Spherical Embeddings for non-Euclidean Dissimilarities

Anonymous CVPR submission

Paper ID 000

Abstract

1. Introduction

Many pattern recognition problems may be posed by defining a way of measuring *dissimilarities* between patterns.

Given a set of samples, the dissimilarity matrix $\mathbf{D}$ defines the problem at hand. If the distances are Euclidean then we can find an equivalent similarity matrix $\mathbf{S}$ which is positive semi-definite. We can identify it with a kernel matrix $\mathbf{S} \equiv \mathbf{K}$ and use a kernel machine to classify the data. The kernel then has an implicit Euclidean space associated with it which can be found by kernel embedding. The similarities in this space are identified with the inner products between the embedded points. However, for many types of distance, the similarity matrix $\mathbf{S}$ is indefinite and cannot be associated with a kernel. The objects represented by $\mathbf{S}$ cannot exist as points in a Euclidean space with the given similarities.

2. Indefinite spaces

We begin with the assumption that we have measured a set of dissimilarities or distances between all pairs of patterns in our problem. This is denoted by the matrix $\mathbf{D}$, where D_{ij} is the distance between i and j. We can define an equivalent set of similarities by using the matrix of squared distances $\mathbf{D}'$, where $D'_{ij} = D^2_{ij}$. This is achieved by identifying the similarities as $-\frac{1}{2}\mathbf{D}'$ and centering the resulting similarities:

$$\mathbf{S} = -\frac{1}{2}(\mathbf{I} - \frac{1}{n}\mathbf{J})\mathbf{D}'(\mathbf{I} - \frac{1}{n}\mathbf{J}) \quad (1)$$

Here $\mathbf{J}$ is the matrix of all-ones, and n is the number of objects. In Euclidean space, this procedure gives exactly the inner-product or kernel matrix for the points.

If $\mathbf{S}$ is positive semi-definite, then the original distances are Euclidean and we can use the kernel embedding to find positions $\mathbf{x}_i$ for the points in Euclidean space:

$$\mathbf{X} = \mathbf{U}_S \mathbf{\Lambda}_S^{\frac{1}{2}} \quad (2)$$

where $\mathbf{U}_S$ and $\mathbf{\Lambda}_S$ are the eigenvector and eigenvalue matrices of $\mathbf{S}$, respectively.

If $\mathbf{S}$ is indefinite, which is often the case, then the objects cannot exist in Euclidean space with the given distances. This does not necessarily mean the the distances are non-metric; metricity is a separate issue. One measure of the deviation from definiteness which has proved useful is the negative eigenfraction (NEF) which measures the fractional weight of eigenvalues which are negative:

$$\text{NEF} = \frac{\sum_{\lambda_i<0} |\lambda_i|}{\sum_i |\lambda_i|} \quad (3)$$

We can measure the non-metricity of the data by counting the number of violations of metric properties. It is very rare to have an initial distance measure which gives negative distance, so we will assume than the distances are all positive. The two measures of interest are then the fraction of triples which violate the triangle inequality (TV) and the degree of asymmetry of the distances (γ):

$$\gamma = \sum_{i \neq j} \frac{|\tilde{d}(i,j) - \tilde{d}(j,i)|}{|\tilde{d}(i,j) + \tilde{d}(j,i)|} \quad (4)$$

where $\tilde{d}(.,.)$ is the dissimilarity scaled so that the average dissimilarity is one.

Previous work has shown that they can be embedded in a non-Riemannian pseudo-Euclidean space[4]. This space uses the non-Euclidean inner product

$$< \mathbf{x}, \mathbf{y} >= \mathbf{x}^T \mathbf{M} \mathbf{y} \quad (5)$$

where

$$\mathbf{M} = \begin{pmatrix} \mathbf{I}_p & 0 \\ 0 & -\mathbf{I}_q \end{pmatrix}$$

The values of -1 correspond to the 'negative' part of the space. The space has a signature (p, q) with p positive dimensions and q negative dimensions.

