Article

Research of SWNMF with New Iteration Rules for Facial Feature Extraction and Recognition

Jing Zhou

School of Mathematics and Computer Science, Jiangnan University, Wuhan 430056, China; zhj131@jhun.edu.cn; Tel.: +86-180-4053-5659

Received: 3 February 2019; Accepted: 6 March 2019; Published: 8 March 2019

Abstract: Weighted nonnegative matrix factorization (WNMF) is a technology for feature extraction, which can extract the feature of face dataset, and then the feature can be recognized by the classifier. To improve the performance of WNMF for feature extraction, a new iteration rule is proposed in this paper. Meanwhile, the base matrix U is sparse based on the threshold, and the new method is named sparse weighted nonnegative matrix factorization (SWNMF). The new iteration rules are based on the smaller iteration steps, thus, the search is more precise, therefore, the recognition rate can be improved. In addition, the sparse method based on the threshold is adopted to update the base matrix U , which can make the extracted feature more sparse and concentrate, and then easier to recognize. The SWNMF method is applied on the ORL and JAFEE datasets, and from the experiment results we can find that the recognition rates are improved extensively based on the new iteration rules proposed in this paper. The recognition rate of new SWNMF method reached 98% for ORL face database and 100% for JAFEE face database, respectively, which are higher than the PCA method, the sparse nonnegative matrix factorization (SNMF) method, the convex non-negative matrix factorization (CNMF) method and multi-layer NMF method.

Keywords: face recognition; new additive iteration rule; threshold sparse; base matrix; coefficient matrix; weighted nonnegative matrix factorization

1. Introduction

The traditional facial feature extraction [1] method is based on the geometric structure, and a standard normalized feature vector is used to describe the structural data of facial organs [2], then the Euclidean distance between organs in the image is calculated to compare which two images have the highest consistency. Facial feature extraction based on geometric features [3] is simple and direct. However, at the same time, this method is too simple, which makes the accuracy of the algorithm unsatisfactory and the reliability unstable. Thus, it has to be combined with the semi-supervised machine learning method [4]. Then the facial feature extraction methods based on local features are proposed, such as the Gabor wavelet [5], HOG (Histogram of Oriented Gradient) [6], SURF (Speeded Up Robust Features), and SIFT (Scale Invariant Feature Transform) [7]. Gabor wavelet convolutes each pixel of sample image in different scales and directions, and then the extracted multi-resolution features can well express the characteristics of face image. However, the dimension of the characteristics extracted by Gabor wavelet is also too high, which leads to the long operation time and ineffectiveness of the algorithm. In order to solve the problem, some local texture information operators with low dimensions are proposed to extract the facial features, such as LBP (Local Binary Pattern) [8] and LTP (Local Ternary Pattern) [9]. The method of LBP can depict different features of local image texture by binarizing the gray value of the central point pixel and its surrounding point pixel. And the core idea of the LTP method is to replace the single threshold form of traditional LBP with a set of threshold intervals and upgrade binary texture pattern in traditional LBP to obtain a ternary texture pattern [10].

The LBP and LTP methods only extract the local features of human face, which cannot describe the global features of human face well. Moreover, HOG is proposed to extract the edge feature of the human face. For an object, the vector of its edge direction is relatively fixed, so its shape (contour) is relatively unchanged. Based on this characteristic, the HOG feature can better reflect the facial edge feature, which is rarely affected by small offset and illumination problems.

All the above methods are adopted to extract features directly from the original face dataset, but the dimension of the original face dataset is usually too high. To reduce the dimension of original dataset, the subspace method was proposed. The main idea of subspace-based facial feature extraction technology is to transform high-dimensional face image features into another low-dimensional space through linear or non-linear spatial transformation. The purpose is that face sample features can be easily classified in the new subspace [11]. PCA (Principal Component Analysis) is the most popular subspace method [12], and the PCA-based “eigenfaces” extraction technology is widely used in face feature extraction [13,14]. Additionally, for any face image, it can be represented by the combination of feature face images, and the feature vector is the corresponding relationship coefficient of feature face combination. Although PCA technology has been used in facial feature extraction and achieved good results, there are still some shortcomings. Firstly, the high-order statistical information extracted by PCA is still insufficient, so this method can't extract enough features for recognition, that is to say, it can't achieve the desired recognition effect. Secondly, there are negative elements in the feature matrix obtained by the PCA method, which weaken the features and then the recognition rate is reduced. Moreover, the negative value also makes the results less explanatory in the non-negative datasets of image, voice, videos and so on.

For these reasons, an unsupervised non-negative matrix factorization method is proposed by Lee and Seung to solve the above problems [15,16]. NMF (Nonnegative Matrix matrix Factorization) is a typical subspace analysis method. The high-dimensional raw data is projected on the base matrix to get the feature vectors with lower dimensions, and then the original image is weighted by the base space image. Additionally, the obtained feature vectors are used to recognition. Compared with the unconstrained matrix decomposition method, NMF has more explanatory and clear physical meaning [17]. Therefore, it has been widely used in image, voice, video and other non-negative datasets to extract the features [18,19]. In recent years, many improved NMF algorithms have been proposed, such as SNMF (sparse nonnegative matrix factorization) method [20], the CNMF (convex nonnegative matrix factorization) method [21], and the multi-layer NMF method [22]. The SNMF method considers the redundant information hidden in the complex data by adding the sparse constraints in the iteration rule. CNMF method is derived from Semi-NMF [23], and it replaces the base matrix with the non-negative convex combination of the original matrix. And the general idea of the multi-layer NMF method is to stack one-layer NMF into multiple layers to learn the hierarchical relationships among features or hierarchical projections.

