International Journal of Multimedia and Ubiquitous Engineering
Vol.10, No.12 (2015), pp.183-194
http://dx.doi.org/10.14257/ijmue.2015.10.12.19

Research on Long Shot Segmentation in Basketball Video

ShengBo Liao [1], Jingmeng Sun[2] and Haitao Yang[3,*]

[1]Beijing Jiaotong University, Beijing 100000, china
[2]Physical Education Department, Harbin Engineering University, Harbin 150000, China
[3]Beijing University of Technology, Beijing 100000, China
daliao168@tom.com

Abstract

In the basketball segmentation in long shots, Gauss filter is adopted to smooth noise in the image firstly. Secondly, the background is separated by the difference of inter-frame and the connected regions are labeled. Thirdly, a strategy is designed to identify basketball by the characteristic of basketball. Finally, the edge deviation is revised and the optimal result is obtained using improved Snakes.

Keywords: *sport video analysis, video segmentation, basketball, long shot*

1. Introduction

In the basketball game field, plentiful cameras are working simultaneously. When recording game videos, workers need only to switch cameras in different positions as to take videos of the game in different visual angles [1-2]. Based on the analysis of basketball features in video shots, videos can be classified to long shots and close shots, as portrayed in Figure 1. Apparently, close shots often refer to close-up shots of persons like players and judgers, and others about penalty and referee holding the ball. Long shots are to present offensive and defensive play in the full court. They are common shooting way used to record such games. Images like dribbling, passing and shooting are taken usually in long shots [3-4].

In basketball game videos [5-6], there are lots of background information. And basketball sport has strong structural features. Hence we use the video segmentation method based on varying region detection, that is, splitting basketball object and background area by detecting changing and unchanged region that each frame of video sequence images trespasses over [7-8]. The flow chart of the method is shown in Figure 1. First of all, utilize Gauss filter to smooth images; then, separate background region with inter-frame difference and mark connected object areas; next, according to basketball region characteristics, formulate “basketball recognition strategy” and keep possible basketball region (*i.e*. candidate basketball region); lastly, apply the improved active contour model to modify edge deviations and build detection area as to get the effective strategy for the optimal solution [9-10].

2. Detection of Varying Regions Based on Inter-Frame Difference

2.1. Threshold Segment of Inter-Frame Difference Image

If the relative movements of cameras and objects are neglected, settings in the basketball game videos are basically unchanging [11]. Since basketball motions are obvious, basketball area in certain frame range is changing noticeably. Based on that

[*] Corresponding Author

ISSN: 1975-0080 IJMUE

feature, we use inter-frame changing information to detect objects in blank areas. The algorithm has the following steps.

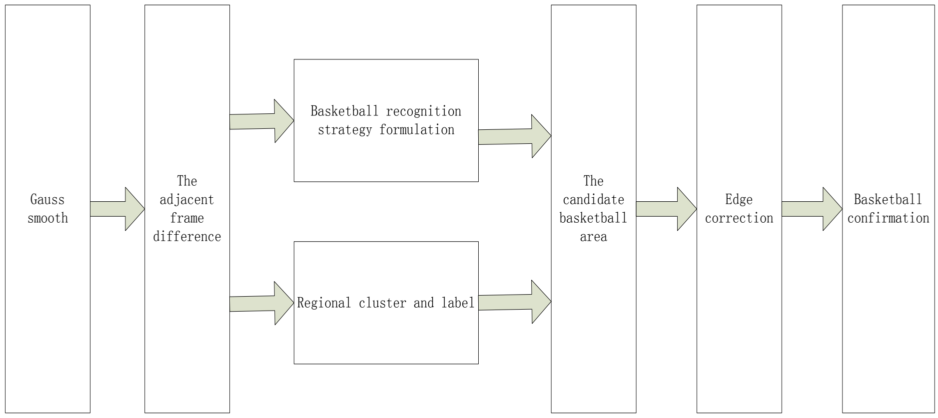


Figure 1. The Basketball Segmentation of Long Shot

2.1.1. Image Smoothing

Video image frames have overlapping noises [12-13]. To enhance the effect of calculating inter-frame difference, firstly we need to preprocess input sequence. Here we utilize Gauss filter to smooth image noises.

