中国科学院自动化研究所
Institute of Automation, Chinese Academy of Sciences

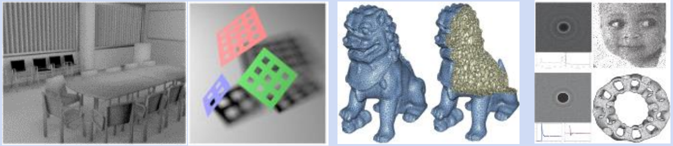
3D Descriptor Design and Learning for Robust Non-rigid Shape Matching

Jianwei Guo 郭建伟

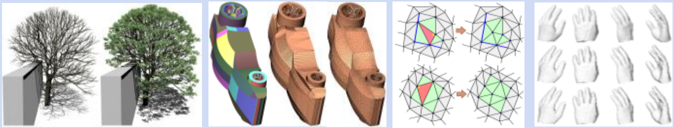
NLPR, Institute of Automation,
Chinese Academy of Sciences

Sept.17, 2020
GAMES Webinar

Team--3D Visual Computing



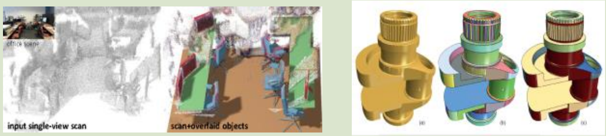
[ACM TOG (SigAsia) 2016] [CAGD 2016] [TVCG 2017]



[CGF 2018] [CAD 2019] [TVCG 2018] [CGF 2019]

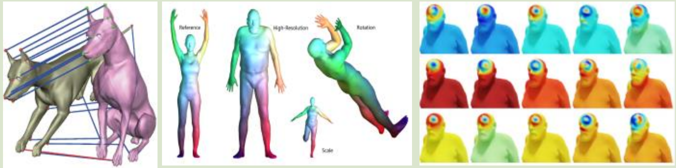
[TVCG 2020] [TVCG 2020] [ACM TOG 2020]

Modeling & Mesh Optimization



[ACM TOG (SIGGRAPH Asia) 2012] [GMP 2006, CAD 2012, JCAD 2018]

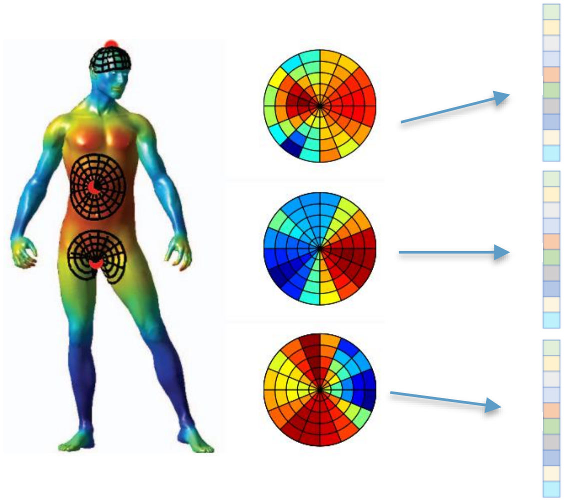
[CAGD 2018] [IEEE TIFS 2018] [CAD 2020]



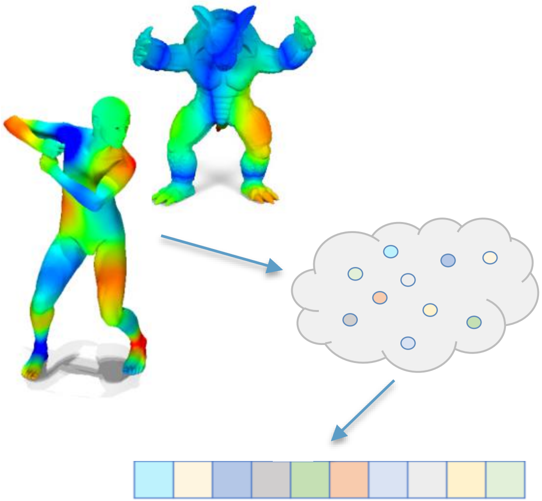
[ECCV 2018] [CVPR 2019] [ACM TOG (SIGGRAPH) 2020]

Shape Analysis

3D Shape Descriptor

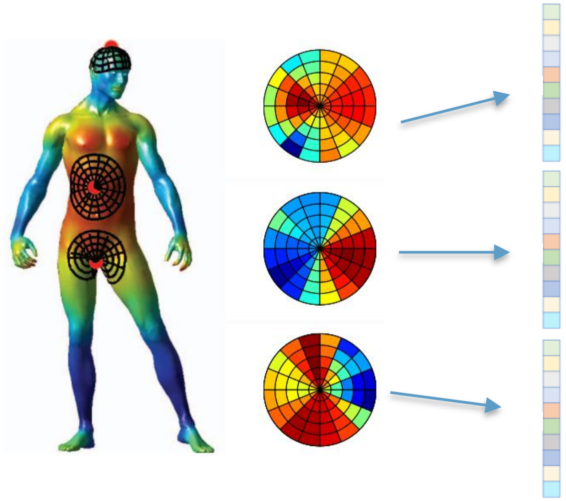


Local descriptor



Global descriptor

3D Shape Descriptor



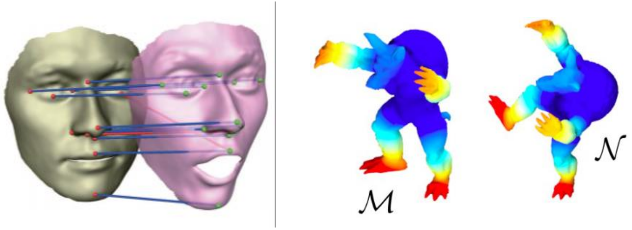
Descriptor design goals:

- **Discriminative:** able to determine if a pair of vertices is similar or different
- **Robust:** work with different discretizations of a surface

Local descriptor

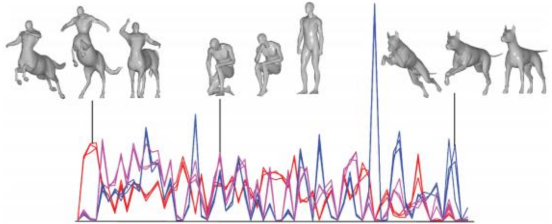
Applications

Shape matching/correspondence



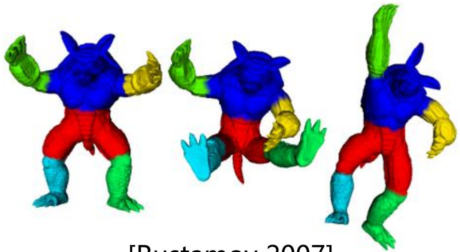
[Ovsjanikov et al. 2017;Wang et al. 2018]

Shape retrieval



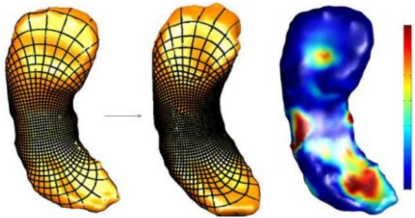
[Ovsjanikov et al. 2009]

Shape segmentation



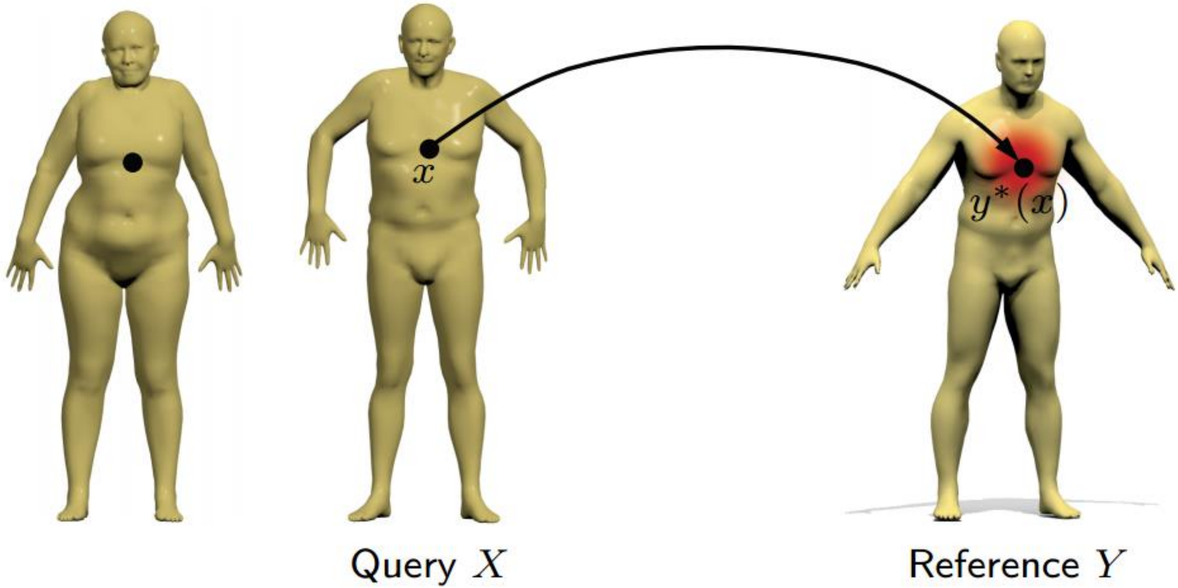
[Rustamov 2007]