However, the above NMF methods are all iterated by the multiplicative iteration rule, and the improvements for the iteration step size are not considered. Actually, it is difficult to extract facial features accurately when the iterative step size is not appropriate, which will result in the low recognition rate of facial datasets. Meanwhile, the respective weights of important information and less important information of the face data are not considered, which will result in the hierarchy of the extracted feature not being clear. Thus, by adding a weight matrix in the iteration rules to improve the hierarchical expression of local features relative to the overall features and improving the iteration step size, a new sparse weighted non-negative matrix factorization (SWNMf) method based on the new iterative step size is proposed in this paper, whose recognition rate is higher than the NMF method, the SNMF [24] method, CNMF method [21], and multi-layer NMF method [22]. Meanwhile, the deep learning methods are also widely used in face feature extraction. CNN (Convolutional Neural Network) is one of the most state-of-the-art in-depth learning methods and the effectiveness of the SWNMf method is also compared with the CNN method in this paper.

2. Image Preprocessing

To ensure that the extracted features are robust to facial changes, face images need to be pretreated.

2.1. Grayscale Normalization

Uneven illumination of the original image can be compensated by the grayscale normalization, thereby the adverse effects of changes in illumination can be overcome for recognition. Given the mean and variance of grayscale, a linear method can be adopted to transform the mean and variance of each grayscale to the given values, and then all face images are in conformity to the same or similar grayscale distribution.

The grayscale distribution matrix of an image of M by N is $I(x, y)$, where $I(x, y)$ is the gray value of the pixel at point (x, y) , where $0 < x < M - 1$ and $0 < y < N - 1$. Thus, the mean μ and variance σ of the image are as shown in Equations (1) and (2), respectively:

$$\mu = \frac{1}{M \times N} \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} I(x, y) \quad (1)$$

$$\sigma^2 = \frac{1}{M \times N} \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} (I(x, y) - \mu)^2 \quad (2)$$

The grayscale value of each image pixel can then be converted using Equation (3), the purpose of which is to give the expected values μ_0 and σ_0 to the mean and variance of the image.

$$\hat{I}(x, y) = \frac{\sigma_0}{\sigma} (I(x, y) - \mu) + \mu_0 \quad (3)$$

where $I(x, y)$ denotes the image matrix before the grayscale normalization, and $\hat{I}(x, y)$ denotes the image matrix after normalization. In addition, the effect of changes in illumination on the recognition accuracy can, thus, be overcome through a normalization of the gray scale.

2.2. Extracting Low-Frequency Information by Wavelet Transform

Wavelet transform is a local transformation in the time and frequency domains. The local characteristics can be extracted from an image and represented effectively using a wavelet transform.

Here, $f(t)$ is set as a square integrable function, and the continuous wavelet transform is defined through Equation (4):

$$w_f(a, b) = \int_{-\infty}^{+\infty} f(t) \varphi_{a,b}(t) dt = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{a}} \varphi\left(\frac{t-b}{a}\right) dt \quad (4)$$

where $a > 0$ is a scaling factor and b is a positional parameter, and $\varphi_{a,b}(t)$ is the mother wavelet:

$$\text{DWT}(j, k_1, k_2) = 2^j \sum_{l_2} \sum_{l_1} f(l_1, l_2) \varphi(2^j l_1 - k_1, 2^j l_2 - k_2) \quad (5)$$

The corresponding two-dimensional discrete wavelet transform is defined through Equation (5), which is used for the original input image, and four sub-images (LL, LH, HL, and HH) can then be produced. In addition, LL is a low-frequency component for the image, which contains most of the information of the original image. When the wavelet transform is performed on LL, a second order wavelet transform is obtained.

A face image in the JAFEE face dataset is used for a wavelet transform, the results of which are shown in Figure 1. From the Figure 1, we can find the extracted low-frequency information can be used as an approximation of the original image, while the noise and other high frequency signals can be greatly suppressed. Meanwhile, a face image constructed based on the low-frequency information

after second order wavelet transform is relatively obscured. Therefore, the low-frequency information obtained from first order wavelet transform is adopted in this paper.

Figure 1. Wavelet transforms of one-layer and two-layer, respectively. The transform results should be listed as: (a) Description of one-layer wavelet transform; and (b) description of the two-layer wavelet transform.

3. SWNMF Method with a New Iterative Rule

Lee and Seung proved the convergence of nonnegative matrix factorization (NMF) [16]. The basic idea of NMF is that for any given nonnegative matrix X with the size of $m \times n$, two nonnegative matrices U and V can be found to satisfy $X = UV$, where U is the base matrix with size of $n \times r$, which represents the features of the image. Additionally, V is the coefficient matrix with size of $r \times m$, which represents the weight coefficient of the image feature. U and V are smaller than the original matrix, so U and V can be regarded as components in V .

