IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 7, 2019, accepted February 5, 2019, date of publication March 6, 2019, date of current version April 2, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2899362

Moving-Object Tracking Algorithm Based on PCA-SIFT and Optimization for Underground Coal Mines

JIANG DAI-HONG[1], DAI LEI[1], LI DAN[1], AND ZHANG SAN-YOU[2,3]

[1]Key Laboratory of Intelligent Industrial Control Technology of Jiangsu Province, Information and Electrical Engineering College, Xuzhou University of Technology, Xuzhou 221000, China
[2]China University of Mining and Technology, Xuzhou 221000, China
[3]Department of Science and Technology, Suzhou Wujiang District Public Security Bureau, Suzhou 215200, China

Corresponding author: Jiang Dai-Hong (daihongjiangr@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61379100, in part by the Major Project of Natural Science Research of the Jiangsu Higher Education Institutions of China under Grant 18KJA520012, in part by the Key Laboratory of Intelligent Industrial Control Technology of the Jiangsu Province Research Project under Grant JSKLIIC201705, and in part by the Xuzhou Science and Technology Plan Project under Grant KC18011.

ABSTRACT In view of the complex and changeable environment in underground coal mines, an improved algorithm based on the principal component analysis-scale invariant feature transform (PCA-SIFT) and mean shift is proposed to address the issues for which existing tracking algorithms are not adequate; for example, when differentiating between moving targets and the background, the tracking in the case of moving objects (e.g., confusion between foreground and background) is not optimal. This results in poor resolution and the inability to deal with very dusty conditions, scale change, and rotation. The proposed feature target tracking model was developed using the scale invariance property of the PCA-SIFT feature-extraction algorithm. Finally, the mean-shift method was used to track moving objects. The experimental results showed that the optimized algorithm for tracking moving objects was significantly better and more robust than the existing algorithm.

INDEX TERMS Target tracking, scale invariant feature transform, mean shift, target detection.

I. INTRODUCTION

Underground monitoring, in environments such as coal mines, involves multiple moving targets. It is necessary to respond instantly to various monitoring anomalies because of the complex and variable nature and circumstances of underground coal mine operations, including frequent accidents and developing links with other systems. Therefore, it is of great importance to develop an algorithm for the timely tracking of moving objects. Research and application of moving-target positioning is an important part of intelligent video perception, and is being studied in China and other parts of the world. Mukherjee *et al.* [1] proposed a position-tracking method to distinguish and track multiple targets in a given scenario, based on the colors of the moving targets. Christiansen *et al.* [2] used the moving target segmentation method based on the points on the contour of the feature to achieve accurate and complete separation of overlapping moving targets. This result was then combined with the tracking model of the extended Kalman filter to realize real-time tracking of the moving target. Niu *et al.* [3] proposed target tracking using the Snake model. This model is a deformable curve defined in the image domain to conform to the target contour by adjusting the natural shape of the Snake. Thus, the target is tracked by way of the motion-estimation method. Akoglu *et al.* [4] introduced a new Snake target-tracking algorithm called Features from Accelerated Segment Test (FAST)-Snake using the improved FAST corner feature-matching to estimate the global affine transformation of the target contour framework and select a projection contour point as the initial contour of the Snake model. He *et al.* [5] first proposed a tracking method based on a three-dimensional (3D) model. Building on Zhan's model, Yuan *et al.* [6] suggested a ground constraint condition that reduces the number of possible target positions by limiting the range of motion of the moving object, thereby improving

The associate editor coordinating the review of this manuscript and approving it for publication was Kun Mean Hou.

the tracking efficiency. To solve the problem of poor accuracy in traditional edge-based 3D visual tracking, Zhang *et al.* [7] proposed a 3D tracking method based on corner features. Other similar algorithms have also been proposed [8]–[11].

The above methods have limitations, especially in application areas such as coal mines, parking lots, and other underground environments. Compared to ground video surveillance, it is difficult to perform image segmentation and estimate target locations owing to limitations in space, poor resolution, uneven illumination, and significantly changing background illumination intensity. Furthermore, the presence of other factors such as dust, moisture, and obstructions severely reduce the quality of surveillance images. Therefore, traditional target-positioning algorithms often fail to produce ideal results or real-time tracking of the target. This hinders the in-depth development and application of intelligent video surveillance. In view of the environment in underground coal mines, Li *et al.* [12] proposed a SIFT algorithm specifically for underground coal mine environments for the detection moving targets. However, the accuracy of their algorithm is reduced when the perspective increases at large detection angles.

