Learning shape correspondence with anisotropic convolutional neural networks (ACNN)

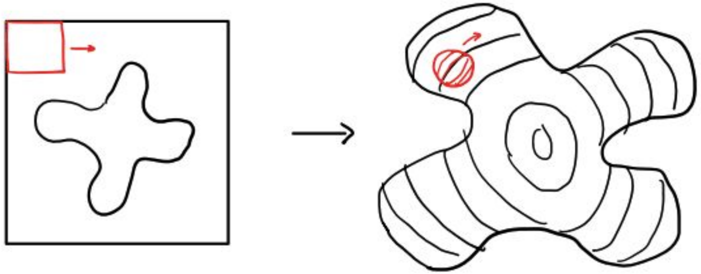
Louisa Cornelis
2/7/24

Agenda

- High Level Overview
- Background
- Architecture
- Related Works
- Results
- Code

High Level Overview

- Extending convolutional neural networks to 3D Geometric data with the goal of learning correspondence between 3D shapes
- Does this by defining an architecture based on anisotropic diffusion kernels, and generalizes convolutions to the non-Euclidean domain
- What is anisotropic: Directionally dependent! Does not behave the same way in every direction



What are we doing?

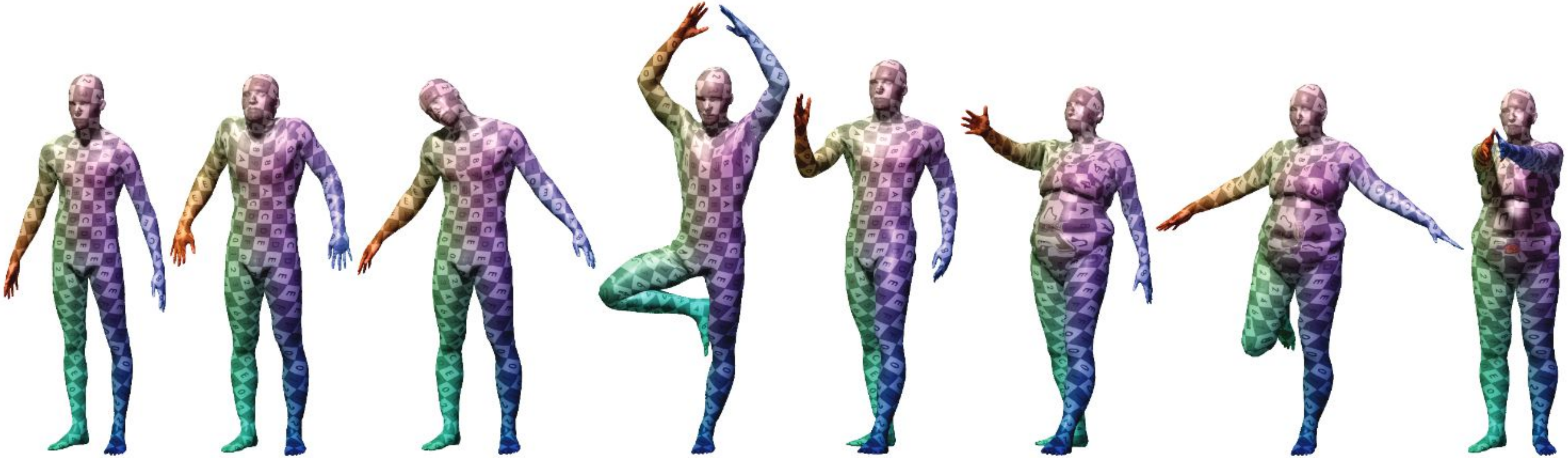
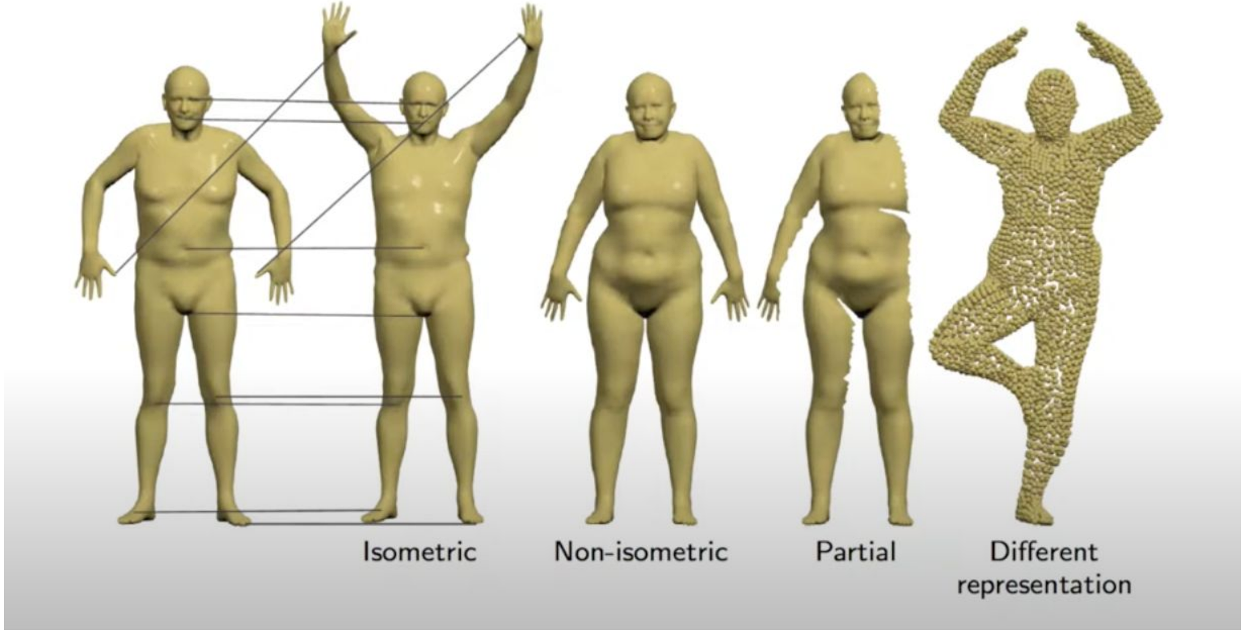


