MATEC Web of Conferences 22, 01010 (2015)
DOI: 10.1051/matecconf/20152201010

An Algorithm of Extracting I-Frame in Compressed Video

Yaling Zhu, Xiaobin Li & Na Chen
College of Software Engineering, Lanzhou Institute of Technology, Lanzhou, Gansu, China

ABSTRACT: The MPEG video data includes three types of frames, that is: I-frame, P-frame and B-frame. However, the I-frame records the main information of video data, the P-frame and the B-frame are just regarded as motion compensations of the I-frame. This paper presents the approach which analyzes the MPEG video stream in the compressed domain, and find out the key frame of MPEG video stream by extracting the I-frame. Experiments indicated that this method can be automatically realized in the compressed MPEG video and it will lay the foundation for the video processing in the future.

Keywords: MPEG; key frame extraction; I-frame; DCT

1 INTRODUCTION

With the increasing development of electronic and information, the digital video data sharply increases every day, for example, electronic library of video data, TV programs, video on demand, video surveillance information and so on. Multimedia information of image, video and audio has gradually become the main form of information media in the field of information processing. Nowadays, according to the scientific statistics, the massive information obtained to human beings and the text information may be less than 20 percent, but the multimedia information may occupy more than 80 percent of the total outcomes [1]. Thus, the research of browsing and searching of the video database is a pressing duty. The technology of shot segmentation and key frame extraction is the foundation of the video analysis and video retrieval based on content. Key frame is the key image to describe a shot, which is composed of one or more I-frames. Thus, the extraction of I-frame in video processing is the most critical and the most basic step.

2 MPEG STANDARD

The MPEG is the abbreviation of Moving Picture Experts Group which is established by the international organization for standardization in 1988. It is committed to draw up the international standard of the moving image compression coding. The MPEG has announced the MPEG-1 standard at first, and then it further introduced the MPEG-2, the MPEG-4, the MPEG-7 and other standards [1, 2].

This paper presents the method called I-frame Extraction Algorithm, which is an approach of directly extracting I-frames from the elementary stream based on the MPEG-2.

2.1 *MPEG-2 standard*

The MPEG-2 which is introduced after MPEG-1 standard stands for the standard of the generic coding for moving pictures and audio information. The MPEG-2 standard includes four parts: the system, the video, the audio and the detection and testing of video, and the audio and system stream.

The MPEG-2 stream is divided into three layers, that is: the basic stream, the packet elementary stream and the multiplexed transport streams, the program stream. The basic stream is composed of video elementary stream (VES, Video Elementary Bit Stream) encoded by video compression standard and audio elementary stream (AES, Audio Elementary Bit Stream) encoded by audio compression standard [2, 3].

2.2 *VES structure*

As shown in figure 1, the VES structure includes six layers: a picture sequence (PS), a group of picture (GOP), a picture, a slice, a macroblock (MB), and a block [2].

2.3 *MPEG-2 frame*

The MPEG-2 standard is used to define three types of frames: the I-frame, the P-frame and the B-frame which uses different compression coding.

I-frame is adopted the DCT to encode with using the correlation of image itself based on intra-frame coding, but it does not need other frames as a reference[4, 5].So, the correlation of time is not considered in the intra-frame coding image. I-frame recorded the main information is the basic frame in the MPEG international standard, and it can be referred to the other types of images.

