

Supplementary Material for Learning Invariant Riemannian Geometric Representations Using Deep Nets

Suhas Lohit
Arizona State University
slohit@asu.edu

Pavan Turaga
Arizona State University
pturaga@asu.edu

1. Concepts from Differential Geometry

In this section, we will define some terms from differential geometry that are necessary for the rest of the paper. For a more comprehensive treatment of matrix manifolds considered here, refer Absil et al.[1] and Edelman et al.[2].

Manifold: A manifold is a topological space that is locally Euclidean i.e., at every point on the manifold $p \in \mathcal{M}$, there exists an open neighborhood H around p , and a mapping ϕ such that $\phi(H)$ is an open subset of \mathbb{R}^n where ϕ is a diffeomorphism. A differentiable manifold is a manifold that has a differentiable structure associated with it. A smooth manifold can be defined similarly.

Tangent and tangent space: At every point on a differentiable manifold, a linear/vector space, called the tangent space, of the same dimension as the manifold can be constructed. Consider a point $p \in \mathcal{M}$. Consider a curve $\alpha(t)$ on the manifold passing through p such that $\alpha(0) = p$. The derivative of this curve at p , $\alpha'(0)$, is the velocity vector, also called the tangent. If one considers all possible curves through this point $\{\alpha_i(t)\}, i = 1, 2, \dots$, then the set of all velocity vectors $\{\alpha'_i(0)\}$ is the tangent space $T_p\mathcal{M}$, at this point. The point at which the tangents are computed is called the pole of the tangent space.

Riemannian metric: A Riemannian metric is function that smoothly associates, to each point $p \in \mathcal{M}$, an inner product on the tangent space $T_p\mathcal{M}$. A smooth manifold equipped with a Riemannian metric is called a Riemannian manifold.

Geodesic: Consider a curve on the manifold $\gamma : [a, b] \rightarrow \mathcal{M}$ such that $\gamma(a) = x$ and $\gamma(b) = y$. The curve that minimizes the functional $E = \int_a^b \|\gamma'(t)\|^2 dt$ is called the geodesic and locally minimizes the path length between two points on the manifold. The norm $\|\cdot\|$ is induced by the Riemannian metric at $\gamma(t)$.

Exponential Map: Given that a unique geodesic $\gamma(t)$ exists locally at $p \in \mathcal{M}$ and $\gamma(0) = p$ and $\gamma'(0) = v \in T_p\mathcal{M}$, the exponential map at p is the function $\exp_p : T_p\mathcal{M} \rightarrow \mathcal{M}$ given by $\exp_p(v) = \gamma(1)$. For a neighborhood, $U \subset T_p\mathcal{M}$ containing 0, it can be shown that \exp_p is

a diffeomorphism, i.e., it has an inverse which is also continuous. The algorithm for computing the exponential map depends both on the manifold of interest and the pole of the tangent space.

Logarithm Map: At least in the neighborhood $U \subset T_p\mathcal{M}$ containing 0, the exponential map has an inverse called the logarithm map $\exp_p^{-1} : \mathcal{M} \rightarrow T_p\mathcal{M}$. This also points to the fact that the pole is an important design choice. The algorithm for computing the logarithm map depends both on the manifold of interest and the pole of the tangent space.

2. Additional Information on Face \rightarrow Illumination Subspace

2.1. Illumination Directions Used for Network Inputs

The 33 illumination directions used for creating inputs for both the training and test sets are shown in the left column of Table 1 and are a subset of the illumination directions used in the Extended Yale Face Database B [3].

2.2. Visualizing output of GrassmannNet-TS

Due to the invariance provided by the Grassmann manifold, the exponential map of the tangent need not return the same point \mathbf{U} whose columns are the principal components (PCs) of the illumination subspace, it is only guaranteed to return a point \mathbf{R} whose columns spans the same subspace as the columns of \mathbf{U} . This means that we cannot use the columns of \mathbf{R} immediately for visual comparison with the ground-truth PCs of the subspace. Therefore we first find an orthogonal matrix \mathbf{Q}^* such that $\mathbf{Q}^* = \arg \min_{\mathbf{Q}} \|\mathbf{U}_{avg} - \mathbf{RQ}\|_F$ and then use $\hat{\mathbf{R}} = \mathbf{RQ}^*$ as the new point that can be used for visualization. \mathbf{Q}^* has a simple closed form expression given by $\mathbf{Q}^* = \mathbf{WY}^T$, where $\mathbf{R}^T \mathbf{U}_{avg} = \mathbf{W}\Sigma\mathbf{Y}^T$ is the singular value decomposition (SVD) of $\mathbf{R}^T \mathbf{U}_{avg}$ [4]. See Figure 1 for visual illustration.