Figure 1: Examples of correspondence on the FAUST humans dataset obtained by the proposed ACNN method. Shown is the texture transferred from the leftmost reference shape to different subjects in different poses by means of our correspondence. The correspondence is nearly perfect (only very few minor artifacts are noticeable). See text for details.

Introduction

- Robust to non isometric deformations
- (div poses of same object that do not preserve geodesic distance),
- Non rigid deformations,
- Partiality
- Different representations of the same object such as:
    - Point clouds
    - Occlusions
    - Partial correspondence



Riemannian Manifolds

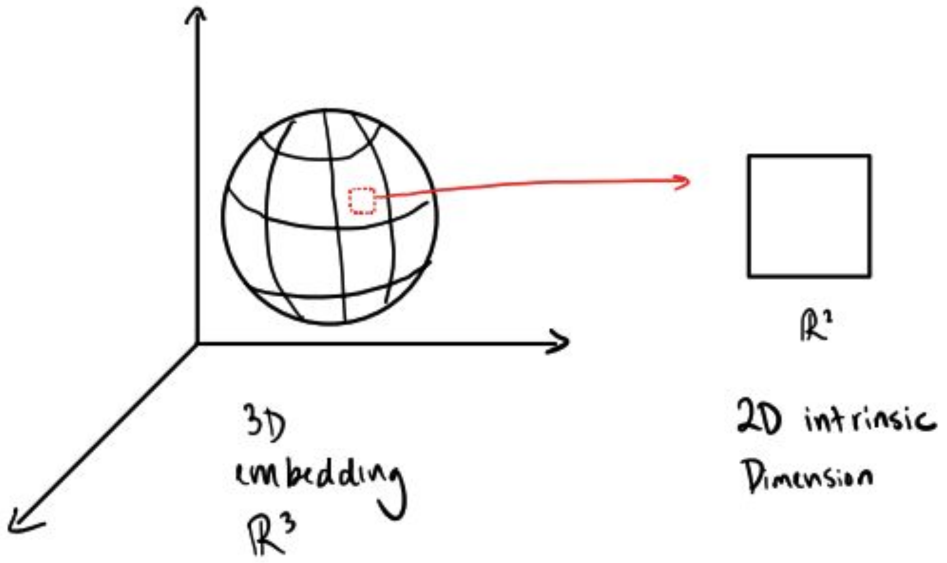
- A real, smooth manifold X equipped with a positive-definite inner product on the tangent space $T_x X$ at each point . x

