Spherical Diffusion Process for Score-Guided Cortical Correspondence via Spectral Attention

Seungeun Lee[1], Sergey Pyatkovskiy[2], Jaejun Yoo[⋆2], and
Ilwoo Lyu[⋆3][0000−0001−5868−9603]

[1] AI Research Team, Klleon, Seoul, South Korea
[2] Graduate School of Artificial Intelligence, UNIST, Ulsan, South Korea
[3] Graduate School of Artificial Intelligence, POSTECH, Pohang, South Korea
`jaejun.yoo@unist.ac.kr,ilwoolyu@postech.ac.kr`

Abstract. Cortical shape correspondence is a crucial problem in medical image analysis, primarily focused on aligning cortical geometric patterns across individuals. This task is particularly challenging due to the intricate geometry of the cortex and the substantial anatomical variability among individuals. In this work, we introduce a novel approach comprising (1) a spherical diffusion process and (2) a spectral attention for robust shape correspondence construction, wherein a score function from the diffusion process guides a deformation to align cortical geometric features on sphere. Specifically, we propose a smooth diffusion process on sphere by introducing a stochastic differential equation in a spherical harmonic space, where we learn the score function that encodes the distribution of subjects. Furthermore, to effectively guide the alignment of cortical geometric patterns using the learned score function, we propose a novel attention mechanism that computes frequency correlations in the spectral domain, enabling efficient conditioning of the score function in this domain. Experimental results demonstrate that our method achieves highly accurate shape correspondence while minimizing the distortions. The code is available at https: //github.com/Shape-Lab/SPHARM-Reg-Diffusion.

Keywords: Cortical correspondence · Spherical diffusion process · Spectral attention.

1 Introduction

Cortical shape correspondence is essential for spatial alignment in shape analysis, matching cortical surfaces across individuals via geometric features. Traditional methods [10,28,24,22,21] use shape deformation, mapping surfaces onto a common space, typically a sphere, and optimizing deformation for feature similarity and smoothness. These methods align geometric features like sulcal depth, assuming structural and functional correspondence. Recently, learning-based approaches [4,26,29,19,18] have gained popularity, leveraging neural networks to

---
⋆ Co-Corresponding Author.

predict deformation fields in an end-to-end manner. While the approaches have demonstrated substantial advances in shape correspondence, they still suffer from the challenge of large anatomical variability across subjects. As a result, finding an optimal and robust correspondence for arbitrary subjects remains an open problem.

Recently, diffusion models [25,13,5,8,14] have been actively applied to various medical imaging analysis. Several studies [16,30,23] have proposed leveraging a score function of diffusion models to compute correspondence between volumetric images on Euclidean space. These studies have found that the score function captures the underlying distribution of source data and encodes the transition density between a pair of images. Consequently, the learned latent features have shown to establish robust correspondence for arbitrary subjects at test time. However, directly applying the diffusion model to non-Euclidean space is non-trivial, as it lacks a proper geometric inductive bias, as discussed in [15,7,20].

In this paper, we propose a novel spherical diffusion model for shape correspondence. The diffusion process on manifold is governed by the Riemannian Laplace-Beltrami operator [15,7,20], which lacks a closed-form solution in high-dimensional spaces, leading to inaccurate and instable approximations [15,7,20]. However, we observe that on sphere, this operator can be explicitly defined using spherical harmonic (SH) coefficients and basis functions via the SH transformation (SHT) [6]. This allows us to estimate the score function in SH space, leveraging geometric features of the cortical surface mapped onto the sphere. Furthermore, we introduce a spectral attention to provide the semantics as a condition to guide shape correspondence. This prevents information dilution while reducing computational overhead of spatial attention, ultimately establishing robust shape correspondence for unseen subjects during training. Our main contributions can be summarized as follows:

– We propose a novel score-based spherical diffusion model by formulating the diffusion process as a heat equation using the spherical harmonic transformation, effectively capturing spherical geometry.
– We leverage the learned score function to model the transition density between cortical features, improving robustness in cortical correspondence.
– We introduce a spectral attention mechanism in the spherical harmonic space, reducing computational overhead compared to spatial attention.
– Extensive experiments demonstrate that our method well establishes the shape correspondence with improved accuracy and reduced distortion.

