Comparing Morse Complexes Using Optimal Transport: An Experimental Study

Mingzhe Li* University of Utah
Carson Storm† University of Utah
Austin Yang Li‡ University of Utah
Tom Needham § Florida State University
Bei Wang ¶ University of Utah

ABSTRACT

Morse complexes and Morse-Smale complexes are topological descriptors popular in topology-based visualization. Comparing these complexes plays an important role in their applications in feature correspondences, feature tracking, symmetry detection, and uncertainty visualization. Leveraging recent advances in optimal transport, we apply a class of optimal transport distances to the comparative analysis of Morse complexes. Contrasting with existing comparative measures, such distances are easy and efficient to compute, and naturally provide structural matching between Morse complexes. We perform an experimental study involving scientific simulation datasets and discuss the effectiveness of these distances as comparative measures for Morse complexes. We also provide an initial guideline for choosing the optimal transport distances under various data assumptions.

Keywords: Morse Complexes, topological data analysis, optimal transport, topology in visualization

Index Terms: Human-centered computing—Visualization——

1 INTRODUCTION

Morse complexes and Morse-Smale complexes are gradient-based topological descriptors of scalar fields. They have numerous applications in scientific visualization, such as feature correspondences [12], feature tracking [15, 16, 23, 26–29], symmetry detection [31], structural change detection [23], and uncertainty visualization [2]. Only a few comparative measures have been developed for these complexes and their variants [12, 23, 29].

On the other hand, recent years have seen the successful application of optimal transport to graph analysis. Namely, the Wasserstein distance [35], Gromov-Wasserstein distance [20, 22], and their variants [5, 33] have been used extensively for graph matching and comparison.

Leveraging recent advances in optimal transport (OT), we apply, *for the first time*, a class of OT-type distances to the comparative analysis of Morse complexes. Contrasting with existing comparative measures, such distances are easy and efficient to compute, and provide explicit structural matching between Morse complexes. Our main contribution is to provide experimental studies to evaluate the effectiveness of OT-type distances in terms of feature correspondences and classification, and to provide an initial guideline for choosing the OT-type distances under various data assumptions.

2 RELATED WORK

Optimal transport for graph matching and comparison. Mémoli first introduced Gromov-Wasserstein (GW) distances [20, 22] for the comparison of metric measure spaces. Theoretical works that expanded the scope of the GW framework (e.g., [6]) and new approaches in optimization [24] have made GW distances a popular tool for studying unregistered graphs. We briefly survey some of the related work here.

Applications of GW distances for graph analysis in a machine learning setting were explored by Xu et al. [38–40], with a focus on scalability and novel uses of the probabilistic nature of the metric. A Riemannian structure on the space of graphs endowed with the GW distance was established in [8], with a view toward statistical analysis of graphs. Connections between GW distances and spectral graph theory were provided in [9], inspired by earlier work of Mémoli in the setting of Riemannian manifolds [21]. Several variants of GW distances have been proposed [7, 30, 36], frequently motivated by applications in graph analysis, such as the fused GW [33, 34] and the partial GW distances [5]; these are discussed in detail below (see Sec. 3). Finally, the GW framework has been successfully applied to compare certain topological descriptors. In particular, it has been used as a tool for summarizing merge tree ensembles [17], feature tracking [19], and comparing merge trees and Reeb graphs endowed with additional topological attributes [10, 11].

Comparative measures for Morse and Morse-Smale complexes have previously focused on comparing the graphs derived from the complexes; see [41] for a survey. Feng et al. [12] studied *feature graphs*, which are 1D skeletons of simplified Morse-Smale complexes, to represent non-rigidly deformed surfaces. They used a minimum-cost matching algorithm to compare feature graphs. Thomas and Natarajan [31] used *augmented extremum graphs* to detect symmetry in scalar fields. They used the geodesic distances between extrema and earth mover's distance between histograms of seed regions. Narayanan et al. [23] defined a distance between extremum graphs by forming a complete graph between all pairs of extrema and computing the maximum distortion of the node sets and the edge sets. To the best of our knowledge, this paper presents the first time OT-type distances are used as comparative measures for Morse complexes.

3 BACKGROUND

We review Morse complexes and their 1D skeletons, referred to as Morse graphs. We also introduce notions of optimal transport (OT) type distances that are applicable in our comparative analysis.

