International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.1 (2016), pp.255-266
http://dx.doi.org/10.14257/ijmue.2016.11.1.25

Research on Close Shot Segmentation in Sports Video

Yang Wang[1], Jingmeng Sun[1,*] Yifei Liu[1] and Yueqiu Han[2]

[1] *Physical Education Department, Harbin Engineering University, Harbin 150000, China*

[2]*College of International Cooperation Education, Harbin Engineering University, Harbin 150000, China*

aoled@126.com

Abstract

The basketball segmentation in close shots is studied. According to the consistency of basketball color, color segmentation is implemented. Then with the help of steerable filters, the method detects image edges and the orientations of them. Finally, based on edges and their orientations, Hough transform is improved to get the circle center and the radius of basketball.

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

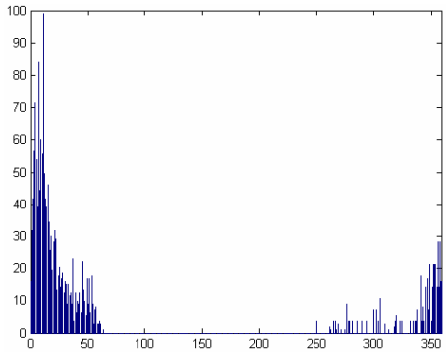
1. Introduction

In the basketball game field, plentiful cameras are working simultaneously. As 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].

Feature of close shots. Figure 1 (a) is a close shot of basketball image, which features analysis found that: firstly, the basketball direction of edge radius is bigger, Figure 3-1 (b) edge extraction for basketball, Sobel edge detection operator obviously, basketball keeps the circular contour; secondly, texture clear and basketball the same shape; detailed information on the signs, such as basketball text lines, destroy the consistency of color distribution, regional basketball tone distribution as shown in Figure1(c); Finally, the basketball movement characteristic is not obvious.

(a)Close Shot (b)Close Shot Image Edge

ISSN: 1975-0080 IJMUE



(c) Close Shot Basketball Hue Histogram

Figure 1. Features Close Shots in Basketball

2. Basketball Segmentation Algorithm in Intra-Frame Close Shots

Firstly, count basketball hue model to finish color segmentation and use steerable filter to make directional edge detection and determine edge points and direction of edge normal according to peak features of edge gradient; next, employ edge points and direction to improve Hough transform and build circle parameter plane of multi-scale radius to search peak points and extract possible circle center; finally, as per the peak value of histogram of radius distribution, determine the optimal radius in multiple scales and respective circle center to finish segmentation of basketball.

3. Color Segment

Color segment means determining possible basketball area as per color features [5]. We choose proper color space as to depict quantitatively image color information.

3.1 Color Space and Conversion

Color space is of a variety, used in different cases for different properties. So far there're two kinds of color space:

(1) RGB color space rather for color display device or printer;

(2) The type aiming at color processing, typical of HSV color space. We'll discuss RGB and HSV later.

3.1.1 RGB Color Space

Trichromatic theory is that almost all colors can be regarded as different combination of the three basic colors: red, green and blue.

$$F = \alpha \cdot R + \beta \cdot G + \gamma \cdot B \quad (1)$$

For that, RGB color space is obtained most easily. RGB color model is a one-unit cube in the three-dimensional orthogonal coordinate color system [6]. It is shown in Figure 2. In the original points in RGB color space, no basic color has illumination, *i.e*. all original points are black. Three basic colors become white only after they reach the maximum brightness. Three basic colors of equivalent low brightness form grey tone. All those points are on the diagonal line of color cube. The diagonal line is called grey line. In the color cube, three angles refer to three basic colors, *i.e*. respectively red, green and blue. The other three angles refer to secondary colors *i.e*. yellow, cyan (blue-green) and pinkish red (pink) [7].