Herein, a new algorithm that combines the strengths of the broadest intra-class variance algorithm with principal component analysis–scale invariant feature transform (PCA-SIFT) and the mean shift method is proposed. The proposed optimizing algorithm first selects a threshold segment by improving Otsu's method [13], then combines it with PCA-SIFT (which has the advantage of scale invariance and anti-noise) and incorporates the scale-invariance feature into the mean-shift tracking method. This resolves the problems of tracking loss inherent in current tracking algorithms, which is a result of weak distinction between moving targets and background, poor tracking efficiency when the scale changes, rotation, and/or noise interference by the moving target. The proposed new algorithm, which has been empirically tested, promises to produce higher accuracy and improved real-time performance.

II. MOVING-TARGET DETECTION BASED ON IMPROVED OTSU'S METHOD

The weak adaptability of traditional methods for moving-target detection in complex environments often results in missing or discontinuous edges of the target image. In the present study, threshold segmentation was conducted by adopting the improved maximum intra-class variance method (Otsu's method [13]) based on inter-frame difference. The proposed method uses morphological filtering to eliminate noise and improve detection in complex environments, and quickly detect the moving target in the video frame. Thus, it provides a sound initial condition for follow-up target tracking.

Otsu's method traditionally calculates the threshold at the entire grey level [14], which requires considerable computational capability to calculate the size of the variance between classes for each grey value. Perfect image separation from a complex background is difficult to achieve. To improve the real-time performance, we propose a new threshold segmentation method for detecting moving targets based on Otsu's method [13]. The specific steps are as follows:

Step 1: Initial threshold segmentation. The initial threshold segmentation is conducted using the average grey value of the entire image. Selection of this value is based on simple image statistics This eliminates the need to analyze complex image histograms. If the grey scale value of an MN grey scale image is (0,1, …, L–1), the initial threshold T0 is given by

$$T_0 = \frac{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)}{M \times N} \tag{1}$$

where f(x,y) is the pixel grey value at image coordinates x, y.

Step 2: Calculate the mean of the target. The initial threshold T0 divides the image into two parts. The part less than T0 is the target area C0. The part greater than T0 is the target area C1. The operation is as follows:

$$C_0 = \{f_{C_0}(x,y) \,|0 \le f(x,y) \le T_0\} \tag{2}$$
$$C_1 = \{f_{C_1}(x,y) \,|L-1 \ge f(x,y) > T_0\}$$
$$f_{C_0}(x,y) = \begin{cases} 1, & 0 \le f(x,y) \le T_0 \\ 0, & f(x,y) > T \end{cases} \tag{3}$$

The average grey scale values of C0 and C1 are calculated using (4) and (5), respectively:

$$T_1 = \frac{\sum_{i=0}^{T_0} f_{C_0}(x,y) * i}{\sum_{i=0}^{T_0} f_{C_0}(x,y)} \tag{4}$$

where i is the threshold value of 0 to T0; and

$$T_d = \frac{\sum_{i=T_0+1}^{L-1} f_{C_1}(x,y) * i}{\sum_{i=T_0+1}^{L-1} f_{C_1}(x,y)} \tag{5}$$

Step 3: Calculate the optimal threshold Td as follows:

$$T_d = \left\{ C = \frac{1}{1+T_0^2}\left(\frac{\sum_{i=0}^{T_0} f_{C_0}(x,y) * i}{\sum_{i=0}^{T_0} f_{C_0}(x,y)} - \frac{\sum_{i=T_0+1}^{L-1} f_{C_1}(x,y) * i}{\sum_{i=T_0+1}^{L-1} f_{C_1}(x,y)}\right)^2, C \ge T_0 \right\} \tag{6}$$

Td is calculated using Otsu's method [13], thus avoiding the search for the optimal threshold over the entire grey scale level. This approach not only reduces the amount of computation, but also maintains the continuity of the target image contour.

(a) and (b) in Figure 1 are the 55th and 56th frames of a narrow-track video of a section of Taoyuan dispatching station. (c) and (d) are the result of color adjustment. Histogram equalization of the original frame is conducted, and then the original frame is filtered using a mean filter. Thus, the contrast ratio is improved, and the image is clearer and easier to handle. A comparison of (e) and (f) shows that the improved foreground area displays the contour of the moving target with greater clarity.