Morse complexes and Morse graphs. We focus on 2D Morse complexes in this paper. Let $f : \mathbb{M} \to \mathbb{R}$ be a Morse function defined on a 2D manifold with boundary $\mathbb{M} \subset \mathbb{R}^2$, with gradient denoted ∇f. A point $x \in \mathbb{M}$ is a *critical point* if $\nabla f = 0$; otherwise, it is a *regular point*. There are three types of critical points: local maxima, local minima, and saddles. The *integral line* of a regular point is the maximal path through the point whose tangent vectors align with the gradient. The *descending manifold* surrounding a critical point contains the point itself and all regular points whose integral lines end at the critical point. Critical points (local minima and saddles) are the 0-cells, integral lines connecting these critical points are the 1-cells, and the rest of the domain makes up the 2-cells of a complex called the *Morse complex* of f. Its 1D skeleton consisting of 0- and 1-cells is referred to as the *Morse graph*.

*e-mail: mingzhe.li@utah.edu
†e-mail: u1120712@utah.edu
‡e-mail: u1364758@utah.edu
§e-mail: tneedham@fsu.edu
¶e-mail: beiwang@sci.utah.edu

For example, the 3D graph of a 2D function f is shown in Fig. 2 (bottom left), where its domain is segmented into nine descending manifolds (surrounding nine local maxima). Its corresponding Morse graph is embedded in the graph of f in cyan; it is also shown in 2D from a top-down viewpoint (top left).

Measure networks. In our context, a Morse graph is modeled as a *measure network* [6]. That is, it is represented as a triple $G = (V, p, W)$, where V is a finite set of nodes sampled from the Morse graph, $p : V \to [0,1]$ is a probability density on V (i.e., $p(x) \geq 0$ for all x and $\sum_{x \in V} p(x) = 1$), and $W : V \times V \to \mathbb{R}$ is a *network function* that captures the relations between pairs of nodes; the particular choice of node density and network function used to represent a Morse graph is explained in Sec. 4. Given a function $F : V \to A$, for some metric space (A, d_A), we may also consider a Morse graph as an *A-attributed measure network*, consisting of the data $G = (V, p, W, F)$. In this paper, we take our attribute space to be $\mathbb{R}^2$ endowed with a Euclidean distance, and the attribute function F assigns a node its location in the domain $\mathbb{M}$.

Given a pair of Morse graphs $G_1 = (V_1, p_1, W_1, F_1)$ and $G_2 = (V_2, p_2, W_2, F_2)$, such that the codomain of F_1 and F_2 is a common metric space (A, d_A), we introduce some notation for the sake of simplicity. Let $n_1 = |V_1|$ and $n_2 = |V_2|$ and enumerate the node sets as $\{x_i\}_{i=1}^{n_1}$ and $\{y_j\}_{j=1}^{n_2}$, respectively. We write $W_1(i,k) := W_1(x_i, x_k)$, $p_1(i) := p_1(x_i)$, and $a_i := F_1(x_i)$. It is convenient to consider W_1 as a matrix in $\mathbb{R}^{n_1 \times n_1}$ and p_1 as a vector in $\mathbb{R}^{n_1}$. We use similar notation for G_2, except we write $b_j := F_2(y_j)$.

We review a number of OT-type distances. They have the common feature of being defined in terms of measure couplings. A *coupling* C between probability measures p_1 and p_2 is a non-negative real-valued matrix representing a joint probability measure on $V_1 \times V_2$ whose row and column marginals agree with p_1 and p_2, respectively. The set of all couplings between p_1 and p_2 is denoted as

$$\mathscr{C} = \mathscr{C}(p_1, p_2) := \{C \in \mathbb{R}_+^{n_1 \times n_2} \mid C\mathbf{1}_{n_2} = p_1, C^T \mathbf{1}_{n_1} = p_2\}, \quad (1)$$

where $\mathbf{1}_n = [1, 1, ..., 1]^T \in \mathbb{R}^n$.

The OT-type distances under consideration are described below. Each OT-type distance is defined by an optimization problem over $\mathscr{C}$. Intuitively, we think of an element of $\mathscr{C}$ as a probabilistic "soft" registration of the vertices of the Morse graphs. The optimization problems favor registrations which preserve intrinsic (node relations) and/or extrinsic (attribute) features as much as possible. Each distance depends on a parameter $q \in [1, \infty)$; the definitions can be extended to $q = \infty$, but we will generally fix $q = 2$ in experiments.

Wasserstein distance [35]. The *q-Wasserstein distance* between G_1 and G_2 is defined by

$$d_q^W(G_1, G_2)^q = \min_{C \in \mathscr{C}} \sum_{i,j} d_A(a_i, b_j)^q C_{i,j}. \quad (2)$$

This distance compares nodes based on attributes, but is agnostic to the network structure encoded by the W-functions.

