

# Perturbation Robust Representations of Topological Persistence Diagrams\*

Anirudh Som<sup>1</sup>\*[0000-0001-7595-3146], Kowshik Thopalli<sup>1</sup>\*[0000-0003-2183-8577],  
Karthikeyan Natesan Ramamurthy<sup>2</sup>[0000-0002-6021-5930],  
Vinay Venkataraman<sup>1</sup>[0000-0001-5601-4549], Ankita Shukla<sup>3</sup>[0000-0002-1878-2667],  
Pavan Turaga<sup>1</sup>[0000-0002-5263-5943]

<sup>1</sup> Geometric Media Lab, Arizona State University  
{asom2,kthopall1,vvenka18,pturaga}@asu.edu

<sup>2</sup> IBM Thomas J. Watson Research Center, knatesa@us.ibm.com

<sup>3</sup> Indraprastha Institute of Information Technology-Delhi, ankitas@iiitd.ac.in

**Abstract.** Topological methods for data analysis present opportunities for enforcing certain invariances of broad interest in computer vision, including view-point in activity analysis, articulation in shape analysis, and measurement invariance in non-linear dynamical modeling. The increasing success of these methods is attributed to the complementary information that topology provides, as well as availability of tools for computing topological summaries such as persistence diagrams. However, persistence diagrams are multi-sets of points and hence it is not straightforward to fuse them with features used for contemporary machine learning tools like deep-nets. In this paper we present theoretically well-grounded approaches to develop novel perturbation robust topological representations, with the long-term view of making them amenable to fusion with contemporary learning architectures. We term the proposed representation as Perturbed Topological Signatures, which live on a Grassmann manifold and hence can be efficiently used in machine learning pipelines. We explore the use of the proposed descriptor on three applications: 3D shape analysis, view-invariant activity analysis, and non-linear dynamical modeling. We show favorable results in both high-level recognition performance and time-complexity when compared to other baseline methods.

**Keywords:** Invariance Learning, Topological Data Analysis, Persistence Diagrams, Grassmann Manifold, Perturbed Topological Signature.

## 1 Introduction

Over the years, tools from topological data analysis (TDA) have been used to characterize the invariant structure of data obtained from a noisy sampling of an underlying metric space [24]. Invariance learning is a fundamental problem in computer vision, since common transformations can diminish the performance

---

\*The first two authors contributed equally.

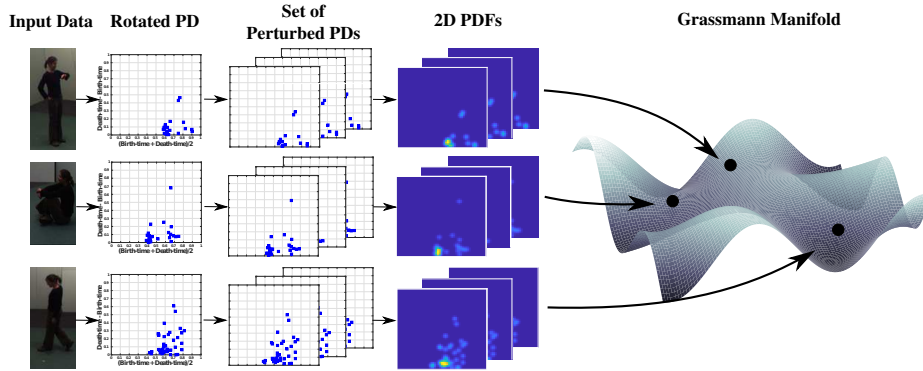


Fig. 1: Illustration of the sequence of steps leading to the proposed Perturbed Topological Signature (PTS) representation. For a given input dataset, the PDs are computed and transformed to maximally occupy the 2D space. A set of perturbed PDs is created, with each perturbed PD having its points displaced by a certain amount about its initial position. For each PD in the set, a 2D PDF is constructed using a Gaussian kernel function via kernel density estimation. The set of 2D PDFs capture a wide range of topological noise for the given input data and are summarized using a subspace structure, equivalent to a point on the Grassmann manifold.

of algorithms significantly. Past work in invariance learning has fallen into one of two classes. The first approach involves ad-hoc choices of features or metrics between features that offer some invariance to specific factors [9]. However, this approach has suffered due to lack of generalizable solutions. The other approach is to increase the training size by collecting samples that capture all the variations of the data, so that the learning algorithm can implicitly marginalize out the variations. A similar effect can be achieved via simple data augmentation [50].

In this context, TDA has emerged as a surprisingly powerful tool to analyze underlying invariant properties of data before any contextual modeling assumptions or the need to extract actionable information kicks in. Generally speaking, TDA seeks to characterize the *shape* of high dimensional data by quantifying various topological invariants such as connected components, cycles, high-dimensional holes, level-sets and monotonic regions of functions defined on the data [24]. Topological invariants are those properties that do not change under smooth deformations like stretching, bending, and rotation, but without tearing or gluing surfaces. We illustrate the connections between topological invariants and learning invariant representations for vision via three applications:

**1) Point cloud shape analysis:** Shape analysis of 3-dimensional (3D) point cloud data is a topic of major current interest due to emergence of Light Detection and Ranging (LIDAR) based vision systems in autonomous vehicles. It has been a difficult problem to solve with contemporary methods (*e.g.* deep learning) due to the non-vectorial nature of the representations. While there is interest in

trying to extend deep-net architectures to point-cloud data [53, 44, 72, 46, 32], the invariance one seeks is that of *shape articulation*, *i.e.* stretching, skewing, rotation of shape that does not alter the fundamental object class. These invariances are optimally defined in terms of topological invariants.

**2) Video analysis:** One of the long-standing problems in video analysis, specific to human action recognition, is to deal with variation in *body type*, *execution style*, and *view-point* changes. Work in this area has shown that temporal self-similarity matrices (SSMs) are a robust feature and provide general invariance to the above factors [34]. Temporal self-similarities can be quantified by scalar field topological constructions defined over video features, leading to representations with encoded invariances not relying on brute-force training data.

**3) Non-linear dynamical modeling:** Many time-series analysis problems have been studied under the lens of non-linear dynamical modeling; including motion-capture analysis, wearable-based activity analysis *etc.* Results from dynamical systems theory (Takens’ embedding theorem [62]) suggest that the placement-invariant property may be related to the topological properties of reconstructed dynamical attractors via delay-embeddings.

One of the prominent TDA tools is persistent homology. It provides a multi-scale summary of different homological features [25]. This multi-scale information is represented using a persistence diagram (PD), a 2-dimensional (2D) Cartesian plane with a multi-set of points. For a point  $(b, d)$  in the PD, a homological feature appears at scale  $b$  and disappears at scale  $d$ . Due to the simplicity of PDs, there has been a surge of interest to use persistent homology for summarizing high-dimensional complex data and has resulted in its successful implementation in several research areas [49, 63, 14, 19, 15, 31, 57, 66]. However, application of machine learning (ML) techniques on the space of PDs has always been a challenging task. The gold-standard approach for measuring the distance between PDs is the *Bottleneck* or the *p-Wasserstein* metric [45, 65]. However, a simple metric structure is not enough to use vector based ML tools such as support vector machines (SVMs), neural networks, random forests, decision trees, principal component analysis and so on. These metrics are only stable under small perturbations of the data which the PDs summarize, and the complexity of computing distances between PDs grows in the order of  $\mathcal{O}(n^3)$ , where  $n$  is the number of points in the PD [11]. Efforts have been made to overcome these problems by attempting to map PDs to spaces that are more suitable for ML tools [5, 12, 52, 48, 51, 3]. A comparison of some recent algorithms for machine learning over topological descriptors can be found in [54]. More recently, topological methods have also shown early promise in improving performance of image-based classification algorithms in conjunction with deep-learning [21].

**Contributions:** Using a novel perturbation framework, we propose a topological representation of PDs called *Perturbed Topological Signature* (PTS). To do this we first generate a set of perturbed PDs by randomly shifting the points in the original PD by a certain amount. A perturbed PD is analogous to extracting the PD from data that is subjected to topological noise. Next, we utilize a 2D probability density function (PDF) estimated by kernels on each of the per-

turbed PDs to generate a smooth functional representation. Finally, we simplify and summarize the end representation-space for the set of 2D PDFs to a point on the Grassmann manifold (a non-constantly curved manifold). The framework described above is illustrated in figure 1. We develop very efficient ML pipelines over these topological descriptors by leveraging the known metrics and statistical results on the Grassmann manifold. We also develop a stability proof of the Grassmannian representations w.r.t. the normalized geodesic distance over the Grassmannian and the *Wasserstein* metrics over PDs. Experiments show that our proposed framework recovers the lost performance due to functional methods, while still enjoying orders of magnitude faster processing times over the classical *p-Wasserstein* and *Bottleneck* approaches.

