International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.1 (2016), pp.55-66
http://dx.doi.org/10.14257/ijmue.2016.11.1.06

Research on Tracking Technology in Basketball Video

Shengbo Liao [1], Miao Wang [2] and Haitao Yang [3,*]

[1] *Beijing Jiaotong University, Beijing 100000, china*
[2] *Harbin Sports Institute, Harbin 150001, china*
[3] *Beijing University of Technology, Beijing 100000, china*
daliao168@tom.com

Abstract

Tracking of basketball is studied. Firstly, the basketball object template and adaptive object model are established by using the basic characteristic of segmented basketball in the former frame. Then an improved block-matching algorithm is proposed for tracking the basketball. Finally, the edge of basketball is revised, and the mechanism for tracking validity is established.

Keywords*: sport video analysis, basketball, object tracking, motion estimation*

1. Introduction

In segmenting the basketball object, fetching features of the object involves huge calculated quantities, which can't satisfy the real-time requirement of algorithm [1-2]. To enhance the efficiency of cutting basketball object, based on the time correlation of basketball motions [3-4], we take full advantage of basic characteristics of segmented basketball in the current frame to split basketball in subsequent frames with the inter-frame tracking technology. Firstly, we use the basketball area divided in the preceding frame to create basketball object template and thus make the model; then, improve existing block matching algorithm with the object model and follow up the position of basketball in the current frame. To solve the tracing result deviations caused by image noises and the profile shrinkage of basketball object, we utilize the peak property of edge gradient to correct. Last, we build the effectiveness detection mechanism for basketball area tracing. If the tracing is effective, use the area to update object template; otherwise, segment the basketball and use the cut basketball area to re-create basketball object template [5].

2. Basketball Object Template and the Object Model

2.1. Basketball Object Template

Due to basketball color consistency and round shape, we use color and shape features of basketball cut in the previous frame to build current basketball object template.

Assume the circle center and radius $(Circle\_x, Circle\_y)_{n-1}$ and $Radius_{n-1}$ of basketball segmented in the $n-1$ frame. The round shape feature $(Circle\_x, Circle\_y, Radius)_n$ of basket object template $Template_n(x,y)$ in the current frame is represented by the circle center and radius of basketball in the $n-1$ frame. It is shown in formula (1).

$$\begin{aligned} Circle\_x_n &= Circle\_x_{n-1} \\ Circle\_y_n &= Circle\_y_{n-1} \\ Radius_n &= Radius_{n-1} \end{aligned} \tag{1}$$

ISSN: 1975-0080 IJMUE

Taking into account the hue H component of HSV color space is robust to illumination, so, the current frame basketball object template is represented in the $n-1$ frame of frame basketball $Hue_{n-1}(x,y)$. It is shown in formula (2).

$$Template_n(x,y) = Hue_{n-1}(x,y),\ (x,y) \in (Circle\_x, Circle\_y, Radius)_n \qquad (2)$$

2.2. Basketball Object Model Based on Statistical Inference

Basketball hue clustering has good consistency. Figure 1 is the hue histogram of basketball in different lights. Obviously, the ball hue basically concentrates between $R_{Fore}:[0,50]$ and $R_{Back}:[300,359]$.

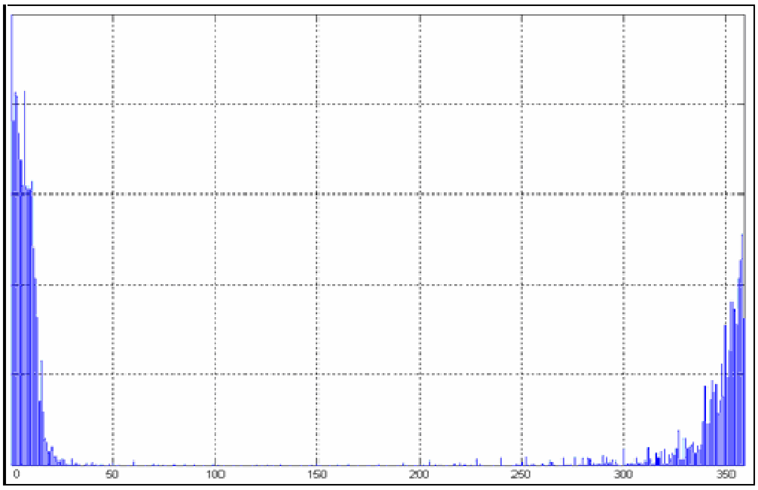


Figure 1. Basketball Color Clustering Results

According to basketball color consistency, we utilize the hue information of object template to create the object model, in the following steps:

