

Research Article

Moving Object Detection Technology of Line Dancing Based on Machine Vision

Liyun Liu 

College of Physical Education, Changsha University of Science & Technology, Changsha, Hunan 410114, China

Correspondence should be addressed to Liyun Liu; llyun@csust.edu.cn

Received 8 March 2021; Revised 6 April 2021; Accepted 20 April 2021; Published 27 April 2021

Academic Editor: Muhammad Babar

Copyright © 2021 Liyun Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, line dancing's moving object detection technology based on machine vision is studied to improve object detection. For this purpose, the improved frame difference for the background modeling technique is combined with the target detection algorithm. The moving target is extracted, and the postmorphological processing is carried out to make the target detection more accurate. Based on this, the tracking target is determined on the time axis of the moving target tracking stage, the position of the target in each frame is found, and the most similar target is found in each frame of the video sequence. The association relationship is established to determine a moving object template or feature. Through certain measurement criteria, the mean-shift algorithm is used to search the optimal candidate target in the image frame and carry out the corresponding matching to realize moving objects' tracking. This method can detect the moving targets of line dancing in various areas through the experimental analysis, which will not be affected by the position or distance, and always has a more accurate detection effect.

1. Introduction

The development of human action recognition can be summarized as movement, action, and activity. So far, human action is in the research stage of action recognition, that is, by extracting some features from the training data, using supervised or unsupervised training as a classification model, extracting features from the new data, and sending them to the model to get the classification results. Human action recognition has developed from human action recognition in the image to human action recognition in the video [1] using input data. Video action recognition is to process and analyze the original graphics or image sequence data collected by sensors and learn and understand the human action and behavior. Generally, based on motion detection and feature extraction, the human motion pattern is obtained through analysis. The mapping relationship between video content and action type description is established so that the computer can understand the video [2, 3].

In recent years, for video action recognition, the primary type of action videos includes daily human actions such as

walking, running, waving, and clapping; sports actions such as diving, skating, and riding; and life actions such as cutting vegetables and washing vegetables. In numerous research studies on human action recognition, few universities and institutions have studied dance actions. Dance is a way of emotional expression to the public through body movements, which is more complex in expression and has a wide variety of movements. There are many dance actions with their own characteristics [4–6]. Therefore, the research of dance action is still in the stage of dance action analysis. In most cases, it is through the posture analysis of the collected dance action, and then it is applied to the performance of animated characters by using animation processing software. Various universities and enterprises have continuously studied human action recognition due to applying action recognition technology in numerous fields. Human action recognition can help the public life to be more colorful and facilitate human beings' production and energy significantly [7].

Due to the disappearance of intangible cultural heritage (because of the lack of inheritors), human action's key information can be preserved through action recognition

technology to preserve intangible cultural heritage relevant actions. For example, dance actions have national characteristics in many areas of China. Critical information of dance actions is obtained and preserved through action recognition technology, which reduces the possibility of interruption of dance inheritance [8]. For dance action recognition, Pan et al. used zero crossings' angular velocity to realize video motion sequence segmentation [9]. Zheng et al. developed a more complex technology and used linear dynamic system output to represent low-level segments (primitives) [10]. Gao showed that motion segmentation is implied in state machine or motion graph representation [11]. However, the low-level segmentation cannot represent a complete action sequence. Talaat proposed three methods to long segment sequences into action subsequences [12] automatically. The first two methods were applied in practice video, which meant that the algorithm traversed all frames from the beginning to the end and created segmentation when breakpoints appear. The first algorithm allocated a cut when the intrinsic dimension of the local model suddenly increases. The second algorithm allocated the cut when the posture of the observed distribution changed. The third algorithm is a batch segmentation process and used the clustering method to form a simple action cluster for the continuous frame sequence of Gaussian mixture model belonging to different elements. Moemen et al. proposed to use the time structure of the vital pose set for modeling [13]. By comparison, the third method proposed by authors in [14] is more flexible (the speed change that allows gaps between adjacent postures is more robust). It is expected that the time structure would be learned automatically rather than designed manually [14].

Machine vision is an emerging and rapidly developing discipline. It is a broad area involving digital signal processing, graphics and image processing, human-computer interaction, pattern recognition, mathematical statistics, computer science, and other fields [15]. It has become an irresistible trend to use machine vision instead of traditional radar, infrared, and other means to capture external information through image and video technology [16]. Object tracking technology is an important research direction in machine vision, which has great practical value in real life. Moving object tracking technology is mainly used to process the video's continuous images, which contains more information about the moving object. Through the analysis of the scene spatial position change information between dynamic sequence images on the time axis, the moving target is extracted from the background, and then the target position is tracked. The target behavior can be analyzed further and understood by calculating the target motion information, such as the centroid position, motion mode, and trajectory route, to achieve a higher-level task. Moving object detection and tracking research is closely related to real life, bringing great convenience to people's lives. Using computer instead of human eyes and brain, people do not have to operate the computer all day to work easily and save workforce and material resources for the company. This technology has high practical value in intelligent transportation, important units, densely populated public areas, banks with high-

security requirements, monitoring and detection of museums, medicine, military, and other aspects. With the continuous improvement of science and technology and computer technology's continued development, video image acquisition and storage costs are significantly reduced, and the processing speed is constantly improved. These are the platform support for the development of video tracking technology. Video tracking technology is in line with the future trend of automation, information, and intelligence. On the other hand, the tracking technology based on machine vision can solve the complex work of processing many video information for a long time, saving workforce and material resources, and bringing convenience to people. Moving object tracking technology is closely related to people's lives, bringing comfort to people's lives. Moreover, target tracking technology is widely used in military science and technology, a broad field related to science and technology and people's livelihood [17]. With the advancement of science and technology, the research in this field will be more mature and perfect. Based on machine vision, this paper studies the moving object detection of line dancing and judges the detection and tracking accuracy and effect of this method.