$$\langle \cdot, \cdot \rangle_{T_x X} : T_x X \times T_x X \to \mathbb{R}$$

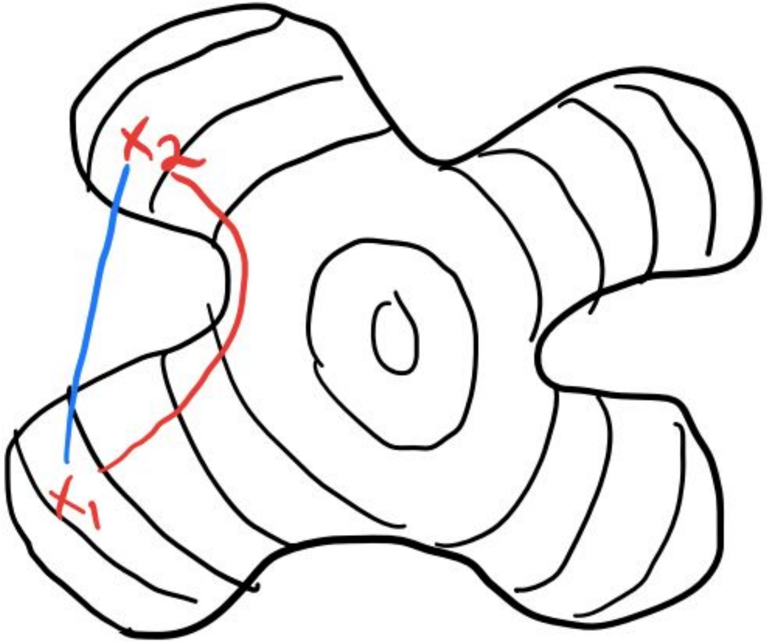
- This inner product is the **Riemannian metric** or **first fundamental form**
- This allows us to define **intrinsic quantities,** which are quantities which are expressible entirely in terms of Riemannian metric, and therefore independent on the way the surface is embedded.
- Intuitively - intrinsic quantities are quantities which exist on the surface of our manifold.

What do we mean by embedding?

- **The Whitney embedding theorem:** Any smooth manifold M of dimension m can be smoothly embedded in $\mathbb{R}^{2m+1}$.



Intrinsic Metrics



Example:
Extrinsic vs Intrinsic Distance between two points:

- Intrinsic stays on the manifold because must be in some direction of tangent space
- Extrinsic exists outside of the manifold

Riemannian Manifolds

- Working with non-Euclidean manifolds
- Prior work has attempted to do this but has treated geometric data as Euclidean structures
  - Can lose significant parts of the object or its fine details
  - Can break topological structure
  - **Euclidean representations are not intrinsic, and vary as the result of pose or deformation of the object.**

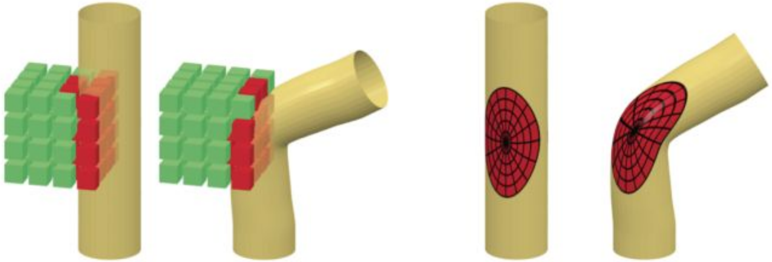


Figure 2: Illustration of the difference between extrinsic and intrinsic deep learning methods on geometric data. Left: extrinsic methods such as volumetric CNNs treat 3D geometric data in its Euclidean representation. Such a representation is not invariant to deformations (e.g., in the shown example, the filter that responds to features on a straight cylinder would not respond to a bent one). Right: in an intrinsic representation, the filter is applied to some data on the surface itself, thus being invariant to deformations.