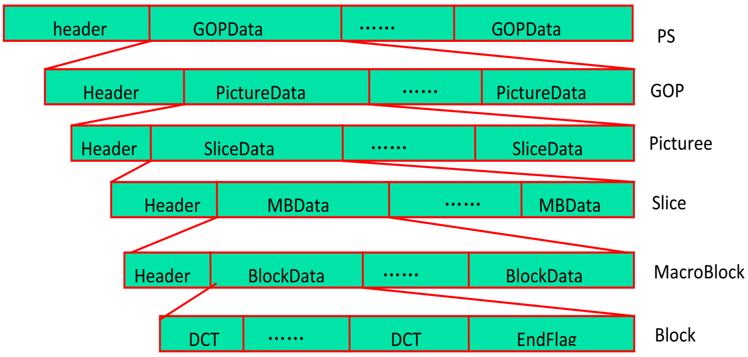

Figure 1. VES structure

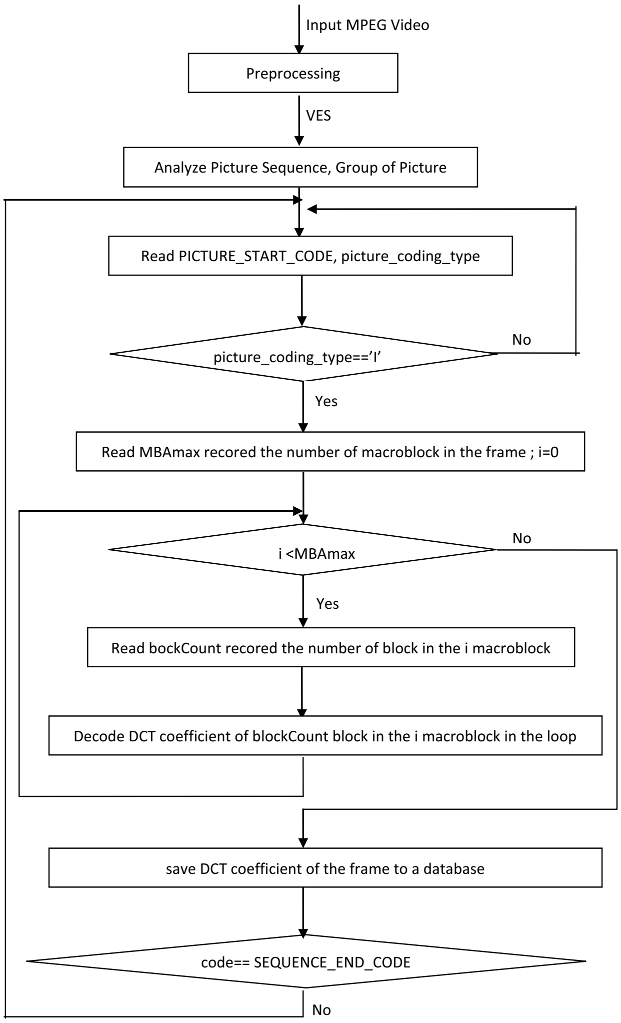
Input MPEG Video
Preprocessing
VES
Analyze Picture Sequence, Group of Picture
Read PICTURE_START_CODE, picture_coding_type
picture_coding_type=='I'
No
Yes
Read MBAmax recored the number of macroblock in the frame ; i=0
i <MBAmax
No
Yes
Read bockCount recored the number of block in the i macroblock
Decode DCT coefficient of blockCount block in the i macroblock in the loop
save DCT coefficient of the frame to a database
code== SEQUENCE_END_CODE
No

Figure 2. Fow chart of algorithm

P-frame and B-frame are the predicted images used for motion compensation. Both of all use inter-frame-coding, but the P-frame is used in the previous I-frame or the previous P-frame to encode, the B-frame requires I-frame or P-frame which comes from two directions that is before and after as a reference[2, 3].

3 DESIGN AND IMPLEMENTATION OF THE ALGORITHM

3.1 *Idea of algorithm*

The main idea of the algorithm can directly read out the header of image sequence, the header of the group of image and the header of picture in turn from the MPEG-2 binary stream without decompression. If the frame type of the header of picture is I-frame, then it can further analyze the slice and the macroblock [6]. Due to the different format of color image, the number of blocks consisting of a macroblock is also not equal. Finally, the approach can parse the data of each block. Each block data is the DCT of the I-frame displayed in the matrix, and then stored in the database. The flow chart of the algorithm is shown in figure 2.

3.2 *Realization of algorithm*

From the main idea of algorithm, the extraction of I-frame from compressed MPEG stream consists of the following steps.

Input: A MPEG video stream.

Output: I-frame and the number of I-frame.

Step1: In order to obtain the data of picture sequence, it needs to find the value of 0X000001E0 from the video binary stream binary that is start tag of the video data, and then return the binary data stream of picture sequence.

Step2: According to the binary data stream obtained by the previous step, it looks for the value of 0X000001B3 marked the start of picture sequence. The value of 0X000001B8 indicated the beginning of the group of picture, and the value of 0X00000100 signed the start of the picture.

Step3: Having a start tag of the picture, it analyzes the data downward layer by layer. The value of picture-coding-type of the binary stream data of picture determines whether the image is I-frame. If it is, it will continue to parse the picture. Otherwise, it will look for the next 0X000001 which is the marker of next picture.

Step4: If it discovers the value of 0X00000101 ~ 0X000001AF signed the beginning of the slice, it will analyze the slice consisted of many macroblock, then obtain the macroblock, and further resolve the data of macroblock[6].