**Outline of the paper:** Section 2 provides the necessary background on topological data analysis and the Grassmannian. Section 3 discusses related work, while section 4 describes the proposed framework and end representation of the PD for statistical learning tasks. Section 5 describes the experiments and results. Section 6 concludes the paper.

## 2 Preliminaries

**Persistent Topology:** Consider a graph  $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$  on the high-dimensional point cloud, where  $\mathcal{V}$  is the set of  $|\mathcal{V}|$  nodes and  $\mathcal{E}$  defines the neighborhood relations between the samples. To estimate the topological properties of the graph’s shape, a simplicial complex  $\mathcal{S}$  is constructed over  $\mathcal{G}$ . We denote  $\mathcal{S} = (\mathcal{G}, \Sigma)$ , where  $\Sigma$  is a family of non-empty level sets of  $\mathcal{G}$ , with each element  $\sigma \in \Sigma$  is a simplex [25]. These simplices are constructed using the  $\epsilon$ -neighborhood rule,  $\epsilon$  being the scale parameter [25]. In TDA, Betti numbers  $\beta_i$  provide the rank of the homology group  $H_i$ . For instance,  $\beta_0$  denotes the number of connected components,  $\beta_1$  denotes the number of holes or loops,  $\beta_2$  denotes the number of voids or trapped volumes, *etc.* They provide a good summary of a shape’s topological features. However, two shapes with same Betti numbers can have very different PDs since PDs summarize the birth vs death time information of each topological feature in a homology group. Birth time ( $b$ ) signifies the scale at which the group is formed and death time ( $d$ ) is the scale at which it ceases to exist. The difference between the death and the birth times is the lifetime of the homology group  $l = |d - b|$ . Each PD is a multiset of points  $(b, d)$  in  $\mathbb{R}^2$  and is hence represented graphically as a set of points in a 2D plane. The diagonal where  $b = d$  is assumed to contain an infinite number of points since they correspond to groups of zero persistence.

We use the Vietoris-Rips (VR) construction denoted by  $\text{VR}(\mathcal{G}, \epsilon)$  to obtain simplicial complexes from  $\mathcal{G}$  for a given scale  $\epsilon$  [24]. An algorithm for computing homological persistence is provided in [25] and an efficient dual variant that uses co-homology is described in [20]. The VR construction obtains the topology of the distance function on the point cloud data. However, given a graph  $\mathcal{G}$ , and a function  $g$  defined on the vertices, it is also possible to quantify the topology induced by  $g$  on  $\mathcal{G}$ . For example, we may want to study the topology of the

sub-level or super-level sets of  $g$ . This is referred to as *scalar field topology* since  $g : \mathcal{V} \rightarrow \mathbb{R}$ . A well-known application of this in vision is in 3D shape data, where the graph  $\mathcal{G}$  corresponds to the shape mesh and  $g$  is a function, such as heat kernel signature (HKS) [60], defined on the mesh [40]. The PD of the  $H_0$  homology group of the super-level sets now describes the evolving segments of regions in the shape. For instance, if we compute the PD of the super-level sets induced by HKS in an *octopus* shape, we can expect to see eight highly persistent segments corresponding to the eight legs. This is because the HKS values are high at regions of high curvature in the shape. In scalar field constructions, the PDs can be obtained efficiently using the Union-Find algorithm by first sorting the nodes of  $\mathcal{G}$  as per their function magnitude and keeping a trail of the corresponding connected components [18].

**Distance Metrics between PDs:** PDs are invariant to rotations, translations and scaling of a given shape, and under continuous deformation conditions are invariant to slight permutations of the vertices [16, 17]. The two classical metrics to measure distances between PDs  $X$  and  $Y$  are the *Bottleneck* distance and the *p-Wasserstein* metric [45, 65]. They are appealing as they reflect any small changes such as perturbations of a measured phenomenon on the shape, which results in small shifts to the points in the persistence diagram. The *Bottleneck* distance is defined as  $d_\infty(X, Y) = \inf_{\eta: X \rightarrow Y} \sup_{x \in X} \|x - \eta(x)\|_\infty$ , with  $\eta$  ranging over all bijections and  $\|\cdot\|_\infty$  is the  $\infty$ -norm. Equivalently, the *p-Wasserstein* distance is defined as  $d_p(X, Y) = (\inf_{\eta: X \rightarrow Y} \sum_{x \in X} \|x - \eta(x)\|_\infty^p)^{1/p}$ . However, the complexity of computing distances between PDs with  $n$  points is  $\mathcal{O}(n^3)$ . These metrics also do not allow for easy computation of statistics and are unstable under large deformations [11].

**Grassmann Manifold:** Let  $n, p$  be two positive integers such that  $n > p > 0$ . The set of  $p$ -dimensional linear subspaces in  $\mathbb{R}^n$  is called a Grassmann manifold, denoted by  $\mathbb{G}_{p,n}$ . Each point  $\mathcal{Y}$  on  $\mathbb{G}_{p,n}$  is represented as a basis, *i.e.* a linear combination of the set of  $p$  orthonormal vectors  $Y_1, Y_2, \dots, Y_p$ . The geometric properties of the Grassmannian have been used for various computer vision applications, such as object recognition, shape analysis, human activity modeling and classification, face and video-based recognition, *etc* [9, 29, 64, 28]. We refer our readers to the following papers that provide a good introduction to the geometry, statistical analysis, and techniques for solving optimization problems on the Grassmann manifold [1, 23, 69, 13, 2].

**Distance Metrics between Grassmann Representations:** The minimal geodesic distance ( $d_G$ ) between two points  $\mathcal{Y}_1$  and  $\mathcal{Y}_2$  on the Grassmann manifold is the length of the shortest constant speed curve that connects these points. To do this, the velocity matrix  $A_{\mathcal{Y}_1, \mathcal{Y}_2}$  or the inverse exponential map needs to be calculated, with the geodesic path starting at  $\mathcal{Y}_1$  and ending at  $\mathcal{Y}_2$ .  $A_{\mathcal{Y}_1, \mathcal{Y}_2}$  can be computed using the numerical approximation method described in [42]. The geodesic distance between  $\mathcal{Y}_1$  and  $\mathcal{Y}_2$  is represented by the following equation:  $d_G(\mathcal{Y}_1, \mathcal{Y}_2) = \text{trace}(A_{\mathcal{Y}_1, \mathcal{Y}_2} A_{\mathcal{Y}_1, \mathcal{Y}_2}^T)$  or  $d_G(\mathcal{Y}_1, \mathcal{Y}_2) = \sqrt{\text{trace}(\theta^T \theta)}$ . Here  $\theta$  is the principal angle matrix between  $\mathcal{Y}_1, \mathcal{Y}_2$  and can be computed as  $\theta = \arccos(S)$ , where  $USV^T = \text{svd}(\mathcal{Y}_1^T \mathcal{Y}_2)$ . To show the stability of the proposed PTS representations in section 4, we use the normalized geodesic distance represented by  $d_{NG}(\mathcal{Y}_1, \mathcal{Y}_2) = \frac{1}{D} d_G(\mathcal{Y}_1, \mathcal{Y}_2)$ , where  $D$  is the maximum possible geodesic distance on  $\mathbb{G}_{p,n}$  [33, 39]. The symmetric directional distance ( $d_\Delta$ ) is another popular metric to compute distances between Grassmann representations with different  $p$  [61, 67]. It is a widely used measure in areas like computer vision [56, 8, 7, 43, 70], communications [55], and applied mathematics [22]. It is equivalent to the chordal metric [71]

and is defined as,  $d_{\Delta}(\mathcal{Y}_1, \mathcal{Y}_2) = (\max(k, l) - \sum_{i,j=1}^{k,l} (y_{1,i}^T y_{2,j})^2)^{\frac{1}{2}}$ . Here,  $k$  and  $l$  are subspace dimensions for the orthonormal matrices  $\mathcal{Y}_1$  and  $\mathcal{Y}_2$  respectively. For all our experiments, we restrict ourselves to distance computations between same-dimension subspaces, *i.e.*  $k = l$ . The following papers propose methods to compute distances between subspaces of different dimensions [61, 67, 71].