Intrinsic Metrics

Intrinsic Gradient:

$$\nabla_X f(x) = \nabla(f \circ \exp_x)(\mathbf{0}),$$

Intrinsic Divergence:

$$\int_X \langle \nabla_X f(x), \mathbf{v}(x) \rangle_{T_x X} dx = -\int_X f(x) \mathrm{div}_X \mathbf{v}(x) dx,$$

Laplace-Beltrami Operator:

$$\Delta_X f(x) = -\mathrm{div}_X (\nabla_X f(x)).$$

Architecture

Define **anisotropic heat kernels** 🔥 that will be used to construct the intrinsic convolution.

General **isotropic form** of heat equation:

$$f(x,t) = H^t f_0(x) = \int_X f_0(\xi) h_t(x,\xi)\, d\xi \,,$$

- f(x,t): temperature at point x at time t.
- $f_0(x)$ initial heat distribution f(x,0)
- $H^t = e^{-t\Delta_X}$ to f_0: Heat operator
- $h_t(x,\xi)$ Is the heat kernel, heat transferred from x to ξ after time t.

Architecture

Anisotropic case: $f_t(x,t) = -\mathrm{div}_X(\mathbf{D}(x)\nabla_X f(x,t)),$

Where $\mathbf{D}(x)$ is the thermal conductivity tensor applied to the intrinsic gradient in the tangent plane. Models heat flow in a way that is position and direction-dependent (anisotropic!).

Specifically, use D such that anisotropic diffusion is driven by surface curvature:

$$\mathbf{D}_{\alpha\theta}(x) = \mathbf{R}_\theta(x)\begin{bmatrix}\alpha & \\ & 1\end{bmatrix}\mathbf{R}_\theta^\top(x),$$

Where $\mathbf{R}_\theta(x)$ rotates with respect to the maximum curvature direction and α controls degree of anisotropy , 1 is isotropic.

Architecture

Get anisotropic heat kernel using anisotropic Laplacian eigenvalues and eigenvectors:

$$\Delta_{\alpha\theta} f(x) = -\mathrm{div}_X (\mathbf{D}_{\alpha\theta}(x) \nabla_X f(x))$$

Where $\{\phi_{\alpha\theta i}, \lambda_{\alpha\theta i}\}_{i \geq 0}$ are its eigenvalues and eigenvectors.
Construct **anisotropic heat kernel** as:

$$h_{\alpha\theta t}(x, \xi) = \sum_{k \geq 0} e^{-t\lambda_{\alpha\theta k}} \phi_{\alpha\theta k}(x) \phi_{\alpha\theta k}(\xi).$$

Architecture

Using this, we can construct the patch operator - a local intrinsic polar representation of the data. These are **intrinsic convolutions!** 🎉

Anisotropic heat kernels act as local spatial weighting functions allowing us to extract local intrinsic representation of our function on the manifold. Define:

$$(D(x)f)(\theta,t) = \frac{\int_X h_{\alpha\theta t}(x,\xi) f(\xi) d\xi}{\int_X h_{\alpha\theta t}(x,\xi) d\xi},$$

- **Weighted average:** Heat equation for function f, over sum of all weights. t is scale of kernel, theta is orientation

Then construct intrinsic convolution:

$$(f * a)(x) = \int a(\theta,t)(D(x)f)(\theta,t) dt d\theta,$$

Where $a(\theta,t)$ are the learnable coefficients.

Intrinsic convolution layer

→ Replaces convolution layer used in classical Euclidean CNNs.

PQ filters (P filters in Q banks), Q banks corresponds to dimension of output. P corresponds to dimension of input.

$$f_q^{\text{out}}(x) = \sum_{p=1}^{P} (f_p^{\text{in}} \star a_{qp})(x), \quad q = 1, \ldots, Q,$$

Where $a_{qp}(\theta, t)$ are the learnable coefficients of the pth filter in the qth filter bank.

