行政院國家科學委員會專題研究計畫成果報告

MPEG-4 影音串流伺服器之設計與實作

The Design and Implementation of an MPEG-4 Media Streaming Server

計畫編號：NSC 90-2213-E-002-091

執行期限：90 年 8 月 1 日至 91 年 7 月 31 日

主持人：陳文進 教授 國立臺灣大學資訊工程研究所

一、中文摘要

隨著寬頻網路的蓬勃發展，影音串流服務（video streaming）不再遙不可及，影音串流技術發展的目的在於如何在複雜的網路媒介下傳送影音資料，用戶端能夠即時接收資料後立刻播放，不需等待整個影片下載完畢，用戶端便不再浪費冗長的等待時間，並且能降低用戶端暫存空間需求。

此計劃的研究課題分為兩類：(1)如何藉由影音串流技術傳送影音資料、(2)如何在有限的頻寬限制下，傳送高品質的影音內容。由於 MPEG-4 標準為了適應網路傳輸的需求，整體的設計已經包含了網路傳輸的各種技術，需要的傳輸頻寬也大大減少，所以本計劃將採用 MPEG-4 標準的各種技術做為發展的主軸，並實作出以下各個相關模組：(1) MPEG-4 檔案編輯器、(2)MPEG-4 媒體格式轉換器、(3)MPEG-4 影音串流伺服器。

關鍵詞：多媒體, MPEG-4, 串流, 網路, 伺服器, 實作。

Abstract

MPEG-4(ISO/IEC 14496) is the state-of-the-art standard for describing the presentation and composition of multimedia contents. It enables the integration of different media data in an objected-based way. The specification of MPEG-4 includes media data access framework, video/audio codec, and synthetic object presentation and integration method of different media data.

The research topics of this project are: (1) how to stream video data in internet, (2) how to send high quality media in the internet with low bandwidth. The design of the MPEG-4 standard includes the needs of streaming media in the network, so in this project we implement the streaming park of the MEPG-4 standard and develop some tools, which are: (1) The MPEG-4 file format editor, (2) The MPEG-4 video format transformer, (3) The MPEG-4 streaming server.

Keywords: MPEG-4, Multimedia, Streaming, Network, Server, Implementation.

二、緣由與目的

Following the great success of MPEG-1 and MPEG-2 in the Audio/Video market, the ISO MPEG group continues to develop their next standard – MPEG-4. MPEG-4 targets on several issues: authoring and encoding of object-oriented scenes, encoding of natural and synthetic media, media streaming, uniform transmission interface in the heterogeneous network environments, and fancy human-machine interactions. Therefore, MPEG-4 is designed to be applicable for a wide spectrum of the presentation applications.

The MPEG-4 standard comes in several parts. The definition of the binary scene description and the management information of the elementary streams are specified in part 1: System. The scene description is based on VRML and specifies the spatial-temporal composition of a MPEG-4 presentation, and the elementary streams convey the coded representation of audiovisual data to the receiver. Then these media objects will be composed into an integrated audiovisual presentation at the receiver. Compression techniques of visual and audible media objects are specified in part 2 and part 3. Other parts of MPEG-4 describe reference software and conformance procedure (part 4 and 5). In part 6, the delivery multimedia integration framework (DMIF) is introduced, which standardizes the transport services features which are available for the MPEG-4 applications.

In this project, our purpose is to research the streaming technologies and to develop some useful systems:

2.1 The MPEG-4 video format transformer

To facilitate MPEG-4 video content creation, we propose our architecture for MPEG-4 video content creator. For the quality and availability of video inputs, we choose the MPEG-2 video clips and captured video to be our video source. This encoder can convert MPEG-2 video clips into MPEG-4 compliant video clips, or capture video frames and convert them into MPEG-4 video clips.

