Framework Design of Sports Image Analysis System Based on Player Behavior Information

Yu Wang, Zhu Mei, Tianbo Yu*

Shenyang Jianzhu University, Shenyang, China

*Correspondence Author

Abstract. Obtaining the data and information urgently needed by coaches and athletes through the processing of visual continuous images of sports is the most effective means and method in the current sport's technical analysis. The sports image photography framework is an important tool for transforming the two-dimensional coordinates of sports technology images into three-dimensional space coordinates. The article proposes a calibration algorithm suitable for the two-plane calibration frame. The calculation is simpler than the DLT algorithm, and it is easy to program. The experimental results show that the system has good tracking effect of moving image key frames, high comprehensiveness, low redundancy and high recall rate.

Keywords: Player information detection; Sports; Image analysis system; Calibration framework.

1. Introduction

Human body movements in sports are mostly complex three-dimensional spatial movements. To accurately and quantitatively describe the movement characteristics of the human body, three-dimensional photography and analysis must be carried out. The stereo camera measurement method is to use two or more cameras to simultaneously shoot the same research object from different angles, and then digitize the flat film or video tape taken to obtain the required spatial coordinates of the three-dimensional motion of the human body, and calculate the relevant Kinematic parameters [1]-[2]. For this purpose, this paper has carried out a theoretical redesign of the frame of the sports image analysis system in order to solve the difficulties of practical work most effectively.

2. Sports video background modeling based on inter-frame difference method and median filter method

2.1 Use the three-frame difference method to calculate the target's motion template

Assuming that the initial video frame containing the moving target in the video sequence is f_k (k=1 in this article), starting from the f_k frame, select continuous n frames of video images as the video sequence for constructing the background model:

$$f_k, f_{k+1}, f_{k+2}, ..., f_{k+n-2}, f_{k+n-1} \tag{1}$$

The above-mentioned original image is converted into a grayscale image, the noise in the image is removed by image smoothing, and then the motion template of the target is calculated, and then the position of the moving target in each frame of video image is estimated. Suppose f_{i-1}, f_i, f_{i+1} is the continuous three frames of video image, $d_{i,i+1}$ is the frame difference image of two adjacent frames of video image f_i, f_{i+1}, and $D_{i,i+1}$ is the motion template obtained by binarizing the difference image, then

$$d_{i,i+1} = \left| f_{i+1} - f_i \right|$$
$$D_{i,i+1} = \begin{cases} 1 & ; d_{i,i+1} \geq T \\ 0 & ; d_{i,i+1} < T \end{cases} \tag{2}$$

Where T is the threshold of binarization? In the same way, $D_{i-1,i}$ is calculated. Perform a logical AND operation on the calculated binary motion template $D_{i-1,i}, D_{i,i+1}$ to obtain a relatively accurate motion template M_i of the moving target in the video frame f_i:

$$M_i = D_{i-1,i} \cap D_{i,i+1} \tag{3}$$

In the same way, the above operations are performed on the n frame images in the video sequence, and then a template sequence $\{M_i\}_{i=k+1,k+2,...,k+n-2}$ containing $(n-2)$ motion templates is obtained. The start frame and the last frame of the video sequence have no corresponding motion templates.

If the target moves continuously in the original video sequence, no processing is required on the motion template sequence $\{M_i\}$. If the moving target is in a relatively static state with the background for a certain period of time, in the obtained frame difference image sequence, there must be a number of consecutive frames of frame difference images in which the value of most pixels is zero [3]. At this time, the corresponding target the motion template is also approximately zero. For such a situation, you only need to find the nearest motion template that contains the moving target, and then use this template to replace those motion templates that are close to zero, that is, the corresponding template when the target is in a static state.

2.2 Fill in the outline of the motion template and mark the position

Perform contour detection on motion template M_i, fill in the detected contour and mark the position. For pixels outside the contour, their position is marked as 0, and pixels located within the contour are marked as 1. Define an array $label(i,x,y)$ to record position mark information, then

$$\begin{cases} label(i,x,y) = 1; & in\ f_i(x,y) \\ label(i,x,y) = 0; & out\ f_i(x,y) \end{cases} \tag{4}$$

Among them, $f_i(x,y)$ is the pixel at any position (x,y) in the video frame f_i.