The rest of this paper is organized as follows. In Section 2, we discuss our proposed moving object detection of line dancing using machine vision. In Section 3, we comprehensively evaluate our proposed work. Finally, the paper is concluded and future research directions are provided in Section 4.

2. Moving Object Detection Method of Line Dancing Based on Machine Vision

The moving object detection of line dancing based on machine vision has two integral parts: background subtraction method and adaptive tracking of moving target. First we discuss the background subtraction method for moving objection detection in line dancing in Section 2.1 followed by adaptive tracking of moving target based on mean-shift in Section 2.2.

2.1. Background Subtraction Method for Moving Object Detection in Line Dancing. After understanding the basic methods of moving object detection, analyzing the applicable scene and the advantages and disadvantages of each method, this paper decided to use the moving object detection method based on background subtraction. Various methods of background modeling are deeply studied, and the advantages and disadvantages of various algorithms are compared. Considering the real-time and accuracy of the system, the improved frame difference background modeling method is adopted.

2.1.1. Improved Frame Difference Background Modeling Method. In this paper, based on the traditional frame difference background modeling, a new initialization method is proposed and its feasibility is proved by experiments. The

steps of improved frame difference background modeling are as follows:

Get Initial Image. This step involves the following methods to get the initial image.

Choosing the Average Method. The principle of choosing the average method is proposed based on the original average method. Two adjacent images are selected. The region with variable pixel value is the moving target region, and the region with stable pixel value is the background region. The traditional adjacent frame difference method is used to distinguish the background region and the foreground region through threshold selection. The pixel of the moving object is set to 0, and the pixel of the background part remains unchanged.

The selection of threshold T can be selected through experiments. Generally, the change of background pixels is small, but it is not completely unchanged. The threshold T can be obtained through experiments according to the actual situation. In order to get a complete background image, we can select N frames in a video and use the adjacent frame difference method to compare the two images before and after to get a continuous background sequence. The process of selecting the average method is as follows:

- (1) The background sequence is obtained by frame difference method. The target region is set to 0, and the background region remains unchanged.

$$b_i(x, y) = \begin{cases} 0, & \text{if } |f_i - f_{i-1}| > T, \\ f_i(x, y), & \text{if } |f_i - f_{i-1}| \leq T. \end{cases} \quad (1)$$

- (2) The number $m(x, y)$ of times that the pixel in $B_i(x, y)$ in the background sequence image is not zero is obtained.
- (3) The average of the sum of background sequence images is as follows:

$$b(x, y) = \frac{1}{m(x, y)} \sum_{i=1}^{N-1} b_i(x, y). \quad (2)$$

- (4) Let iteration parameters be $i = 1$.
- (5) Frame difference method:

$$BW_i = \begin{cases} 1, & \text{if } \text{abs}(I_i - I_{i-1}) \geq T, \\ 0, & \text{if } \text{abs}(I_i - I_{i-1}) < T. \end{cases} \quad (3)$$

The selection of threshold T : T is determined manually, which can be obtained through practical experience. I_i and I_{i-1} are the i -th and $i - 1$ -th images, respectively, and BW_i is the difference image.

Background Update. After the initial background image is obtained by selecting the average value method, the background area and the moving target area of the current

image are distinguished according to the frame difference method. The background part of the moving target area does not need to be updated, and the background area of the previous image remains unchanged. The background region is updated by adding the weights of the background of the previous frame and the image of the current frame.

$$B_i(x, y) = \begin{cases} B_{i-1}(x, y), & \text{if } BW_i(x, y) = 1, \\ (1 - \alpha)B_i(x, y) + \alpha I_i, & \text{if } BW_i(x, y) = 0, \end{cases} \quad (4)$$

- (1) $i = i + 1$, return to step (3), and perform iterative operation.

The average method is selected to segment the moving object, and then the background area is averaged, which reduces the influence of the moving object on the background, and the extracted background effect is more accurate. A few prior video sequences are needed.

When the initial background image is obtained, the appropriate number of frames should be selected for modeling. If the number of selected frames is too small, holes will appear in the results obtained by the average method [18]. If it chooses too many frames, it will not only waste resources but also increase the burden of calculation. Moreover, if the time is too long, factors such as illumination and environmental changes will affect the accuracy of the results. Generally, it can be selected according to the speed of the object and the size of the target, so that all areas in the video image can display the background [19].

2.1.2. Moving Target Recognition. The principle of moving object recognition is that the background image obtained by the improved frame difference background modeling method is combined with the background difference method to obtain the moving object and realize the detection of the moving object. The binary image of the moving target is processed by morphological operation to remove the small target irrelevant to the moving target, fill the holes in the target, and connect the broken part of the target after threshold segmentation to make the image more complete. The main steps of moving target recognition are as follows:

- (1) The improved frame difference background modeling method obtains the background image $B_i(x, y)$.
- (2) Background difference: the difference operation is performed between the current frame image $f_i(x, y)$ and the background image $B_i(x, y)$.

$$D_i(x, y) = |f_i(x, y) - B_i(x, y)|. \quad (5)$$

- (3) Image binarization: the experiment obtains the appropriate threshold T , set "1" if it is greater than the threshold, and set "0" if it is less than the threshold.

$$R_i(x, y) = \begin{cases} 1, & D_i(x, y) \geq T, \\ 0, & D_i(x, y) < T. \end{cases} \quad (6)$$

- (4) For the binary image, the irrelevant pixels are removed by morphological operation to make the target more complete. Flow chart of moving target recognition is shown in Figure 1.

