The Feature Extraction and Recognition of Phone Image Based on Robust Sparse Non-Negative Matrix Factorization

Enjun Yu, Qingwei Ye[*] and Yue Wu

Ningbo Institute of Technology, Zhejiang University, Zhejiang, 315100, China

Information Science and Engineering Institute of Ningbo University, Zhejiang, 315211, China

Abstract: Sparse non-negative matrix factorization algorithm can project image data effectively. It plays an important role in image matching and recognition. In order to improve the effectiveness of SNMF algorithm, which is used in feature extraction of image data with noises, we added a noise term and combined it with SNMF algorithm. Then, we proposed a new sparse optimization objective function and worked out its solution which can guarantee the sparseness of extracted feature and improve the algorithm's immunity against noise at the same time. We name this robust sparse non-negative matrix factorization (RSNMF) algorithm. It is applied for feature extraction and recognition of phone image. The concept of interface image and sub-graph of mobile phone is created. The feature extraction of phone image is used in RSNMF. And the features are put in support vector machine to achieve classification recognition. Experimental results shows that not only phone image data can be large-scale compressed through RSNMF algorithm with good robustness, but also the recognition efficiency is improved by generating sparse matrix as an intermediary target matrix to classification.

Keywords: Feature extraction, phone image, recognition, robust, sparse.

1. INTRODUCTION

Image is an important means of access to information and also a visual form of expressing information. Image processing is a hot topic in the present study, image recognition is an important field of artificial intelligence, and feature extraction is one of the most important issues in image recognition field. Normally, image features can be divided into four categories: visual features, statistical features, transform coefficients features, algebraic features [1]. As non-negative Matrix Factorization (NMF) [2, 3] algorithm is based on the algebraic features of image, and gray image can be expressed in the form of matrix, which reflects an intrinsic property and data structures of the image, so a variety of matrix factorization or various algebraic transforms can be used on gray image. NMF algorithm has obvious advantages: on the one hand, NMF algorithm puts the global data mapped to the local accumulation, which is easy to achieve and explain; on the other hand, the input data is broken down into two parts to achieve dimensionality reduction, without the existence of negative component, it can be applied to solve many practical problems. Therefore, NMF has become one of the most popular tools in signal processing, biomedical engineering, computer vision, image and other research areas which concentrate on multidimensional data processing.

NMF with sparse constraint, as one of the most popular research directions, is focused on making the results of the

*Address correspondence to this author at the Information Science and Engineering Institute of Ningbo University, Zhejiang, 315211, China; Tel: +8613805877051; E-mail: personalfree@163.com

decomposition sparseness, equivalently input data is compressed further, and which is convenient for later processing and application. Currently, there are several widely used sparse non-negative matrix factorization algorithms. Cichocki proposed the enhanced sparse non-negative matrix factorization algorithm [4] by introducing non-linear projection to make the matrix non-negative. Li proposed local NMF [5] and orthogonal restriction. Hoyer construct non-negative sparse coding [6] algorithm by combining sparse coding with NMF, meanwhile a sparse penalty term is added. Liu also proposed SNMF [7] (sparse non-negative matrix factorization) algorithm which is similar to the algorithm Hoyer presented but with different objective function. Hoyer put forward NMFSC [8] (NMF with sparseness constraints) algorithm by using a non-linear projection method to control sparsity precisely. Classic sparse NMF algorithms always pursue progress and innovation of sparse control, but ignore the ability in resisting noise. So they are not suitable iin some practical situations.

Phone image is size fixed, with no specific contours, edges or texture, and contains a large amount of data. It may get deformed or is accompanied with noise. Therefore, this paper puts forward a non-negative matrix factorization algorithm with good robustness (we call it RSNMF short for robust non-negative matrix factorization) based on Liu's SNMF algorithm. We used our algorithm on features extraction of phone image and combined it with support vector machine to do classification and recognition. The whole process includes: proposing the concept of interface image and sub-graph, extraction features of phone image by RSNMF, obtaining feature matrix of phone image, and then using the

feature matrix as input data of SVM to perform classification and recognition.

