Sports Video Structure Analysis and Feature Extraction

Maohua Zhuang
Harbin Institute of Sports, Harbin 150000, China
1023054464@qq.com

Abstract. How to help people to find their favorite sports in the massive video, in order to achieve this goal in the finite state machine (FSM) theory based on in-depth research. Combined with the actual situation of this paper, first, we determined the reasonable FSM model. Then, a fast robust global motion estimation algorithm is used to estimate the global motion of the video sequences, and the foreground is separated from the background by motion compensation.

Keywords: Sports Video, Finite-state Machine, Field Knowledge

1 Introduction

For video programs about sport competitions, TV program producers concluded a complete set of scientific and reasonable production and editing mode based on many years of experience in broadcasting sport matches [1-3]. Almost all videos of sport contests are produced by following that pattern, because only in that mode, can TV audiences enjoy watching sport games to the greatest extent [4-6]. The typical video program compiling model leads to typical sport competition video structure. That makes it possible for us to conduct semantic-based analysis, retrieval and query of sport videos. And thus it’s a significant task to analyze the structural features of sport videos [7].

Like other video data, jumping video data are enormous. For the purpose of effective organization of related videos, it’s required to decompose videos to elementary units. It’s generally accepted that the basic physical unit of video is shot [8-9]. One shot is composed of numerous frame images which are acquired consecutively in time by one camera. The detection of shots is an issue of segmenting videos from the perspective of time domain. To edit shots in different ways, they can join up to form video programs. Different video programs have own uniqueness [10-11]. So it’s necessary to use different methods of shot segmentation for various video programs, to realize the sound decomposition of videos. Then on that basis, further analysis is made in order to perform nonlinear browsing and semantic-based inquiry and retrieval [12-13].
Studies on methods for detecting jumping sport video shot boundaries

The accurate detection of shot boundaries is foundation to make subsequent detection with FSM, which can be considered as pre-treatment. Shot boundary detection is discussed for too many years. Many reliable approaches were proposed to detect abrupt change shots, which, however, proved defective, for example, when fast-moving object in the video frame, or for the explosive scene, the existing detection methods are not effective. The detection methods for gradual change shot need improvements as well. For sport videos, shot boundary detection problem has not been solved, for the reasons as follows:

In sport action video images, there’re two kinds of motion: overall motion, i.e. background movement caused by camera motion; partial motion, i.e. foreground movement caused by sportsmen. The acquisition of accurate global motion parameter is key and foundation to foreground and background separation and sportsman body extraction and movement analysis. Here we use global motion estimation to execute foreground and background split of video frames. Considering characteristics of sport motion videos, the global motion of background caused by camera movement is expressed by parameter affine motion model as:

\[
\begin{aligned}
    x &= ax^t + by^t + e \\
    y &= cx^t + dy^t + f
\end{aligned}
\]  

Current video motion object extraction methods are simply of two types: one based on sequential attribute, cutting motion objects as per video sequential property; the other based on spatial attribute, segmenting motion objects as per image zone or edge information. However, no matter which is used to segment motion objects, visible background and irregular movements of objects will lead to reduction of segmentation accuracy, because both methods split up motion objects and background area by means of motion information. But in movement analysis, static foreground area is easily falsely detected as foreground or background due to manifested background and object’s irregular movements, downgrading the precision of segmentation.

We introduce a new method for the fetch of moving objects based on dynamic background construction. Firstly, the dynamic background construction technology based on foreground separation makes use of multi-frame differences to construct the present background; then by background subtraction, it splits out motional object as to remove noticeable background in segmentation result. Meanwhile, it detects static foreground area as per sequential information and merges it to target area as to get the complete object area, overcoming the impacts of object’s random movements on
segmentation accuracy. Finally, regarding edges of object area as initial position, it applies active contour model which uses color gradient as external energy to get precise profile of moving objects.

Let $I_k$ for the current frame, $I_i(i = k - L, \ldots, k + L)$ is $2L+1$ frame continuous image, the global motion parameters between adjacent frames are $\theta_{k-L+1}, \ldots, \theta_{k+L, k+L-1}$. In order to construct $I_k$ background. To calculate the parameters $\theta^i$ of spatial coordinate theta is aligned to the $I_k$ on $I_i$.

3 Experiment Design and Discussion

Experimental data was collected from television recording sport programs. The video database is very challenging. It lasts 2hours and forty five minutes, totaling 3514 shots and 226936 frame images, including advertisements, sport news, and some different sport competition program fragments. Some are similar video clips like news titles, advertisements; some are repetitive video clips like different car racing matches, same ads in different duration and editing.

The feature library includes two visual features like color and texture and high-level semantic features of motion information included in the moving objects. Low-level features like color are expressed by main color histogram and accumulative histogram, motion information of moving objects in key frames extracted as high-level semantic features of video sequence, which is further divided into video fragments with semantic concepts. To extract motion information, the global movement estimation Konrad algorithm and exterior point filtering algorithm based on Fisher linear discriminant criteria.

The inquiry based on FSM is FSM template established respectively for various sport competition. When query request is sent, inputting video clips to according sport competition FSM will help user find interesting sport program.

Used maximum matching and optimal matching to achieve similarity measurement of video clips

Choose images with target as a group of relative images; then calculate recall and precision ratio based on return results; the higher the recall and precision ratio is, the better performance the retrieval algorithm realizes.

4 Conclusion

In this paper, we first introduce the data structure of sports video, and the video data can be divided into four levels: video, Scene, Shot and Frame. Then, the research results of other researchers are presented, and the segmentation method of motion scene recognition based on global motion is proposed.
References