2.2. Adaptive Tracking of Moving Target Based on Mean Shift. After obtaining the moving object image, in order to realize the moving object detection and tracking of line dancing, the mean-shift algorithm is used to carry out the follow-up research. In recent years, mean-shift algorithm is widely used in target tracking. Some scholars use the Bhattacharyya coefficient as the similarity measure between target model and candidate target and use mean-shift algorithm to search the optimal candidate target, to obtain good tracking effect. Target tracking based on mean shift is a template-based tracking method [20].

2.2.1. Basic Principle of Mean Shift. In the field of machine vision, data analysis is usually carried out in multidimensional space. To estimate the nonparametric density of multidimensional sampled data sets, it needs to know the kernel function used in multidimensional space, that is, the multivariate kernel function.

In d -dimensional space, n points $x_i \in R^d$, $i = 1$ are given; Λ , n and a positive definite $d * d$ bandwidth matrix H , the multivariable kernel density of kernel $K(x)$ are estimated as

$$\hat{f}_{H,K}(x) = \frac{1}{n|H|^{(1/2)}} \sum_{i=1}^n K(H^{-(1/2)}(x - x_i)). \quad (7)$$

Generally, the bandwidth matrix H has two forms: one is diagonal matrix, that is, $H = \text{diag}[h_1^2 \Lambda h_d^2]$; the other is proportional unit matrix, that is, $H = h^2 I$. In this case, (7) can be written as follows:

$$\hat{f}_{h,K}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right). \quad (8)$$

If $K(x)$ is centrosymmetric, $K(x) = C_{k,d} k(\|x\|^2)$ and $K(x)$ is defined in the interval of $x \geq 0$; then $K(x)$ is called the contour function of $K(x)$ and $C_{k,d}$ is the normalized constant coefficient.

After introducing the contour function, the kernel density estimation can be rewritten as follows:

$$\hat{f}_{h,K}(x) = \frac{C_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{x - x_i}{h}\right\|^2\right). \quad (9)$$

This expression is commonly used in mean-shift algorithm to calculate the probability density of eigenvalues. If it wants to know the location of the highest density data in the data set, it can estimate the standard density gradient. Let $g(x) = -k'(x)$, and calculate the error covariance gradient. After simplification, it can get the mean-shift vector $m_{hG}(x)$.

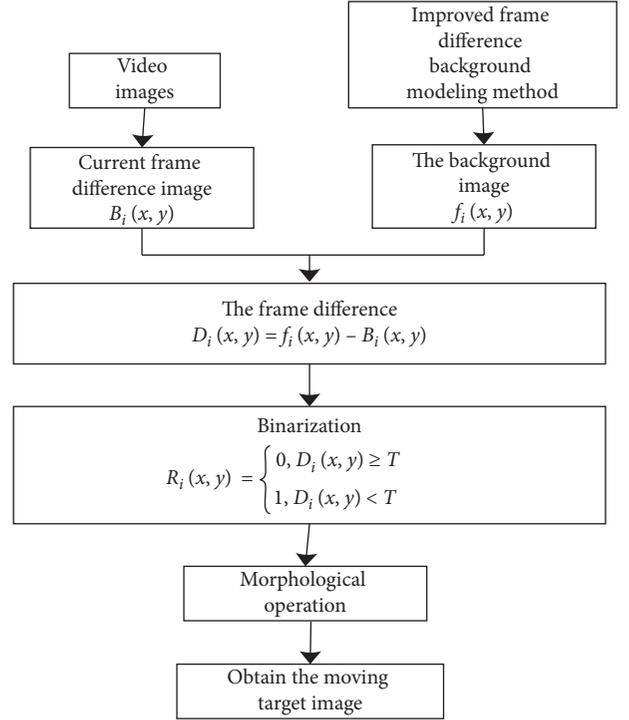


FIGURE 1: Flow chart of moving target identification.

$$m_{hG}(x) = \frac{\sum_{i=1}^n x_i g\left(\|x - x_i/h\|^2\right)}{\sum_{i=1}^n g\left(\|x - x_i/h\|^2\right)}. \quad (10)$$

2.2.2. Practical Application of Mean Shift in Moving Target Tracking. In this section, we discuss the practical application of mean shift in moving target tracking. We mainly focus on the description of target model in this section.

Because the histogram of the target image records the probability of color appearance, which is not affected by the shape change of the target, mean shift selects the color histogram as the feature description of the target. Suppose the center of the target region is y_0 and the pixel position of the target in the image is represented by $\{x_i\}_{i=1 \wedge \Lambda n}$. If the target model is a histogram of m values, the normalized color distribution can be expressed as

$$\hat{q}_u = C \sum_{i=1}^n \delta[b(x_i) - u], \quad (11)$$

where function $b: R^2 \rightarrow \{1 \wedge \Lambda M\}$ is the image of the pixel at x_i to the color index, $\delta(x)$ is the delta function, and C is the normalization coefficient.

