Neurocomputing xx (xxxx) xxxx–xxxx

Contents lists available at ScienceDirect

Neurocomputing

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

Image multi-label annotation based on supervised nonnegative matrix factorization with new matching measurement

Xu Jia[a,*], Fuming Sun[a], Haojie Li[b], Yudong Cao[a], Xing Zhang[a]

[a] School of Electronic and Information Engineering, Liaoning University of Technology, 169 Shiying Street, Guta District, Jinzhou, Liaoning, PR China
[b] School of Software, Dalian University of Technology, Dalian, China

ARTICLE INFO

Communicated by: Yanwei Pang

Keywords:
Nonnegative matrix factorization
Multi-label annotation
Supervised learning
Matching measurement

ABSTRACT

Nonnegative Matrix Factorization (NMF) has been attracting many scholars in the fields of pattern recognition and data mining to study it since its inception. To date, a large number of variant methods have been proposed and applied in image retrieval and image Single-Label Annotation (SLA) successfully. However, the effectiveness of NMF for Multi-Label Annotation (MLA) encounters difficulties and is still an open topic. To meet this goal, this paper proposes a supervised NMF with new matching measurement to enhance MLA accuracy. In contrast with other NMF algorithms with sparse or discriminant constraints, the proposed NMF algorithm implements a supervised training method while integrates feature dimension reduction. What's more, we improve a novel matching measurement function by considering positive and negative samples respectively, which is proved to be more suitable for MLA. In addition, the proposed NMF object function is solved by using the projected gradient method, and image annotation can be achieved. Experiments results on NUSWIDE dataset showed that the proposed algorithm can achieve strong performance compared with existing algorithms in terms of False Rejection Rate (*FRR*) and False Acceptance Rate (*FAR*).

1. Introduction

For image feature analysis and dimensionality reduction, there have been many canonical matrix decomposition methods used previously, such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Independent Component Analysis (ICA) [1]. In these methods, each element in the factorized matrices may be positive or negative after decomposition. From algebraic perspective the factorized positive or negative elements are reasonable, but the negative component in real applications has very little meaning. For example, image pixel values must be non-negative, and similarly, negative value is also meaningless for document statistics. In order to make the result more interpretable, a novel paradigm, called Non-negative Matrix Factorization (NMF), has attracted more attention in recent years. This method was first proposed by *Lee* and *Seung* [2]. Its main difference from the canonical matrix decomposition methods mentioned previously is that non-negativity constraint is imposed on the factorized matrices in NMF, i.e. for a given nonnegative matrix $\boldsymbol{F} \in \boldsymbol{R}^{n\times m}$, two matrices can been found $\boldsymbol{W} \in \boldsymbol{R}^{n\times r}$ and $\boldsymbol{H} \in \boldsymbol{R}^{r\times m}$ which meet Eq. (1) [3],

$$\boldsymbol{F} \approx \boldsymbol{WH}, \quad s.\ t.\quad w_{i\mu},\ h_{\mu j} \ge 0 \tag{1}$$

where $\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1r} \\ w_{21} & w_{22} & \cdots & w_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nr} \end{bmatrix}$, $\boldsymbol{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ h_{r1} & h_{r2} & \cdots & h_{rm} \end{bmatrix}$, $0 < i \le n$, $0 < j \le m$, $0 < \mu \le r$.

In recent years, NMF has been widely used in pattern classification and data mining, such as feature dimension reduction or SLA. But, it may not guaranteed that high classification accuracy can be achieved using the original NMF method, which only requires factorized matrix nonnegative. Therefore, some NMF variants, which imposed other constraints besides non-negativity, were presented [4,5]. However, for MLA problem, there has been little discussion on whether these NMF algorithms are effective or not. To address this issue, we propose a new improved NMF algorithm for image MLA issue in this paper. The main contributions of this paper are summarized as follows: (1) Besides the process of feature dimension reduction, a supervised training process is achieved in the proposed NMF model; (2) In supervised training stage, a new matching measurement method which adopts two different matching functions on positive and negative samples respectively is proposed as constraint term in NMF; (3) Image annotation can be directly implemented by using the proposed NMF without classifier.

