



# Supervised locality discriminant manifold learning for head pose estimation



Yong Liu, Qicong Wang\*, Yi Jiang, Yunqi Lei

Department of Computer Science, Xiamen University, Xiamen 361005, Fujian, China

## ARTICLE INFO

### Article history:

Received 20 March 2013

Received in revised form 6 April 2014

Accepted 10 April 2014

Available online 6 May 2014

### Keywords:

Manifold learning

Supervised learning

Locality discriminant regularization

Head pose estimation

Graph construction

## ABSTRACT

In this paper, we propose a novel supervised manifold learning approach, supervised locality discriminant manifold learning (SLDML), for head pose estimation. Traditional manifold learning methods focus on preserving only the intra-class geometric properties of the manifold embedded in the high-dimensional ambient space, so they cannot fully utilize the underlying discriminative knowledge of the data. The proposed SLDML aims to explore both geometric structure and discriminant information of the data, and yields a smooth and discriminative low-dimensional embedding by adding the local discriminant terms in the optimization objectives of manifold learning. Moreover, for efficiently handling out-of-sample extension and learning with the local consistency, we decompose the manifold learning as a two-step approach. We incorporate the manifold learning and the regression with a learned discriminant manifold-based projection function obtained by discriminatively Laplacian regularized least squares. The SLDML provides both the low-dimensional embedding and projection function with better intra-class compactness and inter-class separability, therefore preserves the local geometric structures more effectively. Meanwhile, the SLDML is supervised by both biased distance and continuous head pose angle information when constructing the graph, embedding the graph and learning the projection function. Our experiments demonstrate the superiority of the proposed SLDML over several current state-of-art approaches for head pose estimation on the publicly available FacePix dataset.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

Head pose estimation is an integral component of multi-view face recognition systems, driver attention monitoring and other human-centered computing applications. It already becomes a hot research topic in computer vision applications [1]. Currently, the state-of-art methods for head pose estimation have adopted intensively the technique of manifold learning and embedding [2]. The technique is based on the idea that the dimensionality of the dataset is only artificially high, and it may be described as a function of only a few underlying parameters. This technique attempts to uncover these parameters in order to find a low-dimensional representations of the data. That is to say, the data points are actually sampled from a low-dimensional manifold that is embedded in a high-dimensional ambient space. For human head pose estimation, a fundamental assumption of using this technique is that face images with varying pose angles are data points that lies on a smooth low-dimensional manifold constrained by the head pose variations. It is always believed that the manifold

models the nonlinear and continuous variations of face appearance with head pose angle, and if learned properly, new face images can then be embedded in the low-dimensional space to estimate the head poses.

The manifold learning includes nonlinear types, such as Isometric Feature Mapping (ISOMAP) [3], Laplacian Eigenmap (LE) [4] and Locally Linear Embedding (LLE) [5], and linear types, such as Locality Preserving Projection (LPP) [6] and Neighborhood Preserving Embedding (NPE) [7], etc. Typically, a quadratic objective derived from a neighborhood graph is set up and solved for its leading eigenvectors.

The nonlinear manifold learning assumes that the underlying structure of the real data is often highly nonlinear and hence cannot be accurately approximated by linear manifolds. It aims to preserve certain geometric properties among neighboring data points in the process of projecting the high-dimensional data to low-dimensional embedding. There are various different geometric properties that include the pairwise geodesic distances (Isometric Feature Mapping (ISOMAP) [3]), the local convexity (Local Linear Embedding (LLE) [5]), the local distance (Laplacian Eigenmap (LE) [4]). By incorporating prior knowledge of class labels, manifold learning methods also perform pattern classification in the feature

\* Corresponding author. Tel.: +86 15980828979.

E-mail address: [qcwang@xmu.edu.cn](mailto:qcwang@xmu.edu.cn) (Q. Wang).

space. Mostly, above algorithms are formulated as convex optimization problems. These models generally assume that the low-dimensional manifold is isometric to a convex subset of Euclidean space, and there may exist problems of high curvature of the manifold and out-of-sample extension for non-isometric manifold [8].

For nonlinear manifold learning a critical issue is the lack of a direct mapping from the input space to the manifold space, which limits the applicability of these methods. One solution to this issue was presented in [9], where the distance matrix was viewed as a kernel and some new points were embedded by using the Nystrom approximation. Unfortunately, it is a rather complex process, and time-consuming does not have a clear interpretation for the LLE [5] case. More recent linear manifold learning approaches (these can be thought as the subspace learning [10]), such as LPP [6] and NPE [7], try to overcome the out-of-sample extension problem by performing a linear approximation of the underlying manifold. The fundamental idea behind these methods is that, using a linear map manifolds can be approximated reasonably well within a local neighborhood, even though they are nonlinear structures. The benefits of linear approximations are some savings in computational time.

The disadvantage of the nonlinear manifold learning makes the out-of-sample extension tricky, while that of linearized approaches can offer only an approximation of the underlying manifold structure. Hence, we propose a novel manifold learning approach to overcome the two issues. We decompose the manifold learning as a two-step approach: graph embedding [11] for the underlying manifold learning and the regression for the learned manifold-based projective function learning. The regression can construct the direct map obtained by using discriminatively Laplacian regularized least square to project the training data and new data. By using the regression as a process of building the projection function rather than direct linear transformation from the input space to the low-dimensional embedding, different kinds of regularizers can be naturally incorporated.