According to the above frame difference method, the background area and the moving target area of the current image are distinguished, so that the pixels near the center of the target model are more reliable than other pixels. Therefore, different weights can be given to the pixels at different positions of the target, the pixels near the center are given a large weight, and the pixels far away from the center are given a small weight. By introducing kernel estimation

into the objective density function, the probability density estimation of the objective model can be expressed as

$$\hat{q}_u = C \sum_{i=1}^n k\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right) \delta[b(x_i) - u], \quad (12)$$

where $k(x)$ is the contour function of the kernel function, and $\|y_0 - x_i/h\|$ in function $k(x)$ is used to eliminate the influence of different sizes of targets. Using condition of $\sum_{u=1}^m \hat{p}_u = 1$, it can get the following results:

$$C = \frac{1}{\sum_{i=1}^n \left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)}. \quad (13)$$

In subsequent frames, the region that may contain the moving target is called the candidate region, and its center coordinate is y , which is also the center coordinate of the kernel function. The pixels in the region are represented by $\{x_i\}_{i=1}^{n_k}$. The description of the candidate region is called target candidate model. The probability density of the eigenvalue $u = 1/\Lambda m$ of the candidate model is as follows:

$$\sum_{u=1}^m \hat{p}_u(y) = C_h \sum_{i=1}^{n_k} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta[b(x_i) - u], \quad (14)$$

where $C_h = (1/\sum_{i=1}^{n_k} k(\|y - x_i/h\|^2))$ is the normalized constant coefficient.

The similarity function describes the degree of similarity between the target model and the candidate target, and the probability distributions of the two models are exactly the same in ideal cases. In the mean-shift algorithm, Bhattacharyya coefficient is selected as the similarity function. The Bhattacharyya coefficient is defined as follows:

$$\hat{p}(y) = p(\hat{p}(y), \hat{q}) = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u}. \quad (15)$$

The value of P is between 0 and 1. The larger the value of $\hat{p}(y)$ is, the more similar the two models are. In the current frame, the candidate model is calculated by calculating different candidate regions, so that the candidate region with the largest $\hat{p}(y)$ is the position of the target in the current frame.

2.2.3. Implementation Process of Moving Object Detection and Tracking. The most likely position of the target in the current frame is the candidate region that makes $\hat{p}(y)$ maximum. In order to maximize $\hat{p}(y)$, the position y_0 of the target center of the previous frame in the current frame is the starting position, and the optimal matching position is found, whose center is y . Firstly, the target candidate model $\hat{p}(y_0)$ is calculated, and the Taylor expansion of (15) is performed at $\hat{p}(y_0)$. The Bhattacharyya coefficient can be approximately expressed as

$$\rho(\hat{p}(y), \hat{q}) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(y_0) \hat{q}_u} + \frac{C_h}{2} \sum_{i=1}^{n_k} \omega_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right), \quad (16)$$

where

$$\omega_i = \sum_{u=1}^m \delta[b(x_i - u)] \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}}. \quad (17)$$

It can be seen from (16) that only the second term changes with y , so that

$$f_{h,k} = \frac{C_h}{2} \sum_{i=1}^{n_k} \omega_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right). \quad (18)$$

To maximize $\rho(\hat{p}(y), \hat{q})$ is equivalent to maximizing $f_{h,k}$.

By analyzing (17), it can see that this expression is similar to the kernel density estimation function, but with one more weight ω_i , it can make $f_{h,k}$ take the maximum value by means shift iteration. The specific steps are as follows:

- (1) Suppose the distribution of the target model is $\{\hat{q}_u\}_{u=1}^{\Lambda m}$ and the estimated position of the target is \hat{y}_0 .
- (2) Initialize the target position of the current frame with \hat{y}_0 , calculate the distribution $\{\hat{p}_u(\hat{y}_0)\}_{u=1}^{\Lambda m}$, and estimate the Bhattacharyya coefficient: $\rho[\hat{p}(y_0), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_0) \hat{q}_u}$.
- (3) Calculate the weight according to (17).
- (4) Calculate the new position of the target according to the mean-shift vector: $\hat{y}_1 = \sum_{i=1}^{n_h} x_i \omega_i g(\|\hat{y}_0 - \hat{x}_i/h\|^2) / \sum_{i=1}^{n_h} \omega_i g(\|\hat{y}_0 - \hat{x}_i/h\|^2)$.
- (5) Update $\{\hat{p}_u(\hat{y}_0)\}_{u=1}^{\Lambda m}$ and estimate $\rho[\hat{p}(y_1), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_1) \hat{q}_u}$.
- (6) When $\rho[\hat{p}(y_1), \hat{q}] < \rho[\hat{p}(y_0), \hat{q}]$, $\hat{y}_1 \leftarrow (1/2)(\hat{y}_0 + \hat{y}_1)$.
- (7) If $\|\hat{y}_1 - \hat{y}_0\| < \varepsilon$, end; otherwise $\hat{y}_0 \leftarrow \hat{y}_1$, go to step 1.

The process of target tracking is to get the maximum Bhattacharyya coefficient through the gradient descent search of mean-shift vector, so as to find the best matching position of the target in the next frame.

3. Experimental Results and Discussion

Taking a dance troupe in a certain city as the research object, 10 dance performers from the dance troupe are invited to cooperate with the experiment. They are asked to carry out dance performance and use cameras to collect images of their dance actions.

The background effect obtained by the proposed method is compared with that obtained by traditional unimproved frame difference background modeling. The results are shown in Figure 2.