Step5: The format of the color image identifies the number of blocks included in each macroblock. Therefore, we can read out the data of each block of every macroblock of I-frame, and store the data in database with SQL Server, then analyze them.

Step6: If it meets the value of 0X000001B7 labeled at the end of the sequence, the algorithm will terminate. Otherwise, it will return to step 2 to read out the next group of picture.

The pseudo code of the algorithm is shown as follows:

```
#define VIDEO_START_CODE 000001E0
#define SEQUENCE_START_CODE 000001B3
#define SEQUENCE_END_CODE 000001B7
#define GROUP_START_CODE 000001B8
#define PICTURE_START_CODE 00000100
#define SLICE_START_CODE_MAX 00000101
#define SLICE_START_CODE_MIN 000001AF
#define START_CODE_PREFIX 000001
static int video_sequence (int*Bitstream_Framenumber) {
  while(1) {
  find the next start flag and save to a variable named code;
  switch (code) {
  case SEQUENCE_HEADER_CODE:deocede_sequence_header(); break;
  case GROUP_START_CODE: decode_group_of_pictures_header(); break;
  case PICTUTURE_START_CODE:decode_picture_header();return 1;break;
  case SEQUENCE_END_CODE:return 0;break; }
   ……
  if (picture_coding_type==’I’){
  obtain MBAmax saved the tota number of the macro block in the picture
  SS: find the next start code;
   if (code < SLICE_START_CODE_MAX && code>SLICE_START_CODE_MIN)
   for(;;) {//to analyze MBAmax slice
  call the method named decode_slice() in the loop;
  {  ……
  call the method named decode_macroblock() in the loop;
  {  ……
  call the method named decode_block() in the loop;
  if (getBits(16)==START_CODE_PREFIX)
  goto SS;}
  }
sava DCT coefficient of a frame to the database;}
}
```

4 EXPERIMENTAL RESULTS

The algorithm is implemented in a console with vc6.0 and SQLServer2000 using lots of trial [7]. In order to evaluate the algorithm proposed in this paper, we establish 3G video's database which includes five different types of video, that is: animation, sports, stories, news and views [8, 9]. Some of the experimental results are shown in the Table 1.

Table 1. Number of frames extracted from different types of video

Video Type	Video Size	The total nubmer of frames	The number of I-frame
Sport	19.2M	1270	90
Animation	61.7M	4665	359
View	294M	21488	1632
News	23.4	1665	128
Story	256M	18414	1364

5 CONCLUSION

Video processing based on I-frames can not only keep the basic and original video information, but also conduct the information processing on the original compressed video within the margin of error which is greatly reducing the amount of data. Experimental results show that I-frames automatically extracted from compressed MPEG video can be effectively realized.

ACKNOWLEDGEMENTS

This work was financially supported by the Teaching Reform Project of Lanzhou Institute of Technology, China (Study on the reform of the teaching mode of software technology courses based on CDIO) and the Scientific Research Foundation of the Higher Education Institutions of Hainan Province, China (No. Hjkj 2013-53).

REFERENCES

[1] Ke Shen. & Edward J. Delp. 1996. A Fast Algorithm for Video Parsing Using MPEG Compressed Sequences. Proceedings of International Conference on Image Processing (ICIP'96), Lausann.

[2] Internation Standard. ISO/IEC13818-2 Generic coding of moving pictures and associated audio information: video second edition 2000.

[3] Di Zhong. & S.F.Chang. 2004. Real-time view recognition and event detection for sports video. *J. Vis. Commun. Image R*. (15): 330-347.

[4] N. Babaguchi, Y. Kawai. Y. Yasugi & T. Kitahashi. 2000. Linking Live and Replay Scenes in Broadcasted Sports Video. Proc. *ACM International Workshop on Multimedia Information Retrieval*, (11): 216-227.

[5] Hari Sundaram. 2002. *Segmentatio, Structure Detection and Summarization of Multimedia Sequences*. University of Columbia.

[6] Dan Tao. 2004. *The Research and Implementation of Keyframe Extracion Methods in Content-Based Video Retrieval System*. University of Columbia.

[7] Yizhen Zhang. & Tao Liu. 2002. Visual C++ Implementation of MPEG/JPEG Encoding and decoding technology. Posts and Telecommunications Press.

[8] John S.Boreczky, Lawrence A. Rowe. 1996. Comparison of Video Databases IV. Proc, SPIE 2670, 170-179.

[9] Arun Hampapur. 1995. Design Video Data Management Systems. *The University of Michigan*.