Gromov-Wasserstein distance [20]. The *q-Gromov-Wasserstein distance* (GW) is defined as

$$d_q^{GW}(G_1, G_2)^q = \min_{C \in \mathscr{C}} \sum_{i,j,k,l} |W_1(i,k) - W_2(j,l)|^q C_{i,j} C_{k,l} \quad (3)$$

This distance compares the network structures of the Morse graphs encoded by the W-functions, but has no dependency on attributes.

Fused Gromov-Wasserstein distance [34]. With a trade-off parameter $\alpha \in [0,1]$, the *q-Fused Gromov-Wasserstein distance* (FGW) between attributed Morse graphs G_1 and G_2 is defined as

$$d_{q,\alpha}^{FGW}(G_1, G_2)^q = \min_{C \in \mathscr{C}} \sum_{i,j,k,l} [(1-\alpha) d_A(a_i, b_j)^q + \alpha |W_1(i,k) - W_2(j,l))|^q] C_{i,j} C_{k,l}. \quad (4)$$

We have $d_{q,0}^{FGW} = d_q^W$ and $d_{q,1}^{FGW} = d_q^{GW}$, whereas $\alpha \in (0,1)$ yields a distance which depends on both network and attribute structures. See the supplementary materials for a simple example illustrating the differences between these distances.

The distances defined above can be sensitive to outliers, since they are forced to match the full masses of p_1 and p_2. These constructions can be made more robust by using *relaxed couplings*. That is, given $m \in [0,1]$, we may consider measures on $V_1 \times V_2$ with relaxed coupling constraints and a total mass given by m; that is, we define

$$\mathscr{C}_m = \mathscr{C}_m(p_1, p_2) := \{C \in \mathbb{R}_+^{n_1 \times n_2} \mid C\mathbf{1}_{n_2} \leq p_1, C^T \mathbf{1}_{n_1} \leq p_2, \mathbf{1}_{n_1}^T C \mathbf{1}_{n_2} = m\}. \quad (5)$$

Each of the OT-type distances d_q^W, d_q^{GW} and $d_{q,\alpha}^{FGW}$ has an associated "partial" version, which depends on an additional mass parameter m. These are obtained by replacing optimization over $\mathscr{C}$ in the above definitions with optimization over $\mathscr{C}_m$. Intuitively, elements of $\mathscr{C}_m$ still represent soft registrations between nodes, but with extra flexibility due to their ability to ignore some proportion of mass. Respectively, the partial versions of the OT-type distances are denoted to as $d_{q,m}^{pW}$ for *partial Wasserstein distance* (pW) [5], $d_{q,m}^{pGW}$ for *partial Gromov-Wasserstein distance* (pGW) [5] and $d_{q,\alpha,m}^{pFGW}$ for *partial Fused Gromov-Wasserstein distance* (pFGW) [19]. See the supplementary materials for formal definitions of the partial OT-type distances and a discussion of metric properties of all distances.

4 Method

To compare Morse graphs using the OT-type distances discussed in Sec. 3, we first model them as measure networks $G = (V, p, W)$. To emphasize well-connected nodes essential to the Morse graphs, we chose a p measure based on the degrees of the nodes following [40], in which $p(v) = \frac{\deg(v)}{\sum_i \deg v_i}$; this heuristic choice gives extra importance to more highly connected nodes when searching for an optimal registration. To capture the intrinsic pairwise relations between nodes in the Morse graph, we define $W(v,w)$ to be the shortest geodesic path length between v and w in V, following [19]. Finally, to model a Morse graph as an A-attributed measure network, the attribute function assigns a node v its location (v_x, v_y) in the domain.

To examine the structural alignment between Morse graphs, we apply a color transfer based on the coupling matrix between a source (reference) and a target Morse graph. We first assign a color to each node in the source using a gradient colormap based on node position. For each of the distances described above, we compute a minimizing coupling C. To each node y_j in the target, we assign the color corresponding to the node $\text{argmax}_{x_i} C(i,j)$ in the source. We then evaluate the quality of the alignment via visual inspection.

Implementation. We compute the Morse graphs using TTK (1.1.0) [32] and ParaView (5.10.1) [1]. The computation of the six OT-type distances is extended from [19], which is available on GitHub [18]. Our implementation for the FGW (thus Wasserstein and GW) distances follows [33, 34] and Python Optimal Transport (POT) [13], which uses conditional gradient method for optimization. For partial OT-type distances, we add dummy nodes to carry over the probability mass to be ignored, following [5].