This inner-product induces a norm, or distance measure:

$$|\mathbf{x}|^2 =< \mathbf{x}, \mathbf{x} >= \mathbf{x}^T \mathbf{M} \mathbf{x} = \sum_{i_+} x_i^2 - \sum_{i_-} x_i^2 \quad (6)$$

We can then write the similarity as

$$\mathbf{S} = \mathbf{U}_S |\mathbf{\Lambda}_S|^{\frac{1}{2}} \mathbf{M} |\mathbf{\Lambda}_S|^{\frac{1}{2}} \mathbf{U}_S^T \quad (7)$$

where the negative part of the space corresponds to the negative eigenvalues of $\mathbf{K}$, and the kernel embedding as

$$\mathbf{X} = \mathbf{U}_S |\mathbf{\Lambda}_S|^{\frac{1}{2}} \quad (8)$$

So the pseudo-Euclidean embedding reproduces precisely the original distance and similarity matrices. But, while the pseudo-Euclidean embedding reproduces the original distance matrix, it introduces a number of other problems. The embedding space is non-metric and points in the space can have negative distances to each other. Locality is not preserved in such a space. *more about problems* In order to overcome these problems, we would like to embed the points in a space with a metric distance measure which produces indefinite similarity matrices; this means that the space must be curved.

3. Riemannian Manifolds

In this paper, we use Riemannian manifolds to embed a set of objects. On the manifold, distances are measured by geodesics (the shortest curve between points), so we will employ manifolds where the geodesics are easy to compute. The manifold must also be curved in order to produce an indefinite similarity matrix. Two prime candidates for the embedding are the *elliptic* manifold and the *hyperbolic* manifold.

An n-dimensional Riemannian space is defined by its metric tensor g_{ij} in some local coordinate system $\{u_1, u_2 \dots u_n\}$. This can be related to an infinitesimal distance element in the space by

$$ds^2 = \sum_{ij} g_{ij} du_i du_j \quad (9)$$

The metric tensor must be positive definite, and any metric tensor defines a particular Riemannian space. However, this definition is not always the most convenient - alternatively we can define a manifold as a subspace embedded in a larger flat (Euclidean or pseudo-Euclidean) space. The embedding then implies a particular metric for the space. Note that the converse is not true; there may be many different embeddings with the same metric and therefore the same Riemannian space. Finally, the geodesic distance in a Riemannian space is a metric distance, but in general it is non-Euclidean. However, finding the geodesic between two points on the manifold is not easy; it involves solving a set of coupled second-order differential equations. There are, however, manifolds which are non-Euclidean but on which it is easy to find the geodesics. Two examples are furnished by the elliptic and hyperbolic manifolds.

4. Elliptic Geometry

Elliptic geometry is the geometry on the surface of a hypersphere. The hypersphere can be straightforwardly embedded in Euclidean space; for example the embedding of a sphere in three dimensions is well known:

$$\mathbf{x} = (r \sin u \sin v, r \cos u \sin v, r \cos v)^T \quad (10)$$

This embedding implies a particular metric tensor:

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 \\ &= r^2 \sin^2 v du^2 + r^2 dv \end{aligned} \quad (11)$$

The embedding of an $(n-1)$-dimensional sphere in n-dimensional space is a straightforward extension of this. We can define the surface implicitly using the constraint

$$\sum_i x_i^2 = r^2 \quad (12)$$

For the hypersphere to be a Riemannian space, we should have a positive definite metric tensor g. This is equivalent to the statement $ds^2 > 0$ for any infinitesimal movement in the surface. We have

$$ds^2 = \sum_i dx_i^2 \quad (13)$$

which clearly must be positive for any values of dx_i. This surface is curved and has a constant sectional curvature of $K = 1/r^2$ everywhere.

The geodesic distance between two points in curved space is the length of the shortest curve lying in the space and joining the two points. For an elliptic space, the geodesic is a great circle on the hypersphere. The distance is the length of the arc of the great circle which joins the two points. If the angle subtended by two points at the centre of the hypersphere is θ_{ij}, then the distance between them is

$$d_{ij} = r\theta_{ij} \quad (14)$$

With the coordinate origin at the centre of the hypersphere, we can represent a point by a position vector $\mathbf{x}_i$ of length r. Since the inner product is $< \mathbf{x}_i, \mathbf{x}_j >= r^2 \cos\theta_{ij}$ we can also write

$$d_{ij} = r \cos^{-1} \frac{< \mathbf{x}_i, \mathbf{x}_j >}{r^2} \quad (15)$$

The elliptic space is metric but clearly not Euclidean. It is therefore a good candidate for representing points which

