2013 Fourth World Congress on Software Engineering

Design and Implementation of a Real-time Video Stream Analysis System Based on FFMPEG

Lei Xiaohua, Jiang Xiuhua, Wang Caihong
School of Information Engineering
Communication University of China
Beijing, China

Abstract—**this paper introduces the design and implementation of a FFMPEG-based stream analysis system. The system supports real-time analysis of media data from a variety type of transmission protocols, media container formats and video / audio coding standards. The system breaks through the limitation of traditional stream analysis system that only supports a specific data format, and is more suitable for analysis of Internet media stream. First, the paper introduces the structure of the system. Then, the paper implements the FFMPEG-based Decoding Section. Meanwhile, it introduces FFMPEG's important data structures and FFMPEG's decoding process. Finally, the paper implements the Analysis Section. Method mentioned in this paper can be used as the basis of more complex stream analysis system.**

Keywords- Stream analysis; FFMPEG; real-time;

I. Introduction

With the increasing of network bandwidth, and the development of video compression technology, multimedia is becoming more and more popular on the Internet. Video and audio are gradually replacing the traditional information transmission method such as text and picture [1]. In the era of the Internet, multimedia data varies from transmission protocols, media container formats, and the video / audio coding standards. This is different from digital TV and IPTV.

Digital TV, or IPTV, have independent transmission network, so they have unified transmission method; and they also have unique program source, so their multimedia data have unified media container format and video / audio coding standard. Thus implementing a stream analysis system that focuses on digital TV or IPTV is relatively simple, only need to analyze a specific transmission method, media container format, and video / audio coding standard. For example, IPTV typically use RTP or UDP as its transmission method, TS as its media container format, H.264 as its video coding standard [2]; digital TV typically use QAM modulation as its transmission method, TS as its media container format, MPEG2 as its video coding standard [3].

Now the videos on the Internet are different. Different video service providers use different type of the transmission protocols, media container formats, as well as video / audio coding standards. Transmission methods include a variety type of network protocols such as HTTP, RTMP, MMS, RTP [4]; media container formats include FLV, MP4, AVI, TS, MKV, RMVB, etc. [5]; video coding standards include H.264, VC-1, VP8, MPEG4, MPEG2, RV4 etc. [6]. Therefore, it is not appropriate to use system that focus on a particular transmission method, media container format, and video / audio coding standard to analyze today's Internet media stream.

The stream analysis system described in this paper is designed to provide a unified solution to analysis a wide range of media data formats. In order to adapt to the vastly different transmission protocols, media container formats, as well as video / audio coding standards, the system uses FFMPEG as its kernel. FFMPEG is an open source, very powerful multimedia framework that supports multiple transmission protocols, media container formats, as well as video / audio coding standards and it provides a unified data structure to store the information extracted from multimedia data, thus effectively solves the difficulty in analysis of wide range of media data formats. In addition, FFMPEG provides highly efficient codec algorithm that can meet requirements of real-time video / audio analysis [7]. The system in this paper has two kinds of functions: Simple Analysis and Detailed Analysis. Simple Analysis obtains basic information of each frame along with the playback of the media stream, includes: frame type (I, P, or B), PTS, DTS, etc.; Detailed Analysis conducts a deep analysis of a single frame of video, including each macroblock's quantization parameter, partition mode, motion vector, reference frame. The system uses Microsoft VC++ 2010 MFC to show the result of the analysis. Our implementation result shows that the system achieves real-time decoding and analysis of media stream.

II. Design of System Structure

A. Structure

The structure of the system is shown in Figure 1. The system is divided into two parts: the Decoding Section and the Analyzing Section.

Decoding Section's main function is to decode and play the input media stream. It contains four modules: Decoding Protocol Module; Decoding Format Module; Decoding Video / Audio Module and Synchronization of Video and Audio Module. The data is first putted into Decoding Protocol Module; and after a series of processing, it is played out after Synchronization of Video and Audio Module.

Analyzing Section's main function is to obtain data from the Decoding Video Module to analysis the media stream. There are two analysis modules: the Simple Analysis Module and the Detailed Analysis Module. Simple Analysis Module does the real-time analysis the basic information of each video frame. Detailed Analysis Module makes deep analysis of a

978-1-4799-2883-5/13 $31.00 © 2013 IEEE
DOI 10.1109/WCSE.2013.38

video frame. Since the processing method is more complex, Detailed Analysis Module is not applicable to analyze each frame of video real-timely.

