Technical Report

Title: Manifold learning and Random Projections for multi-view object recognition

Authors: Grigorios Tsagkatakis¹ and Andreas Savakis²

¹Center for Imaging Science, Rochester Institute of Technology, PhD Student
²Professor and Head of Computer Engineering, Rochester Institute of Technology

Purpose: discussion of the outcomes of the work supported by the CIS-Kodak micro fund.

Resulted publications:


Abstract

Recognizing objects from different viewpoints is a challenging task. One approach for handling this task is to model the appearance of an object under different viewing conditions using a low dimensional subspace. Manifold learning describes the process by which this low dimensional embedding can be generated. However, manifold learning is an unsupervised method and thus gives poor results on classification tasks. To address this impediment of manifold learning, we investigated the combination of manifold learning and distance metric learning for the generation of a representation that is both discriminative and informative. In the proposed system, initial dimensionality reduction is achieved using random projections, a computationally efficient and data independent linear transformation. Distance metric learning is then applied to increase the separation between classes and improve the accuracy of
nearest neighbor classification. Finally, a manifold learning method is used to generate a mapping between the randomly projected data and the low dimensional manifold. We demonstrate that this approach is effective for multi view face recognition. Furthermore, we show that random projections can be applied as an initial step without significantly affecting the classification accuracy. Although we provide results regarding the classification of human faces from different viewpoints, the same approach can be used for the classification of generic objects under different viewing conditions.

I. Introduction

Images and video convey a wealth of information, yet, extracting a higher level of understanding from data is a computationally complex task. With added constraints, such as processing power, memory and bandwidth, the problem becomes even more challenging, and the high dimensional nature of the data may not be optimal for modeling and computational purposes. Thus, dimensionality reduction becomes a necessary step before more elaborate computer vision algorithms can be applied.

Dimensionality reduction is often achieved by principal components analysis (PCA). Even though PCA is a widely used method, certain limitations of the method, such as data dependence and linearity restrictions, are becoming more evident and suggest the investigation of more appropriate dimensionality reduction techniques. Manifolds offer a novel approach in signal representation that has received considerable attention in recent years [11]. Compared to linear techniques, manifold learning can identify non-linear structures and project the data in a low dimensional space, while preserving key properties such as geodesic distances or the local neighborhood structure.

A variety of signals can be reliably modeled using manifold models [15]. Manifold learning was successfully applied for face recognition based on Laplacianfaces [1]. The face manifold was identified by applying locality preserving projections (LPP), and then nearest neighbor classification was used for recognition. The method achieved higher recognition accuracy compared to Eigenfaces [16] (PCA based) and Fisherfaces [17] (discriminant based) that model faces using linear structures.

![Figure 1: Block diagram of the proposed system](image-url)
Our proposed method is composed of three modules for face representation and classification as shown in Figure 1: (a) random projections (RPs), (b) distance metric learning (DML) and (c) manifold learning (ML). RPs is a data-independent linear dimensionality reduction technique that preserves the manifold structure of the data [7] and is applied as a preprocessing step. DML is then applied in order to generate a distance metric that will increase the separability between classes and improve classification accuracy. Once the new distance metric is learned, it is used during the generation of the distance matrix in ML. During testing a nearest neighbor classifier is employed in manifold space and utilizes the learned distance for classification.

II. Manifold Learning

The goal of manifold learning is to reduce the dimensionality of the input data while preserving its geometric properties. Methods like Isometric Feature Embedding (Isomap) [2] and Local Linear Embedding (LLE) [3] are examples of graph embedding manifold learning approaches. In these methods, the data is described through a distance matrix (adjacency graph) and dimensionality reduction is achieved by applying eigenanalysis on this distance matrix.

A problem with manifold learning methods, such as Isomap and LLE, is the lack of a direct mapping from the input space to the manifold space. Instead, they provide a mapping of the input data to the manifold space. This is called the out-of-sample extension problem and can limit the applicability of these methods for new data [12].

More recent methods try to overcome this problem by performing a linear approximation of the manifold based on the idea that even though manifolds are non-linear structures, they can be approximated reasonably well within a small neighborhood using a linear map. The benefits of linear approximations are mostly expressed in terms of savings in computational time, although in some cases, e.g. locality preserving projections, very promising results have been reported.