WNMF is developed on the basis of NMF. The main feature of WNMF is to add a weight matrix P when updating U and V , which makes the weight of important information larger and the weight of relatively less important information smaller, so as to improve the expression of local features and then the image features can be better expressed.

3.1. New Iteration Rule

The objective function of weighted nonnegative matrix factorization is defined as the weighted Euclidean distance, which is described as Equation (6):

$$J(V) = \frac{1}{2} \|X - UV\|_P^2 = \frac{1}{2} \sum_i (p_i [x_i - (Uv)_i]^2) = \frac{1}{2} (x - Uv)^T D_p (x - Uv) \quad (6)$$

where p_i is the element value of weighted matrix, and $D_p = \text{diag}(p)$. Let v^t be the current approximation of the minimizer of $J(v)$, then $J(v)$ can be rewrite as the following quadratic form:

$$J(v) = J(v^t) + (v - v^t)^T \nabla_v J(v^t) + \frac{1}{2} (v - v^t)^T (U^T D_p U) (v - v^t) \quad (7)$$

where $\nabla_v J(v^t)$ is explicitly given by:

$$\nabla_v J(v^t) = -U^T D_p (x - Uv^t) \quad (8)$$

Then one auxiliary function $Z(v, v^t)$ for $J(v)$ is constructed as Equation (9), which satisfies the conditions as Equation (11). Meanwhile, v is updated by Equation (12) [25]:

$$Z(v, v^t) = J(v^t) + (v - v^t)^T \nabla_v J(v^t) + \frac{1}{2} (v - v^t)^T L(v^t) (v - v^t) \quad (9)$$

$$\text{where } L(v^t) = \text{diag} \left(\frac{U^T(D_p(UV^t))}{V^t} \right) \frac{(U^T(D_p X) + U^T(D_p(UV^t)))_{rj}}{(U^T(D_p X))_{rj}} \quad (10)$$

$$Z(v, v^t) \geq J(v), Z(v, v) = J(v) \quad (11)$$

$$v^{t+1} = \arg \min_v Z(v, v^t) \quad (12)$$

When $Z(v, v^t)$ is constructed to satisfy the conditions as Equation (11) and J is updated by Equation (12), the objective function $J(v)$ can satisfy the Equation (13), which illustrate that $J(v)$ is a non-incremental and convergent function:

$$J(v^{t+1}) \leq Z(v^{t+1}, v^t) \leq Z(v^t, v^t) = J(v^t) \quad (13)$$

Blondel-Ho have defined the $G(v, v^t)$ in literature [25], and have proved that $G(v, v^t) \geq J(v)$. in literature [25], so we only need prove that $Z(v, v^t) \geq G(v, v^t)$, then $Z(v, v^t) \geq J(v)$ can be proved.

In order to prove $Z(v, v^t) \geq G(v, v^t)$, based on Equation (9) and the $G(v, v^t)$ described in literature [25] as Equation (14), we only need prove Equation (15):

$$G(v, v^t) = J(v^t) + (v - v^t)^T \nabla_v J(v^t) + \frac{1}{2} (v - v^t)^T D(v^t) (v - v^t) \quad (14)$$

$$L(v^t) \geq D(v^t) = \text{diag} \left(\frac{U^T(D_p(UV^t))}{V^t} \right) \quad (15)$$

By the description of $L(v^t)$ as Equation (10) and $D(v^t)$ as Equation (15), we can derive the following:

$$L(v^t) = \text{diag} \left(\frac{U^T(D_p(UV^t))}{V^t} \right) \frac{(U^T(D_p X) + U^T(D_p(UV^t)))_{rj}}{(U^T(D_p X))_{rj}} = \text{diag} \left(\frac{U^T(D_p(UV^t))}{V^t} \right) (1 + (U^T(D_p(UV^t)))_{rj} / (U^T(D_p X))_{rj}) \geq \text{diag} \left(\frac{U^T(D_p(UV^t))}{V^t} \right) = D(v^t) \quad (16)$$

Based on Equation (16), we can get $Z(v, v^t) \geq G(v, v^t) \geq J(v)$, that is, the $Z(v, v^t)$ constructed as Equation (9) is proved to satisfy the conditions as Equation (11), meanwhile, v is updated with Equation (12), then Equation (13) can be obtained, which illustrates that the objective function J is non-increasing and convergent.