2   Methods

We formulate spherical deformation for cortical shape correspondence (Sec. 2.1), describe the base architecture to estimate a deformation field (Sec. 2.2), introduce a novel spherical diffusion model to estimate a score function (Sec. 2.3), and present a spectral attention mechanism, enabling efficient conditioning of the score function, and enhancing robustness for the shape correspondence (Sec. 2.4). Fig. 1 illustrates a schematic overview of the proposed method.
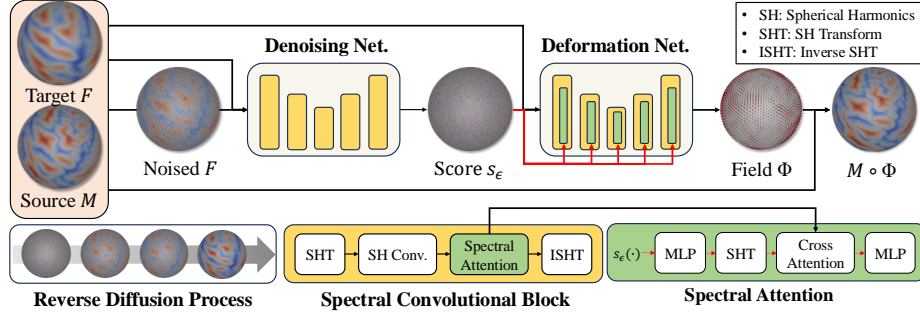
Fig. 1: Overall process of our proposed method. Given source and target geometric feature M, F, we first perform reverse process of spherical diffusion model, using denoising network to estimate score s_ϵ on SH space. We pass s_ϵ to the deformation network g_ψ along with attention mechanism on spectral domain, while taking the M and F as inputs to g_ψ. At the end, the estimated Φ is multiplied to M to align with F.

2.1   Problem Formulation

The shape correspondence aims to estimate a deformation field $\Phi : \mathbb{S}^2 \to \mathbb{S}^2$ to align a source geometric feature $M : \mathbb{S}^2 \to \mathbb{R}$ to a target geometric feature $F : \mathbb{S}^2 \to \mathbb{R}$, mapped onto a sphere with N uniformly sampled points $(\theta, \phi) \in [0, \pi] \times [-\pi, \pi]$, such that $F(\theta, \phi) = M(\Phi(\theta, \phi)$ for $\forall (\theta, \phi) \in \mathbb{S}^2$.

2.2   Underlying Learning-based Cortical Correspondence

We adopt the base architecture of [19,18] as the deformation network g_ψ which takes M and F as input to estimate Φ. g_ψ is jointly optimized over the similarity and regularization loss functions. The similarity term minimizes the L_2 loss to align $M \circ \Phi$ to F:

$$\mathcal{L}_{sim}(F, M \circ \Phi) = \frac{1}{N} \sum_{i=1}^{N} \left(F(\theta_i, \phi_i) - M \circ \Phi(\theta_i, \phi_i) \right)^2 . \tag{1}$$

The regularization term aims to reduce distortion of Φ by minimizing the arc length changes, between the deformation grids before and after deformation for spherical location $\mathbf{x} = (\theta, \phi)$ with its neighborhoods $\mathcal{N}_{\mathbf{x}} \subset \mathbb{S}^2$:

$$\mathcal{L}_{reg}(\Phi) = \frac{1}{2} \sum_{\mathbf{y} \in \mathcal{N}_{\mathbf{x}}} (\cos^{-1}(\mathbf{x}^\mathsf{T}\mathbf{y}) - \cos^{-1}(\Phi(\mathbf{x})^\mathsf{T}\Phi(\mathbf{y})))^2. \tag{2}$$

Based on this baseline g_ψ, we leverage score function learned from diffusion process on $\mathbb{S}^2$ manifold as a prior to achieve robust shape correspondence.