### 3 Prior Art

PDs provide a compact multi-scale summary of the different topological features. The traditional metrics used to measure the distance between PDs are the *Bottleneck* and *p-Wasserstein* metrics [45, 65]. These measures are stable with respect to small continuous deformations of the topology of the inputs [16, 17]. However, they do poorly under large deformations. Further, a feature vector representation will be useful that is compatible with different ML tools that demand more than just a metric. To address this need, researchers have resorted to transforming PDs to other suitable representations [5, 12, 52, 48, 51, 3]. Bubenik proposed persistence landscapes (PL) which are stable and invertible functional representations of PDs in a Banach space [12]. A PL is a sequence of envelope functions defined on the points in PDs that are ordered on the basis of their importance. Bubenik’s main motivation for defining PLs was to derive a unique mean representation for a set of PDs which is not necessarily obtained using Fréchet means [45]. Their usefulness is however limited, as PLs can provide low importance to moderate size homological features that generally possess high discriminating power.

Rouse *et al.* create a simple vector representation by overlaying a grid on top of the PD and count the number of points that fall into each bin [52]. This method is unstable, since a small shift in the points can result in a different feature representation. This idea has also appeared in other forms, some of which are described below. Pachauri *et al.* transform PDs into smooth surfaces by fitting Gaussians centered at each point in the PD [48]. Reininghaus *et al.* create stable representations by taking a weighted sum of positive Gaussians at each point above the diagonal and mirror the same below the diagonal but with negative Gaussians [51]. Adams *et al.* design persistence images (PI) by defining a regular grid and obtaining the integral of the Gaussian-surface representation over the bins defined on each grid [3]. Both PIs and the multi-scale kernel defined by Reininghaus *et al.* show stability with respect to the Wasserstein metrics and do well under small perturbations of the input data. They also weight the points using a weighting function, and this can be chosen based on the problem. Prioritizing points with medium lifetimes was used by Bendich *et al.* to best identify the age of a human brain by studying its arterial geometry [10]. Cohen-Steiner *et al.* suggested prioritizing points near the death-axis and away from the diagonal [16].

In this paper, we propose a unique perturbation framework that overcomes the need for selecting a weighting function. We consider a range of topological noise realizations one could expect to see, by perturbing the points in the PD. We summarize the perturbed PDs by creating smooth surfaces from them and consider a subspace of these surfaces, which naturally becomes a point on the Grassmann manifold. We show the effectiveness of our features in section 5 for different problems using data collected from different sensing devices. Compared to the *p-Wasserstein* and *Bottleneck* distances, the metrics defined on the Grassmannian are computationally less complex and the representations are independent of the number of points present in the PD. The proposed PTS representation is motivated from [28], where the authors create a

subspace representation of blurred faces and perform face recognition on the Grassmannian. Our framework also bears some similarities to [5], where the authors use the square root representation of PDFs obtained from PDs.

## 4 Perturbed Topological Signatures

In this section we go through details of each step in our framework’s pipeline, illustrated in figure 1. In our experiments we transform the axes of the PD from  $(b, d) \rightarrow (\frac{b+d}{2}, d-b)$ , with  $b \leq d$ .

**Create a set of Perturbed PDs:** We randomly perturb a given PD to create  $m$  PDs. Each of the perturbed PDs has its points randomly displaced by a certain amount compared to the original. The set of randomly perturbed PDs retain the same topological information of the input data as the original PD, but together capture all the probable variations of the input data when subjected to topological noise. We constrain the extent of perturbation of the individual points in the PD to ensure that the topological structure of the data being analyzed is not abruptly changed.

**Convert Perturbed PDs to 2D PDFs:** We transform the initial PD and its set of perturbed PDs to a set of 2D PDFs. We do this via kernel density estimation: by fitting a Gaussian kernel function with zero mean, standard deviation  $\sigma$  at each point in the PD, and then normalizing the 2D surface. The obtained PDF surface is discretized over a  $k \times k$  grid similar to the approach of Rouse *et al.* [52]. The standard deviation  $\sigma$  (also known as bandwidth parameter) of the Gaussian is not known a priori and is fine-tuned to get best results. A multi-scale approach can also be employed by generating multiple surfaces using a range of different bandwidth parameters for each of the PDs and still obtain favorable results. Unlike other topological descriptors that use a weighting function over their functional representations of PDs [51, 3], we give equal importance to each point in the PD and do not resort to any weighting function. Adams *et al.* prove the stability of persistence surfaces obtained using general and Gaussian distributions ( $\phi$ ), together with a weighting function ( $f$ ), with respect to the 1-*Wasserstein* distance between PDs in [3, Thm. 4, 9]. For Gaussian distributions, both  $L_1$  and  $L_\infty$  distances between persistence surfaces  $\rho_B, \rho_{B'}$  are stable with respect to 1-*Wasserstein* distance between PDs  $B, B'$ ,  $\|\rho_B - \rho_{B'}\|_1 \leq \sqrt{\frac{10}{\pi}} \frac{1}{\sigma} d_1(B, B')$ .

**Projecting 2D PDFs to the Grassmannian:** Let  $\rho(x, y)$  be an unperturbed persistence surface, and let  $\rho(x+u, y+v)$  be a randomly shifted perturbation. Under assumptions of small perturbations, we have using Taylor’s theorem:

$$\rho(x+u, y+v) - \rho(x, y) \approx [\rho_x, \rho_y][u, v]^T \quad (1)$$

Now, in the following, we interpret  $\approx$  as an equality, enabling us to stack together the same equation for all  $(x, y)$ , to get a matrix-vector form  $\overline{\rho_{pert}^{u,v}} - \overline{\rho} = [\overline{\rho_x}, \overline{\rho_y}]_{N \times 2} [u, v]_{2 \times 1}^T$ , where the overline indicates a discrete vectorization of the 2D functions. Here,  $N$  is the total number of discretized samples from the  $(x, y)$  plane. Now consider the set of all small perturbations of  $\rho$ , i.e.  $span(\overline{\rho_{pert}^{u,v}} - \overline{\rho})$ , over all  $[u, v] \in \mathbb{R}^2$ . It is easy to see that this set is just a 2D linear-subspace in  $\mathbb{R}^N$  which coincides with the column-span of  $[\overline{\rho_x}, \overline{\rho_y}]$ . For a more general affine-perturbation model, we can show that the required subspace corresponds to a 6-dimensional (6D) linear subspace, corresponding to the column-span of the  $N \times 6$  matrix  $[\rho_x, \rho_y, x\rho_x, x\rho_y, y\rho_x, y\rho_y]$ . More details on this can be found in the supplement. In implementation, we perturb a given PD several times using random offsets, compute their persistence surfaces, use singular value decomposition

(SVD) on the stacked matrix of perturbations, then select the  $p$  largest left singular vectors, resulting in a  $N \times p$  orthonormal matrix. Also, we vary the dimension of the subspace across a range of values. Since the linear span of our matrix can be further identified as a point on the Grassmann manifold, we adopt metrics defined over the Grassmannian to compare our perturbed topological signatures.

