

RTP 기반 실시간 MPEG Audio 전송을 위한 PAD 모듈 설계 및 구현

권장우*, 김익형**, 김수진**, 김정철**

*동명정보대학교 컴퓨터공학과, **팜캐스트 기술

Design and Implementation of PAD Module for the Real-time MPEG Audio Transmission based on RTP

Jang-Woo Kwon*, IgHyeong Kim**, SuJin Kim**, JungChul Kim**

*Dept. of Computer Engineering, Tongmyong Univ. of Information Technology

**PalmCast Technology

요 약

멀티미디어 데이터의 효과적인 네트워크 전송에 대한 연구와 투자가 계속적으로 증가하고 있다. 본 논문에서는 실시간 전송 프로토콜인 RTP(Real-time Transfer Protocol)를 기반으로 MPEG 오디오 데이터를 실시간 전송하기 위한 PAD(Packet Assembler/Disassembler) 모듈을 설계, 구현하였다. RTP 기반 MPEG 오디오의 PAD 구현 방법은 MPEG 오디오의 계층에 관계없이 전송하는 방법과 MPEG Layer-3에 특화된 방법 등의 두 가지 방법이 있으며, 본 논문에서는 범용성을 중시하여 전자의 방식을 채택, 구현하였다.

구현한 PAD 모듈은 MPEG-1 및 MPEG-2 오디오 포맷을 지원한다.

1. 서론

컴퓨팅 기술의 발전과 더불어 인터넷 환경의 급속한 성장은 과거의 텍스트 및 스틸 사진 중심의 정보 교환에서 오늘날 음향 및 영상 등의 실시간 정보 교환으로의 기술적 발전 및 사용자 요구가 증가하고 있는 상황이다. 현재 서비스되고 있는 수많은 인터넷 방송 시스템이나 동영상 응용 시스템은 대부분이 TCP/IP 프로토콜을 기반으로 많은 네트워크 리소스를 사용하며, 또한 각기 다른 전송 포맷을 사용함으로써 이기종 시스템간의 호환성에 문제점을 드러내고 있다.

본 연구의 목적은 멀티미디어 데이터의 실시간 전송을 목적으로 개발된 RTP를 이용하여 네트워크 자원의 손실을 최소화하고 이기종 시스템간의 호환성을 확보하는데 있다. 실시간 전송 프로토콜인 RTP는 UDP를 기반으로 음향 및 영상 등의 멀티미디어 데이터 전송에서 선행되어야 할 실시간 정보처리에 초점을 맞추고 있다[1]. 따라서 RTP를 기반으로 MPEG 오디오 데이터의 실시간 전송을 위해 데이터 흐름을 분리된 패킷들로 나누고, 수신측에서 패킷을 다시 재조립하는 PAD(Packet Assembler/Disassembler) 모듈의 설계가 필요하다. 이와 같은 모듈을 구현하는 데에는 MPEG1, 2 오디오의 계층에 관계없이 오디오 포맷을

지원하는 방법[4]과 MPEG Layer-3 계층에만 특화된 데이터 포맷을 지원[5]하는 두 가지 전송방식이 있다. 본 연구에서는 전자의 방식에 따라 모듈을 구현함으로써 다양한 시스템에 결합할 수 있도록 하였다.

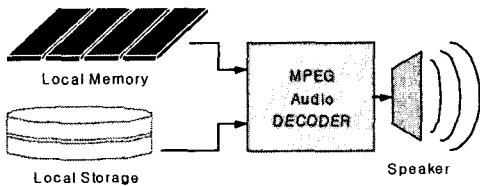
본론에서는 RTP 기반의 MPEG 오디오 PAD 모듈의 설계 및 구현을 위해 실제 데이터 전송을 담당하는 RTP 프로토콜을 소개하고, MPEG 오디오 재생 시스템과 MPEG 오디오 실시간 전송을 위한 PAD 기법을 설명한다. 그리고 실제 PAD 모듈의 내부 설계 구조에 대해서 살펴보고, 마지막으로 구현한 모듈의 실험을 통해서 본 연구에 대한 결론을 맺는다.

