Available online at www.sciencedirect.com
ScienceDirect
Neurocomputing 71 (2008) 1824–1831
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Total variation norm-based nonnegative matrix factorization for identifying discriminant representation of image patterns

Taiping Zhang, Bin Fang*, Weining Liu, Yuan Yan Tang, Guanghui He, Jing Wen

College of Computer Science, Chongqing University, Chongqing 400044, PR China

Available online 29 February 2008

Abstract

The low-rank approximation technique of nonnegative matrix factorization (NMF) is emerging recently for finding parts-based structure of nonnegative data based on minimizing least-square error (L_2 norm). However, it has been observed that the proper norm for image processing is the total variation norm (TVN) other than the L_2 norm, and image denoising methods applying TVN can preserve clearer local features, such as edges and texture than L_2 norm. In this paper, we propose a robust TVN-based NMF algorithm for identifying discriminant representation of image patterns. We provide update rule in optimality search process and prove mathematically convergence of the iteration. Experimental results show that the proposed TVNMF is more effective to describe local discriminant representation of image patterns than NMF.

Keywords: Nonnegative matrix factorization; Total variation norm; Discriminant representation of image patterns

1. Introduction

Data analysis is to reveal as low as possible dimensional structure of patterns observed in high dimensional spaces [11,5,13,20–22]. A fundamental problem in data analysis is to find a suitable low-rank representation of the data, an optimal low-rank representation typically makes latent structure in the data explicit.

PCA and ICA [3,23] as data analysis methods aim at learning holistic, not parts-based, representation of data, where low-rank representation is achieved by discarding least significant components. The resulting components are global interpretations but these methods are unable to extract basis components manifesting local features such as image edges. Whereas local features are important for pattern recognition and classification, hence learning local parts-based representation of visual patterns has been a research hotspot in computer vision for a long time.

Nonnegative matrix factorization (NMF) as a learning technique for local parts-based representation of patterns is recently developed for data analysis with increasingly popular in dimension reduction, compression, feature extraction and computer vision applications [1,24]. The NMF method is designed to capture alternative structures inherent in the data, and possibly to provide more biological insight. For example, NMF can yield a decomposition of human faces into parts reminiscent of features such as lips, eyes, nose, etc.

However, it has been observed that the proper norm for image processing is the total variation norm (TVN) and not the L_2 norm [17,16], and the TVN approach can preserve finer scale image features, such as edges and texture than L_2 norm in the application of image denoising [2].

In this paper, we propose a robust algorithm of NMF by minimizing TVN instead of original L_2 norm (Euclidean distance) to identify discriminant image patterns. Compared with L_2 norm-based NMF approach, the TVN-based NMF algorithm is able to present parts-based discriminant representation of image patterns. Reconstruction simulation for image patterns by the proposed method also showed advantage over original NMF technique.

This paper is structured as follows. In Section 2 we describe the general NMF. Section 3 discusses how to incorporate TVN minimization into the NMF algorithm. Section 4 provides experimental results that verify our algorithm. Finally, conclusion is given in Section 5.

*Corresponding author. Tel.: +86 23 65112784; fax: +86 23 65102502.
E-mail addresses: tpzhang@cqu.edu.cn (T. Zhang), fb@cqu.edu.cn (B. Fang).

0925-2312/$ - see front matter
doi:10.1016/j.neucom.2008.01.022

2. Nonnegative matrix factorization

NMF is a linear, nonnegative data representation technique. Given the nonnegative $n \times m$ matrix V and the constant rank r, the NMF [9] finds a nonnegative $n \times r$ matrix W and another nonnegative $r \times m$ matrix H such that the product WH approximate to V, that is

$$V \approx WH \quad \text{or} \quad V_{ij} \approx \sum_{k=1}^{r} W_{ik} H_{kj}.$$

This can be interpreted as follows: each column of matrix W contains a basis vector while each column of H contains the weights needed to approximate the corresponding column of V. Hence the product WH can be regarded as a compressed form of the data in V. The constant rank of the factorization W and H is generally chosen such that $(n+m)r < nm$.