Figure 2(a) is the background of traditional frame difference background modeling method, and Figure 2(b) is the image of improved frame difference background modeling. It can be seen that the initial image of the traditional method contains the target information, and the shadow of the object motion is obvious when modeling the background, so the modeling is not accurate. The improved modeling method

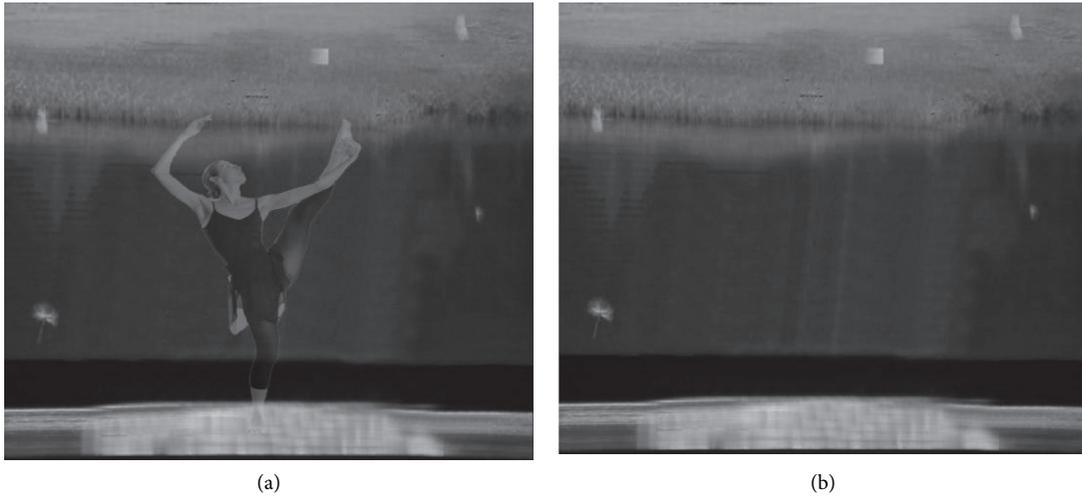


FIGURE 2: Background modeling results. (a) Traditional methods of background modeling. (b) The method of background modeling in this paper.

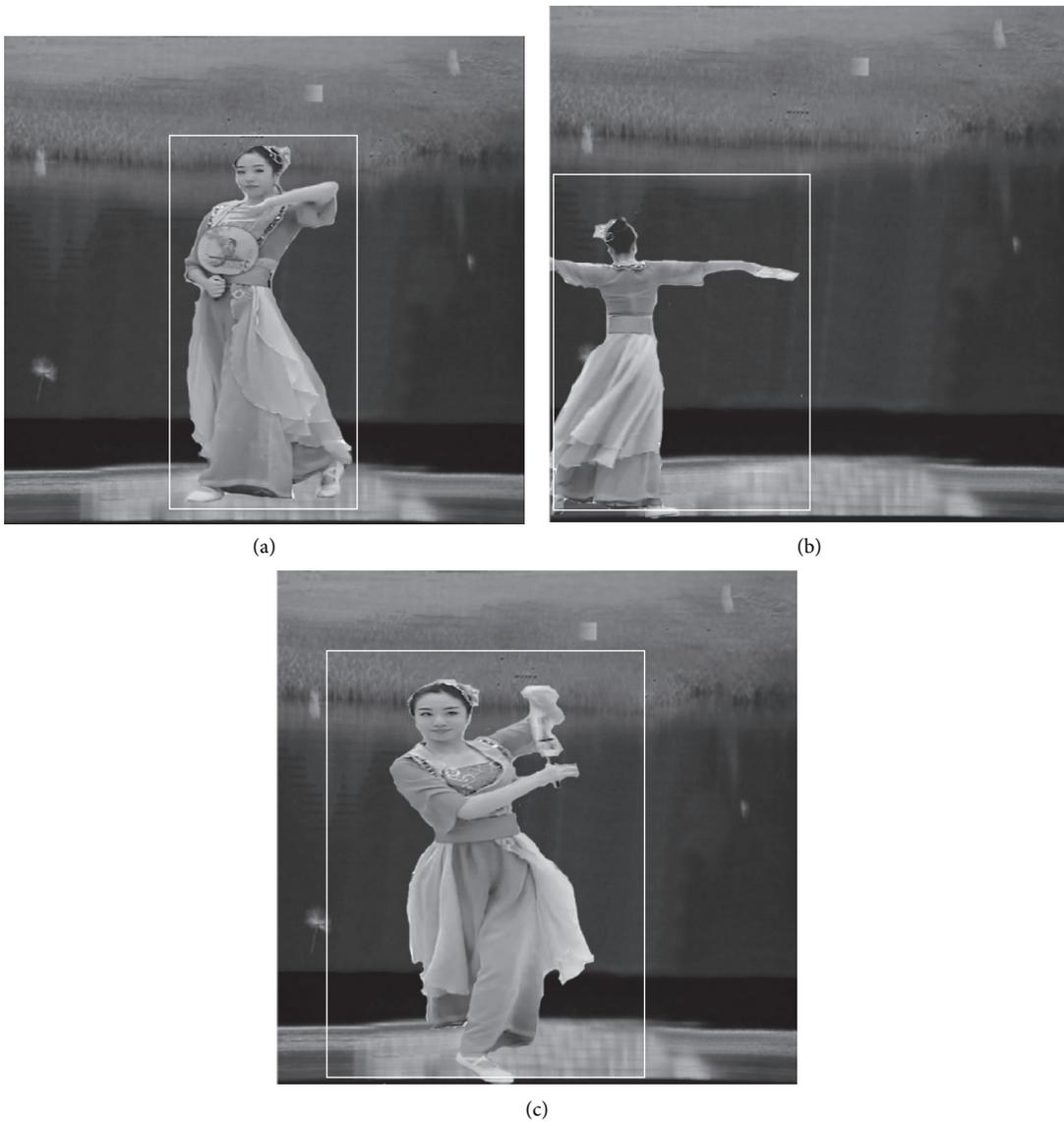


FIGURE 3: Detection effect of moving target in line dancing under simple background. (a) Starting test. (b) The dancer is on the edge of the picture. (c) The dancer approaches the camera.