2.2 The MPEG-4 file format editor

The MPEG-4 standard proposes MP4 file format to be media data storage format. According to MP4 specification, media bitstream can be divided into two parts: media description information (metadata) and media data. The media description information is stored in a “moov” atom, and the media data is stored in a “mdat” atom. Without any GUI application assistance, we can not create MP4 file in an instinctive way.

To facilitate the MP4 file creation and management, we implemented an MP4 file editor. The encoded MPEG-4 video data can be transformed into MP4 file track easily. The media samples of the media track can be analyzed and the statistical information can be used to schedule the media samples streaming plan.

2.3 The MPEG-4 Streaming Server

Conforming to the MPEG-4 file format and media data streaming definition, we propose the architecture design and implementation of a MPEG-4 streaming server. In the streaming server, we propose the design and implementation of media data segmentation, SL packet generation, user commands, and the access unit timing information. In the client site, we propose the design and implementation of the access unit reconstruction, SL packet lost detection mechanism, and integration with open source ISO MPEG-4 video decoder. With media streaming, users can review the media presentation at the streaming client in real-time.

三、結果與討論

3.1 The MPEG-4 video format transformer

3.1.1 System introduction

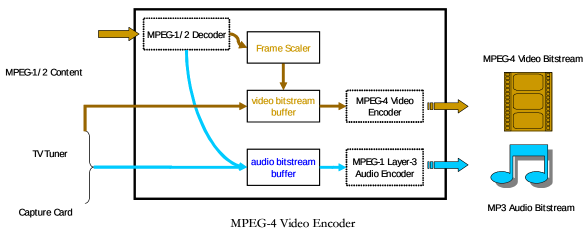

Figure 1: The data flow of the MPEG-4 video format transformer.

The MPEG-4 standard proposes a high quality and bandwidth efficient video codec. The quality and bandwidth requirement of the MPEG-4 video content is suitable for video streaming on broadband network (1.5 Mbps ~ 8 Mbps). Conforming to the popular and practical decoder support, we adapt MPEG-4 video and MPEG-1 Layer-3 audio to be the format of our output bit streams. Our architecture (based on Microsoft VFW / DirectShow filter architecture) integrates four modules. Output of the MPEG-4 video encoder conforms to MPEG-4 Visual Advance Simple Profile (Visual Profile).

For high video quality and availability of video inputs, we choose the MPEG-2 video streams and capture card to be our video input source. First, the internal buffer manager will invoke MPEG-2 video decoder (developed based on the MSSG video decoder library) to decode one video frame or read one frame from capture input. After the uncompressed video frame is obtained from the decoder, the internal buffer manager invokes external MPEG-4 encoder to encode this video frame. After this video frame is encoded into MPEG-4 compliant video bitstream, the internal buffer manager outputs this encoded frame to an MP4 file.

3.1.2 System result

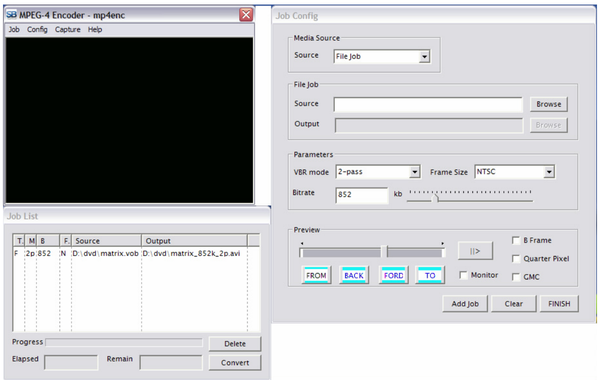

Figure 2: The appearance of our system.

The transcoding phase of the MPEG-4 video encoder can be processed in 1:1 time length issue at AMD XP-2000+ system (2-pass, 320x240). It includes MPEG-2 video bitstreams decoding, the AC3 audio streams decoding, MPEG-4 video frames encoding and A/V multiplexing. Figure 2 shows the user interface of the MPEG-4 Video Encoder.

