Pattern Recognition 52 (2016) 135–147

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

On incremental semi-supervised discriminant analysis

Tejas Indulal Dhamecha, Richa Singh, Mayank Vatsa*

IIIT Delhi, Okhla Phase III, New Delhi 110020, India

ARTICLE INFO

Article history:
Received 27 May 2015
Received in revised form
19 September 2015
Accepted 24 September 2015
Available online 9 October 2015

Keywords:
Incremental learning
Semi-supervised learning
Discriminant analysis
Face recognition

ABSTRACT

In various pattern classification problems, semi-supervised discriminant analysis has shown its effectiveness in utilizing unlabeled data to yield better performance than linear discriminant analysis. However, many of these semi-supervised classifiers operate in batch-mode and do not allow to incrementally update the existing model, which is one of the major limitations. This paper presents an incremental semi-supervised discriminant analysis algorithm, which utilizes the unlabeled data for enabling incremental learning. The major contributions of this research are (1) utilizing large unlabeled training set to estimate the total scatter matrix, (2) incremental learning approach that requires updating only the between-class scatter matrix and not the total scatter matrix, and (3) utilizing manifold regularization for robust estimation of total variability and sufficient spanning set representation for incremental learning. Using face recognition as the case study, evaluation is performed on the CMU-PIE, CMU-MultiPIE, and NIR-VIS-2.0 datasets. The experimental results show that the incremental model is consistent with the batch counterpart and reduces the training time significantly.

1. Introduction

Discriminant analysis (DA) [1] based classifiers have found their utility in wide range of problems such as image retrieval [2] and face recognition [3]. Linear discriminant analysis (LDA) [4] and its variants have been efficiently used in various pattern classification problems [5–8]. Some of the most interesting successors of LDA are kernel LDA [9], maximum margin criterion based discriminant function [10], and graph-embedding [6,11] based algorithms.

The formulations of DA techniques, typically, require labeled training data. In certain applications, such as image retrieval and object classification, it is difficult to obtain large labeled data. However, large amount of unlabeled data is easily available. To address this aspect, researchers introduced semi-supervised learning in discriminant analysis [12–15]. The paradigm utilizes labeled as well as large amount of unlabeled training data to learn the model [16]. Semi-supervised learning is very important in addressing the labeled data related limitation, as it learns a model from labeled as well as large amount of unlabeled training data. Therefore, semi-supervised learning approaches have been proposed in discriminant analysis. Generally, existing semi-supervised incremental learning algorithms first learn the model using labeled data, which is followed by classification of unlabeled data [14,15,17]. Either a new classification model is learned or existing model is updated using the confidently classified unlabeled data samples. Therefore, these set of algorithms create pseudolabeled data from unlabeled data, and use an existing supervised learning framework. However, this approach requires to iteratively learn the model which might be time consuming. Cai et al. [12] proposed semi-supervised discriminant analysis (SSDA) by utilizing unlabeled data for learning the regularized total scatter matrix. The regularization is performed using graph Laplacian of unlabeled training set which encodes the manifold assumption. Few other related semi-supervised learning algorithms are summarized in Table 1. Semi-supervised learning can be utilized in various research areas ranging from bioinfomatics, speech recognition, natural language parsing, and spam filtering [16,18].

Both the types of discriminant analysis algorithms, supervised (e.g. LDA) or semi-supervised (e.g. SSDA), are usually trained in batch mode. In many real world applications, it is likely that whole labeled training set is not available before hand; rather the training data is obtained incrementally. The batch learning algorithms have a major limitation related to very limited provision for updating discriminant components by incorporating the newly available training samples only. To obtain a new model, the discriminant classifier has to be learned from the merged data, i.e. both original and incremental training data. Since the core of every discriminant analysis objective function contains an eigenvalue decomposition problem, learning a new classifier from merged data has cubic time complexity. Further, as SSDA encodes the data in the form of graphs, graph adjacency matrix has to be obtained

* Corresponding author. Tel.: +91 9654653404; fax: +91 11 26907405.
E-mail addresses: tejasd@iiitd.ac.in (T.I. Dhamecha), rsingh@iiitd.ac.in (R. Singh), mayank@iiitd.ac.in (M. Vatsa).

http://dx.doi.org/10.1016/j.patcog.2015.09.030
0031-3203/

Table 1
Literature review of the related research pertaining to incremental learning and semi-supervised learning algorithms related to discriminant analysis.

Algorithm	Year	Description
Semi-supervised leaning		
SSDA [12]	2007	Semi-supervised discriminant analysis based on LDA which uses unlabeled set for estimating total scatter
SELF [13]	2010	A semi-supervised extension of local Fisher discriminant analysis that preserves the global structure of the unlabeled data
SSGDA [14]	2011	Confidently classified unlabeled data samples are utilized with pseudolabels in generalized discriminant analysis
Byun [15]	2012	Utilizes pseudolabels of only those unlabeled samples that are expected to reduce errors
Incremental leaning		
IPCA [19]	2000	Proposed algorithm for merging eigenpaces of total scatter matrices
ILDA-Pang [20]	2005	Incrementally updates between- and within-class scatter
GSVD-ILDA [21]	2008	Incremental version of generalized singular value decomposition based LDA [50]
LS-ILDA [22]	2009	Formulates ILDA in terms of least square solution by incrementally updating total scatter of mean centered data matrix
IDCC [17]	2010	Incremental discriminant canonical correlation analysis by adapting the sufficient spanning set based merging of eigenspace [19]
ILDA [23,24]	2007, 2011	Merging eigenspaces of between- and within-class scatter of existing and new batch for updating model
I-CLDA [25,26]	2012	Incremental complete linear discriminant analysis utilizing QR decomposition to obtain orthonormal projection directions
ISDA (subclass) [27]	2012	Extension of ILDA [24] to incremental subclass discriminant analysis
ILDA-KT [28]	2012	Addressing concept drift in incremental learning using knowledge transfer
LS-LDA-CD [29]	2013	Addresses concept drift issue in least square LDA [22]
Chunk-IDR/QR [30]	2015	A time-efficient version of IDR-QR [31]
ILDA/QR [32]	2015	Utilization of QR decomposition of data matrix for incremental learning
Proposed ISSDA	–	Extension of ILDA to semi-supervised discriminant analysis with reduced time complexity