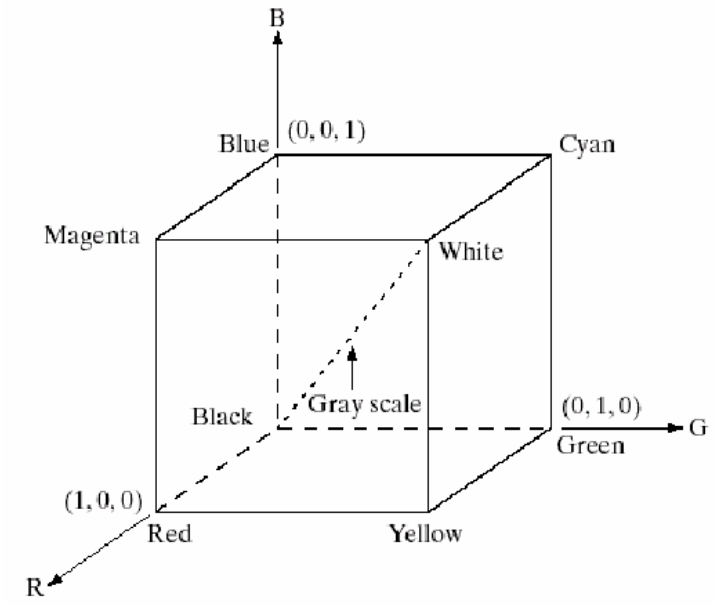


Figure 2. RGB Color Space Model

RGB color space is mostly used in video capture device, CRT monitoring equipment and other raster graphic device. Its color distribution is not uniform, impossible to measure with Euclidean distance the chromatic aberration of two colors in chromaticity diagram [8], *i.e*. distance in RGB color space doesn't represent the color similarity in human vision and correlation exists in the three color components of RGB. So for color processing, it's not advisable to use directly RGB color space.

3.1.2 HSV Color Space and Conversion

HSV color space represents color signal as these three components: hue, saturation and brightness. Hue is what's generally called color, determined by the optical wavelength of mixed spectrum; saturation refers to hue purity of color; brightness is the degree of color light and shade. From the point of color's spectrum, brightness means the intensity of light.

HSV color space is of upside-down cone shape; vertical axis is brightness V and horizontal axis is saturation S; hue H is rotation angle around axis V, as shown in Figure 3. Apparently, H is in the range [0-360]; red corresponds to 0; green to 120 and blue to 240; S and V is in the range 0-1. Of that, the apex of cone is corresponding to black; the center of conical top surface (*i.e*. V=1, S=0, H undefined) corresponds to white color.

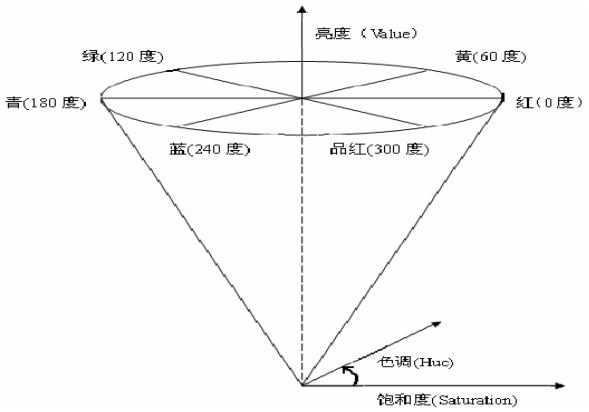


Figure 3. HSV Color Space

Make three components of RGB color space respectively R, G and B; HSV's respectively H, S and V. The conversion formula is put as follows [9]:

$$
\begin{aligned}
&V = \max(R,G,B)\\
&S = mm / V \quad mm = \max(R,G,B) - \min(R,G,B)\\
&r = \frac{V-R}{mm}, g = \frac{V-G}{mm}, b = \frac{V-B}{mm}\\
&h = \begin{cases} 5+b & R = \max(R,G,B) \ \ and \ \ G = \min(R,G,B)\\ 1-g & R = \max(R,G,B) \ \ and \ \ G \neq \min(R,G,B)\\ 1+r & G = \max(R,G,B) \ \ and \ \ B = \min(R,G,B)\\ 3-b & G = \max(R,G,B) \ \ and \ \ B \neq \min(R,G,B)\\ 5-r & B = \max(R,G,B) \ \ and \ \ G = \min(R,G,B)\\ 3+g & Otherwise \end{cases}\\
&H = h.60^{\circ}
\end{aligned}
\tag{2}
$$

HSV color space has these merits:

(1) HSV color space is directly corresponding to three elements of human's color vision features, which are independent;