5 Results

We study various OT-type distances using synthetic and real-world datasets. The first instance of each dataset is used as the source and shown in Fig. 1 (1st and 2nd rows); see supplementary materials for details on the datasets, parameter tuning, and runtime analysis.

5.1 An Overview With A Synthetic Dataset

We first use the synthetic Sinusoidal dataset to compare the behaviors of six OT-type distances introduced in Sec. 3, where the target

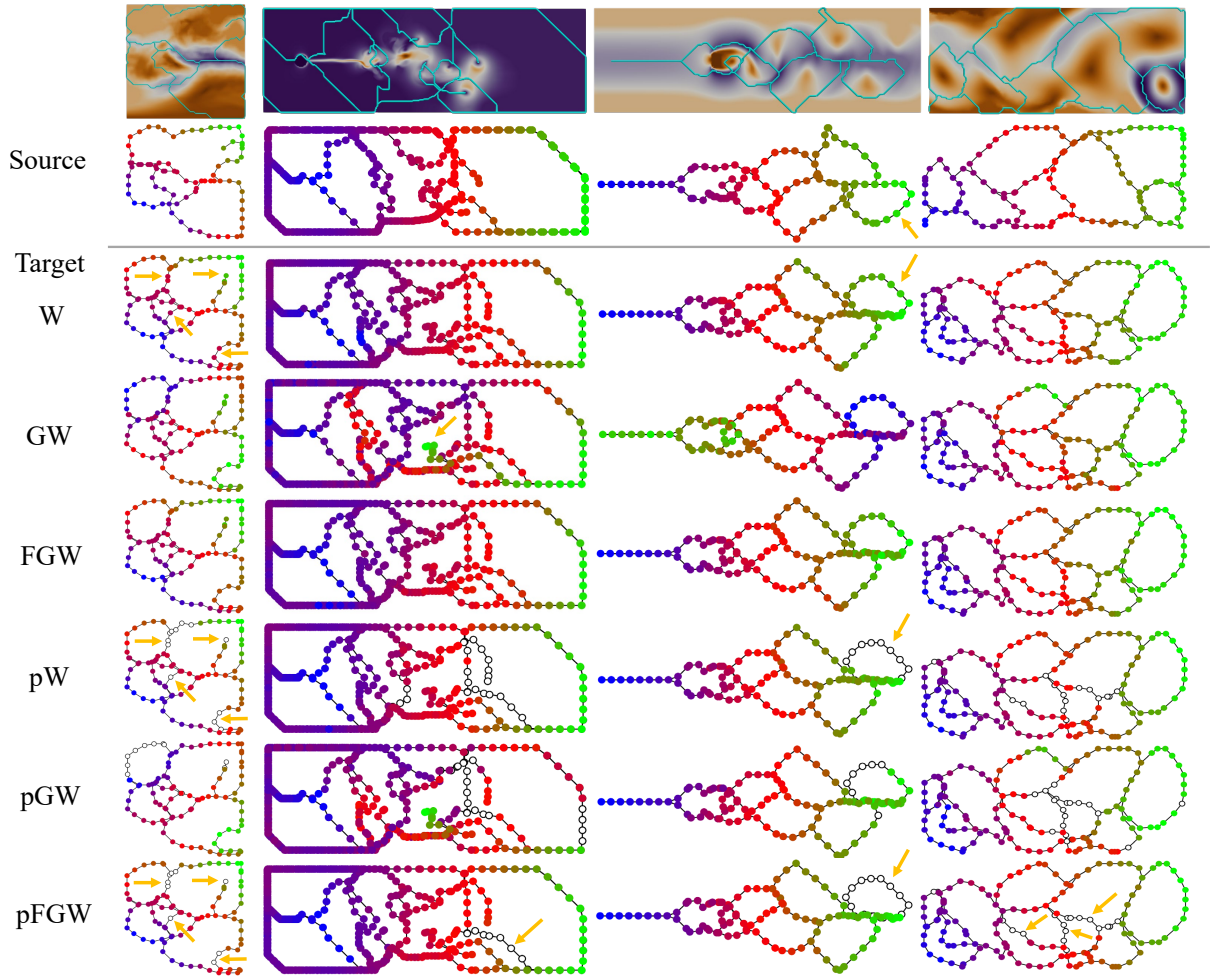

Figure 1: Structure alignments between the source and the target across all OT-type distances (from left to right) for the Wind, HeatedCylinder, NavierStokes, and RedSea datasets, respectively. W: Wasserstein distance.