2.2.1 Calculate hue mean μ_n and variance δ_n of the template in R_{Fore} and R_{Back}; make R_{Fore} mean μ_n^F and variance δ_n^F; make R_{Back} mean μ_n^B and variance δ_n^B;

Take R_{Fore} for instance and estimate the hue mean and variance.

Count hue distribution $Hist[H]\ H \in [0,359]$ of object template $Template_n(x,y)$, *i.e.* number of pixels with the same hue in the template, as seen in Figure 2; then, normalize $Hist[H]$ by formula (3) and meanwhile eliminate scattered hue points to have the histogram of hue distribution of basketball object template. It is shown in Figure 2 (b).

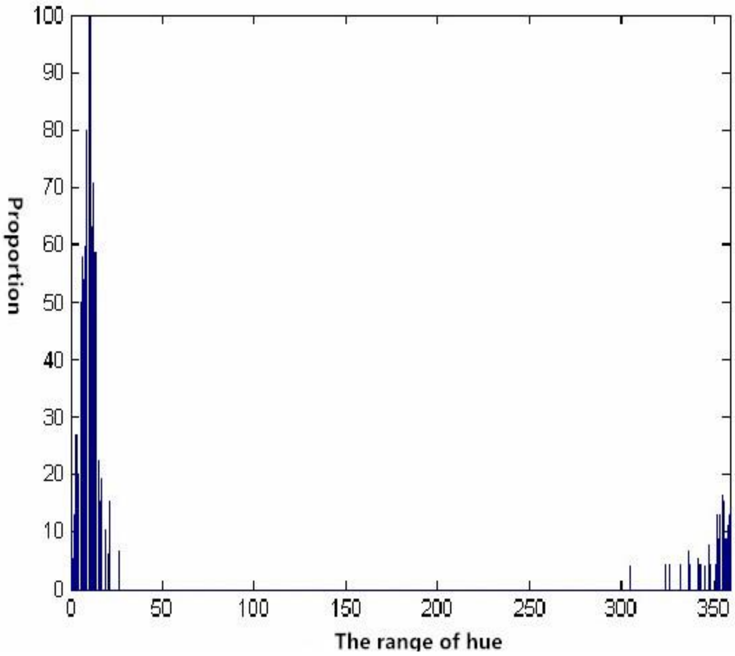
$$Hist[H] = \begin{cases} 100.Hist[H]/Max\_H & Hist[H] \geq 5 \\ 0 & Hist[H] < 5 \end{cases} \qquad (3)$$

Calculate hue mean μ_n^F and variance δ_n^F of the template in R_{Fore}, It is shown in formula (4).

$$\begin{cases} \mu_n^F = \sum_{H=0}^{50} H.Hist[H]/Num_F \\ \delta_n^F = \sqrt{\sum_{H=0}^{50} (H-\mu_n^F)^2.Hist[H]/Num_F} \end{cases} \qquad (4)$$

Similarly, Calculate hue mean μ_n^B and variance δ_n^B of the template in R_{Back} . It is worth noting that, long shot basketball color only is on R_{Back} . Therefore the hue mean and variance of basketball in long shot of the object template are $\mu_n^B = \delta_n^B = 0$. At this time, only need to calculate μ_n^F , δ_n^F .

