

A Method for Detecting Surface Defects in Insulators Based on RPCA

Wei Hu^{1(✉)}, Hongyu Qi², Zhenbing Zhao^{2(✉)}, and Leilei Xu²

¹ Dispatching and Control Center, State Grid Information and Telecommunication Branch, Beijing 100761, China
whu@sgcc.com.cn

² North China Electric Power University, Baoding 071003, China
zhaozhenbing@ncepu.edu.cn

Abstract. To solve such problems of existing methods for detecting surface defects in insulators as monotonous detectable category and long processing time, this paper presents a simple but effective detection approach based on Robust Principal Component Analysis (RPCA). Firstly, our method is based on insulator strings. We divide insulator string image into multiple insulators images, and then use these images as a test set. Secondly, due to the insulator has the characteristic of strong similarity, we decompose an insulator into a non-defective low-rank component and a defective sparse component by RPCA, and detect whether the insulator is defective. Through the Probabilistic Robust Matrix Factorization (PRMF) algorithm, the operation efficiency is improved. Furthermore, we verify the feasibility and effectiveness of the method by taking a great deal of experimental data.

Keywords: Detection · Insulator · Defect · RPCA · PRMF

1 Introduction

Insulators play an important role in transmission lines. However, it is vulnerable and can be thwarted by a fracture or other faults. These faults even cause power failure and serious economic losses or casualties. Therefore, monitoring the condition of insulators is of great significance.

As per literature [1, 2], the damage of insulators can be determined by the minimum law of the vertical gray statistic chart. Works [1, 2] realizes the damage detection of insulator without cover, but when the insulator is occluded partially, their method will destroy the minimum law of the insulator and lead to miscalculation. Work [3] divides an insulator into 10 parts, defines the standard of Contract-Mean-Variance (CMV) curve by three texture parameters of each part, and then detects whether the

This work was supported in part by the National Natural Science Foundation of China under grant number 61401154, by the Natural Science Foundation of Hebei Province of China under grant number F2016502101 and by the Fundamental Research Funds for the Central Universities under grant number 2015ZD20.

insulator is broken. However, when the insulator is divided, the damage is distributed on two adjacent parts. Under this condition, the stability of the CMV curve will be increased, thus probably causing miscalculation. Work [4] uses the edge slope feature model of crack to detect crack and obtains satisfactory results from the smooth insulators. But when the insulator surface is unsmooth, the method may cause miscalculation. Work [5] extracts the crack of binary image vertical projection, horizontal projection and two-order moment invariant of geometric moment invariants as the feature value, and detects the category of crack by Adaptive Resonance Theory (ART) network. The features extracted by this method is only applicable to the classification of the crack and it is monotonous. In [1–5], the number of fault categories is so limited, and the criterion of insulator fault is monotonous. They cannot deal with a variety of defects [6].

This paper presents a simple but effective approach to detect surface defects of insulators based on RPCA. The approach decomposes the image matrix of insulator into a low-rank component and a sparse component effectively. The low-rank image and the sparse image correspond to the non-defective component and the defective component, respectively. The advantages of the method are that the defects can be identified by separating the sparse component image and their categories can be determined. Moreover, we have reduced the computational complexity relatively.

2 Low-Rank Matrix Restoration Based on RPCA

RPCA is also called sparse and low-rank matrix decomposition [7] and it is evolved from the Principal Component Analysis (PCA) algorithm. PCA is the method that makes high-dimensional data in a low-dimensional space. It reduces the data dimension. Meanwhile, it can preserve more details in source images. The formula is as follows:

$$\begin{aligned} & \min \|\mathbf{E}\|_F \\ \text{s.t. } & \text{rank}(\mathbf{A}) \leq r, \mathbf{A} + \mathbf{E} = \mathbf{D} \end{aligned} \quad (1)$$

where \mathbf{A} is a low-rank component, \mathbf{E} is a sparse-error component, \mathbf{D} is original input data and each column represents the data of an observed image. $\|\cdot\|_F$ is the Frobenius matrix norm. However, there is a difficult problem in most error matrices, which is to recover low-rank matrices from high-dimensional data of observation matrices. These error matrices do not belong to the independent distribution of Gaussian noise. Moreover, PCA needs to know the dimensionality of low-dimensional feature space. But these problems are difficult to solve generally.