2.3 Background modeling of median filter method

For video sequence $f_{k+1}, f_{k+2}, ..., f_{k+n-2}$, look at the position mark of pixel $f_i(x,y)$ at the same coordinate (x,y) in each frame one by one. If $label(i,x,y) = 0$, the pixel $f_i(x,y)$ is considered to be in the background area, and the pixel is reserved. If $label(i,x,y) = 1$, the pixel $f_i(x,y)$ is considered to be located in the moving target area, and the pixel is discarded at this time. Take the average or median of the retained pixels as the background value at the (x,y) coordinate in the constructed background image [4]. However, it should be noted that when the average value is taken, the front sights that are misjudged as the background will have an adverse effect on the background modeling. Therefore, in general, the median value is more robust. This article uses the median value as the background value of the constructed background image.

3. Morphological image processing

3.1 Expansion and corrosion

Dilation is to merge the two sets by applying vector addition to the set elements. Dilation $\delta_B(X)$ is the set of the sum of all possible vectors of operands X and B. The expansion operation is expressed as:

$$\delta_B(X) = \left\{ p \in \varepsilon^2 : p = x + b, x \in X \text{and} b \in B \right\} \tag{5}$$

Among them, B is an expanded structural element, and X is a binary image. The result of the expansion operation is shown in Figure 1:

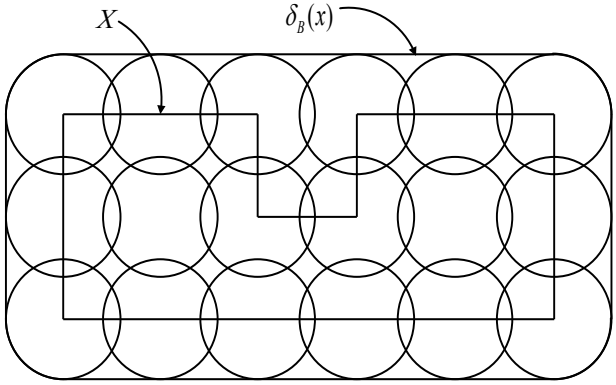


Figure 1. Expansion

It can be seen from Figure 1 that the expansion operation causes X to expand into a rectangle with rounded corners, and fills in the original groove. The main function of the expansion operation is to increase the number of pixels of the target and the size of the object to fill holes and cracks. Dilation has the effect of expanding the image [5]. If the size of the object is to remain basically unchanged, it needs to be used in conjunction with the corrosion operation.

Corrosion is a dual operation of expansion, which merges two sets by applying vector subtraction to set elements. The corrosion operation is expressed as:

$$\varepsilon_B(X) = \left\{ p \in \varepsilon^2 : p + b \in X, \text{For each} b \in \mathrm{B} \right\} \tag{6}$$

The above formula shows that the corrosion operation has been applied to each point p of the image, and the result of the corrosion is composed of all the points p satisfying that $p+b$ belongs to X. The result of the erosion operation is shown in Figure 2. It can be seen from the figure that the erosion operation will shrink the boundary of the original D to the inside, so it is often used to eliminate irrelevant details in the image, such as removing isolated small points and disconnecting unnecessary smallness. Connection etc.

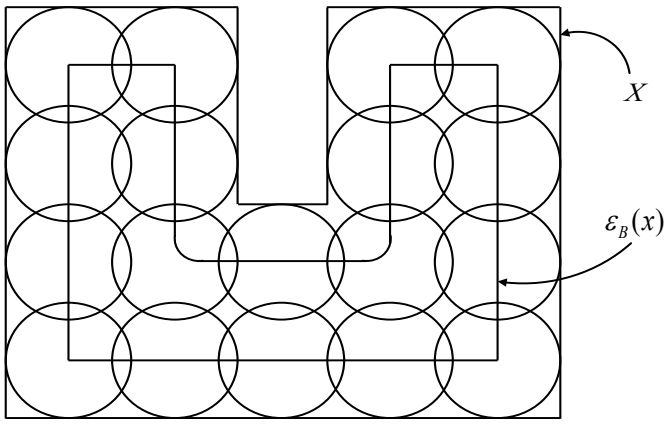


Figure 2. Corrosion

3.2 Open operation and close operation

In practical applications, corrosion and expansion are often used in combination as needed. From this, two other basic morphological operations can be derived-closed operation and open operation.