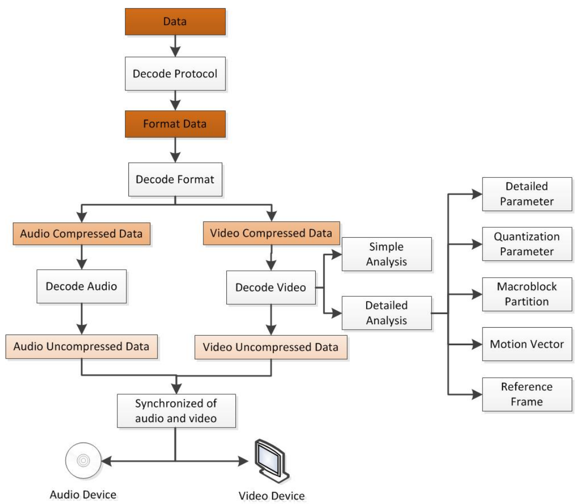

Figure 1. System Structure

B. Decoding Protocol

Decoding Protocol Module is the first module processing the input data. The function of the module is to decode input protocol data to media container format data. The input data always uses many types of transmission protocols, for example HTTP, RTMP, or MMS etc. These protocols typically contain some signaling data that control media stream's transmission and playback. These data are not a part of video stream. The function of Decoding Protocol is to remove the protocol signaling data, and process data to the form of standard media container format. For example, after the process of the Decoding Protocol, RTMP protocol data converts to FLV format data [8].

C. Decoding Format

Decoding Format Module is the second module. The function of the module is to separate media container format data into audio encoded data and video encoded data. There are many kinds of media container format, for example TS, FLV, AVI, MP4, MKV, RMVB, etc., and its function is to put encoded video data and audio data into a certain format of file [5]. Therefore, it is necessary to decode media container format to get the video / audio encoded data before actually decoding the video / audio. For example, after the process of Decoding Format, FLV format data is separated to the H.264-encoded video data and AAC-encoded audio data [9].

D. Decoding Video / Audio

Decoding Video / Audio Module are the third module. The function of the module is to decode video / audio encoded data into uncompressed video / audio raw data. Audio coding standard contains AAC, MP3, AC-3, etc. [10]; video coding standard contains H.264, MPEG2, VC-1, etc. Decoding Video / Audio are the most important and complex part in the system. After processing of Decoding Module, the encoded video data is converted to uncompressed pixel data, such as YUV420P, RGB, etc. [11]; encoded audio data is converted to uncompressed audio sample data, such as PCM data [10].

E. Synchronization of Video and Audio

Synchronization of Video and Audio Module is the last module. The function of the module is to synchronize uncompressed video and audio data based on parameters that extracted from Decoding Format Module, and send video / audio data to the video / audio device to play.

F. Simple Analysis

Simple Analysis Module makes real-time analysis of the video stream accompany with its playback. Considering real-time requirements, the module makes a simple analysis when system is decoding a video frame, extracts parameters such as frame type, size, PTS, DTS. The chapter IV will be described it in detail.

G. Detailed Analysis

Detailed Analysis Module makes deep analysis of the video frame accompany with video's playback. When system is decoding a video frame, the module can carry out a detailed analysis of this frame, such as the each macroblock's quantization parameter, partition mode, motion vector, and reference frame. The chapter IV will be described it in detail.

III. Implementation of Decoding Section

The system development environment is Microsoft VC++ 2010. It uses FFMPEG as its kernel to process media stream. This chapter is an overview of its important data structures and its process of decoding media stream.

A. FFMPEG's Data Structure [7] [12]

FFMPEG's most important data structure in the process of decoding is shown in Figure 2. They can be divided into 4 kinds of usage:

1) Decoding Protocol

AVIOContext, URLProtocol, URLContext stores the media stream's protocol type and state. URLProtocol stores the type of input media protocol.

2) Decoding Format

The AVFormatContext store the information contained in the media container format; AVInputFormat stores the type of input media stream's media container format.

3) Decoding Video / Audio

Each AVStream store the information of a video / audio stream; each AVStream contains an AVCodecContext to stores the decoding information of video / audio stream; each AVCodecContext contains an AVCodec, which stores the information of video / audio decoder.

4) Storing Data

AVPacket stores the information about encoded video / audio data; while AVFrame stores the information about decoded uncompressed video / audio data.