**Stability of Grassmannian metrics w.r.t. Wasserstein:** The next natural question to consider is whether the metrics over the Grassmannian for the perturbed stack are in any way related to the Wasserstein metric over the original PDs. Let the column span of  $X = [\bar{\rho}_x, \bar{\rho}_y]$  be represented by  $\mathcal{X}(\rho)$ . Let  $\rho_1, \rho_2$  be two persistence surfaces, then  $\mathcal{X}(\rho_1), \mathcal{X}(\rho_2)$  are the subspaces spanned by  $X_1 = [\bar{\rho}_{1,x}, \bar{\rho}_{1,y}]$  and  $X_2 = [\bar{\rho}_{2,x}, \bar{\rho}_{2,y}]$  respectively. Following a result due to Ji-Guang [33], the normalized geodesic distance  $d_{\text{NG}}$  between  $\mathcal{X}_1$  and  $\mathcal{X}_2$  is upper bounded as follows:  $d_{\text{NG}}(\mathcal{X}_1, \mathcal{X}_2) \leq \|X_1\|_F \cdot \|X_1^\dagger\|_2 \cdot \frac{\|\Delta X\|_F}{\|X_1\|_F} = \|X_1^\dagger\|_2 \cdot \|\Delta X\|_F$ . Here,  $\|X^\dagger\|_2$  is the spectral norm of the pseudo-inverse of  $X$ ,  $\|X\|_F$  is the Frobenius norm, and  $\Delta X = X_1 - X_2$ . In the supplement, a full derivation is given, showing  $\|\Delta X\|_F^2 \leq \frac{10}{\pi} \frac{2}{\sigma^6} d_1^2(B_1, B_2) + 2 \frac{\mathcal{K}^2}{\sigma^4} k_{max}^2 N$ , where  $d_1(B_1, B_2)$  is the 1-*Wasserstein* metric between the original unperturbed PDs,  $k_{max}$  is the maximum number of points in a given PD (a dataset dependent quantity),  $N$  refers to the total number of discrete samples from  $[0, 1]^2$  and  $\mathcal{K} = \frac{1}{(\sqrt{2\pi}\sigma)^2}$ . This is the critical part of the stability proof. The remaining part requires us to upper bound the spectral norm  $\|X^\dagger\|_2$ . The spectral-norm of the pseudo-inverse of  $X$ , i.e.  $\|X^\dagger\|_2$ , is simply the inverse of the smallest singular-value of  $X$ , which in turn corresponds to the square-root of the smallest eigenvalue of  $X^T X$ . i.e.  $\|X^\dagger\|_2 = \sigma_{max}(X^\dagger) = \frac{1}{\sigma_{min}(X)} = \frac{1}{\sqrt{\lambda_{min}(X^T X)}}$ .

Given that  $X = [\bar{\rho}_x, \bar{\rho}_y]$ ,  $X^T X$  becomes the 2D structure-tensor of a Gaussian mixture model (GMM). While we are not aware of any results that lower-bound the eigenvalues of a 2D GMMs structure-tensor, in the supplement we show an approach for 1D GMMs that indicates that the smallest eigenvalue can indeed be lower-bounded, if the standard-deviation  $\sigma$  is upper-bounded. For example, a non-trivial lower-bound is derived for  $\sigma < 1$  in the supplement. It is inversely proportional to the number of components in the GMM. We used  $\sigma = 0.0004$  for all our experiments. The approach in the supplement is shown for 1D GMMs, and we posit that a similar approach applies for the 2D case, but it is cumbersome. In empirical tests, we find that even for 2D GMMs defined over the grid  $[0, 1]^2$ , with  $0 < \sigma < 1$ , the spectral-norms are always upper-bounded. In general, we find  $\|X^\dagger\|_2 \leq k/\sqrt{g(\sigma)}$ , where  $g(\sigma)$  is a positive monotonically decreasing function of  $\sigma$  in the domain  $[0, 1]$ , and  $k$  is the number of components in the GMM (points in a given PD). If we denote  $k_{max}$  and  $\sigma_{max}$  as the maximum allowable number of components in the GMM (max points in any PD in given database) and the maximum standard deviation respectively, an upper bound readily develops. Thus, we have

$$d_{\text{NG}}(\mathcal{X}_1, \mathcal{X}_2) \leq \frac{k_{max}}{\sqrt{g(\sigma_{max})}} \sqrt{\frac{10}{\pi} \frac{2}{\sigma^6} d_1^2(B_1, B_2) + 2 \frac{\mathcal{K}^2}{\sigma^4} k_{max}^2 N} \quad (2)$$

Please refer to the supplement for detailed derivation and explanation of the various constants in the above bound. We note that even though the above shows that the normalized Grassmannian geodesic distance over the perturbed topological signatures is stable w.r.t the 1-*Wasserstein* metric over PDs, it still relies on knowledge of the maximum number of points in any given PD across the entire dataset  $k_{max}$ , and also on the sampling of the 2D grid.



## 5 Experiments

In this section we first show the robustness of the PTS descriptor to different levels of topological noise using a sample of shapes from the SHREC 2010 dataset [41]. We then test the proposed framework on three publicly available datasets: SHREC 2010 shape retrieval dataset [41], IXMAS multi-view video action dataset [68] and motion capture dataset [4]. We briefly go over the details of each dataset, and describe the experimental objectives and procedures followed. Finally, we show the computational benefits of comparing different PTS representations using the  $d_G$  and  $d_\Delta$  metrics, over the classical  $p$ -Wasserstein and *Bottleneck* metrics used between PDs.

### 5.1 Robustness to Topological Noise

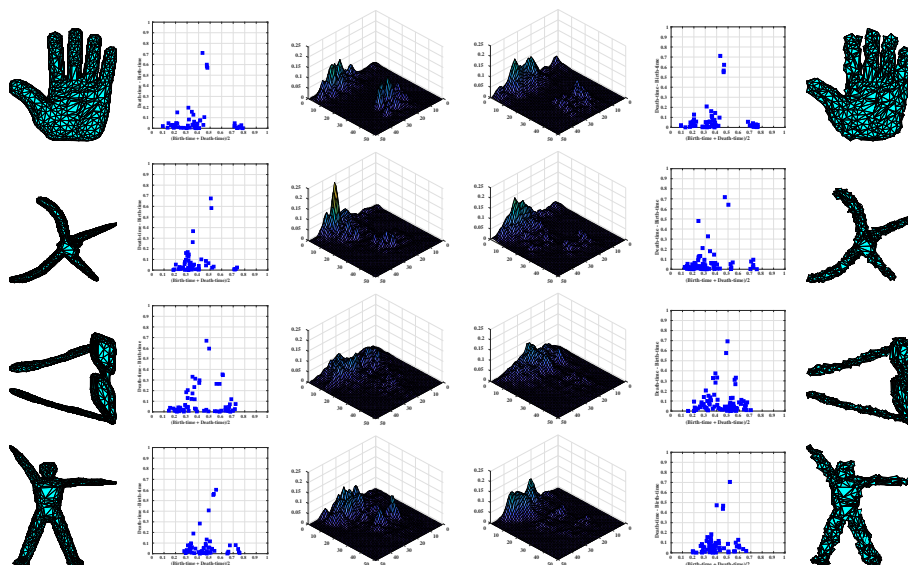


Fig. 2: Illustration of PD and PTS representations for 4 shapes and their noisy variants. Columns 1 and 6 represent the 3D shape with triangular mesh faces; columns 2 and 5 show the corresponding 9<sup>th</sup> dimension SIHKS function-based PDs. columns 3 and 4 depict the PTS feature of the PD for the original and noisy shapes respectively. A zero mean Gaussian noise with standard deviation 1.0 is applied on the original shapes in column 1 to get the corresponding noisy variant in column 6. The PTS representation shown is the largest left singular vector (reshaped to a 2D matrix) obtained after applying SVD on the set of 2D PDFs and lies on the  $\mathbb{G}_{1,n}$  space.

We conduct this experiment on 10 randomly chosen shapes from the SHREC 2010 dataset [41]. The dataset consists of 200 near-isometric watertight 3D shapes with articulating parts, equally divided into 10 classes. Each 3D mesh is simplified to 2000 faces. The 10 shapes used in the experiment are denoted as  $\mathcal{S}_i$ ,  $i = 1, 2, \dots, 10$ . The minimum bounding sphere for each of these shapes has a mean radius of 54.4 with standard deviation of 3.7 centered at (64.4, 63.4, 66.0) with coordinate-wise standard

deviations of (3.9, 4.1, 4.9) respectively. Next, we generate 100 sets of shapes, infused with topological noise. Topological noise is applied by changing the position of the vertices of the triangular mesh face, which results in changing its normal. We do this by applying a zero-mean Gaussian noise to the vertices of the original shape, with the standard deviation  $\sigma$  varied from 0.1 to 1 in steps of 0.1. For each shape  $\mathcal{S}_i$ , its 10 noisy shapes with different levels of topological noise are denoted by  $\mathcal{N}_{i,1}, \dots, \mathcal{N}_{i,10}$ .