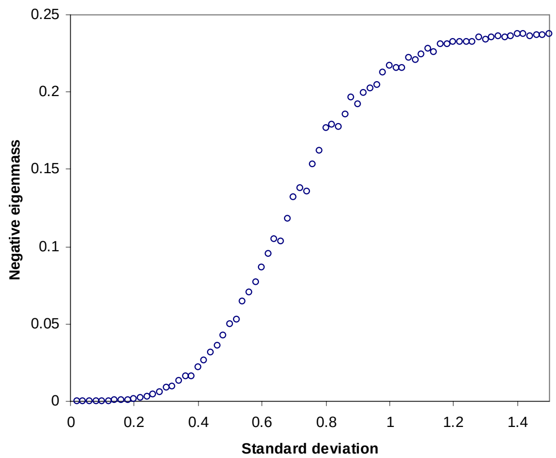

Figure 1. Negative eigenvalue fraction of points on an elliptic surface

produce indefinite kernels. The first question we wish to answer is, to what extent do the points in a curved space produce indefinite similarities? To answer this question, we have constructed the similarity matrices of points in these spaces. The points are generated via a parameterisation (Eqn 10) and drawing points from the parameters via a normal distribution. The indefinite nature of the similarity can be characterised by the negative eigenfraction (NEF) (Eqn. 3)

Figure 1 shows the NEF for points on a unit hypersphere with varying standard deviation. The curved manifold produces significant negative eigenfraction, up to 24%.

4.1. Embedding in Elliptic space

Given a distance matrix $\mathbf{D}$, we wish to find the set of points in an elliptic space which produce the same distance matrix. Since the curvature of the space is unknown, we must additionally find the radius of the hypersphere. We have n objects of interest, and therefore we would normally look for an $n-1$-dimensional Euclidean space. Since we have freedom to set the curvature, we must look for a $n-2$-dimensional elliptic space embedded in the $n-1$-dimensional Euclidean space.

We begin by constructing a space with the origin at the centre of the hypersphere. If the point positions are given by $\mathbf{x}_i, i = 1 \dots n$, then we have

$$< \mathbf{x}_i, \mathbf{x}_j >= r^2 \cos\theta_{ij} = r^2 \cos\frac{d_{ij}}{r} \tag{16}$$

Next, we construct the matrix of point positions $\mathbf{X}$, with each position vector as a column. Then we have

$$\mathbf{X}\mathbf{X}^T = \mathbf{Z} \tag{17}$$

where $Z_{ij} = r^2 \cos d_{ij}/r$. Since the embedding space has dimension $n-1$, $\mathbf{X}$ consists of n points of dimension $n-1$ and $\mathbf{Z}$ should then be an n by n matrix which is positive semi-definite with rank $n-1$. In other words, $\mathbf{Z}$ should have a single zero eigenvalue, with the rest positive. We can use this observation to determine the radius of curvature. Given a radius r and a distance matrix $\mathbf{D}$, we can construct $\mathbf{Z}(r)$ and find the smallest eigenvalue λ_0. By minimising the magnitude of this eigenvalue, we can find the optimal radius.

$$r^* = \arg\min_r |\lambda_0 \left[\mathbf{Z}(r)\right]| \tag{18}$$

In practice we locate the optimal radius via search. The smallest eigenvalue can be determined efficiently using the power method without the expense of the full eigendecompositon. Give the optimal radius, the embedding positions are determined via the full eigendecomposition:

$$\mathbf{Z}(r^*) = \mathbf{U}_Z \mathbf{\Lambda}_Z \mathbf{U}_Z \tag{19}$$

$$\mathbf{X} = \mathbf{U}_Z \mathbf{\Lambda}_Z^{\frac{1}{2}} \tag{20}$$

If the points truly lie on a hypersphere, then this is sufficient. However, in general this is not the case, the optimal smallest eigenvalue λ_0 will be less than zero, and there will be residual negative eigenvalues. The embedding then is onto a 'hypersphere' of radius r, but embedded in a pseudo-Euclidean space. In order to obtain points on the hypersphere, we must correct the recovered points. The traditional method in kernel embedding is to discard the negative eigenvalues; here that will not suffice as this will change the length of the vectors and constraint 12 will be violated. In the next section we present a solution to this problem.