NMF has become increasingly popular in machine learning, signal processing, pattern recognition and computer vision applications [1,24]. One reason for the popularity is that NMF can produce parts-based discriminant representation where PCA and ICA are difficult to interpret such results. Another is that NMF codes naturally favor sparse, the entries of W and H are generally combined additively (not subtracted). Hence, these constraints might be useful for extracting parts-based representation of image patterns with low feature dimensionality [7,8].

In order to find an approximate factorization $V \approx WH$, the optimal choice of matrices W and H are defined to be those nonnegative matrices that minimize the squared error (Euclidean distance) between V and WH as follows:

$$E(W,H) = \| V - WH \|^2 = \sum_{i,j} (V_{ij} - (WH)_{ij})^2.$$

Although the minimization problem is convex for W and H, respectively, it is not convex for both simultaneously. To deal with the problem, the authors [10] devised an iteration rule that is somewhat simpler to implement and showed good performance:

$$H_{a\mu} \leftarrow H_{a\mu} \frac{(W^{\mathrm{T}} V)_{a\mu}}{(W^{\mathrm{T}} WH)_{a\mu}}, \quad W_{ia} \leftarrow W_{ia} \frac{(VH^{\mathrm{T}})_{ia}}{(WHH^{\mathrm{T}})_{ia}}. \tag{1}$$

In fact, the update rule for optimality search procedure is based on gradient descent computation.

3. NMF with minimizing TVN

In this section, we describe the basic idea to incorporate TVN into NMF framework for extraction local discriminant representation of image patterns in details, and derive the whole algorithm including iteration rule and convergence proof for real-world image processing.

3.1. Total variation norm

We denote the Euclidean space $R^{N \times N}$ by X to define the discrete TVN, where $N \times N$ is the size of two-dimensional matrices of images. Now we introduce a discrete gradient operator. If $u \in X$, the gradient ∇u is a vector given by

$$(\nabla u)_{i,j} = ((\delta_x u)_{i,j}, (\delta_y u)_{i,j})$$

with

$$(\delta_x u)_{i,j} = \begin{cases} u_{i+1,j}, -u_{i,j}, & i < N, \\ 0, & i = N, \end{cases}$$

$$(\delta_y u)_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j}, & j < N, \\ 0, & j = N. \end{cases}$$

The TVN of u is defined by

$$J(u) = \sum_{1 \leqslant i,j \leqslant N} |(\nabla u)_{i,j}|$$

with $|x| = \sqrt{x_1^2 + x_2^2}$ for every $x = (x_1, x_2) \in R^2$.

The TVN has been introduced in image processing first in Refs. [17,16]. The authors pointed out that the proper norm for image processing is the TVN and not the L_2 norm. Methods on TVN have been widely used to solve denoising problems, and remain one of the active areas of research in mathematical image processing.

TVN has two fundamental properties: edge-preserving and multiscale additive signal decomposition, the theoretical justification for the properties was provided in [19]. Compared with L_2 norm, the TVN is able to preserve finer scale image features such as edges and texture while denoising [2]. This fine scale details offer advantages to pattern recognition. Fig. 1 demonstrates the comparison result by applying these two techniques for image denoising. It is obvious that restoration image by TVN approximation looks better than the "same" L_2 approximation. The "same" means approximation procedure subject to the same constraints [17].

3.2. NMF by minimizing TVN

In order to gain desired characteristics, several researchers suggested extension and modification of the original NMF model, including auxiliary constraints on W and H or penalty terms to enforce auxiliary constraints, and extension of the cost function for original problem. Li [12] noted that NMF could only found global features from the ORL face image database and suggested an extension they call local nonnegative matrix factorization (LNMF). Hoyer [7,8] extended the NMF to include the option to control sparseness explicitly.

In this paper, we replace the cost function (least-square error minimization) of basis model with minimizing TVN for identifying local detailed representation in image patterns. The model can be mathematically expressed as