A 17-dimensional scale-invariant heat kernel signature (SIHKS) spectral descriptor function is calculated on each shape [36], and PDs are extracted for each dimension of this function resulting in 17 PDs per shape. The PDs are passed through the proposed framework



Fig. 3: Sample SHREC 2010 shapes used to test robustness of PTS feature to topological noise.

to get the respective PTS descriptors. The 3D mesh, PD and PTS representation for 4 of the 10 shapes (shown in figure 3) and their respective noisy-variants (Gaussian noise with standard deviation 1.0) is shown in figure 2. In this experiment, we evaluate the robustness of our proposed feature by correctly classifying shapes with different levels of topological noise. Displacement of vertices by adding varying levels of topological noise, interclass similarities and intraclass variations of the shapes make this a challenging task. A simple unbiased one nearest neighbor (1-NN) classifier is used to classify the topological representations of the noisy shapes in each set. The classification results are averaged over the 100 sets and tabulated in table 1. We also compare our method to other TDA-ML methods like PI [3], PL [12], PSSK [51] and PWGK [38]. For PTS, we set the discretization of the grid  $k = 50$ . For PIs we chose the linear ramp weighting function, set  $k$  and  $\sigma$  for the Gaussian kernel function, same as our PTS feature. For PLs we use the first landscape function with 500 elements. A linear SVM classifier is used instead of the 1-NN classifier for the PSSK and PWGK methods. From table 1, the 2-*Wasserstein* and *Bottleneck* distances over PDs perform poorly even at low levels of topological noise. However, PDs with 1-*Wasserstein* distance and PTS representations with  $d_G$ ,  $d_\Delta$  metrics show stability and robustness to even high noise levels. *Nevertheless, the average time taken to compare two PTS features using either  $d_G$  or  $d_\Delta$  is at least two orders of magnitude faster than the 1-*Wasserstein* distance as seen in table 1. We also observe that compari-*

Method	$\mathcal{N}_{i,1}$	$\mathcal{N}_{i,2}$	$\mathcal{N}_{i,3}$	$\mathcal{N}_{i,4}$	$\mathcal{N}_{i,5}$	$\mathcal{N}_{i,6}$	$\mathcal{N}_{i,7}$	$\mathcal{N}_{i,8}$	$\mathcal{N}_{i,9}$	$\mathcal{N}_{i,10}$	Average Accuracy (%)	Average Time Taken ( $10^{-4}$ sec)
PD (1- <i>Wasserstein</i> )	100.00	100.00	100.00	99.90	100.00	99.80	99.60	99.00	96.60	94.40	98.93	256.00
PD (2- <i>Wasserstein</i> )	97.50	98.00	98.10	97.20	97.20	96.00	94.40	92.80	90.30	88.50	95.00	450.00
PD ( <i>Bottleneck</i> )	99.90	99.90	99.90	99.20	99.40	98.60	97.10	96.90	94.30	92.70	97.79	36.00
PI ( $L_1$ )	100.00	100.00	100.00	99.70	98.10	93.70	83.20	68.30	56.00	44.90	84.39	0.31
PI ( $L_2$ )	99.90	99.50	98.60	97.40	93.10	88.50	82.90	69.70	59.40	49.90	83.89	0.26
PI ( $L_\infty$ )	89.10	83.00	80.20	78.90	78.40	69.90	68.60	64.00	61.90	56.80	73.08	0.12
PL ( $L_1$ )	99.20	99.70	99.00	98.50	98.50	97.30	95.90	92.30	89.10	84.50	95.40	0.74
PL ( $L_2$ )	99.10	99.70	98.90	98.50	98.30	96.90	95.60	92.10	89.00	84.30	95.24	0.76
PL ( $L_\infty$ )	98.90	99.60	98.80	98.40	98.30	96.50	94.80	91.70	88.70	83.80	94.95	0.09
PSSK - SVM	100.00	100.00	100.00	100.00	100.00	100.00	91.60	90.00	89.80	89.00	96.04	4.55
PWGK - SVM	100.00	100.00	100.00	100.00	100.00	99.90	99.40	95.90	87.50	73.30	95.60	0.17
PTS ( $d_G$ )	100.00	100.00	100.00	100.00	100.00	99.90	99.80	98.80	96.80	93.60	<b>98.89</b>	<b>2.30</b>
PTS ( $d_\Delta$ )	100.00	100.00	100.00	100.00	100.00	99.90	99.90	99.30	97.10	94.10	<b>99.03</b>	<b>1.60</b>

Table 1: Comparison of 1-*Wasserstein*, 2-*Wasserstein*, *Bottleneck*,  $d_\Delta$  and  $d_G$  methods for correctly classifying the topological representations of noisy shapes to their original shape.

son of PIs, PLs and PWGK is an order of magnitude faster than comparing PTS features. However, these methods show significantly lower performance compared to the proposed feature, at correctly classifying noisy shapes as the noise level increases.

## 5.2 3D Shape Retrieval

In this experiment, we consider all 10 classes consisting of 200 shapes from the SHREC 2010 dataset, and extract PDs using 3 different spectral descriptor functions defined on each shape, namely: heat kernel signature (HKS) [60], wave kernel signature (WKS) [6], and SIHKS [36]. HKS and WKS are used to capture the microscopic and macroscopic properties of the 3D mesh surface, while SIHKS descriptor is the scale-invariant version of HKS.

Using the PTS descriptor we attempt to encode invariances to shape articulations such as rotation, stretching, skewing. For the task of 3D shape retrieval we use a 1-NN classifier to evaluate the performance of the PTS representation against other methods [12, 51, 3, 40, 38]. A linear SVM classifier is used to report the classification accuracy of the PSSK and PWGK methods. Li *et al.* report best results after carefully selecting weights to normalize the distance combinations of their BoF+PD and ISPM+PD methods. As in [40], we also use the three spectral descriptors and combine our PTS representations for each descriptor. PIs, PLs and PTS features are also designed the same way as described before. The results reported in table 2 show that the PTS feature (with subspace dimension  $p = 1$ ) alone using the  $d_{\Delta}$  metric achieves an accuracy of 99.50 %, outperforming other methods. The average classification result of the PTS feature on varying the subspace dimension  $p = 1, 2, \dots, 25$  is  $98.42 \pm 0.4$  % and  $98.72 \pm 0.25$  % using  $d_{\Delta}$  and  $d_G$  metrics respectively, thus displaying its stability with respect to the choice of  $p$ .

Method	BoF [40]	SSBoF [40]	ISPM [40]	PD (Bottleneck) [40]	PD (1-Wasserstein)	PD (2-Wasserstein)	BoF+PD [40]	ISPM+PD [40]	PI ( $L_1$ ) [3]	PI ( $L_2$ ) [3]	PI ( $L_{\infty}$ ) [3]	PL ( $L_1$ ) [12]	PL ( $L_2$ ) [12]	PL ( $L_{\infty}$ ) [12]	PSSK (SVM) [51]	PWGK (SVM) [38]	PTS ( $d_G$ )	PTS ( $d_{\Delta}$ )
1-NN Accuracy (%)	97.00	97.50	97.50	98.50	98.50	98.50	98.50	99.00	88.50	87.50	89.50	95.00	95.00	95.00	98.50	99.00	99.00	99.50

Table 2: Comparison of the classification performance of the proposed PTS descriptor with other baseline methods [40] on the SHREC 2010 dataset.

## 5.3 View-invariant Activity Analysis

The IXMAS dataset contains video and silhouette sequences of 11 action classes, performed 3 times by 10 subjects from five different camera views. The 11 classes are as follows - *check watch, cross arms, scratch head, sit down, get up, turn around, walk, wave, punch, kick, pick up*. Sample frames across 5 views for 2 actions are shown in figure 4. We consider only the silhouette information in the dataset for our PTS representations. For each frame in an action sequence we extract multi-scale shape distributions which are referred to as A3M, D1M, D2M and D3M, over the 2D silhouettes [58]. The multi-scale shape distribution feature captures the local to global changes in different geometric properties of a shape. For additional details about this feature, please see: [58, 59, 47].

For  $n$  frames in an action sequence and  $b$  bins in each shape distribution at a certain scale, an  $n \times b$  matrix representing the action is obtained. Treating the  $n$  frames as nodes, scalar field topological PDs are calculated across each column, resulting in  $b$  PDs. PDs capture the structural changes along each bin in the distributions. We select 5 different scales for the multi-scale shape features, giving us  $5b$  PDs per action which are passed through the proposed pipeline resulting in  $5b$  PTS features. PTS features try to encode the possible changes with respect to view-point variation, body-type and execution style. To represent the entire action as a point on the Grassmannian, we select the first two largest singular vectors from each of the  $5b$  PTS descriptors, apply SVD and choose 20 largest components.