Based on the updating rule of v described as Equation (12), the new iteration rule of v can be derived as Equation (17):

$$\begin{aligned} v^{t+1} &= \arg \min_h Z(v, v^t) = v^t - \nabla J(v^t) / L(v^t) \\ &= v^t + \text{diag} \left(\frac{V^t}{U^T(D_p(UV^t))} \right) \frac{(U^T(D_p X))_{rj}}{(U^T(D_p X) + U^T(D_p(UV^t)))_{rj}} U^T(D_p(x - Uv^t)) \\ &= v^t + v^t \frac{(U^T(D_p X))_{rj}}{(U^T(D_p(UV^t)))_{rj}} \frac{(U^T(D_p X) - U^T(D_p(UV^t)))_{rj}}{(U^T(D_p X) + U^T(D_p(UV^t)))_{rj}} \end{aligned} \quad (17)$$

From Equation (17), we can the new iteration step for v is set as α , which is described as Equation (18). Meanwhile, the traditional iteration step for v is defined as Equation (19):

$$\alpha = \text{diag} \left(\frac{V^t}{U^T(D_p(UV^t))} \right) \frac{(U^T(D_p X))_{rj}}{(U^T(D_p X) + U^T(D_p(UV^t)))_{rj}} \quad (18)$$

Similarly, the new iteration rule of u described as Equation (19) can be get, and the new iteration step for u is set as β , which is described as Equation (20):

$$\begin{aligned} u^{t+1} &= u^t + \text{diag} \left(\frac{U^t}{(D_p(U^t V)) V^T} \right) \frac{((D_p X) V^T)_{ir}}{((D_p X) V^T + (D_p(U^t V)) V^T)_{ir}} (D_p(x - u^t V)) V^T \\ &= u^t + u^t \frac{((D_p X) V^T)_{ir}}{((D_p(U^t V)) V^T)_{ir}} \frac{((D_p X) V^T)_{ir} - ((D_p(U^t V)) V^T)_{ir}}{((D_p X) V^T + (D_p(U^t V)) V^T)_{ir}} \end{aligned} \quad (19)$$

$$\beta = \text{diag}\left(\frac{U^T}{(D_p(U^T V))V^T}\right) \frac{((D_p X)V^T)_{ir}}{((D_p X)V^T + (D_p(U^T V))V^T)_{ir}} \quad (20)$$

The idea of proving the convergence of objective function of WNMf method is adopted to prove the convergence of the SWNMf method. Firstly, a quadratic auxiliary function is found as shown in Equation (9), which satisfies the conditions described as Equation (11). Then the minimum value of the auxiliary function is get and plugged into the Equation (12), and the value of the next base matrix can be obtained. Then the minimum value of the auxiliary function at this value is taken as the value of the next base matrix. If we continue to use Equation (12) based on the found auxiliary function, we can satisfy Equation (13). Thus, we can prove that J is an non-increasing function and it is convergent, and the updating rule can be derived as Equation (17). From the Equation (17), we can get the parameter of iteration step, which is described as Equation (18). Similarly, the other iteration step parameter of Equation (20) can also be decided by the iteration rule described as Equation (19).

Based on the above deriving process, the objective function J has been proved convergent when the matrix U and V are updated by the new iterative rules described as Equations (17) and (19). Thus, the SWNMf algorithm is convergent and terminated when the target function (6) reaches the minimum value, or the number of iterations reaches the upper limit. The initial value of U and V of the iterative algorithm can be randomly generated based on the dimensions of U and V . The optimal U and V can be obtained when the iteration process is finished. Therefore, the new SWNMf method can be used to extract the facial features.

3.2. Sample Weighting and Sparse Constraints

In order to improve the recognition rate, considering the difference of importance for each sample, the samples can be weighted according to the sample quality, then the above weighted non-negative matrix factorization method is adopted. For example, if the eyes of some samples are closed, that is, the key characteristic eyes are not obvious, the weights of the samples are set low. Additionally, the weights of some samples, whose expressions are too exaggerated, leading to the deformation of the key features such as eye and mouth, which is not conducive to recognition and classification, are also set low. On the contrary, the weights of the clear face images in face dataset are set higher, and the weights of non-exaggerated face images will be set higher.

In addition, in the new iterative rules of SWNMf method, no sparse constraints are added to the base matrix U and, thus, redundant information remains in the dataset. Therefore, the threshold judgment is adopted as the sparse constraint for the base matrix U during the additive iterations of U and V .

A threshold is selected when using Equations (17) and (19) for an iteration. In addition, when U is produced during each iteration, we will reset all data in the columns of base matrix U based on the threshold. If the data are greater than the threshold, the values are reset to 1, otherwise, the data are reset to 0. Then the base matrix is transformed to 0 and 1 matrix, which describes the important features as 1, whereas describes the weight of the less important features as 0, and the facial features can be expressed more concentrated and clear. In this way, the region representing the facial features can be intensely extracted, and the interference of other unrelated areas is reduced. Meanwhile, the number of computations is reduced.

Figure 2a–d shows visual images of the base matrices obtained from the training samples based on the PCA, threshold SNMF, CNMF, Multi-layer NMF and the new additive iterative SWNMf method, respectively. Since the initial values of U and V are random values between 0 and 1, the threshold is set as a value between 0 and 1. Usually, the threshold is smaller, the more redundant information can be eliminated, and the data are more centralized. Here, the threshold is set to 0.01.

Figure 2. Visual images of the optimal base matrices of the various NMF methods. (a) Description of the base matrix image for PCA; (b) description of the base matrix image for threshold SNMF; (c) description of the base matrix image for CNMF; (d) description of the base matrix image for Multilayer-NMF; and (e) description of the base matrix image for SWNMF with new iterative rules.

