Master SIF - REP (17/20)
Advanced tools for Digital Image Processing II

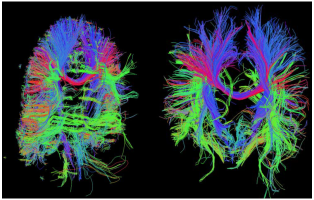
Thomas Maugey
thomas.maugey@inria.fr

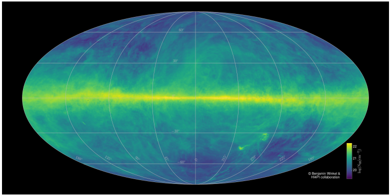
UNIVERSITÉ DE RENNES 1

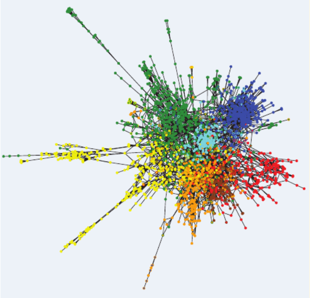
informatics mathematics Inria

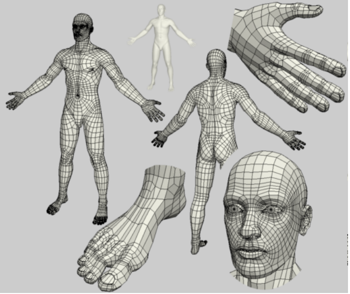
Fall 2019

Image lying on irregular domain



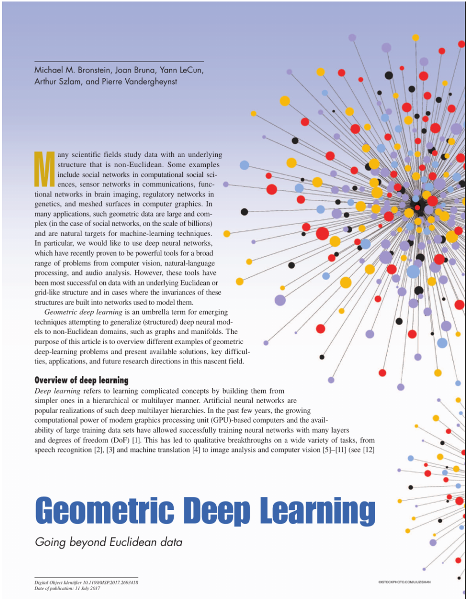






How to perform Deep learning on such data ?

A deep review of the field



Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst

Many scientific fields study data with an underlying structure that is non-Euclidean. Some examples include social networks in computational social sciences, sensor networks in communications, functional networks in brain imaging, regulatory networks in genetics, and meshed surfaces in computer graphics. In many applications, such geometric data are large and complex (in the case of social networks, on the scale of billions) and are natural targets for machine-learning techniques. In particular, we would like to use deep neural networks, which have recently proven to be powerful tools for a broad range of problems from computer vision, natural-language processing, and audio analysis. However, these tools have been most successful on data with an underlying Euclidean or grid-like structure and in cases where the invariances of these structures are built into networks used to model them.

Geometric deep learning is an umbrella term for emerging techniques attempting to generalize (structured) deep neural models to non-Euclidean domains, such as graphs and manifolds. The purpose of this article is to overview different examples of geometric deep-learning problems and present available solutions, key difficulties, applications, and future research directions in this nascent field.

Overview of deep learning

Deep learning refers to learning complicated concepts by building them from simpler ones in a hierarchical or multilayer manner. Artificial neural networks are popular realizations of such deep multilayer hierarchies. In the past few years, the growing computational power of modern graphics processing unit (GPU)-based computers and the availability of large training data sets have allowed successfully training neural networks with many layers and degrees of freedom (DoF) [1]. This has led to qualitative breakthroughs on a wide variety of tasks, from speech recognition [2], [3] and machine translation [4] to image analysis and computer vision [5]–[11] (see [12]

Geometric Deep Learning

Going beyond Euclidean data

Digital Object Identifier 10.1109/MSP.2017.2693418
Date of publication: 11 July 2017

CNN on graphs, what's the problem ?