Surface registration

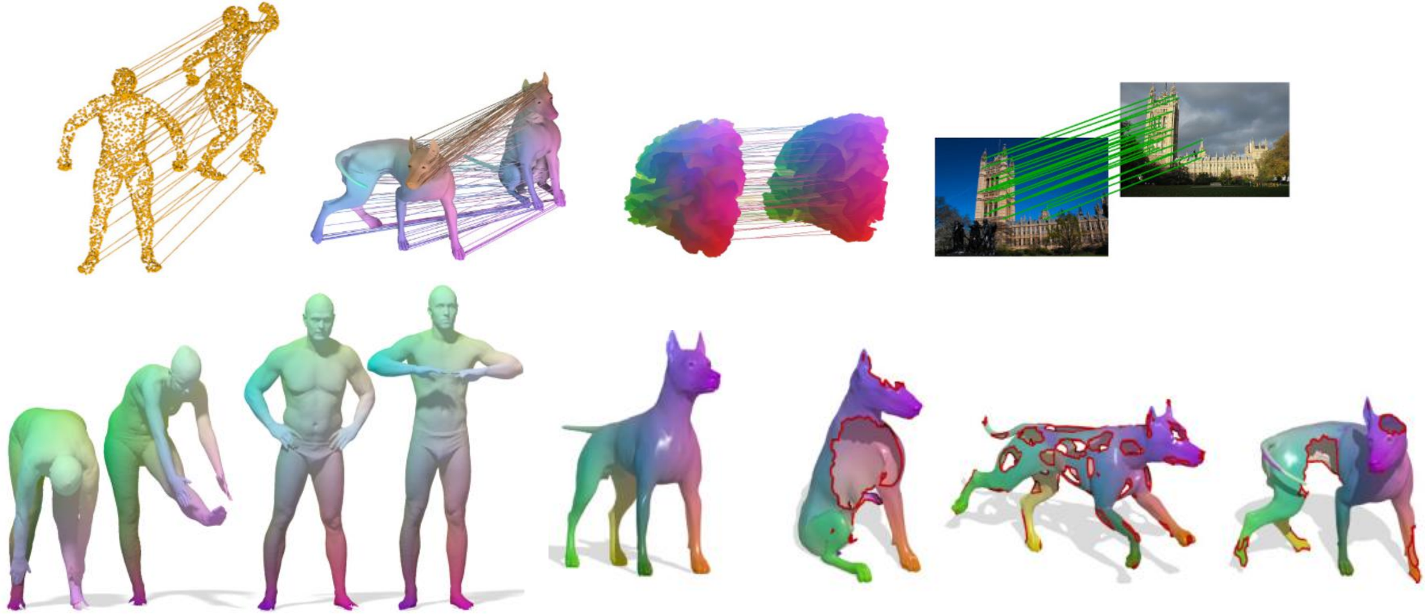


[Lui et al. 2010]

Shape matching/correspondence



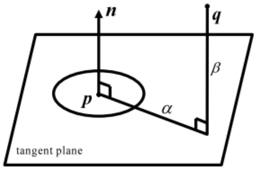
Shape matching/correspondence



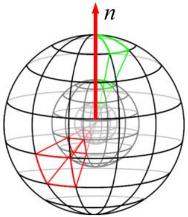
Van Kaick et al.: *A survey on shape correspondence*. CGF, 2011.
Ovsjanikov et al.: *Computing and processing correspondences with functional maps*. SIGGRAPH ASIA 2016 Courses.

Related Work

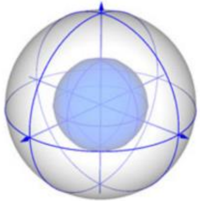
❖ **Spatial domain approaches**



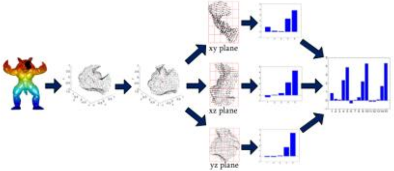
SI [Johnson and Hebert 1999]



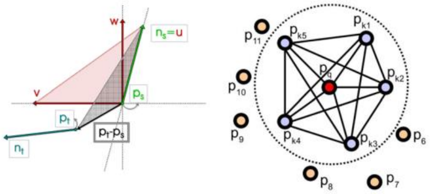
3DSC [Frome et al. 2004]



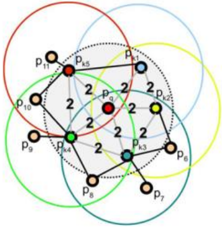
SHOT [Tombari et al. 2010]



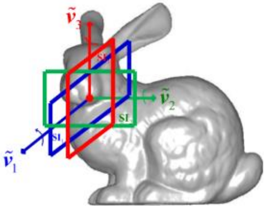
RoPS [Guo et al. 2013]



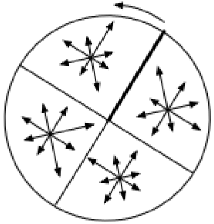
PFH [Rusu et al. 2008]



FPFH [Rusu et al. 2009]



TriSI [Guo et al. 2015]

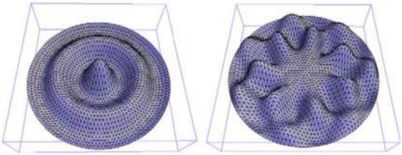


Mesh-HOG [Zaharescu et al. 2009]

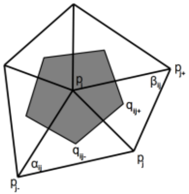
Guo et al.: *A comprehensive performance evaluation of 3D local feature descriptors*. IJCV, 2016.

Related Work

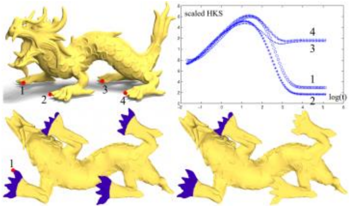
❖ Spectral domain approaches



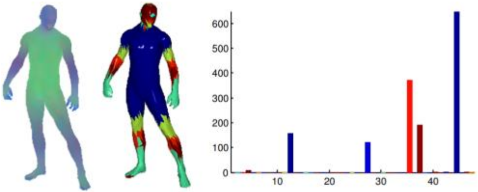
Shape-DNA [1999Reuter et al. 2006]



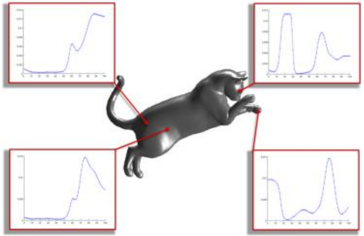
GPS [Rustamov 2007]



HKS [Sun et al. 2010]



Scale-invariant HKS [Bronstein and Kokkinos 2010]

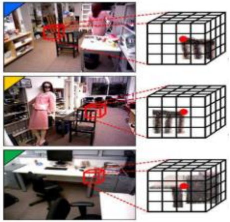


WKS [Aubry et al. 2011]

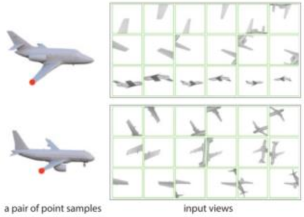
DTEP [Melziet al. 2018]

Related Work

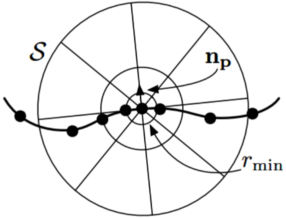
❖ Deep learning approaches



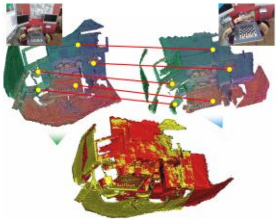
[Zeng et al. 2017]



[Huang et al. 2018]

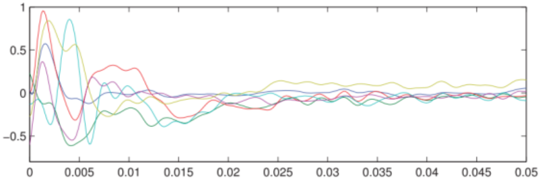


CGF [Khoury et al. 2017]

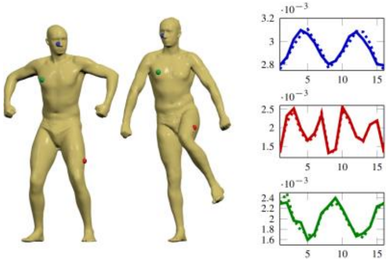


PPFNet [Deng et al. 2018]

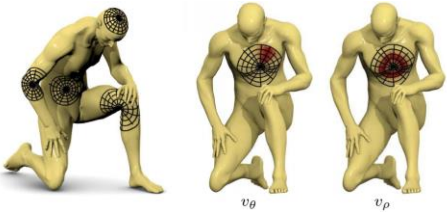
- Mesh-aware



OSD [Litman and Bronstein 2014]



[Boscaini et al. 2015, 2016]



[Masci et al. 2018]

Related Work

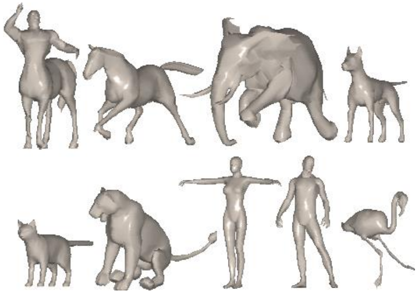
❖ Non-learned descriptors

- \+ **pros**: : WKS and HKS are robust
- \- **cons** : not as discriminative as supervised descriptors

❖ Supervised descriptors

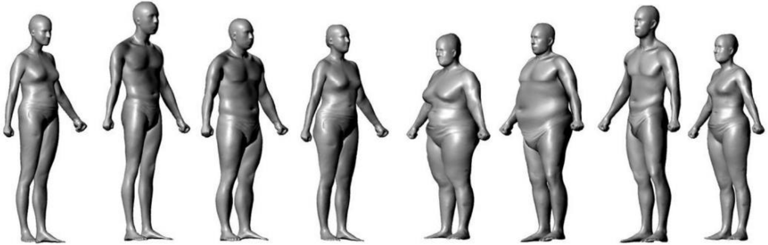
- \+ **pros:** GNNs compute discriminative descriptors
- \- **cons** : less robust to different surface discretizations

Datasets



TOSCA [Bronstein et al. 2008]

SCAPE [Anguelov et al. 2005]



SPRING [Yang et al. 2014]

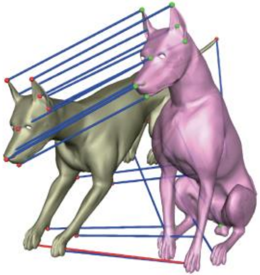
FAUST [Bogo et al. 2014]

Our Work

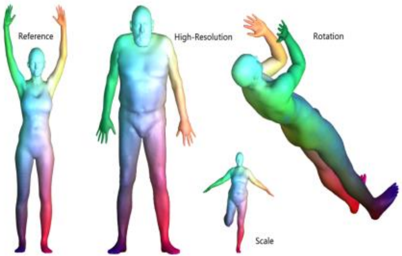
❖ Contributions

— Two new non-learned features, namely, *Local Point Signature* (LPS) and *Wavelet Energy Decomposition Signature* (WEDS). They exhibit high resilience to changes in mesh resolution, triangulation, scale, and rotation.