2. RTP 기반의 MPEG 오디오 PAD 모듈

2.1 RTP (Real-time Transfer Protocol)

RTP(Real-time Transfer Protocol)는 오디오, 비디오의 데이터를 실시간으로 멀티캐스트(Multicast) 및 유니캐스트(Unicast) 방식으로 전송하는데 적합한 전송 프로토콜이다[1]. RTP는 네트워크 자원의 예약을 하지 않으며 실시간 전송 서비스에 대한 QoS 역시 보장하지 않는다. QoS 보상은 데이터의 전송 감시 및 데이터 인식 등의 최소한의 제어를 지원하는 제어 프로토콜인 RTCP(Real-time Transfer Control Protocol)에 의해 결정된다[2]. RTP와 RTCP는 OSI 7 Layer에서 전송 계층(Transport Layer), 네트워크 계층(Network Layer)에 의존하여 설계되었다.

2.2 이상적인 MPEG 오디오 재생 시스템

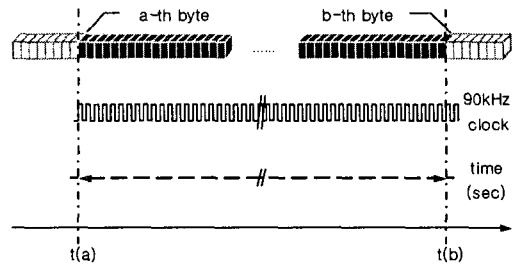


[그림 1] 이상적인 MPEG 오디오 재생 시스템

그림 1은 이상적인 MPEG 오디오 재생 시스템을 보여준다. 이상적인 재생 시스템은 지역 메모리, 또는 저장 매체에 기록된 데이터와 복호기 간의 전송 시간 및 전송오류가 거의 없는 시스템을 말한다[3].

본 연구의 RTP 기반의 MPEG 오디오 전송 모듈은 이상적인 MPEG 오디오 재생 시스템에 네트워크 상의

전송 시간에 대한 보상을 더하여 구현하였다.

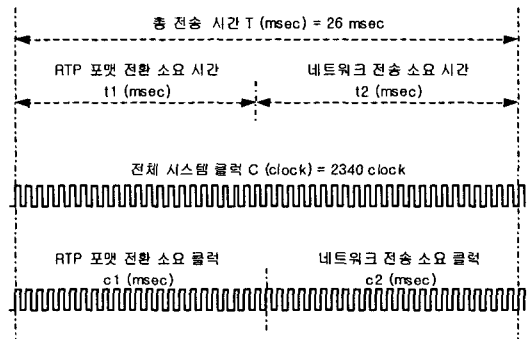


[그림 2] 이상적인 MPEG 오디오 재생 시스템의 time과 clock의 관계

그림 2는 이상적인 MPEG 오디오 재생 시스템에서의 동작 시간(time)과 동작 클럭(clock)의 관계를 나타낸다. 저장 매체에서 복호기로 데이터가 이동하는 관점에서 a번째 바이트의 데이터부터 b번째 바이트의 데이터까지 전송된다고 볼 때, 시스템의 동작 클럭이 90kHz, 비트율이 128kbps라고 가정하면, 1byte를 전송하는데 소요되는 시간 t는[3],

$$t(\mu s) = 1000000(\mu s) * 8 / 128000 \text{ bits} \approx 62.5 \mu s$$

이다. 즉 각 바이트 당 62.5μs의 소요시간동안 전송이 이루어져야 한다.



[그림 3] PAD 모듈의 MPEG 오디오 프레임 전송과 재생 시간

그림 3은 본 논문에서 구현한 PAD 모듈에서 MPEG 오디오 프레임 전송과 재생시간과의 관계를 나타낸 것이다. 128kbps, 44.1kHz의 MPEG 오디오에서 하나의 프레임을 전송한다고 가정할 경우, 프레임 길이는 417 ~ 418 bytes이다. 총 전송 시간, T는 26ms이며 클럭으로는 2340클럭이다[3]. 이 시간은 이상적인 재생기에서의 하나의 프레임에 대한 재생 시간과 같다.

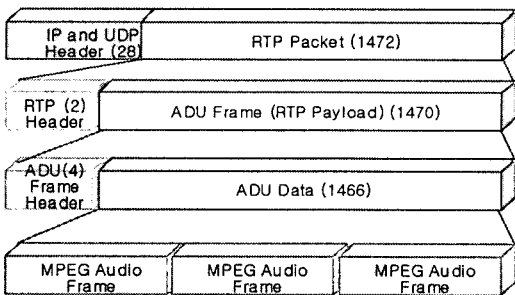
본 연구에서 구현한 모듈은 총 재생 시간, 즉 총 전송 시간, $T(ms)$ 에서 RTP 포맷으로의 전환에 따른 소요 시간, $t1(ms)$ 을 감산하여 얻은 시간동안 지연한 후 생성된 ADU 프레임들을 전송한다.