2.3   Spherical Diffusion Model

Definition. Let us consider a input F, and a condition M. Our spherical diffusion model aims to learn a score function that captures the transition infor-

mation from M to F during diffusion process. To introduce a smooth diffusion process with a geometric inductive bias on the $\mathbb{S}^2$ manifold, we construct the spherical heat equation as a stochastic differential system (SDE) over time $t \in \{0, 1, \cdots, T\}$:

$$\frac{\partial F(\theta, \phi, t)}{\partial t} = \Delta_{\mathbb{S}^2} F(\theta, \phi, t), \tag{3}$$

where $\Delta_{\mathbb{S}^2}$ is the Laplace-Beltrami Operator on $\mathbb{S}^2$. Furthermore, a spherical signal can be re-expressed using the SHT as a linear combination of spherical coefficients $c_l^k \in \mathbb{R}$ and SH basis functions $Y_l^k \in \mathbb{R}$ with SH degree l and order k. Since the LBO of the SH basis function has a known analytic form [6], the SDE on $\mathbb{S}^2$ is reformulated as follows:

$$\frac{\partial F(\theta, \phi, t)}{\partial t} = \Delta_{\mathbb{S}^2} \sum_{l=0}^{\infty} \sum_{k=-l}^{l} c_l^k(t) Y_l^k(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{k=-l}^{l} -l(l+1) c_l^k(t) Y_l^k(\theta, \phi). \tag{4}$$

Forward Process. Therefore, the diffusion process is represented in the SH space using the aforementioned known heat kernel. That is, the generalized Riemannian SDE [15,20] can be specialized to $\mathbb{S}^2$ using the spherical harmonic heat kernel. Instead of perturbing F, the c_l^k are perturbed along with the SDE:

$$dc_l^k(t) = -l(l+1) c_l^k dt + \sigma_t dW_l^k(t), \tag{5}$$

where $W_l^k(t) \in \mathbb{R}$ is an independent Brownian motion [25], and σ_t controls the noise magnitude which is calculated as $\sqrt{1 - e^{-2l(l+1)t}}$. By solving the above differential equation through numerical integration, the coefficient c_l^k at time t is obtained as follows:

$$c_l^k(t) = c_l^k(0) e^{-l(l+1)t} + \sigma_t w_l^k, \tag{6}$$

where $w_l^k \sim \mathcal{N}(0, \mathbf{I})$ is Gaussian noise. This formulation ensures that high-frequency components decay faster than low-frequency ones, progressively transforming the signal into isotropic Gaussian noise over time.

Reverse Process. The reverse process restores the original signal by solving the time-reversed SDE [1]. To obtain the transition density from M to F, We model the reverse process as conditional SDE. In high-dimensional manifolds, there is no closed-form solution for solving the time-reverse SDE because the basis functions that constitute the heat kernel are unknown, as discussed in [15,20]. However, on $\mathbb{S}^2$ manifold, the heat kernel and its basis functions are known in the form of SH, allowing the closed-form solution of the time-reverse SDE to be computed as the inverse function of the heat kernel. The solution at the previous step $t-1$ is computed as follows:

$$c_l^k(t-1) = l(l+1) c_l^k(t) + \nabla_{c_l^k} \log p_t(c_l^k(t) | \hat{c}_l^k(t)) + \sigma_t \zeta, \tag{7}$$

where $\hat{c}_l^k(t)$ is conditional SH coefficient of M, $\zeta \sim \mathcal{N}(0, \mathbf{I})$ is time-reversed Gaussian noise, and $\nabla_{c_l^k} \log p_t(c_l^k(t)|\hat{c}_l^k(t))$ is a conditional score function modeling the gradient of the log-density.