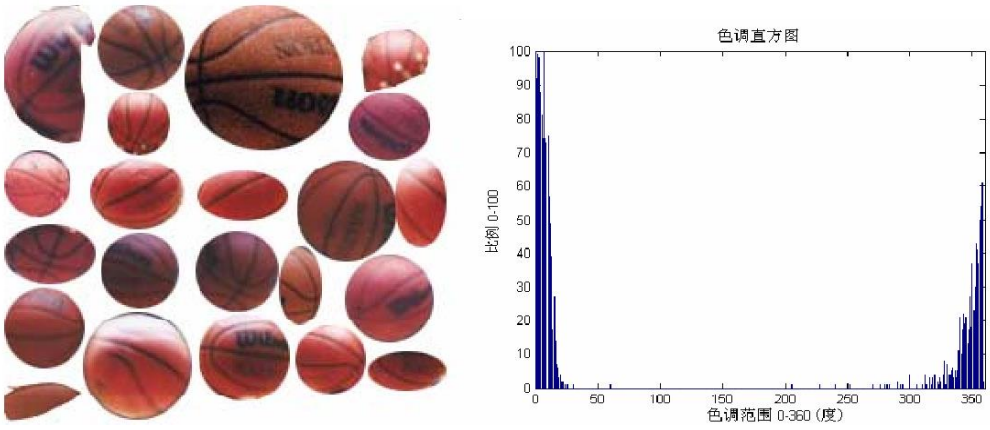
(2) HSV color space is one well-proportioned color space; perceivable color difference is directly proportional to Euclidean distance of points in HSV color space coordinate [10].

Although HSV color space accords with human visual features and each component is independent, each component is not all very stable. Generally, H is the least stable; S is less stable. S but H changes along with light, very slightly. Component H is the most stable. In normal cases, H is closely of normal distribution and variance is very small, but H changes obviously with light.

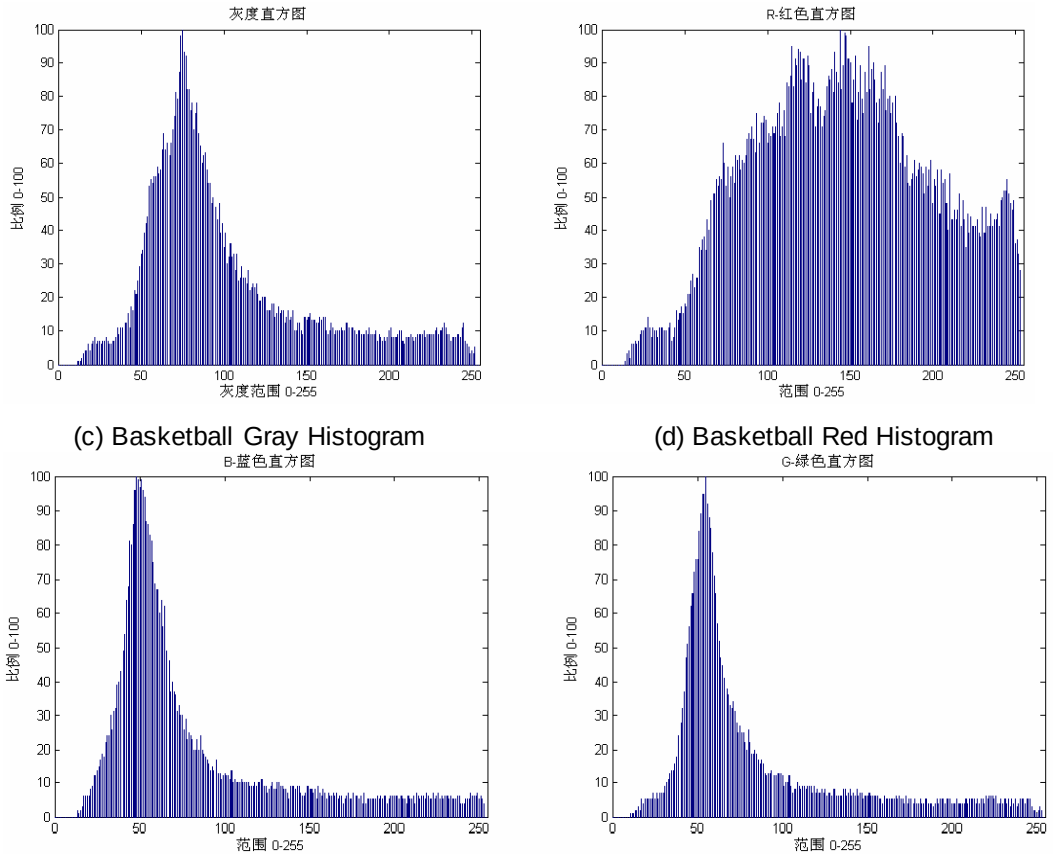
3.2 Selection of Color Space and Statistics of Basketball Color Features

Studies found that RGB space approximates human eye's quantification of colors. So it's extensively used in color imaging device and display equipment. But in basketball segmentation, RGB space is not the optimal color model.

