Research Article

Target Recognition, Localization, and Motion Control of Soccer Robot

Changsheng Zhu and Min Gong

College of General Education, Nanjing University, Nanjing, Guangxi 530200, China

Correspondence should be addressed to Changsheng Zhu; 19401012@masu.edu.cn

Received 29 May 2022; Revised 26 June 2022; Accepted 6 July 2022; Published 15 July 2022

Academic Editor: Haibin Lv

Copyright © 2022 Changsheng Zhu and Min Gong. 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.

The realization of target tracking includes many technologies, among which the most critical and core element is the target tracking algorithm. Based on this, this paper first uses an omnidirectional vision image enhancement algorithm to identify the target and then designs an adaptive target recognition algorithm based on SVM to identify the actual target. The results show that, in the case of different illuminance, the fluctuation range of the four evaluation indexes after image enhancement is smaller than that of the original image, indicating that the brightness distribution of the image after enhancement is more uniform, which eliminates interference factors for the subsequent image recognition and tracking. The difference of recognition rate between the two algorithms is small when the illumination value is medium. When the illumination value is high or low, the recognition rate of the adaptive algorithm is much higher than that of the threshold method.

1. Introduction

The football robot vision module consists of two parts, the image acquisition part and the image processing part. The image acquisition is mainly completed by the camera. After the image is acquired, the most important thing is to extract the useful information contained in the image, which requires that further analysis and processing measures are taken for image information, which involves some algorithms for dedrying and feature extraction [1]. Therefore, whether useful information can be obtained depends largely on the performance of these algorithms. The performance evaluation indicators of image processing algorithms mainly include execution efficiency, robustness, and complexity. Background interference or chromatic aberration are often encountered in the game, which requires image processing algorithms to have strong anti-interference to these uncertain factors, so as to adapt to various complex environments [2, 3].

The research field of the football robot vision module mainly involves technologies such as target recognition and target tracking. Through the research and improvement of related technologies, on the one hand, the existing theoretical system can be improved, and on the other hand, the application scope of recognition and tracking technology can be expanded [4]. In recent years, artificial intelligence has become more and more popular, and one of the most important aspects is image recognition and tracking, which plays an important role in realizing the autonomous operation of machines. Every year, a large number of scholars pour into this field [5]. With the continuous updating and improvement of algorithms, the application scope of this technology has also extended from the initial field of machine vision to many other fields such as industrial inspection, geological exploration, autonomous driving, and remote sensing images. Therefore, the research on image recognition and tracking technology has also played a positive role in promoting the progress of other research directions. Figure 1 shows the improvement of target positioning technology and tracking algorithm for humanoid soccer robots [6].

2. Literature Review

RoboCup has established the category of the mid-sized group since its inception, and the RoboCup Medium-Sized
League (MSL) has been held since 1997. The field of competition is set on an indoor football field with a size of 18 m * 12 m, which is appropriately reduced relative to the actual football field. Two teams, each team is allowed to have a maximum of six football robots, using orange balls to play. Human intervention is not allowed during the game, unless there is an unexpected situation such as a robot foul or a robot failure. For the shape of the robot, it is stipulated that the maximum diameter cannot exceed 50 cm, and each team can design hardware and software completely autonomously [7].

In the RoboCup competition, the competition of the medium group is more watchable and the competition is more intense than the competition of other groups. In these years of competition, there have also been many strong teams. In the first few years of the competition, the German team and the Japanese team performed relatively well. Later, Portugal and the Netherlands began to catch up slowly. In recent years, in the competition, the goal team is often the champion and runner-up. Some colleges and universities also participated. Although they started late, there were also teams that performed well. These teams have their own strengths [8].