Denoising Network & Training. The score follows a complicated marginal probability density, making it difficult to model directly, as discussed in [13,25]. Therefore, we approximate $\nabla_{c_l^k} \log p_t(c_l^k(t)|\hat{c}_l^k(t))$ using a neural network $s_\epsilon(\cdot)$ given $c_l^k(t)$ and $\hat{c}_l^k$. To compute the score function with $c_l^k(t)$ and $\hat{c}_l^k(t)$ of F and M, we utilize SPHARM-Net [12], which is well-suited for our purpose as it transforms spherical signals into the SH space and performs convolution operations. For training, we modify the Riemannian score matching objective [15,20] for $\mathbb{S}^2$ with approximated score c_l^k / σ_t^2:

$$\mathcal{L}_{score}(\epsilon) = \mathbb{E}_{t,\hat{c}_l^k,c_l^k}[||s_\epsilon(c_l^k, \hat{c}_l^k, t) + \frac{c_l^k}{\sigma_t^2}||^2]. \tag{8}$$

In result, we can train a denoising network $s_\epsilon(\cdot)$ using the simplified score matching objective on the sphere to obtain the conditional score function that encodes the transition density. The denoising network s_ϵ and deformation network g_ψ are jointly optimized by minimizing the combination of $\mathcal{L}_{score}$, $\mathcal{L}_{sim}$, and $\mathcal{L}_{reg}$:

$$\arg\min_{\psi,\epsilon} \mathcal{L}_{score}(\epsilon) + \lambda_0 \mathcal{L}_{reg}(\psi) + \lambda_1 \mathcal{L}_{sim}(\psi), \tag{9}$$

where λ_0, λ_1 are balancing weight factors between the loss functions.

2.4   Spectral Attention

From the score-based conditional spherical diffusion model, we learn the distribution of the source data and obtain the output of the denoising network $s_\epsilon(\cdot)$ at the last time step, which encodes the transition density from the source feature M to the target feature F. To estimate Φ for the robust correspondence, we provide this as a condition to g_ψ. However, its simple concatenation as an input to g_ψ may dilute the information as propagated through the network. To effectively reflect $s_\epsilon(\cdot)$ in g_ψ, we incorporate a cross-attention mechanism into each layer of the warp module. In conventional approaches, a cross-attention is computed over spatial sampling points. This unfortunately requires 1.6B combinations for 40K vertices of the unit sphere, which is computationally demanding. To address this issue, we make the attention mechanism work in the spectral domain. Firstly, we map $s_\epsilon(\cdot)$ to $E \in \mathbb{R}^{(L+1)^2 \times d_e}$ with MLP, where $L = 40$ is SH bandwidth. Following the convention in a general attention mechanism [27] with a hidden dimension d_h, we let its query, key, and value be $Q \in \mathbb{R}^{(L+1)^2 \times d_h}$, $K \in \mathbb{R}^{(L+1)^2 \times d_h}$, and $V \in \mathbb{R}^{(L+1)^2 \times d_h}$, respectively. The proposed attention mechanism is then performed over the E and intermediate feature $H^{(i)} \in \mathbb{R}^{(L+1)^2 \times d^{(i)}}$ of g_ψ's i-th layer in SH space. Specifically, these values are calculated via matrix multiplication:

$$Q = H^{(i)} \cdot W_Q^{(i)}, \ K = E \cdot W_K^{(i)}, \ V = E \cdot W_V^{(i)}, \tag{10}$$

Table 1: Ablation study for spherical diffusion and spectral attention. Each component improves the shape correspondence, preserving geometric properties.

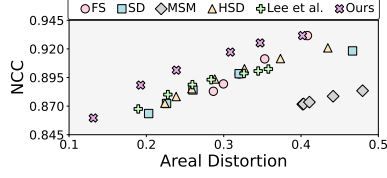| Method	Accuracy		Areal Distortion		
	MSE$\downarrow$	NCC$\uparrow$	Mean$\downarrow$	Median$\downarrow$	$P_{95}\downarrow$
Ablated Model	0.307	0.899	0.289	0.225	0.775
+ Spherical Diffusion	0.301	0.904	0.282	0.214	0.755
+ Spectral Attention	**0.294**	**0.910**	**0.275**	**0.204**	**0.750**

Fig. 2: Comparison at multiple regularization levels for each baseline and ours (x-axis: areal distortion, y-axis: NCC).