4.2. Approximation of the points

For a general set of distances, the points do not lie on a hypersphere, and need correction to lie on the manifold. We pose the problem as follows: The task is to find a point-position matrix $\mathbf{X}$ on the elliptic manifold which minimises the Frobenius distance to the Euclidean-equivalent matrix $\mathbf{Z}$. Given the manifold radius r, determined by the method in the previous section, we begin with the normalised matrix $\hat{\mathbf{Z}} = \mathbf{Z}/r^2$. The problem is then

$$\begin{aligned} \min_{\mathbf{X}} & \quad |\mathbf{X}\mathbf{X}^T - \hat{\mathbf{Z}}| \\ \mathbf{x}_i^T\mathbf{x}_i & = 1 \end{aligned} \tag{21}$$

This can be simplified by observing in the usual way that the Frobenius norm is invariant under an orthogonal similarity transform, so given $\hat{\mathbf{Z}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$, we apply $\mathbf{U}$ as an orthogonal similarity transform to get

$$\min_{\mathbf{X}} |\mathbf{U}^T\mathbf{X}\mathbf{X}^T\mathbf{U} - \mathbf{\Lambda}| \tag{22}$$

which has a solution $\mathbf{X} = \mathbf{U}\mathbf{D}$ where $\mathbf{D}$ is some diagonal matrix, giving

$$\min_{\mathbf{D}} |\mathbf{D}^2 - \mathbf{\Lambda}| \tag{23}$$

Of course, $\mathbf{D}^2 = \mathbf{\Lambda}$ is a solution if all the eigenvalues are positive, and this is the case if the points lie precisely on a hypersphere. In the general case, there will be negative eigenvalues and we must find a minimum of the constrained optimisation problem. Let $\mathbf{d}$ be the vector of squared diagonal elements of $\mathbf{D}$, i.e. $d_i = D_{ii}^2$ and U_s be the matrix of squared elements of $\mathbf{U}$, $U_{sij} = U_{ij}^2$. Then we can write the problem as

$$\begin{aligned} \min_{\mathbf{d}} \quad & (\mathbf{d} - \lambda)^T (\mathbf{d} - \lambda) \\ d_i \quad & > \quad 0 \\ \mathbf{U}_s \mathbf{d} \quad & = \quad \mathbf{1} \end{aligned} \tag{24}$$

While this is a quadratic problem, and can be solved by quadratic programming, the solution actual has a simple form which can be found by noting that the matrix $\mathbf{U}_s$ should have rank $n-1$ and hence one singular value equal to zero. First we make the following observations: The vector of eigenvalues λ is an absolute minimiser of this problem, i.e. $\mathbf{d} = \lambda$ minimises the Frobenius norm and satisfies constraint 2, but not constraint 1. Secondly, $\mathbf{d} = \mathbf{1}$ satisfied both constraints since $\sum_i U_{ij}^2 = 1$ (as $\mathbf{U}$ is orthogonal). These observations, and the fact that $\mathbf{U}_s$ is rank $n-1$, means that the general solution to the second constraint is

$$\mathbf{d} = \mathbf{1} + \alpha(\lambda - \mathbf{1}) \tag{25}$$

It only remains then to find the value of α which satisfies the first constraint and minimises the criterion. Since the criterion is quadratic, the solution is simply given by the largest value of α for which the first constraint is satisfied.

4.3. Hyperbolic geometry

As we previously observed, the pseudo-Euclidean(pE) space has been used to embed points derived from indefinite kernels. The pE space is clearly non-Riemannian as points may have negative distances to each other. However, it is still possible to define a sub-space which is Riemannian. As an example, take the 3D pE space with a single negative dimension (z) and the 'sphere' defined by

$$< \mathbf{x}, \mathbf{x} >= x^2 + y^2 - z^2 = -r^2 \tag{26}$$

This space is called *hyperbolic*.

