J. Vis. Commun. Image R. 56 (2018) 207–219

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

# Discriminative kernel-based metric learning for face verification

Siew-Chin Chong [a], Thian-Song Ong [a], Andrew Beng Jin Teoh [b,*]

[a] Faculty of Information Science and Technology, Multimedia University, Malaysia
[b] School of Electrical and Electronic Engineering, Yonsei University, South Korea

ARTICLE INFO

Article history:
Received 22 March 2017
Revised 17 August 2018
Accepted 22 September 2018
Available online 22 September 2018

Keywords:
Face verification
Metric learning
Kernel machine
Discriminant analysis

ABSTRACT

This paper outlines a simplistic formulation for doublet constrained discriminative metric learning framework for face verification. The Mahalanobis distance metric of the framework is formulated by leveraging the within-class scatter matrix of the doublet and a quadratic kernel function. Unlike existing metric learning methods, the proposed framework admits efficient solution attributed to the convexity nature of the kernel machines. We demonstrate three realizations of the proposed framework based on the well-known kernel machine instances, namely Support Vector Machine, Kernel Ridge Regression and Least Squares Support Vector Machine. Due to wide availability of off-the-shelf kernel learner solvers, the proposed method can be easily trained and deployed. We evaluate the proposed discriminative kernel-based metric learning with two types of face verification setup: standard and unconstrained face verification through three benchmark datasets. The promising experimental results corroborate the feasibility and robustness of the proposed framework.



## 1. Introduction

Over the years, face verification has captivated huge interest among the researchers. Various techniques of face verification have been actively studied and invented due to the application demands such as nonintrusive access control in security systems, verification for business transactions, military and law-enforcement applications and others. A face verification system will decide whether a pair of test faces is from the client (correct identity to be claimed as) or the imposter (fault identity to be assigned to). Human faces are immensely challenging due to the significant variations in appearance, which may be caused by varying poses, aging, lighting, expression and others. Moreover, face images captured in an unconstrained environment (the factors of variations in appearance are not controlled such as photos taken from news articles) increase the difficulty in verification task, not to mention the situation of applying restricted protocol [51], which usually consist of large intra-class variations and no additional information about the person.

In addition, limited training data are available in the real-world environment and it may be impractical for those face verification techniques that rely heavily on full label training information. The main reasons are the waste in the computational costs and storage and the difficulty of obtaining full information from a given individual in certain situations. Contrary to current deep learning approach that exploits gigantic outside training data [2,30], the restricted protocol place the emphasis to examine the learning capacity of the innovated algorithms rather than the performance gain attributed to the training data volume. Hence, it is important to better simulate the real-world scenarios by inventing the face verification techniques without relying on the facial side information. In practice, the unconstrained yet restricted face verification remains beneficial, but not limited to the applications of, (1) Alzheimer's patient face verification: Alzheimer's patients who lost their way home do not have the ability to remember their personal information due to memory lost. The facial image is the only data to verify the person against the government database. (2) Surveillance face verification: The facial image of a murder/robber suspect caught on a CCTV camera with no extra information of the individual is verified against the government database. The Boston Marathon bombing case in 2013 [66] is an example where the law enforcement failed to match the unconstrained face image of the bombing suspect against the government database without any additional information of the bombing suspect.

Metric learning plays an imperative role in contemporary face verification as well as in many machine learning problems [1–3]. Metric learning dedicates to learn a Mahalanobis distance metric $d_M = \mathbf{x}_j^T\mathbf{M}\mathbf{x}_j$ where $\mathbf{M}$ is the Mahalanobis matrix, from the training samples in measuring the similarity score between $\mathbf{x}_i$ and $\mathbf{x}_j$ by enhancing the similarity of matched pairs and suppressing the

* Corresponding author.
E-mail addresses: chong.siew.chin@mmu.edu.my (S.-C. Chong), tsong@mmu.edu.my (T.-S. Ong), bjteoh@yonsei.ac.kr (A.B.J. Teoh).

https://doi.org/10.1016/j.jvcir.2018.09.017
1047-3203/

similarity of the mismatched pairs. Various metric learning formulations can be established based on the objective functions that designed for the specific task [4–8].

The problem of learning a good metric from training data is of critically essential. A good metric learning method should be able to reduce the influence of non-informative dimensions while emphasize on relevant dimensions [9]. There are two criteria to be considered when learning the Mahalanobis matrix. The first criterion is that the labels of the training data should be as weak as possible. Due to the fact that it is always difficult to obtain strong label such as class label of the training data in practice, it is preferable to design a metric learning algorithm that is based on the data pair labels (similarity or dissimilarity), which is more commonly available. For methods that depend much on the additional information, for example, the identity information of a person would be impractical in some applications such as verifying a stranger whose identity is not stored in the data bank or an intruder who repeatedly abuses the system. The second criterion in designing a good metric learning algorithm is that the algorithm should be computationally efficient. In this regard, the metric learning algorithm should not be heavy in computation.