At the beginning of the robot football competition, because the technical research of the participating teams was still in its infancy, in order to reduce the difficulty of the competition and allow the competition to proceed normally, the organizers set the competition rules relatively loosely. Compared with the past, there have been certain breakthroughs, and in order to make the robot adapt to a more complex environment, the organizer has revised the robot’s competition rules, making the rules more and more similar to the rules of human competition. Therefore, the current competition environment is more complex; lighting changes, visual blind spots, and visual blur caused by movement are all key factors that affect the accuracy of the visual system’s recognition of objects. Therefore, the main research purpose of the vision module of the football robot is to make the vision module recognize the target accurately in most cases, not just to complete the recognition task by precalibrating the color. In response to this robustness problem, domestic and foreign researchers have done a lot of research on it. Particularly in recent years, many new algorithms have been proposed in the field of machine vision, and the target features used in target recognition have gradually diversified, such as edge features, linear features, and HOG features, rather than just a single color feature to identify objects. Among the adopted recognition methods, the two most frequently used methods are the template matching method and the classifier method. Shruti and Deka proposed a target recognition algorithm based on optical flow field, which assumes that the gray gradient is basically constant or the brightness is constant. At this stage, the technology based on the first-order spatiotemporal gradient is still the mainstream of the optical flow recognition algorithm, and many new algorithms have also emerged [9]. Husari and Seshadrinath developed an optical flow field-based robot, which imitates the visual mechanism of bees and can walk in corridors independently. However, the calculation algorithm of the optical flow field is complex and time-consuming, and when the distinction between the background and the target is small, or the image quality is not particularly high, the accuracy of the algorithm for the detection and recognition of the target will be greatly reduced [10]. The four-wheeled mobile robot introduced by Cai and Zheng is also a nonomnidirectional robot. Common omnidirectional mobile robots include one-wheel, three-wheel, and four-wheel. It can fully reflect the maneuverability of omnidirectional mobile robots in occasions where space is not very abundant. Omnidirectional robots generally use omnidirectional wheels. Compared with ordinary wheels, omnidirectional wheels can slide sideways, which determines that omnidirectional mobile robots have good maneuverability and can move at any speed and any angular velocity on the plane [11]. Xie et al. showed how to achieve the calibration of a parabolic reflection panorama system from the imaging of three lines in space or two sets of parallel lines in an omnidirectional view [12]. Wu et al. developed a new panoramic vision system, which improves the imaging shortcomings of the traditional omnidirectional vision system. It combines hyperbolic mirrors, vertical proportional mirrors, and horizontal proportional mirrors. The combination of mirrors improves the problem of severe image distortion in traditional imaging systems [13]. Zeng et al. proposed a recognition method based on the use of real-time adjustment of calibrated colors. This method adjusts color information in real time according to the principle of color drift and combined it with the degree of illumination change, so as to achieve robustness to changes in light intensity. Sexual purpose [14]. Li et al. proposed a method...
of using the Gabor filter for target recognition of images based on the YUV color model. The two models used are the YUV color model fast and the Gabor filter model. Recognition accuracy [15]. Shen and Yan developed a classification-biased recognition method based on PCA and SIFT algorithms, which integrates traditional PCA and SIFT algorithms, thereby reducing the number of feature descriptors in the SIFT algorithm. And use the K-nearest neighbor classification algorithm for target feature matching, thereby improving the recognition accuracy, but the real-time performance needs to be improved [16]. Zhang et al. proposed an algorithm based on color lookup table and radial model matching, which is simple and fast and still has strong stability in complex environments [17].

The realization of target tracking includes many technologies, among which the most critical and core element is the target tracking algorithm. The algorithms that are often used at present are the frame difference method, mean shift algorithm, particle filter, multi-feature fusion, etc. [18, 19]. These algorithms are mainly based on one of the basic algorithms such as detection, matching, filtering, etc., or the fusion of several different algorithms, so as to achieve the effect of complementary advantages [20].

Based on this, this paper first uses the omnidirectional vision image enhancement algorithm to identify the target and then designs the SVM-based adaptive target recognition algorithm to identify the actual target.