Table 2: **Bold**: best. *Blue*: $q < 0.05$. Our method significantly outperforms the comparison methods in terms of registration accuracy and distortion.

Method	Accuracy			Areal Distortion			Edge Distortion		
	MSE$\downarrow$	NCC$\uparrow$	Dice$\uparrow$	Mean$\downarrow$	Median$\downarrow$	$P_{95}\downarrow$	Mean$\downarrow$	Median$\downarrow$	$P_{95}\downarrow$
FS	0.313	0.898	0.874	0.322	0.267	0.792	0.136	0.109	0.349
SD	0.309	0.898	0.873	0.320	0.250	0.861	0.153	0.117	0.418
MSM	0.311	0.891	0.863	0.573	0.404	1.703	0.257	0.196	0.708
HSD	0.305	0.898	0.872	0.306	0.234	0.830	0.146	0.116	0.385
Lee *et al.*	0.307	0.899	0.871	0.289	0.225	0.775	0.144	0.112	0.384
Ours	**0.294**	**0.910**	**0.880**	**0.275**	**0.204**	**0.750**	**0.131**	**0.103**	**0.349**

where $W_Q^{(i)} \in \mathbb{R}^{d^{(i)} \times d_h}$, $W_K^{(i)} \in \mathbb{R}^{d_e \times d_h}$, and $W_V^{(i)} \in \mathbb{R}^{d_e \times d_h}$ are learnable projection matrices. We then perform attention operations, $softmax(QK^{\top}) \cdot V$, which generates an $(L+1)^2$-by-d_h matrix. Finally, we define another MLP that recovers the attention back into an $(L+1)^2$-by-$d^{(i)}$ matrix. It is noteworthy that the proposed approach can significantly speed up the attention operations regardless of the spherical tessellation (*i.e.*, only $(L+1)^4$ combinations required).

3   Experimental Setup

Dataset. We used the HCP dataset [11] (1,113 scans) for training only and the Mindboggle dataset [17] (101 scans) with manually annotated parcellation labels of 32 regions for testing.[1] Cortical surfaces were reconstructed with their invertible spherical mappings by a standard FreeSurfer pipeline [9]. The spheres were re-tessellated via icosahedral subdivision of level 6, with $N = 40,962$ [2].

Evaluation. We measured the Dice score, mean squared error (MSE), and normalized cross-correlation (NCC) to assess the alignment accuracy, with higher Dice score, NCC, and lower MSE indicating better alignment. We also measured areal [24] and edge [21] distortion to evaluate shape distortion, with the lower values indicating the better alignment while preserving the surface topology.

Implementation Details. For comparison with previous cortical shape correspondence methods [10,28,24,21,19], we adopted their settings using sulcal depth as the geometric feature and the Freesurfer *fsaverage* atlas. Notably, our model is not restricted to this feature and supports others. For denoising network, we

---

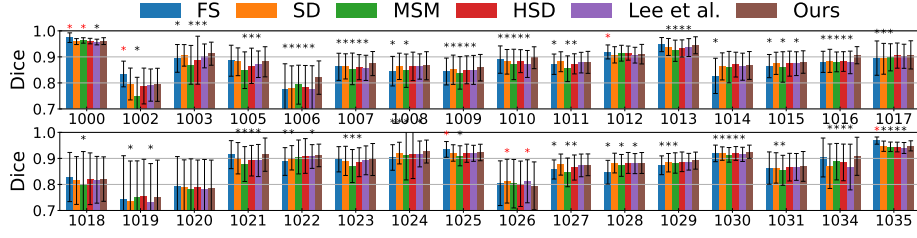[1] https://mindboggle.readthedocs.io/en/latest/labels.html