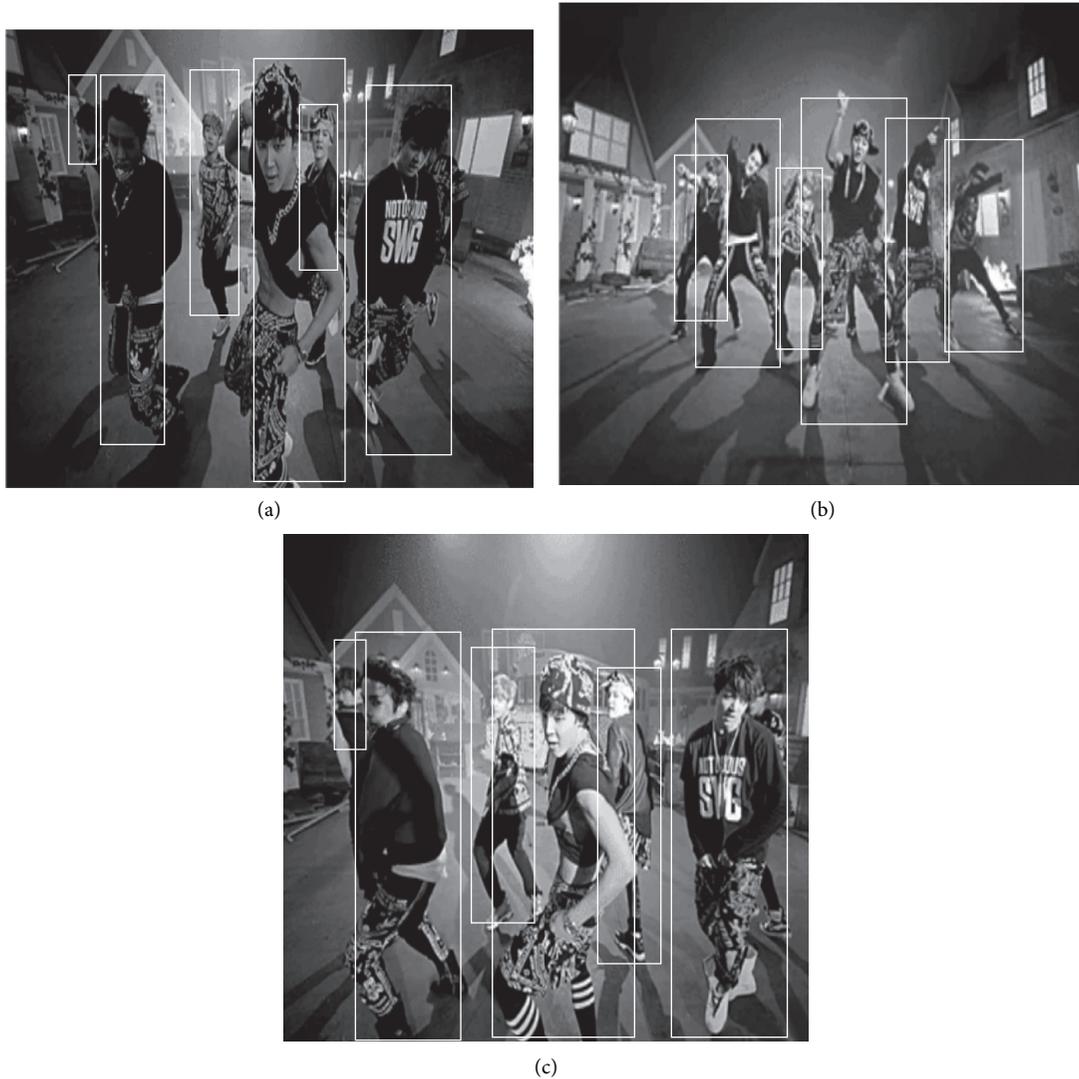


FIGURE 4: Detection results of moving objects in line dancing under complex background. (a) Camera lens nearly. (b) Camera pulls away. (c) Actor displacement.

reduces the influence of motion information, and the background effect is ideal.

A simple stage background is selected; the dancers are invited to make different dance actions and are required to be in the normal position, the edge of the screen, and the position close to the camera, respectively, so as to verify the effect of moving object detection and tracking of line dancing by the proposed method. The test results are shown in Figure 3.

It can be seen from Figure 3 that when the line dancing target moves, the proposed method can predict the centroid position of the target and search and match the target in a window neighborhood of the prediction point in the next frame. It can be seen from Figure 3(a) that the method in this paper can accurately locate the object of line dancing. As can be seen from Figure 3(b), when the target object is at the edge of the field of view of the camera, a part of the target object can still be located. As can be seen from Figure 3(c), when the moving object is

close to the camera and the size changes, the detection and tracking window will also change.

Under the influence of complex background, the moving target detection results of line dancing select the complex stage background with bright color. The webcam is fixed, and the proposed method is used to detect and track the line dancing movement of each dancer. The results are shown in Figure 4.

It can be seen from Figure 3 that the proposed method can detect the moving object of line dancing in complex background and obtain more accurate detection and tracking results.