This surface has a parameterisation given by

$$\mathbf{x} = (r \sin u \sinh v, r \cos u \sinh v, r \cosh v)^T \tag{27}$$

As before, there is a particular metric tensor associated with this embedding:

$$\begin{aligned} ds^2 &= dx^2 + dy^2 - dz^2 \\ &= r^2 \sinh^2 v du^2 + r^2 dv \end{aligned} \tag{28}$$
$$\tag{29}$$

and so the metric tensor is

$$g = r^2 \begin{pmatrix} \sinh^2 v & 0 \\ 0 & 1 \end{pmatrix} \tag{30}$$

The metric tensor is positive definite, and so the surface is Riemannian and distances measured on the surface are metric, even though the embedding space is non-Riemannian.

We can extend this hyperbolic space to more dimensions. Firstly, we take the case when there is just one negative dimension, z in the embedding space.

$$\sum_i x_i^2 - z^2 = -r^2 \tag{31}$$

It can be shown that $ds^2 \geq 0$ and so the hyperbolic surface is Riemannian. If there is more than one negative dimension, the surface is no longer Riemannian as it is possible to obtain $ds^2 < 0$. The hyperbolic space is therefore restricted to any number of positive dimensions but just one negative dimension.

Finally, the sectional curvature of this space, as with the hypersphere, is constant everywhere. In this case, the curvature is negative and given by $K = -1/r^2$. For the hyperbolic space, the geodesic is the analogue of a great circle. While the notion of angle in Euclidean space is intuitive, it is less so in pE space. However, we can define such a notion from the inner product. The inner product is defined as

$$\begin{aligned} < \mathbf{x}_i, \mathbf{x}_j > &= \sum_k x_{ik} x_{jk} - z_i z_j && (32) \\ &= -|\mathbf{x}_i||\mathbf{x}_j| \cosh \theta_{ij} && (33) \end{aligned}$$
$$\tag{34}$$

which in turn defines the notion of hyperbolic angle. From this angle, the distance between two points in the space is

$$d_{ij} = r\theta_{ij} \tag{35}$$

With the coordinate origin at the centre of the hypersphere, we can represent a point by a position vector $\mathbf{x}_i$ of length r. Since the inner product is $< \mathbf{x}_i, \mathbf{x}_j >= r^2 \cosh \theta_{ij}$ we can also write

$$d_{ij} = r \cosh^{-1} - \frac{< \mathbf{x}_i, \mathbf{x}_j >}{r^2} \tag{36}$$

Figure 2 shows the NEF for points on a unit hyperbolic surface with varying standard deviation. The curved manifold produces significant negative eigenfraction, up to 16%.

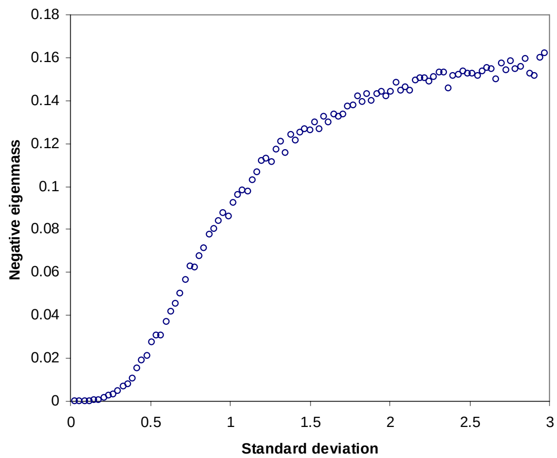

Figure 2. Negative eigenvalue fraction of points on a hyperbolic surface

4.4. Embedding in Hyperbolic space

In hyperbolic space, we have

$$< \mathbf{x}_i, \mathbf{x}_j >= -r^2 \cosh \theta_{ij} = r^2 \cosh \frac{d_{ij}}{r} \tag{37}$$

with the inner product defined by Eqn 5. Constructing $\mathbf{Z}$ as before, we get

$$\mathbf{XMX}^T = \mathbf{Z} \tag{38}$$

Again we have an embedding space of dimension $n-1$, but $\mathbf{Z}$ is no longer positive semi-definite. In fact, $\mathbf{Z}$ should have precisely one negative eigenvalue (since the hyperbolic space has just one negative dimension) and again a single zero eigenvalue. We must now minimise the magnitude of the second smallest eigenvalue:

$$r^* = \arg\min_r |\lambda_1 [\mathbf{Z}(r)]| \tag{39}$$

The embedded positions become

$$\mathbf{X} = \mathbf{U}_Z |\mathbf{\Lambda}_Z|^{\frac{1}{2}} \tag{40}$$

As with the elliptic embedding, in general the points do not lie on the embedding space and there will be residual negative eigenvalues.