(a) The White Annular Region - Object Template



(B) Hue Histogram of Normalized Template

Figure 2. Basketball Hue Clustering Results

2.2.2 Create object model

Based on the above result, we create basketball object model $Model_{Object}$

$$Model_{Object} : H \in \{[H_1, H_2] \cup [H_3, H_4]\} \tag{5}$$

Choose $\alpha = 1$ and the choice of object model $Model_{Object}$ interval is accurate; see Figure 3(b), where white area stands for in basketball area, pixel dots which are accordant with $Model_{Object}$ hue; apparently, what's irrelevant like words, stripes, trademark on the basketball and human hands is all removed.

(a) The White Ring - Basketball Object Template

(b) The White Area - Pixel Hue Range of Object Model

Figure 3. The Selection Effect of Object Model

3. Tracking of Basketball Movements

3.1 Block Matching Algorithm

Block matching method was developed by Jain in 1981 [6]. Its basic idea is to cut images into numerous small blocks which are mutually disjoined but spread over the whole image plane; movements like translation, rotation and scaling of all objects between original and reference images are realized by the translation of those blocks; while relative positions of dots inside blocks remain the same between two images, *i.e.* each dot inside blocks has the same motion vector. Make sub-block $MB_k(i, j), i = 0,1,\ldots, M - 1, J = 0,1\ldots, N - 1$ in the k th frame; M and N is respectively length and width of matching block. Search in the $k - 1$ frame the most similar matching block to k.

In the block matching method, block shape doesn't affect matching results; but block size does, because in the method, all movements in the image are based on "block". Hence it's ensured as much as possibly that each block comes from one single object and each pixel dot inside the block has the same motion vector. The block size should possibly be minimal, to make consistent the motion vector of each dot inside the block. However, if the block size is smaller, it's easily influenced by noises. So it's important to choose properly the size of block.

Normally, block matching method has two questions about:

(1) How to define similarity of two blocks, *i.e.* criteria of block matching

(2) How to find the most similar block, *i.e.* search strategy

3.1.1 Matching Criteria

Matching criteria are rules used for measuring the relevancy of two image blocks. The selection of such rules affects directly the complexity of searching process and accuracy of estimating motion vector. We introduce some common block matching criteria in the following.

(1) Cross correlation function method

The definition of the cross correlation function:

$$CCF(i,j)=\frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}B_k(m,n)B_{k-1}(m+i,n+j)}{\left[\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}B_k^2(m,n)\right]^{1/2}\left[\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}B_{k-1}^2(m+i,n+j)\right]^{1/2}} \tag{6}$$

(2)Mean square error

The definition of the mean square error:

$$MSE(i,j)=\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}[B_k(m,n)B_{k-1}(m+i,n+j)]^2 \tag{7}$$

(3)The absolute difference of sum

The definition of the absolute difference of sum:

$$SAD(i,j)=\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|B_k(m,n)B_{k-1}(m+i,n+j)| \tag{8}$$

We know that multiply operation by computer is more complicated and the calculated amount is far bigger than additive operation. Generally, it's preferable to use addition instead of multiplication if possible. In the three methods, since the sum of absolute error is computed not by complex multiplying, the workload is relatively less. So it can be applied the most widely.

3.1.2 Search Strategy

Generally speaking, motion vectors are highly concentrated near the center of search window [7], in other words, the relativity SAD of image blocks has monotonicity. Based on that feature, many strategies were proposed for searching block movements.

(1)Full search method

Of all motion searching algorithms, full search method is of the best performance, because it implements block matching calculation of every dot in the whole search window and gets the optimal matching point according to SAD minimum principle. As the method searches point by point, its calculated amount is big; if system processing speed is not fast enough, it's difficult to realize the synchronization of multimedia system with it. The performance of any other algorithm is not better than full search algorithm; but other methods can reduce workload by cutting down points in the search window. However, the performance is surely declined, increasing processing rate at the expense of performance. In recent years, quick search methods have been extensively concerned. There are many effective algorithms.