An image of the optimal base matrix obtained by the new additive iterative SWNMF method is shown in Figure 2e, which can accurately reflect the facial features and expression, and make the facial feature data more concentrated and clear. The images of the base matrices obtained by the other four methods are shown in Figure 2a–d, and their feature information is obviously insufficiently clear and concentrated. Therefore, compared with the other methods, the SWNMF method proposed in this paper can make the data of base matrix U sparser, which can result in the feature information being more accurate and concentrated thus, the recognition based on the corresponding coefficient matrix V for weight U will be more accurate.

As shown in Table 1, the recognition rate is further improved based on the new SWNMF, and is significantly higher than the threshold SNMF method described elsewhere [24].

Table 1. Comparison of recognition rate of threshold SNMF and SWNMF with different rank r .

r	20	35	55	75	175	200
SNMF	81%	85%	90%	88%	84%	80%
SWNMF	83.5%	90%	92%	95%	98%	98%

4. Classification Based on a Support Vector Machine

A support vector machine (SVM) is essentially a binary classifier. In this paper, however, it is a typical multi-classification problem used to classify faces of multiple categories. Although an SVM can handle multiple classification problems using both one-to-one and one-to-many strategies, the results of a one-to-one strategy are more accurate, and thus this particular SVM strategy is adopted. The S classes of samples can be divided into two categories, and $S(S - 1)/2$ classifiers are then constructed.

Forty faces were used in the experiment, and the one-to-one strategy was applied to construct 780 classifiers. The coefficient matrix V of the face sample set and the corresponding category label set were set as the input training set of the SVM classifier. The parameters of each correct binary

classification of training set were stored, and a multiple classification parameter file was obtained. This file can be called to obtain the parameter information of the multi-classifier to classify the test set.

A category voting matrix is set for the dual-classification determination of each sample. For example, if the sample is determined to belong to class k , then 1 is added to the k_{th} column value of the category voting matrix, i.e., one ticket is voted for the k column. In contrast, if the sample is determined to be of class q by the classifier, then 1 is added to the q_{th} column of the category voting matrix, i.e., one ticket is voted for the q column. The number of columns with the most votes in the statistical matrix is the category number of the sample.

Based on the following experimental results, the one-to-one SVM classifier is proved that it can correctly recognize the face with different facial expressions.

5. Experimental Results and Analysis

Since the facial features differ greatly between Europeans and Asians, the JAFEE face database of Japanese female models and ORL library of European face models provided by Cambridge University are selected to identify the SWNMF effectiveness. The SWNMF method was adopted to identify a human face in the two face databases, respectively. The JAFEE database is composed of 200 facial images for ten Japanese women, each person has twenty faces of different facial expressions. Each image has 256 gray levels, and the size of the image is 256×256 . The first ten images of the ten women were selected as training samples, and the last ten images were used as test samples. Meanwhile, there are forty people in the ORL gallery, with ten faces for each person and, thus, there are 400 faces in total. Each face has 256 grayscale levels, and the size is 112×92 . The facial expressions and facial details of each person differ, including smiling and not smiling, eyes opening and closing, wearing and not wearing glasses, and so on. The facial postures are also different, and variations in the rotation angles of the depth and plane can reach 20° , and the variation in the size of the facial image reaches 10%. The first five images were randomly selected for training, and the remaining five images were used for testing, and thus the training and test galleries have 200 images, respectively.

5.1. Comparison of SWNMF with multiple iteration NMF methods and PCA

The experiments for the face recognition rate with different values of r based on threshold SNMF [24], CNMF [21], multi-layer NMF [22], of multiplicative iterative rules, PCA [12] and the SWNMF based on the new additive iteration rules proposed in this paper are executed on the JAFEE database and ORL database, respectively. The results of comparison experiments are shown in Figures 3–6. In Figures 3–6, comparison curves for various NMF methods and PCA method with continuous changes in the value of r are given.

Figure 3. Continuous variation of recognition rate for ORL database with increasing r .

Figure 4. Continuous variation of recognition rate for JAFEE database with increasing r for five methods.

Figure 5. Error of reconstruction for the ORL database.

Figure 6. Error of reconstruction for the JAFEE database.

The recognition rate for ORL database of the SWNMF method proposed in this paper is given in Figure 3. From Figure 3, we can find that the recognition rate of the SWNMF method proposed in this paper is consistently and significantly higher than the other four methods with different values of r . In addition, the highest recognition rate of the SWNMF method is 98%, which is 77% higher than the highest recognition rate of PCA, 8% higher than the threshold SNMF method, 35.5% higher than the CNMF method and 3.5% higher than the multi-layer NMF.

As can be seen from Figure 3, the recognition rate of the ORL database based on the SWNMF method continues increasing when r increases. When $r = 175$, the highest recognition rate of 98% is obtained, and when r continues to increase, the recognition rate remains unchanged at 98%. For the threshold SNMF method, when r increases to 55, the highest recognition rate is 90%, when r continues to increase, the recognition rate decreases instead. For multi-layer NMF, the highest recognition rate is 94.5% when $r = 175$, and when r continues to increase, the recognition rate also decreases. For CNMF,

the highest recognition rate is 67% when $r = 20$, and when r continues to increase, the recognition rate also decreases. Meanwhile, for PCA, the highest recognition rate is only 21% when $r = 20$, and when r continues to increase, the recognition rate decreases. Thus, for the ORL database, the recognition rates of PCA, threshold SNMF, CNMF, and multi-layer NMF are lower than the SWNMF method.