Based on the above PCA problems, Wright et al. [8] proposed RPCA to solve the question that data in \mathbf{A} have been seriously damaged. In other words, they transformed the PCA problem into the RPCA problem. We describe the RPCA problem abstractly as follows: when we know the observation matrix \mathbf{D} and $\mathbf{A} + \mathbf{E} = \mathbf{D}$, although we do not know \mathbf{A} and \mathbf{E} , but we know that \mathbf{A} is a low-rank component and \mathbf{E} is a sparse component in which nonzero elements can be arbitrarily large. And then we try to restore \mathbf{A} in this condition. The formula can be described as follows:

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{E}} \quad & \text{rank}(\mathbf{A}) + \lambda \|\mathbf{E}\|_0 \\ \text{s.t.} \quad & \mathbf{A} + \mathbf{E} = \mathbf{D} \end{aligned} \quad (2)$$

where the rank of matrix \mathbf{A} is represented by $\text{rank}(\mathbf{A})$, and $\|\cdot\|_0$ denotes the L_0 norm, λ denotes the weight of noise. If we can solve the RPCA problem by the correct λ , (\mathbf{A}, \mathbf{E}) can be accurately recovered as well. Nevertheless, Eq. (2) belongs to the NP-Hard and Strictly Non-convex Optimization problem, and there is no efficient solution [9, 10].

We can turn the RPCA problem into a convex optimization problem which can be solved easily by making the stress relaxation of Eq. (2). In other words, we use the L_1 norm instead of the l_0 norm and use the kernel norm $\|\mathbf{A}\|_* = \sum_i \sigma_i(\mathbf{A})$ of \mathbf{A} instead of the rank (\mathbf{A}) . Then we turn Eq. (2) into Eq. (3). And then we turn the problem into a convex optimization problem which is easy to solve relatively. The result as Eq. (3): under certain conditions, (\mathbf{A}, \mathbf{E}) can be solved uniquely and recovered ideally.

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{E}} \quad & \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1 \\ \text{s.t.} \quad & \mathbf{A} + \mathbf{E} = \mathbf{D} \end{aligned} \quad (3)$$

Where $\|\mathbf{A}\|_* = \sum_i \sigma_i(\mathbf{A})$, $\sigma_i(\mathbf{A})$ represents the i^{th} singular value of matrix \mathbf{A} , $\|\mathbf{E}\|_1 = \sum_{ij} \sigma_{ij}(\mathbf{A})$ represents the sum of the absolute values of elements in matrix \mathbf{E} .

The relaxation method is mainly to replace the non-convex l_0 norm into the l_1 norm which can measure the sparsity function easily. Thus, we can solve the RPCA problem by using the Convex optimization or non-linear programming to approximate the original problem [11]. In practical applications, with an increasing in detection frames, the input matrix \mathbf{D} increases rapidly. The difficulty of computer processing and the time of processing is increasing as well. Hence the processing method for real-time updating is essential.

3 PRMF Algorithm

With the increasing application of RPCA in computer vision and image processing, scholars have proposed a series of algorithms to solve the RPCA problem, such as Singular Value Decomposition (SVD) [12], Accelerated Proximal Gradient (APG) [13], Alternating Direction Method (ADM) [14] and Augmented Lagrange Multiplier (ALM) [15]. However, with an increase in the sequence of video frames, the number of the input matrix is increasing rapidly and it extends the calculation time and affects the efficiency. To solve this problem, this paper uses RPCA based on PRMF [16] to realize real-time processing of input sequence and our method can solve the above problems effectively.

3.1 Matrix Decomposition

\mathbf{Y} is a matrix with outliers and $\mathbf{Y} = [y_{ij}] \in R^{m \times n}$. The matrix decomposition of \mathbf{Y} is expressed as follows:

$$\min_{\mathbf{U} \in \mathbb{R}^{m \times r}, \mathbf{V} \in \mathbb{R}^{n \times r}} \left\| \mathbf{W} \otimes (\mathbf{Y} - \mathbf{UV}') \right\|_a^a \quad (4)$$

where $\mathbf{W} = [w_{ij}]$ is a matrix of size $m \times n$ pixels and only contains 0, 1. When the element y_{ij} in matrix \mathbf{Y} is not damaged, $w_{ij} = 1$; when y_{ij} is damaged, $w_{ij} = 0$.

a is the coefficient of the loss function: when $a = 1$, it corresponds to the L_1 norm; when $a = 2$, it corresponds to the L_2 norm.