2. ROBUST SPARSE NON-NEGATIVE MATRIX FACTORIZATION

2.1. Non-Negative Matrix Factorization

NMF algorithm was proposed by D.D. Lee and H.S. Seung in 1999, its essence is a matrix decomposition projection technology. The basic principle of NMF is expressed as follows: for a $n \times m$ dimensions non-negative matrix, finding a $n \times r$ dimensions non-negative matrix W and a $r \times m$ dimensions non-negative matrix H, so that

$$\underset{n \times m}{V} \approx \underset{n \times r}{W} \underset{r \times m}{H} \tag{1}$$

Where r is called characteristic dimension, whose value is determined by $r \leq \dfrac{mn}{m+n}$.

Each column of the objective matrix V represents n dimensions vector of a sample picture. In $W = [w_1 w_2 ... w_r]$, W is called the base matrix, its physical meaning can be considered as a basic set about the linear representation of v_i. $H = [h_1, h_2, ..., h_r]$, H is called the coefficient matrix, h_i represents the i-th sample's projection coefficients on W. Generally speaking, the decomposition of V cannot strictly satisfy formula (1) but can be carried out through approximate decomposition with infinitesimal error. Therefore, we can define an appropriate objective function and by minimizing the objective function we may find the closest W and H to satisfy formula (1), which are all proposed by Lee and Sueng in reference [3]. One takes the square of Euclidean distance between V and WH as the objective function, which has the advantage of simple and intuitive. Another objective function is based on the K-L divergence, which has the advantage of easy evaluation of the optimization result.

2.2. Sparse Non-Negative Matrix Factorization (SNMF)

NMF with sparse constraint is based on NMF with constraints of making the result matrix W or H sparse, or making both W and H sparse. The objective function of SNMF algorithm that Liu *et al* proposed is based on K-L divergence.

$$O(V \| WH) = \sum_{i,j} (X_{ij} \log \frac{X_{ij}}{(WH)_{ij}} - X_{ij} + (WH)_{ij}) + \lambda \sum_{i,j} H_{ij}$$

$$s.t. \quad W, H \geq 0, \sum_i W_{ir} = 1 \tag{2}$$

In formula (2), the item which contains λ is the penalty term of sparsity. $\lambda \geq 0$, according to practical experience, the value of λ would be from 0.1 to 1. Liu *et al.* also sets the iterative rules of W, H:

$$H \leftarrow \frac{H .* (W^T * (V ./ (WH)))}{1 + \lambda} \tag{3}$$

$$W \leftarrow W .* ((V ./ (W * H)) * H^T) ./ (1_{m \times 1} * \sum H^T) \tag{4}$$

$$W_{ir} \leftarrow \frac{W_{ir}}{\sum_i W_{ir}} \tag{5}$$

Where .* and ./ represents the dot product and dot division respectively, namely the corresponding elements multiply or divide each other; * represents matrix multiplication.

Bin Shen [9] proposed Robust NMF algorithm based on the influence of noise, He added a noise term in the objective function but without sparse constraint. The main idea of Bin Shen's algorithm is: in an objective function, V is replaced by V-E, add a function item about E, where E refers to noise. By combining SNMF algorithm and Bin Shen's algorithm, we proposed a new algorithm called RSNMF. The new objective function we defined in this paper is as follows.

$$O(V \| WH) = \sum_{i,j} ((X - E)_{ij} \log \frac{(X - E)_{ij}}{(WH)_{ij}} - (X - E)_{ij} + (WH)_{ij})$$

$$+ \lambda \sum_{i,j} H_{ij} + \alpha \sum_{i,j} E_{ij}$$

$$s.t. \quad W, H \geq 0 \ , \sum_i W_{ir} = 1 \tag{6}$$

In formula (6), α is the tradeoff factor between sparsity of E and image reconstruction error; it controls the weight of E and $V\text{-}E \geq 0$ ensures that the raw data is non-negative. We use the iterative rules of SNMF algorithm to minimize the objective function $O(V \| WH)$ about W, H, and get derivative of W, H respectively. In this case E can be regarded as a constant term, as shown in formula (7). Therefore, the objective function in this paper can be obtained by replacing V as V-E in iterative formula of SNMF algorithm.

$$O(V \| WH) = \sum_{i,j} ((X - E)_{ij} \log \frac{(X - E)_{ij}}{(WH)_{ij}} - (X - E)_{ij} + (WH)_{ij}) \tag{7}$$

$$+ \lambda \sum_{i,j} H_{ij} + cons \tan t$$

As the objective function with respect to W, H, E is not convex, it is difficult to find a global optimum. Then a local optimization iteration method is used to iteratively update W, H and E. According to the reference [9] and the iterative formula of W, H in SNMF, the specific process of RSNMF is as follows.