The remainder of this paper is organized as follows. In Section 2, some related works are addressed. In Section 3, an improved NMF

* Corresponding author.
E-mail address: gbjdjiaxu@163.com (X. Jia).

http://dx.doi.org/10.1016/j.neucom.2016.09.052
Received 22 August 2015; Received in revised form 17 December 2015; Accepted 25 September 2016
Available online xxxx
0925-2312/

method with new matching measurement for MLA is proposed. Section 4 gives a projected gradient algorithm to solve the proposed NMF objective function. We describe experiments that were conducted on multi-label image data sets to prove the effectiveness of the proposed algorithm in Section 5. Finally, conclusions are drawn in Section 6.

2. Related work

In recent years, many algorithms based on NMF have been applied in SLA [6,7]. Initially, the basic NMF which is only required to meet Eq. (1) was used, where the columns of factorized matrices $\boldsymbol{W}$ and $\boldsymbol{H}$ represent the feature basis and coefficient vectors respectively [8,9], and these coefficient vectors can be seen as new feature vectors. However, as mentioned previously, it is hard to acquire good annotation results only using basic NMF with sole non-negativity constraint. Therefore, additional constraint $P(W, H)$ needs to be added as regularization term in NMF as shown in Eq. (2).

$$J(W, H) = \min\{\frac{1}{2}\|\boldsymbol{F} - \boldsymbol{WH}\|_F^2 + \lambda P(W, H)\}, \quad s.\ t. \quad w_{i,\mu},\ h_{\mu,j} \geq 0 \tag{2}$$

In general, the constraints can be divided into the following four classes for image classification or annotation:

2.1. Sparse NMF

Sparse NMF requires sufficient sparsity of the factorized matrices $\boldsymbol{W}$ and $\boldsymbol{H}$, i.e. the l_0-norm of both matrices must be minimized as much as possible. Of course, the l_0-norm solution is a NP-hard problem and difficult to work with [10]. So, this problem is often replaced by solving the l_1-norm according to compressed sensing theory [11]. A previous study [12] enforced the sparse constraint on the basis matrix $\boldsymbol{W}$, which meant that each basis vector represents a partial-based or local-based image feature. Thereafter, it could detect the materials presented in images. Similar to sparse constraints, Laplacian regularization was presented to preserve the dependent properties of locality and similarity between the instances in [13].

2.2. Orthogonal NMF

In order to reduce the redundancy between different feature bases, Orthogonal NMF was proposed in [14]. It requires that the basis vectors are orthogonal, namely $\boldsymbol{W}^T\boldsymbol{W} = \boldsymbol{I}$. In order to preserve the local neighborhood geometrical structure within the reduced space, *Wen* proposed a neighborhood-preserving orthogonal projective non-negative matrix factorization (NPOPNMF) method for feature extraction [15]. Besides image classification, Orthogonal NMF can also be applied in image clustering [16], and two new methods, i.e., the EM-like algorithm and the augmented Lagrangian approach were introduced to solve Orthogonal NMF function.

2.3. Discriminant NMF

For image classification and annotation, supervised learning is more common than unsupervised learning. So, in order to maximize the distance between the images annotated by different labels, a discriminant constraint can be often added in NMF. In practical applications, such as face recognition, *Xue* [17] and *Zafeiriou* [18] integrated linear discriminant analysis into the NMF algorithm. However, Discriminant NMF often does not guarantee convergence to a stationary limit point. It was addressed in [19] and a novel DNMF method using projected gradient is presented. Besides classification performance and convergence, the computation complexity also needs to be considered. Thus, a modified Spectral Projected Gradient was proposed in [20], which considerably alleviated the computational problem.