The experimental results show that in the case of obstacles, the detection and tracking of moving objects in line dancing are improved. The experimental results are shown in Figure 5.

It can be seen from Figure 5 that when the detected object is partially occluded, even if only a few parts are not occluded, moving object detection of line dancing can still be

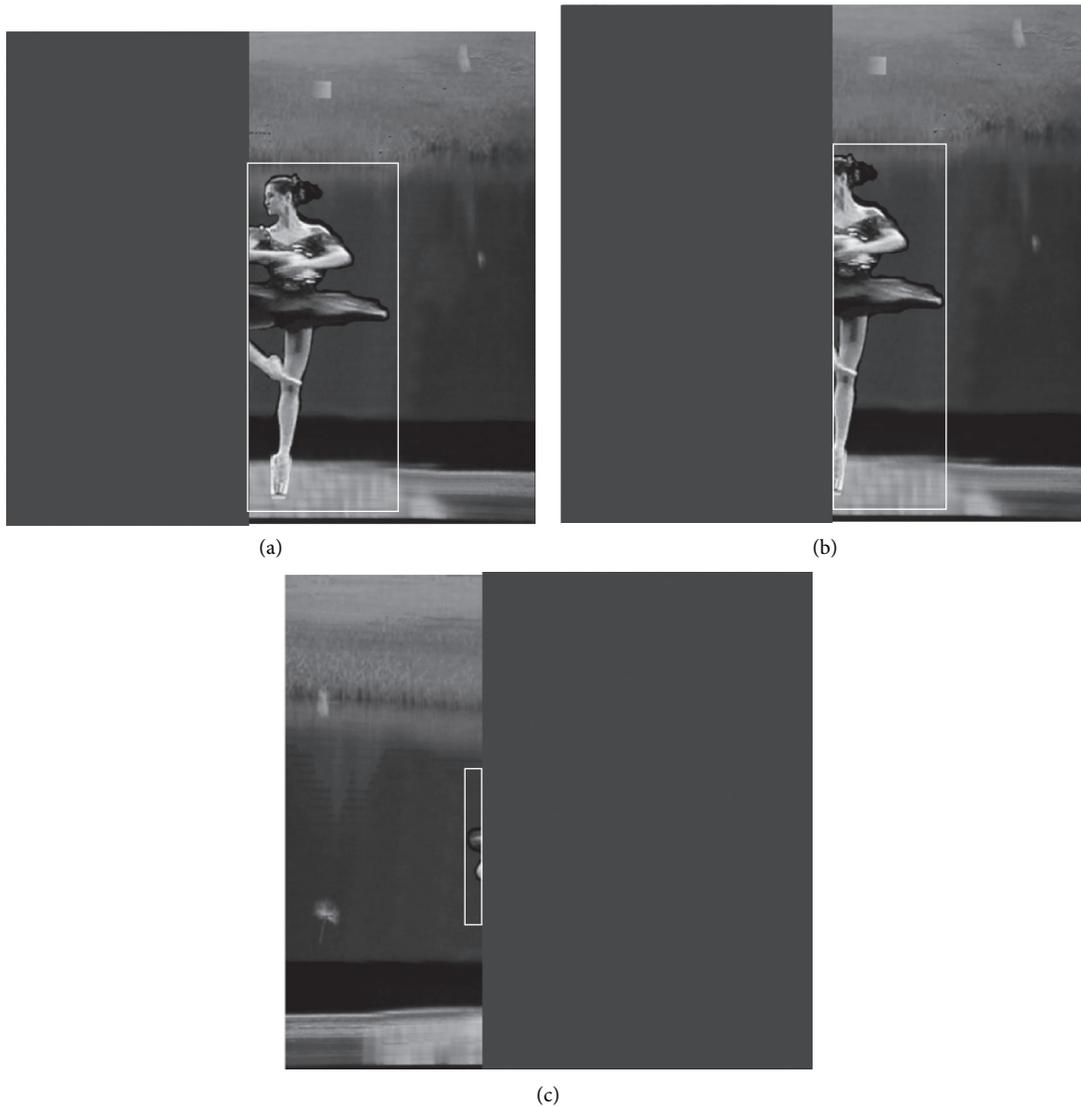


FIGURE 5: Detection results in the case of obstacle occlusion. (a) 1/3 Keep out. (b) 2/3 Keep out. (c) Remove the cover.

TABLE 1: Recognition rate of different human body regions (%).

Detection method	The upper part of the body	The lower part of the body	The upper part of the body + the lower part of the body	Hold body
RGB methods	72.31	71.58	84.54	79.15
Flow methods	85.05	84.24	90.66	85.46
This paper's methods	96.99	96.42	98.51	97.34

realized. When the detected object is completely occluded, the detection and tracking frame will be reduced to a small point. Once the target to be detected is exposed, an edge will be detected immediately, which shows that the detection efficiency of the proposed method is high.

In this paper, the recognition rate of RGB recognition method, flow recognition method, and the proposed method are used to compare the different parts of the upper body, lower body, and the whole body. From the recognition rate obtained by different methods, it can be seen that the

recognition rate of the upper body or the lower body is relatively close, but the recognition rate of the lower body is slightly higher. The recognition rate of the combination of human body regions is the highest. The recognition rate of the direct recognition of human body regions is lower than that of the combination of human body regions. The reason is that the number of upper body or lower body movements is close, and the lower body movements are more. In the application of the whole-body region of the human, when the movement is only the upper body region or the lower

body region, the two regions will affect each other, and the dance action naming often uses the lower body and upper body action naming, respectively, and the separate training can improve the accuracy of recognition. The recognition effect of different features extracted from different human regions is shown in Table 1.