The recognition rate for JAFEE database of the SWNMF method proposed in this paper is given in Figure 4. It can be seen from Figure 4 that the recognition rate of the SWNMF method proposed in this paper is consistently and significantly higher than the other methods with different values of r . In addition, the highest recognition rate of the SWNMF method is 100%, and it is 67% higher than the PCA method, 8% higher than the threshold SNMF method, 22% higher than the CNMF method, and 4.5% higher than the highest recognition rate of the Multi-layer NMF.

As can be seen from Figure 4, the recognition rate of the JAFEE dataset based on the SWNMF method continues increasing when r increases. When $r = 175$, the highest recognition rate of 100% is obtained, and when r continues to increase, the recognition rate remains unchanged at 100%. For PCA, the highest recognition rate is only 33% when $r = 20$, and when r continues to increase, the recognition rate decreases. For CNMF, the highest recognition rate is 78% when $r = 35$, and when r continues to increase, the recognition rate also decreases. For the threshold SNMF method, when r increases to 55, the highest recognition rate is 92%. However, when r continues to increase, the recognition rate decreases instead. And for Multi-layer NMF, the highest recognition rate is 95.5% when $r = 175$, and when r continues to increase, the recognition rate also decreases. For the JAFEE database, the recognition rates of PCA, threshold SNMF, CNMF, and multi-layer NMF, are lower than the SWNMF method, and the recognition rates also decrease instead when r increases.

The SWNMF method adopts the new iteration steps proposed in this paper, from the definition of the new iteration step of Equation (18) and the original iteration step definition of WNMF in literature [25], we can see that the new iteration step is smaller than the traditional iteration step, so the search accuracy can be improved. In addition, the threshold-sparse rule is added into the iteration rule of base matrix, which make the extracted feature data sparser and concentrated, as well as easier to recognize, thus a higher recognition rate can be achieved. However, the iteration rules of SNMF, CNMF, and multi-layer NMF method are still the traditional multiplicative iteration rules based on the traditional iteration steps, thus, the recognition rate cannot be improved. Moreover, the PCA, CNMF, and multi-layer NMF methods haven't sparse the base matrix, then there are still a lot of redundant information in the base matrix, thus, the features are not obvious to recognize. In addition, the accuracy of multi-layer NMF is related to the number of hidden layers. Too many layers will lead to too long running time.

After the base matrix is obtained during the experiment, the image can be reconstructed based on the base matrix. In addition, the reconstruction error is smaller, and the image can be reconstructed more accurately. Additionally, the errors of the reconstructed face image of the above five methods can be calculated under different values of r , which can further verify the accuracy of the extracted features for different methods. The errors of the reconstruction for ORL and JAFEE databases based on the five methods are shown as Figures 5 and 6, respectively.

As shown in Figures 5 and 6, when the image is reconstructed by the base matrix, the errors of the PCA, CNMF, the threshold SNMF and Multi-layer NMF method are larger than the SWNMF. That is, only the new additive iterative SWNMF method proposed in this paper can guarantee both a high recognition rate and high-quality image reconstruction. Therefore, considering the comprehensive factors of the recognition rate and accuracy of an image reconstruction, the method proposed in this paper is significantly better than the other methods.

Based on the new additive iterative SWNMF method, considering the recognition rate and image reconstruction, $r = 175$ is selected, and facial recognition software based on MATLAB is realized. MATLAB is a commercial mathematical software produced by MathWorks company in the united states. It is used in advanced technical computing language and interactive environment for algorithm development, data visualization, data analysis and numerical calculation. The program results are

shown in Figures 7–9. Randomly picking out the faces of different people from the two face databases, and for cases such as eyes open or closed, and whether or not glasses are worn, the software consistently results in a correct classification. For the ORL database, a person can be recognized, and their face correctly recognized with a probability of 0.98, as shown in Figures 7–9. For JAFEE database, a person can be recognized, and their face correctly recognized with a probability of 1, as shown in Figure 9. In addition, the recognition rate for the face category is also shown in the text box in Figures 7–9.

Figure 7. Correct face recognition when eyes are open or closed.

Figure 8. Correct face recognition when the man is wearing glasses or not.

Figure 9. Different facial expressions of the same person are correctly recognized. (a) Description of the correct recognition result for one kind of facial expression of ORL data set; and (b) Description of the correct recognition result for another kind of facial expression of ORL data set. (c) Description of the correct recognition result for one kind of facial expression of JAFEE data set; (d) Description of the correct recognition result for another kind of facial expression of JAFEE data set.

As shown from the results of the software simulation, a significantly high face recognition rate can be achieved using the new additive iterative SWNMF method.