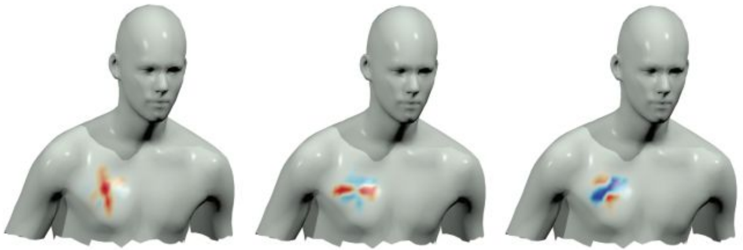


Figure 4: Examples of filters in the first IC layer learned by the ACNN (hot and cold colors represent positive and negative values, respectively).

Related Works

Extends prior work:

- **Geodesic CNN (GCNN)**
  - MASCI J., BOSCAINI D., BRONSTEIN M. M., VANDERGHEYNST P.: Shapenet: Convolutional neural networks on non euclidean manifolds. arXiv:1501.06297 (2015)
- **Localized Spectral CNN (LSCNN)**
  - BOSCAINI D., MASCI J., MELZI S., BRONSTEIN M. M., CASTELLANI U., VANDERGHEYNST P.: Learning class-specific descriptors for deformable shapes using localized spectral convolutional networks. CGF 34, 5 (2015), 13–23.
- **Anisotropic Diffusion Descriptors (ADD)**
  - [BMR∗16] BOSCAINI D., MASCI J., RODOLÀ E., BRONSTEIN M. M., CREMERS D.: Anisotropic diffusion descriptors. Computer Graphics Forum 35, 2 (2016).

Results

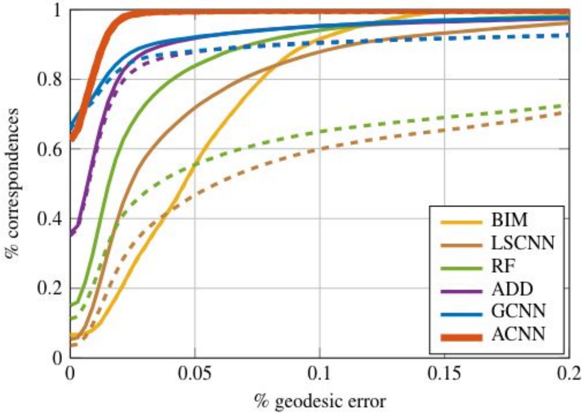


Figure 6: Performance of different correspondence methods on FAUST meshes. Evaluation of the correspondence was done using the symmetric (solid) and asymmetric (dashed) Princeton protocol.

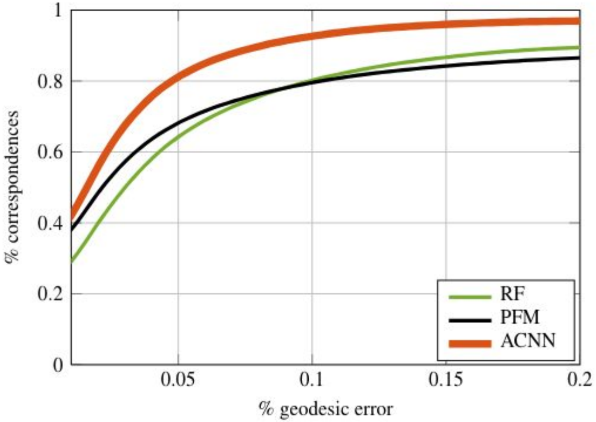


Figure 7: Performance of different correspondence methods on SHREC'16 Partial (cuts) meshes. Evaluation of the correspondence was done using the symmetric Princeton protocol.