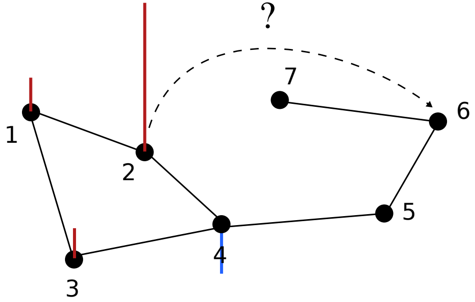
CNN $\rightarrow$ Convolution $\rightarrow$

CNN on graphs, what's the problem ?

CNN $\rightarrow$ Convolution $\rightarrow$ Translation

CNN on graphs, what's the problem ?

CNN → Convolution → Translation



CNN on graphs, what's the problem ?

CNN $\rightarrow$

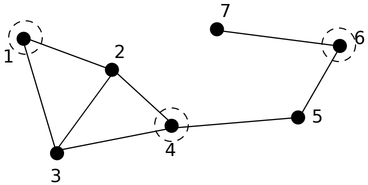
CNN on graphs, what's the problem ?

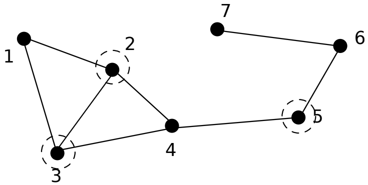
CNN → Pooling (downsampling)

CNN on graphs, what's the problem ?

CNN → Pooling (downsampling)

50% of the nodes ?





CNN on graphs, what's the problem ?

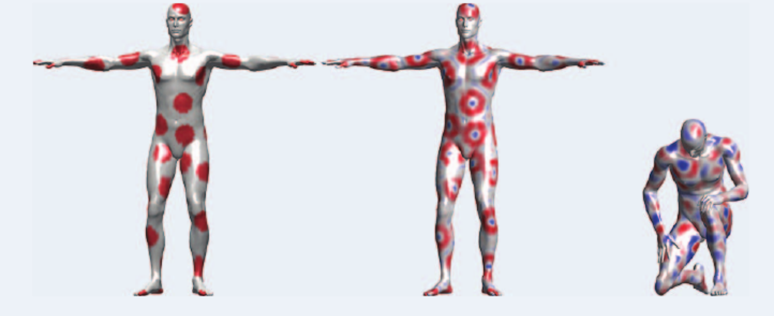
CNN $\rightarrow$

CNN on graphs, what's the problem ?

CNN $\rightarrow$ Fixed grid

CNN on graphs, what's the problem ?

CNN $\rightarrow$ Fixed grid



From left to right: a signal, frequency-domain edge detection, same detection applied when the topology slightly changes

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

Table of Contents

Table of Contents

Filtering

In classical signal processing:

In graph signal processing:

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Spectral definitions

Filtering

In classical signal processing:

$$\hat{f}_{\text{out}}(\omega) = \hat{f}_{\text{in}}(\omega)\hat{h}(\omega)$$

In graph signal processing:

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Filtering

In classical signal processing:

$$\hat{f}_{\text{out}}(\omega) = \hat{f}_{\text{in}}(\omega)\hat{h}(\omega)$$

In graph signal processing:

$$\hat{f}_{\text{out}}(\lambda_l) = \hat{f}_{\text{in}}(\lambda_l)\hat{h}(\lambda_l)$$

which gives

$$f_{\text{out}}(n) = \sum_{l=0}^{N-1} \hat{f}_{\text{in}}(\lambda_l)\hat{h}(\lambda_l)u_l(n)$$

it can also be written as

$$\mathbf{f}_{\text{out}} = \hat{h}(\mathbf{L})\mathbf{f}_{\text{in}}, \quad \text{with} \quad \hat{h}(\mathbf{L}) = \mathbf{U}\begin{bmatrix} \hat{h}(\lambda_0) & & 0 \\ & \ddots & \\ 0 & & \hat{h}(\lambda_{N-1}) \end{bmatrix}\mathbf{U}^\top$$

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Convolution

In classical signal processing:

In graph signal processing:

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Spectral definitions

Convolution

In classical signal processing:

$$(f * h)(t) = \int_{\mathbb{R}} f(\tau)h(t-\tau)d\tau$$

which can be written as

$$(f * h)(t) = \int_{\mathbb{R}} \hat{f}(\omega)\hat{h}(\omega)e^{2i\pi\omega t}d\omega$$

In graph signal processing:

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Spectral definitions

Convolution

In classical signal processing:

$$(f * h)(t) = \int_{\mathbb{R}} f(\tau)h(t-\tau)d\tau$$

which can be written as

$$(f * h)(t) = \int_{\mathbb{R}} \hat{f}(\omega)\hat{h}(\omega)e^{2i\pi\omega t}d\omega$$