Step 1: Set all the elements of E to 0 and initialize W, H to random non-negative matrices;

Step 2: Fix E, optimize with respect to W, H, Iteratively do the following:

$$H \leftarrow \frac{H .* (W^T * ((V - E) ./ (WH)))}{1 + \lambda} \tag{8}$$

$$W \leftarrow W. * (((V - E)./(W * H)) * H^T)$$
$$./ (1_{m \times 1} * \sum H^T) \tag{9}$$

$$W_{ir} \leftarrow \frac{W_{ir}}{\sum_i W_{ir}} \tag{10}$$

Step 3: Fix W, H, optimize with respect to E

$$E_{ij} = \begin{cases} (V - WH)_{ij}, & |V - WH|_{ij} > \sqrt{\alpha} \\ 0, & otherwise \end{cases} \tag{11}$$

Step 4: Check convergence of O, if it converges, return to W, H and E, otherwise go back to step 2.

Studying SNMF algorithm is our starting point, in which we added noise reduction processing. For a smooth noise we use the rules given in formula (11) to update E, so the algorithm can handle the situation while images are affected by some points with abnormal value. Robustness is enhanced without limits of different forms of noise, therefore the availability of latter processed images is increased. Moreover, the sparse constraint is kept and only the projection coefficient H is sparse. This algorithm can adapt to larger image distortion and noise. When facing large amount of data, we can convert these high dimensional sample to coefficient matrix H, as a result, the storage space can be greatly reduced.

2.3. Phone Image Feature Extraction Algorithm Flow

Based on the above theories, we propose the feature extraction algorithm based on sparse non-negative matrix factorization for the single phone interface image. The algorithm process is as follows (Feature Extraction By RSNMF, simply called FE-RSNMF):

FE-RSNMF Algorithm

Input: set of phone template images $T = \{T_1, T_2, ...T_M\}$, M indicates the total number of template image; single phone interface image V.

Output: Feature matrix s with size $1 \times M$, its each element represents whether this single interface image V contains template image T_i, while 1 represent yes, 0 represent no.

1. Decomposing each T_i by RSNMF, namely $T_i = w_i h_i$;

2. Calculating the projection of V on w_i based on the size of T_i, Projection formula: $h_i' = (w_i^T w_i)^{-1} w_i^T T_i$;

3. Calculating the Euclidean distance $E = 1/2 \|h_i' - h_i\|_2^2$ between h_i' and h_i, finding the minimum value of E after full scanning and
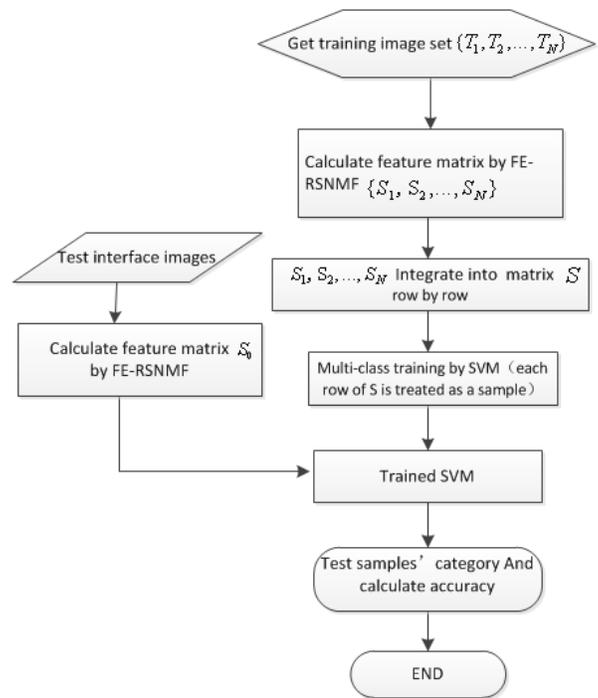

Fig. (1). Recognition flow.

projection of phone interface image. If $E_{min} \leq \delta$ (δ is set to a dimensionless value) $s(i) = 1$, otherwise $s(i) = 0$, $s(i)$ indicate the i-th element of s.