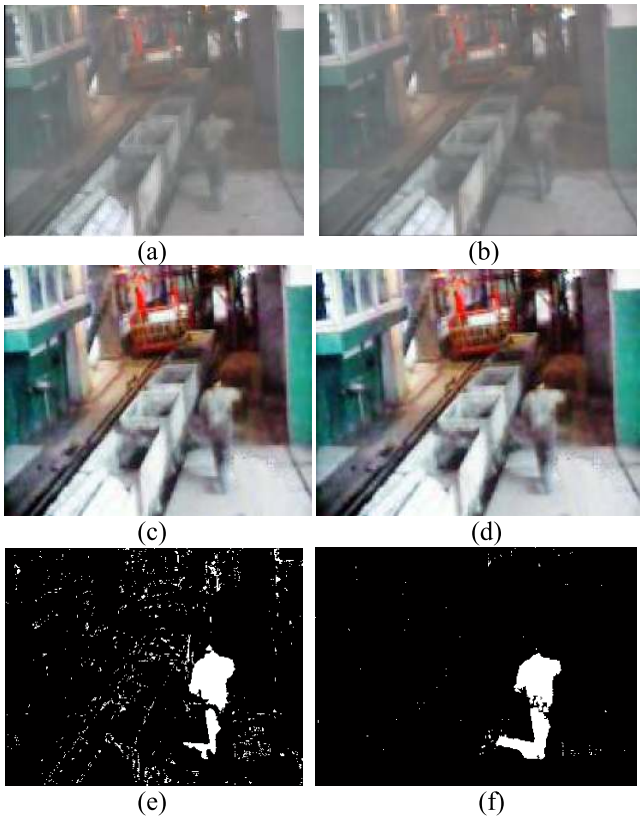


FIGURE 1. Test results of improved OTSU. (a) 55th frame. (b) 56th frame. (c) 55th Color adjustment. (d) 56th Color adjustment. (e) Foreground before. (f) Foreground after.

III. TARGET-TRACKING ALGORITHM BASED ON PCA-SIFT AND MEAN SHIFT

The mean-shift algorithm is often applied in the field of target tracking. However, it is relatively rarely used in the target color feature description when the kernel function histogram model is directly used. This is because it leads to tracking deviation or even target loss.

The PCA-SIFT algorithm is a detection method that uses scale-invariant key points based on a scale space property. It uses the PCA-SIFT feature vector of the image key point for matching. The feature points extracted by this algorithm are scale and rotation invariant. Compared with the ordinary SIFT algorithm, PCA-SIFT greatly reduces calculation time because of its variable dimension. Furthermore, PCA-SIFT is more robust in terms of change of perspective, affine transformation, and noise shift [15]. Therefore, by combining the advantages of these two algorithms, the proposed method incorporates the scale-invariance feature into the mean-shift tracking method to produce a new intelligent vision-positioning algorithm. This incorporation improves the accuracy of positioning and real-time performance of the algorithm.

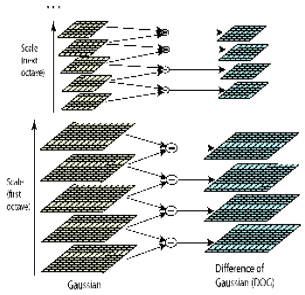


FIGURE 2. DOG scale space.

The specific process in each phase is described in detail below.

A. ESTABLISH SCALE SPACE

The scale space of the image can be expressed as the convolution of the scale space at different scales with the Gaussian kernel, defined as follows:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{7}$$

in which $G(x, y, \sigma)$ is a scale-variable Gaussian function given by

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)}/2\sigma^2 \tag{8}$$

where (x, y) are the spatial coordinates and σ is the scale space factor. Figure 2 shows the difference of Gaussians (DOG) scale space. A Gaussian pyramid is constructed by convolving the image $I(x, y)$ with the Gaussian kernel $G(x, y, \sigma)$ for different values of σ. To detect the stable key points in the scale space effectively, the adjacent scales of images k and k–1 are differentiated, as shown in Figure 3, and the Gaussian difference scale space is obtained, which is defined as follows:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \tag{9}$$

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, (\mathrm{k}-1)\sigma)) * I(x, y) \\ &= L(x, y, (k-1)\sigma) - L(x, y, \sigma) \end{aligned} \tag{10}$$