field is generated from the source field with added noise. We use a color transfer to highlight the structural alignment between the source and the target Morse graphs across all distances. As shown

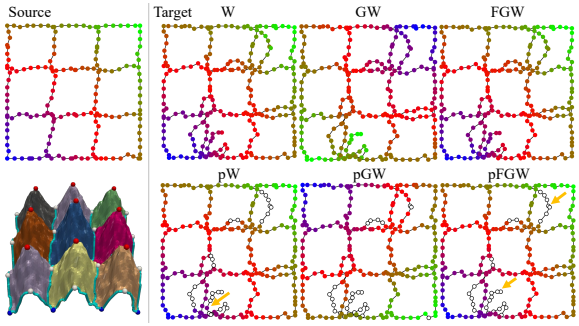

Figure 2: Sinusoidal dataset: structural alignments between the source and the target across all OT-type distances.

in Fig. 2 (top right), both the Wasserstein (W) and the FGW distance show reasonable structural alignments, even with the added noise in the target domain. However, using the GW distance, we observe that the bottom left of the source is aligned with the top right of the target. This is due to the fact that the GW distance focuses on the intrinsic relations among nodes in the Morse graph and is unaware of the symmetry within the dataset. Therefore, the GW distance may result in flipped or rotated alignments.

On the other hand, the partial OT-type distances aim to be more robust by relaxing the couplings and thereby ignoring certain noisy features. As shown in Fig. 2 (bottom right), certain nodes in the target are hollow, which are ignored by partial OT during optimization. In particular, using the pFGW distance, a number of 1-cells in the target (indicated by orange arrows) that deviate from the source are ignored in the alignment, indicating its robustness against noise.

Both pW and pGW distances are able to ignore subsets of the noisy features. The result from pW is slightly worse than that of pFGW. For example, in the relaxed coupling of pW, a node (indicated by the orange arrow) becomes isolated because its neighboring nodes are ignored. This is because the pW distance does not consider the intrinsic relations between nodes in the Morse graph. The pGW distance, on the other hand, results in an upside-down flipped alignment. Overall, for the Sinusoidal dataset, the pFGW distance performs the best for feature correspondences, followed by the pW distance.

5.2 Real-World Datasets: Feature Correspondences

We now examine the OT-type distances using real-world datasets in Fig. 1. For a time-varying dataset, we use an earlier time step as the source and a later time step as the target, that is, time steps 1 and 10 for Wind [37], 800 and 879 for HeatedCylinder [25], and 60 and 63 for NavierStokes [4] datasets, respectively. For the ensemble dataset RedSea [14], we compare ensemble members 1 and 4.

Wind dataset. For full OT-type distances, the Wasserstein and the FGW distances align the global features reasonably well, resulting in

a smooth color transfer between the source and the target. The GW distance produces a top-down flipped alignment due to symmetry. Four noticeable structural differences exist between the source and target (indicated by orange arrows). Among the partial OT-type distances, pFGW distance performs the most robustly as it ignores these four differences by treating them as noisy features. The pW distance ignores all four features but also ignores two more nodes on the top boundary. Again, pGW suffers from symmetry in the data. In summary, partial OT is useful when we are interested in capturing global similarities while ignoring local differences.

Heated Cylinder dataset. The differences between the source and the target are minor. Both Wasserstein and FGW distances produce reasonable alignments, whereas the GW distance produces some noticeable misalignments in the center of the domain (indicated by an orange arrow). Using partial OT, the pFGW distance performs the best by ignoring the additional 1-cell in the target during alignment (indicated by orange arrows). In comparison, the pW and pGW distances perform less robustly than the pFGW distance. For instance, some 1-cells are incorrectly ignored, also possibly due to inconsistent sampling density between the source and the target.

Navier Stokes dataset. Comparing the source and the target, we notice one additional loop on the top right corner of the target. Both Wasserstein and FGW distances match such an additional loop in the target to a bottom-right loop in the source (indicated by orange arrows). This is undesirable but understandable, as these two distances are forced to match the full masses on the nodes between the source and the target. With partial OT, the pW and pFGW distances could ignore the additional loop in the coupling matrices. The pGW distance also ignores a part of the loop.

Red Sea dataset. The RedSea dataset is an ensemble dataset. As a result, the target has a very different structure from the source. Nevertheless, we aim to align these two Morse graphs as much as possible. Results using the Wasserstein and the FGW distances are similar. While it is trickier to tune the parameter for partial OT (in comparison with time-varying datasets), we observe that pW and pFGW distances ignore 1-cells at the center of the target to form bigger holes to better align with the source (that contains bigger holes and fewer loops).