(2) Three-step search approach (TSS)

T. Koga *et al.* proposed three-step search method in 1981 [8-9]. It is an efficient fast search algorithm. It keeps almost all functions of full search method, but its calculated amount takes up only 10% of the latter. It is broadly applied in video conference and video phone. Three-step method makes initial step length T and shortens gradually the search step length. The search each time is based on the last search results, at the step length of 3 x 3 pixels and search accuracy of 1 pixel. Figure 4 shows the whole procedure of TSS: firstly, use window center as the center; make step length 4 and search eight points nearby; SAD minimal corresponding to the optimal matching point; secondly, use the above optimal matching point as center, step length being halved to 2 and continue searching eight points round to have the matching point; thirdly, likewise make step length 1 to get the best matching point as thus to get motion vector for predicting images. TSS searches totally 25 points; while full search method does 225 points, running time decreasing considerably but performance weakening a little bit.

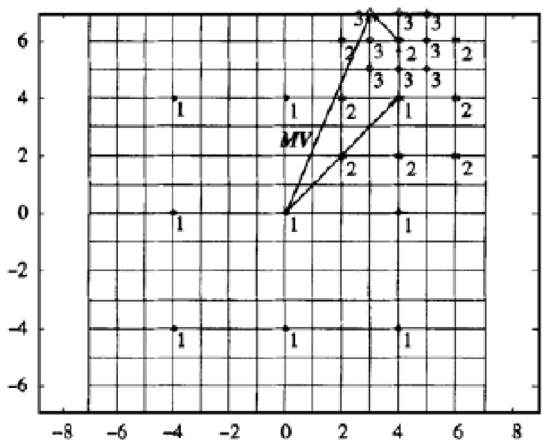


Figure 4. Three-Step Search Approach

(3) Two Dimension Logarithm method (TDL)

This method tracks the smallest SAD point through fast search (Figure 5). Starting from motion vector (0,0), we choose five points distributing in cruciform to constitute point group of each search as to get the smallest SAD point. If the smallest point appears in one edge of cruciform point group, the next search should center around the point, with the same step length; if the smallest point appears in the center of crisscross, the next search centers around it, with the step length cut by half; if in the searching process, the central point of new cruciform point group lies in one edge of search window, the step length is halved. Repeat the operation till the step length becomes 1; by now the smallest matching point with SAD is the optimal one. This method is slightly inferior to TSS but the searching time is apparently shortened. If in search window, there is only one SAD minimal point, *i.e.* only one valley point, the search result is the optimal matching point; if there're several valley points, the search may go to one, which is otherwise mistaken as the best matching point, and now the search stops. Therefore, its performance is not as good as TSS, but speed is much quicker.

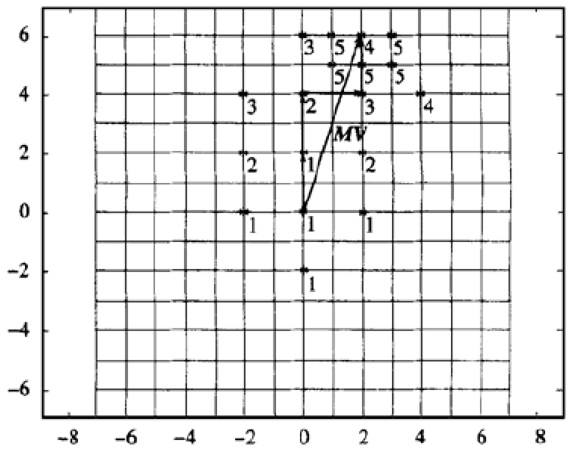
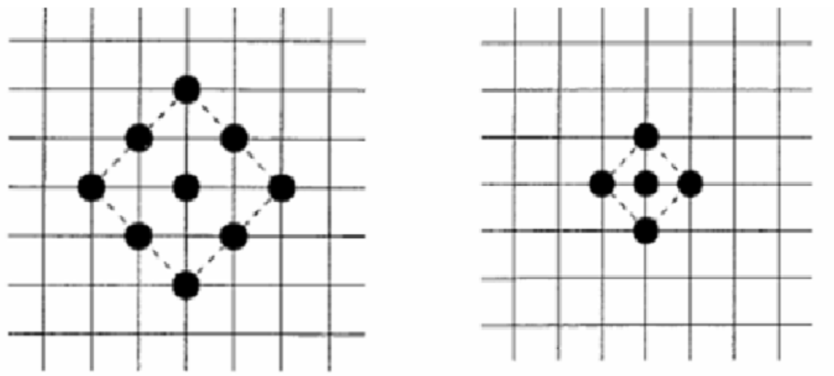