Define two-dimensional Gauss function as:

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

Where, The gradient vector is $\nabla G=\begin{bmatrix}\partial G/\partial x\\ \partial G/\partial y\end{bmatrix}$, can improve the speed by decomposition method, the two filter convolution template that ∇G is decomposed into Two one-dimensional line filter

$$\begin{aligned}\frac{\partial G}{\partial x}&=kx\exp[-\frac{x^2}{2\sigma^2}]\exp[-\frac{y^2}{2\sigma^2}]=h_1(x)h_2(y)\\ \frac{\partial G}{\partial y}&=ky\exp[-\frac{y^2}{2\sigma^2}]\exp[-\frac{x^2}{2\sigma^2}]=h_1(y)h_2(x)\end{aligned} \tag{2}$$

In actual processing, Gauss function performs discrete approximation. We'll take 5x5 Gauss mask to realize Gauss smoothing with standard convolution. The size of mask template affects directly the smoothing effect. The bigger the template is, the better result smoothing will have. But at the same time, edges become obscure, which will have effects on edge detection. We choose $\sigma=1.5, k=1$ and get the Gaussian template $G(i,j)$, it is shown in Figure 2.

Set input video sequence frame size $W\times H$ (W and H is width and height of frame image); the grey value $f_n(x,y), 0\le x\le W-1, 0\le y\le H-1$ of pixel dot (x,y) of the nth frame. After the nth frame image is filtered by Gauss function, we get the result:

$$g_n(x,y)=\sum_{i=-2}^{2}\sum_{j=-2}^{2}[G(i+2,j+2)\cdot f_n(x+i,y+j)] \tag{3}$$

2	4	5	4	2
4	9	12	9	4
5	12	15	12	5
4	9	12	9	4
2	4	5	4	2

Figure 2. Gauss Function of Discrete Approximation

2.1.2. Frame Difference Image and its Binaryzation

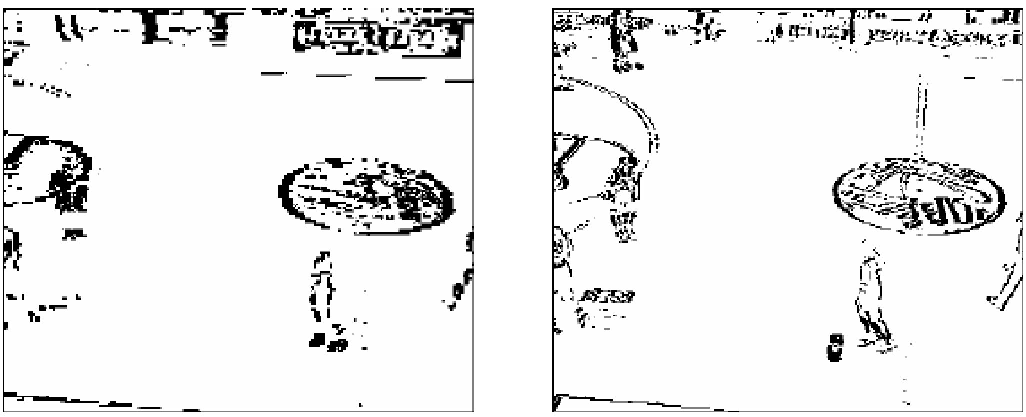
Compute the frame difference D_n of image sequence g_n between the front frame g_{n-1} after smoothing filter, which can be expressed as:

$$D_n(x,y) = \left| g_n(x,y) - g_{n-1}(x,y) \right| \tag{4}$$

By the analysis of the front, the frame difference image will have a high gray value in sports area, and the pixels of the background is dark. Therefore, set the threshold T on the frame difference image D_n is binaryzation, *i.e.*

$$D_n(x,y) = \begin{cases} 0 & D_n(x,y) < T \\ 255 & D_n(x,y) \geq T \end{cases} \tag{5}$$

In HSV color space, all components are not very stable. Comparing images of gray frame difference and hue frame difference, we find the latter brings about many noises and can’t check out basketball object (Figure 3 (a)). Gray frame difference can detect well the basketball and many background areas are removed, as seen in Figure 3 (b).