3.2 The MPEG-4 file format editor

3.2.1 System introduction

The MP4 file format is defined in the MPEG-4 version 2 specification. It is designed to integrate the media data and metadata of a MPEG-4 presentation in a flexible, extensible, self-contained file format; which facilitates the content interchange, management, editing and presentation.

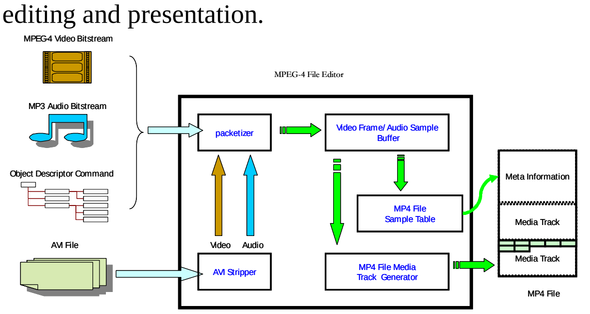

Figure 3: The architecture of the MP4 file format editor.

We implemented one MP4 file parser based on the ISO/IEC JTC1/SC29/WG11 N2801 subpart 4. To facilitate MP4 file creation and management, we integrated our MP4 file parser with a graphical user interface. It provides the basic functionalities to manage MP4 files, such as adding/deleting MP4 media track, or converting AVI files to MP4 files. We can add several video bit streams into the one MP4 file, and this mechanism can provide the scalability for video streaming. Without storing different quality of video contents into different MP4 files, the streaming server can select media tracks among MP4 file dynamically to satisfy different quality requirements from client site.

3.2.2 System result

The architecture of the MP4 file editor is illustrated in Figure 3. To extract the access unit in the media stream, we implemented the packetizer module to segment access units from the media streams. Our packetizer module can recognize the MPEG-4 video, MPEG-1 Layer-3 audio and OD command encoded bitstream. After access units are extracted from the media stream, we put these access units in the sample buffer. Then we update the MP4 file sample table information and store these samples in the media data atom ('mdat' atom). The user interface of the MP4 file editor is shown as Figure 4. The atom information is displayed in the right Information window. The left window is the tree structure of the MP4 file.

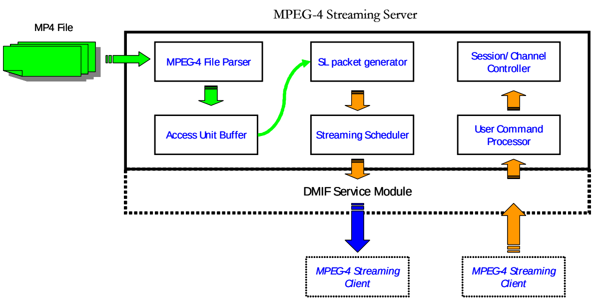

Figure 5: The architecture of the MPEG-4 streaming server.

The architecture of our MPEG-4 streaming framework is illustrated in Figure 5. Our MPEG-4 streaming framework is divided into two parts, a streaming server and a streaming client. The network protocol stack is based on the DMIF Server implementation. There are two kinds of media streams between the streaming server and the client tools, upchannel and downchannel. The downchannel stream is the stream from the streaming server to the streaming clients, and the upchannel stream is the stream from the client tools to the streaming server. The upchannel streams are responsible for forwarding acknowledge or request to the streaming server.

The MPEG-4 Streaming Server consists of the MPEG-4 File Parser, the Access Unit Buffer, SL Packet Generator, User Command Processor and the Streaming Scheduler.

The User Command Processor is responsible for processing the requests from the client and the user commands are forwarded and processed at the Session/Channel Controller. The MP4 file parser is discussed above. For each client application, the MPEG-4 streaming server creates a new Session structure. The session structure keeps the information about the status of the channels which belong to the session. The basic playback control commands can set the flag in the session structure.