For Eq. (4), we obtained the following formula by the regularization procedure.

$$\min_{\mathbf{U}, \mathbf{V}} \left\| \mathbf{W} \otimes (\mathbf{Y} - \mathbf{UV}') \right\|_a^a + \frac{\lambda_u}{2} \|\mathbf{U}\|_2^2 + \frac{\lambda_v}{2} \|\mathbf{V}\|_2^2 \quad (5)$$

In Eq. (5), λ_u and λ_v are the parameters of regularization, and they are greater than zero.

3.2 Matrix Decomposition Based on Likelihood Probability

From the perspective of Bayesian probability, the problem in Eq. (5) can be regarded as a Maximum A Posteriori problem. To improve robustness, we use the L_1 norm to decompose a matrix. The model of matrix decomposition based on probability is as follows:

$$\mathbf{Y} = \mathbf{UV}' + \mathbf{E} \quad (6)$$

$$u_{ij} | \lambda_u \sim N(u_{ij} | 0, \lambda_u^{-1}) \quad (7)$$

$$v_{ij} | \lambda_v \sim N(v_{ij} | 0, \lambda_v^{-1}) \quad (8)$$

where $\mathbf{E} = [e_{ij}]$ is the error matrix of size $m \times n$ pixels and each of its elements e_{ij} follows a separate Laplace distribution $L(e_{ij} | 0, \lambda)$. We get the following expression:

$$p(\mathbf{E} | \lambda) = \left(\frac{\lambda}{2}\right)^{mn} \exp\{-\lambda \|\mathbf{E}\|_1\} \quad (9)$$

To calculate \mathbf{U} and \mathbf{V} , we set λ_u , λ_v and λ to fixed values and then use the MAP based on Bayes' Rule, and we can get:

$$p(\mathbf{U}, \mathbf{V} | \mathbf{Y}, \lambda_u, \lambda_v) \propto p(\mathbf{Y} | \mathbf{U}, \mathbf{V}, \lambda) p(\mathbf{U} | \mathbf{Y}, \lambda_u) p(\mathbf{V} | \mathbf{Y}, \lambda_v) \quad (10)$$

$$\log p(\mathbf{U}, \mathbf{V} | \mathbf{Y}, \lambda_u, \lambda_v) = -\lambda \|\mathbf{Y} - \mathbf{UV}'\|_1 - \frac{\lambda_u}{2} \|\mathbf{U}\|_2^2 - \frac{\lambda_v}{2} \|\mathbf{V}\|_2^2 + C \quad (11)$$

where C is a constant independent of \mathbf{U} , \mathbf{V} . Thus, the problem of solving the maximum value of $\log p(\mathbf{U}, \mathbf{V} | \mathbf{Y}, \lambda_u, \lambda_v)$ is equivalent that of solving the minimum problem of the following equation:

$$\min_{\mathbf{U}, \mathbf{V}} \|\mathbf{Y} - \mathbf{UV}'\|_1 + \frac{\lambda'_u}{2} \|\mathbf{U}\|_2^2 + \frac{\lambda'_v}{2} \|\mathbf{V}\|_2^2 \quad (12)$$

In Eq. (12), $\lambda'_u = \lambda'_u/\lambda$ and $\lambda'_v = \lambda'_v/\lambda$.