(a) Tone Frame Difference (b) Gray Frame Difference

Figure 3. Gray Frame Difference and Tone Frame Difference Image

2.2. Clustering of Connected Regions

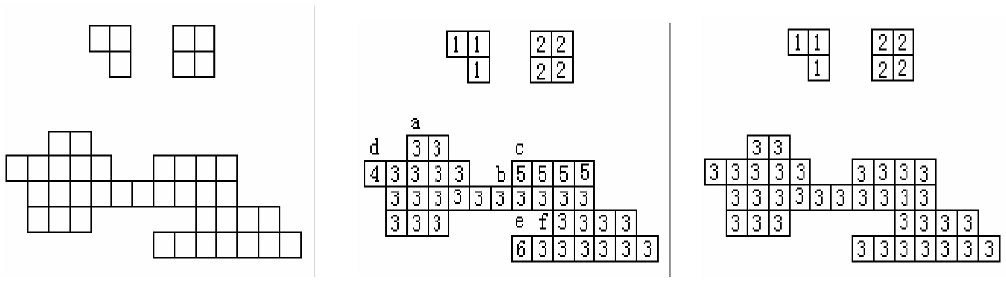
From Figure3 (b), we note that although most non-basketball areas are wiped out, there are still many irrelevant regions. We'll depend on the statistic information of the acreage and shape of each connected region to design "basketball recognition strategy". At first, we adopt region clustering method to tag each connected region of gray frame difference image D as to acquire the parameter of the region. The clustering includes two steps:

2.2.1. Mark Number of Every Target Pixel

Firstly in frame difference image D_n, scan from left to right and top to bottom. When the gray value of a pixel dot in the target area is 0, detect eight neighborhoods of the dot. If they are not marked, then label the pixel dot with a new value; or, use the smallest labeling value of the eight neighborhoods to label the dot. In Figure 4 (b), when dot is scanned, since its eight neighborhoods are not marked, the labeling value is 3 because the labeling value of one preceding target pixel dot is 2. Pixel point c, d and e is similar to a. When b is scanned and two pixel dots of its eight neighbors are marked, whose value is respectively 3 and 5, so the labeling value of b is 3. Pixel dot f is similar to b. In the third target area in Figure 4 (b), there are several dots with labeling value (3, 4, 5, 6). So it's necessary to cluster those dots and make their labeling value all 3.

2.2.2. Clustering of Every Target Area

In frame difference image D_n, scan progressively from top to bottom (firstly from left to right and then right to left). When there is a target pixel dot, detect its eight neighborhoods. If the smallest labeling value of them is bigger than that of the pixel, the pixel's value should be replaced with the smallest value. When scanning the whole picture is over, scan progressively in the same manner from bottom to top till the labeling value of all target pixel dots won't change any more. In Figure4 (c), three target areas are labeled with 1, 2, and 3. Clearly, all connected regions in the image are given only one label.



(a) Connected Region (b) Scan Area (c) Region Clustering

Figure 4. Connected Region Label

Figure 5 (a) is the difference of continuous frame gray $D_n(x, y)$, set $C_n(x, y)$ as a result of connected region label. It is shown in Figure 5 (b). Established array of structures $Object(m), m \in [1, MaxTag]$. Description $C_n(x, y)$ features of each label area.

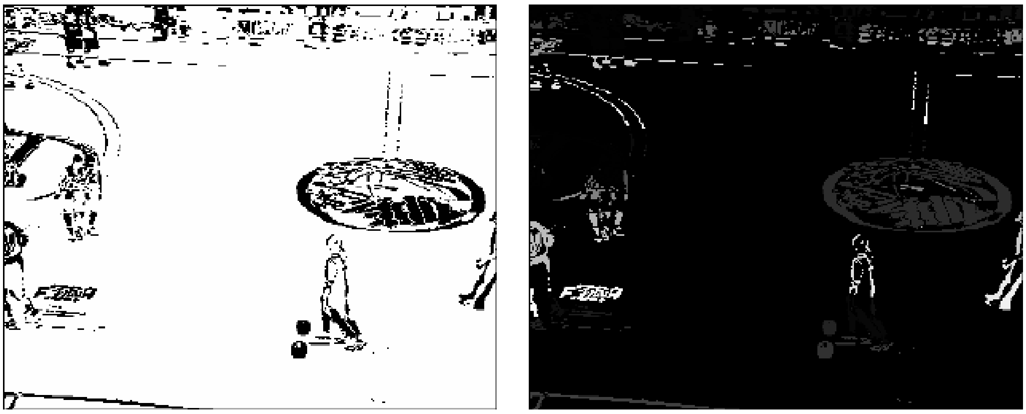
Where, MaxTag is the label maximum, structure parameters are defined as follows:

```
Typedef struct ParameterRegion {
        int     Tag;
        int     Acreage;
        int     Width;
```



```
        int       Height;
        int       NumObject;
    } Region
```