Corrosion first and then expansion is called open operation:

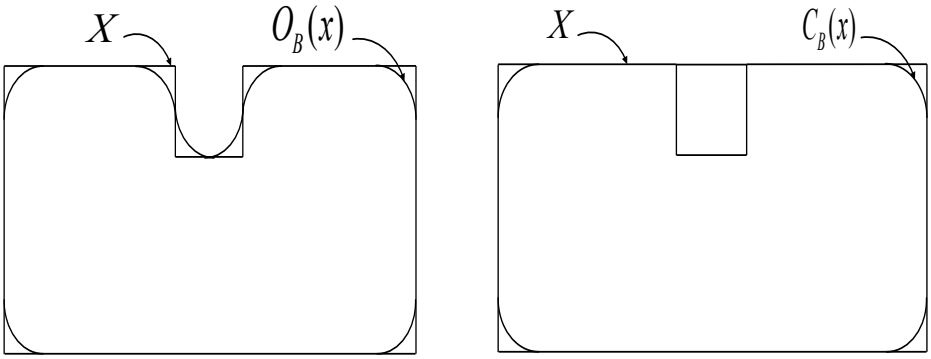
$$O_B(X) := \delta_B[\varepsilon_B(X)] \tag{7}$$

First expansion and then corrosion is called a closed operation:

$$C_B(X) := \varepsilon_B[\delta_B(X)] \tag{8}$$

Close operation and open operation are a pair of dual transformations. There is a big difference between the results of opening and closing operations on the same graph: the result of the opening

operation can eliminate the details of the image that are smaller than the structural elements, that is, some isolated points and small targets, and the corners of the graph become smooth, The straight part of the boundary will be consistent with the original X, and the local shape of the object will basically remain unchanged, as shown in Figure 3(a). Like the open operation, the closed operation can also smooth the contours of the object, but the difference is that the closed operation fills in the small holes inside the target and connects adjacent targets [6]. The closed operation results in a rectangle with rounded corners, as shown in Figure 3(b).



(a) Open operation (b) Close operation

Figure 3. Open and close operations

4. Accuracy analysis of double plane frame

4.1 Experimental design of frame accuracy analysis

Taking the dual-plane frame as the research object, applying the principle of three-dimensional camera measurement, two JVC9800 cameras (shooting frequency 50Hz) A and B are respectively placed at the positions shown in Figure 4 (the picture is quoted from ISO 12233-Resolution and spatial frequency responses). The connection of the main optical axis of the two cameras A and B to the shooting area are respectively 15.6m and 16.4m, and the height of the camera is 1.15m. Place a double-plane frame and a pole length of 740.0mm in the shooting area respectively for shooting [7]. The experimental test diagram is shown in Figure 4.

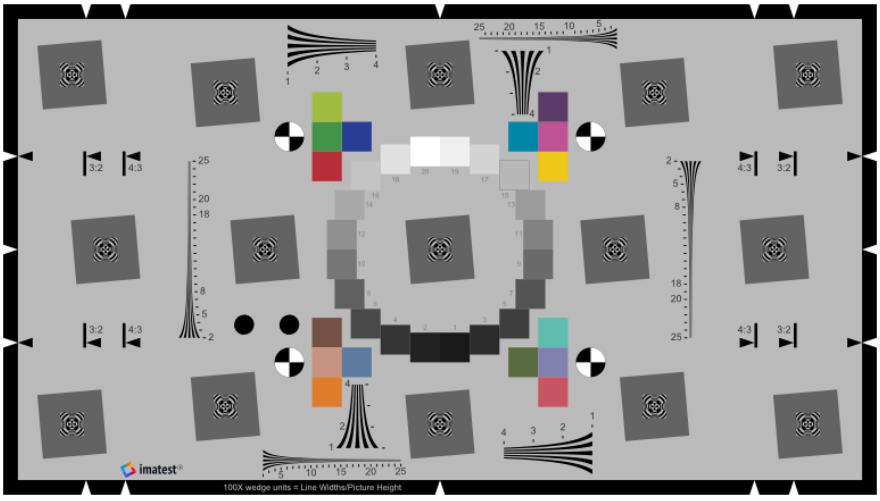


Figure 4. Frame accuracy verification test chart

4.2 Image acquisition part

Image collection is to read a picture from a high-speed video tape into the computer without distortion. In this process, the key technology is how to store the read-in picture in the computer in a

large capacity. The entire image acquisition part includes the following key components and working procedures.

