Field Convolutions for Surface CNNs: Supplement

Thomas W. Mitchel
Johns Hopkins University
tmitchel@jhu.edu

Vladimir G. Kim
Adobe Research
vokim@adobe.com

Michael Kazhdan
Johns Hopkins University
misha@cs.jhu.edu

A. Field convolutions commute with isometries

Here we offer a detailed proof of the claim made in Section 4 that field convolution commutes with the action of isometries. That is, given any $X \in \Gamma(TM)$ and filter $f \in L^2(\mathbb{C})$, if $\Psi : M \to N$ is an isometry, then

$$d\Psi\left[\big(X * f\big)(p)\right] = \left[d\Psi(X) * f\right]\big(\Psi(p)\big). \tag{1}$$

To see this, consider surfaces M and N and any two points $p \in M$ and $p' \in N$. Let $\mathcal{N} \subset M$ and $\mathcal{N}' \subset N$ be $\epsilon-$balls about the points and suppose that $\mathcal{N}$ and $\mathcal{N}'$ are isometric. That is, there exists a map $\Psi : M \to N$ taking p to p' and satisfying $\forall q_i \in \mathcal{N}$,

$$d\left(q_0,\, q_1\right) = d\left(q_0',\, q_1'\right), \quad q_i' = \Psi(q_i) \in \mathcal{N}',$$

where $d\left(\,\cdot\,,\,\cdot\,\right)$ is the geodesic distance.

Let $X' \in \Gamma(TN)$ be the push-forward of X under $d\Psi$ where, using the tangent space representation of Knoppel *et al.* [3], $X'\big|_{p'} = \rho'_{p'}\, e^{i\phi'_{p'}}$. For any two points $a, b \in M$, denote the logarithm of a with respect to b and the change in angle resulting from the parallel transport along the shortest geodesic from b to a as $\log_b a = r_{ba}\, e^{i\theta_{ba}}$ and φ_{ab}, respectively. It follows that $\forall q \in \mathcal{N}$ [2],

$$\begin{aligned} \rho'_{q'} &= \rho_q \quad \text{and} \quad \phi'_{q'} = \phi_q + \psi_q, \\ r_{q'p'} &= r_{qp} \quad \text{and} \quad \theta_{q'p'} = \theta_{qp} + \psi_q, \\ \varphi_{p'q'} &= \varphi_{pq} + \psi_p - \psi_q, \end{aligned}$$

where ψ_p is the angle of rotation corresponding to the action of the differential $d\Psi\big|_p$, taking vectors in T_pM to $T_{p'}N$. (Recall that as Ψ is an isometry, $d\Psi$ is an orthogonal transformation.) Then, in the expression for the field convolution

$$(X' * f)\big|_{p'} = \int_M \rho_{q'} e^{i(\phi_{q'} + \varphi_{p'q'})}\, f\left(r_{q'p'} e^{i\left(\theta_{q'p'} - \phi_{q'}\right)}\right)\, dq'$$

we have

$$\begin{aligned} \rho'_{q'}\, e^{i\left(\phi'_{q'} + \varphi_{p'q'}\right)} &= \rho_q\, e^{i(\phi_q + \varphi_{pq} + \psi_p)}, \\ r_{q'p'}\, e^{i\left(\theta_{q'p'} - \phi'_{q'}\right)} &= r_{qp}\, e^{i(\theta_{qp} - \phi_q)}, \end{aligned}$$

with the measures dq and dq' satisfying $dq' = dq$ since $d\Psi\big|_q$ is an orthogonal transformation. From the definition of field convolution, this gives $(X' * f)\big|_{p'} = e^{i\psi_p}\,(X * f)\big|_p$ which is equivalent to $d\Psi\left[\big(X * f\big)(p)\right] = \left[d\Psi(X) * f\right]\big(\Psi(p)\big)$ as desired.

B. Learned Gradients

In practice, inputs to surface CNNs are often scalar features, such as the raw 3D positions of points. To lift such features to a vector field, we use a learnable operation analogous to a weighted gradient calculation. For any function $\xi \in L^2(V)$ we learn the magnitude and direction of its "gradient" separately, with respect to compactly supported radially isotropic filters $f_1,\, f_2 \in L^2(\mathbb{C})$. That is, we learn the vector field $\Phi_{f_1} : V \to \mathbb{C}$ and scalar field $P_{f_2} : V \to \mathbb{R}$ with

$$\Phi_{f_1}(p) = e^{i\beta} \sum_{q \in \mathcal{N}_p} w_q\, \left(\xi(q) - \xi(p)\right)\, f_1(r_{pq})\, e^{i\theta_{pq}}, \tag{2}$$

$$P_{f_2}(p) = \sum_{q \in \mathcal{N}_p} w_q\, \xi(q)\, f_2(r_{pq}) \tag{3}$$

with w_q, r_{pq}, θ_{pq} defined as in Equation (7) (the latter two parameters corresponding to $\log_p q$) and β a learnable rotational offset. Using these, we define the "gradient" of ξ with respect to f_1 and f_2 as the vector field

$$P^2_{f_2}(p)\frac{\Phi_{f_1}(p)}{\|\Phi_{f_1}(p)\|} \tag{4}$$

While this approach ensures that scalar features are passed directly to vector fields, we do not consider it to be a critical part of our framework and it can be replaced by a linear layer with only a small decrease in performance.