The target area of Acerage is the total number of pixels of the label value; Width is the target zone width, *i.e.*, the pixel position of label the target area left, right is difference, Height is the target region length, *i.e*. label pixels most below, on the target zone is difference; NumObject is a basketball and point target zone, including whether the decision conditions of basketball point will be based on the basketball tone model



(a) Gray Frame Difference Image (b) Clustering Results of Target Area

Figure 5. The Target Area Map after Pretreatment

3. Basketball Recognition

Add up features of basketball in long shots to plot "basketball recognition strategy". Referring to parameters like region shape and square meter, eliminate non-basketball regions which don't meet decision conditions.

3.1. Basketball Recognition Strategy

According to the feature that in long shots, basketball region features are of good consistency with hue clustering, we design "basketball recognition strategy" as follows:

3.1.1. Removal of Fragments

In Figure5 (a), there're many "fragment" areas, whose size is smaller than 25 pixels. Statistics reveal that in long shots, basketball area size is bigger than 25 pixels. So the area whose size is smaller than 25 can be considered as "fragment" area and should be removed. It is shown in formula (6)

$$Object(m).Tag = 0 \qquad Object(m).Acerage < 25 \tag{6}$$

3.1.2. Elimination of Hue

The hue clustering of basketball in long shots is consistently good. The hue is concentrated. So we establish the $Model_{Hue}$. It is shown in formula (7)

$$Model_{Hue} : H \in [0,40] \tag{7}$$

3.1.3. Regional Directionality

Although in long shots the edge directionality of basketball is bad, its region directionality is good, *i.e.* region Height basically close to Width. Here, Width is location variation between the leftmost labeled pixel and the rightmost in the target area; Height is position difference between the uppermost labeled pixel and the lowest in the target region. It is shown in formula (8)

$$\begin{cases} Object(m).Width = x\_Right - x\_Left \\ Object(m).Height = y\_Bottom - y\_Top \end{cases} \quad (x, y) \in Object(m) \tag{8}$$

3.2. Extraction of Candidate Basketball Region Features

After treated by basketball recognition strategy, most irrelevant regions are cleared. Then sum up regions whose Tag is not 0 (which is called candidate basketball area). Build the structural body which can effectively describe features of candidate basketball areas. Parameters are designed as follows:

```
Typedef struct CandidateRegion{
        int    Tag;
        int    Circle_x;
        int    Circle_y;
        int    Radius；
        double   EnergeInside;
        double   EnergeContour;
    } Basket;
```

To fetch the basketball in long shots, we still use circle center and radius to characterize the basketball. So the extraction of basketball region features in long shots is divided into the fetch of circle center and radius.

3.2.1. Track of Region Edge Contours

To extract region’s center and radius, it’s required to track region edge contours. Traditional tracking methods are reptile and raster scanning method. Those methods must repeat several times to finally get results. Since the times of repetition can’t be well controlled artificially, the tracing result may not be accurate. Sometimes, there may be reiterative tracing of one local area, causing the program to fall into dead cycle.

Here we utilize the improved edge detection approach to fetch edge profile.