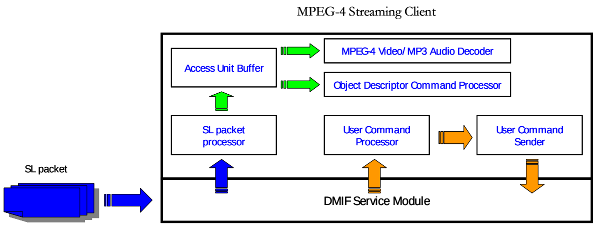

Figure 6: The architecture of the MPEG-4 streaming client.

The MPEG-4 Streaming Client consists of the SL Packet processor, the Access Unit Buffer, User Command Processor and the ISO MPEG-4 video decoder. The SL Packet process assembles the SL packets to the access units. We have implemented an AU buffer manager class providing the ByteStream interface to the ISO MPEG-4 video decoder. After an access unit is assembled, the SL packet processor will put the access unit in the AU buffer. The AU buffer is implemented as a FIFO queue. And the ISO MPEG-4 video decoder uses the AU buffer with the ByteStream interface.

3.3.2 System result

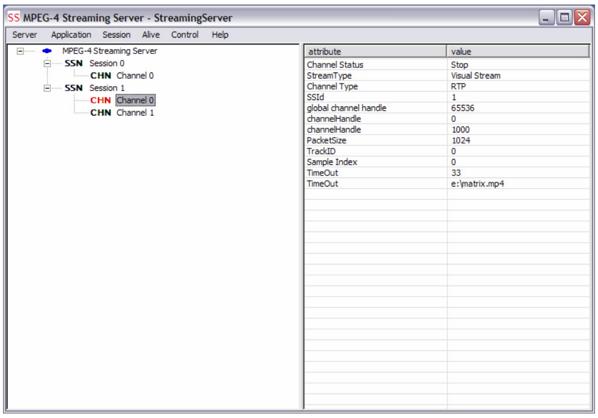

Figure 7: The appearance of the streaming server.

The Streaming Server user interface is illustrated in Figure 7. The Session window displays the status of the current client connections. There is a tree structure for current connections. The channels of each session are displayed in a tree structure manner. The Information window displays the status of every channel connection.

四、計劃成果自評

In this project, we research the kernel technology of the MPEG-4 streaming and develop some useful tool to facilitate the process of creating the MPEG-4 video and streaming the media via the internet. Currently, our system is useful to process the video only MPEG-4 media, but we cannot stream the scene data with network. We will still work on some advanced research topics about the MPEG-4 streaming technology.

五、参考文獻

[1] 劉登榮, "MPEG-4 媒體伺服器架構設計", 國立台灣大學資訊所博士班論文, July 1992.

[2] International Standard ISO/IEC 14496-1, Information Technology – Generic Coding of Audio-Visual Objects – Part 1: System, International Organization for Standardization, 1999.

[3] International Standard ISO/IEC 14496-2, Information Technology – Generic Coding of Audio-Visual Objects – Part 2: Visual, International Organization for Standardization, 1999.

[4] International Standard ISO/IEC 14496-3, Information Technology – Generic Coding of Audio-Visual Objects – Part 3: Audio, International Organization for Standardization, 1999.

[5] International Standard ISO/IEC 14496-6, Information Technology – Generic Coding of Audio-Visual Objects – Part 6: DMIF, International Organization for Standardization, 1999.

[6] International Standard ISO/IEC 14772-1:1997 – VRML 97: The Virtual Reality Modeling Language.

[7] ISO/IEC IS 11172, Information Technology – Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mbit/s – Part 2: Coding of Moving Picture Information, Part 3: Audio, International Organization for Standardization, 1991.

[8] ISO/IEC IS 13818, Information Technology – Generic Coding of Moving Pictures and Associated Audio Information – Part 2: Video, Part 3: Audio, International Organization for Standardization, 1994.