To perform multi-view action recognition, we train non-linear SVMs using the Grassmannian RBF kernel,  $k_{rp}(\mathcal{X}_i, \mathcal{Y}_i) = \exp\left(-\beta \|\mathcal{X}_i^T \mathcal{Y}_i\|_F^2\right)$ ,  $\beta > 0$  [30]. Here,  $\mathcal{X}_i, \mathcal{Y}_i$  are points on the Grassmannian and  $\|\cdot\|_F$  is the Frobenius norm. We set  $\beta = 1$  in our implementations. Junejo *et al.* train non-linear SVMs using the  $\chi^2$  kernel over the SSM-based descriptors and follow a one-against-all approach for multi-class classification [34]. We follow the same approach and use a joint weighted kernel between their SSM kernel and our kernel, *i.e.*  $\chi^2 + \lambda \cdot k_{rp}$ , where  $\lambda = 0.1, 0.2, \dots, 1.0$ . The SSM-based descriptors are computed using the histogram of gradients (HOG), optical flow (OF) and fusion of HOG, OF features. The classification results are tabulated in table 3. Apart from reporting results of PTS representations obtained using the multi-scale shape distributions, we also show recognition results of PTS feature computed over the HOG descriptor (PTS-HOG). We see significant improvement in the results by fusing different PTS features with the SSM-based descriptor. We also tabulate the mean and standard deviation values for all classification results obtained after varying  $\lambda$  from 0.1 to 1.0 and subspace dimension  $p$  from 1 to 10. These results demonstrate the flexibility and stability associated with the proposed PTS topological descriptor.

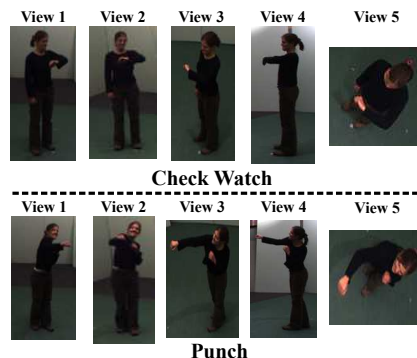


Fig. 4: Sample frames for “check watch” and “punch” action sequences from five views in the IX-MAS dataset.

Method	Same Camera Accuracy (%)		Any-To-Any Accuracy (%)	
	Best	Mean±SD	Best	Mean±SD
SSM-HOG [34]	67.30	-	52.60	-
PTS-HOG	51.31	-	41.24	-
SSM-HOG + PTS-HOG	69.01	-	55.13	-
SSM-HOG + PTS-A3M	73.15	72.06±1.14	58.36	56.96±1.05
SSM-HOG + PTS-D1M	74.25	73.26±1.53	59.26	57.67±1.19
SSM-HOG + PTS-D2M	74.92	74.22±1.36	59.77	58.19±1.03
<b>SSM-HOG + PTS-D3M</b>	<b>76.18</b>	<b>73.72±1.13</b>	<b>60.33</b>	<b>58.72±1.11</b>
SSM-OF [34]	66.60	-	53.80	-
SSM-OF + PTS-A3M	72.02	70.25±1.06	58.85	57.48±0.93
SSM-OF + PTS-D1M	73.67	71.62±1.17	59.56	57.81±1.05
SSM-OF + PTS-D2M	73.45	72.53±1.12	60.60	59.05±1.11
<b>SSM-OF + PTS-D3M</b>	<b>74.41</b>	<b>72.21±1.03</b>	<b>61.51</b>	<b>59.33±1.13</b>
SSM-HOG-OF [34]	76.28	-	61.25	-
SSM-HOG-OF + PTS-A3M	79.30	78.05±0.71	64.93	63.58±0.65
SSM-HOG-OF + PTS-D1M	79.61	79.03±0.96	65.39	64.27±0.65
SSM-HOG-OF + PTS-D2M	79.86	79.35±0.76	65.70	64.62±0.83
<b>SSM-HOG-OF + PTS-D3M</b>	<b>81.12</b>	<b>79.49±0.99</b>	<b>66.16</b>	<b>64.99±0.79</b>

Table 3: Comparison of the recognition results on the IXMAS dataset. Results are presented for two combinations of train camera  $X$  and test camera  $Y$ . “Same Camera” denotes  $X=Y$ ; “Any-To-Any” implies any combination of  $X, Y$ .

#### 5.4 Dynamical Analysis on Motion Capture Data

This dataset consists of human body joint motion capture sequences in 3D, where each sequence contains 57 trajectories (19 joint trajectories along 3 axes). There are 5 action classes - *dance, jump, run, sit and walk*, with each class containing 31, 14, 30, 35 and 48 sequences respectively.  $H_1$  homology group PDs are computed over the reconstructed attractor for each trajectory, resulting in 57 PDs per action [5] and the corresponding PTS feature is also extracted. We report the average classification performance over 100 random splits, with each split having 25 random test samples (5 samples from each class) and remaining 133 training samples. For SVM classification, we train non-linear SVMs using the projection kernel,  $k_p(\mathcal{X}_i, \mathcal{Y}_i) = \|\mathcal{X}_i^T \mathcal{Y}_i\|_F^2$  [29].

The results are tabulated in table 4. PTS features have a classification accuracy of 85.96 % and 91.92 % using the 1-NN and SVM classifier respectively. While these results are slightly lower than the 1-*Wasserstein* metric, the proposed descriptor with the  $d_\Delta$  metric is more than 2 orders of magnitude faster. Topological properties of dynamic attractors for analysis of time-series data has been studied and applied to tasks such as wheeze detection [27], pulse pressure wave analysis [26] and such applications are surveyed in [37]. We ask our readers to refer to these papers for further exploration.

Method	Accuracy (%)	Average Time Taken ( $10^{-4}$ sec)
PD (1-Wasserstein) NN [73]	93.68	22.00
Hilbert Sphere NN [5]	89.87	590.00
Hilbert Sphere PGA+SVM [5]	91.68	-
<b>PTS (<math>d_\Delta</math>) - NN</b>	85.96	0.19
<b>PTS - SVM</b>	91.92	-

Table 4: Comparison of classification performance and the average time taken to compare two feature representations on the motion capture dataset.

#### 5.5 Time-complexity of Comparing Topological Representations

All experiments are carried out on a standard Intel i7 CPU using Matlab 2016b with a working memory of 32 GB. We used the Hungarian algorithm to compute the *Bottleneck* and *p-Wasserstein* distances between PDs. Kerber *et al.* take advantage of the geometric structure of the input graph and propose geometric variants of the above metrics, thereby showing significant improvements in runtime performance when comparing PDs having several thousand points [35]. However, extracting PDs for most real datasets of interest in this paper does not result in more than a few hundred points. For example, on average we observe 71, 23, 27 points in each PD for the SHREC 2010,

Dataset	Average Number of Points in PD	Average Time Taken ( $10^{-4}$ sec)					Subspace Dimension ( $p$ ) of PTS Feature
		1-Wasserstein	2-Wasserstein	Bottleneck	$d_G$	$d_\Delta$	
SHREC 2010 [41]	71	256.00 (Kerber <i>et al.</i> [35]: 219.00)	450.00 (Kerber <i>et al.</i> [35]: 237.00)	36.00 (Kerber <i>et al.</i> [35]: 295.00)	<b>2.30</b>	<b>1.60</b>	10
IXMAS [68]	23	16.00	16.00	3.43	<b>2.21</b>	<b>0.68</b>	20
Motion Capture [4]	27	22.00	22.00	2.72	<b>0.24</b>	<b>0.19</b>	1

Table 5: Comparison of the average time taken to measure distance between two PDs using the 1-Wasserstein, 2-Wasserstein and Bottleneck metrics, and between two PTS features using  $d_G$  and  $d_\Delta$  metrics. The time reported is averaged over 3000 distance calculations between the respective topological representations for all three datasets used in section 5.

IXMAS and motion capture datasets respectively. The Hungarian algorithm incurs similar computations in this setting as shown in table 5. The  $d_G$  and  $d_\Delta$  metrics used to compare different PTS representations (grid size  $k = 50$ ) are fast and computationally less complex compared to the *Bottleneck* and *p-Wasserstein* distance measures. The average time taken to compare two topological signatures (PD or PTS) for each of the datasets is tabulated in table 5. The table also shows the average number of points seen per PD and the subspace dimension  $p$  used for the PTS representation.