In graph signal processing:

$$(f * h)(n) := \sum_{l=0}^{N-1} \hat{f}(\lambda_l)\hat{h}(\lambda_l)u_l(n)$$

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Spectral definitions

Translation

In classical signal processing:

In graph signal processing:

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Spectral definitions

Translation

In classical signal processing:

$$(\mathcal{T}_\tau f)(t) = f(t-\tau)$$

which can be written as

$$(\mathcal{T}_\tau f)(t) = (f * \delta_\tau)(t)$$

In graph signal processing:

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Translation

In classical signal processing:

$$(\mathcal{T}_\tau f)(t) = f(t-\tau)$$

which can be written as

$$(\mathcal{T}_\tau f)(t) = (f * \delta_\tau)(t)$$

In graph signal processing:

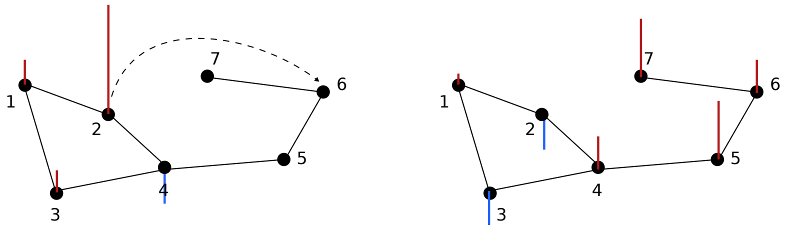
$$(\mathcal{T}_k f)(n) = \sqrt{N}(f * \delta_k)(n)$$

which becomes

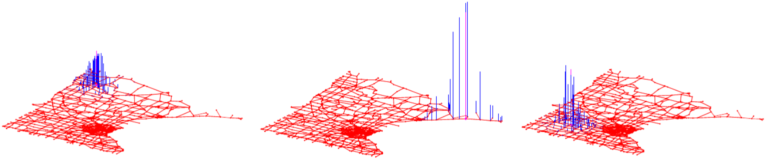
$$(\mathcal{T}_k f)(n) = \sqrt{N} \sum_{l=0}^{N-1} \hat{f}(\lambda_l) u_l(k) u_l(n)$$

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Spectral definitions, a good solution?

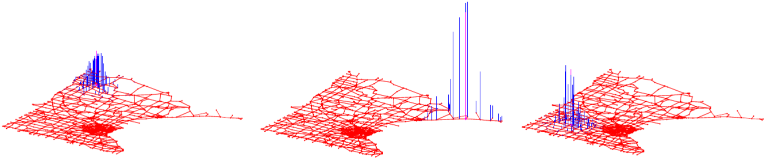


Spectral definitions, a good solution?



[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Spectral definitions, a good solution?



The spectrum is the same, however, the spatial shape is different. It can be a problem

[D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 83-98, May 2013.]

Table of Contents

Recall of Shannon-Nyquist theorem

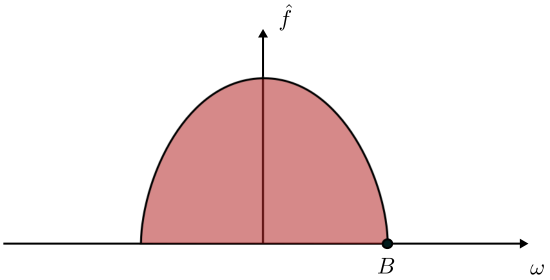
Let us consider a signal f that contains no frequencies higher than B:

$$\forall \omega, \quad \text{s.t. } |\omega| > B, \quad \text{then} \quad \hat{f}(\omega) = 0.$$

Recall of Shannon-Nyquist theorem

Let us consider a signal f that contains no frequencies higher than B:

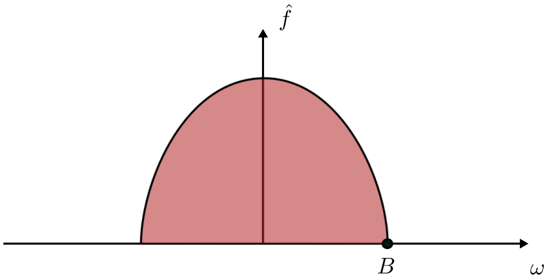
$$\forall \omega, \quad \text{s.t. } |\omega| > B, \quad \text{then} \quad \hat{f}(\omega) = 0.$$