B. KEY POINT POSITIONING

The detected local extreme points require further precise positioning to become feature points. A 3D quadratic function is used to fit the curve of DOG scale-space to accurately determine the position and scale of the key points. Simultaneously, it is necessary to remove key points that have low contrast, eliminate the unstable edge response points, and improve anti-noise ability, to enhance the stability of target tracking. This is accomplished as follows:

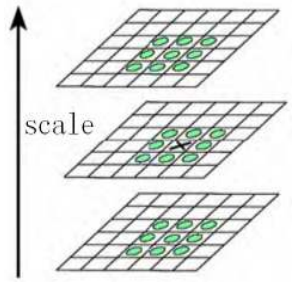


FIGURE 3. Regional extremum detection.

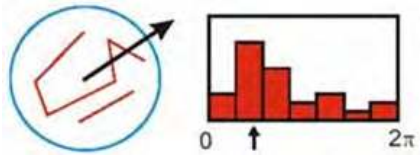


FIGURE 4. Main direction.

Quadratic Taylor expansion of $D(x, y, \sigma)$:

$$D(X) = D + \frac{\partial D^T}{\partial X}X + \frac{1}{2}X^T\frac{\partial^2 D}{\partial X^2}X \tag{11}$$

Taking the derivative of X and setting the first derivative in (13) to 0, the extreme point is obtained:

$$X' = -\frac{\partial^2 D^{-1}}{\partial X^2}\frac{\partial D}{\partial X} \tag{12}$$

Taking the derivative of X' the extreme point in (13), we have the following:

$$D(X') = D + \frac{1}{2}\frac{\partial D^T}{\partial X}X' \tag{13}$$

After the exact position and scale of the feature points are obtained, the description operator is set invariant to the image rotation, as shown in Figure 4. The size and direction of the gradient at the position are calculated as follows:

$$m(x,y) = \sqrt{[L(x+1,y)-L(x-1,y)]^2+[L(x,y+1)-L(x,y-1)]^2} \tag{14}$$

$$\theta(x,y) = \tan^{-1}\{[L(x,y+1)-L(x,y-1)]/[L(x+1,y)-L(x-1,y)]\} \tag{15}$$

where L is the scale of each key point. As shown in Figure. 5, the coordinate axis is rotated in the direction of the SIFT feature region, and an 8 × 8 window is taken as the center of the key point position of the feature region. Eight direction gradient histograms are calculated on each 4 × 4 small block, and one key point is composed of 4 × 2× 2 seed points with eight direction vector information. Lowe uses 4 × 4 small windows in practical applications, and each feature point is characterized by a 128-dimensional vector. Finally, to remove the effects of illumination changes, the feature vectors are normalized.

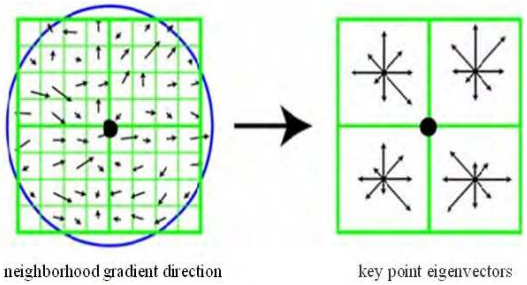


FIGURE 5. SIFT feature vector.

C. PCA DIMENSION REDUCTION

Although the SIFT algorithm adapts to various affine transformations, the SIFT feature vector can be as large as 128 dimensions. This high dimensionality significantly affects the real-time performance of the coal mine safety monitoring system. In this paper, the principal component analysis (PCA) in statistics is combined with SIFT to reduce the number of dimensions [16], [17]. The PCA method based on the variable covariance matrix is used to process, compress and extract the information; therefore, the first few principal components with a high contribution rate contain most of the information of the feature vector. Replacing the main component of the sample with the first few irrelevant principal component values results in a high cumulative contribution rate and a dramatic reduction of vector dimension. Furthermore, this also results in a more efficient feature matching. Therefore, the real-time performance is improved.