On the other hand, kernel machines [10–13] for classification and regression have had a huge impact to machine learning community over the past several decades [14,15]. Kernel machines have been analyzed from the metric learning perspective lately and the intrinsic relationship of them is explored. A number of efficient algorithms have emerged, such as using metric learning notion in classifier design [16] and integrating support vector machine in metric learning [17].

Distance metric learning for face verification is the focus of this paper. In order to mitigate the problems of the existing distance metric learning algorithms for face verification, specifically with unconstrained images, a new distance metric learning framework is put forward. In this paper, the metric learning algorithm and the widely available kernel machines had coalesced into a simplistic metric learning framework for face verification to garner the advantages of both worlds. The Mahalanobis distance metric of this work is framed by leveraging the within-class scatter matrix of the doublet and a quadratic kernel function. To be specific, doublets are constructed from the weak label training pairs and reformulated as a quadratic kernel function through the popular kernel machine instances such as SVM, LSSVM and KRR. As such, the metric learning problem is transformed to classification or regression problem. Three realizations of the proposed framework are developed based on the well-known kernel machine instances such as Support Vector Machine [10], Least Squares Support Vector Machine [11] and Kernel Ridge Regression [12]. The decision to propose three different realizations is mainly due to the weaknesses and strengths of the kernel machine instances, which can be further improved and absorbed respectively, as discussed in Section 4.7.

In Section 2, we will briefly discuss the existing metric learning algorithms and kernel machines designed for face domain. It is followed by the motivation and contributions of this work in Section 3. The formulation of the three proposed realizations based on the proposed framework is given in Section 4. Section 5 reports the experimental results through the use of three publicly available datasets, including two standard datasets and one unconstrained dataset. Finally, a concluding remark and future work are presented in Section 6.

## 2. Related work

This section is primarily for the background of this study. An insight of various kernel machine learning and distance metric learning approaches in face verification is discussed. In addition, the intrinsic relationship among the kernel machine instances and the distance metric learning are elaborated.

### 2.1. Kernel machine learning

Kernel machine instances such as Support Vector Machine (SVM) [10], Least Square Support Vector Machine (LSSVM) [11], Kernel Ridge Regression (KRR) [12] and others [13,14,18–20], have a different learning paradigm than that of the metric learning. Among kernel machines, SVM is one of the most prominent classifiers due to its excellent predictive performance. The maximum-margin decision hyperplanes of SVM are determined based on a subset of the training data, called support vectors. SVM solves a convex quadratic programming problem that involves inequality constraints. SVM is convex but lacking of closed form solution. Thus, iterative procedure is required. However, from the metric learning point of view, the margin of SVM resembles between-class distance but the within-class distance is ignored [16].

LSSVM reformulates the SVM in least square notion and thus only involves the equality constraints. The solution can be simplified by merely solving a set of linear equations that yields close-form solution. Extensive empirical studies [21] have shown that LSSVM is comparable to SVM in terms of generalization performance. Both SVM and LSSVM can be made non-linear by adopting Mercer's compliant kernel functions.

KRR is one of the classical methods for regression, which integrates the ridge regression model with the kernel functions. KRR is quite similar to support vector regressor [22], except that different constraint objective function is being optimized. Furthermore, KRR is equivalent with regularization networks and Gaussian processes when the bias is zero [23].

### 2.2. Metric learning

In this section, we give a brief account of metric learning techniques dedicated to face verification as well as the prior arts that relevant to our proposed method, as tabulated in Table 1. For extensive review of metric learning, we refer the readers to [24].

Linear Discriminant Analysis (LDA) [25] is a classical tool for face recognition. LDA can be viewed as a supervised metric learning method over the intra-personal linear subspace, on its objective in minimizing the average distance between the similar pairs. LDA is found has connection with other machine learning tools such as LDA is equivalent to a least squares regression to the labels [32], and penalized LDA is indeed equivalent to ridge regression [33]. Recently, a variant of LDA devoted for weak label has been proposed by Kan et al. [34] to estimate the within-class and between-class scatter matrices through the use of side-information.

Another celebrate instance is Large Margin Nearest Neighbor (LMNN) that proposed by Weinberger et al. [4]. LMNN strives to preserve the consistency in the neighborhood of data and to keep a large margin at the boundaries of different categories. However, LMNN leverages full label information which could be restrictive in practice. In addition, LMNN might prone to overfitting due to the absence of regularizer. Shen et al. [26] improves LMNN by mere using weak label. Another extended version of LMNN is presented by Kumar et al. [35] for transformation invariant classification.

On the other hand, Information Theoretic Metric Learning (ITML) [27] is formulated as a special instance of Bregman optimization problem to maximize the differential entropy of a multivariate Gaussian subject to constraints on the associated Mahalanobis distance. The theory of LogDet divergence regulazation being introduced in ITML is important in Mahalanobis distance metric learning methods [63]. Hieu et al. [5] develop a fast