Recall of Shannon-Nyquist theorem

Let us consider a signal f that contains no frequencies higher than B:

$$\forall \omega, \quad \text{s.t. } |\omega| > B, \quad \text{then} \quad \hat{f}(\omega) = 0.$$



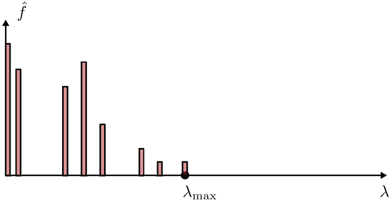
This signal can be sampled at a frequency of $2B$ and fully recovered.

Extension to graphs

Let us consider a signal $\mathbf{f}$ defined on a graph $\mathcal{G}$ that is bandlimited with a bandwidth $\lambda_{\max}$:

$$\forall\ \lambda_l > \lambda_{\max}, \quad \hat{f}(\lambda_l) = 0$$

The set of bandlimited signals $\lambda_{\max}$ with bandwidth is called the **Paley-Wiener space** $PW_{\lambda_{\max}}(\mathcal{G})$.



A **uniqueness set** for $PW_{\lambda_{\max}}(\mathcal{G})$ is a subset of vertices $\mathcal{S} \subset \mathcal{V}$ for which

$$\forall f, g \in PW_{\lambda_{\max}}(\mathcal{G}),\ f(\mathcal{S}) = g(\mathcal{S}) \Rightarrow f = g$$

The smallest **uniqueness set** for $PW_{\lambda_l}(\mathcal{G})$ has a size of l

Recovering the missing samples

Let $f \in PW_{\lambda_n}$, and $\mathcal{S}$ be a minumum uniqueness set.
Estimate the graph Fourier coefficients

$$\mathbf{f} = \begin{bmatrix} \mathbf{f}(\mathcal{S}) \\ \mathbf{f}(\mathcal{S}^c) \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1(\mathcal{S}) & \dots & \mathbf{u}_N(\mathcal{S}) \\ \mathbf{u}_1(\mathcal{S}^c) & \dots & \mathbf{u}_N(\mathcal{S}^c) \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}(1) \\ \vdots \\ \mathbf{f}(\hat{n}) \\ \mathbf{0} \end{bmatrix}$$

Thus

$$\begin{bmatrix} \hat{\mathbf{f}}(1) \\ \vdots \\ \hat{\mathbf{f}}(n) \end{bmatrix} = (\tilde{\mathbf{U}}_n(\mathcal{S}))^{-1}\mathbf{f}(\mathcal{S})$$

with $\tilde{\mathbf{U}}_n$ being the n first eigenvectors.
Finally,

$$\mathbf{f}(\mathcal{S}^c) = \tilde{\mathbf{U}}_n(\mathcal{S}^c) \begin{bmatrix} \hat{\mathbf{f}}(1) \\ \vdots \\ \hat{\mathbf{f}}(n) \end{bmatrix}$$

[Narang, S. K., Gadde, A., Sanou, E., and Ortega, A. (2013, December). Localized iterative methods for interpolation in graph structured data. In 2013 IEEE Global Conference on Signal and Information Processing (pp. 491-494). IEEE.]

Graph sampling

In the equation $\begin{bmatrix} \hat{\mathbf{f}}(1) \\ \vdots \\ \hat{\mathbf{f}}(n) \end{bmatrix} = (\tilde{\mathbf{U}}_n(\mathcal{S}))^{-1}\mathbf{f}(\mathcal{S})$, the matrix $\tilde{\mathbf{U}}_n(\mathcal{S})$ should be invertible.

Sampling algorithms consist in finding the set $\mathcal{S}$ for which, $\tilde{\mathbf{U}}_n(\mathcal{S})$ is invertible

Initialize: $\mathcal{S} \leftarrow \mathcal{V}_i$ where i is the index of any nonzero element offirst eigenvector
for $m = 2 \rightarrow n$ **do**
  Compute $\mathbf{x} = null(\tilde{\mathbf{U}}_m(\mathcal{S}))$
  Compute $\mathbf{b} = \tilde{\mathbf{U}}_m(\mathcal{S}^c)\mathbf{x}$
  $i \leftarrow argmax_i(|\mathbf{b}(i)|)$
  $\mathcal{S} \leftarrow \mathcal{S} \cup \mathcal{S}_c(i)$
end

[D. E. Tzamarias, P. Akyazi, and P. Frossard, âĂA novel method for sampling bandlimited graph signals,âĂİ in Proceedings of EUSIPCO, no. CONF, 2018.]