At first, apply Robert edge detection operator to extract Basket edges of candidate $Basket(n)$ region. The edge amplitude is $R(i, j)$. It is shown in formula (9).

$$R(i, j) = \left| f(i, j) - f(i+1, j+1) \right| + \left| f(i+1, j) - f(i, j+1) \right| \tag{9}$$

3.2.2. Extraction of Circle Center of Candidate Basketball Region

It’s all known that circle center is the central point of round shape edges. So when region edges are determined, it’s possible to calculate the central point of $Basket(n)$, $n \in [0, MaxR-1]$ boundaries of candidate basket area to extract the circle center; then reckon the coordinate of all edge points $Edge(x_i, y_i)$ and $(Basket(n)_{sum\_x}, Basket(n)_{sum\_y})$ of candidate basketball area $Basket(n)$. It is shown in formula (10).

$$\begin{cases} Basket(n)_{Sum\_x} = \sum_{i=0}^{Max\_n-1} x_i \\ Basket(n)_{Sum\_y} = \sum_{i=0}^{Max\_n-1} y_i \end{cases} \quad (x_i, y_i) \in Basket(n) \tag{10}$$

3.2.3. Extraction of Radius of Candidate Basketball Region

According to the feature that all round contour points are in the same distance to the circle center, count totally the average distances between all edge points of the candidate $Basket(n)$ region to the center circle as to define radius. Next, estimate the distance from all edge points $Edge(x_i, y_i)$ in the area Basket to the circle center $(Basket(n).Circle\_x, Basket(n).Circle\_y)$ and $Basket(n)_{sum\_radius}$. It is shown in formula (11).

$$Basket(n)_{Sum\_Radius} = \sum_{i=0}^{Max\_n-1} \sqrt{(x_i - Basket(n).Circle\_x)^2 + (y_i - Basket(n).Circle\_y)^2} \tag{11}$$

Then, get the average distance from edge point $Edge(x_i, y_i)$ of all regions to the corresponding region center $Basket(n).Radius$. It is shown in formula (12).

$$Basket(n).Radius = {Basket(n)_{Sum\_Radius}} \Big/ {Max\_n} \tag{12}$$

4. Correction of Basketball Region Edges

The interferences like image noise and motion blur lead to deviations in the segmented basketball region.

Here, we utilize the characteristics of gradient peak of basketball edges to correct such divergences. The circle center and radius of candidate $Basket(n)$ region Basket is $(Basket(n).Circle\_x, Basket(n).Circle\_y)$ and $Basket(n).Radius$. We describe the idea of revising basketball edge contours in the instance of one $Basket(n)$ region Basket.

Figure 6 is a modification model. For easier treatment by the program, we need to discretize edges of $Basket(n)$. To do that, by using axis 0 (*i.e.*, direction of arrow) as rotating axis, we discretize edges of round area at the interval θ of rotation direction.

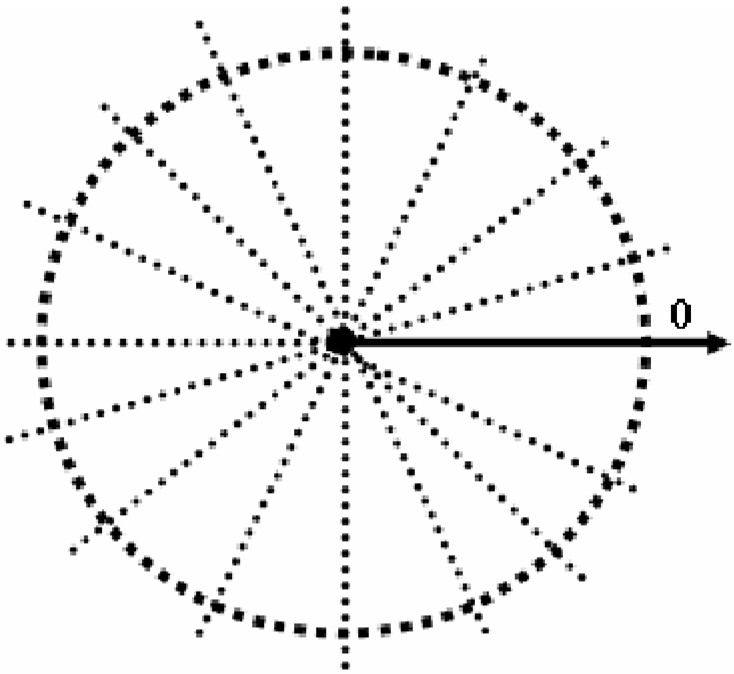


Figure 6. Round Edge Contour Correction

Since in long shots, the basketball radius is small. To save computational workload, the rotation interval $\theta = 5$, that is, basketball boundaries are discretized to 36 dots, of which in the direction φ of one dot (x_0, y_0) , many adjacent dots are selected; then correct the position of the dot based on the feature of edge energy peak. Generally speaking, no severe changes happen to the profile of continuous inter-frame basketball. So it's required to choose appropriate neighboring dots to avoid over-correction of edges. Considering that basketball contour shrinkage Num_S is smaller than the increase Num_W , we choose the neighboring dot coordinate (x_k, y_k) of $Num_S = 2$ $Num_W = 3$. It is shown in formula (13).