3.3 PRMF Algorithm Based on RPCA

The general equation for RPCA has been presented in the previous section:

$$\min_{\mathbf{A}} \|\mathbf{A}\|_* + \lambda \|\mathbf{D} - \mathbf{A}\|_1 \quad (13)$$

We have known that Eq. (12) is the matrix decomposition equation based on probability and Eq. (13) is the general equation of RPCA, and there are some similarities between them. Literature [17] proposes a theorem of matrix factorization. For any matrix $\mathbf{Z} \in R^{m \times n}$, we have the following expression:

$$\|\mathbf{Z}\|_* = \min_{\mathbf{U}, \mathbf{V}, \mathbf{Z}=\mathbf{UV}'} \frac{1}{2} (\|\mathbf{U}\|_2^2 + \|\mathbf{V}\|_2^2) \quad (14)$$

When $\lambda'_u = \lambda'_v = \lambda_r$ and $r = k$, we can get the following results by substituting Eq. (12) for Eq. (14):

$$\begin{aligned} \min_{\mathbf{U}, \mathbf{V}} \|\mathbf{Y} - \mathbf{UV}'\|_1 + \frac{\lambda'_u}{2} \|\mathbf{U}\|_2^2 + \frac{\lambda'_v}{2} \|\mathbf{V}\|_2^2 \\ = \min_{\mathbf{U}, \mathbf{V}} \|\mathbf{Y} - \mathbf{UV}'\|_1 + \lambda_r \|\mathbf{UV}'\|_* \\ = \min_{\mathbf{X}, \text{rank}(\mathbf{X})=k} \|\mathbf{Y} - \mathbf{X}\|_1 + \lambda_r \|\mathbf{X}\|_* \end{aligned} \quad (15)$$

Through the above analysis, we established the RPCA method based on the PRMF. Next, we will use our method to solve the problems of surface defects detection of insulators effectively.

4 Analysis of Insulators Surface Defects Detection

4.1 Experiment on Public Datasets

Background modeling is a common application in computer vision processing, with the purpose of separating the dynamic part from the static background. The dynamic part includes moving objects and illumination variations. The static part of surveillance video is generally unchangeable, which can be regarded as the low-rank component \mathbf{A} . Moving objects and illumination variations change over time and they can be regarded as sparse component \mathbf{E} . This paper selects 200 frames of size 144×176 pixels from public datasets as test samples. As shown in Fig. 1:



Fig. 1. Foreground extraction results based on PRMF-RPCA, (a) original image frame, (b) the low-rank component, (c) the sparse component

From the perspective of computation time, we compared the PRMF-RPCA algorithm with several common RPCA algorithms. From Table 1, we found that PRMF-RPCA algorithm improves the efficiency effectively.

Table 1. The running time of several common RPCA algorithms

	PRMF	APG	IALM	ALM
Time/s	40.45	632.43	119.54	135.93

4.2 Surface Defects Detection of Insulator Strings

Insulators are an indispensable element for transmission lines. The good state of insulators is a prerequisite for power transmission line operation. But insulators can still be damaged by the wind and the sun in a natural environment. Harsh environments will lead to weathered oxidation, breakage and cracks of insulator strings. [18]. This paper presents a method for detecting surface defects in insulators based on PRMF-RPCA. The experiment used 8 insulator images from aerial video as an example. The experimental results are shown in Fig. 2.

As is shown, Fig. 2(a) displays the images of an insulator string. We can find that the insulator string is oxidized into holes in the first image. In the third image, the insulator has a little breakage. Other images exhibit different defects. Figure 2(b) shows the non-defect low-rank component. Figure 2(c) a defective sparse component. From Fig. 2, we can find that except for the defective part, the sparse component is in white and the rest is in black. The rate of defect detection of insulator strings is close to 100%.

4.3 Insulator Surface Defects Detection

To further detect tiny defects of the insulator string, our method is based on insulators. Insulators share a strong similarity which is expressed as the linear low-rank component in the image matrix. The defects on the insulator surface are different generally, as expressed by non-zero elements in the sparse matrices. Therefore, based on the above analysis, we use RPCA to detect surface defects of insulators.

