SHOT DETECTION

a) Algorithm Choosing

Noticing that histogram-based algorithm works well on shot detection, I decided to use it to solve the problem. To figure out which specific histogram-based algorithm is better, I have tried the following methods:

1) Histogram

One threshold is used. An xx-bin (the amount of bins varied in different color spaces) histogram is computed over the entire frame. The difference measure is the sum of the absolute bin-wise histogram differences. A shot boundary is declared if the histogram difference between consecutive frames exceeds a threshold.

Reference:

1. *"Comparison of video shot boundary detection techniques" J.S. Boreczky, L.A. Rowe;*
2. *MMAI Slides.*

Color space:

RGB gray-level :	64 bins
RGB :	4*4*4 bins
HSV :	18*3*3 bins
YIQ:	16*4*4 bins

2) Region Histogram

Two thresholds are used. Each frame is divided into 16 blocks. An xx-bin histogram is computed for each region. Histogram differences are computed (use the same evaluation method in histogram) for each region between consecutive frames. If the number of region differences that exceed the difference threshold is greater than the count threshold, a shot boundary is declared.

Reference:

1. *"Comparison of video shot boundary detection techniques" J.S. Boreczky, L.A. Rowe;*
2. *MMAI Slides.*

Color space

RGB gray-level :	64 bins
RGB :	4*4*4 bins

3) Dealing with Gradual transition

The frames in gradual transitions in the testing video data have great differences between each other (even greater than the cut transition threshold), so the running method mentioned in reference 1 (Twin Comparison in reference 2) is useless in the experiment. To duplicate too many frames detected in one gradual transition, I simply choose the last frame if there are a few consecutive frames.

b) Algorithm Comparison

Non-region VS. Region

According to reference 1, the region method is better than non-region method. But, in my experiment, I found that since region method need 2 thresholds, it's difficult to adjust the thresholds to get the best result. And it's a little sensitive to movements in some cases. In addition, the non-region method is obviously faster than the region one.

Color Space Comparison

According to my experiment results, the gray-level histogram is the simplest and fast one and works not bad (even the worst among all the methods I have tried, it got both 100% precision and recall on video 1&3). Compared to 2D color space, 3D color space gets apparently better performance on all the testing videos. And HSV and YIQ color space work a little better than RGB. The difference between HSV and YIQ is hard to detect.

Conclusion

After comparing the results of all these algorithms, taking the accuracy and efficiency into consideration, I will choose the HSV or YIQ histogram method to solve this problem. However, I noticed that all the algorithms work poor in cases such like light changing, fast moving, bombing and etc. in one shot (e.g. video2, 5 and 7) since both color and object of the consecutive frames will have great change in these situations.

C) Genre-independent or dependent

	Genre	Transition Types	Frame Counts	Average Shot Length (Frame)
1	News	Cut	829	69
2	Anime	Fade out	1079	37
3	MV	Cut	1199	75
4	Anime	Fade in & out	1019	113
5	Anime	Cut	900	32
6	Trailer	Dissolves	930	37
7	Ad	Cut	1767	44
8	MV	Dissolves	1409	30

According to the table, we can see that the transition type of the video is genre-independent, and the shot length is also genre-independent.

The performance and the threshold of my algorithms are genre-independent but depend on transition types. They work well on all the videos but work better on cut transition than gradual transition. And the thresholds of videos in same transition type are almost same.