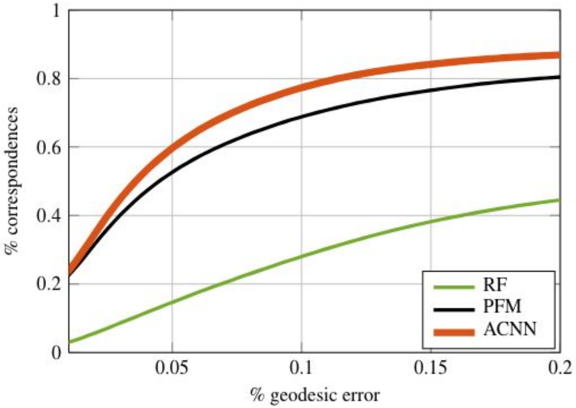
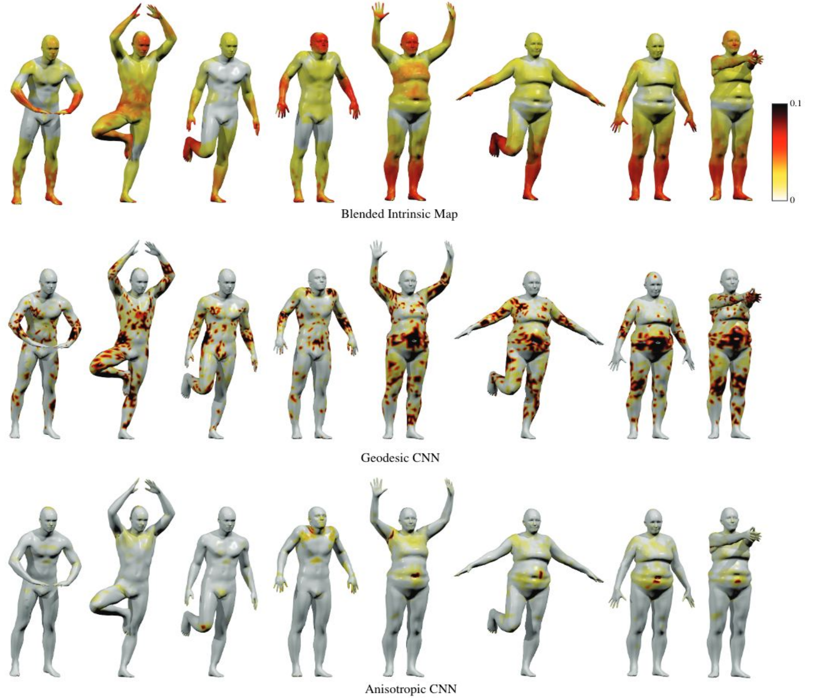


Figure 8: Performance of different correspondence methods on SHREC'16 Partial (holes) meshes. Evaluation of the correspondence was done using the symmetric Princeton protocol.

Results



Blended Intrinsic Map

Geodesic CNN

Anisotropic CNN

Figure 5: Pointwise geodesic error (in % of geodesic diameter) of different correspondence methods (top to bottom: Blended intrinsic maps, GCNN, ACNN) on the FAUST dataset. For visualization clarity, the error values are saturated at 10% of the geodesic diameter. Hot colors correspond to large errors. Note the different behavior of different approaches: BIM produces large distortions with very few accurate matches; GCNN produces many near-perfect matches but also many matches with large distortion; ACNN produces very few matches with large distortion and many near-perfect matches.