— Two supervised frameworks to transform the non-learned features to more discriminative descriptors



[ECCV 2018]



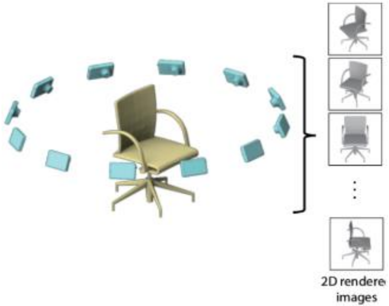
[CVPR 2019]

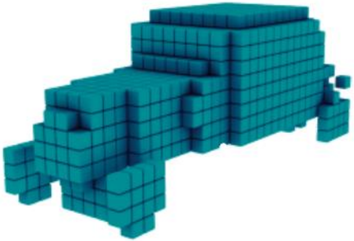
[CVMJ 2020]

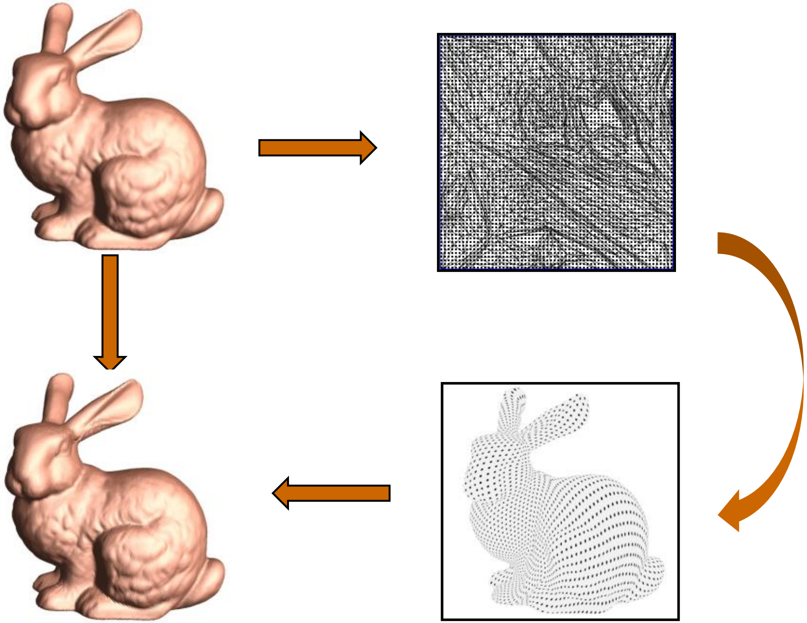
[ACM TOG (SIGGRAPH) 2020]

1-Descriptor Learning using Geometry Images

❖ Motivation



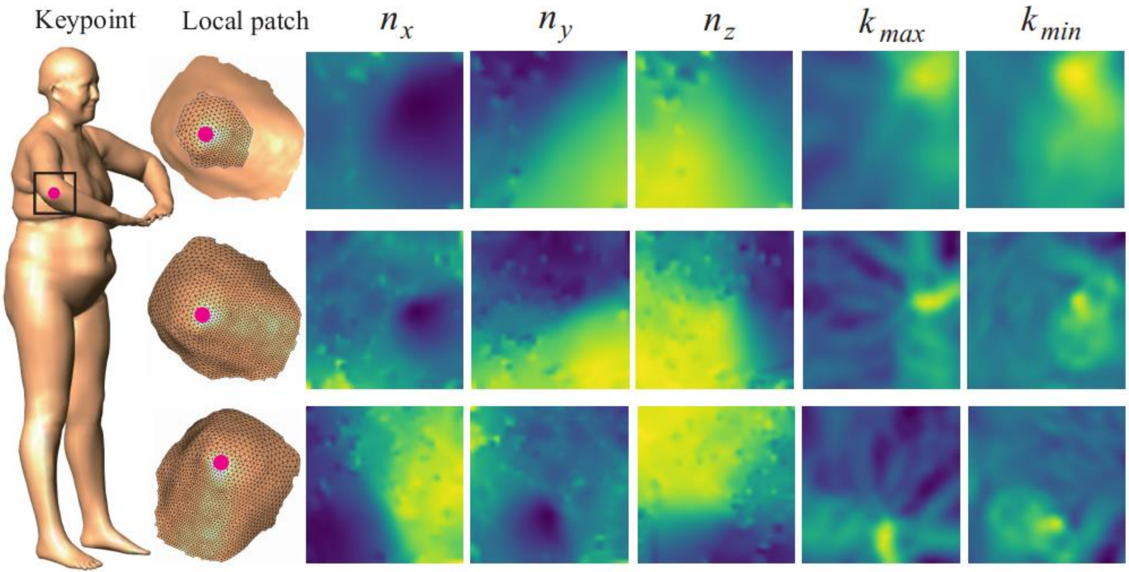




Geometry images [Gu et al. 2002]

1-Descriptor Learning using Geometry Images

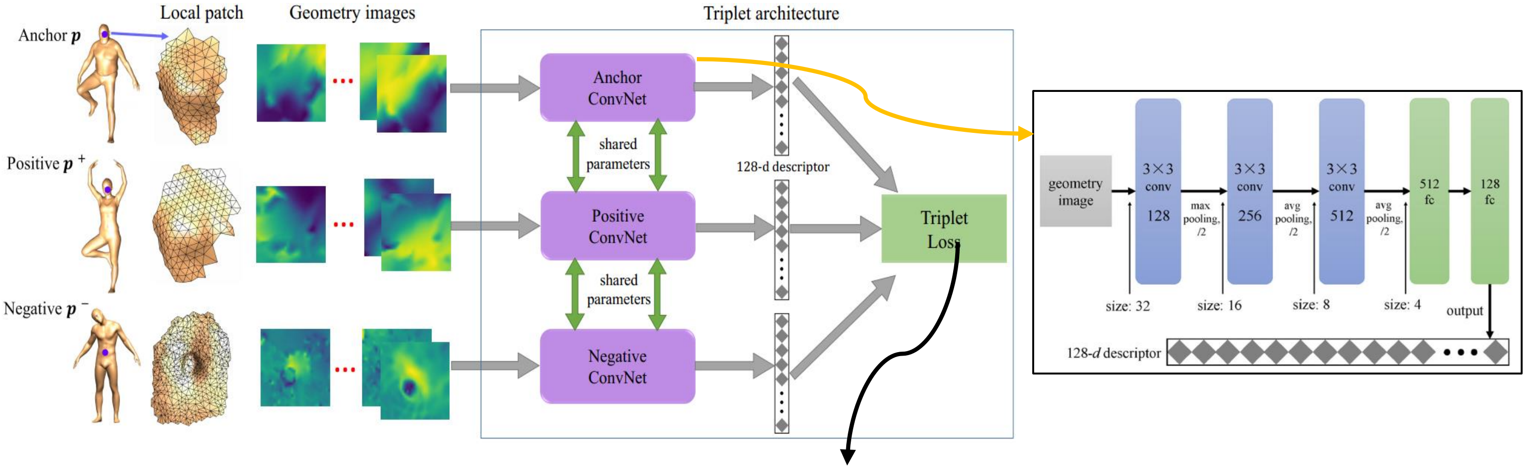
❖ Motivation



Desbrun et al.: *Intrinsic parameterizations of surface meshes*. CGF, 2002

1-Descriptor Learning using Geometry Images

❖ Triplet Neural Network



$$L_{Min-CV} = \lambda \frac{\sigma(D_{pos})}{\mu(D_{pos})} + \sum_{i=1}^{N} \left[D_{pos}^{i} - D_{neg}^{i} + \alpha \right]_{+}$$

1-Descriptor Learning using Geometry Images

Triplet Loss

- Classic triplet loss:

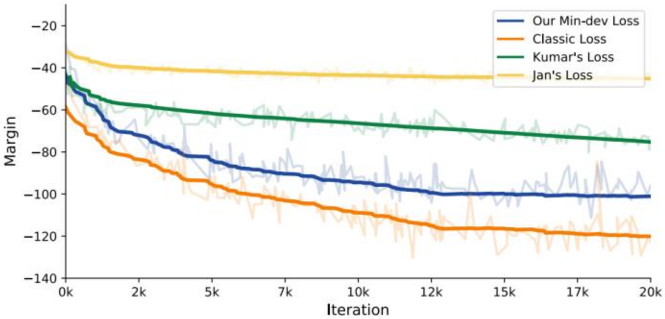
$$L = \sum_{i=1}^{N} [D_{\text{pos}}^i - D_{\text{neg}}^i + \alpha]_+$$

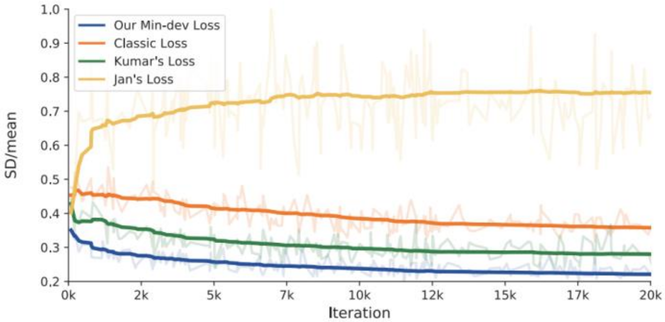
$$D_{\text{pos}}^i = D(f(\boldsymbol{p}_i), f(\boldsymbol{p}_i^+))$$

$$D_{\text{neg}}^i = D(f(\boldsymbol{p}_i), f(\boldsymbol{p}_i^-))$$

- Min-Coefficient of Variation (Min-CV) loss:

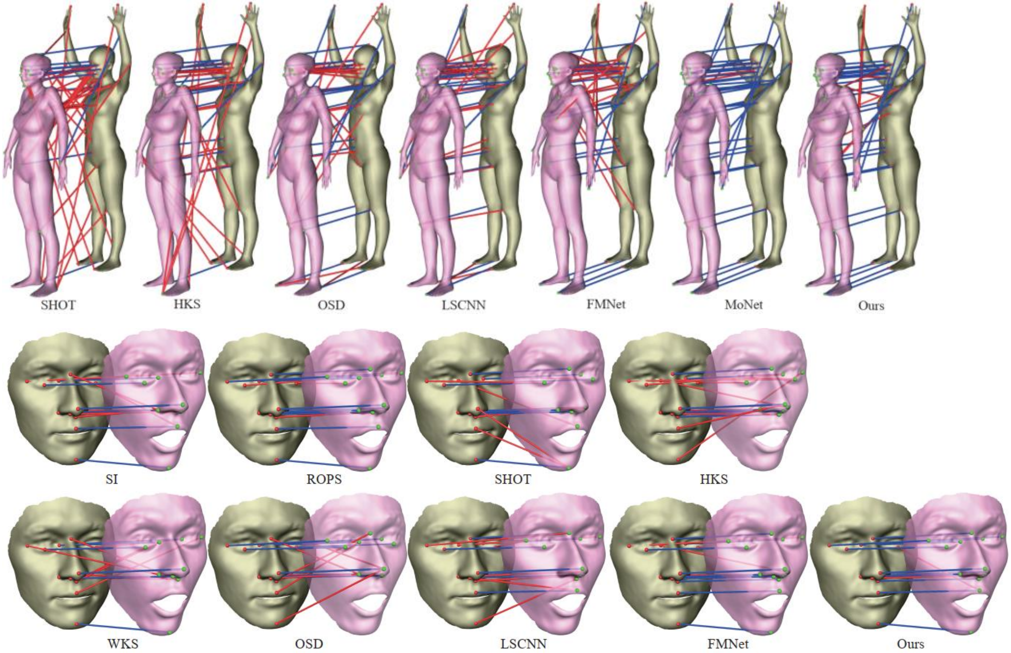
$$L_{Min-CV} = \lambda \frac{\sigma(D_{pos})}{\mu(D_{pos})} + \sum_{i=1}^{N} \left[D_{pos}^i - D_{neg}^i + \alpha\right]_+$$





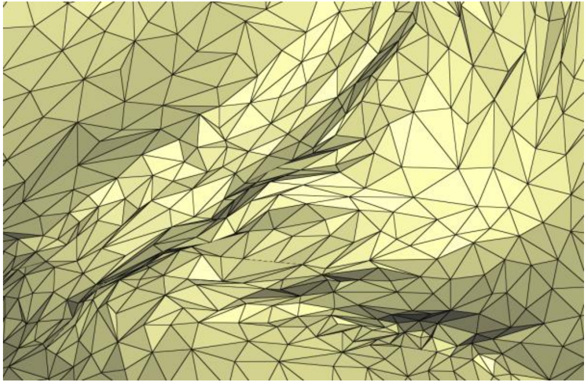
1-Descriptor Learning using Geometry Images

- **Results: shape matching**

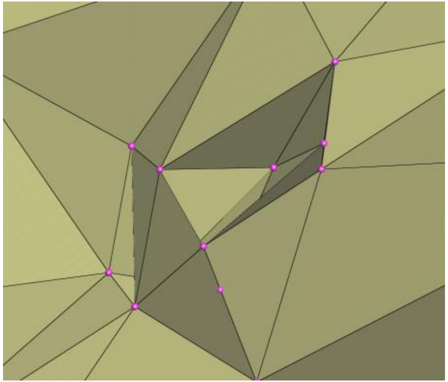


1-Descriptor Learning using Geometry Images

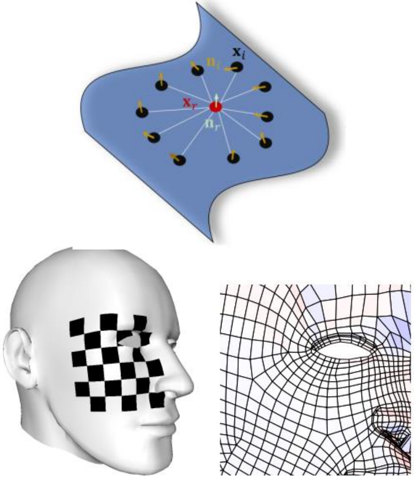
❖ Discrete Geodesic Polar Coordinates (DGPC)



ill-shaped triangles



degenerate triangles



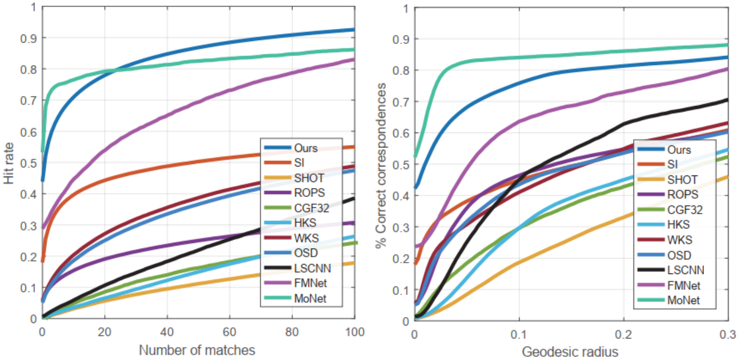
DGPC [*Melvær and Reimers 2012*]

Melvær and Reimers. *Geodesic polar coordinates on polygonal meshes*. CGF, 2012

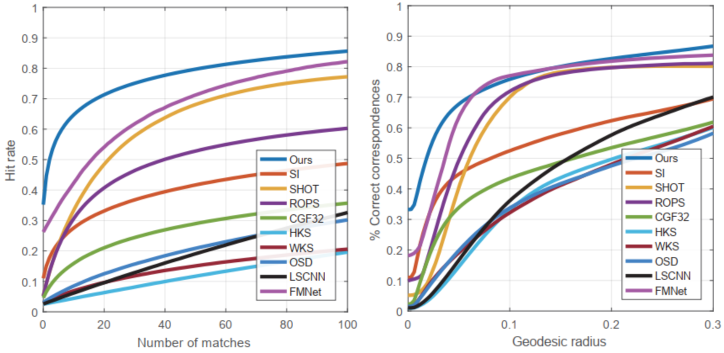
1-Descriptor Learning using Geometry Images