Figure 4(a) is basketball legend under different lights. Count its hue histogram, grey scale histogram, red histogram, green and blue histograms. Clearly, only hue has robustness to light, with good clustering performance. It is shown in Figure4 (b)-(f). Hence, we take HSV color space to make segment of hue.



(a) Different Illumination Basketball (b) Basketball Hue Histogram



(c) Basketball Gray Histogram (d) Basketball Red Histogram

(e) Basketball Blue Histogram (f) Basketball Green Histogram

Figure 4. Basketball Color Clustering

3.3 Implementation of Color Segmentation

Through statistics and analysis of a certain number of basketball color characteristics, we find basketball hue range basically satisfies:

$$H_{Model} \in [(0,30) \bigcup (330,359)] \quad (3)$$

Make hue value j of each pixel in original image. Color segmentation process is as follows:

(1) Coarse segmentation of color

Use basketball hue model in equation 3 to maintain possible basketball area in primary color; other irrelative area is made white color, it is shown in equation 4

$$Hue[i][j] = \begin{cases} Hue[i][j] & Hue[i][j] \in H_{Model} \\ 255 & Hue[i][j] \notin H_{Model} \end{cases} \quad (4)$$

It is seen in Figure 5(b). Evidently, most non-basketball region dots are removed, but meanwhile there're many "fragments", which are reserved due to their hue is similar to basketball area. On contrary to basketball area, such "fragment" areas are very small. They are eliminated based on the statistical number of pixels in such areas.

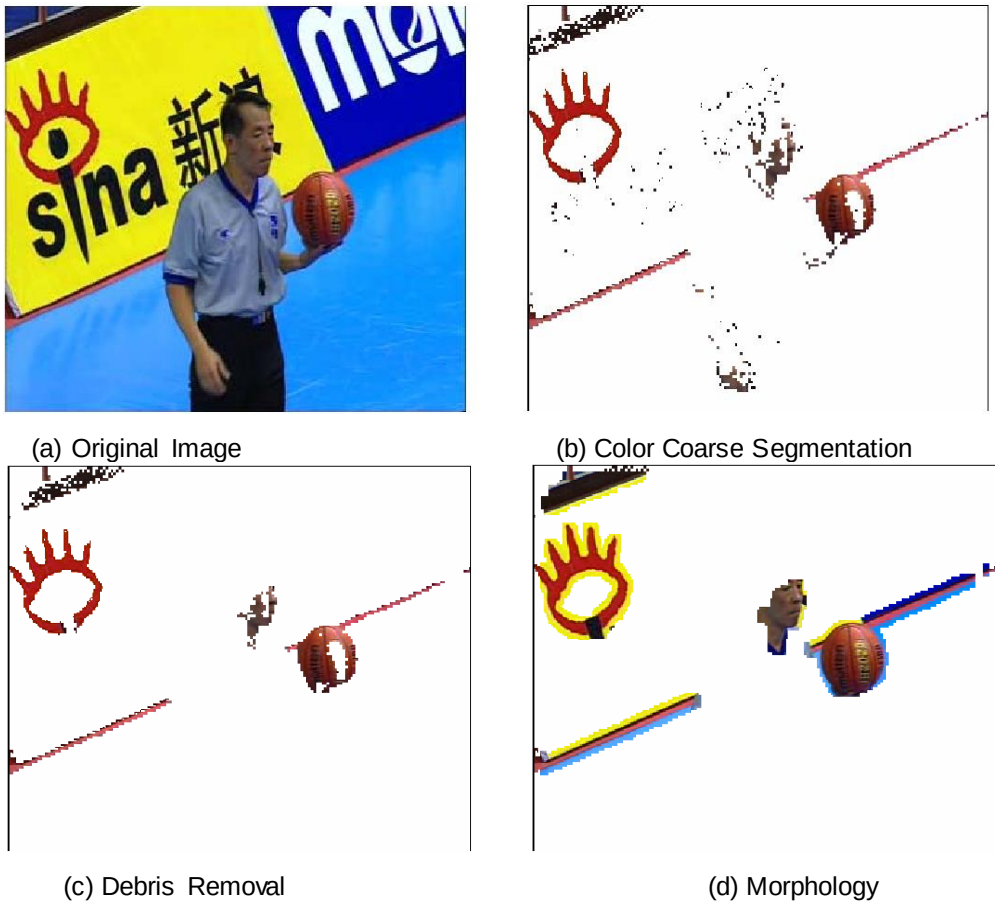


Figure 5. Basketball Color Segmentation

(2) Removal of fragments

Sum up the pixel number $Num_{adjacency}$ included in each connected area $R_{adjacency}$ in Figure 5(b). We learnt that "fragment" areas are generally smaller than 50 pixels. At this moment, if Num in $R_{adjacency}$ is below 50, remove it (make white), *i.e.*