5.2. Comparison of SWNMF with CNN

CNN is the popular deep learning technology. For comparing the performance of SWNMF and CNN, the deep learning toolbox of MATLAB for CNNs is applied to the ORL dataset. The facial features are extracted by the convolution and pooling layers, and are classified by the full connection layer. The filter numbers for each convolutional layer are 6, 12, and 6, respectively, and the filter size is 5×5 . The CNN is adopted for the ORL database, and based on the experiments, we find that the recognition rate of the CNN can be increased when the number of iterations increases. When the iteration number increases from 200 to 2000, the recognition rate increases from 2.5% to 84%, which is shown as Table 2. Meanwhile, the recognition time increases rapidly. Comparing the SWNMF method, we obtain the results in Table 3. From Table 3 we can find that the highest recognition rate of CNN is 84% when the iteration number arrived at 2000 is obviously too time-consuming. However, the recognition rate of SWNMF can reach 98% within 89 seconds.

Table 2. Recognition rate and recognition time for CNN.

Iteration Numbers	Recognition Rate	Recognition Time
200	2.5%	55 s
1000	76%	270 s
2000	84%	520 s

Table 3. Comparison of recognition rate and recognition time of CNN and SWNMF.

	Recognition Rate	Recognition Time
SWNMF	98%	89 s
CNN	84%	520 s

As can be seen from Table 3, when the recognition rate of deep network reaches a high value, it takes a longer time than the SWNMF method. The reason is that the deep learning requires enough learning and training processes to accurately extract features, which can improve the recognition rate, however, the SWNMF method can accurately extract features within a shorter time and it can effectively improve the recognition rate.

When the depth of network layer is further deepened, even if the recognition rate can be further improved, the learning and training time will be longer. Thus, the efficiency of the SWNMF method is still higher than the deep neural network. Since the total number of datasets used in this paper is not large, that is, there is not much information provided for deep network to learn. In this case, the use of deep neural network has no advantage. In the case of insufficient sample information, the SWNMF method has great advantages in recognition rate and recognition time compared with the deep neural network. In many practical application scenarios, it is impossible to collect many samples, and then the SWNMF method can be adopted to achieve higher efficiency than deep learning, such as higher recognition rate and shorter recognition time.

6. Conclusions

A new additive iterative rule is proposed for facial extraction and recognition based on nonnegative matrix factorization, and a sparse constraint is applied to the base matrix U by setting the threshold value during the additive iteration process. First, the face image is normalized to reduce the influence of light, and a discrete wavelet transform is adopted to extract the low-frequency information of the image to reduce the noise. Second, the new additive iteration rules of the WNMF method are derived to update U and V , and then the threshold sparse constraint is added on the base matrix U during the iterative process, which can effectively extract the facial features. Finally, the testing face datasets are decomposed in the optimized feature subspace represented by the base matrix U , and then the weight coefficient matrix V of the features can be recognized by the SVM.

With a different r , comparative experiments on the face recognition rate and the error of face image reconstruction respectively based on the PCA, threshold SNMF, CNMF, multi-layer NMF, and new additive iterative SWNMF method were conducted. As shown from the experimental results, the new additive iterative SWNMF method has the highest recognition rate, which is 77%, 8%, 35.5%, and 3.5% higher than that of PCA, threshold SNMF, CNMF, and multi-layer NMF, respectively. Moreover, the new additive iterative SWNMF method has the least error of reconstruction for face reconstruction, compared with the PCA, threshold SNMF, CNMF, and multi-layer NMF methods.

At the same time, deep learning methods are also widely used in face feature extraction. CNN is the most state-of-the-art deep learning method. Therefore, the effectiveness of the SWNMF method proposed in this paper is also compared with CNN. From the experimental results, we can find that the SWNMF method is more efficient and it can extract features accurately within the shorter time, thus, it does not require the sufficient learning, and is more suitable for scenarios with insufficient sample information.

In conclusion, compared with the traditional NMF methods, such as CNMF, SNMF, and multi-layer NMF method, the SWNMF method can achieve a higher recognition rate and better image reconstruction effect. At the same time, compared with the current state-of-the-art deep learning method, the SWNMF method is more efficient, and can achieve higher a recognition rate without a large number of samples and sufficient learning.

Author Contributions: Conceptualization, methodology, and software realization were performed by J.Z. Validation and formal analysis were finished by T.W.

Funding: This research was funded by the research Project of Provincial Teaching Reform in Hubei Province, grant number 2017301.