$$\begin{cases} x_k = x_0 + k \cdot \cos\varphi \\ y_k = y_0 + k \cdot \sin\varphi \end{cases} \tag{13}$$

5. Identification of Basketball

Normally, only part of the candidate basketball area is actual basketball region. In the paper the improved active contour model is applied to determine the basketball in the candidate region.

5.1. Improved Active Contour Model

The active contour model is renamed Snake, an image segmentation technology proposed by Kass *et al*. in 1987 to treat rigid or not rigid objects. By solving the local extremum of energy function of transformable curve, we can adjust the natural shape of Snake as to agree with the outline of objects. Snake energy includes the internal and external forces, the former limiting the shape of objects, in the role of smoothing profile; the latter guiding Snake to move towards image feature.

The active contour model shows a unified solution to the visual problem of a wide range, no need to create relative model for specific objects. When applied for image cut, the model is superior over traditional methods in terms of accuracy, robustness and practicability.

Make Snake's changeable curve $v(s) = \{x(s), y(s)\}$, of which $s \in [0,1]$, s is the length of normalized curve. The total energy function of Snake. It is shown in formula (14).

$$E_{snake} = \int_0^1 [E_{inside}(v(s)) + E_{contour}(v(s))]ds$$
$$= \alpha \cdot \left|v_s(s)\right|^2 + \beta \cdot \left|v_{ss}(s)\right|^2 + \gamma \cdot E_{image}(v(s)) \tag{14}$$

Snake can't satisfy the need of instantaneity, because apparently, when the deformable curve of Snake approximates gradually the contour of object, computation of its internal energy $E_{inside}(v(s))$ and external energy $E_{contour}(v(s))$ becomes more complicated and it requires iterations, thus the active contour model is challenged with tremendous calculated amount.

Starting from the practical tracing of basketball, we revised the calculation method of Snake energy in formula (14). Firstly, basketball edge contour is round and the contour smoothness is fixed. It's no necessity to compute $v_{ss}(s)$. Make region energy E_{region} to replace the last item in (14). After analysis, we get the improved total energy function of Snake: It is shown in formula (15).

$$E_{snake} = \alpha \cdot \left|v_s(s)\right| + \beta \cdot \oint_S E_{region} ds \tag{15}$$

5.2. Fetch of Regional Optimal Solution

Generally very few candidate basketball region $Basket(n)$, $n \in [0, MaxR - 1]$ is the true basketball region. So it's required to build the rule for fetching regional optimal solution. Additionally, basketball temporally disappears in the video for being covered by the player. To acquire the solution should take into account circumstances when basketball doesn't show in the video. By calculating Snake energy in candidate basketball region, we confirm the effectiveness of regional segmentation.

For convenient computer processing and real demands, we need to discretize round edges. We design 36 of edge discrete point (x_k, y_k) based on interval angle $\theta = 5$.

Firstly, calculate gradient energy $E_{grad}(n)$ of $Basket(n)$; It is shown in formula (16).

$$E_{Grad}(n) = \sum_{k=0}^{35} \left|I(x_k, y_k) - I(x_k + 1, y_k + 1)\right| + \left|I(x_k + 1, y_k) - I(x_k, y_k + 1)\right| \Big/ 36 \tag{16}$$

Secondly, sum up the number of dots conforming to basketball hue model $Model_{Hue}$ in the region $Basket(n)$, which is defined as $Basket(n)$ energy $E_{region}(n)$; then, standardize them, It is shown in formula (17).

$$E_{region}(n) = 100 \cdot E_{region}(n) \Big/ N_{region}(n) \tag{17}$$

as to calculate the total energy function of Snake; It is shown in formula (18).