Table 6 shows the variation of the average time taken to compare PTS features on varying the grid size ( $k$ ) of the 2D PDF. Here too the average time is reported after averaging over 3000 distance calculations between PTS features computed from PDs of the SHREC 2010 dataset. We observe that the time taken to

	Average Time Taken ( $10^{-4}$ sec)										
Grid size ( $k$ )	5	10	20	40	60	80	100	200	300	400	500
PTS ( $d_G$ )	0.72	0.73	0.89	1.31	1.48	2.28	5.53	8.35	18.40	32.88	47.07
PTS ( $d_\Delta$ )	0.20	0.33	0.84	0.72	1.00	1.85	4.32	7.70	17.69	31.56	46.68

Table 6: Comparison of the average time taken to measure distance between two PTS features using  $d_G$  and  $d_\Delta$  metrics w.r.t. variation in grid size  $k$ . The time reported is averaged over 3000 distance calculations between the topological representations for the SHREC 2010 dataset.

compare two PTS features with a grid size  $k = 500$  is two orders of magnitude greater than the time obtained for PTS features using  $k = 5$ . However, these times are still smaller than or on par with the times reported using *p-Wasserstein* and *Bottleneck* distances between PDs as seen in table 5. For all our experiments we set  $k = 50$  for our PTS representations and as shown in table 5, the times reported for  $d_\Delta$  and  $d_G$  are at least an order of magnitude faster than *Bottleneck* distance and two orders of magnitude faster than the *p-Wasserstein* metrics.

## 6 Conclusion and Discussion

We believe that a perturbed realization of a PD computed over a high-dimensional shape/graph is robust to topological noise affecting the original shape. Based on the type of data and application, topological noise can imply different types of variations, such as: articulation in 3D shape point cloud data; diversity in body structure, execution style and view-point pertaining to human actions in video analysis, *etc.* In this paper, we propose a novel topological representations called PTS that is obtained using a perturbation approach, taking first steps towards robust invariant learning with topological features. We obtained perturbed persistence surfaces and summarized them as a point on the Grassmann manifold, in order to utilize the different distance metrics and Mercer kernels defined for the Grassmannian. The  $d_G$  and  $d_\Delta$  metrics used to compare different Grassmann representations are computationally cheap as they do not depend on the number of points present in the PD, in contrast to *Bottleneck* and *p-Wasserstein* metrics, which do. The PTS feature offers flexibility in choosing the weighting function, kernel function and perturbation level. This makes it easily adaptable to different types of real-world data. It can also be easily integrated with various ML tools, which is not easily achievable with PDs. Future directions include fusion with contemporary deep-learning architectures to exploit the complementarity of both paradigms. We expect that topological methods will push the state-of-the-art in invariant representations, where the requisite invariance is incorporated using a topological property of an appropriately redefined metric space. Additionally, the proposed methods may help open new feature-pooling options in deep-nets.

## Acknowledgments

This work was supported in part by ARO grant number W911NF-17-1-0293 and NSF CAREER award 1452163.

## References

1. Absil, P.A., Mahony, R., Sepulchre, R.: Riemannian geometry of grassmann manifolds with a view on algorithmic computation. *Acta Applicandae Mathematica* **80**(2), 199–220 (2004)
2. Absil, P.A., Mahony, R., Sepulchre, R.: Optimization algorithms on matrix manifolds. Princeton University Press (2009)
3. Adams, H., Emerson, T., Kirby, M., Neville, R., Peterson, C., Shipman, P., Chepushtanova, S., Hanson, E., Motta, F., Ziegelmeier, L.: Persistence images: A stable vector representation of persistent homology. *Journal of Machine Learning Research* **18**(8), 1–35 (2017)
4. Ali, S., Basharat, A., Shah, M.: Chaotic invariants for human action recognition. In: IEEE 11th International Conference on Computer Vision (ICCV). pp. 1–8 (2007)
5. Anirudh, R., Venkataraman, V., Natesan Ramamurthy, K., Turaga, P.: A riemannian framework for statistical analysis of topological persistence diagrams. In: The IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 68–76 (2016)
6. Aubry, M., Schlickewei, U., Cremers, D.: The wave kernel signature: A quantum mechanical approach to shape analysis. In: IEEE International Conference on Computer Vision Workshops (ICCV Workshops). pp. 1626–1633 (2011)
7. Bagherinia, H., Manduchi, R.: A theory of color barcodes. In: IEEE International Conference on Computer Vision Workshops (ICCV Workshops). pp. 806–813 (2011)
8. Basri, R., Hassner, T., Zelnik-Manor, L.: Approximate nearest subspace search. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(2), 266–278 (2011)
9. Begelfor, E., Werman, M.: Affine invariance revisited. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). vol. 2, pp. 2087–2094. IEEE (2006)
10. Bendich, P., Marron, J.S., Miller, E., Pieloch, A., Skwerer, S.: Persistent homology analysis of brain artery trees. *The Annals of Applied Statistics* **10**(1), 198–218 (2016)
11. Bertsekas, D.P.: A new algorithm for the assignment problem. *Mathematical Programming* **21**(1), 152–171 (1981)
12. Bubenik, P.: Statistical topological data analysis using persistence landscapes. *The Journal of Machine Learning Research* **16**(1), 77–102 (2015)
13. Chikuse, Y.: *Statistics on special manifolds*, vol. 174. Springer Science & Business Media (2012)
14. Chintakunta, H., Gentimis, T., Gonzalez-Diaz, R., Jimenez, M.J., Krim, H.: An entropy-based persistence barcode. *Pattern Recognition* **48**(2), 391–401 (2015)
15. Chung, M.K., Bubenik, P., Kim, P.T.: Persistence diagrams of cortical surface data. In: International Conference on Information Processing in Medical Imaging. pp. 386–397. Springer (2009)

16. Cohen-Steiner, D., Edelsbrunner, H., Harer, J.: Stability of persistence diagrams. *Discrete & Computational Geometry* **37**(1), 103–120 (2007)
17. Cohen-Steiner, D., Edelsbrunner, H., Harer, J., Mileyko, Y.: Lipschitz functions have  $L_p$ -stable persistence. *Foundations of Computational Mathematics* **10**(2), 127–139 (2010)
18. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: *Introduction to Algorithms*, 2nd edition. The MIT Press (2001)
19. Dabaghian, Y., Mémoli, F., Frank, L., Carlsson, G.: A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Computational Biology* **8**(8), 1–14 (2012)
20. De Silva, V., Morozov, D., Vejdemo-Johansson, M.: Dualities in persistent (co) homology. *Inverse Problems* **27**(12), 124003 (2011)
21. Dey, T.K., Mandal, S., Varcho, W.: Improved Image Classification using Topological Persistence. In: Hullin, M., Klein, R., Schultz, T., Yao, A. (eds.) *Vision, Modeling & Visualization*. The Eurographics Association (2017). <https://doi.org/10.2312/vmv.20171272>
22. Draper, B., Kirby, M., Marks, J., Marrinan, T., Peterson, C.: A flag representation for finite collections of subspaces of mixed dimensions. *Linear Algebra and its Applications* **451**, 15–32 (2014)
23. Edelman, A., Arias, T.A., Smith, S.T.: The geometry of algorithms with orthogonality constraints. *SIAM Journal on Matrix Analysis and Applications* **20**(2), 303–353 (1998)
24. Edelsbrunner, H., Harer, J.: *Computational topology: an introduction*. American Mathematical Society (2010)
25. Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. *Discrete & Computational Geometry* **28**(4), 511–533 (2002)
26. Emrani, S., Gentimis, T., Krim, H.: Persistent homology of delay embeddings and its application to wheeze detection. *IEEE Signal Process. Lett.* **21**(4), 459–463 (2014). <https://doi.org/10.1109/LSP.2014.2305700>, <https://doi.org/10.1109/LSP.2014.2305700>
27. Emrani, S., Saponas, T.S., Morris, D., Krim, H.: A novel framework for pulse pressure wave analysis using persistent homology. *IEEE Signal Process. Lett.* **22**(11), 1879–1883 (2015). <https://doi.org/10.1109/LSP.2015.2441068>, <https://doi.org/10.1109/LSP.2015.2441068>
28. Gopalan, R., Taheri, S., Turaga, P., Chellappa, R.: A blur-robust descriptor with applications to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**(6), 1220–1226 (2012)
29. Hamm, J., Lee, D.D.: Grassmann discriminant analysis: a unifying view on subspace-based learning. In: *Proceedings of the International Conference on Machine Learning (ICML)*. pp. 376–383. ACM (2008)
30. Harandi, M.T., Salzmann, M., Jayasumana, S., Hartley, R., Li, H.: Expanding the family of grassmannian kernels: An embedding perspective. In: *European Conference on Computer Vision (ECCV)*. pp. 408–423. Springer (2014)
31. Heath, K., Gelfand, N., Ovsjanikov, M., Aanjaneya, M., Guibas, L.J.: Image webs: Computing and exploiting connectivity in image collections. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2010)
32. Hofer, C., Kwitt, R., Niethammer, M., Uhl, A.: Deep learning with topological signatures. *arXiv preprint arXiv:1707.04041* (2017)
33. Ji-guang, S.: Perturbation of angles between linear subspaces. *Journal of Computational Mathematics* pp. 58–61 (1987)