3. Research Methods

3.1. Homomorphic Filter Image Enhancement Algorithm Based on Mallat Wavelet Transform. Homomorphic filtering is generally performed in the frequency domain, so the omnidirectional image must first be converted from the spatial domain to the frequency domain, but the spatial resolution of the transformed image cannot reach the expected effect, because there is no local feature of the image optimization, which makes the image easily blurred, and the contrast improvement in the target area is not obvious [21]. Wavelet transform is a signal analysis theory based on Fourier transform, but Fourier transform does not have the ability of local analysis. Compared with the Fourier transform, the wavelet can analyze images from space and frequency. Local transformation can effectively extract local information. It solves the problem of local information extraction that is difficult to solve by Fourier transform. There are many kinds of wavelets. For the purpose of simplifying the algorithm, the Mallat wavelet is selected for wavelet table transformation, and it is applied to the homomorphic filtering algorithm [22, 23].

The specific implementation steps of the filtering enhancement algorithm based on the Mallat wavelet transform are listed as follows:

1. Design the mask according to the characteristics of the omnidirectional image, and perform the bit AND operation on it with the initial image, so as to delete the noninformation area and reduce the size of the area that needs to be processed in the subsequent steps

2. Transform the trimmed image to the HSI color coordinate system, and list the luminance map I separately from it, let \( I = f(x, y) \)

3. Take the logarithm of the luminance function \( f(x, y) \) to get \( \ln f(x, y) \)

4. Perform Marat frequency domain decomposition on \( \ln f(x, y) \), the decomposition level is 3, and obtain the wavelet coefficients of the high-frequency part \( D^1, D^2, \) and \( D^3 (j = 1, 2, 3) \) and the low-frequency part \( A_3 \)

5. Filter the frequency domain part after wavelet decomposition by using a filter

6. Reconstructing the filtered frequency domain components

7. Take the antilog of the restored image obtained after reconstruction, thereby generating an enhanced image \( g(x, y) \)

The flowchart of the homomorphic filtering image enhancement algorithm based on the Mallat wavelet is shown in Figure 2.

3.2. Adaptive Target Recognition Algorithm. In this paper, the orange football is selected as the recognition target of the football robot, and the color is used as the main target feature to recognize the ball [24]. The most traditional method is the threshold method. This method calibrates the limited colors in the field and establishes a look-up table. According to the table, the target of the specified color is identified. The main features are simple and short. The disadvantage is that the adaptability is poor, and the color drifts to a certain extent due to uneven lighting in the game. In this case, the stability of the traditional threshold method cannot be guaranteed. Although image enhancement can partially improve the color changes caused by factors such as lighting changes and textures, such changes still exist. The support vector machine (SVM) algorithm has strong adaptability and learning ability, so this algorithm is selected as the recognition model in this paper [25].

SVM, also called the support vector machine algorithm, is a classifier algorithm developed from the generalized portrait algorithm. SVM belongs to the category of supervised learning and is a generalized linear classifier for binary classification of data. The classifier calculates the empirical risk according to the loss function and applies regularization constraints to the classification model to optimize the structural risk. Compared with similar learning algorithms, this method requires a smaller sample size and higher learning efficiency. It is good at dealing with classifier problems and has been widely used in practice.

Since the IV color space is constructed for target recognition in this paper, the two feature quantities constitute a two-dimensional plane. Since the color composition in the scene is relatively simple, the samples in the two-
dimensional plane have a high probability of being linearly separable. Therefore, the classifier is mainly designed on the basis of linear separability, while the computational complexity is small when linearly inseparable. The specific algorithm is as follows:

(1) Suppose the known training set is as

\[
T = \begin{bmatrix}
I_1 & V_1 & y_1 \\
I_2 & V_2 & y_2 \\
\vdots & \vdots & \vdots \\
I_l & V_l & y_l \\
\end{bmatrix},
\]

(1)

where \(x_i = [I_i, V_i]\) is the training sample, that is, the pixel in the enhanced image, \(y = [y_1, y_2, \ldots, y_l]^T\) is the sample label, the positive sample is set to 1, and the negative sample is set to -1.