4. $i \leftarrow i + 1$, until $i = M$.

2.4. Phone Image Recognition Algorithm Flow After Feature Extraction

Make one part of the samples of phone interface image as training object and another part as testing object. In training process, first of all, calculate the feature matrix s_j (j=1,...,N,N represents the whole number of training images) by using the FE-RSNMF algorithm, and then integrate s_j row by row to matrix S and perform SVM [10, 11] multi-class training [12, 13] with input S. For the test object, go through the same feature matrix calculation process, and then feed it into the trained SVM to test. Finally, calculate the error rate. Specific process is shown in Fig. (**1**).

3. APPLICATION OF RSNMF ALGORITHM

3.1. A Short Introduction About Phone Image

Image recognition is an important process of phone automated testing. This process simulates the eyes of test engineers to identify text and image information through grabbing images on the LCD screen, to estimate the test results. It mainly relates to image acquisition and comparative analysis. Improving the matching accuracy rate is the emphasis direction to improve the efficiency of automated testing, as
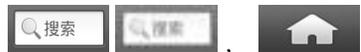
Fig. (2). Feature of interface images.

the phone image database is always large. The characteristics of phone interface image: the same function used in the different interfaces will share some similar characteristics, as shown in Fig. (**2**) is the two phone interfaces of all the functional categories of anti-virus software interfaces. Three kinds of framed sub-graphs below the pictures can be used as a database of feature map for anti-virus software interfaces. The explanation of red annotation is the starting point of mobile phone image processing in this paper.

The collected mobile phone interface images can be defined and classified according to the content on interface sub-graphs. For example: phone call interfaces usually have the following fixed sub-graphs: [icon], [icon], [icon], [icon], [icon]. Internet interfaces often have the following fixed sub-graphs [icon], [icon], [icon], [icon], [icon], and so on. Phone call interfaces often contain different interfaces such as answering the phone, phone calls, phone call records *etc*. Although the function of each interface is not the same, but they belong to the same category, which includes the same fixed sub-graphs and the position of these sub-graphs in the interface is fixed, such as the power logo image is generally in the upper right corner. Mobile phone interface definition and classification are done on the basis of these sub-graphs. Extract features of

sub-graphs by RSNMF, then compute feature matrix of phone interface image. Based on the above two steps, the interface images can be classified and identified by SVM at last.

3.2. Generation of Training and Testing Samples

Phone interface images are obtained by the camera, and converted to gray-scale images, for performing corresponding preprocessing. Then every interface image's size is fixed as 480 * 800. We select 6 classes of phone interface images, and each class has 25 phone interface images, making a total of 150 interface images, as shown in Fig. (**3**). The 6 classes represent 6 functions, namely main interface, messaging interface, text editing interface, internet interface, antivirus interface, and camera interface. Select 120 interface images as the training set, and the remaining 30 as testing set. We obtain 15 sub-graphs templates from these standard sample images whose positions are fixed in the interface images, and these sub-graph templates are the potential basis of feature matrix calculation for interface image, as shown in Fig. (**3**).

3.3. The Impact of Noise Factor α

Study in this area of phone image feature extraction is quite a little. Innovation of this paper is to start from the non-negative part of image data, then make the existed SNMF algorithm better by adding noise processing term.

Fig. (3). Generated 6 categories of representative interface images.

Fig. (4). Set of sub-graph template images.

Experiments: Firstly, calculate feature matrix S of the 120 interface images for training by using FE-RSNMF algorithm respectively and the dimension of S is integrated to 120 * 15. Then take S as the input of multi-classification SVM (one-to-many) and select Gaussian function as the kernel function, then perform the test. All the test processes are completed under the same conditions of feature matrix calculation and parameters of SVM are consistent.