Figure 5. Two Dimension Logarithm Method

(4) Diamond search algorithm (DS)

In general, the shape and size of search template have effects on the speed and precision of motion estimation algorithm. If small search template is chosen, it may fall into local optimal; if big template is chosen, it may not search out the best point. Diamond search method is known for its template shape [10-11]. It has two kinds of search template: large diamond search pattern (LDSP) and small diamond search patter (SDSP), as shown in Figure 6.

Starting from motion vector (0,0), we use nine points distributed in LDSP to compose point group and find the smallest SAD point. If such a point appears in one of eight points in rhombic shape, the next search should center on it and still use LDSP point group; if such a point appears in the center of diamond shape, it means being close to the best matching point and the next search centers on it, in the use of SDSP. In the SDSP, the matching point whose SAD is minimal is the best one. Figure 7 presents the whole procedure of DS.



(a) Large Diamond Search Pattern (b) Small Diamond Search Patter

Figure 6. Diamond Search Patter

Diamond search algorithm has own features, *e.g.* analyzing the basic rules about motion vector in video images; choosing big and small search patterns. At first, we use LDSP to search. Since step length is big and search range is wide, we can perform rough positioning to avoid the search process from falling into local minimal; when the coarse positioning ends, it's thought that the optimal point locates in the rhombic area formed by eight neighboring points; then, we use SDSP to position accurately, avoiding big changes to the search. For now, we say its performance is better than other rapid search algorithms.

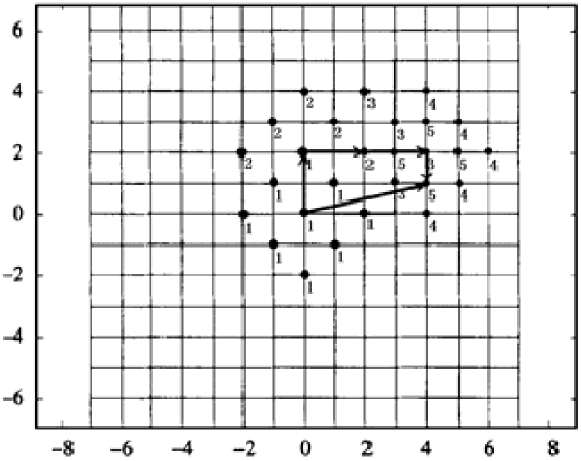


Figure 7. Diamond Search Algorithm

3.2. Tracking Basketball by the Improved Block Matching Method

Block matching is easily achieved and requires little calculated amount. Also it has favorable image matching ability. We're thinking if it's possible to improve it as for tracking basketball object. This is why we designed the idea of basketball tracing. We'll improve it from the aspect of matching criteria and search strategy.

3.2.1 Improved Object Matching Criteria

The usual block matching criteria like cross-correlation function (CCF), mean square error (MSE) and sum of absolute differences (SAD) are all based on regional location correlation. It requires that each pixel in the matching block makes translational movement of the same nature. But in basketball match videos, in addition to translation, basketball can self-rotate in three-dimensional space, leading to incoherent motions of each pixel in basketball area and unpredictability of hue variation in the area (Figure 8). As a result, in basketball movement tracking, the matching criteria will become infeasible.

(a)177 帧 (b) 179 帧 (c) 181 帧

Figure 8. Basketball Rotary Motion

To get rid of interferences by basketball self-rotation, we designed statistical matching criteria irrelevant to regional location, *i.e.* sum of object points (SOP) matching criteria. We'll introduce briefly SOP calculation method.

When looking for the most similar block to object model in current frame, we make (i, j) the motion vector of the current search point against the initial point (*i.e.* the center of object template $Template_n$. By now we define $SOP(i, j)$ as the number of pixels which accord to object model $Model_{Object}$ hue in the matching block. The size of matching block is the round area of object template $Template_n$.