(2) The hyperplane equation is \((w \cdot x) + b = 0\). In order for the classification surface to correctly classify all samples and have the largest classification interval, the minimum value of \(1/2\|w\|^2\) must be obtained under the constraint of the condition \(y_i([w \cdot x_i] + b) \geq 1\)\(i = 1, 2, \ldots, l\).

To this end, the following Lagrange function can be defined, where \(a_i > 0\), as in

\[
L(w, b, a) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{l} a_i y_i ([w \cdot x_i] + b) + \sum_{i=1}^{l} a_i, a_i \geq 0, i = 1, \ldots, n.
\]

(2)

Taking \(w\) and \(b\) in the above formula as separate independent variables to obtain partial derivatives, and simplifying the obtained formula and making it equal to 0, formula (3) is obtained:

\[
w = \sum_{i=1}^{l} a_i y_i x_i, \sum_{i=1}^{l} a_i y_i = 0.
\]

(3)

Then, put formula (3) into formula (2), eliminate and simplify \(w\) and \(b\) to obtain the dual problem, and solve the maximum value of the following formula as

\[
\max Q(a) = \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} a_i a_j y_i y_j ([x_i \cdot x_j])
\]

(4)

s.t. \(\sum_{i=1}^{l} a_i y_i = 0, a_i \geq 0, i = 1, \ldots, n.\)

(5)

Get the optimal solution \(a^* = [a_1^*, a_2^*, \ldots, a_l^*].\)

(3) Substitute \(a^*\) into Equation (3) to get \(a^*\) to select a positive component \(a_i^*\) of \(a^*\), and substitute into the following equation to obtain \(b^*\) as in

\[
b^* = -\frac{1}{2} \left( \max_{n_{y_i}=1} (w^* \cdot x_i) + \min_{n_{y_i}=1} (w^* \cdot x_i) \right).
\]

(6)

(4) \(w^*\) and \(b^*\) are obtained in formulas (5) and (6), respectively, from which the classification hyperplane function \((w^* \cdot x) + b^* = 0\) can be obtained, and thus, the decision function can be obtained as

\[
f(x) = \text{sgn} \left( (w^* \cdot x) + b^* \right) = \text{sgn} \left( \sum_{i=1}^{l} a_i^* y_i [x_i \cdot x] + b^* \right).
\]

(7)

Among them, \(x_i\) is the support vector, and \(x\) is the sample to be classified.

(5) Target recognition

Step 1. Obtain the sample points to be tested in real time, that is, all the pixels on the scan line with the image center as the starting point and the site boundary as the end point, and take the pixels corresponding to the scan line as the samples to be classified.

Step 2. Use the sample points obtained in real time as the input of the trained SVM classifier to perform classification operations. Classify according to the output result of the decision function, that is, formula (7). When \(f(x) = 1\), the pixel is classified as orange, and when \(f(x) = -1\), the pixel is classified as nonorange. That is, the final target recognition result is the sample point of \(f(x) = 1\). Figure 3 shows the flow chart of the target recognition algorithm in this paper.
4. Analysis of Results

4.1. Evaluation of Homomorphic Filter Image Enhancement Algorithm Based on Mallat Wavelet Transform

4.1.1. The Brightness Is Uniform.

\[
L = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} I(x, y). \tag{8}
\]

Among them, M and N represent the size of the circle and direction, respectively.

4.1.2. Standard Deviation. The standard deviation is used to judge the contrast of the image. This metric is positively correlated with contrast.

\[
S = \sqrt{\frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (I(x, y) - \bar{L})^2}. \tag{9}
\]

4.1.3. Entropy. Here, entropy is used to judge the richness of image details. The larger the value, the more details.

\[
E = -\sum_{i=0}^{255} P_i \log P_i, \tag{10}
\]

where \( P_i \) is the probability of the \( i \)th gray level.