Let each feature point in the image be represented by p feature description vectors. Let n denote the number of feature points, i.e., dimensions, that constitute a sample matrix. Through the sample PCA, the p feature description vectors that affect each other are reduced in terms of their dimension. The process is performed in the following steps:

Step 1: Let $x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^T, i = 1, 2, \ldots n.$ be a sample of $X = (X_1, X_2, \ldots, X_n)^T$ whose dimension is n, then the sample covariance matrices is as follows:

$$\sum = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})(x_k - \bar{x})^T \tag{16}$$

Of which

$$\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p)^T, \quad x_j = \frac{1}{n}\sum_{i=1}^{n}x_{ij},\ i = 1, 2, \ldots p$$

Step 2: The feature value of sample covariance matrices $\sum$ is $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots \geq \hat{\lambda}_p \geq 0$, and the orthogonal unitized feature vector is $\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_p, \hat{e}_i = (\hat{e}_{i1}, \hat{e}_{i2}, \ldots, \hat{e}_{ip})$,

$$\hat{e}_i\hat{e}_i^T = 1, \quad i = 1, 2, \ldots p \tag{17}$$

Step 3: The principal component of the ith sample is computed as follows:

$$y_i = \hat{e}_{i1}x_1 + \hat{e}_{i2}x_2 + \ldots + \hat{e}_{ip}x_p, \quad i = 1, 2, \ldots p \tag{18}$$

Substituting the n observations of X, $x_k = (x_{k1}, x_{k2}, \ldots, x_{kp})^T, k = 1, 2, \ldots n$, the n observations of the principal

component of the ith sample are obtained: $y_{ki}(k = 1, 2, \ldots n)$, which is the score of the ith principal component.

Step 4: The sample covariance of y_i and $y_j = 0$, $i \neq j$, total variance of the sample $= \sum_{i=1}^{p} \hat{\lambda}_i$, and the principal component contribution rate of the ith sample of X is $\hat{\lambda}_i / \sum_{i=1}^{p} \hat{\lambda}_k (i = 1, 2, \ldots p)$. The cumulative contribution rate of the main components of the first m samples is $\sum_{i=1}^{m} \hat{\lambda}_i / \sum_{k=1}^{p} \hat{\lambda}_k$.

D. ESTABLISH A TARGET-TRACKING FEATURE MODEL

Assuming $t_i(t_{ix}, t_{iy})$ is the coordinate position of the ith pixel in the target model, the initial frame-center of the tracked target is d_0, and the current frame-center is d_1. The initial frame and the current frame target tracking feature models are as follows:

$$\hat{q}_u = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{t_i - d_1}{h}\right\|^2\right) \delta\left[b(t_i) - u\right] \tag{19}$$

$$\hat{p}_u(d_1) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{t_i - d_1}{h}\right\|^2\right) \delta\left[b(t_i) - u\right] \tag{20}$$

where n is the total number of pixels, $k(\|x\|^2)$ is the kernel function, h is the template radius, and functions b and δ are used to determine whether the color value of the ith pixel belongs to the eigenvalue u and whether the normalization coefficients C and C_h satisfy the following equations:

$$C = \frac{1}{\sum_{i=1}^{n} k\left(\left\|\frac{t_i - d_0}{h}\right\|^2\right)} \tag{21}$$

$$C_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{t_i - d_1}{h}\right\|^2\right)} \tag{22}$$

The most critical step in the mean-shift target tracking algorithm is to find the matching parameters such that the Bhattacharyya similarity coefficient is maximized, which means that the location of the region closest to the target model is pinpointed. The Bhattacharyya similarity coefficient is defined as follows:

$$\hat{\rho}(t) \equiv \rho[\hat{p}(t), \hat{q}] = \sum_{u=1}^{m} \sqrt{\hat{p}_u(t)\hat{q}_u} \tag{23}$$

In the above formula, m represents the number of image pixels.

The Bhattacharyya similarity coefficient is maximized if the search target in the current frame starts from the target center position d_1 of the previous frame. The Taylor expansion of (24) at d_1 is as follows:

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{b=1}^{m} \sqrt{\hat{p}_b(\hat{y}_0)\hat{q}_b} + \frac{1}{2} \sum_{b=1}^{m} \hat{p}_b(y) \sqrt{\frac{\hat{q}_b}{\hat{p}_b(\hat{y}_0)}} \tag{24}$$

In (24), the pixel in the center of the search window in the previous frame image is denoted by y_0, and the current target optimal target position is denoted by y.