Results

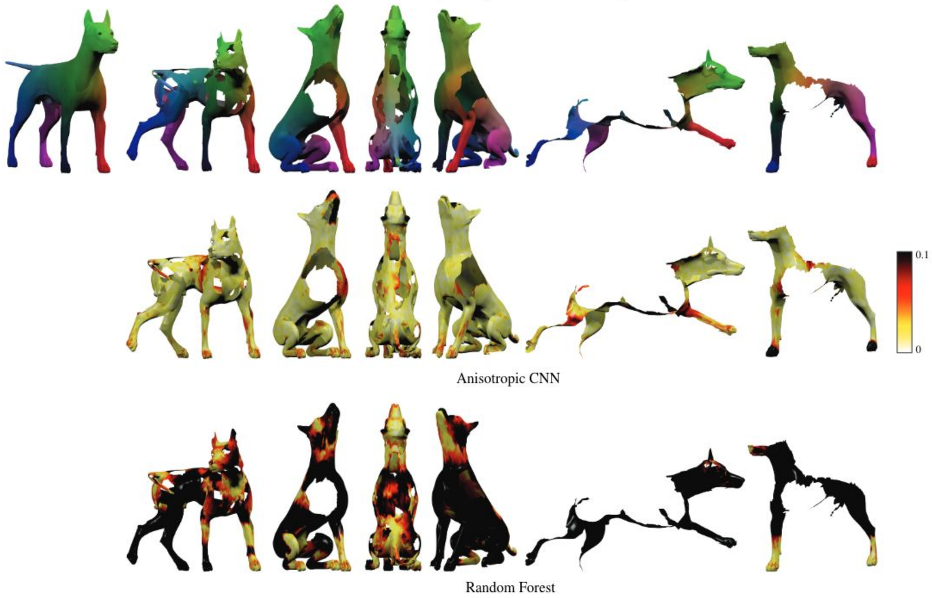


Figure 10: Examples of partial correspondence on the dog shape from the SHREC'16 Partial (holes) dataset. First row: correspondence produced by ACNN. Corresponding points are shown in similar color. Reference shape is shown on the left. Second and third rows: pointwise geodesic error (in % of geodesic diameter) of the ACNN and RF correspondence, respectively. For visualization clarity, the error values are saturated at 10% of the geodesic diameter. Hot colors correspond to large errors.

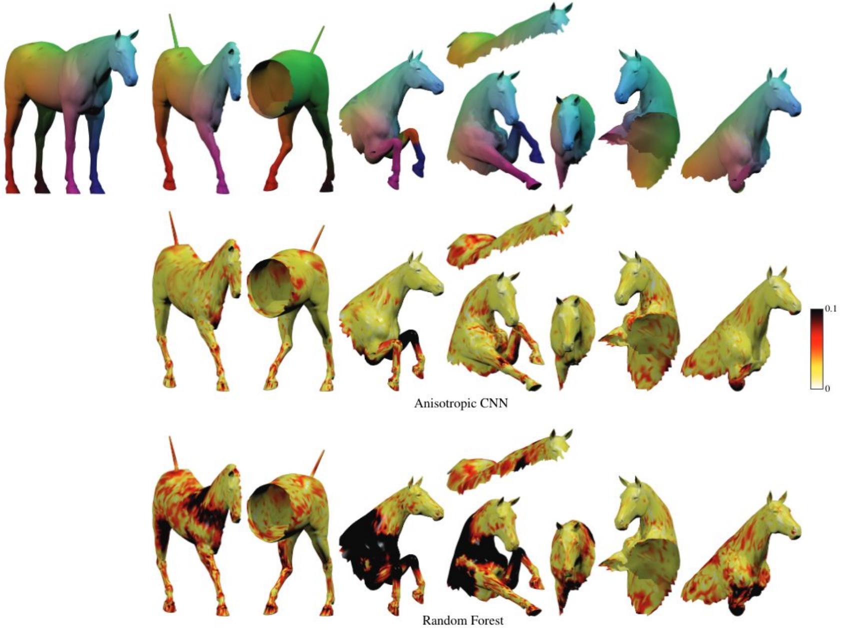


Figure 9: Examples of partial correspondence on the horse shape from the SHREC'16 Partial (cuts) dataset. First row: correspondence produced by ACNN. Corresponding points are shown in similar color. Reference shape is shown on the left. Second and third rows: pointwise geodesic error (in % of geodesic diameter) of the ACNN and RF correspondence, respectively. For visualization clarity, the error values are saturated at 10% of the geodesic diameter. Hot colors correspond to large errors.

References

Boscaini, Davide, et al. "Learning shape correspondence with anisotropic convolutional neural networks." Advances in neural information processing systems 29 (2016).

Boscaini, Davide. "Spotlight video for the paper "Learning shape correspondence with anisotropic CNNs."" YouTube, uploaded by Davide Boscaini, 21 November 2016, https://www.youtube.com/watch?v=tXT41yZTCMA .

Tubbenhauer, Daniel. "What is...Whitney's embedding theorem?" YouTube, uploaded by VisualMath, 3 September 2022, https://www.youtube.com/watch?v=KWAOgpDI8mA