4. Conclusion

This paper mainly studies the moving object detection method of line dancing based on machine vision. We used two approaches to achieve this objective. An improved frame difference background modeling is used to attain moving object detection, and the mean-shift method is used to achieve moving object tracking. We took the real-life choreographer as an example for our experimental results to verify our approach and prove that both our approaches can detect the moving objects in any area of the camera lens and not receive the restriction of occlusion or distance camera. It can be seen that the method can obtain good detection results of moving objects of line dancing [19, 20].

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was supported by the Project of the Ideological and Political Research of University Course in Hunan Province: Research and Practice of the Ideological and Political Teaching of the Course “Line Dance” and the study result of the Project of the Hunan Educational Science Plan: Study on the Optimizing of Physical Education in University under the Impact of Epidemic Situation (XJK20CGD012).

References

- [1] W. Huang and H. Xu, “Development of six-dof welding robot with machine vision,” *Modern Physics Letters B*, vol. 32, no. 5, Article ID 1840079, 2018.
- [2] X. Pan, Z. Liu, and G. Zhang, “Line structured-light vision sensor calibration based on multi-tooth free-moving target and its application in railway fields,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–10, 2020.
- [3] C. Guo, J. Qian, and D. Reynaerts, “Deterministic removal strategy for machine vision assisted scanning micro electrochemical flow cell,” *Journal of Manufacturing Processes*, vol. 34, no. 08, pp. 167–178, 2018.
- [4] P. Trsli, A. Weir, J. Riordan, E. Omerdic, and G. Dooly, “Vision-based localization system suited to resident underwater vehicles,” *Sensors*, vol. 20, no. 2, p. 529, 2020.
- [5] M. F. Seyed, M. M. Reza, and N. Mahdi, “Flying small target detection in ir images based on adaptive toggle operator,” *IET Computer Vision*, vol. 12, no. 4, pp. 527–534, 2018.
- [6] P. Hai-Hong, V. Ngoc-Son, N. Vu-Lam, and Q. Mathias, “Action recognition based on motion of oriented magnitude patterns and feature selection,” *IET Computer Vision*, vol. 12, no. 5, pp. 735–743, 2018.
- [7] B. Zhong, H. Huang, and E. Lobaton, “Reliable vision-based grasping target recognition for upper limb prostheses,” *IEEE Transactions on Cybernetics*, vol. 99, pp. 1–13, 2020.
- [8] Y. T. Chang, L. F. Sun, J. X. Pu, and H. F. Yu, “Air target identification using multiple sensors based on modified evidence support,” *Computer Simulation*, vol. 37, no. 07, pp. 394–398, 2020.
- [9] M. Pan, A. Liu, Y. Yu et al., “Radar hrrp target recognition model based on a stacked cnn-bi-rnn with attention mechanism,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 99, pp. 1–14, 2021.
- [10] J. Zheng, H. Fan, Q. Zhang, G. Yin, and Z. Li, “Magnetic anomaly target recognition based on svd and svms,” *IEEE Transactions on Magnetics*, vol. 55, no. 9, pp. 1–8, 2019.
- [11] X. Gao, “Design and implementation of marine automatic target recognition system based on visible remote sensing images,” *Journal of Coastal Research*, vol. 115, p. 277, 2020.
- [12] M. Talaat, I. Arafa, and H. M. B. Metwally, “Advanced automation system for charging electric vehicles based on machine vision and finite element method,” *IET Electric Power Applications*, vol. 14, no. 13, pp. 2616–2623, 2020.
- [13] M. Y. Moemen, H. Elghamrawy, S. N. Givigi, and A. M. Nouredin, “3-d reconstruction and measurement system based on multi-mobile robot machine vision,” *IEEE Transactions on Instrumentation and Measurement*, vol. 99, p. 1, 2020.
- [14] A. I. Khalyasmaa, M. D. Senyuk, and S. A. Eroshenko, “High-voltage circuit breakers technical state patterns recognition based on machine learning methods,” *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1747–1756, 2019.
- [15] S. Abouzahir, M. Sadik, and E. Sabir, “Bag-of-visual-words-augmented histogram of oriented gradients for efficient weed detection,” *Biosystems Engineering*, vol. 202, no. 3, pp. 179–194, 2021.
- [16] J. Zhang, J. Zhu, H. Deng, Z. Chai, M. Ma, and X. Zhong, “Multi-camera calibration method based on a multi-plane stereo target,” *Applied Optics*, vol. 58, no. 34, p. 9353, 2019.
- [17] Z. Gao, Y. Gu, Y. Lv, Z. Xu, and Q. Wu, “A monocular vision measurement system of three-degree-of-freedom air-bearing test-bed based on fccsp,” *Photonic Sensors*, vol. 8, no. 2, pp. 119–126, 2018.
- [18] G. Maier, F. Pfaff, M. Wagner et al., “Real-time multitarget tracking for sensor-based sorting,” *Journal of Real-Time Image Processing*, vol. 16, no. 6, pp. 2261–2272, 2019.
- [19] Y. Wang, H. Liu, M. Guo, X. Shen, B. Han, and Y. Zhou, “Image recognition model based on deep learning for remaining oil recognition from visualization experiment,” *Fuel*, vol. 291, no. 3, Article ID 120216, 2021.
- [20] Z. Wang, G. Yu, B. Zhou, P. Wang, and X. Wu, “A train positioning method based-on vision and millimeter-wave radar data fusion,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–11, 2021.