In the actual tracking process, the displacement of the target in the continuous frame is minimal; hence, (24) can be approximated as

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{b=1}^{m} \sqrt{\hat{p}_b(\hat{y}_0)\hat{q}_b} + \frac{C_h}{2} \sum_{b=1}^{nh} w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \tag{25}$$

in which

$$w_i = \sum_{\mu=1}^{m} \sqrt{\frac{\hat{q}_b}{p_b(\hat{y}_0)}} \delta[B(x_i) - \mu] \tag{26}$$

where $B(x_i)$ is the characteristic value of the position coordinates x_i, mapped from the feature space. The mean-shift vector for the target trace is the extreme value of (25):

$$M_{h,G}(t) = d_c - d_1 = \frac{\sum_{i=1}^{n_h} G\left(\left\|\frac{t_i - d_1}{h}\right\|^2\right) w_i t_i}{\sum_{i=1}^{n_h} G\left(\left\|\frac{t_i - d_1}{h}\right\|^2\right) w_i} - d_1 \tag{27}$$

Continuous iteration of $d_1 \leftarrow d_c$ gradually approaches the optimal value, finally converging to the local maximum, and thus obtaining accurate target positioning.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

To verify the effectiveness of the algorithm, an experimental verification and analysis was performed.

Experimental environment: We used a Pentium 4 CPU (1.80 GHz), with 2 GB of storage, and 256 MB of memory. The programming software was Visual C ++ 6.0, and the operating system was Microsoft Windows XP. Equation (28) was used to calculate the deviation of the ith frame tracking, and the tracking results were analyzed. Thus, we define the tracking bias as follows:

$$err_i = \sqrt{(d_x - \sigma_x)^2 + (d_y - \sigma_y)^2} \tag{28}$$

where d_x and d_y are the center positions of the ith frame tracking target. σ_x, σ_y are the center positions of the real target. err_i is the tracking bias of the ith frame. The smaller the value of err_i the more accurate the tracking.

The image of the 300 × 300 underground locomotive shown in Fig. 6 is selected for experiment. The points in a 12 × 12 area are selected to describe the feature points. The pixel grayscale value in the gray-scaled region is used as the sample feature description vector for sample PCA, and the result is shown in Table 1.

It can be seen from Table 1 that the feature vector dimension is as high as 144 dimensions; however, the first eight principal components have covered most of the information of the feature vector, and the cumulative contribution rate of the sample principal component is as high as 95%. Experiments show that combining SIFT with the sample principal

FIGURE 6. Feature points of extracted locomotive image.

TABLE 1. Principal component contribution rate of locomotive image.

Principal component	Feature value	Feature value variance	Principal component contribution rate (%)	Cumulative contribution rate (%)
1	4.41439183	2.36945738	0.4013	0.4013
2	2.04493446	0.80303153	0.1859	0.5872
3	1.24190293	0.29798645	0.1129	0.7001
4	0.94391647	0.22911305	0.0858	0.7859
5	0.71480343	0.29822424	0.065	0.8509
6	0.41657919	0.01050982	0.0379	0.8888
7	0.40606937	0.10042222	0.0369	0.9257
8	0.30564715	0.06445848	0.0278	0.9535

components greatly reduces the feature vector dimension and improves the matching efficiency.

A sample video at a coal mine auxiliary wellhead was selected for the tracking experiment. The mean shift algorithm and the PCA-SIFT combined with mean shift algorithm proposed herein were used for tracking and comparison. The video (image pixel size 320 × 240) collected at the coal mine dispatching station shows that the illumination contrast of the monitoring screen was low, and the picture was blurred, as it was mixed with noise. The results are shown in Figure 7(a) and (b). Four images (20th, 30th, 40th, and 50th frames) were tracked. Initially there was no clear difference between the two algorithms; however, after the 30th frame, the SIFT-based mean-shift tracking algorithm began to deviate significantly due to the distant sight line, low contrast, noise, and obstruction. Given that the SIFT-based mean-shift tracking algorithm uses color features and the feature description is quite simple, the target was finally lost. In comparison, the algorithm that combines PCA-SIFT with the Mean Shift algorithm detects and locates moving targets better. To better verify the accuracy and stability of the proposed algorithm, the tracking errors of different algorithms were calculated. The comparison curve is shown in Figure. 8. It can be seen from the figure that after 30 frames, the error of SIFT-based mean shift algorithm increases significantly. The proposed algorithm is more robust, and is better at overcoming the problems of tracking failure caused by illumination changes, noise interference, and occlusion in complex environments.