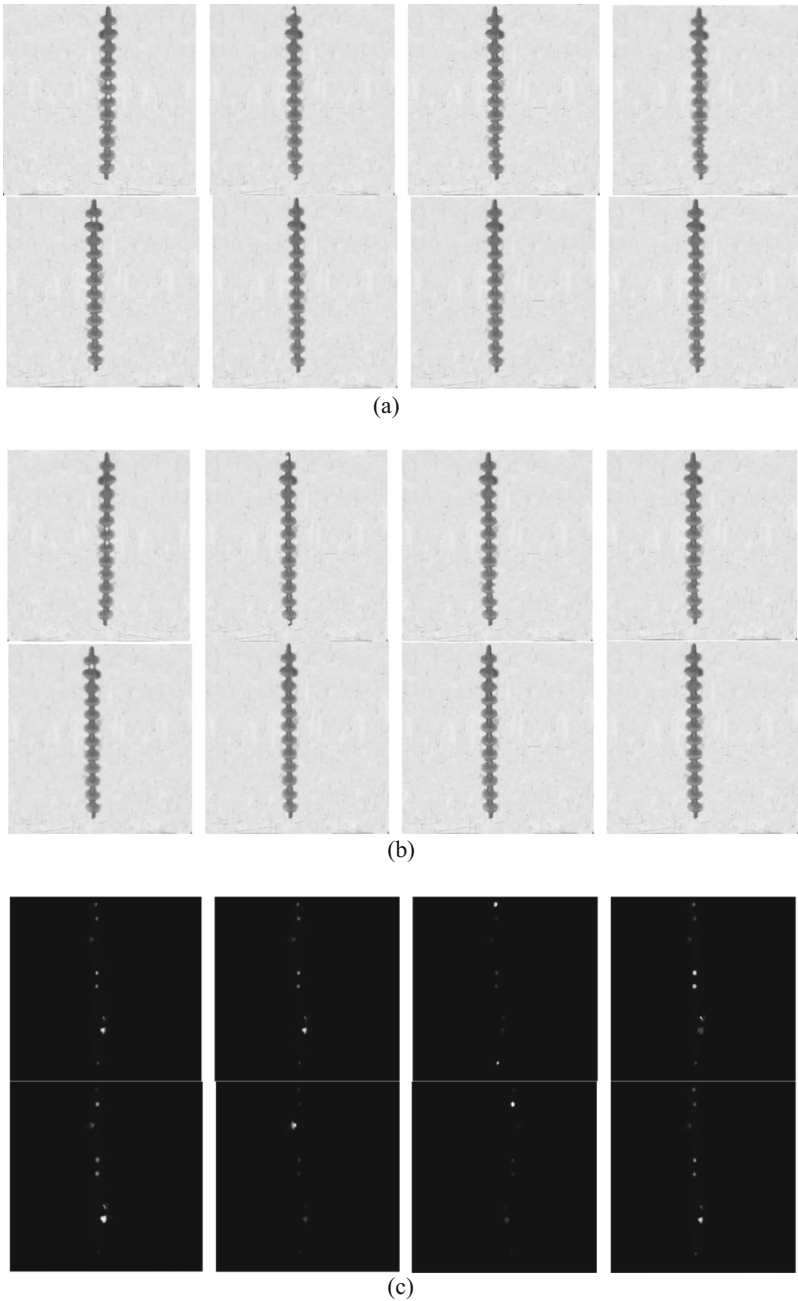


Fig. 2. Surface defects detection results of insulator, (a) the to-be-detected insulator string, (b) the low-rank component, (c) the sparse component

We convert a set of insulator images into a set of column vectors: X_1, X_2, \dots, X_N . And then we generate matrix $\mathbf{D} = [X_1, X_2, \dots, X_N] \in R^{m \times n}$ by concatenating all the vectors and using \mathbf{D} as the input of RPCA, where m is the number of pixels per image and n is the number of input images. To verify the feasibility and practicability of the principle and algorithm of PRMF-RPCA proposed in this paper, we select 20 images of insulators of size 256×256 pixels. As shown in Fig. 3: the type of defects includes noise, cracks, and string breakage.