2.4. NMF on manifold

Manifold learning methods can discover the intrinsic geometrical structure of a high-dimensional data space. For image clustering, a Hyper-graph regularized NMF was proposed in [21], which incorporated a manifold regularizer into standard NMF framework, leading to novel performance.

All of the above NMF methods have been applied to image clustering or annotation, and achieved good results. Among these methods, the discriminant NMF is obviously superior for supervised learning, because it imposes a discriminant constraint on the coefficient matrix $\boldsymbol{H}$ besides non-negativity, which means that the similarity between images with different labels would be as low as possible. After factorization, all column vectors in the factorized coefficient matrix $\boldsymbol{H}$ can be considered as new feature vectors, these new features need to be input into a classifier and output annotation results. In other words, image annotation task needs classification process besides matrix factorization. Clearly, it will take a certain time. In order to improve real-time performance, therefore, the proposed algorithm implemented both feature dimensional reduction and classification simultaneously through solving the NMF function, where a new matching measurement has been added as regularization term for MLA. To the best of our knowledge, the problem has not been discussed to date. Below, we focus our efforts to target the issues.

3. Improved NMF with new matching measurement

Convex function optimization problem is often used to resolve the NMF with the objective function rewritten as Eq. (2). Clearly, for MLA problem, the proposed NMF algorithm should meet the following two conditions: (1) The proposed constraint term should contribute to image classification; (2) The objective function must be convex. Based on the above analysis, for the objective Eq. (2), it is important to verify whether the constraint term $P(W, H)$ is reasonable and effective. For NMF methods with sparse, orthogonal and manifold constraints, the factorized matrices reflect some important properties of the original matrix from different perspectives, but no guarantee can be made that these properties will be helpful for MLA. It has been observed that the discriminant NMF has advantages over other methods, the acquired feature vectors have high similarity within the same class and dissimilarity between different classes, where the similarity and dissimilarity are often measured by Euclidean distance or Manhattan distance. However, it has not been confirmed whether these distance measurements are suitable for image annotation.

Note that the real-time performance of MLA is another key issue. From Section 2, a proper classifier is often needed when annotating an image using NMF. No doubt designing a classifier is a trouble thing. It is a very natural idea, shall we only use NMF model to annotate image directly without designing a classifier? We will solve this problem.

For the MLA problem, there is a set $\{l_1, l_2, \ldots, l_r\}$ of class labels, where r is the number of image labels in the database. Let χ be the instance space, which can be a finite dimensional space R^n, and the image $\boldsymbol{I}_i$ can be represented as vector $\boldsymbol{f}_i$, $\boldsymbol{f}_i \in \chi$, and annotated as $\boldsymbol{L}_i = [l_{1i}, l_{2i}, \ldots, l_{ri}]^T$, where if the label l_k is assigned to $\boldsymbol{I}_i$, $l_{ki} = 1$, otherwise $l_{ki} = 0$. The purpose of MLA is to learn a map function $\boldsymbol{f} \to \boldsymbol{L}$ from the training samples, and predict the labels vector $\boldsymbol{L}_j$ accurately and quickly for given image $\boldsymbol{I}_j$. To achieve this goal, in the proposed NMF algorithm we will improve the meaning of factorized matrix $\boldsymbol{H}$, i.e. each element h_{ki} in matrix will not represent new image feature, but decide whether the image can be annotated by label i.

Assumed that there is a set of vectors of training samples $\boldsymbol{F} = [\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_m]$, which are acquired by extracting features from m training images, and their labels $\boldsymbol{L} = [\boldsymbol{L}_1, \boldsymbol{L}_2, \ldots, \boldsymbol{L}_m]$ are obtained as prior knowledge according to supervised learning theory. In the proposed NMF for MLA, Ideally, when $\boldsymbol{F}$ is decomposed as $\boldsymbol{F} \approx \boldsymbol{WH}$, the columns of $\boldsymbol{W}$ and $\boldsymbol{H}$ are seen as basis vectors and label vectors