❖ Results：shape correspondence

- Cumulative match characteristic (CMC)
- Cumulative geodesic error（CGE）



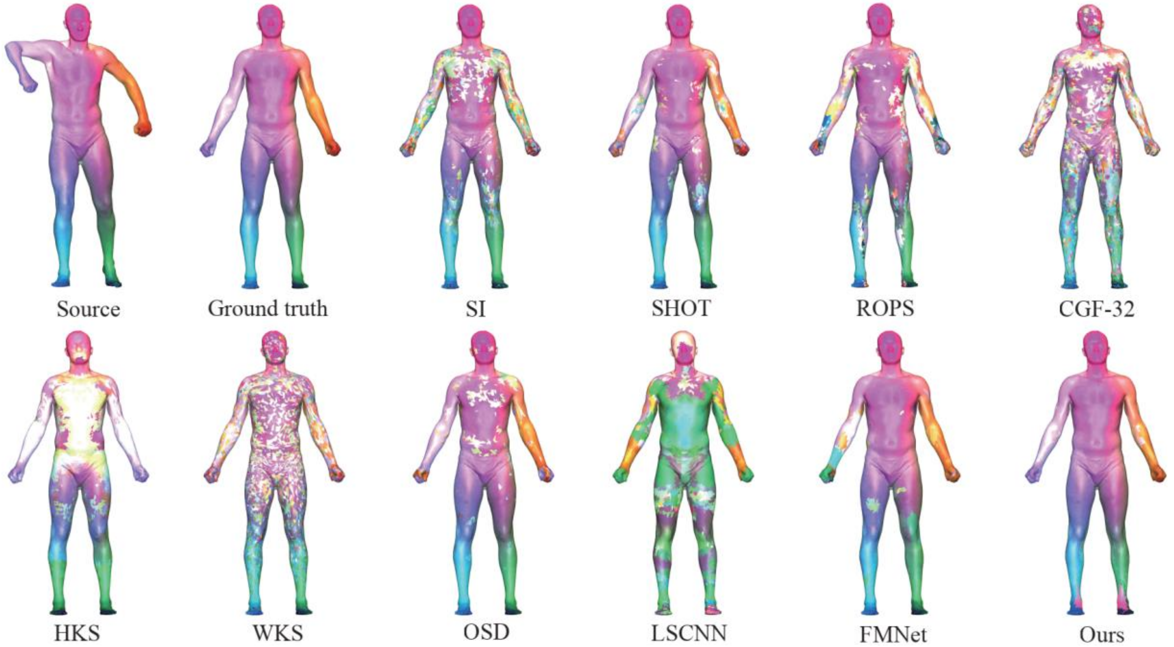
FAUST



SPRING

1-Descriptor Learning using Geometry Images

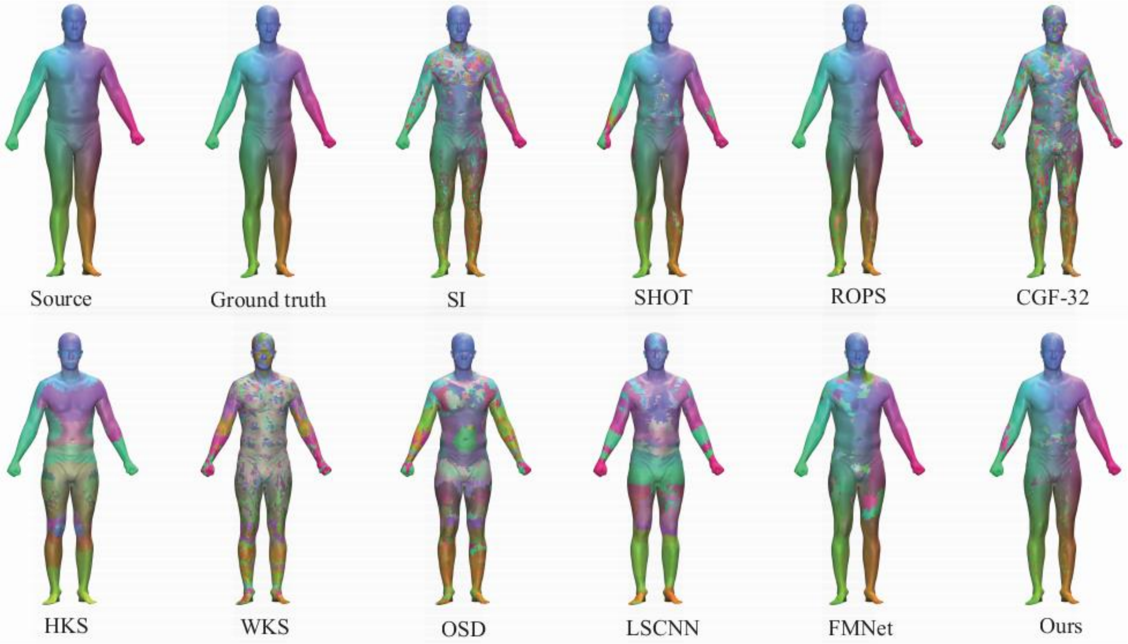
❖ **Results：shape correspondence**



SCAPE

1-Descriptor Learning using Geometry Images

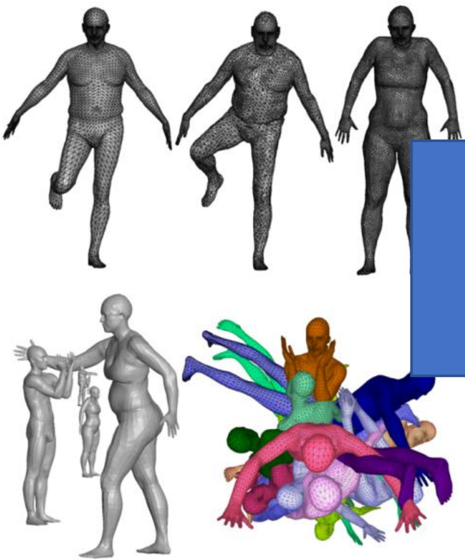
❖ **Results：shape correspondence**



FAUST

2-Robust Local Spectral Descriptor

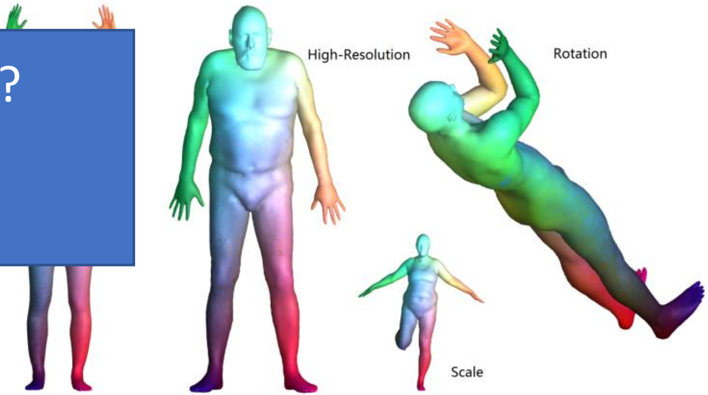
❖ Motivation



Resolutions, triangulations, scales, rotations

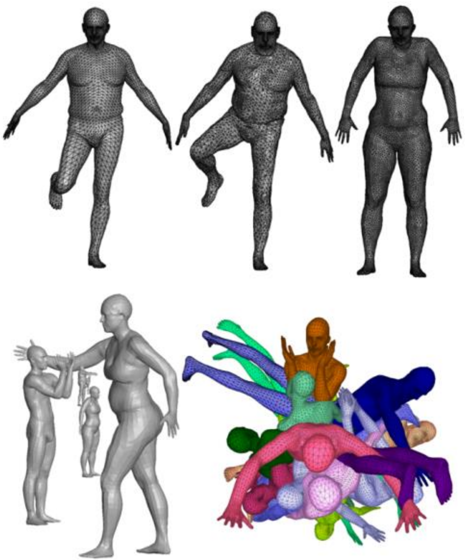
Approach?

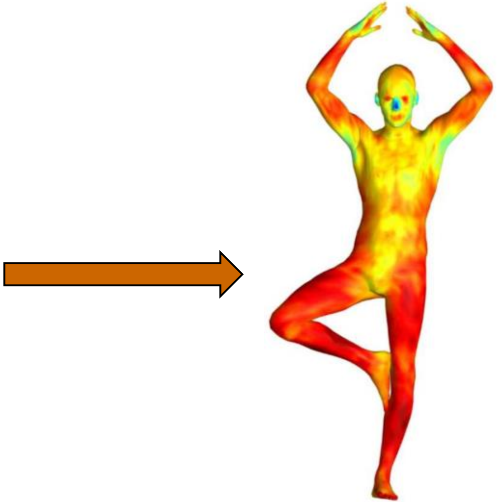
Dataset?



2-Robust Local Spectral Descriptor

❖ **Motivation**





Dirichlet energy

$$f : S \rightarrow R$$

$$E(f) = \int_S \left| \nabla f(v) \right|^2 dv = \int_S f(v) \Delta f(v) dv$$

2-Robust Local Spectral Descriptor

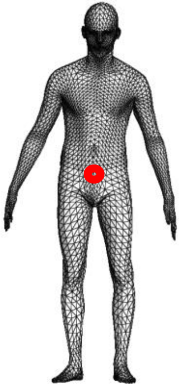
❖ Local Point Signature

$$E(f)=\int_S |\nabla f(v)|^2 dv = \int_S f(v)\Delta f(v)dv$$

$$\Delta f = -\mathrm{div}(\nabla f)$$

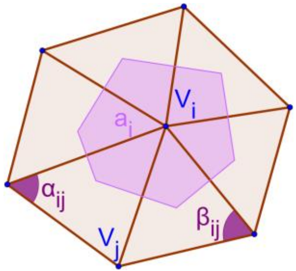
Laplace–Beltrami operator

$$\mathbf{L}_{ij}=\begin{cases} -\dfrac{\cot\alpha_{ij}+\cot\beta_{ij}}{2a_i} & \text{if } i, j \text{ are adjacent} \\ \sum_k \dfrac{\cot\alpha_{ik}+\cot\beta_{ik}}{2a_i} & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}$$



Mesh M

$\tilde{f}: V \rightarrow R$



$$\Delta\tilde{f}(v_i)=\frac{1}{a_i}\sum_{j\in N(v_i)}\frac{\cot\alpha_{ij}+\cot\beta_{ij}}{2}\left(\tilde{f}(v_i)-\tilde{f}(v_j)\right)$$

$$\mathbf{L}\Phi_i=\lambda_i\Phi_i, i=0,1,\ldots,k-1 \qquad \{\lambda_i \mid i=0,1,\ldots,k-1\}$$

2-Robust Local Spectral Descriptor

Local Point Signature

$$\mathbf{T}\Phi_i = \lambda_i \mathbf{A}\Phi_i, i = 0,1,\ldots,k-1 \qquad \langle \Phi_i, \Phi_j \rangle_{\mathbf{A}} = \Phi_i^T \mathbf{A}\Phi_j$$

$$\tilde{E}(\tilde{f}) = \tilde{f}^{\mathrm{T}} \mathbf{A}\mathbf{L}\tilde{f} = \sum_{j=0}^{N-1} \sigma_j^2 \lambda_j \qquad \sigma_j = \langle \tilde{f}, \Phi_j \rangle_{\mathbf{A}} = \tilde{f}^T \mathbf{A}\Phi_j$$

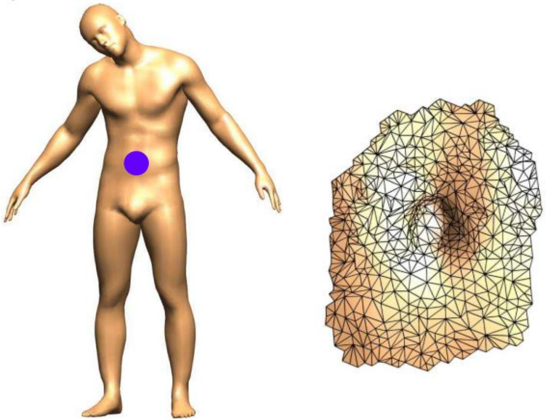
Spectral coefficients

$$\tilde{E}(\tilde{F}) = \sum_{i=1}^{d} \sum_{j=0}^{N-1} \sigma_{ij}^2 \lambda_j = \sum_{j=0}^{N-1} \lambda_j \sum_{i=1}^{d} \sigma_{ij}^2 \qquad \tilde{F} = (\tilde{f}_1, \tilde{f}_2, \ldots, \tilde{f}_d) : V \to R^d$$

$$sf = \left(\lambda_1 \sum_{i=1}^{d} \sigma_{i1}^2, \lambda_2 \sum_{i=1}^{d} \sigma_{i2}^2, \ldots, \lambda_{N-1} \sum_{i=1}^{d} \sigma_{iN-1}^2 \right)$$

2-Robust Local Spectral Descriptor

❖ Local Point Signature



- Build a local patch mesh around a vertex

- Compute Laplacian eigenvectors and eigenvalues

- Compute spectral coefficients

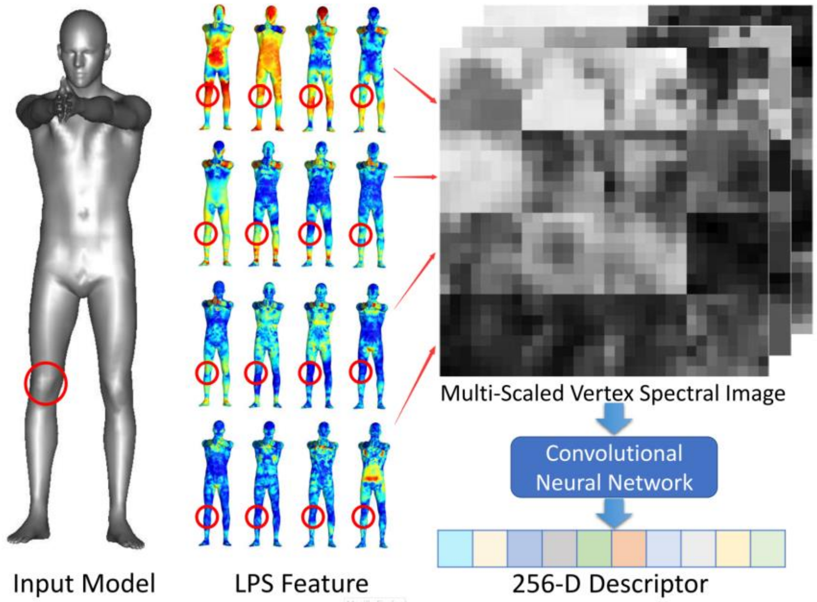
$$\sigma_j = \left\langle \tilde{f}, \Phi_j \right\rangle_{\mathbf{A}} = \tilde{f}^T \mathbf{A} \Phi_j \qquad X = (x_1, x_2, x_3) : V \to R^3$$