2.3 MPEG 오디오를 위한 패킷화(Packetizing)

미디어 데이터를 RTP 전송용 포맷으로 전환하는 것을 패킷타이징(패킷 분해; Packet Assembler; Packetizing)이라 정의하고, RTP 전송용 포맷의 데이터를 원래의 원본 미디어 데이터로 전환하는 것을 디패킷타이징(패킷 조립; Packet Disassembler; Depacketizing)이라고 한다[1]. 따라서 PAD(Packet Assembler/Disassembler)는 MPEG 오디오 데이터의 실시간 전송을 위해 데이터 흐름을 분리된 패킷들로 나누고 수신측에서 패킷을 재조립하는 방식이라 말할 수 있다.

본 논문에서는 PAD를 구현하는 방법으로 MPEG 오디오 포맷에서 각 계층을 모두 지원하는 방식을 채택하여 구현했다. MPEG 오디오 데이터는 최소 프레임 단위로 전송되며 ADU(Application Data Unit, RTP Payload라고도 한다)라는 프레임에 포함된다.

그림 4는 RTP 전송 포맷을 나타낸다[4]. MPEG 오디오만 전송되는 단일 음향 재생 세션에서는 ADU Frame Header의 값이 모두 0이 되지만, MPEG 시스템 영역의 MPEG 비디오 등의 영상과 더불어 재생될 때는 영상과의 동기화를 위해 값이 재설정된다[4].



[그림 4] RTP의 ADU 프레임 포맷

RTP Header에는 전송되는 데이터의 재생 시간(Presentation Time)에 해당하는 타임 스탬프(Time Stamp)와 Payload의 데이터 타입을 나타내는 Payload Type, ADU의 시간적인 연속성을 표현하는 식별 비트가 포함된다.

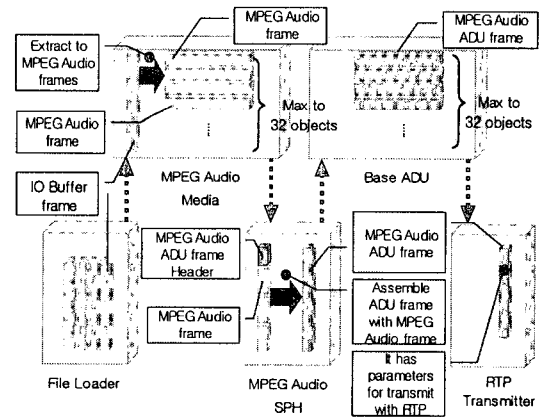
UDP를 기반으로 하는 데이터 전송에서 전송 간의 손실을 막기 위해 IP, UDP를 합한 전체 패킷의 용량을 이더넷(Ethernet)의 프레임의 최대 허용 용량인 1500 바이트로 제한하므로 IP Header 부분을 제외하고 실제 RTP 패킷은 1472바이트 이하로 제한된다[7].

전송 패킷의 용량을 감안하면 비트율이 128kbps, 샘플링 주파수가 44.1kHz인 MPEG 오디오는 한번에 세 개의 MPEG 오디오 프레임을 전송할 수 있다.

전송 시에는 한번에 3개의 프레임이 클라이언트에게 도달된 후 3개의 프레임의 재생시간만큼 지연되고 다음 3개의 프레임이 도착 하게 된다.

클라이언트에서는 일정한 Buffering 영역을 두어 도달 시간 내에 도착하지 않은 프레임에 대한 재생 오류 문제를 제거 할 수 있다.

2.4 PAD 모듈 설계

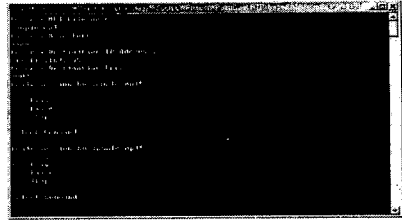


[그림 5] PAD 모듈 구조

그림 5는 RTP 기반의 실시간 MPEG 오디오 전송을 위한 PAD 구조를 나타낸다. 각 세션은 다섯 부분으로 구성되며 원본 데이터를 로딩하는 Source 부분, 데이터의 포맷이나 구조의 전환을 수행하는 Transform 부분, 데이터의 재생 및 네트워크 상으로의 전송을 위한 Render 부분, 미디어 데이터를 저장하는 Media 부분, 그리고 ADU 프레임을 저장하는 ADU 부분으로 구성된다.