Table of Contents

Graph Coarsening

Coarsening: From an initial graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{L}\}$ with N nodes and a signal $\mathbf{x}$, build a new coarsened graph $\mathcal{G}_c$ with N_c nodes:

$$\mathbf{x}_c = \mathbf{P}\mathbf{x}$$
$$\tilde{\mathbf{x}} = \mathbf{P}^{+}\mathbf{x}_c$$

where $\mathbf{P} \in \mathbb{R}^{N_c \times N}$ are matrices with more columns than rows and $\mathbf{P}^{+}$ the pseudo-inverse.

$$\mathbf{L}_c = \mathbf{P}^{\mp} \cdot \mathbf{L} \cdot \mathbf{P}^{+} \qquad \mathbf{L} \approx \mathbf{P}^{\top} \cdot \mathbf{L}_c \cdot \mathbf{P}$$

[A. Loukas and P. Vandergheynst, ''Spectrally approximating large graphs with smaller graphs,'' arXiv preprint arXiv:1802.07510, 2018]

Graph Coarsening

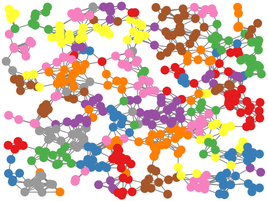
With the constraint that $\mathbf{L}_c$ is a graph Laplacian, we have:

$$\mathbf{P}(r,i) = \begin{cases} \frac{1}{\|\mathcal{V}^{(r)}\|} & \text{if } v_i \in \mathcal{V}^{(r)} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$\mathbf{P}^{+}(i,r) = \begin{cases} 1 & \text{if } v_i \in \mathcal{V}^{(r)} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Graph Coarsening

Results of the recursive coarsening available in
`https://github.com/loukasa/graph-coarsening`



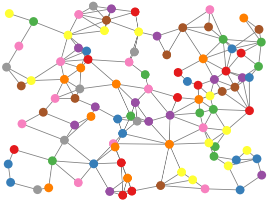
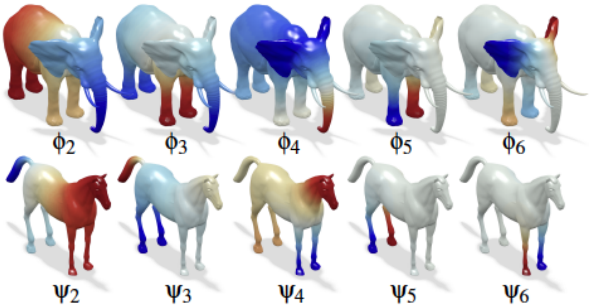


Table of Contents

Motivations

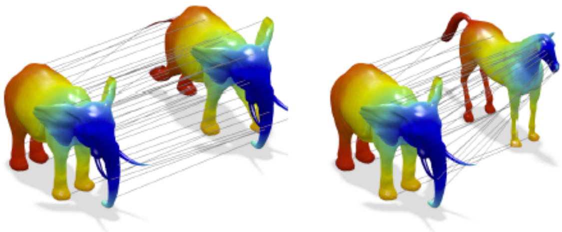
For different graph structure $\mathbf{L}_1$ and $\mathbf{L}_2$, the graph transform $\mathbf{\Phi}$ and $\mathbf{\Psi}$ may drastically vary



[Kovnatsky, A., Bronstein, M. M., Bronstein, A. M., Glashoff, K., and Kimmel, R. (2013, May). Coupled quasiâĂŘharmonic bases. In Computer Graphics Forum (Vol. 32, No. 2pt4, pp. 439-448). Oxford, UK: Blackwell Publishing Ltd.]