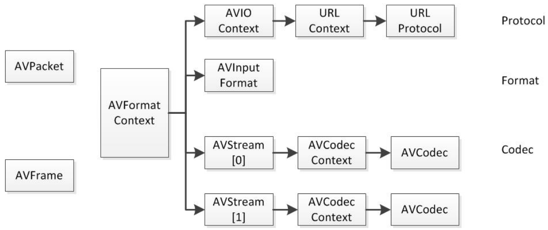

Figure 2. Data Structure

B. *Decoding Process [7] [13] [14]*

Decoding Process of the system can be divided into 2 steps:

1) Pre-processing.

As shown in Figure 3, Pre-processing step includes: Initialization, Decoding Protocol, Decoding Format, and Getting AVPacket.

2) Decoding.

As shown in Figure 4, Decoding step includes: Decoding AVPacket, Synchronization of the decoded video / audio data, and Sending decoded data to video / audio devices.

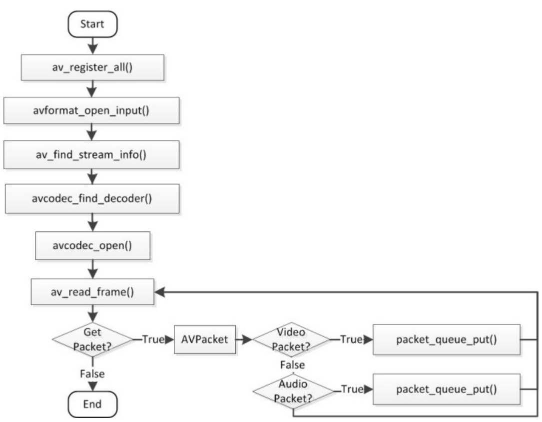

Figure 3. Pre-processing Step

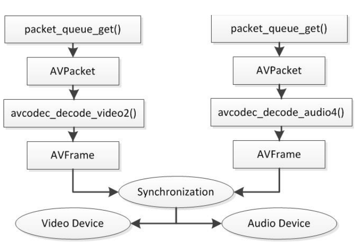

Figure 4. Decoding Step

IV. IMPLEMENTATION OF ANALYSIS SECTION

Analysis Section of the system analyzes video frame's parameters from the Decoding Module introduced in chapter III and displays the result graphically in the system interface. The structure of the Analysis Section is shown in Figure 5. The Data Module extracts information from AVPacket, AVFrame and related data structures in Video Decoding Module. The system first stores related information in the Data Module, and then it calls Analysis Module for the analysis; finally it calls Drawing Module and Output Module to display the analysis result to the system interface.

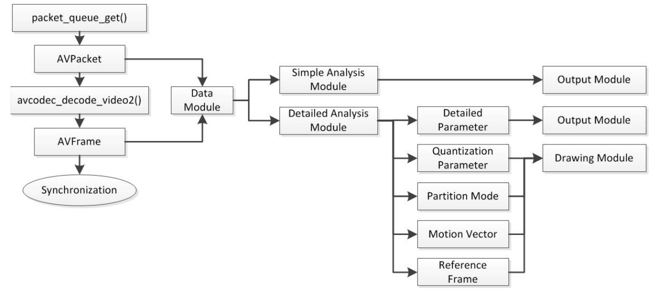

Figure 5. Analysis Section

A. *Simple Analysis Module*

Simple Analysis Module makes real-time analysis of the video stream accompany with its playback. Considering real-time requirements, the module just extract some basic parameters when the system is decoding each frame of video. The module analyzes the video frame's parameters such as frame type, PTS, and DTS.

B. *Detailed Analysis Module*

1) Detailed Parameters

Detailed Parameters contains more parameters of each frame of video such as display time, macroblock's number, and motion vector's number etc.

2) Analysis Panel

Analysis Panel is used to display the detailed information of each macroblock in the system interface. It contains the following analysis: the Quantization Parameter Analysis, Partition Mode Analysis, Motion Vector Analysis, and Reference Frame Analysis. The following sections will describe these functions.

3) Quantization Parameter Analysis

Quantization Parameter Analysis Module is used for the analysis and graphical display each macroblock's quantization parameter in a frame. Different quantization parameter value has different level of gray background. The quantization parameter value is shown on each macroblock above.

4) Partition Mode Analysis

Partition Mode Analysis Module is used for the analysis and graphical display each macroblock's partition mode in a frame. Different partition mode has different background color. The cut lines of each macroblock are shown on each macroblock above.

5) Motion Vector Analysis

Motion Vector Analysis Module is used for the analysis and graphical display each macroblock's motion vectors in a frame. Motion vector is drawn as line shown on each macroblock above.