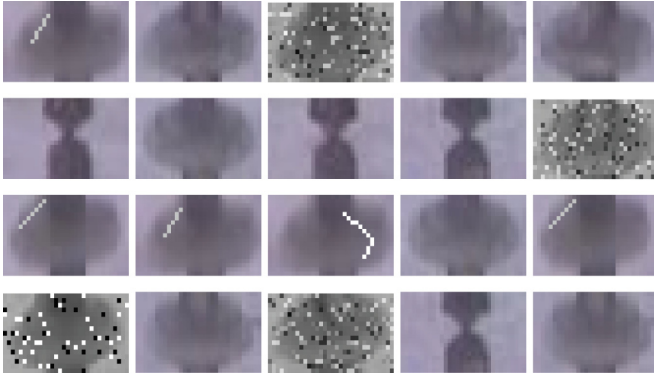


Fig. 3. Insulator images

The experimental results of using the PRMF-RPCA method are shown in Fig. 4: (a) is the non-defective low-rank component \mathbf{A} , (b) is the defective sparse component \mathbf{E} .

The experimental results show that we can detect the surface defects of insulator from 20 similar images by PRMF-RPCA algorithm successfully. From Figs. 3 and 4(b) we can see that the defects information which got by PRMF-RPCA is the same as the result of manual detection exactly and it is marked with white clearly. The non-defective information is displayed as a black picture in Fig. 4(b).

From Table 2 we can see that our method's running time of 20 images is 24.60 s; the average time is about 1.2 s, but the detection time of APG is 850.32 s, IALM is 122.52 s, and ALM is 168.5 s. From a real-time perspective, our approach has made significant progress over other methods.

We can distinguish the categories of insulator surface defects simply by the sparsity. But the different defects of insulators are expressed by sparse component, which is inadequate. As the experimental results shown in Fig. 4, we found that the gray value of sparse component of normal insulator is low, and then the image is smoother. But the gray value of string breakage changes greatly, and the smoothness of string breakage is poor.

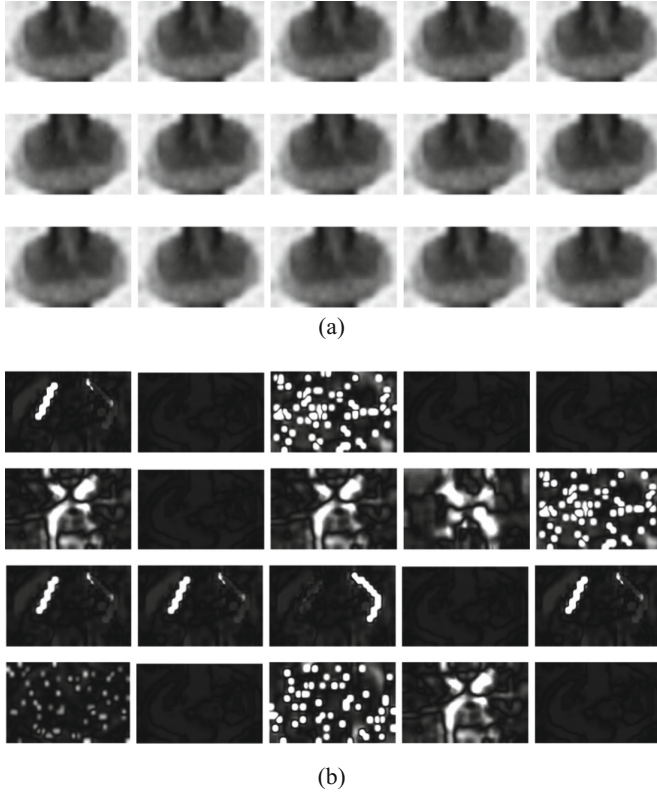


Fig. 4. Detection results of surface defects of insulator, (a) low-rank matrix image, (b) sparse matrix image

Table 2. Detection time of insulator surface defects in several common algorithms

	PRMF	APG	IALM	ALM
Time/s	24.60	850.32	122.52	168.46

$$S_p = \frac{1}{M} \sum_{i=0}^{M-1} (x_i - \bar{x})^2 \quad (16)$$

In Eq. (16), x_i is the value of each pixel of image, \bar{x} is the mean of the image. To distinguish the categories of surface defects of insulators better, we introduce a measure of smoothing to detect the categories of insulators. The smoothness of each category is shown in Table 3: we can conclude that the smoothness of normal insulators is smaller and its image is smoother, but the smoothness of the crack is relatively smaller. Because the noise is dispersed in insulators and it is heterogeneous, and the smoothness of the noise is larger than others'. Thus, we can detect the insulator surface defects better by the smoothness.