4.5. Approximation of the points

A similar procedure may applied for hyperbolic space as for the elliptic space. Exactly one of the d_i's must be negative (the one corresponding to the most negative element of λ). Let d_1 be the component of $\mathbf{d}$ corresponding to the negative dimension. We then have, as before

$$\begin{aligned} \min_{\mathbf{d}} \quad & (\mathbf{d} - \lambda)^T (\mathbf{d} - \lambda) \\ d_1 \quad & < \quad 0 \\ d_i^* \quad & > \quad 0, i \neq 1 \\ \mathbf{U}_s \mathbf{d} \quad & = \quad -\mathbf{1} \end{aligned} \tag{41}$$

Now we have a global minimiser of $\mathbf{d} = \lambda$ which satisfies the final constraint and a second solution of the constraint is given by $\mathbf{d} = -\mathbf{1}$. We must therefore find the optimal value for α in

$$\mathbf{d} = -\mathbf{1} + \alpha(\lambda + \mathbf{1}) \tag{42}$$

The solution is more complicated than in the elliptical case, due to the constraint $d_1 < 0$. This means that it is possible that there is no solution. However, in all cases we have examined, there is a set of solutions. If a solution exists, the optimal point will lie on one of the two boundaries of the feasible region.

5. Results

We have applied our spherical embedding techniques to three different datasets. These datasets give rise to indefinite similarities. These are the Chickenpieces data[5, 1], the CoilYork data[6] and the DelftGestures[3] set.

5.1. Chickenpieces

The Chickenpieces data is a useful set for the study of indefinite similarities, because there is a set of parameters which can be varied to change the level of indefiniteness. In this case, we use a cost of 45 and varying values of L. Figure 3 shows the negative eigenfraction(NEF) of the data, which is a measure of indefiniteness. We also measure the residual NEF of the matrix $\mathbf{Z}$, which measures the negative eigenvalues of $\mathbf{Z}$ not explained by the spherical embedding. For the elliptic embedding, this is precisely Eqn. 3 applied to $\mathbf{Z}$, but for the hyperbolic embedding we ignore the largest negative eigenvalue in the top sum, which is allowed in this geometry.

The assymmetry of the data varies with L, but reaches a maximum of 0.08 for $L = 40$, and the fraction of triangle violations is less than 0.01% in all cases. The data is therefore highly non-Euclidean, but only very slightly non-metric.

Figure 4 shows the two-dimensional kernel embedding of the points (i.e. the largest positive dimension of the pseudo-Euclidean embedding) and two views of the elliptic embedding. The point distributions are quite different, particularly for the red class.

Secondly, we used the 1-NN classifier to measure the classification accuracy on the data. The 1-NN is chosen to give a comparison between the original embeddings and

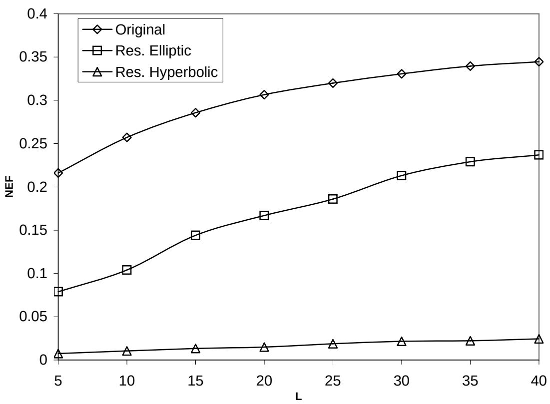

Figure 3. NEF of the original data and of the similarity matrices $\mathbf{Z}$ for the ChickenPieces data

NEF	Res. Elliptic	Res. Hyperbolic
0.258	0.129	0.0099

Table 1. NEF for the original data and embeddings of the CoilYork data.

spherical space because it is one of very few classifiers which do not rely either explicitly or implicitly on a underlying manifold. For comparison purposes, we also show results against a couple of Euclidean embeddings; the positive subspace ('Pos Sp') which is the kernel embedding discarding negative eigenvalues, and the absolute space ('Abs sp') which is the kernel embedding were negative eigenvalues are instead considered to be positive (absolute space). The error-rates are estimated using 10-fold cross-validation. The error-rates are shown in Figure 5.