- LPS is derived from Dirichlet energy and expressed as follows

$$LPS = \left(\lambda_1 \sqrt{\sum_{i=1}^{3} \sigma_{i1}^2}, \lambda_2 \sqrt{\sum_{i=1}^{3} \sigma_{i2}^2}, \ldots, \lambda_{16} \sqrt{\sum_{i=1}^{3} \sigma_{i16}^2} \right)$$

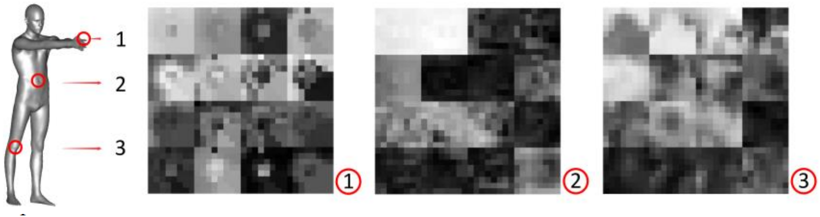
2-Robust Local Spectral Descriptor

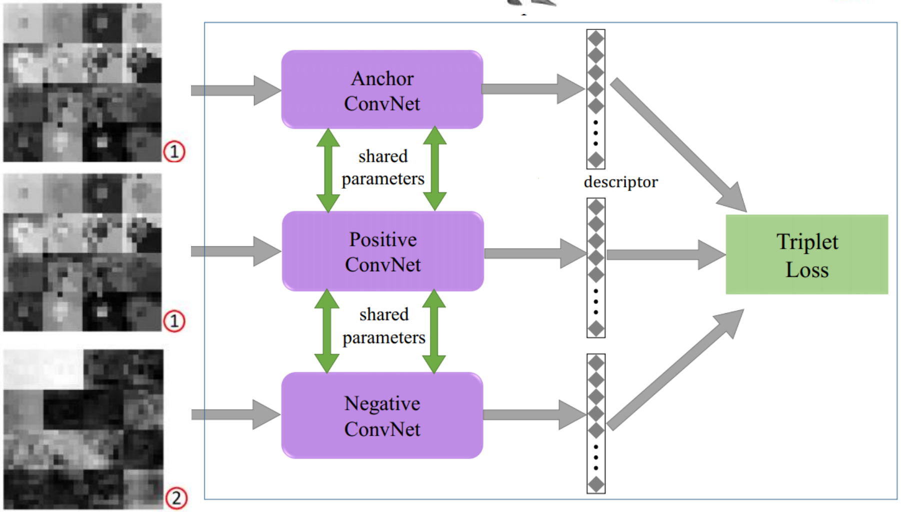
❖ **Descriptor learning**



2-Robust Local Spectral Descriptor

- Descriptor learning

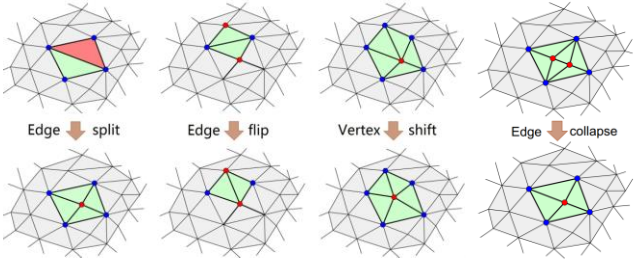


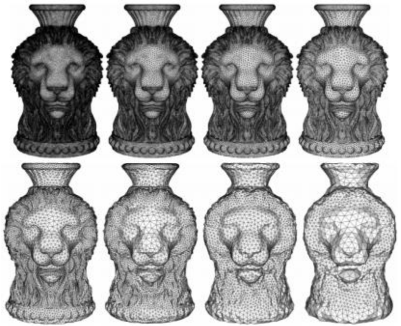


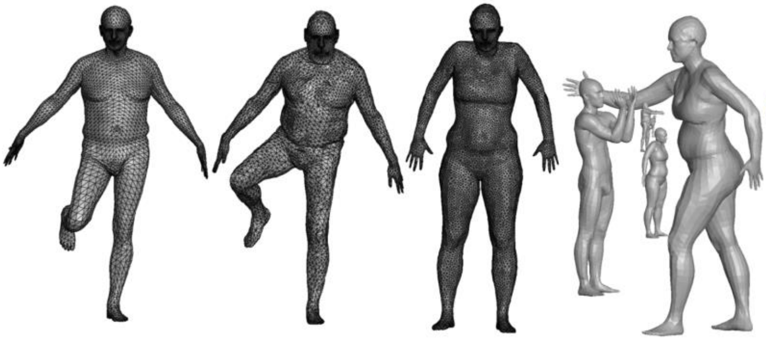
2-Robust Local Spectral Descriptor

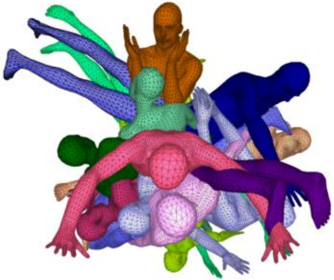
❖ Dataset

Discrete optimization method (Wang et al. 2019)



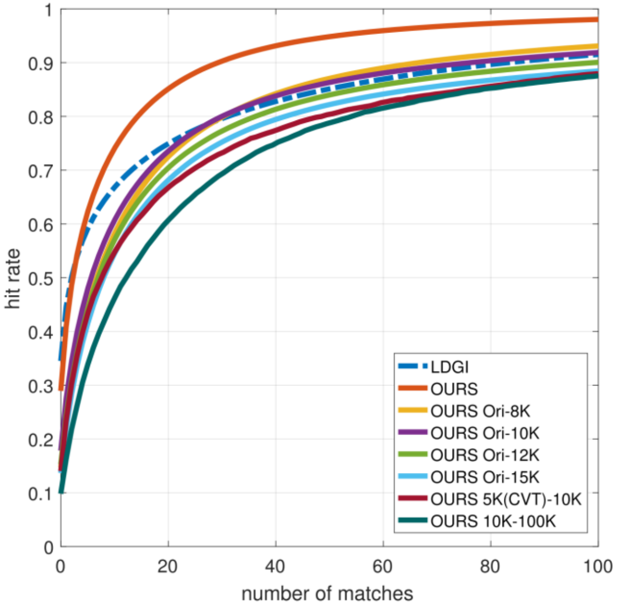


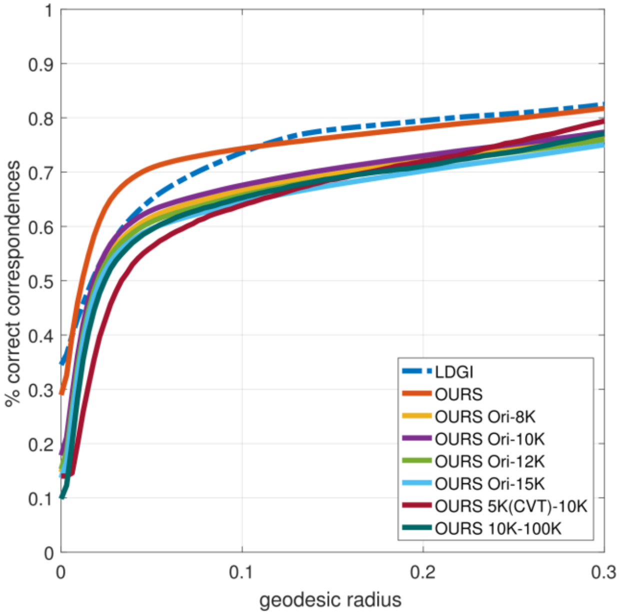




2-Robust Local Spectral Descriptor

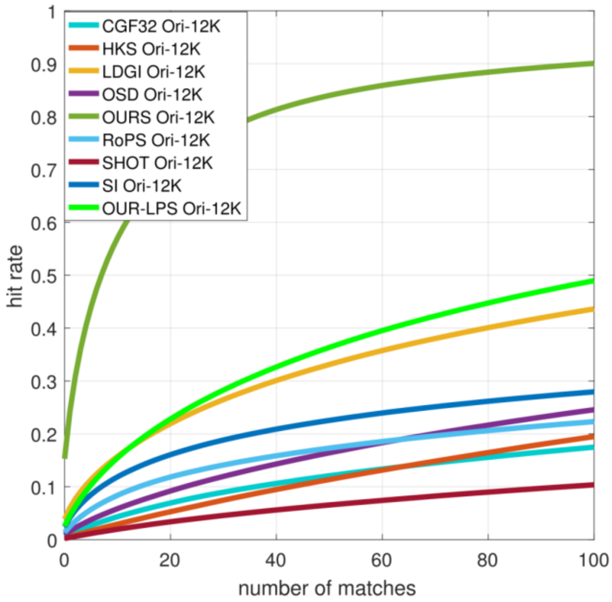
❖ **Results: Robust to resolution and triangulation**