$$Hue[i][j] = \begin{cases} Hue[i][j] & (Num_{adjacency} < 50) \bigcap ((i,j) \in R_{adjacency}) \\ 255 & (Num_{adjacency} \geq 50) \bigcap ((i,j) \in R_{adjacency}) \end{cases} \quad (5)$$

(3) Morphological processing

With regards to the hollow area of object in Figure 5(c), we utilize morphological expansion operator to fill up the hollow area. The processing result is put in Figure 5(d). Noticeably, the object area is completely retained and most non-basketball areas in the image are removed. Hence, the effect of background area is reduced on basketball segment, and also the number of pixels for processing is decreased. The algorithm speed is thus improved.

4. Directional Edge Detection Based on Steerable Filter

4.1 Edge Detection Algorithms and Performance Comparison

Traditional image edge detection algorithms fetch edge information from image's high-frequency components. Differential operation is a main way for edge detection and extraction. These two differential operations are often adopted:

(1) First-order difference operation, *i.e*. gradient operator, typical of Robert operator, Sobel operator, PreWitt operator and Canny;

(2) Second-order differential operation with Laplacian operator as representative.

Here we introduce common edge detection operators and compare their performance.

4.1.1 Edge Detection Operator

Robert edge detection operator, *i.e*. cross-over gradient operator, is calculated by the equation:

$$R(x,y)=\sqrt{f_x^{\,2}+f_y^{\,2}} \qquad \begin{cases} f_x = f(x,y)-f(x+1,y+1) \\ f_y = f(x,y+1)-f(x+1,y) \end{cases} \tag{6}$$

Sobel edge detection operator, the calculation formula is as follows:

$$S(x,y)=\sqrt{f_x^{\,2}+f_y^{\,2}}$$

$$\begin{aligned} f_x &= [f(x-1,y-1)-2f(x-1,y)+f(x-1,y+1)]-[f(x+1,y-1)-2f(x+1,j)+f(x+1,y+1)] \\ f_y &= [f(x-1,y-1)-2f(x,y-1)+f(x+1,y-1)]-[f(x-1,y+1)-2f(x,y+1)+f(x+1,y+1)] \end{aligned} \tag{7}$$

Canny edge detection operator [11], *i.e*. uses the optimized numerical method to get the optimal edge detection template of specified edge type. The operator has good signal to noise ratio and detection precision so that it's applied widely. Canny edge detection is carried out in the following steps:

(1) Use Gaussian filter to smooth images, *i.e*. choose appropriate Gaussian filter function $G(x,y)$'s standard variance σ and domain size; employ $G(x,y)$ to perform convolution operation of image $f(x,y)$ to get smooth image $s(x,y)$ S; smooth filter can inhibit effectively noises;

(2) Use gradient like Prewitt operator to get the partial derivative G of image gray in two directions as to obtain gradient size |G| and direction θ;

$$|G|=\sqrt{G_x^{\,2}+G_y^{2}} \qquad \theta = Arc\tan\left(\frac{G_y}{G_x}\right)$$

(3) Make non-maximum suppression of gradient magnitude; if the edge gradient $G(x,y)$ G of image pixel (i,j) in image $f(x,y)$ is smaller than edge strength of two adjacent pixels in gradient direction, it's thought that (i,j) is not edge point and should be removed; non-maximum suppression images still have lots of fake edges caused by noises and fine textures, which require further detection and connection of edges by dual threshold algorithm.