Table 3. Smoothness values of insulation surface defects

Defect category	Normal	String breakage	Crack	Noisy
Smoothness	0.0007	0.0667	0.0273	0.1195

5 Conclusion

To detect insulator surface defects from a mess of aerial images, the traditional artificial method is difficult to feed the actual requirements of the engineering. Subsequent scholars have proposed a variety of defects detection methods of insulator surface and they have achieved some results. However, most of their methods have plenty of shortcomings and can only detect categories monotonously.

In this paper, for the similarities of insulators, we decompose the defective insulator into the low-rank component and the sparse component by PRMF-RPCA. Specifically, we introduce a measure of smoothing to detect the categories of insulators. The experimental results show that the method improves the efficiency and ability of surface defects detection of insulators, and realizes the real-time updating detection of insulators.

References

1. Yang, H.M., Liu, Z.G., Han, Y., et al.: Foreign body detection between insulator pieces in electrified railway based on affine moment invariant. *Power Syst. Technol.* **37**(8), 2297–2302 (2013)
2. Yang, H.M., Liu, Z.G., Han, Z.W., et al.: Foreign body detection between insulator pieces in electrified railway based on affine moment invariant. *J. China Railw. Soc.* **35**(4), 30–36 (2013)
3. Zhang, X.Y., An, J.B., Chen, F.M.: A method of insulator fault detection from airborne images. In: 2010 Second WRI Global Congress on Intelligent Systems (GCIS), vol. 2, pp. 200–203 (2010)
4. Sun, J.: Research on crack detection of porcelain insulators based on image detection. North China Electric Power University, Baoding, China (2008)
5. Liu, G.H., Jiang, Z.J.: Recognition of porcelain bottle crack based on modified ART-2 network and invariant moment. *Chin. J. Sci. Instrum.* **30**(7), 1420–1425 (2009)
6. Yao, M.H., Li, J., Wang, X.B.: Solar cells surface defects detection using RPCA method. *Chin. J. Comput.* **36**(9), 1943–1952 (2013)
7. Jean, J.H., Chen, C.H., Lin, H.L.: Application of an image processing software tool to crack inspection of crystalline silicon solar cells. In: *Machine Learning and Cybernetics (ICMLC)*, vol. 4, pp. 1666–1671, Singapore (2011)
8. Wright, J., Peng, Y.G., Ma, Y., et al.: Robust principal component analysis: exact recovery of corrupted low-rank matrices via convex optimization. *Adv. Neural Inf. Process. Syst.* **87**(4), 20:3–20:56 (2009)
9. Chen, M.M., Lin, Z.C., Shen, X.Y.: Algorithm and implementation of matrix reconstruction. University of Chinese Academy of Sciences, Beijing (2010)
10. Candès, E.J., Li, X., Ma, Y., et al.: Robust principal component analysis? *J. ACM (JACM)* **58**(3), 11 (2011)

11. Jin, J.Y.: Separation of image based on sparse representation. Xidian University, Xi'an, China (2011)
12. Cai, J.F., Candès, E.J., Shen, Z.: A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.* **20**(4), 1956–1982 (2010)
13. Peng, Y., Ganesh, A., Wright, J., et al.: RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(11), 2233–2246 (2012)
14. Ke, Q., Kanade, T.: Robust L1 norm factorization in the presence of outliers and missing data by alternative convex programming. In: 2005 IEEE Conference on Computer Vision and Pattern Recognition, pp. 739–746. IEEE (2005)
15. Lin, Z., Chen, M., Ma, Y.: The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. arXiv preprint [arXiv:1009.5055](https://arxiv.org/abs/1009.5055) (2010)
16. Wang, N., Yao, T., Wang, J., Yeung, D.-Y.: A probabilistic approach to robust matrix factorization. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012. LNCS, vol. 7578, pp. 126–139. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33786-4_10
17. Mazumder, R., Hastie, T., Tibshirani, R.: Spectral regularization algorithms for learning large incomplete matrices. *J. Mach. Learn. Res.* **11**(Aug), 2287–2322 (2010)
18. Cui, K.B.: Research on the key technologies in insulator defect detection based on image. North China Electric Power University, Beijing (2016)