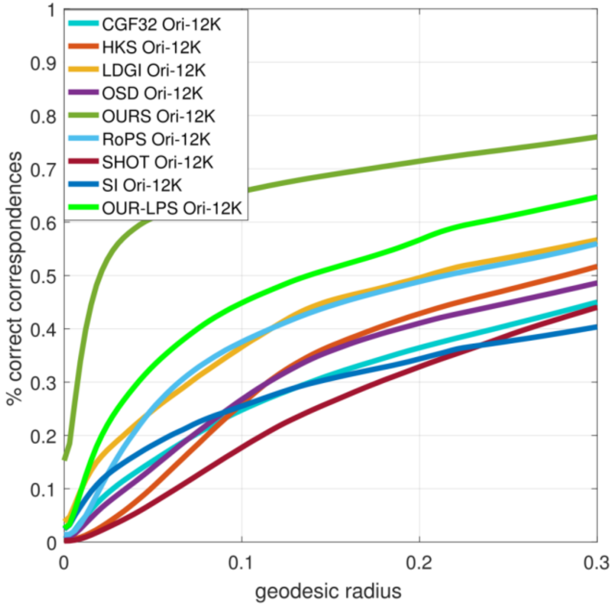




2-Robust Local Spectral Descriptor

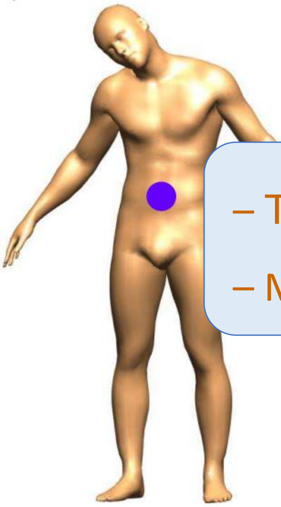
❖ Results: Comparison





3-Descriptor Learning using Multiscale GCNs

❖ Motivation

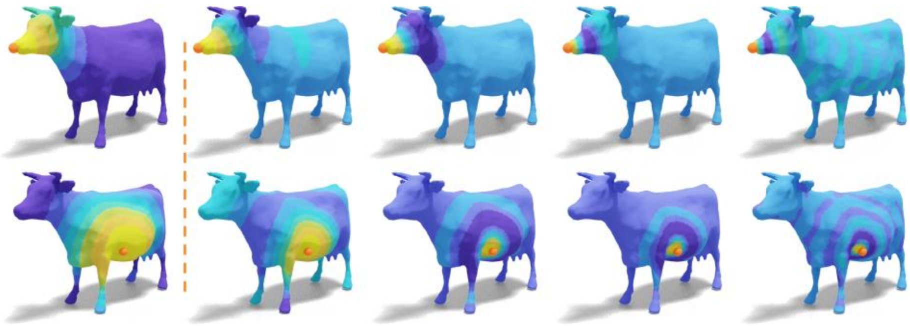


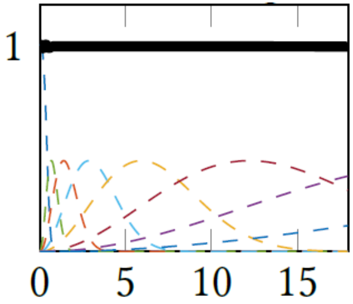
- – Time-consuming
- – Miss global information

cut a geodesic disk

3-Descriptor Learning using Multiscale GCNs

❖ Graph wavelets





Wavelet filter functions

$$\xi_v = \sum_{j=0}^{N-1} a(v)\, h(\lambda_j)\, \phi_j(v)\, \boldsymbol{\phi}_j$$

$$\boldsymbol{\psi}_{t,v} = \sum_{j=0}^{N-1} a(v)\, g(t\lambda_j)\, \phi_j(v)\, \boldsymbol{\phi}_j$$

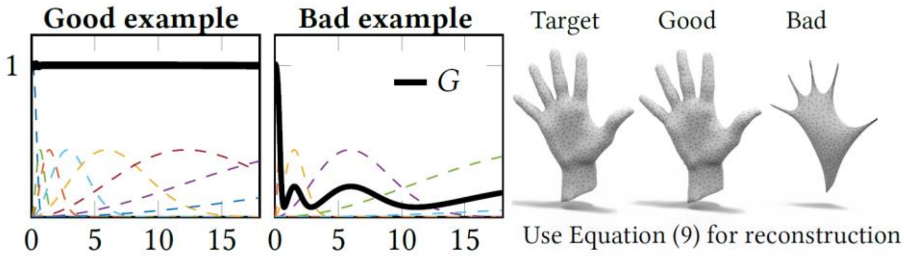
Scaling functions

Wavelet functions

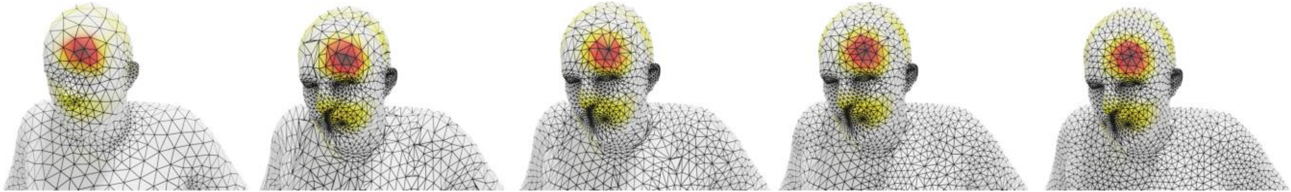
- Multiscale property of wavelet functions

3-Descriptor Learning using Multiscale GCNs

- **Graph wavelets**
  - Reconstruction property of multiscale wavelet functions



  - Robustness to the change of resolution and triangulation



3-Descriptor Learning using Multiscale GCNs

Wavelet Energy Decomposition Signature (WEDS)

Reconstruct discrete Dirichlet energy

$$E(\mathbf{f}) = \mathbf{f}^{\mathrm{T}}\mathbf{L}\mathbf{f} = \sum_{j=0}^{N-1} \lambda_j \left(\sum_{m=0}^{K} \sum_{v} \gamma_j (t_m, v) \right)^2$$

Restructure into a sum per vertex

$$E(\mathbf{X}) = \sum_{i=1}^{d} \sum_{j=0}^{N-1} \lambda_j \sum_{m=0}^{K} \sum_{v} \gamma_{ij}(t_m, v)\omega_{ij} = \sum_{m=0}^{K} \sum_{v} \sum_{j=0}^{N-1} \lambda_j \sum_{i=1}^{d} \gamma_{ij}(t_m, v)\omega_{ij},$$

Collect the local energy using multiscale wavelets

$$WEDS_{t_s}(v) = \left\{ \sum_{x} \psi^{*}_{t_s, v}(x) \varepsilon_{t_m}(x) \right\}, m \in [0, K]$$

3-Descriptor Learning using Multiscale GCNs

- **Wavelet Energy Decomposition Signature (WEDS)**

3-Descriptor Learning using Multiscale GCNs

❖ Multiscale Graph Convolution Network

- Convolutions on graphs are defined as:

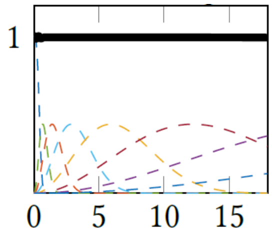
$$\mathbf{x}^{\text{out}} = \mathbf{y} *_{w} \mathbf{x}^{\text{in}} = \Phi\left(\left(\Phi^T \mathbf{y}\right) \odot \left(\Phi^T \mathbf{x}^{\text{in}}\right)\right) = \Phi \mathbf{w}_\theta \Phi^T \mathbf{x}^{\text{in}}$$

- ChebyNet approximates $\mathbf{w}_\theta$ using m-order polynomials:

$$\mathbf{w}_\theta = \sum_{m=0}^{K-1} \theta_m \operatorname{diag}\left(\{\lambda_j\}_{j=0}^{k-1}\right)^m = \sum_{m=0}^{K-1} \theta_m \boldsymbol{\Lambda}^{\boldsymbol{m}}$$

- MGCN approximates $\mathbf{w}_\theta$ using multiscale wavelet filter basis:

$$\mathbf{w}_\theta = \sum_{m=0}^{K} \theta_m \operatorname{diag}\left(\{g_{t_m}(\lambda_j)\}_{j=0}^{k}\right) = \sum_{m=0}^{K} \theta_m g_{t_m}(\boldsymbol{\Lambda})$$

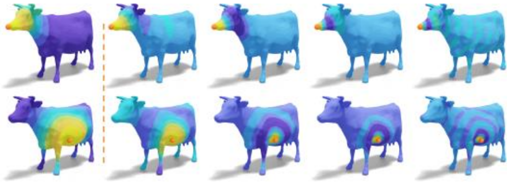


3-Descriptor Learning using Multiscale GCNs

❖ Multiscale Graph Convolution Network

- The multiscale convolution can be simplified as follows:

$$\mathbf{x}^{\text{out}} = \mathbf{y} *_w \mathbf{x}^{\text{in}} \approx \sum_{t_s \in S_{t_s}} \theta_{t_s} \overline{\Psi_{t_s}}^T \mathbf{x}^{\text{in}},$$