Apparently, SOP matching criteria is one statistical matching algorithm having nothing with block regional position. They can eliminate complicated influencing factors like object rotation and partial cover.

3.2.2Improved Diamond Search Path

Search path is a major factor affecting the performance of block matching algorithm. Also, it's a key technology in movement search algorithms. Of many search algorithms, diamond search algorithm has high efficiency and good prediction result, not limited by the size of search area. So it's widely used in international standards like MPEG-4.

Diamond search algorithm selects LDSP and SDSP, as seen in Figure 9, SDSP labeled as ③, ① and ② form LDSP.

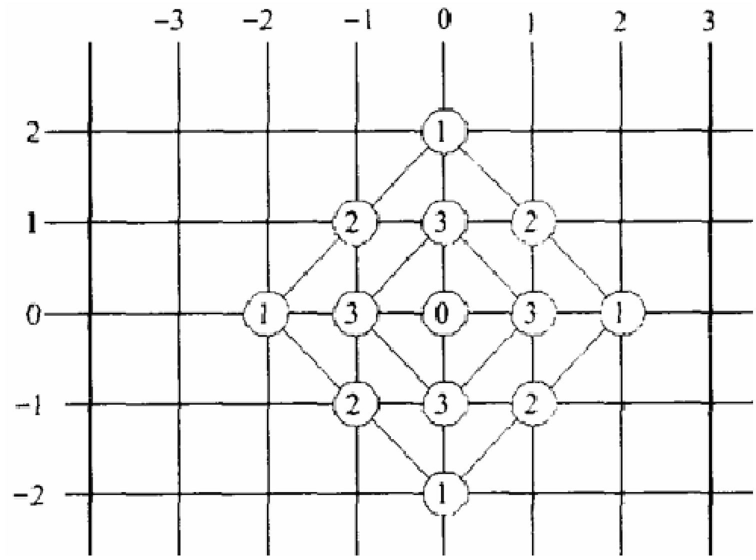


Figure 9. Diamond Search Pattern

Based on the strong correlation and center bias of basketball motion vector, we'll upgrade the traditional diamond movement search algorithm as to reduce search times and enhance operation speed, in the following steps:

(1) Beginning with the initial point, we use five points in ① to form search point group and find the biggest SOP point;

(2) If the optimal point is one point of ①, compute SOP of its two neighboring points marked with ②; make the biggest SOP point as center and still use points marked with ① to constitute search point group;

(3) Repeat (1) and (2) till the best point appears in the center of rhombus, suggesting it's close to the optimal matching point; the next search should center on it and utilizes LDSP (*i.e.* points marked with ③), of which the maximum SOP point is the optimal matching one.

Here we divided traditional large diamond search pattern into label ① and ②. Evidently, the search points are fewer than the traditional diamond movement search algorithm.

With the improved diamond movement search method, the search speed was enhanced and we got better basketball tracing result. But video noises and basketball profile shrinkage had some impacts on the generation of tracking results, causing minor divergences to tracking result, as found in Figure 10(a). Multi-frame cumulated deviations made basketball tracking invalid. For the goal here, we corrected basketball contour.

(a) Tracking Error of Contour (b) Contour Correction

Figure 10. Tracking Results of Basketball

4. Experimental Analysis and Results

The hardware and software platform of basketball tracking experiments are shown in Table1.

Table 1. The Software and Hardware Experiment Platform

Cpu	Dual core2.4
Memory	2GB
OS	WIN7
Programing language	VC++6.0
Image format	352*288 bmp

The experiment chose five videos of two matches as testing data: the 22th FIBA Asia Championship China mainland VS China Taipei and Japan VS Philippines. The experimental results are listed in Table2.