5.3 Classification: Wasserstein vs. GW Distances

We observe that the Wasserstein distance typically produces better alignments than the GW distance for the evolving scalar field datasets in Sec. 5.2. We now discuss potential scenarios when the GW distance outperforms the Wasserstein distance.

We introduce two synthetic datasets for classification tasks. For the RotatingGaussian dataset, we generate a mixture of two Gaussian functions with equal bandwidth and rotate the mixture at 100 evenly sampled angles (referred to as the *Rotating Binary Gaussians*); similarly, we generate 100 rotations of a mixture of three Gaussian functions (referred to as the *Rotating Trinary Gaussians*). For the RandomGaussian dataset, we generate 100 instances of randomly generated mixtures of two and three Gaussian functions with random centers and varying bandwidths, referred to as *Random Binary Gaussians* and *Random Trinary Gaussians*, respectively. We introduce random noises to increase the complexity of each dataset.

We apply multi-dimensional scaling (MDS) [3] to visualize the GW, FGW, and Wasserstein distances for each dataset, as shown Fig. 3. Blue points represent data instances from Binary Gaussians while orange points are from Trinary Gaussians. For the RotatingGaussian dataset, all three distances produce visible clusters and clearly separate the two classes. The periodic behaviors of the dataset are clearly visible as loops from the MDS projection using the Wasserstein distance as it is sensitive to geometry. For the RandomGaussian dataset, we still observe some clustering effects using GW and FGW distances, but without clear separations between the classes.

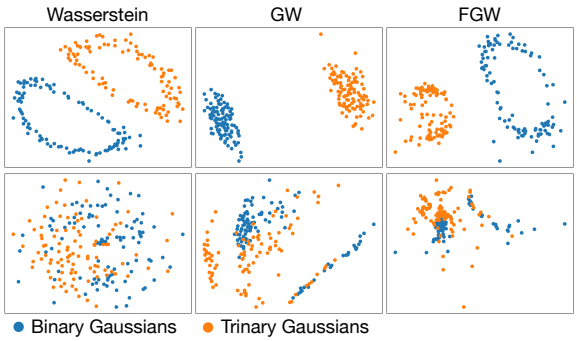

Figure 3: MDS projections for the RotatingGaussian (top) and RandomGaussian (bottom) datasets.

We then apply a k-Nearest Neighbors (kNN) classifier to explore the classification accuracies using these distances. For each dataset, we use an 80-20 split of training and testing data. We apply cross-validation to set $k = 3$. We report the test accuracy and the F1-score for each class for all three distances, as shown in Tab. 1. We observe better classification results using a GW distance in comparison with the Wasserstein distance. The property that GW distances are insensitive to the geometric locations and solely measure the difference between intrinsic graph structures becomes an advantage in this specific task.

Dataset	Distance	Accuracy	F1-score Binary Gaussians	F1-score Trinary Gaussians
RotatingGaussian	Wasserstein	1.00	1.00	1.00
	GW	1.00	1.00	1.00
	FGW	1.00	1.00	1.00
RandomGaussian	Wasserstein	0.70	0.76	0.60
	GW	**0.85**	**0.86**	**0.83**
	FGW	0.70	0.77	0.57

Table 1: Test accuracy and F1-score for kNN classifiers ($k = 3$) using the Wasserstein, GW, and FGW distances.

6 Guidelines and Future Work

The Wasserstein distance optimizes the coupling by minimizing the Euclidean distances between matched points, whereas the GW distance solely preserves the intrinsic node relations. Combining these two components, the FGW distance preserves both Euclidean proximity and intrinsic similarity during alignment. Both the Wasserstein and FGW distances perform reasonably well in aligning the global structures between the source and the target. In comparison, the GW distance is less robust due to a lack of geometric information to "anchor" the aligned regions. If full structural alignment is required, then the FGW distance is typically the top choice for feature correspondence and comparison tasks, followed by the Wasserstein distance. If intrinsic graph structures are more important than extrinsic geometry, such as for certain classification tasks, then the GW distance outperforms the distances with a Wasserstein component.

On the other hand, distances based on partial OT are generally more robust to noise and outliers than their full OT counterparts. If partial structural alignment is allowed, the pFGW distance typically performs the best for feature correspondences, followed by the pW distance. Aside from being flexible, partial OT also helps highlight the structural differences between the source and the target.