Azimuth	Elevation	Azimuth	Elevation
0	0	0	90
0	-20	-35	65
0	20	35	65
0	-35	-50	-40
0	45	50	-40
-5	-5	-60	-20
-5	5	60	-20
5	-5	-70	-35
5	5	-70	0
-10	0	-70	45
-10	-20	70	-35
10	0	70	0
10	20	70	45
15	20	-85	-20
-15	20	-85	20
-20	-10	85	-20
-20	10	85	20
-20	-40	-95	0
20	-10	95	0
20	10	-110	-20
20	-40	-110	15
25	0	-110	40
-25	0	-110	65
-35	-20	110	-20
-35	15	110	15
-35	40	110	40
35	-20	110	65
35	15	-120	0
35	40	120	0
-50	0	-130	20
50	0	130	20
-60	20		
60	20		

Table 1: Illumination directions (azimuth and elevation in degrees) used to generate illumination subspaces. As network inputs for both training and test sets, we only use the 33 illumination conditions shown in the left column of the table.

2.3. More results on F2IS

In this section, we provide additional results for the F2IS application. Tables 2, 3, 4 shows additional visual results for subspace dimensions $p = 3, 4, 5$ respectively and using Fréchet mean \mathbf{U}_{Fr}^d as the pole for GrassmannNet-TS. Similar results are observed with \mathbf{U}_{Tr}^d as the pole. Table 5 shows the histograms of the subspace distances obtained on the test set using the two proposed frameworks. These results demonstrate that GrassmannNet-TS is undoubtedly the better and the right way to solve the problem.

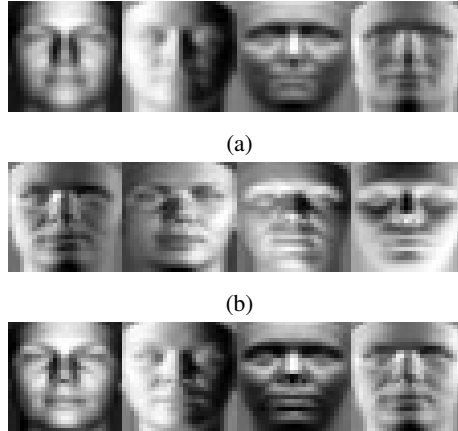


Figure 1: Images illustrating F2IS for $d = 4$ (a) Top 4 PCs of the training set i.e., \mathbf{U}_{Tr}^4 . (b) Output of exponential map for GrassmannNet-TS. (c) Output of GrassmannNet-TS after rotation by \mathbf{Q}^* .

References

- [1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.
- [2] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM journal on Matrix Analysis and Applications*, 20(2):303–353, 1998.
- [3] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001.
- [4] G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.

Input	Ground truth PC's	Output of baseline n/w	Output of GrassmannNet-TS
		 $D_G = 0.5766$	 $D_G = 0.4854$
		 $D_G = 0.7095$	 $D_G = 0.4787$
		 $D_G = 0.4429$	 $D_G = 0.4009$
		 $D_G = 1.0904$	 $D_G = 0.3042$
		 $D_G = 0.9012$	 $D_G = 0.3118$
		 $D_G = 0.9244$	 $D_G = 0.3294$

Table 2: Test results for six input images using $d = 3$. From the figures, we can clearly observe that the GrassmannNet-TS (with the Fréchet mean of the training set as the pole) framework performs much better than the baseline that attempts to regress directly to the PC's. The numbers below the output images indicate the subspace distance from the ground truth (lower the better).

















Input	Ground truth PC's	Output of baseline n/w	Output of GrassmannNet-TS
		 $D_G = 0.6750$	 $D_G = 1.1580$
		 $D_G = 1.2047$	 $D_G = 0.6674$
		 $D_G = 0.6225$	 $D_G = 0.3653$
		 $D_G = 1.5900$	 $D_G = 0.9339$

Table 3: Additional test results for six input images using $d = 4$. From the figures, we can clearly observe that the GrassmannNet-TS (with the Fréchet mean of the training set as the pole) framework performs much better than the baseline that attempts to regress directly to the PC's. The numbers below the output images indicate the subspace distance from the ground truth (lower the better).

