Table 2. Tracking Results of Basketball

Video clips	The length of the sequence	Basketball area	Processing time (seconds)	Tracking speed (frames/ second)
1	3213	150-200	116.1	27.7
2	2725	150-200	101.2	26.9
3	983	200-300	39.1	25.2
4	730	400-600	29.7	24.5
5	226	2000	12.5	18.1
Total	7877		298.6	26.4

From the Table2, we see when basketball area is bigger than 2000 pixels (Figure11(a)), the tracking speed is about 18 frames/second; when the area is about 500 pixels (Figure11(b)), the speed is around 24 frames/second; when the area is below 200 pixels, the speed is approx. 27 frames/second (Figure11(c)). Clearly, with increasingly bigger basketball area, basketball movement tracing speed got slower. After statistics of two matches, we found that video fragments where the basketball area was over 200 pixels took about only 3% of the playing time. So we think if the tracking speed can reach 25 frames/second, the real-time requirements can be met.

(a) 2109 Pixels (b) 517 Pixels (c) 193 Pixels

Figure 11. Tracking of Basketball

5. Conclusion

This paper used the hue characteristics of basketball object template, established an adaptive object model for object tracking of basketball. The model can effectively eliminate the non-basketball area, thereby reducing the interference area has nothing to do with the basketball to the algorithm. From The matching criteria and search path aspects, improved traditional block matching algorithm. The improved method is to improve the computation speed at the same time, effectively eliminate the interference caused by the rotating movement of basketball. Starting from the actual basketball tracking, improved calculation method of active contour model is designed to measure energy, extracting the effective detection index of basketball. The experimental results show that this method can effectively track the basketball object in complex background, and the basketball part occlusion, motion blur and basketball's rotation of complex motion have good robustness.

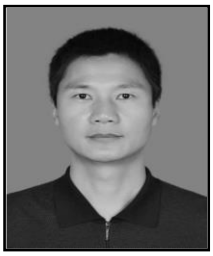
Acknowledgement

This work was supported by The Fundamental Research Funds for the Central Universities. No. HEUCF151601.

References

[1] Hong, "Research on human joint motion tracking technology", Harbin Engineering University, **(2008)**.
[2] Z. Dongmei, "Analysis of training in Basketball Teaching in Colleges and universities sports tactical consciousness", The sports science and technology, vol. 36, **(2014)**, pp. 116.
[3] B. Linhai, "Tracking technology of video based human motion", Shenyang University of Aeronautics and Astronautics, **(2011)**.
[4] W. Yanbin and Z. Wei, "Study on the ergonomics of human motion tracking technology", ergonomics, vol. 2, **(2007)**, pp. 29-31.
[5] L. Xiaoyue and M. Yan, "The research of moving object detection and tracking algorithms", Journal of Hebei Union University (Natural Science Edition), vol. 1, **(2015)**, pp. 65-70.
[6] J. R. Jain and A. K. Jain, "Displacement measurement and its application in inter-frame image coding", IEEE Trans. Communication, vol.COM-29, **(1981)**, pp. 1799-1808.
[7] Li R., Zeng B. and Liou M. L., "A new three-step search algorithm for fast black motion estimation", IEEE Trans Circuits System Video Tech, vol. 4, **(1994)**, pp. 438-442.
[8] D. Xu, C. Bailey and R. Sotudeh, "An Improved Three-step Search Block-matching Algorithm for Low Bit-rate Video Coding Applications", In: Int. Symposium on Signals, Systems, and Electronics, **(1998)**, pp. 178-181
[9] T. Kga, "Motion-compensated inter-frame coding for video conferencing", in Proc. Nat. Tel-communication Conf. New Orleans, **(1981)**, pp. G5.3.1-G5.3.5.
[10] A. C. N. Alan and B. Zeng, "A New Fast Motion Estimation Algorithm Based on Search Window Sub-Sampling and Object Boundary Pixel Block Matching", Image Processing, ICIP 98, **(1998)**, pp. 605-608.
[11] S. Zhu and K. K. Ma, "A New Diamond Search Algorithm for Fast Block Matching Motion Estimation", IEEE Trans Image Processing, vol. 9, no. 2, **(2000)**, pp. 287-290.

Author



ShengBo Liao, He received his B.S degree from Beijing Sport University, and received his M.S degree from Beijing Sport University. He is an associate professor in Beijing Jiaotong University. His research interests include Physical education and training.