In the objection function, α is treated in weights of noise term, so α is called noise factor. According to the analysis of Zhang Lijun [14], if α is large enough, the error

when doing large image reconstruction will increase as well, RSNMF will degenerate to SNMF. In fact, α makes the noise term E sparse in some way, it is a tradeoff between reconstruction and sparsity, so we discuss the experimental results while the value of α ranging from 0 to 1 seen in Fig. (**4**). We added different intensity of noise to test the impact of α with a different feature dimension r. r is the intermediate parameter of NMF algorithm, *e.g.* in this paper $r \leq 14$. As H is sparse already, if r is too small, lots of image's information will be lost. Conversely, if r is too large, the recognition rate would be reduced. So we set r=8,10,12 for analysis. While r is fixed, we added Gaussian white noises

(mean value is 0, variance respectively as 0.002,0.004,0.006) to the interface image to do the experiment, as all the images are obtained from the interface image.

CONCLUSION

This paper proposes a feature extraction algorithm of phone image based on sparse non-negative matrix factorization and noise reduction processing is added on the basis of the original sparse constraint, and hence, the algorithm can handle the images with abnormal point value. Therefore, the robustness is enhanced without limits of different forms of noise. To verify the robustness of the algorithm, the improved algorithm is used in feature extraction of phone image, combining with support vector machine for classification and recognition, finally taking a comparison with the classical sparse non-negative matrix factorization algorithm. Experimental result shows that the feature extraction method can better retain the basic information of phone image noise, has a stronger ability in resisting the noise, and achieves high recognition accuracy as well. However, if the image data dimension is too large, the speed of algorithm needs to be further enhanced.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   X. Wu, D. Zhang, and K. Wang, "Fisherpalms based palmprint recognition," *Pattern Recognition Letters*, vol. 24, no. 15, pp. 2829-2838, 2003.

[2]   D. Lee, and H.S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788-791, 1999.

[3]   D. Lee, and H.S. Seung, *"Algorithms for Non-Negative Matrix Factorization,"* *MIT Press*, 2001, pp. 556-562.

[4]   A. Cichocki, S. Amari, R. Zdunek, R. Kompass, G. Hori, and Z. He, "Extended SMART algorithms for non-negative matrix factorization," *Lecture Notes in Artificial Intelligence*, vol. 4029, pp. 548-562, 2006.

[5]   S.Z. Li, X.W. Hou, H.J. Zhang, and Q.S. Cheng "Learning spatially localized, parts-based representation," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2001*, vol. 2001, no. 1, pp. 207-212, 2001.

[6]   P.O. Hoyer, "Non-negative sparse coding," *Proceedings of the 2002 12th IEEE Workshop on Neural Networks for Signal Processing*, vol. 2002, pp. 557-565, 2002.

[7]   W. Liu, N. Zheng, and X. Lu, "Non-negative matrix factorization for visual coding," *Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, vol. 2003, no. 3, pp. 293-296, 2003.

[8]   P.O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *Journal of Machine Learning Research*, vol. 2004, no. 5, pp. 1457-1469.

[9]   X. Chen, L. Gu, S.Z. Li, and H.J. Zhang "Learning representative local features for face detection," In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2001, no. 1, pp. 1126-1131, 2001.

[10] B. Shen, Z. Datbayev, and O. Makhambetov, "Direct robust Non-Negative Matrix Factorization and its application on image processing," In: *2012 6th International Conference on IEEE Application of Information and Communication Technologies (AICT)*, 2012, pp. 1-5.

[11] L. Cai, and Y.S. Zhu, "Time-frequency spectra recognition based on sparse non-negative matrix factorization and support vector machine," *Acta Automatica Sinica*, vol. 35, no. 10, pp. 1272-1277, 2009.

[12] V.N. VAPNIK, *"The Nature of Datatistical Learning Theory,"* New York: Springer-Verlag, 1995.

[13] C.W. Hsu, and C.J. Lin, "A comparison of methods for multi – class support vector machines," *IEEE Transactions on Neural networks*, vol. 6, no. 13, pp. 415-425, 2002.

[14] L. Zhang, Z. Chen, and M. Zheng, "Robust non-negative matrix factorization," *Frontiers of Electrical and Electronic Engineering in China*, vol. 6, no. 2, pp. 192-200, 2011.