Fig. 3: Region-wise Dice score comparison. The statistical significance is reported after FDR [3] at $q = 0.05$. Our method significantly improves the score for a total of 17, 17, 24, 15, and 18 out 32 regions against FS [10], SD [28], MSM [24], HSD [21], and Lee *et al.* [19], respectively. Meanwhile, degraded performance is revealed in a total of 5, 1, 2, and 1 out of 32 regions compared to FS, SD, MSM, and Lee *et al.*. *: improved performance. *: degraded performance. For brevity, the annotation label numbers of these regions follow the DKT protocol; see [17] for the complete nomenclature.
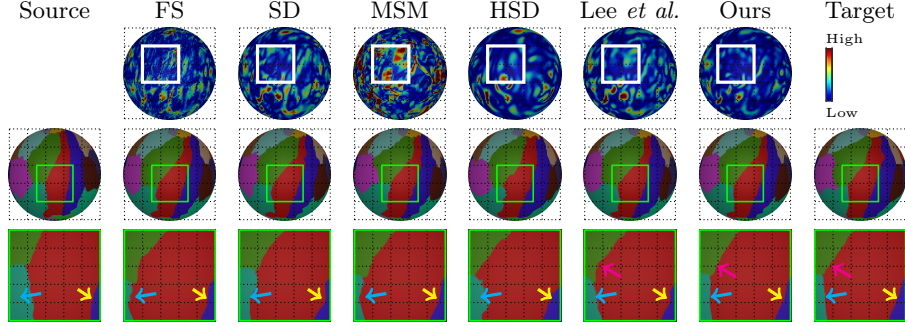

Fig. 4: Visual inspection of an example subject: exponentiated areal distortion (*1st row*), and parcellation maps (*2nd–3rd rows*). Our method particularly reduces the distortion at the supramarginal region (*white* box), along with anatomical parcellation alignment at the boundaries of the postcentral area (*cyan, yellow* & *magenta* arrow).

set SH bandwidth to 40. In spectral attention, for the i-th layer of g_ψ, we set $d_e = d^{(i)}$ and $d_h = 2 \cdot d^{(i)}$.

Baselines. We compared our method to public cortical shape correspondence methods, including FreeSurfer (FS) [10], Spherical Demons (SD) [28], MSM [24], HSD [21] and Lee *et al.*'s. [19]. To achieve fair comparison across all methods, we ran parameter optimization and report performance across all runs: FS with distance factors $[1, 3, 5, 7]$, SD with smoothing iterations $[1, 3, 5, 7, 9]$, MSM with regularization weight $[0.2, 0.3, 0.5, 0.7, 0.9]$, HSD with distortion regularization $[1/8^2, \cdots, 1/20^2]$, and Lee *et al.* with smoothing weight $[1/40, \cdots, 1/10]$.

4   Results

Ablation Study. To validate the effectiveness of our proposed method, we conducted ablation studies on the spherical diffusion and spectral attention. Starting from the base architecture, we sequentially added each modules. We report the results of this experiment in Table 1. We observed that spherical diffusion improved alignment by 1.95% and reduced distortion by 2.42%. Similarly, spectral attention enhanced alignment by 2.33% and reduced distortion by 2.42%. Each component demonstrated a significant performance improvement, indicating that our method establishes robust correspondence on unseen subjects during training. To ensure fair comparison, we tuned the weighting factors of regularization terms for baseline methods (see Sec. 3) and also varied the balancing weight λ_0 in our method within [10, 15, 20, 25, 30, 35] while fixing $\lambda_1 = 1$. Fig. 2 plots the alignment accuracy of all methods, varying the level of regularization. We observed a consistent trade-off between accuracy and distortion across all methods. Nevertheless, our method demonstrated superior the accuracy when all methods exhibited similar levels of distortion, indicating that our method has higher model capacity than others.