In particular, Locality Sensitive Discriminant Analysis (LSDA) [12] has been proposed recently to exploit both geometric and discriminant information simultaneously in manifold learning. LSDA [12] incorporates discriminant information based on the graph Laplacian and demonstrated better performance than some other manifold learning methods. It constructs intra-class graph and inter-class graph to model the local neighborhood relationship of the data according to their labels. By making the margin maximal among data points of different classes in each local neighborhood, it can perform the discriminative ability for classification in the reduced subspace. Intuitively, we incorporate the discriminant knowledge as the supplement into the objective functions of manifold learning in order to find the optimal low-dimensional embedding for estimating head poses accurately.

Furthermore, in the regression stage of the proposed two-step approach, we develop discriminatively Laplacian regularized least square, and this is partly inspired by discriminatively regularized least square [13]. We directly embed the discriminative information as well as the local geometry of the data based on the graph Laplacian in the regularization term such that it can explore as much underlying knowledge as possible.

Meanwhile, some head pose estimation methods are proposed based on manifold learning. Raytchev applied ISOMAP-based manifold learning technique for user-independent pose estimation [14]. Fu and Huang presented an appearance-based strategy for head pose estimation using supervised Graph Embedding (GE) analysis [15]. To incorporate the pose labels that are usually available during the training phase, Balasubramanian proposed the Biased Manifold Embedding framework (BME) [16] for head pose estimation, which uses biased distance measurement to determine  $k$  nearest neighbors such that the head pose angle information can

be incorporated as the prior knowledge for estimating head pose more accurately. BME [16] uses a Generalized Regression Neural Network (GRNN) to learn the nonlinear mapping for dealing with out-of-sample data points, and applies linear multivariate regression to estimate the pose. In [17], by incorporating prior knowledge of head pose angle, BenAbdelkader proposed Supervised Manifold Learning (SML) framework for head pose estimation, and uses nonlinear mapping, cubic smooth splines and support vector regression, to estimate head pose angle from embedded face images. SML [17] performs better than other head pose estimation methods by incorporating continuous head pose angle information in the process of graph construction and graph embedding. They all demonstrated their effectiveness for head pose estimation. However, their methods fail to efficiently handle out-of-sample extension data points. In addition, they use a nonlinear mapping (e.g. GRNN and cubic smooth splines) to estimate the head poses.

Instead, we propose a novel manifold learning approach, supervised locality discriminant manifold learning (SLDML), which decomposes the manifold learning as a two-step approach: the graph embedding stage and the regression stage. Firstly, we construct an intra-class graph  $G_w$  and an inter-class graph  $G_b$  according to their head pose labels. In this way, the geometric structures and discriminant knowledge of the data can be accurately characterized by the two graphs. Based on the graph Laplacian, we then add the local discriminant term constructed by minimizing the margin of data points from  $G_w$  and maximizing the margin of data points from  $G_b$  in the optimization objective of manifold learning. Secondly, we develop discriminatively Laplacian regularized least squares which map the data to the low-dimensional reduced space for directly out-of-sample extension more effectively. Moreover, we incorporate continuous head pose angle information into all stages of the manifold learning. Meanwhile, for a new face image, we first embed it in the low-dimensional space where we determine its  $k$  nearest neighbors, and then estimate the head pose angles.

The rest of this paper is organized as follows. Section 2 describes our supervised locality discriminant manifold learning process. Experiments on public dataset are presented to show the robustness and superiority of our method over other state-of-the-art methods in Section 3. Finally, Section 4 concludes our work with a summary and introduces the future work.

## 2. Proposed method

In this section, we introduce supervised locality discriminant manifold learning approach. Specifically, for a given data set of  $n$  points in  $D$ -dimensional space, we denote  $\mathbf{X} = [(\mathbf{x}_1, z_1), (\mathbf{x}_2, z_2), \dots, (\mathbf{x}_n, z_n)]$ ,  $\mathbf{x}_i \in \mathbb{R}^D$ , where  $z_i$  is label (i.e. head pose angle) of  $\mathbf{x}_i$ . We wish to reduce the dimensionality of this data, and assume only that the data lies on a  $d$ -dimensional manifold embedded into  $\mathbb{R}^D$ , where  $d < D$ . Moreover, we assume that the manifold is given by a single coordinate chart  $f$ . We can now describe the problem formally: given data points  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^D$  that lie on a  $d$ -dimensional manifold  $M$  that can be described by a single coordinate chart  $f: M \mapsto \mathbb{R}^d$ , find  $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n] \in \mathbb{R}^d$ , where  $\mathbf{y}_i = f(\mathbf{x}_i)$ ,  $i = 1, \dots, n$ . Note that  $\mathbf{y} = [y_1, y_2, \dots, y_n] \in \mathbb{R}^1$ . Solving this problem is referred to as manifold learning, since we are trying to “learn” a manifold from a set of points. Currently, manifold learning can be cast in terms of graph construction and graph embedding based on the specific intrinsic graph that encodes certain geometric properties of the data [11].

### 2.1. Biased graph construction

Graph construction is mostly predefined artificially, and  $k$  nearest neighbors criterion is a common way to construct the graph for

Download English Version:

<https://daneshyari.com/en/article/402334>

Download Persian Version:

<https://daneshyari.com/article/402334>

[Daneshyari.com](https://daneshyari.com)