Input	Ground truth PC's	Output of baseline n/w	Output of GrassmannNet-TS
		 $D_G = 0.7797$	 $D_G = 0.5966$
		 $D_G = 1.5355$	 $D_G = 0.6170$
		 $D_G = 1.6760$	 $D_G = 0.4420$
		 $D_G = 1.6703$	 $D_G = 0.4939$

Table 4: Additional test results for $d = 5$. From the figures, we can clearly observe that the GrassmannNet-TS (with the Fréchet mean of the training set as the pole) framework performs much better than the baseline that attempts to regress directly to the PC's. The numbers below the output images indicate the subspace distance from the ground truth (lower the better).

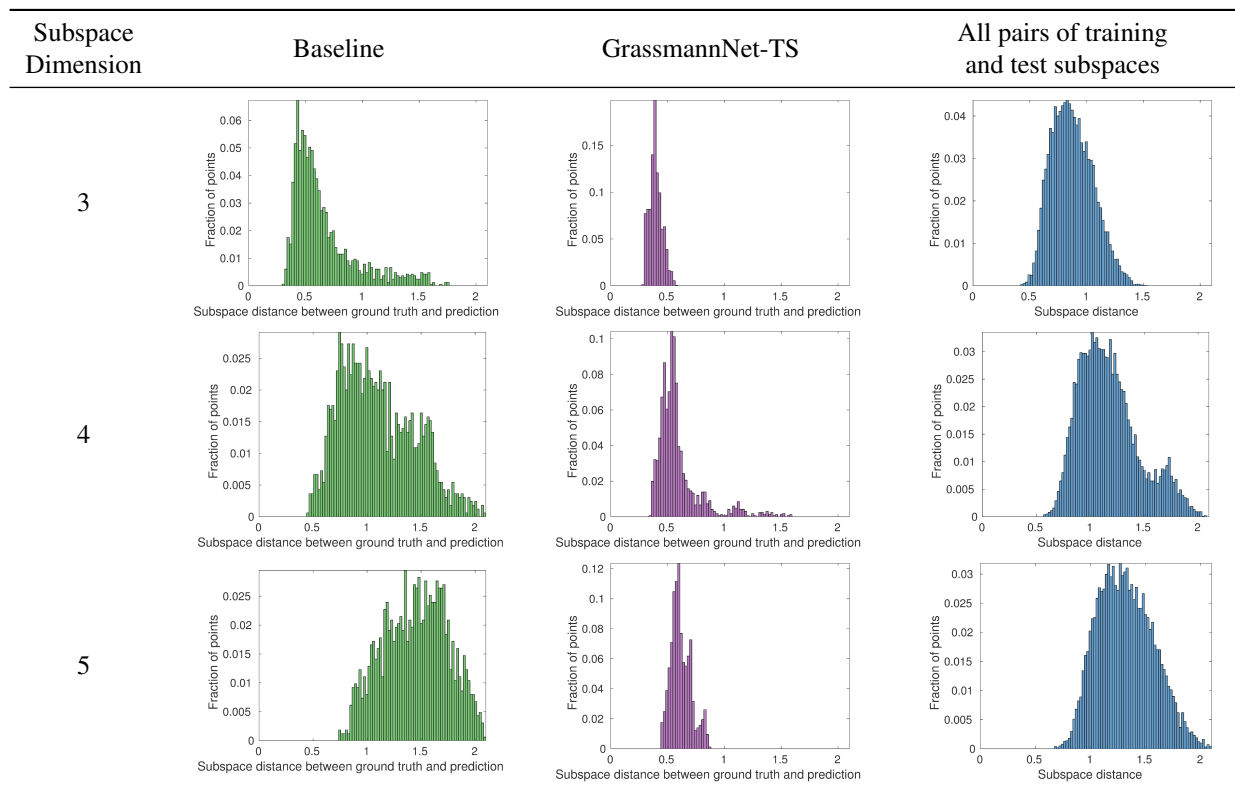


Table 5: Histograms of test results. From the plots, we can clearly observe that the GrassmannNet-TS (with the Fréchet mean of the training set as the pole) framework performs much better than the baseline that attempts to regress directly to the PC's. The last column shows the histograms of subspace distances between training and test subspaces.