4.1.4. Average Gradient. This indicator is used to evaluate local features such as contrast and texture. The larger the value, the higher the image definition.

\[
G = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} \sqrt{\left(\frac{\partial I(x, y)}{\partial x}\right)^2 + \left(\frac{\partial I(x, y)}{\partial y}\right)^2}. \tag{11}
\]

By comparing the evaluation indicators in Tables 1 and 2, it can be seen that the fluctuation range of the four evaluation indicators after image enhancement is smaller than that of the original image under different illumination conditions, indicating that the brightness of the image after enhancement processing is the distribution that is more uniform, and interference factors are eliminated for subsequent image recognition and tracking. When the illumination is the same, the standard deviation and the mean brightness of the enhanced image are larger than the same index of the image without enhancement processing, reflecting the increase of contrast; moreover, the entropy and average gradient index of the enhanced image have a certain degree of increase. It reflects that the image contains more local features and more perfect details, which makes the image more recognizable.

4.2. Evaluation of Adaptive Object Recognition Algorithms. After adjusting the parameters of the football robot, start it, and adjust the brightness of the indoor fluorescent lamps and the degree of opening and closing of the curtains to change the illuminance. Then, turn on the camera and collect 100 omnidirectional pictures at various positions in the field, and take 30 of them as the test set and 70 as the sample set. Set the number of scan lines \( N = 360 \) and \( M = 30 \). First, the image is enhanced and preprocessed using the Marat-based homomorphic filtering algorithm mentioned above. Then, perform color model conversion on the obtained enhanced image to obtain an IV color model. In the offline state, the biSCAN scan line algorithm is used to extract the sample vector \([I, V]\) in the picture and input it into the SVM for training. When recognizing the target in a real-time state, the classification sample vector \([I, V]\) is extracted from the image by the biSCAN scan line method, and the trained SVM model is used to classify the sample, so as to obtain the recognition result of the target ball, as shown in Table 3.

By comparing the algorithm in this paper with the threshold method, the results in Figure 4 are obtained.

It can be seen from Figure 4 that the difference between the recognition rates of the two recognition algorithms is small under the condition of moderate illumination. When the illumination value is high or low, the recognition rate of the algorithm is much higher than that of the threshold method.

Overall, the target recognition algorithm proposed in this paper has a higher recognition rate than the threshold method. And it can maintain high stability under different conditions. The threshold method is less adaptable and only
suitable for a specific situation. Therefore, the algorithm proposed in this paper is more adaptable and can meet the needs of target recognition.

5. Conclusion

Due to the particularity of imaging, the omnidirectional vision system has a large deformation of the obtained omnidirectional image, and these factors increase the difficulty of target recognition and tracking to a certain extent. However, most of the general target recognition and tracking algorithms are designed for images based on ordinary perspective imaging, so they cannot be directly applied to omnidirectional vision systems. Moreover, in the soccer robot game scene, the color is highly coded and the target is often in a nonstationary state. Therefore, considering the characteristics of omnidirectional vision and the game environment, this paper studies the recognition and tracking of soccer balls in this application scene. The main research results of this paper are as follows:

(1) A filtering algorithm based on Marat transform for image enhancement is proposed. This method combines the Mallat wavelet transform with the traditional homomorphic filtering to enhance the image. Then, use the Marat wavelet transform instead of the Fourier transform to transform it into the frequency domain and decompose it to obtain the wavelet coefficients of each high frequency and low frequency, and then, use the Butterworth type homomorphic filter function to filter each wavelet coefficient in the frequency domain. Then, the wavelet coefficients are reconstructed and inversely transformed to the spatial domain, and then, the transformed image is indexed to achieve the purpose of enhancement. According to the experimental results, it can be seen that the image enhancement method performs well in terms of robustness and adaptability and can meet the actual needs.

(2) The biSCAN algorithm and the SVM classifier are combined for target recognition. According to the intraclass scatter and interclass scatter matrix as the criterion, the traditional color space is improved, and an IV color model is designed for the football robot application scene. Based on this, the biSCAN scanning ray method is used to select the field. Then, the trained SVM model is used in an offline state to identify footballs under different light intensities and finally compared with the traditional threshold method. The experimental results show that the light adaptability and recognition rate of this method are greatly improved.

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 they have no conflicts of interest.

References