5.2. CoilYork

The CoilYork data is a set of dissimilarity measurements between four objects from the COIL database. Firstly, feature points are extracted from the object images, and then these points are used to construct a graph for each image[6]. The graph matching algorithm of Gold and Ranguranjan[2] is then used to generate a dissimilarity score between the graphs. Table 1 summarises the indefiniteness of the data and the two embeddings. The asymmetry coefficient is 0.009 and there is just one triangle violation from more than 23 million triples.

In Figure 6, we show learning curves for the CoilYork data. These error-rates are obtained using 10-fold cross-validation on training sets of varying sizes. As before, 'Pos Sp' is the positive subspace and 'Abs sp' is the absolute space.

5.3. DelftGestures

The DelftGestures are a set of dissimilarities generated from a sign-language interpretation problem. The gestures

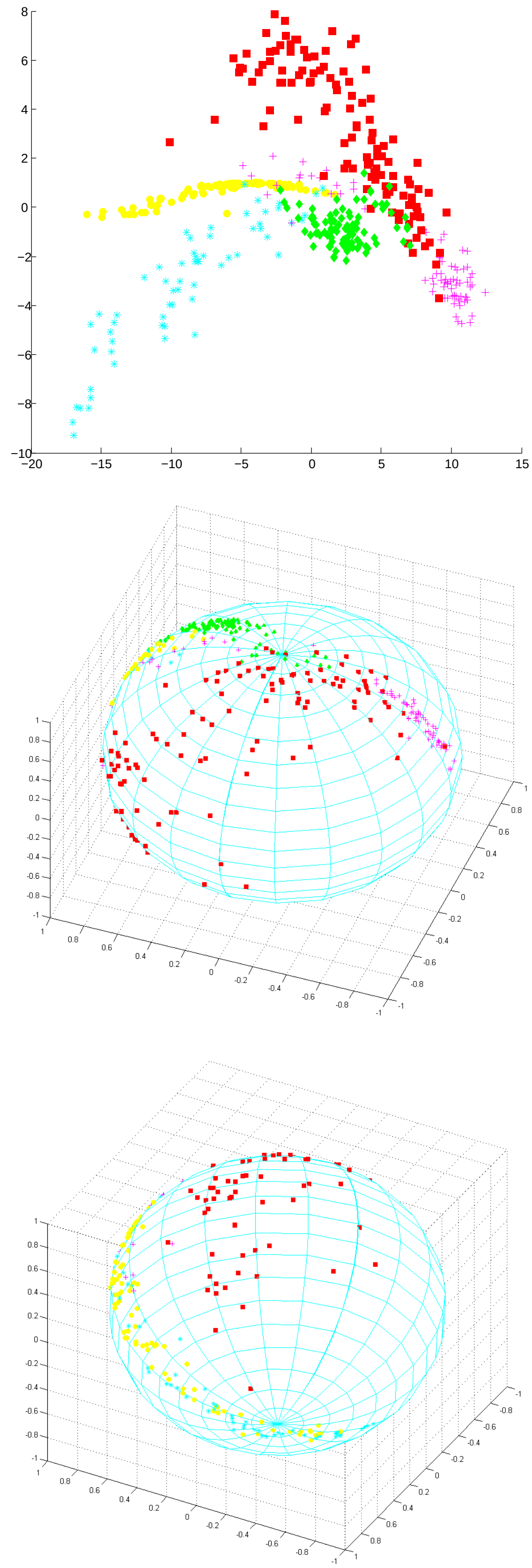

Figure 4. The 2D kernel embedding and two views of the 2D elliptic embedding of the Chickenpieces data with L=25