(4) Detect and link edges with dual threshold algorithm

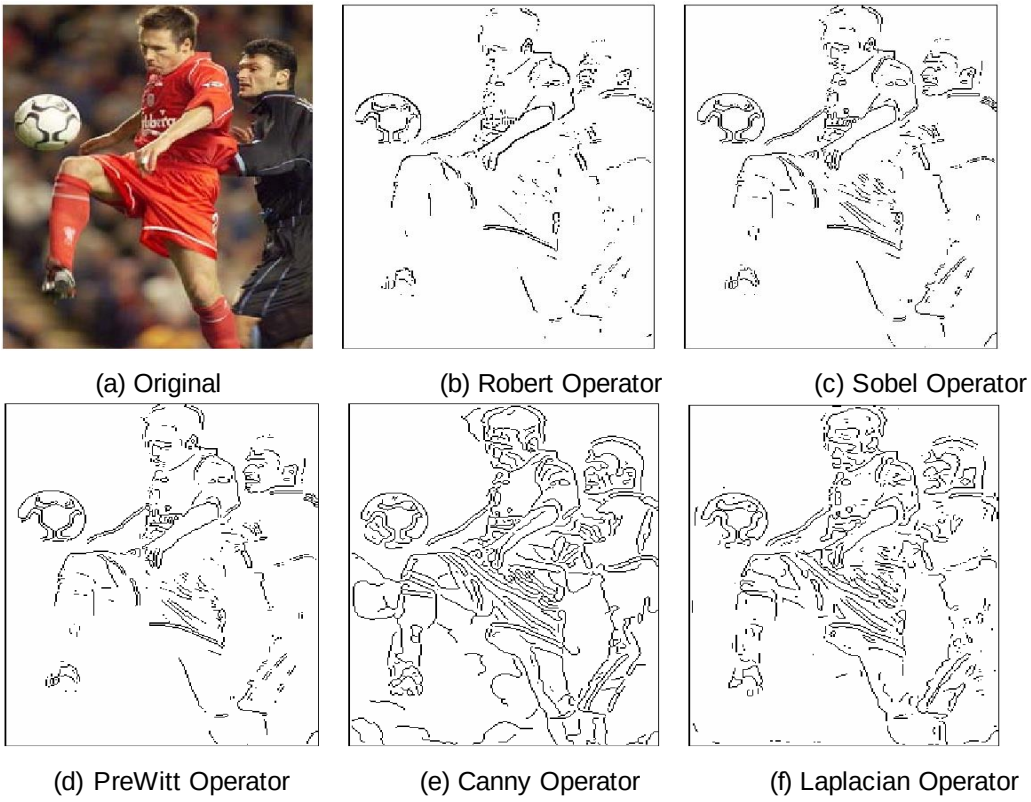
Dual threshold algorithm [12] sets threshold τ_1 and τ_2 $(\tau_2 > \tau_1)$; and sets dual threshold for non-maximum suppression images to acquire high-threshold and low-threshold detection results. Then, in high-threshold detection results, join up edge contour and link to end points; in low-threshold detection results, search edge points till all gaps of high-threshold detection results are linked up. Fixed threshold value can't meet practical needs. On the basis of dual threshold algorithm, we use dynamic threshold method to overcome the shortcoming of missing edges by Canny algorithm when detecting edges of images of uneven gray scales. The method has ability of detecting fuzzy edges.

Laplacian operator is second-order differential edge operator with no relation to edge direction. It's defined as:

$$\nabla^2 f(i,j)=f(i+1,j)+f(i-1,j)+f(i,j+1)+f(i,j-1)-4f(i,j) \tag{8}$$

4.2 Performance Comparison of Edge Detection Operators

We use all the above edge detection operators to treat Figure6 (a). The result image is shown in Figure 6 (b) - (f)



(a) Original (b) Robert Operator (c) Sobel Operator

(d) PreWitt Operator (e) Canny Operator (f) Laplacian Operator

Figure 6. Getting the Edge Detection Results of Different Image Operator

After analysis, we find Robert operator is simple, with detection performance of edges in vertical and horizontal directions better than in oblique orientation and higher positioning precision of detection. However, it's susceptible to noises. Sobel operator can bring about better edge detection effect and can smooth noises. But, the resulted edges are too coarse and positioning precision of detection is lower. Prewitt, Robinson and Kirsch operator are similar to Sobel operator. Canny operator is obviously superior over other first-order differential edge detection operators in terms of positioning precision and noise immunity. The edge extracted by Laplacian operator with second derivative for zero cross-over is one pixel unit. The obtained results require no refinement and edge positioning precision is guaranteed. But second derivative operators are affected more easily by noises than first-order derivative operators.