To verify the robustness of the tracking algorithm in the various scenarios underground, ten videos are obtained from different scenes such as a dispatching station, head and a transformer substation. Each video is taken as a group, and ten segments are obtained in each group. The underground

FIGURE 7. Comparison of tracking results. (a) Mean shift; (b) proposed algorithm.

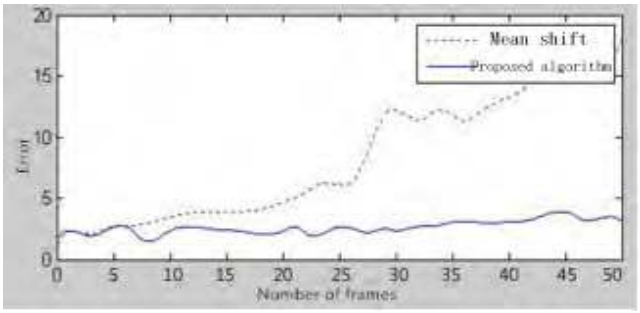


FIGURE 8. Comparison of the tracking error of the two algorithms.

rigid or non-rigid moving targets in these obtained short videos are tracked using different methods. In terms of algorithm efficiency, Table 2 shows a comparison of algorithm processing frame efficiency. It can be seen from the table that the PCA-SIFT algorithm proposed herein takes less time to process the same number of frames in a video. Table 3 shows the comparison of algorithm tracking accuracy. The proposed algorithm is based on PCA-SIFT feature detection algorithm. Therefore, the overall real-time performance of the algorithm is better than that of the Mean Shift tracking algorithm, which is based on color

FIGURE 9. Tracking comparison of different algorithms in coal mine substation. (a) Mean shift algorithm. (b) PCA-SIFT algorithm.

TABLE 2. Comparison of frame processing efficiency.

Algorithm	Frame	Time consumption (ms)
Mean shift	100	306
Proposed	100	256

TABLE 3. Contrast of tracking accuracy.

	Mean-shift algorithm	Proposed algorithm
Actual number of moving targets	100	100
Detected moving targets	83	95
Accurate recognition of moving targets	75	90
Recognition accuracy	75%	90%

histogram features. Moreover, the target detection rate and recognition rate are higher. This greatly improves the tracking accuracy.

V. CONCLUSION

Herein, a new tracking algorithm based on differentiation between frames is proposed. The proposed algorithm exploits the PCA-SIFT improved maximum intra-class variance method to detect targets. Additionally, using feature vectors, a target tracking feature model was established, and the mean-shift method was applied to position the moving target. The empirical results show that the proposed algorithm improved the tracking effect in complex environments, including changes in rotation and scale of the moving target, noise interference, occlusion rotation, and unstable illumination. The algorithm makes a significant contribution in achieving robust and effective tracking of moving targets and to promote safe operation in underground coal mines.