34. Junejo, I.N., Dexter, E., Laptev, I., Perez, P.: View-independent action recognition from temporal self-similarities. *IEEE transactions on Pattern Analysis and Machine Intelligence* **33**(1), 172–185 (2011)
35. Kerber, M., Morozov, D., Nigmatov, A.: Geometry helps to compare persistence diagrams. *Journal of Experimental Algorithmics (JEA)* **22**(1), 1–4 (2017)
36. Kokkinos, I., Bronstein, M., Yuille, A.: Dense scale invariant descriptors for images and surfaces. Ph.D. thesis, INRIA (2012)
37. Krim, H., Gentimis, T., Chintakunta, H.: Discovering the whole by the coarse: A topological paradigm for data analysis. *IEEE Signal Process. Mag.* **33**(2), 95–104 (2016). <https://doi.org/10.1109/MSP.2015.2510703>, <https://doi.org/10.1109/MSP.2015.2510703>
38. Kusano, G., Hiraoka, Y., Fukumizu, K.: Persistence weighted gaussian kernel for topological data analysis. In: *International Conference on Machine Learning (ICML)*. pp. 2004–2013 (2016)
39. Li, C., Shi, Z., Liu, Y., Xu, B.: Grassmann manifold based shape matching and retrieval under partial occlusions. In: *International Symposium on Optoelectronic Technology and Application: Image Processing and Pattern Recognition* (2014)
40. Li, C., Ovsjanikov, M., Chazal, F.: Persistence-based structural recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1995–2002 (2014)
41. Lian, Z., Godil, A., Fabry, T., Furuya, T., Hermans, J., Ohbuchi, R., Shu, C., Smeets, D., Suetens, P., Vandermeulen, D., et al.: Shrec’10 track: Non-rigid 3d shape retrieval. *Eurographics Workshop on 3D Object Retrieval (3DOR)* **10**, 101–108 (2010)
42. Liu, X., Srivastava, A., Gallivan, K.: Optimal linear representations of images for object recognition. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (2003)
43. Luo, D., Huang, H.: Video motion segmentation using new adaptive manifold denoising model. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2014)
44. Masci, J., Boscaini, D., Bronstein, M., Vandergheynst, P.: Geodesic convolutional neural networks on riemannian manifolds. In: *IEEE International Conference on Computer Vision Workshops (ICCVW)*. pp. 37–45 (2015)
45. Mileyko, Y., Mukherjee, S., Harer, J.: Probability measures on the space of persistence diagrams. *Inverse Problems* **27**(12) (2011)
46. Monti, F., Boscaini, D., Masci, J., Rodolà, E., Svoboda, J., Bronstein, M.M.: Geometric deep learning on graphs and manifolds using mixture model cnns. *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* (2017)
47. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Shape distributions. *ACM Transactions on Graphics (TOG)* **21**(4), 807–832 (2002)
48. Pachauri, D., Hinrichs, C., Chung, M.K., Johnson, S.C., Singh, V.: Topology-based kernels with application to inference problems in alzheimer’s disease. *IEEE transactions on Medical Imaging* **30**(10), 1760–1770 (2011)
49. Perea, J.A., Harer, J.: Sliding windows and persistence: An application of topological methods to signal analysis. *Foundations of Computational Mathematics* **15**(3), 799–838 (2015)
50. Rahmani, H., Mian, A., Shah, M.: Learning a deep model for human action recognition from novel viewpoints. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017)

51. Reininghaus, J., Huber, S., Bauer, U., Kwitt, R.: A stable multi-scale kernel for topological machine learning. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015)
52. Rouse, D., Watkins, A., Porter, D., Harer, J., Bendich, P., Strawn, N., Munch, E., DeSena, J., Clarke, J., Gilbert, J., et al.: Feature-aided multiple hypothesis tracking using topological and statistical behavior classifiers. In: SPIE Defense+Security (2015)
53. Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G.: The graph neural network model. *IEEE Transactions on Neural Networks* **20**(1), 61–80 (2009)
54. Seversky, L.M., Davis, S., Berger, M.: On time-series topological data analysis: New data and opportunities. In: DiffCVML 2016, held in conjunction with IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2016, Las Vegas, NV, USA, June 26 - July 1, 2016. pp. 1014–1022 (2016)
55. Sharafuddin, E., Jiang, N., Jin, Y., Zhang, Z.L.: Know your enemy, know yourself: Block-level network behavior profiling and tracking. In: IEEE Global Telecommunications Conference (GLOBECOM 2010). pp. 1–6 (2010)
56. da Silva, N.P., Costeira, J.P.: The normalized subspace inclusion: Robust clustering of motion subspaces. In: IEEE International Conference on Computer Vision (ICCV). pp. 1444–1450. IEEE (2009)
57. Singh, G., Memoli, F., Ishkhanov, T., Sapiro, G., Carlsson, G., Ringach, D.L.: Topological analysis of population activity in visual cortex. *Journal of Vision* (2008)
58. Som, A., Krishnamurthi, N., Venkataraman, V., Ramamurthy, K.N., Turaga, P.: Multiscale evolution of attractor-shape descriptors for assessing parkinson’s disease severity. In: IEEE Global Conference on Signal and Information Processing (GlobalSIP) (2017)
59. Som, A., Krishnamurthi, N., Venkataraman, V., Turaga, P.: Attractor-shape descriptors for balance impairment assessment in parkinson’s disease. In: IEEE Conference on Engineering in Medicine and Biology Society (EMBC). pp. 3096–3100 (2016)
60. Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. In: *Computer Graphics Forum*. vol. 28, pp. 1383–1392. Wiley Online Library (2009)
61. Sun, X., Wang, L., Feng, J.: Further results on the subspace distance. *Pattern Recognition* **40**(1), 328–329 (2007)
62. Takens, F.: Detecting strange attractors in turbulence. In: *Dynamical Systems and Turbulence*, vol. 898, pp. 366–381 (1981)
63. Tralie, C.J., Perea, J.A.: (quasi) periodicity quantification in video data, using topology. *SIAM Journal on Imaging Sciences* **11**(2), 1049–1077 (2018)
64. Turaga, P., Veeraraghavan, A., Srivastava, A., Chellappa, R.: Statistical computations on grassmann and stiefel manifolds for image and video-based recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(11), 2273–2286 (2011)
65. Turner, K., Mileyko, Y., Mukherjee, S., Harer, J.: Fréchet means for distributions of persistence diagrams. *Discrete & Computational Geometry* **52**(1), 44–70 (2014)
66. Venkataraman, V., Ramamurthy, K.N., Turaga, P.: Persistent homology of attractors for action recognition. In: IEEE International Conference on Image Processing (ICIP). pp. 4150–4154. IEEE (2016)
67. Wang, L., Wang, X., Feng, J.: Subspace distance analysis with application to adaptive bayesian algorithm for face recognition. *Pattern Recognition* **39**(3), 456–464 (2006)

68. Weinland, D., Boyer, E., Ronfard, R.: Action recognition from arbitrary views using 3d exemplars. In: IEEE International Conference on Computer Vision (ICCV). pp. 1–7. IEEE (2007)
69. Wong, Y.C.: Differential geometry of grassmann manifolds. Proceedings of the National Academy of Sciences **57**(3), 589–594 (1967)
70. Yan, J., Pollefeys, M.: A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate. In: European Conference on Computer Vision (ECCV). pp. 94–106. Springer (2006)
71. Ye, K., Lim, L.H.: Schubert varieties and distances between subspaces of different dimensions. SIAM Journal on Matrix Analysis and Applications **37**(3), 1176–1197 (2016)
72. Yi, L., Su, H., Guo, X., Guibas, L.: Syncspecnn: Synchronized spectral cnn for 3d shape segmentation. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) (2017)
73. Zomorodian, A.: Fast construction of the vietoris-rips complex. Computers & Graphics **34**(3), 263–271 (2010)