Locality preserving projections (LPP) [1] is the optimal linear approximation of the LLE algorithm and can be described by four steps. First, the adjacency graph $G$ is built, which is similar to LLE and Isomap. Then, the Laplacian of the graph is estimated. The Laplacian is a “normalization” of the graph that can provide more insights into its structure, and is given by $L = D - W$, where $D_{ii} = \sum_j W_{ij}$ and $W$ is the affinity matrix of the graph $G$. Using the Laplacian, the following generalized eigenvalue problem is solved by Eq. (1)

$$XLX^T a = \lambda XDX^T a$$  \hspace{1cm} (1)
where $X$ is the data matrix. In the final step, the linear mapping that provides the embedded $y_i$ of the input point $x_i$ is generated by forming a matrix $A$. Each column of $A$ corresponds to an eigenvector as shown in Eq. (2)

$$x_i \mapsto y_i = A^T x_i, A = (a_0, a_1, ..., a_i)$$

Although manifold learning is a powerful method for unsupervised modeling, it does not take class information into account during the embedding. In other words, preserving the structure of the manifold, which is the goal in manifold learning, is independent of the classification, since neighboring points might belong to different classes, e.g. faces of different individuals. The problem was identified in [4] where it was coined “classification-oriented multi-manifolds learning”. A few supervised manifold learning approaches have been presented in the literature, such as supervised LLE [10], Large Margin Component Analysis [5] and Local Discriminant Embedding [6].

Supervised manifold learning algorithms exploit the class label information during training in order to increase the accuracy. However, once training is performed, the mapping and the classification are usually applied as in unsupervised methods. To handle this shortcoming, we investigate the introduction of distance learning before the application of manifold learning. The benefit of the proposed approach is that class label information is utilized in both the manifold learning and the subsequent classification.

### III. Distance Metric Learning

Typically, the distance between input data is measured using the Euclidian distance. However, when additional information about the data is available, different distance metrics can provide a more meaningful connection between the data. Distance metric learning (DML) has received considerable attention in the past few years and a number of different DML approaches have been presented. In supervised DML, the objective is to learn a new distance that will satisfy the pairwise constrains provided by the class labels. Formally, the distance between two data points $x$ and $y \in \mathbb{R}^n$ is given by

$$d_M(x, y) = \|x - y\|^2_M = (x - y)^T M (x - y)$$

where $M \in \mathbb{R}^{n \times n}$ is the distance metric matrix. Alternatively, $M$ can be written as $M = L^T L$ in which case Eq. (3) becomes

$$d_M(x, y) = (L(x - y))^T (L(x - y))$$

and equivalently, the new distance metric is the Euclidean distance between the linearly transformed data.
A well-known example is the Mahalanobis distance where the inverse of the covariance matrix is used as the matrix $M$ in Eq. (4). In this work, we focus on a successful DML algorithm, the Large Margin Nearest Neighbor (LMNN) [8]. LMNN generates a new distance metric with the additional constraint that points from the same class are separated by a large margin from points in different classes. Formally, assume $x_i$ is an input point associated with class $y_i$ and let $x_i$ be a true neighbor, associated with the same class $y_i$ and $x_j$ an imposter associated with a different class $y_j$. The objective of LMNN is to learn a distance matrix $M$, or an equivalent linear transformation $L$ such that:

$$\|L(x_i - x_j)\|^2 \leq \|L(x_i - x_j)\|^2 + 1$$

The objective is achieved through a convex optimization procedure.

In the proposed system, DML is applied first in order to identify a distance that satisfies the class constrains. The learned distance, or the equivalent linear transform, is then used in the generation of the adjacency matrix of the LPP algorithm. Once the manifold embedding is learned, classification of new data points is achieved by measuring the learned distance between the embedded training samples and the new data point. Experimental results suggest that the properties of the learned distance, i.e. keeping the true neighbors close and the impostures far away, are preserved through the embedding. The class is identified by the labels of the nearest neighbors.

IV. Random Projections

Due to the optimization nature of DML, an initial dimensionality reduction, such as PCA, is usually applied before training. We investigate the use of RPs primarily for two reasons. First, RPs is a linear technique and thus offers fast application through matrix multiplications. Furthermore, unlike PCA, RPs is data-independent, i.e. it is does not require training on a particular dataset and thus allows better generalization.