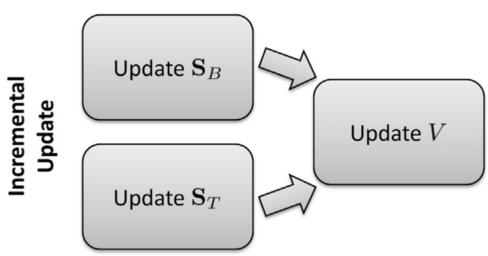

Fig. 1. Traditional incremental discriminant approaches, such as Kim et al. [23,24] and Lamba et al. [27], update between-class and overall variability. New eigenmodels of $\mathbf{S}_B$ and $\mathbf{S}_T$ are learned from incremental batch, which are merged with corresponding existing eigenmodels. Discriminating components $\mathbf{V}$ are obtained from merged eigenmodels of $\mathbf{S}_B$ and $\mathbf{S}_T$.

from the merged data to update the model. This presents an additional challenge for learning a new model. The challenge can be addressed using incremental learning [22–24] where new training samples are incrementally incorporated into the classification model. The motivation of incremental learning for discriminant analysis is to be able to *update* the existing model using the newly available training samples with significantly less time complexity. Some existing contributions pertaining to incremental learning are summarized in Table 1. Kim et al. [23,24] and Lamba et al. [27] utilized the eigenspace merging algorithm [19] to formulate incremental linear discriminant analysis (ILDA) and incremental subclass discriminant analysis, respectively. As shown in Fig. 1, both the algorithms use the new training samples to update the between-class scatter matrix and the total scatter matrix individually, and learn the discriminating components. Liu et al. [22] proposed an incremental learning algorithm based on the least square formulation of LDA. Time complexity of the algorithm proposed in [22] is less than that of [23,24], however the space complexity of the former is more as it requires to store the entire data matrix or total scatter matrix as part of the classification model.

To mitigate the above mentioned challenges, this paper presents an incremental semi-supervised discriminant analysis (ISSDA) algorithm. We address the problem with two reasonable assumptions: large unlabeled training data is available offline and labeled data is received incrementally. The proposed algorithm aims at reducing the computational complexity of the incremental update process by utilizing the unlabeled dataset for robust data statistics estimation. The major contributions of this paper are:

- showcasing that large unlabeled training set can be leveraged to efficiently estimate the total scatter matrix,
- utilization of manifold regularization of robust estimation of total scatter matrix, and
- sufficient spanning set representation [23,24] based incremental learning approach which requires to update only the between-class scatter matrix and not the total scatter matrix.

The effectiveness of the proposed algorithm is evaluated for face recognition application. The performance is evaluated by comparing the accuracy, time and consistency of the proposed incremental algorithm with the corresponding batch learning model. Evaluations to understand the effects of the manifold regularizer and unlabeled data size are also performed. Further, the effect of updating the model with incremental batch consisting of samples of new classes is also studied.

2. Incremental semi-supervised discriminant analysis

Discriminant analysis based approaches have a fundamental objective of maximizing the inter-class variation and minimizing the intra-class variation. In case of linear discriminant analysis, inter-class variability is modeled in terms of between-class scatter matrix $\mathbf{S}_B$ and intra-class variability is modeled in terms of within-class scatter matrix $\mathbf{S}_W$ [4]:

$$\mathbf{S}_B=\sum_{i=1}^{c} n_i(\boldsymbol{\mu}_i-\boldsymbol{\mu})(\boldsymbol{\mu}_i-\boldsymbol{\mu})^T \quad \text{and} \tag{1}$$

$$\mathbf{S}_W=\sum_{i=1}^{c}\sum_{x_j\in\mathbf{X}_i}(x_j-\boldsymbol{\mu}_i)(x_j-\boldsymbol{\mu}_i)^T \tag{2}$$

Here, c is the number of classes, n is the total number of samples, n_i is the number of samples in i^{th} class, $\mathbf{X}=[\mathbf{X}_1\ldots\mathbf{X}_c]$ is the data matrix, $\mathbf{X}_i$ is the set of samples belonging to i^{th} class, $\boldsymbol{\mu}_i(=(1/n_i)\sum_{x_j\in\mathbf{X}_i}x_j)$ is the mean of i^{th} class, and $\boldsymbol{\mu}(=(1/n)\sum_{i=1}^{c}\sum_{x_j\in\mathbf{X}_i}x_j)$ is the mean of all the data samples. The objective is to find the set of projection directions $\mathbf{V}$ such that,