Comparison to Baselines. Table 2 summarizes the overall performance in registration accuracy and distortion. To ensure fair comparison, we selected models with similar registration accuracy ($\sim$0.3 MSE) and compared distortion reduction. Our method outperformed all baselines, particularly in reducing shape distortion with a statistical significance. It is noteworthy that the generalization performance of our method is comparable to that of traditional iterative optimization-based methods, which have been widely used and maintained for decades. In other words, unlike traditional methods that perform a repetitive optimization process for each individual subject, our neural network-based approach achieves improved alignment accuracy while preserving low distortion, even on test subjects unseen during training. Additionally, region-wise Dice scores (Fig. 3) show significant improvement in most of the 32 regions, with minor degradation in a few regions.

   This suggests that the score function from the diffusion process further refines region-wise alignment, while capturing both semantics and latent trajectories to the target. For qualitative evaluation, we computed group-average and sample-specific areal distortion maps, as well as parcellation maps (Figs. 4 and 5). The results confirm improved alignment in that our approach well-establishes the shape correspondence on the unseen subjects during training.

5   Conclusion

In this work, we proposed a novel framework for cortical shape correspondence that integrates a spherical diffusion process and a spectral attention mechanism. We defined a specialized diffusion process on $\mathbb{S}^2$ by establishing heat equation based on SHT. By learning a score function in SH space, our method effectively

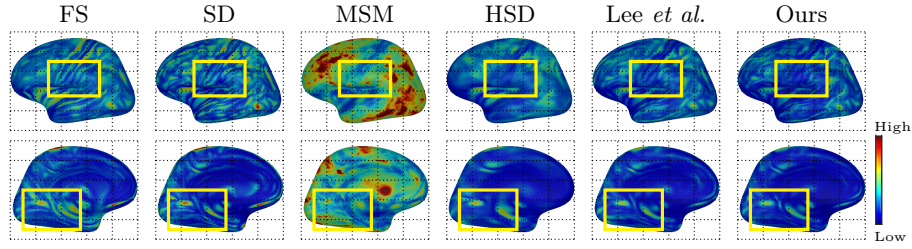FS   SD   MSM   HSD   Lee *et al.*   Ours

Fig. 5: The group average maps of exponentiated areal distortion across participants. Our method overally decreases the distortion, notably in the postcentral, supramarginal, parahippocampal and lingual regions (*yellow*).

captured the distribution of cortical shapes and guided geometric alignment in a principled manner. Additionally, spectral attention mechanism enabled efficient conditioning of the score, enhancing robustness of correspondence construction. Experimental results demonstrated the effectiveness of our approach in achieving highly accurate cortical shape correspondence while preserving geometric properties. These findings suggested that our method provides a promising direction for improving cortical shape analysis in medical imaging applications.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Anderson, B.D.: Reverse-time diffusion equation models. Stochastic Processes and their Applications **12**(3), 313–326 (1982)
2. Baumgardner, J.R., Frederickson, P.O.: Icosahedral discretization of the two-sphere. SIAM Journal on Numerical Analysis **22**(6), 1107–1115 (1985)
3. Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) **57**(1), 289–300 (1995)
4. Cheng, J., Dalca, A.V., Fischl, B., Zöllei, L., Initiative, A.D.N., et al.: Cortical surface registration using unsupervised learning. NeuroImage **221**, 117161 (2020)
5. Choi, J., Kim, S., Jeong, Y., Gwon, Y., Yoon, S.: Ilvr: Conditioning method for denoising diffusion probabilistic models. arXiv preprint arXiv:2108.02938 (2021)