$$E_{snake}(n) = \alpha \cdot E_{Grad}(n) + E_{region}(n) \tag{18}$$

Since the weight of results regarding gradient energy $E_{Grad}(n)$ and regional energy $E_{region}(n)$ is not identical, we need design α to make $E_{snake}(n)$ achieve the best result. As it's seen, $E_{snake}(n)$ is bigger, indicating probably the region is basketball object. So we design threshold σ to obtain the optimal solution and decide whether basketball exists in the video. We employ equation (19) to identify if the region is basketball area.

$$\begin{cases} Basket(n): basket & E_{snake}(n) \geq \sigma \\ Basket(n): non-basket & E_{snake}(n) < \sigma \end{cases} \tag{19}$$

6. Experimental Results

Make $\alpha = 1.0$ and decision threshold $\sigma = 100$. In Figure 7, white circular area is the candidate basketball region. The relevant parameters of two areas are shown in Table 1.

Table 1. The Candidate Basketball Regional Parameters

The candidate regional parameters	Region 1	Region 2
Center and radius	(876,256),6	(98,289),6
Regional energy	98	50
Gradient energy	54.3	45.7
total energy function	147.5	97.5
Region semantic	basket	Players arm

They are connected with each other. However we didn't make right the weight α and decision threshold σ. The fetch result was not satisfactory. In the future work, we will have to add testing sample to set up properly α and σ. Figure 7 (b) is the result of final segmentation of basketball.

(a) Candidate Basketball Area (b) Optimal Solution Extraction

Figure 7. Extraction of Regional Optimal Solution

7. Conclusion

The paper analyzed the segmentation of basketball in videos. In line with different basketball features in long-shot videos, we designed relative methods for basketball segmentation. For cutting of basketball in long shots, we used Gauss filter to smooth image noises; then split up background areas with the use of inter-frame difference; next, we devised “basketball recognition strategy” based on basketball region features, by which, the candidate basketball region was conserved; to further on, we corrected edge deviations by referring to peak features of edge energy. Finally, we reached the optimal solution of the candidate basketball region by the advantage of improved active contour model.

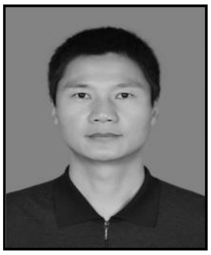
Acknowledgments

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

References

[1] Z. Guangyu, “Research on sports video content analysis based on the information of player behavior”, Harbin Institute of Technology, **(2009)**.
[2] H. Ye, “Graph theory based image segmentation technology research”, Xi'an Electronic and Science University, **(2011)**
[3] L. Peng, “Target detection and tracking in soccer video”, Nanjing University of Science and Technology, **(2006)**.
[4] S. Jian, “Study on the technology of video human objects segmentation”, Shanghai Jiao Tong University, **(2007)**.
[5] W. Yunfeng, “Research on video object segmentation technique”, Guangdong University of Technology, **(2014)**.
[6] L. Jun, “Study of low quality fingerprint image segmentation technology”, Zhejiang Sci.-Tech University, **(2013)**.
[7] N. Z. Xing, “Soccer video modeling and content analysis method research”, Xi'an Electronic and Science University, **(2012)**.
[8] H. Xiong, “Research and implementation of video content structuring techniques”, China PLA National University of defense technology, **(2001)**.
[9] L. Yue, “Research on video shot detection and classification based on content”, Shandong Normal University, **(2010)**.
[10] B. Jiang and S. Y. Lao, “White light, TOLLARI Sabrina, MARSALA Christophe., a real time recognition algorithm for sports video advertising brand”, Journal of automation, vol. 04, **(2011)**, pp. 418-426.
[11] N. Zhenxing, L. Jie and G. X. Bo, “The semantic color extraction and semantic shot segmentation for soccer video”, Journal of Xi'an Electronic and Science University, vol. 04, **(2010)**, pp. 613-618.
[12] Y. Junqing and H. Yue, “Detection and recognition of video film scene scale”, Computer applications, vol. 12, **(2009)**, pp. 3422-3426.
[13] L. Dian, “Generation of soccer video slow motion detection and video summary”, North China University of Technology, **(2008)**.

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.