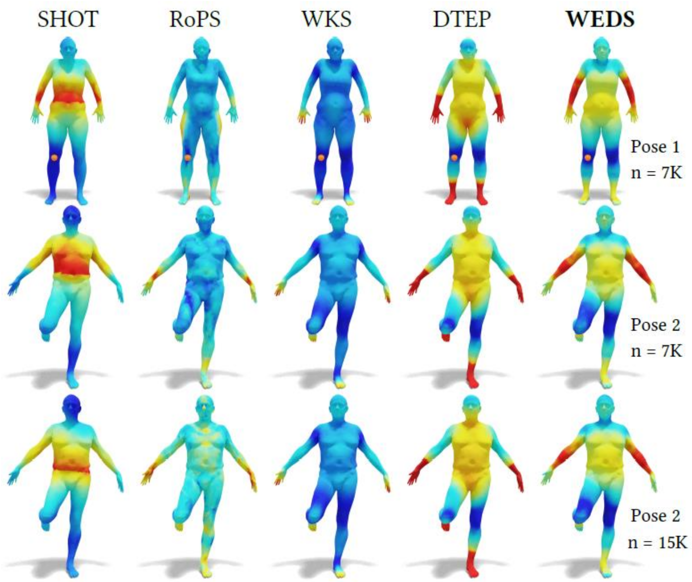


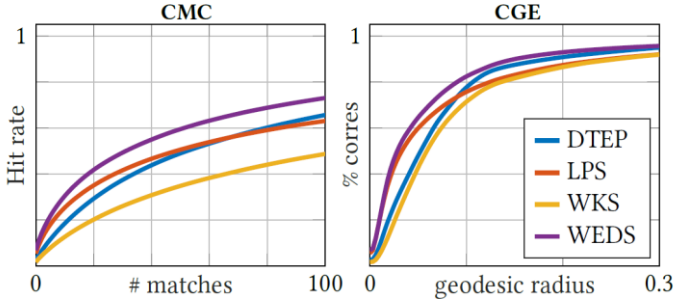
Convolve local and global information

Robustness to discretization

3-Descriptor Learning using Multiscale GCNs

❖ **Results: performance of WEDS**

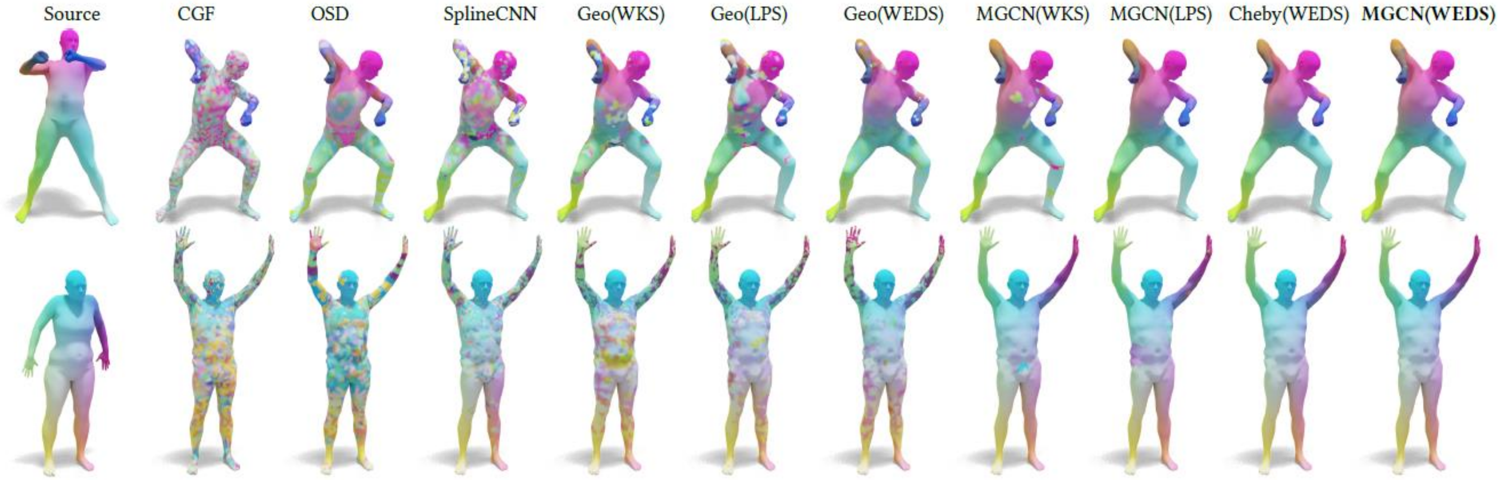




Descriptors	Dataset	
	FAUST(6890)	SCAPE(12.5K)
SI	352 / 153	380 / 251
SHOT	381 / 262	271 / 196
RoPS	346 / 252	267 / 187
HKS	511 / 396	507 / 409
WKS	335 / 118	260 / 72
LPS	325 / 97	227 / 71
DTEP	312 / 89	267 / 76
WEDS	**287 / 69**	**225 / 66**

3-Descriptor Learning using Multiscale GCNs

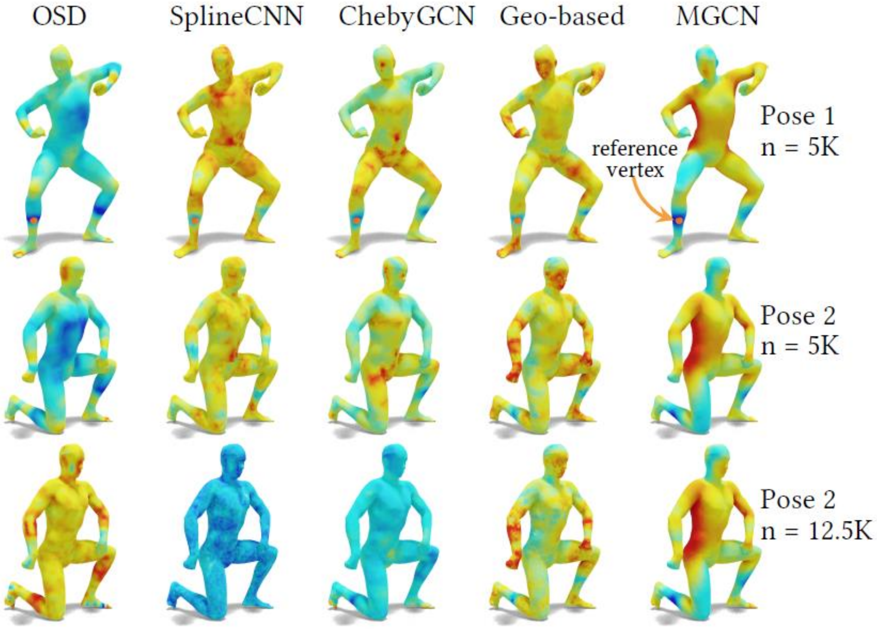
- Results: shape correspondence



FAUST

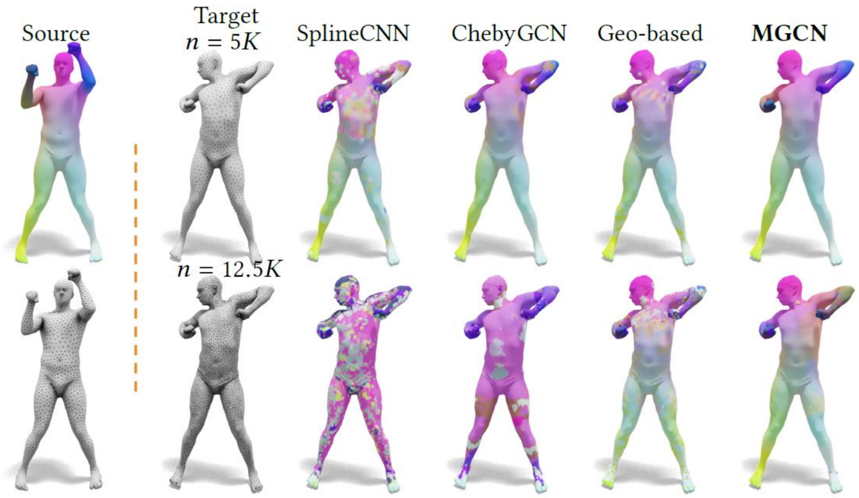
3-Descriptor Learning using Multiscale GCNs

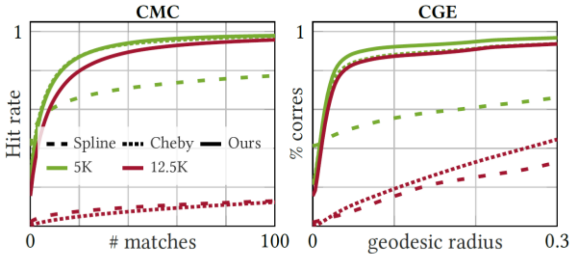
❖ Results: robust to resolution and triangulation



3-Descriptor Learning using Multiscale GCNs

❖ Results: robust to resolution and triangulation



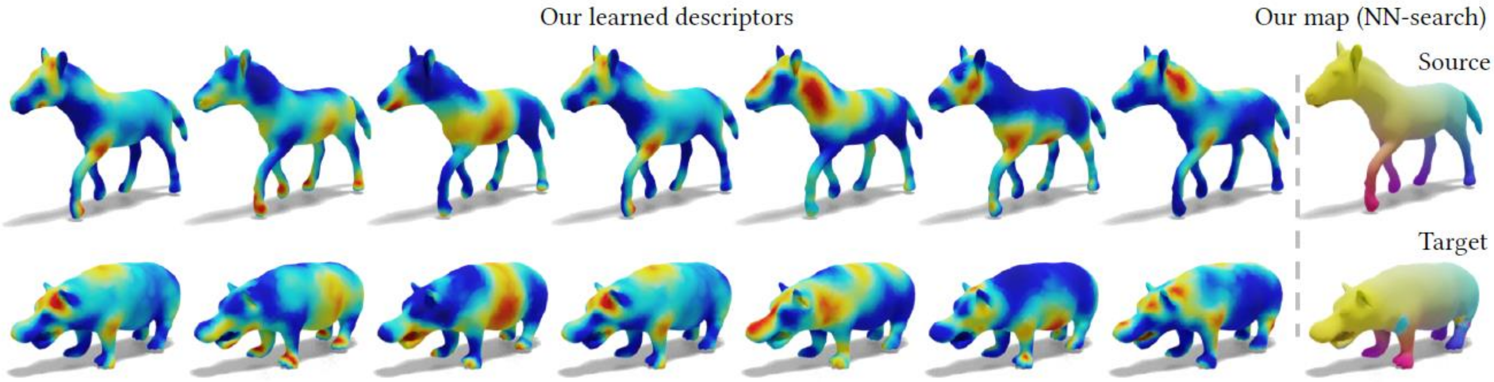


Descriptors	Network	#Resolution	
		SCAPE 5K	SCAPE 5K-12.5K(Test)
Histogram	CGF	374 / 264	428 / 323
-	OSD	259 / 94	835 / 742
-	SplineCNN	297 / 180	503 / 353
WEDS	ChebyGCN	68 / 29	458 / 332
WKS	Geo-based	185 / 72	192 / 80
LPS	Geo-based	175 / 68	182 / 74
WEDS	Geo-based	163 / 65	179 / 71
WKS	MGCN	67 / 46	98 / 64
LPS	MGCN	54 / 26	82 / 44
WEDS	MGCN	**48 / 17**	**73 / 39**

Extended FAUST

3-Descriptor Learning using Multiscale GCNs

Results: non-isometric shapes



Shape Pairs	BIM	SplineCNN + NN	MGCN + NN
Average (8×7 pairs)	59	225	**45**
Cow, Wolf	37	219	**18**
Tiger, Dog	39	227	**38**
Fox, MaleLion	104	221	**51**
Horse, Hippo	**76**	244	80

Conclusions

❖ Contributions

— Two non-learned features: *Local Point Signature* (LPS) and *Wavelet Energy Decomposition Signature* (WEDS).

— Two supervised frameworks to transform the non-learned features to more discriminative descriptors

❖ Future work

— Extend to point clouds and triangle soups

— Industrial applications

Acknowledgements

Co-authors:

Dong-Ming Yan, CASIA

Yiqun Wang, CASIA

Hanyu Wang, University of Maryland-College Park

Jing Ren, Peter Wonka, KAUST

CCF-腾讯犀牛鸟科研基金

Code and Data:

LDGI: https://github.com/jianweiguo/local3Ddescriptorlearning

LPS: https://github.com/yiqun-wang/LPS

MGCN: https://github.com/yiqun-wang/MGCN

Thank you!