Research on Coal and Rock Image Recognition Based on 2DPCA Algorithm

Yao Xu *
JiLin Technology College of Electronic Information, Jilin 132013, China

*Corresponding author email: zuodongwangyi@126.com

Abstract. The PCA algorithm in coal and rock recognition had some fundamental flaws such as large computation and low efficiency. In this paper, a coal and rock image recognition method based on 2DPCA algorithm was proposed. The experimental results showed that this method could not only reduce the dimension of the original feature, but also reduce the complicated calculation and improve the recognition accuracy effectively.

Keyword: 2DPCA algorithm; Coal and rock recognition; Characteristic matrix.

1. Introduction
In recent years, with the deepening of coal and rock analysis, a variety of identification methods, such as natural gamma ray method, vibration analysis method based on coal seam, radar detection, infrared temperature of coal seam interface recognition and sound detection etc., are applied for coal and rock vitrinite recognition. These methods have certain application value, but there are also some limitations, such as gamma ray method can only apply to rock and coal containing low content of radioactive elements, while sound test exists a large difference to the coal and rock medium.

PCA (principal component analysis) is one of the commonly used methods in research. It can effectively reduce the dimension of the original matrix and the interference of other components. However, when PCA is applied to coal and rock analysis, it is necessary to transform the two-dimensional image matrix, the dimension of the image increases greatly and the feature extraction is difficult subsequently. In order to solve the above problems, according to the calculation principle of PCA, this paper adopts a 2DPCA method for identification. The method is simpler and the recognition rate is better than PCA.

2. Principle of 2DPCA algorithm
If X represented the m dimensional unitary column vectors, the idea of 2DPCA is to project the image m×n matrix A onto X by linear transformation Y =AᵀX. So the n-dimensional column vector Y called the projection eigenvector of image A is acquired.

Assuming that there are M training images A₁, A₂, ... AM as m×n matrix, and the total dispersion matrix G of the training mode is
\[ G = \frac{1}{M} \sum_{i=1}^{M} (A_i - \bar{A})(A_i - \bar{A})^T \]  

Among them, \( \bar{A} = \frac{1}{M} \sum_{i=1}^{M} A_i \) is the mean matrix of the training mode population, and it is simple to prove that G is a nonnegative definite matrix of \( m \times m \).

Define criterion function: \( J(x) = X^T G X \). In order to maximize the dispersion of the image matrix eigenvectors obtained after the projection on X, the unit vector of the criterion function is maximized and called the optimal projection vector. Correspondingly, the optimal projection vector is in accordance with the unit eigenvector of the maximum eigenvalue of G. Generally speaking, a single optimal projection direction is not enough in the case of a large number of samples. We need to find a set of optimal projection vectors \( X_1, X_2, \ldots, X_d \) that satisfied the orthonormal condition and maximum the criterion function.

\[ P = \{ X_1, X_2, \ldots, X_d \} \] is set and called the optimal projection matrix. For the known image sample A, after the transformation of \( Y_k = A^T X_k \), the projection feature vector \( Y_1, Y_2, \ldots, Y_d \) is obtained and called the principal component of image sample A. Based on the obtained principal components, the feature matrix of image sample A or feature graph \( B = [ Y_1, Y_2, \ldots, Y_d ] \) can be constructed as: \( B = A^T [X_1, X_2, \ldots, X_d] = A^T P \). \( \bar{A}_i \) is the mean vector matrix of the i training image samples, and the characteristic matrix of \( \bar{A}_i \) is \( \bar{B}_i \). Then the nearest neighbor classifier is used for pattern classification of image A.

### 3. Identification method of coal vitrinite based on 2DPCA

#### 3.1. Training

Assuming that there are N coal rock images in the training library represented by \( I_i \in \mathbb{R}^{M \times N} \) matrix, where the range of i is \( (1, N) \). The training steps of coal rock images are as follows:

- **Step1**: the average of all coal rock images in the training library are calculated as:
  \[ A = \frac{1}{N} \sum_{i=1}^{N} I_i \]  

- **Step 2**: the difference between the mean image and the coal image in the training library is acquired, and the standardized image is calculated as:
  \[ Y_i = I_i - A , \quad \forall i \]  

- **Step 3**: the total dispersion matrix of the training pattern is calculated
  \[ G = \frac{1}{N} \sum_{i=1}^{N} Y_i \cdot Y_i^T \]  

- **Step 4**: the eigenvalues and eigenvectors of G are calculated, and the optimal projection matrix P is obtained.

- **Step 5**: the feature matrix B of the image sample is constructed.