Using time-varying datasets, these OT-type distances can be used in feature tracking. In particular, partial OT may be used to detect structural changes over time. Our comparative analysis could be easily extended to study Morse-Smale complexes and their variants such as the extremal graphs. This is left for future work.

Acknowledgments

This project was partially funded by NSF IIS-2145499, NSF IIS-1910733, NSF DMS-2107808, and DOE DE-SC0021015.

REFERENCES

[1] J. P. Ahrens, B. Geveci, and C. C. Law. ParaView: An end-user tool for large-data visualization. In *The Visualization Handbook*, 2005. 2

[2] T. Athawale, D. Maljovec, L. Yan, C. R. Johnson, V. Pascucci, and B. Wang. Uncertainty visualization of 2D Morse complex ensembles using statistical summary maps. *IEEE Transactions on Visualization and Computer Graphics*, 28(4):1955–1966, 2022. 1

[3] I. Borg and P. J. Groenen. *Modern multidimensional scaling: Theory and applications*. Springer Science & Business Media, 2005. 4

[4] S. Camarri, M. Buffoni, A. Iollo, and M. V. Salvetti. Simulation of the three-dimensional flow around a square cylinder between parallel walls at moderate Reynolds numbers. In *XVII Congresso di Meccanica Teorica ed Applicata*, 2005. 3

[5] L. Chapel, M. Z. Alaya, and G. Gasso. Partial optimal tranport with applications on positive-unlabeled learning. *Advances in Neural Information Processing Systems*, 33:2903–2913, 2020. 1, 2

[6] S. Chowdhury and F. Mémoli. The Gromov-Wasserstein distance between networks and stable network invariants. *Information and Inference: A Journal of the IMA*, 8(4):757–787, 2019. 1, 2

[7] S. Chowdhury, D. Miller, and T. Needham. Quantized Gromov-Wasserstein. In *Proceedings of the Machine Learning and Knowledge Discovery in Databases*, pages 811–827, 2021. 1

[8] S. Chowdhury and T. Needham. Gromov-Wasserstein averaging in a Riemannian framework. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 842–843, 2020. 1

[9] S. Chowdhury and T. Needham. Generalized spectral clustering via Gromov-Wasserstein learning. In *International Conference on Artificial Intelligence and Statistics*, pages 712–720, 2021. 1

[10] J. Curry, H. Hang, W. Mio, T. Needham, and O. B. Okutan. Decorated merge trees for persistent topology. *Journal of Applied and Computational Topology*, 6(3):371–428, 2022. 1

[11] J. Curry, W. Mio, T. Needham, O. B. Okutan, and F. Russold. Topologically attributed graphs for shape discrimination. *arXiv preprint arXiv:2306.17805*, 2023. 1

[12] W. Feng, J. Huang, T. Ju, and H. Bao. Feature correspondences using Morse Smale complex. *The Visual Computer*, 29(1):53–67, 2013. 1

[13] R. Flamary, N. Courty, A. Gramfort, M. Z. Alaya, A. Boisbunon, S. Chambon, L. Chapel, A. Corenflos, K. Fatras, N. Fournier, L. Gautheron, N. T. Gayraud, H. Janati, A. Rakotomamonjy, I. Redko, A. Rolet, A. Schutz, V. Seguy, D. J. Sutherland, R. Tavenard, A. Tong, and T. Vayer. POT: Python optimal transport. *Journal of Machine Learning Research*, 22(78):1–8, 2021. 2

[14] I. Hoteit. The IEEE SciVis Contest. https://kaust-vislab.github.io/SciVis2020/, 2020. 3

[15] J. Kasten, J. Reininghaus, I. Hotz, and H.-C. Hege. Two-dimensional time-dependent vortex regions based on the acceleration magnitude. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2080–2087, 2011. 1

[16] A. Kuhn, W. Engelke, M. Flatken, H.-C. Hege, and I. Hotz. Topology-based analysis for multimodal atmospheric data of volcano eruptions. *In Topological Methods in Data Analysis and Visualization V: Theory, Algorithms, and Applications (Proceedings of TopoInVis 2017)*, 2017. 1

[17] M. Li, S. Palande, L. Yan, and B. Wang. Sketching merge trees for scientific data visualization. *arXiv preprint arXiv:2101.03196*, 2021. 1

[18] M. Li, C. Storm, A. Y. Li, T. Needham, and B. Wang. Optimal transport distances for Morse complexes. https://github.com/tdavislab/MCOpt, 2023. 2