In the extraction in edge directions, Laplacian operator is second-order differential edge operator with no relation to edge directions. So it can't reflect directional continuity of basketball edges. Although first-order differential operators like Sobel, Prewitt and Canny can get edge gradient direction, but the calculation is not accurate. Hence, with direct application of the above edge detection operator, it's impossible to take full advantage of the feature that in close shots, basketball has good edge directionality. For that, we use steerable filter which can accurately calculate edge direction to detect directional edges.

4.3 Steerable Filter

Steerable filter uses basic filters for linear combination to realize filter bank structure of any directional filtering. Being capable to rotate freely, the filter can detect accurately

features of target images such as edges, veins and singular points. Figure 7 gives the common structure of steerable filter.

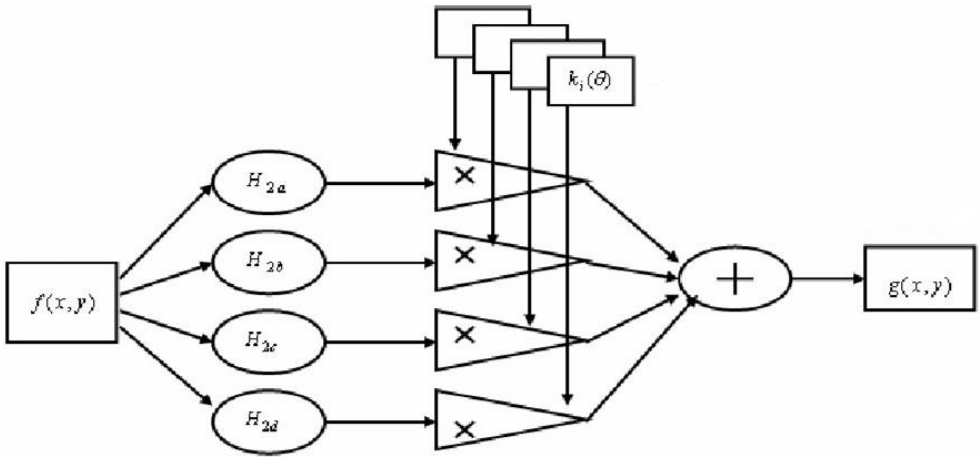


Figure 7. The Common Structure of Steerable Filter

To accord to different edge features, steerable filter has multiple forms. It's found that Gaussian function derivative filter apparently has peak value in the current pixel, fit for extracting isolated points and lines. Gaussian function derivative Hilbert transform filter is of staircase distribution, applicable to extract ladder-like edges. With increase of Gaussian function derivative order, the filter has obvious oscillation. Comparatively, second order filter is more stable.

Considering that basketball's internal characteristics are similar and its edges are of stair shape, we choose Gaussian function second-order Hilbert steerable filter $H_2^{\theta}(x,y)$. In the Figure3, $H_{2a}, H_{2b}, H_{2c}, H_{2d}$ is a group of basic filter; $k_a(\theta), k_b(\theta), k_c(\theta), k_d(\theta)$ is a group of weighted coefficients relating to angle θ. Figure 8 is three-dimensional illustration in $H_2^{\theta}(x,y)$ direction in different θ. Apparently, with changing θ, $H_2^{\theta}(x,y)$ direction varies.