C. Feature Matching Experiments

Here we provide a detailed explanation of how our feature matching experiments are performed in Section 6.5. Each pair in the SHREC 2019 Correspondence Dataset [1] consists of a *model* mesh V_M and a *scene* mesh V_S, with the dense ground-truth correspondence mapping the latter to the former. We randomly generate correspondences $C_{SM} = \{(s_i, m_i)\} \subset V_S \times V_M$ and non-correspondences $N_{SM} = (V_S \times V_M) \setminus C_{SM}$ by selecting 2048 points on both the model and the scene mesh using farthest point sampling, mapping the sampled scene points to the model mesh using the ground truth correspondence, and associating each mapped scene point to the geodesically nearest sampled point on the model.

In training, the objective of the network is to make the outputs for corresponding and non-corresponding pairs as similar and dissimilar as possible, respectively [4, 5]. To this end we use a twin network, wherein each mesh in a pair is fed to the same network which learns a compact 16-dimensional descriptor F at each point. Specifically, for each pair in each epoch, we randomly subsample 512 pairs of corresponding and non-corresponding points, $P_{SM} = C_{SM}^{512} \cup N_{SM}^{512}$ and minimize the twin loss [8]

$$L(P_{SM}) = \sum_{(s,m)\in P_{SM}} \alpha_{s,m} \|F_S(s) - F_M(m)\|^2 + (1 - \alpha_{s,m}) \max\left(0, 5 - \|F_S(s) - F_M(m)\|^2\right), \tag{5}$$

where $\alpha_{s,m} = 1$ if $(s, m) \in C_{SM}$ or is set to a random variable between 0 and 0.2 otherwise.

We compute precision-recall curves as follows. Given a sampled point in the scene mesh $s \in V_S$, we sort all sampled model points based on descriptor distance, giving $\{m_1, \ldots, m_K\} \subset V_M$, with

$$\|F_S(s) - F_M(m_i)\| \leq \|F_S(s) - F_M(m_{i+1})\|,$$

for $1 \leq i \leq K - 1$. We define $\mathcal{M}_p \subset V_M$ to be the set of sampled model points that are valid matches with p, which consists of all sampled model points whose ground-truth correspondence lies within a geodesic ball of radius 0.05 about p. While this corresponds to a slightly more relaxed definition of correspondence, we find that all methods perform better maintaining a stricter notion of correspondence during training. Then, following [7, 6] the precision $\mathcal{P}_p$ and recall $\mathcal{R}_p$ assigned to p are defined as functions of the top r model keypoints,

$$\mathcal{P}_p(r) = \frac{\left|\mathcal{M}_p \cap \{m_i\}_{i\leq r}\right|}{r}, \tag{6}$$

$$\mathcal{R}_p(r) = \frac{\left|\mathcal{M}_p \cap \{m_i\}_{i\leq r}\right|}{|\mathcal{M}_p|}. \tag{7}$$

References

[1] R. M. Dyke, C. Stride, Y.-K. Lai, P. L. Rosin, M. Aubry, A. Boyarski, A. M. Bronstein, M. M. Bronstein, D. Cremers, M. Fisher, T. Groueix, D. Guo, V. G. Kim, R. Kimmel, Z. Lähner, K. Li, O. Litany, T. Remez, E. Rodolà, B. C. Russell, Y. Sahillioğlu, R. Slossberg, G. K. L. Tam, M. Vestner, Z. Wu, and J. Yang. Shape correspondence with isometric and non-isometric deformations. In Silvia Biasotti, Guillaume Lavoué, and Remco Veltkamp, editors, *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2019. 1

[2] Jean H Gallier and Jocelyn Quaintance. *Differential Geometry and Lie Groups: A Computational Perspective*, volume 12. Springer Nature, 2020. 1

[3] Felix Knöppel, Keenan Crane, Ulrich Pinkall, and Peter Schröder. Globally optimal direction fields. *Transactions on Graphics*, 32(4), 2013. 1

[4] Roee Litman and Alexander M Bronstein. Learning spectral descriptors for deformable shape correspondence. *Transactions on Pattern Analysis and Machine Intelligence*, 36(1):171–180, 2013. 2

[5] Jonathan Masci, Davide Boscaini, Michael Bronstein, and Pierre Vandergheynst. Geodesic convolutional neural networks on riemannian manifolds. In *International Conference on Computer Vision*, pages 37–45, 2015. 2

[6] Thomas W Mitchel, Szymon Rusinkiewicz, Gregory S Chirikjian, and Michael Kazhdan. Echo: Extended convolution histogram of orientations for local surface description. In *Computer Graphics Forum*. Wiley Online Library, 2020. 2

[7] Philip Shilane, Patrick Min, Michael Kazhdan, and Thomas Funkhouser. The Princeton shape benchmark. In *Proceedings Shape Modeling Applications*, pages 167–178, 2004. 2

[8] Zhiyu Sun, Yusen He, Andrey Gritsenko, Amaury Lendasse, and Stephen Baek. Embedded spectral descriptors: learning the point-wise correspondence metric via siamese neural networks. *Journal of Computational Design and Engineering*, 7(1):18–29, 2020. 2