[19] M. Li, X. Yan, L. Yan, T. Needham, and B. Wang. Flexible and probabilistic topology tracking with partial optimal transport. arXiv preprint arXiv:2302.02895, 2023. 1, 2

[20] F. Mémoli. On the use of Gromov-Hausdorff distances for shape comparison. *Eurographics Symposium on Point-Based Graphics*, pages 81–90, 2007. 1, 2

[21] F. Mémoli. Spectral Gromov-Wasserstein distances for shape matching. In *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, pages 256–263. IEEE, 2009. 1

[22] F. Mémoli. Gromov-Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics*, 11(4):417–487, 2011. 1

[23] V. Narayanan, D. M. Thomas, and V. Natarajan. Distance between extremum graphs. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, pages 263–270, 2015. 1

[24] G. Peyré, M. Cuturi, and J. Solomon. Gromov-Wasserstein averaging of kernel and distance matrices. *Proceedings of the 33rd International Conference on Machine Learning, PMLR*, 48:2664–2672, 2016. 1

[25] S. Popinet. Free computational fluid dynamics. *ClusterWorld*, 2(6), 2004. 3

[26] J. Reininghaus, J. Kasten, T. Weinkauf, and I. Hotz. Efficient computation of combinatorial feature flow fields. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1563–1573, 2012. 1

[27] B. Rieck, F. Sadlo, and H. Leitte. Persistence concepts for 2D skeleton evolution analysis. In *Topological Methods in Data Analysis and Visualization V: Theory, Algorithms, and Applications*. Springer, 2020. 1

[28] A. Schnorr, D. N. Helmrich, H. Childs, T. W. Kuhlen, and B. Hentschel. Feature tracking utilizing a maximum-weight independent set problem. In *IEEE 9th Symposium on Large Data Analysis and Visualization (LDAV)*, pages 6–15, Vancouver, BC, Canada, 2019. 1

[29] A. Schnorr, D. N. Helmrich, D. Denker, T. W. Kuhlen, and B. Hentschel. Feature tracking by two-step optimization. *IEEE Transactions on Visualization and Computer Graphics*, 26(6), 2020. 1

[30] T. Séjourné, F.-X. Vialard, and G. Peyré. The unbalanced Gromov-Wasserstein distance: Conic formulation and relaxation. *Advances in Neural Information Processing Systems*, 34:8766–8779, 2021. 1

[31] D. M. Thomas and V. Natarajan. Detecting symmetry in scalar fields using augmented extremum graphs. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2663–2672, 2013. 1

[32] J. Tierny, G. Favelier, J. A. Levine, C. Gueunet, and M. Michaux. The Topology ToolKit. https://topology-tool-kit.github.io/, 2021. 2

[33] T. Vayer, L. Chapel, R. Flamary, R. Tavenard, and N. Courty. Fused Gromov-Wasserstein distance for structured objects. *Algorithms*, 13(9):212, 2020. 1, 2

[34] T. Vayer, N. Courty, R. Tavenard, and R. Flamary. Optimal transport for structured data with application on graphs. *International Conference on Machine Learning*, pages 6275–6284, 2019. 1, 2

[35] C. Villani. *Optimal Transport: Old and New*, volume 338. Springer, 2009. 1, 2

[36] C. Vincent-Cuaz, R. Flamary, M. Corneli, T. Vayer, and N. Courty. Semi-relaxed Gromov-Wasserstein divergence with applications on graphs. In *International Conference on Learning Representations*, 2022. 1

[37] F. Vitart, C. Ardilouze, A. Bonet, A. Brookshaw, M. Chen, C. Codorean, M. Déqué, L. Ferranti, E. Fucile, M. Fuentes, et al. The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, 98(1):163 – 173, 2017. 3

[38] H. Xu, J. Liu, D. Luo, and L. Carin. Representing graphs via Gromov-Wasserstein factorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1):999–1016, 2022. 1

[39] H. Xu, D. Luo, and L. Carin. Scalable Gromov-Wasserstein learning for graph partitioning and matching. *Advances in Neural Information Processing Systems*, pages 3046–3056, 2019. 1

[40] H. Xu, D. Luo, H. Zha, and L. Carin. Gromov-Wasserstein learning for graph matching and node embedding. *International Conference on Machine Learning*, pages 6932–6941, 2019. 1, 2

[41] L. Yan, T. B. Masood, R. Sridharamurthy, F. Rasheed, V. Natarajan, I. Hotz, and B. Wang. Scalar field comparison with topological descriptors: Properties and applications for scientific visualization. *Computer Graphics Forum*, 40(3):599–633, 2021. 1