6) Reference Frame Analysis

Reference Frame Analysis Module is used for the analysis and graphical display each macroblock's reference frame in a frame. This function is available for some new video coding standards such as H.264 [15]. Different reference frame index value has different level of gray background. The reference frame index value is shown on each macroblock above.

V. RESULT OF IMPLEMENTATION

A. *Main Interface*

Main Interface of the system is shown in Figure 6. The system is a subsystem of the "Internet Multimedia QoE Evaluation and Monitoring System". As shown in the figure, when the system is decoding video file located in "E:\movie\jiangnan.flv", the main interface shows the parameters extracted from data structures in FFMPEG.

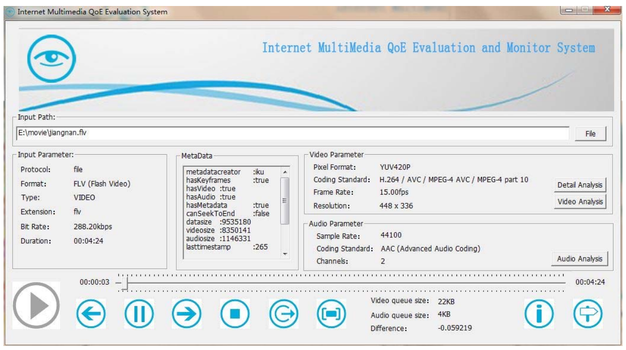

Figure 6. Main Interface

B. *Video Window*

Video Window is shown in Figure 7. After decoding the media stream, uncompressed video and audio data are played in the window.

Figure 7. Video Window

C. *Simple Analysis*

Simple Analysis Module's interface is shown in Figure 8. Chart on left records the decoded frame's size and type. List on the right shows the PTS, DTS, and other information of each decoded frame.

D. *Detailed Analysis*

Detailed Analysis Module's interface is shown in Figure 9. Analysis Panel is on the left; it divides a video frame into the macroblock unit and can graphically display parameters of each macroblock in a video frame. The drop-down list above is used to select parameters that Analysis Panel need to analysis. Detailed parameters are shown above and below the Analysis Panel. The statistic information about macroblock partition mode and quantization parameter is on the right of Analysis Panel.

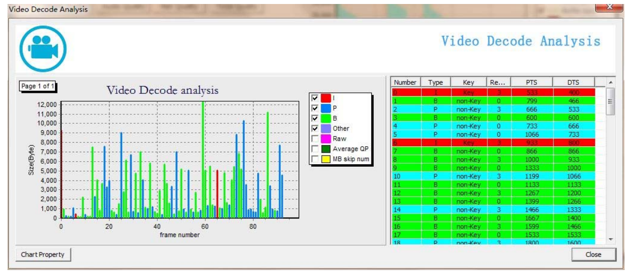

Figure 8. Simple Analysis

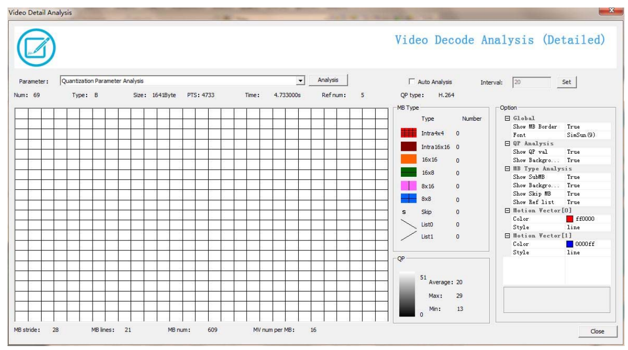

Figure 9. Detailed Analysis

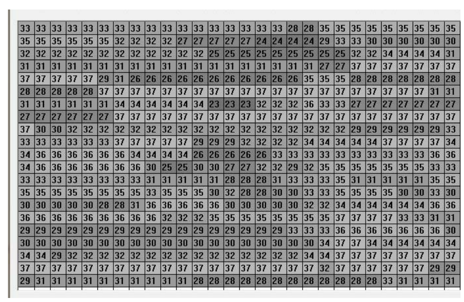

Figure 10. Quantization Parameter Analysis

1) *Quantization Parameter Analysis*

Quantization Parameter Analysis result shown in Figure 10. As shown in the figure, most of macroblock's quantization parameters are between 25 and 37. Different quantization parameter has different level of gray background.