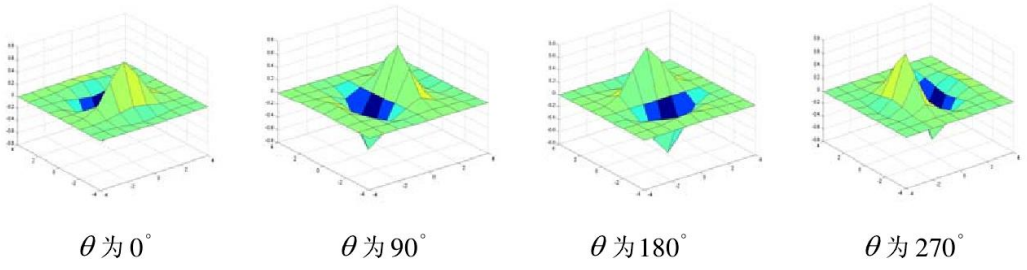


Figure 8. The 3D Graphic of $H_2^{\theta}(x,y)$ Direction in θ

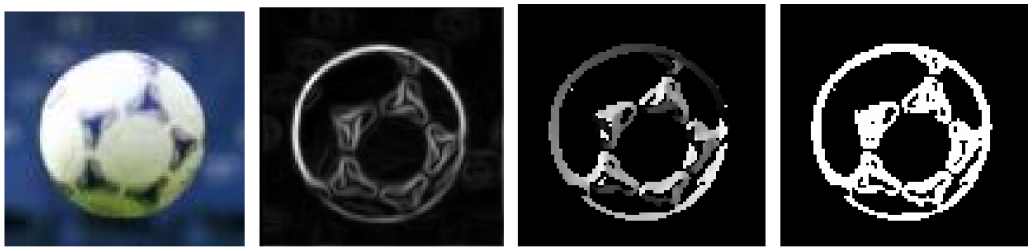
4.4 Detection of Directional Edges

We use Gaussian function second-order Hilbert steerable filter $H_2^{\theta}(x,y)$ to make directional edge detection and meanwhile get image edges and direction, as follows:

For every pixel (x_i, y_i), i=0,1,…,M-1,j=0,1,…,N-1, in the image, we can get its gradient energy $Grad(x_i, y_i, \theta)$ in each direction by steerable filter.

$$Grad(x_i, y_j, \theta) = \left| f(x_i, y_j) * H_2^{\theta}(x_i, y_j) \right| \quad \theta = 0,1,2,...,359 \tag{9}$$

Figure 10 is the directional diagram of the image and its edge energy and edge normal. In Figure10(c), edge normal angle $\theta(x_i, y_i)$ magnifies gradually with round contour edges and it changes in the range [0-359]. It means that steerable filter can fetch well directional features of image edges.



(a) Original Image (b) Edge Energy (c) The Normal Direction (d) Edge Points

Figure 10. The Direction and Edge Point Detection, Energy Direction Variable Filter

According to peak feature of edge energy, we can define edge points and have edge image $Edge(x_i, y_i)$. The extraction is done in following steps:

(1) Set threshold C_{th}; if point (x_i, y_i)'s edge energy $E E(x_i, y_i) \geq C_{th}$, the step length of edge normal angle $\theta(x_i, y_i)$ along with (x_i, y_i) is 1 and we can decide eight neighboring points (x_k, y_k); noticeably, k=0 corresponding to current pixel (x_i, y_i);

(2) Calculate $E(x_k, y_k)$, k=-4,-3,…,0,…,3,4 where, non-integer pixel's energy is obtained by the linear interpolation of energies of neighborhood pixels;

(3) Modify edge image values of related pixels.

$$Edge(x_i, y_j) = \begin{cases} 1 & E(x_i, y_j) = \max\{E(x_{i+k}, y_{j+k}),\ k = \{-4,...,4\} \\ 0 & otherwise \end{cases} \tag{10}$$

5. Experimental Results

Make $\alpha = 1.0$ and decision threshold $\sigma = 100$. In Fig11, white circular area is the candidate basketball region. The relevant parameters of two areas are shown in Table1

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 11 (b) is the result of final segmentation of basketball.

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

Figure 11. Extraction of Regional Optimal Solution

6. Conclusion

The paper analyzed the segmentation of basketball in videos. In line with different basketball features in close-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.

Acknowledgement

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] R. C. Gonzalez and R. E. Woods, "Digital Image Processing (second Edition)", Prentice Hall, **(2002)**.
[6] K. R. Castleman, "Digital image processing", electronic industry press, Beijing, **(1998)**.
[7] X. Xu, "System research", image retrieval based on visual features of Zhejiang University, Hangzhou, **(1999)**.
[8] J. R. Smith, "Integrated spatial and feature image systems: Retrieval, compression and analysis", Ph.D. dissertation, Columbia Univ., New York, **(1997)**.
[9] Z. Ming, "2D visual object segmentation", Zhejiang University, Hangzhou, **(2004)**.
[10] J. Canny, "A Computational Approach to Edge Detection", IEEE Trans PAMI-8, vol. 6, **(1986)**, pp. 679-698.
[11] Meer P. and Georgescu B., "Edge detection with embedded confidence", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 23, no. 12, **(2001)**, pp. 1351-1365.
[12] W. Li and F. Ming, "A Canny algorithm improved edge detection method based on computing technology and automation", vol. 22, no. 1, **(2003)**, pp. 24-26.
[13] W. T. Freeman and E. H. Adelson, "The Design and Use of Steerable Filters", IEEE Trans. on PAMI, vol. 13, no. 9, September **(1991)**, pp. 891-906.

Author

Wang Yang, He is a lecturer at Physical Education Department of Harbin Engineering University. He is in the research of Sport Economy.