4.3 Image card

It is proposed to use BSP1900 image card. The card can manage up to 1000M×8B image memory, can realize 768×576 full-screen sampling and display, and the horizontal sampling density can reach 1024 pixels. It adopts bus-type structure management, the expansion module is extremely convenient, it also has a video bus, can form a high-speed image processing system with other image processing cards, and is easy to operate and low in price.

4.4 Image correction

That is to correct the horizontal, vertical scale and other distortions of many images input to the computer. Its purpose is to make many frames on the display screen can be at the same position and level, which can be absolutely coincident, which is extremely helpful to reduce the calculation error.

4.5 Image preprocessing

It mainly includes two working procedures; one is to compress the image data entered by digital conversion equipment; the other is to emphasize certain characteristics of the image [8]. These tasks are very important, and it is even more important than the subsequent specific processing tasks, because the accuracy of the calculation of image picking points has a very close relationship with this.

4.6 Frame accuracy analysis experimental results

The camera parameters solved by the linear equation are shown in Table 1. The actual value of the benchmark length, the calculated value of the benchmark length and the relative error data obtained after two-plane frame calibration and three-dimensional reconstruction are shown in Table 2. It can be seen from the experimental data that the developed two-plane calibration frame has a relative error of less than 5% in the three-dimensional measurement, which can meet the needs of human motion parameter measurement.

Table 1. Camera parameters solved by linear equation

	Camera A	Camera B
p11	-1.5	-1
p12	-1.15	1.15
p13	0	-0.04
p14	401	343
p21	0.14	-0.35
p22	-0.15	-0.1
p23	-1.9	-1.64
p24	536.8	542
p31	0	0
p32	0	0
p33	0	0

Table 2. Frame accuracy analysis test data

Direction	Three-dimensional coordinates		Benchmark length/ mm		Relative error/%
	Point A	Point B	Calculated	Actual value	
x	(89, 28.3, 12.2)	(66.8, 45.9, 75.6)	710.0	740	4.05
y	(38.9, 14.5, 92.6)	(-30.1, 10.6, 108.8)	769.5	740	3.99
z	(21.99, -7.81, 0.66)	(23.2, -11.13, 75)	767.7	740	3.74

5. Conclusion

The establishment of a feasible sports image analysis system is essentially a very complicated task, and it needs to incorporate many high-tech. This article only starts from the specific practice of technical analysis and explores the overall structure of the sports image analysis system. It requires a closer combination of sports researchers and researchers in other disciplines to improve and improve sports images. Analyze the function and practical value of the system.

References

[1] Huang Xiaolong. Research on Feature Detail Recognition of Astronomical Photoelectric Images Based on Machine Learning. Automation and Instrumentation, vol. 4, pp.4-11, June 2020.

[2] Lu Laibing, Bruce Lee. The clustering and evolution of artificial intelligence technology in the field of international sports. Journal of Shandong Institute of Physical Education, vol. 36, pp. 12-19, March 2020.

[3] Li Jing, Fan Wenliang, Lei Ziqiao, et al. Recognition of Parkinson's patients based on machine learning model of magnetic resonance diffusion tensor imaging. Chinese Medical Equipment, vol. 36, pp.45-47, October 2021.

[4] Ye Chunming, Zhao Shengwen, Yang Xiuhong, et al. Analysis and prediction of young athletes' ability to respond to the new crown pneumonia epidemic based on machine learning. Journal of Physical Education, vol. 27, pp. 68-70, March 2020.

[5] Du Xiuli, Jiang Jiawei, Xu Zigang, et al. Research on seismic performance index calibration of shallow-buried rectangular frame subway station structure. China Civil Engineering Journal, vol. 52, pp.10-19, October 2019.

[6] Jia Chunyan, Cui Li. Research on Mathematical Model for Estimating Direction Angle of Motion Blurred Image. Mathematics in Practice and Knowledge, vol. 49, pp.82-88, December 2019.

[7] Li Xiang, Yang Cankun, Zhou Chunping, et al. Research review of target motion information extraction from high-resolution optical satellite images. Remote Sensing for Land and Resources, vol. 31, pp.99-103, March 2019.

[8] Chen Mei, Che Shangyue. Research on tracking control strategy of uncalibrated visual servo robot. Control Engineering, vol. 26, pp.55-62, June 2019.