also be used to encode the characteristics of the shape itself with suitable choices of function. Letting $f \coloneqq \delta_{\mathrm{x}}$, the delta Dirac centered at point x, and $h_i = K_{t_i}$ the heat ("Gaussian") kernel for some diffusion time t_i relates to the *Heat Kernel Signature* (9) via $\mathrm{HKS}_{t_i}(\mathrm{x},\mathrm{x}) \triangleq (K_{t_i} * \delta_{\mathrm{x}})(\mathrm{x})$. The HKS is invariant w.r.t. isometric transformations of Ω. To recover additional geometric information, the signed arc-length $\mathrm{MCFS}_{t_i} \triangleq \int_0^{t_i} \langle \partial_t X | \vec{n}(X)\rangle dt$ traveled over the mean curvature flow is readily available. Together they capture much of the intrinsic and extrinsic local geometrical information (for $t_i \to 0$, the former relates up to first order to the Gaussian curvature and the later to the mean curvature at x).

Deep Spectral Kernels (DSKs). Deep Spectral Kernels (DSKs) are trainable, hierarchical constructs formed by stacking SWT and kernel non-linearities at multiple scales. DSKs are built in fine-to-coarse fashion. Each layer outputs a set of activation maps that record the response to (learned) templates. The activation maps output by layer $l-1$ are transported to the subsequent, coarser geometry (scale σ_l) to form the input of the lth layer. The SWT of these input maps over σ_l-neighborhoods is computed, densely at every point x on the mesh. This functional signature summarizes information about joint/co-occurrence of finer-grained structures within the σ_l-neighborhood. This signature is concatenated with the structural signature $(\mathrm{HKS}_{\sigma_l}(\mathrm{x}), \mathrm{MCFS}_{\sigma_l}(\mathrm{x}))$ at scale σ_l, and together form a multiscale representation of the neighborhood of x at all scales below σ_l. Layer l outputs activation maps that capture the similarity between a given σ_l-neighborhood and a bank of learned templates. The similarity between a given signature and a prototype signature is given by a (learned) multivariate Gaussian RBF (the mean is interpreted as the prototype/template, the covariance parameters account for suitable combinations and weighting of various features). Subsequent layers operate in the same fashion. The parameters in the deep kernel construction are the templates at each scale and their number, and the covariance matrices in the Gaussian RBFs. The hyperparameters are the filterbanks as well as the number and choice of layers/scales.

Learning DSKs. The parameters of the DSK can be trained layer-wise in an unsupervised manner, by k-Means or Variational Bayes EM approaches. The procedure is maybe best understood by analogy to unsupervised learning of a stack of RBMs (4), an effective early strategy for pre-training of auto-encoder networks.

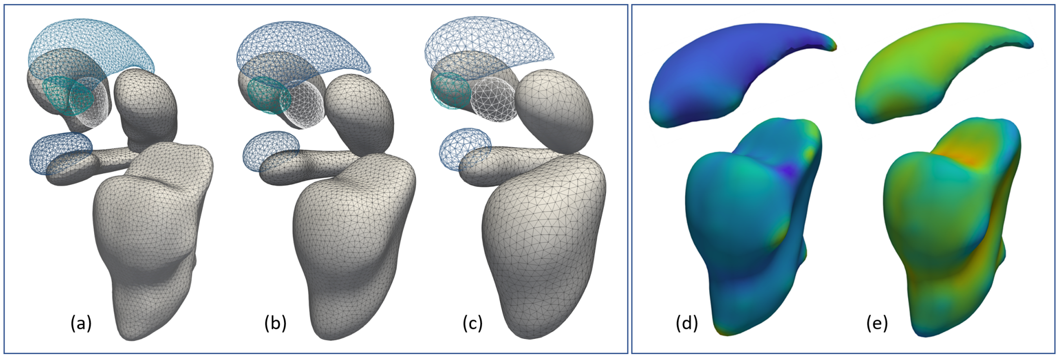

Figure 1: (Left) Hierarchical shape coarsening, at three different scales (a. $\sigma = 0.5$mm, b. $\sigma = 3$mm, c. $\sigma = 6$mm). From top to bottom and left to right, the subcortical structures are: caudate, putamen, accumbens, pallidum, thalamus, amygdala, hippocampus, brain stem. (Right) Spectral Signature for the first layer of the Deep Kernel, displayed for the caudate and stem. The local geometry around each point is summarised via 2 scalar values: (d) Heat Kernel Signature, (e) Mean Curvature Flow Signature. The behaviour especially differs along geometric ridges and at saddle points.

3 Experiments & Discussion

We experimented with a dataset of 100 training/test subjects from the Brain Biobank dataset, containing 15 subcortical structures (Fig 1). We trained a DSK of 3 layers (a. $\sigma = 0.5$mm, b. $\sigma = 3$mm, c. $\sigma = 6$mm) with 8 learned prototypes at each level. The output of the last two layers (normalised to discard surface/volume information) forms the shape signature. The mean signature per class is computed, and test points are assigned to the closest signature. We achieved class-wise accuracies of 0.73 (stem), 0.99 (accu), 0.46 (amyg), 0.55 (caud), 0.87 (hipp), 0.67 (pall), 0.45 (puta), 0.74 (thal).

Future work will investigate end-to-end training strategies, and strategies to decode the shape signature to reconstruct the mesh.

References

[1] Bronstein, A.M., Bronstein, M.M., Guibas, L.J., Ovsjanikov, M.: Shape Google: Geometric words and expressions for invariant shape retrieval. ACM Trans Graph 30(1), 1 (2011) 1

[2] Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P.: Geometric deep learning: going beyond euclidean data. IEEE Signal Processing Magazine 34(4), 18–42 (2017) 2

[3] Hammond, D.K., Vandergheynst, P., Gribonval, R.: Wavelets on graphs via spectral graph theory. Applied and Computational Harmonic Analysis 30(2), 129–150 (2011) 2

[4] Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. science 313(5786), 504–507 (2006) 1, 2

[5] Konukoglu, E., Glocker, B., Criminisi, A., Pohl, K.M.: WESD–Weighted Spectral Distance for measuring shape dissimilarity. IEEE Trans. Pattern Anal. Mach. Intell. 35(9), 2284–2297 (2013) 1

[6] Le Folgoc, L., Nori, A.V., Criminisi, A.: Spectral kernels for probabilistic analysis and clustering of shapes. In: International Conference on Information Processing in Medical Imaging. pp. 67–79. Springer (2017) 1

[7] Raviv, D., Bronstein, M.M., Bronstein, A.M., Kimmel, R.: Volumetric heat kernel signatures. In: ACM Workshop on 3D Object Retrieval. pp. 39–44. ACM (2010) 1

[8] Reuter, M., Wolter, F.E., Peinecke, N.: Laplace–Beltrami spectra as Shape-DNA of surfaces and solids. Computer-Aided Design 38(4), 342–366 (2006) 1

[9] Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. In: Computer graphics forum. vol. 28, pp. 1383–1392. Wiley Online Library (2009) 2

[10] Wachinger, C., Golland, P., Kremen, W., Fischl, B., Reuter, M., ADNI, et al.: Brainprint: a discriminative characterization of brain morphology. NeuroImage 109, 232–248 (2015) 1