RPs is based on the Johnson–Lindenstrauss (JL) [9] lemma which states that, given $1>\epsilon>0$ and an integer $n$, select $k$, such that $k \geq O(\epsilon^2 \log n)$, then for every set $P$ of $n$ points in $R^d$ there exists a Lipchitz function $f: R^d \rightarrow R^k$ such that for all $x, y$ in $P$:

$$(1 - \epsilon)\|x - y\|^2 \leq \|f(x) - f(y)\|^2 \leq (1 + \epsilon)\|x - y\|^2$$

As a result, the complexity of algorithms that depend on the dimensions of their input data, such as DML, can be significantly decreased. What is even more surprising is that very efficient methods exist for generating the projection matrices.
A random projection from the original d-dimensional space to the reduced k-dimensional space can be achieved by multiplying the input vectors with a matrix $R$, where $R$ is a $d \times k$ matrix. In [9] it was shown that the elements $r_{ij}$ of $R$ can be drawn as independent and identically distributed random variables from the following distribution:

$$r_{ij} = \begin{cases} 1 \text{ with probability } 1/6 \\ 0 \text{ with probability } 2/3 \\ -1 \text{ with probability } 1/6 \end{cases}$$

The distribution in Eq. (7) is notably efficient, since it discards 2/3 of the data that correspond to multiplication by zero. Furthermore, it can be implemented using fixed point arithmetic operations consisting of only additions and subtractions if the scaling coefficient is factored out.

Even though, RPs were initially proposed for subspace dimensionality reduction, the approach was extended to manifold modeled signals in [7] where it was shown that manifold properties, such as geodesic distances, curvature of the paths, etc., are also preserved through the projection.

V. Experimental Results

Two sets of experiments were carried out. First we investigated how random projections affect the manifold learning algorithm. Then we tested how well manifold modeling with learned distance performed on face recognition, based on the system shown in Fig. 1.

The face recognition performance was measured on the AT&T dataset [13], which consists of 40 subjects showing 10 images per subject at varying conditions. For each scenario i.e. number of images per individual, 50 different random splits of the dataset were applied and the mean recognition accuracy is reported. The only preprocessing on the images was the normalization of each face vector.

Figure 2 presents the recognition accuracy of nearest neighbor classification when LPP was performed in random projections space versus the original space. The results suggest that the recognition accuracy achieved by applying LPP in the original dimensions is not compromised after dimensionality reduction via RPs when the dimensionality reduction in RP space is kept above 300.
In the second experiment, we investigated the recognition accuracy of the proposed system using a learned distance in a 500-dimensional random projected space. Table 1 shows the results of our proposed method compared to standard methods for face recognition as reported in [14]. We note that LPP generates the distance matrix using same class neighbors only. The first column of the table corresponds to the number of images per individual that were used for training while the rest were used for testing. We observe that combining DML and LPP significantly increases the recognition accuracy.

<table>
<thead>
<tr>
<th>Train Img/Individual</th>
<th>PCA</th>
<th>LDA</th>
<th>LPP</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>86.3</td>
<td>92.8</td>
<td>93.1</td>
<td>98.2</td>
</tr>
<tr>
<td>4</td>
<td>82.1</td>
<td>89.4</td>
<td>90.4</td>
<td>91.4</td>
</tr>
<tr>
<td>3</td>
<td>76.6</td>
<td>84</td>
<td>86.6</td>
<td>87.7</td>
</tr>
<tr>
<td>2</td>
<td>66.9</td>
<td>71.5</td>
<td>76.1</td>
<td>83.7</td>
</tr>
</tbody>
</table>

VI. Conclusions

In this report, we describe an object recognition system that combines random projections for efficient dimensionality reduction, distance metric learning for increased classification performance, and manifold learning for embedding the data in a low dimensional manifold. Results suggest that random projections transformation does not affect classifier performance and that the combination of distance metric learning and manifold learning can increase the recognition accuracy of a classifier. The proposed scheme can be useful for face recognition and other applications, and may be extended to different combinations of distance metric learning and manifold learning approaches.
References