2) *Partition Mode Analysis*

Partition Mode Analysis result is shown in Figure 11. As shown in the figure, most of macroblocks use 16x16 partition mode, while some of macroblocks use 8x16, 16x8, 8x8, or 4x4 partition mode. They are marked in a different background color and different form of cut line.

3) *Motion Vector Analysis*

Motion Vector Analysis result is shown in Figure 12. As shown in the figure, most of motion vectors are on the Left-

Bottom part of the frame, which reflects that there are more fierce movements here. The motion vector is drawn in the form of red line in the corresponding position.

4) *Reference Frame Analysis*

Reference Frame Analysis result is shown in Figure 13. As shown in the figure, most of macroblocks use frame which index number is "0" as reference frame, while little number of macroblocks use frame which index number is between "1" and "3". Different reference frame index number has different level of gray background.

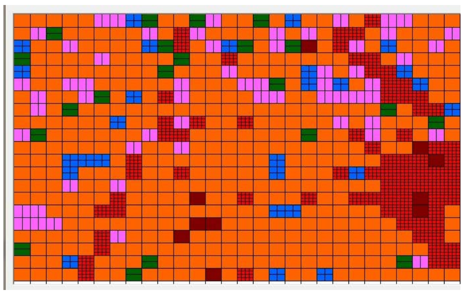

Figure 11. Partition Mode Analysis

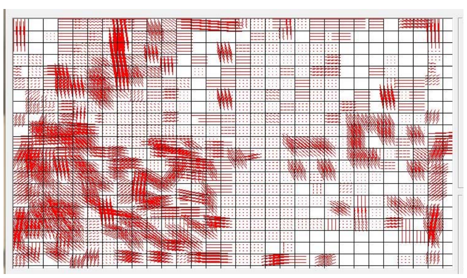

Figure 12. Motion Vector Analysis

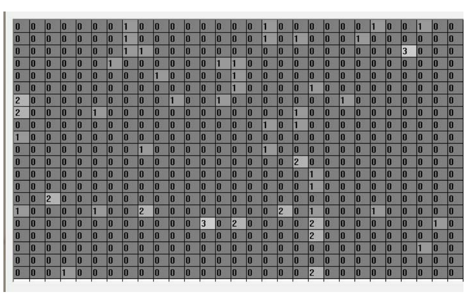

Figure 13. Reference Frame Analysis

VI. CONCLUTION AND PROSPECT

This paper introduces the design and implementation of FFMPEG-based real-time video stream analysis system. The system support varies type of transmission protocols, media container formats and video / audio coding standard; and it achieves real-time stream analysis requirement. However, the function of stream analysis may not complete. We need to do more work focus on the system's Analysis Section in the future. The method mentioned in this paper can be used as the basis of design more complex real-time stream analysis system.

ACKNOWLEDGMENT

This work was supported by the "Research of Network multimedia QoE evaluation and monitoring system" Project in Communication University of China.

REFERENCES

[1] China Electronics Standardization Institute, CCBN2012 Proceedings. Beijing: China Standard Press, 2012, pp. 218–227..

[2] X.Wan, H.Zhang, Y.Ou, Y.Zhang. IPTV Technology and Operations. September 2010.

[3] X.Jiang, Y.Zhang. The Principle and Application of Digital Television Broadcasting. September 2007.

[4] Wikipedia, Comparison of streaming media systems. http://en.wikipedia.org/wiki/Comparison_of_streaming_media_systems.

[5] Wikipedia. Comparison of container formats. http://en.wikipedia.org/wiki/Comparison_of_container_formats

[6] Wikipedia. Comparison of video codecs. http://en.wikipedia.org/wiki/Comparison_of_video_codecs

[7] The FFmpeg developers. FFmpeg Documentation. May 2013

[8] Adobe Systems Inc., RTMP Specification. June 2009.

[9] Adobe Systems Inc., Video File Format Specification Version 10. November 2008

[10] Audio file format. http://en.wikipedia.org/wiki/Audio_format

[11] Uncompressed video. http://en.wikipedia.org/wiki/Uncompressed_video

[12] Shuliang Yang. FFMPEG source code analysis. http://files.cnblogs.com/mcodec/ffdoc.7z

[13] Martin Bohme. How to Write a Video Player in Less Than 1000 Lines. http://dranger.com/ffmpeg/

[14] M.Song,J.Sun, X.Fu, W.Xiong. "Design and Implementation of Media Player Based on Android". 6th WiCOM, September. 2010

[15] Houjie Bi. A New Generation of Video Coding Standard --H.264. November 2009