Acknowledgments: This paper is sponsored by the Research Project of Provincial Teaching Reform in Hubei Province (No. 2017301).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Abudarham, N.; Shkiller, L.; Yovel, G. Critical features for face recognition. *Cognition* **2019**, *182*, 73–83. [[CrossRef](#)] [[PubMed](#)]
2. Hassaballah, M.; Aly, S. Face recognition: Challenges, achievements and future directions. *IET Comput. Vis.* **2015**, *9*, 614–626. [[CrossRef](#)]
3. Ghimire, D.; Lee, J.; Li, Z.N.; Jeong, S. Recognition of facial expressions based on salient geometric features and support vector machines. *Multimed. Tools Appl.* **2017**, *76*, 7921–7946. [[CrossRef](#)]
4. Zhang, L.; Zhang, D.; Sun, M.M.; Chen, F.M. Facial beauty analysis based on geometric feature: Toward attractiveness assessment application. *Expert Syst. Appl.* **2017**, *82*, 252–265. [[CrossRef](#)]
5. Dora, L.; Agrawal, S.; Panda, R.; Abraham, A. An evolutionary single Gabor kernel based filter approach to face recognition. *Eng. Appl. Artif. Intell.* **2017**, *62*, 286–301. [[CrossRef](#)]
6. Xiang, Z.; Tan, H.; Ye, W. The Excellent Properties of a Dense Grid-Based HOG Feature on Face Recognition Compared to Gabor and LBP. *IEEE Access* **2018**, *6*, 29306–29319. [[CrossRef](#)]
7. Ali, N.; Bajwa, K.B.; Sablatnig, R.; Chatzichristofis, S.A.; Iqbal, Z.; Rashid, M.; Habib, H.A. A Novel Image Retrieval Based on Visual Words Integration of SIFT and SURF. *PLoS ONE* **2016**, *11*, 1–20. [[CrossRef](#)] [[PubMed](#)]
8. Werghi, N.; Tortorici, C.; Berretti, S.; Bimbo, A.D. Boosting 3D LBP-based face recognition by fusing shape and texture descriptors on the mesh. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 964–979. [[CrossRef](#)]
9. Wang, J.; Zhang, R.; Wu, T.T.; Ok, S.; Lee, E. Face Recognition Based on Improved LTP. *ISMEMS* **2017**, *134*, 6–10.
10. Tan, X.; Triggs, B. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans. Image Process.* **2010**, *19*, 1635–1650. [[PubMed](#)]
11. Yaman, M.A.; Subasi, A.; Rattay, F. Comparison of Random Subspace and Voting Ensemble Machine Learning Methods for Face Recognition. *Symmetry* **2018**, *10*, 1–19. [[CrossRef](#)]
12. Admane, A.; Sheikh, A.; Paunikar, S.; Jawade, S.; Wadbude, S.; Sawarkar, M.J. A Review on Different Face Recognition Techniques. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **2019**, *5*, 207–213.
13. Tatepamulwar, C.B.; Pawar, V.P.; Khamitkar, S.D.; Fadewar, H.S. Technique of Face Recognition Based on PCA with Eigen-Face Approach. *Comput. Commun. Signal Process.* **2019**, *810*, 907–918.
14. Senthilkumar, R.; Gnanamurthy, R.K. A comparative study of 2D PCA face recognition method with other statistically based face recognition methods. *J. Inst. Eng. (India) Ser. B* **2016**, *97*, 425–430. [[CrossRef](#)]
15. Lee, D.; Seung, H. Learning the parts of objects by non-negative matrix factorization. *Nature* **1999**, *401*, 788–791. [[CrossRef](#)] [[PubMed](#)]
16. Lee, D.; Seung, H. Algorithms for non-negative matrix factorization. *Adv. Neural Inf. Process. Syst.* **2001**, *13*, 556–562.
17. Arefin, M.M.N. Face Reconstruction Using Non-Negative Matrix Factorization and ℓ_1 Constrained Optimization. In Proceedings of the 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, 10–12 January 2019; Volume 5, pp. 1–9.

18. Zhu, J.X.; Hu, H.; He, X. Moving Target Detection Method Based on Non-negative Matrix Factorization of Sliding Window. *Comput. Technol. Dev.* **2017**, *27*, 20–24.
19. Virtanen, T.; Raj, B.; Gemmeke, J. Active-set Newton algorithm for non-negative sparse coding of audio. *IEEE Int. Conf. Acoust. Speech Signal Process.* **2017**, *1*, 3092–3096.
20. Sabzalain, B.; Abolghasemi, V. Iterative Weighted Non-smooth Non-negative Matrix Factorization for Face Recognition. *Int. J. Eng.* **2018**, *31*, 1698–1707.
21. Cui, G.; Li, X.; Dong, Y. Subspace clustering guided convex nonnegative matrix factorization. *IEEE Trans. Signal Process.* **2018**, *292*, 38–48. [[CrossRef](#)]
22. Song, H.A.; Kim, B.K.; Xuan, T.L.; Lee, S.Y. Hierarchical feature extraction by multi-layer non-negative matrix factorization network for classification task. *Neurocomputing* **2015**, *165*, 63–74. [[CrossRef](#)]
23. Wang, D.; Gao, X.; Wang, X. Semi-Supervised Nonnegative Matrix Factorization via Constraint Propagation. *IEEE Trans. Cybern.* **2016**, *46*, 233–244. [[CrossRef](#)] [[PubMed](#)]
24. Lin, Q.; Li, J.; Yong, J.P.; Liao, D.A. Improved face recognition method based on NMF. *Comput. Sci.* **2012**, *39*, 243–245.
25. Blondel, V.; Ho, N.D.; Dooren, P.V. Algorithms for Weighted Non-Negative Matrix Factorization. *Submit. Publ.* **2007**, *3*, 1–13. Available online: <https://www.ime.usp.br/~jstern/miscellanea/seminario/Blondel10.pdf> (accessed on 3 February 2019).

© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).