A priori step

Set correspondences



And register them in matrices $\mathbf{F}$ and $\mathbf{G}$

Joint Laplacian diagonalization

Solve the problem

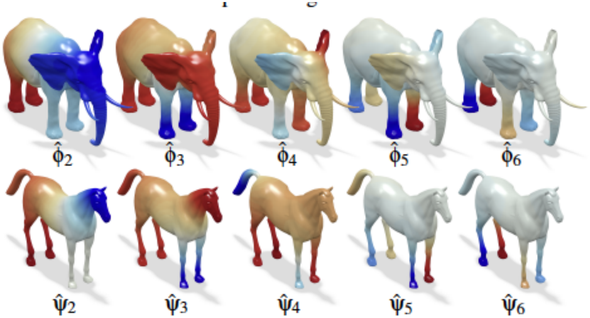
$$\min_{\hat{\mathbf{\Phi}},\hat{\mathbf{\Psi}}} = off(\hat{\mathbf{\Phi}}^\top \mathbf{L}_1 \hat{\mathbf{\Phi}}) + off(\hat{\mathbf{\Psi}}^\top \mathbf{L}_2 \hat{\mathbf{\Psi}}) + \mu||\mathbf{F}^\top \hat{\mathbf{\Phi}} - \mathbf{G}^\top \hat{\mathbf{\Psi}}||$$

$$\text{s.t.} \quad \hat{\mathbf{\Phi}}^\top \hat{\mathbf{\Phi}} = \mathbf{I} \text{ and } \hat{\mathbf{\Psi}}^\top \hat{\mathbf{\Psi}} = \mathbf{I}$$

Solved with

- Rotation and permutation of the original eigenvectors
- Gradient descent

Results


$\hat{\Phi}_2$
$\hat{\Phi}_3$
$\hat{\Phi}_4$
$\hat{\Phi}_5$
$\hat{\Phi}_6$
$\hat{\Psi}_2$
$\hat{\Psi}_3$
$\hat{\Psi}_4$
$\hat{\Psi}_5$
$\hat{\Psi}_6$

Results

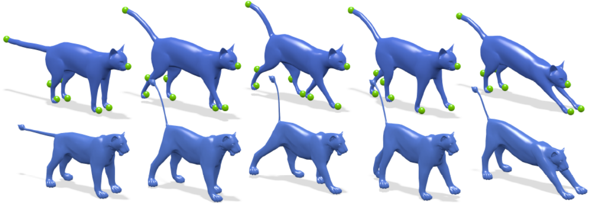
Moving object

Results

Two similar shapes

Results

Application for motion generation



Results

Texture transfer

[Monti, F., Boscaini, D., Masci, J., Rodola, E., Svoboda, J., and Bronstein, M. M. (2017). Geometric deep learning on graphs and manifolds using mixture model CNNs. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5115-5124).]

Focus of today: *Deep Sphere*

Abstract:

- What are the contributions ? Answering which challenge ?
- What are the state-of-the art methods ?
- What is the problem studied in the experiments ?
- What's the structure of the paper ?

[N. Perraudin, M. Defferrard, T. Kacprzak, R. Sgier, DeepSphere: Efficient spherical convolutional neural network with HEALPix sampling for cosmological applications, Astronomy and Computing, Volume 27, 2019, Pages 130-146, ISSN 2213-1337]

Focus of today: *Deep Sphere*

Introduction (paragraph 2):

- What are the three characteristics of a good convolution ?

Introduction (paragraph 3):

- Explain how a standard 2D CNN would be developed for spherical data analysis ?
- Why is it written that the obtained convolutions are not equivariant/invariant to rotation ?
- What's the limitation to spherical Fourier transforms ?

Introduction (paragraph 4):

- What are the main ideas of the proposed scheme ?

Focus of today: *Deep Sphere*

Method (paragraph 1):

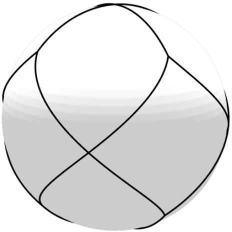
- Among all classical CNN operations, which ones need to be adapted to the sphere ?

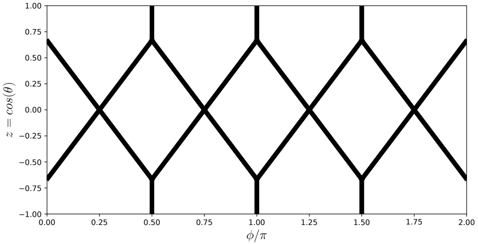
Method (paragraph 2):

- Why computing the convolution in the spectral domain is computationally inefficient ?
- What is author's proposition for that ?

HEALPix sampling

- The pixelization starts with partitioning the sphere into 12 equal-area regions (base resolution).