According to the above formula, the characteristic matrix B of the license plate image in the training library is calculated

#### 3.2. Identify

The distance between \( B_i = [ Y_1^{(i)}, Y_2^{(i)}, \ldots, Y_d^{(i)} ] \) and \( B_j = [ Y_1^{(j)}, Y_2^{(j)}, \ldots, Y_d^{(j)} ] \) is defined as:

\[ d(B_i, B_j) = \sum_{k=1}^{d} \| Y_k^{(i)} - Y_k^{(j)} \| \]  

Here the \( \| Y_k^{(i)} - Y_k^{(j)} \| \) represents the Euclidean distance between two principal component vectors \( Y_k^{(i)} \) and \( Y_k^{(j)} \). The characteristic matrix of training samples are expressed as \( B_1, B_2, \ldots, B_M \), and the
training sample category set is expressed as $W_k$, $k=1,2..., N$. M is the total number of training samples, and N is the single number of training samples. Assuming that the eigenmatrix of any test sample is represented as B, if $d(B,B_j) = \min d(B,B_i), j=1,2,...M$, so the test result is $B \in W_k$.

4. Experimental results and analysis

The collected pictures were divided into training samples and test samples. 50 typical single coal and rock images of $128 \times 128$ size were taken as test charts. In the experiment, each image was firstly divided into a processing window of $9 \times 9$. After pretreatment to remove the binder and triviality, the recognition results were shown in Table 1.

| Number of samples | Accuracy rate (%) |
|-------------------|-------------------|
| 50                | 92.75             |
| 30                | 91.35             |
| 10                | 90.50             |

As can be seen from the experimental data in Table 1, when the number of feature vectors is certain, the reliability of coal rock image recognition is related to the number of training samples. In other words, the more training samples, the higher the recognition accuracy.

At the same time, the time-consuming of PCA and 2DPCA under different training sample is measured and shown in Table 2. It is clear that the time consumption of 2DPCA method is obviously less in the feature extraction process, and the time difference between two methods is increased with the increasing number of training samples.

| Method  | Training sample number |
|---------|------------------------|
|         | 1         | 2         | 3         | 4         | 5         |
| PCA     | 6.72s     | 11.35s    | 22.37s    | 35.66s    | 53.80s    |
| 2DPCA   | 6.40s     | 10.01s    | 14.55s    | 22.17s    | 31.78s    |

5. Conclusion

The 2DPCA algorithm is applied to coal and rock image recognition, and through experimental comparison, it can be found that:

1) Compared with PCA algorithm, the matrix singular value decomposition and tedious calculation used in image feature extraction can be avoided based on 2DPCA algorithm.

2) The correct recognition rate of 2DPCA algorithm is related to the selection of training samples and test samples. The performance of 2DPCA algorithm is better than that of PCA algorithm under the same conditions.

The experimental results show that the 2DPCA algorithm has a better recognition rate and can be applied to coal and rock image recognition. In further research, the 2DPCA algorithm combined with other coal and rock recognition methods should be considered for improving the recognition rate further.

References
[1] Wang Ying. Study on coal and rock image identification. [J]. China Energy and Environmental Protection. 2019(4):139-142.
[2] Liu Zhongchao Liu Yongjun. Status analysis and development direction of coal and rock identification [J]. Journal of Nanyang Institute of Technology. 2018(4):26-30.
[3] Sun Jiping Yang Kun. The invention relates to a method of coal and rock image feature extraction and recognition. [J]. Industry and Mine Automation. 2017(5):1-5.
[4] Wang Shuisheng. Coal and rock identification control system of shearer based on vibration characteristic analysis. [J]. Industry and Mine Automation. 2015(5): 83-87.

[5] Wang Peizheng. The invention relates to a classification method of coal and rock macerals vitriol based on PCA and rbf-svm. [J]. Journal of China Coal Society. 2017(4): 977-984.

[6] Li Yang, Zhou Dongmei. Application of dpca in face recognition algorithm. [J]. Computer Knowledge and Technology. 2015(9): 193-194.

[7] Wang Fei. Application of 2DPCA in face recognition algorithm. [J] Information and computers. 2016(3): 117-119.

[8] Zeng Min. Research on face recognition technology based on 2DPCA on android platform. [D]. Southwest Jiaotong University. 2016.

[9] Jiang Na. Face recognition algorithm based on improved 2DPCA and correlation vector machine. [D]. Harbin university of science and technology. 2015.

[10] Zhang Danqing. Two-dimensional principal component analysis (pca) and its application. [D]. Xidian university. 2019.

[11] Wang Xia Luo Liumin Yang Jing. Face detection based on an improved 2DPCA algorithm. [J]. Wireless Internet Technology. 2018(10): 137-139.

[12] Feng Fei Jiang Baohua Liu Pei Chen Yujie. Improved application of 2DPCA algorithm in face recognition. [J]. Computer Science. 2017(S2): 267-268+311.

[13] Chen Huawei Wu Lushen Yuan Xiaocui. Design of 2DPCA algorithm and its application in fastener recognition. [J]. Computer Engineering and Design. 2016(2): 424-428.

[14] Ding Biwei. Face recognition based on principal component analysis (pca). [D]. Shandong normal university. 2019.