6. Chung, M.K., Dalton, K.M., Shen, L., Evans, A.C., Davidson, R.J.: Weighted fourier series representation and its application to quantifying the amount of gray matter. IEEE transactions on medical imaging **26**(4), 566–581 (2007)
7. De Bortoli, V., Mathieu, E., Hutchinson, M., Thornton, J., Teh, Y.W., Doucet, A.: Riemannian score-based generative modelling. Advances in neural information processing systems **35**, 2406–2422 (2022)
8. Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis. Advances in neural information processing systems **34**, 8780–8794 (2021)
9. Fischl, B.: Freesurfer. Neuroimage **62**(2), 774–781 (2012)
10. Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M.: High-resolution intersubject averaging and a coordinate system for the cortical surface. Human brain mapping **8**(4), 272–284 (1999)
11. Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., et al.: The minimal preprocessing pipelines for the human connectome project. Neuroimage **80**, 105–124 (2013)
12. Ha, S., Lyu, I.: Spharm-net: Spherical harmonics-based convolution for cortical parcellation. IEEE Transactions on Medical Imaging **41**(10), 2739–2751 (2022)
13. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems **33**, 6840–6851 (2020)
14. Ho, J., Salimans, T.: Classifier-free diffusion guidance. arXiv preprint arXiv:2207.12598 (2022)
15. Huang, C.W., Aghajohari, M., Bose, J., Panangaden, P., Courville, A.C.: Riemannian diffusion models. Advances in Neural Information Processing Systems **35**, 2750–2761 (2022)
16. Kim, B., Han, I., Ye, J.C.: Diffusemorph: unsupervised deformable image registration using diffusion model. In: European Conference on Computer Vision. pp. 347–364. Springer (2022)
17. Klein, A., Tourville, J.: 101 labeled brain images and a consistent human cortical labeling protocol. Frontiers in neuroscience **6**, 171 (2012)
18. Lee, S., Lee, S., Ryu, S., Lyu, I.: Spharm-reg: Unsupervised cortical surface registration using spherical harmonics. IEEE Transactions on Medical Imaging (2025)
19. Lee, S., Ryu, S., Lee, S., Lyu, I.: Unsupervised learning of cortical surface registration using spherical harmonics. In: International Workshop on Shape in Medical Imaging. pp. 65–74. Springer (2023)
20. Lou, A., Xu, M., Farris, A., Ermon, S.: Scaling riemannian diffusion models. Advances in Neural Information Processing Systems **36**, 80291–80305 (2023)
21. Lyu, I., Kang, H., Woodward, N.D., Styner, M.A., Landman, B.A.: Hierarchical spherical deformation for cortical surface registration. Medical image analysis **57**, 72–88 (2019)
22. Lyu, I., Kim, S.H., Seong, J.K., Yoo, S.W., Evans, A., Shi, Y., Sanchez, M., Niethammer, M., Styner, M.A.: Robust estimation of group-wise cortical correspondence with an application to macaque and human neuroimaging studies. Frontiers in neuroscience **9**, 210 (2015)
23. Qin, Y., Li, X.: Fsdiffreg: Feature-wise and score-wise diffusion-guided unsupervised deformable image registration for cardiac images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 655–665. Springer (2023)
24. Robinson, E.C., Jbabdi, S., Glasser, M.F., Andersson, J., Burgess, G.C., Harms, M.P., Smith, S.M., Van Essen, D.C., Jenkinson, M.: Msm: a new flexible framework for multimodal surface matching. Neuroimage **100**, 414–426 (2014)

25. Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B.: Score-based generative modeling through stochastic differential equations. arXiv preprint arXiv:2011.13456 (2020)
26. Suliman, M.A., Williams, L.Z., Fawaz, A., Robinson, E.C.: A deep-discrete learning framework for spherical surface registration. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 119–129. Springer (2022)
27. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)
28. Yeo, B.T., Sabuncu, M.R., Vercauteren, T., Ayache, N., Fischl, B., Golland, P.: Spherical demons: fast diffeomorphic landmark-free surface registration. IEEE transactions on medical imaging **29**(3), 650–668 (2009)
29. Zhao, F., Wu, Z., Wang, F., Lin, W., Xia, S., Shen, D., Wang, L., Li, G.: S3reg: superfast spherical surface registration based on deep learning. IEEE Transactions on Medical Imaging **40**(8), 1964–1976 (2021)
30. Zhuo, Y., Shen, Y.: Diffusereg: Denoising diffusion model for obtaining deformation fields in unsupervised deformable image registration. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 597–607. Springer (2024)