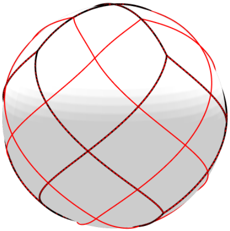


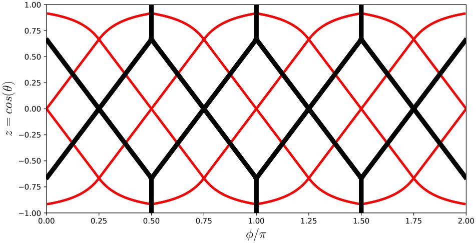


cylindrical projection

HEALPix sampling

- The pixelization starts with partitioning the sphere into 12 equal-area regions (base resolution).
- Finer pixelization is achieved by dividing each region into 4 equal-area regions.

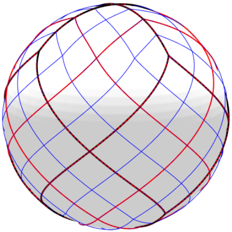


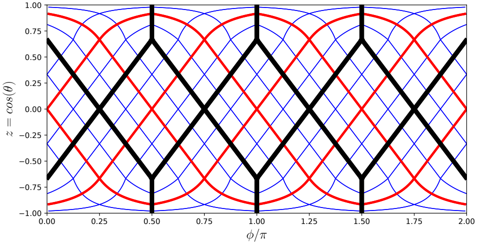


cylindrical projection

HEALPix sampling

- The pixelization starts with partitioning the sphere into 12 equal-area regions (base resolution).
- Finer pixelization is achieved by dividing each region into 4 equal-area regions.
- Hierarchical partitioning is repeated to reach the desired resolution.





cylindrical projection

All pixel centers are placed on rings of constant latitude.

Focus of today: *Deep Sphere*

2.2 Graph construction:

- What should be retained from this section ?
- Comment this figure:

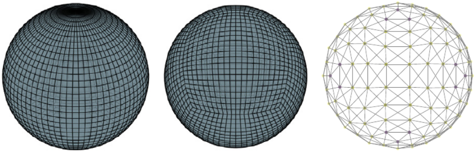


Fig. 6. Some pixelizations of the sphere. Left: the equirectangular grid, using equiangular spacing in a standard spherical-polar coordinate system. Middle: an equiangular cubed-sphere grid, as described in Ronchi et al. (1996). Right: graph built from a HEALPix pixelization of half the sphere ($N_{side} = 4$). By construction, each vertex has eight neighbors, except the highlighted ones which have only seven.[4]
Source: Left and middle figures are taken from Boomsma and Frellsen (2017)

Focus of today: *Deep Sphere*

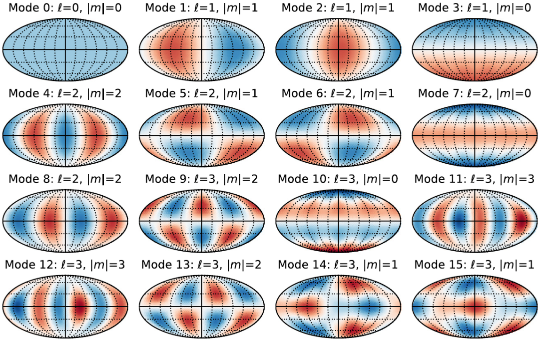
2.3 : Graph Fourier basis

- Retain that

$$\mathbf{L} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\top$$

- The transformed of a signal $\mathbf{f}$ is given by

$$\hat{\mathbf{f}} = \mathbf{U}^\top\mathbf{f}$$



2.4 : Convolution on graphs

- Explain

$$(h(\mathbf{L})\mathbf{f})_i = < h(\mathbf{L})\boldsymbol{\delta}_i, \mathbf{f} >$$

2.5 : Efficient convolutions

- (paragraph 1) What's the problem with (1) ?

$$h(\mathbf{L})\mathbf{f} = \mathbf{U}h(\boldsymbol{\Lambda})\mathbf{U}^\top\mathbf{f}$$

- (paragraph 2) Why is the proposed convolution less complex ?

$$h_\theta(\mathbf{f}) = \sum_{k=0}^{K} \theta_k \mathbf{L}^k \mathbf{f}$$

- (paragraph 2) What is the role of $\mathbf{L}^k$?
- (paragraph 2) In 2D convolution, it its 1 coefficient per pixel. How is it for the proposed convolution ?