REFERENCES

[1] S. Mukherjee, A. Ahmed, D. P. Dogra, S. Kar, and P. P. Roy. (2018). "Fingertip detection and tracking for recognition of air-writing in videos." [Online]. Available: https://arxiv.org/abs/1809.03016
[2] R. H. Christiansen, J. Hsu, M. Gonzalez, and S. L. Wood, "Monocular vehicle distance sensor using HOG and Kalman tracking," in *Proc. 51st Asilomar Conf. Signals, Syst., Comput.*, Nov. 2017, pp. 178–182.
[3] S. Niu, Q. Chen, L. de Sisternes, Z. Ji, Z. Zhou, and D. L. Rubin, "Robust noise region-based active contour model via local similarity factor for image segmentation," *Pattern Recognit.*, vol. 61, pp. 104–119, Jan. 2017.
[4] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: A survey," *Data Mining Knowl. Discover*, vol. 29, no. 3, pp. 626–688, May 2015.
[5] K. He, X. Cao, Y. Shi, D. Nie, Y. Gao, and D. Shen, "Pelvic organ segmentation using distinctive curve guided fully convolutional networks," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 585–595, Aug. 2018.
[6] X. Yuan, L. Wu, and Q. Peng, "An improved Otsu method using the weighted object variance for defect detection," *Appl. Surf. Sci.*, vol. 349, pp. 472–484, Sep. 2015.
[7] H. Zhang, J. Li, M. Wang, and H. Li, "Image edge detection based on fusion of wavelet transform and mathematical morphology," in *Proc. 11th Int. Conf. Comput. Sci. Educ. (ICCSE)*, Aug. 2016, pp. 981–984.
[8] M. G. Arvanitidou, M. Tok, A. Glantz, A. Krutz, and T. Sikora, "Motion-based object segmentation using hysteresis and bidirectional inter-frame change detection in sequences with moving camera," *Signal Process. Image Commun.*, vol. 28, no. 10, pp. 1420–1434, Nov. 2013.
[9] C. C. Lin and M. Wolf. "Detecting moving objects using a camera on a moving platform," in *Proc. Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 460–463.
[10] C. Yuan, I. Schwab, F. Recktenwald, and H. A. Mallot, "Detection of moving objects by statistical motion analysis," in *Proc. Int. Conf. Intell. Robots Syst.*, Oct. 2018. pp. 1722–1727.
[11] Y. Ye, C. Song, Y. Liu, and H. Tang, "Dynamic video object detection with single PTU camera," in *Proc. Conf. Visual Commun. Image Process.*, Nov. 2011, pp. 1–4.
[12] D. Li, J. S. Qian, and Y. L. Chai, "The object detecting in dangerous areas for coal mine underground," *J. China Coal Soc.*, vol. 36, no. 3, pp. 527–532, Mar. 2011.
[13] M. Huang, W. Yu, and D. Zhu. "An improved image segmentation algorithm based on the Otsu method," in *Proc. 13th ACIS Int. Conf. Softw. Eng., Artif. Intell., Netw. Parallel/Distrib. Comput.*, Aug. 2012, pp. 135–139.
[14] M. Hu, M. Li, and R. G. Wang, "Application of an improved Otsu algorithm in image segmentation," *J. Electron. Meas. Instrum.*, vol. 24, no. 5, pp. 443–449, 2010.
[15] Y. Bastanlar, A. Temizel, and Y. Yardımcı, "Improved PCA-SIFT matching for image pairs with scale difference," *Electron. Lett.*, vol. 46, no. 5, pp. 346–348, Mar. 2010.
[16] D. H. Jiang and G. Hua, "Research on image enhancement method based on adaptive immune genetic algorithm," *J. Comput. Theor. Nanosci.*, vol. 12, pp. 1–9, Jan. 2015.
[17] G. Hua and D. H. Jiang, "A new method of image denoising for underground coal mine based on the visual characteristics," *J. Appl. Math.*, vol. 2014, Apr. 2014, Art. no. 362716. doi: 10.1155/2014/362716.

JIANG DAI-HONG was born in Hunan, China, in 1969. She graduated from the China University of Mining and Technology, in 2015. She is currently a Professor and has a doctorate in communication and information systems.

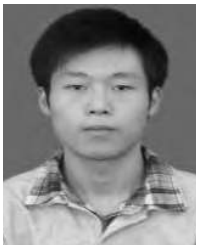
From 1995 to 2017, she was with the Xuzhou University of Technology. She has authored two books and more than ten articles, and she holds the patents for four inventions. She is a member of the China Computer Federation. Her main research interests include intelligent computation and database technology.

Dr. Jiang received the Second Prize for natural science and technology progress by universities in Jiangsu Province, the Second Prize at the Xuzhou Science and Technology Progress Awards, the Second Prize at the Science and Technology Progress Awards of the China Federation of Logistics and Purchasing, and the Xuzhou City Invention Patent Excellence Award.

DAI LEI was born in Jiangsu, China, in 1973. He graduated from Zhejiang University, in 2005. He received the master's degree in computer science and technology from the China University of Mining and Technology, where he is currently pursuing the Ph.D. degree and also an Associate Professor.

From 1996 to 2017, he was with the Xuzhou University of Technology. He has authored ten articles. He is a member of the China Computer Federation. His main research interests include graphic image processing and mobile computing.

LI DAN was born in Xuzhou, Jiangsu, China. He graduated from the China University of Mining and Technology. He is currently an Associate Professor with the Information and Electrical Engineering College, Xuzhou Institute of Technology. His research interest includes communication and information systems.



ZHANG SAN-YOU was born in 1989. He graduated from Soochow University, in 2013. He received the master's degree in computer application technology from the China University of Mining and Technology, where he is currently pursuing the Ph.D. degree. He is also with the Suzhou Wujiang District Public Security Bureau. His main research interest includes computer vision.