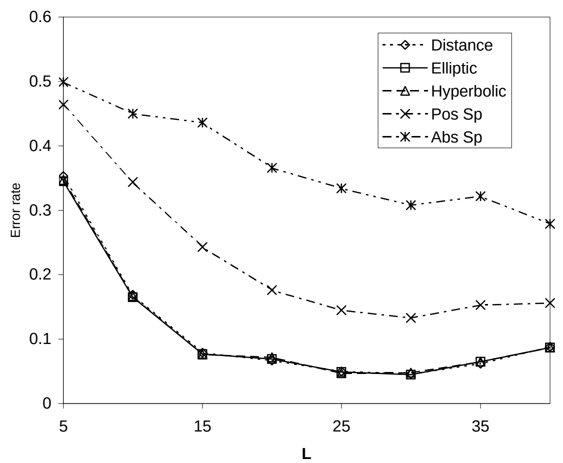

Figure 5. Estimated error rates for the Chickenpieces data

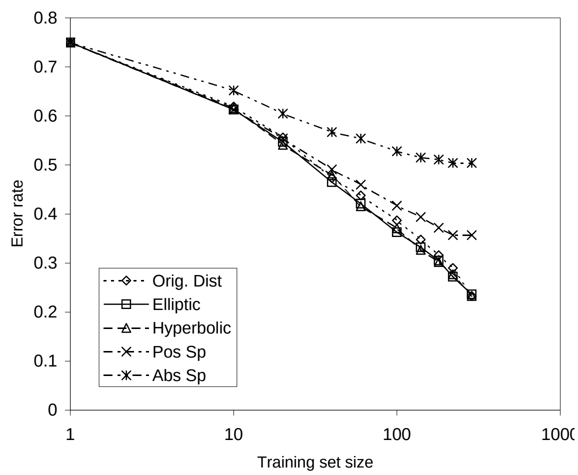

Figure 6. Estimated error rate for the CoilYork data

NEF	Res. Elliptic	Res. Hyperbolic
0.309	0.214	0.022

Table 2. NEF for the original data and embeddings of the Delft-Gesture data.

are measured by two video cameras observing the positions the two hands in 75 repetitions of creating 20 different signs. The dissimilarities are computed using dynamic time warping procedure on the sequence of positions[3]. Table 2 shows the NEF for the data and the two embeddings. The data is completely symmetric, and the fraction of triangle violations is 4×10^{-6}, so for all normal purposes the data is metric but significantly non-Euclidean.

In Figure 7, we show learning curves for the DelftGestures data. These error-rates are obtained using 10-fold cross-validation on training sets of varying sizes. As before, 'Pos Sp' is the positive subspace and 'Abs sp' is the absolute space.

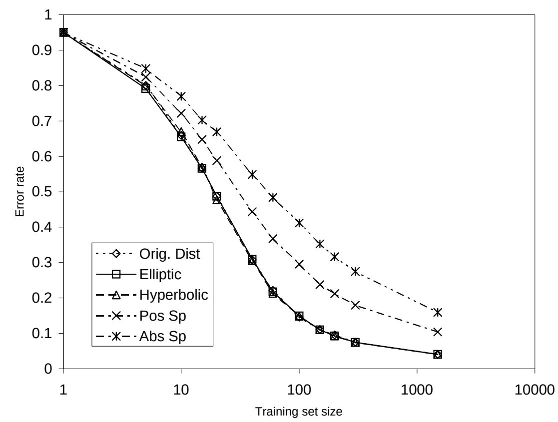

Figure 7. Estimated error rate for the DelftGestures data

6. Conclusion

References

[1] G. Andreu, A. Crespo, and J. Valiente. Selecting the toroidal self-organizing feature maps (tsofm) best organized to object recognition. In *ICNN'97*, pages 1341–1346, 1997. 5

[2] S. Gold and A. Rangarajan. A graduated assignment algorithm for graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:377–388, 1996. 6

[3] J. Lichtenauer, E. A. Hendriks, and M. J. T. Reinders. Sign language recognition by combining statistical dtw and independent classfication. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30:2040–2046, 2008. 5, 7

[4] E. Pekalska and R. P. W. Duin. *The dissimilarity representation for pattern recognition*. World Scientific, 2005. 1

[5] E. Pekalska, A. Harol, R. P. W. Duin, B. Spillmann, and H. Bunke. Non-euclidean or non-metric measures can be informative. In *SSPR/SPR*, pages 871–880, 2006. 5

[6] B. Xiao and E. R. Hancock. Geometric characterisation of graphs. *ICIAP 2005*, LNCS 3617:471–478, 2005. 5, 6