그림 5에서 File Loader는 Media로 stream data를 전달하고 Media에서 MPEG 오디오 프레임으로 전환한다. MPEG Audio SPH는 MPEG Audio Media로부터 MPEG 오디오 프레임을 읽어 MPEG 오디오 기반의 ADU 프레임을 생성하고 Base ADU에 저장한다. RTP

Transmitter는 Base ADU로부터 ADU 프레임을 받아 각 ADU 프레임마다 설정된 전송 시간 및 Payload Type 등의 파라미터를 참조하여 네트워크 상으로 전송한다. Media와 ADU는 최대 32개의 단위 Media 및 ADU 프레임 객체를 저장할 수 있도록 하여 동시에 부하가 집중되는 ADU 생성이나 Frame 생성에서 손실되는 CPU 점유 시간을 보상 받을 수 있다. Media 부분은 MPEG 비디오나 H.263 등 각각의 코덱에 따라 저장되는 프레임의 단위가 다르게 되며, ADU는 Payload Type에 따른 파일 포맷에 관계없이 다양한 ADU 프레임을 받아들일 수 있다.

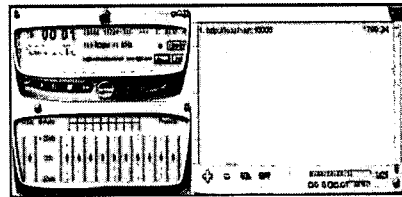


[그림 7] MPEG 오디오 전송 서버 실행 화면

그림 7은 MPEG 오디오 단일 전송의 실행 화면을 보여주며 PLAY, PAUSE, STOP 등의 PAD 모듈의 제어 부분을 가능하게 하는 예제이다.

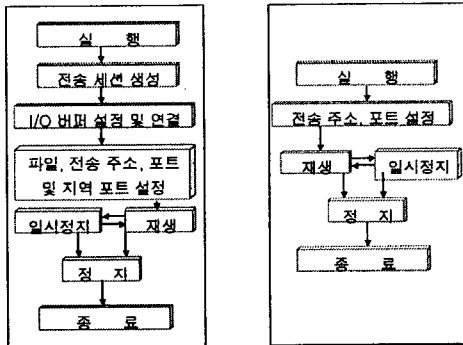
3. 구현 및 고찰

실시간 MPEG 오디오 전송을 위한 PAD 모듈의 실행은 Windows 2000 Server 환경 하에서 펜티엄III Celeron 534MHz, 메모리 256MB, VC++ 이용하여 구현하였다.



[그림 8] 플레이어 재생 화면

그림 8은 MPEG 오디오 단일 전송을 통해 데이터가 전송 중일 때 윈앰프의 RTP plug-in 모듈에서 실제 전송 받은 데이터를 재생하는 화면을 보여주고 있다.

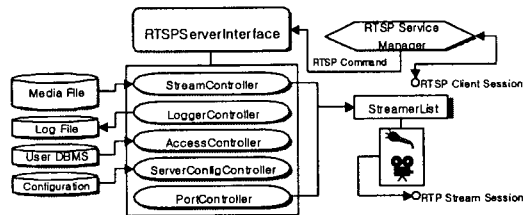


(a) MPEG Audio Streaming Server (b) MPEG Audio Streaming Client

[그림 6] MPEG 오디오 전송 서버 및 클라이언트 시스템 블록도

그림 6은 PAD 모듈의 실행을 위한 미디어 전송 서버와 클라이언트의 구조를 나타낸다. 서버 시스템은 실시간 전송을 위해 RTSP(Real Time Streaming Protocol) 기반의 실시간 스트리밍 전송 서버를 이용했고, 클라이언트는 현재 대중화되어 있는 Nullsoft의 윈앰프 플레이어와 Nullsoft에서 제공하는 MP3를 위한 RTP 송신을 할 수 있는 plug-in 모듈을 사용하였다. 그림 6(a)는 MPEG 오디오 파일 전송 서버의 실행 순서를 나타내고, 그림 6(b)는 클라이언트의 실행 순서를 보여준다.

본 논문에서 구현한 PAD 모듈을 실제 서버 소프트웨어에 사용하기 위해서는 그림 9와 같은 구조로 장착하여 사용하게 된다. Media Server에게로 다수의 RTSP클라이언트가 접속을 수행한 후 RTP Audio/Video 전송 세션을 요청하면 Media Server는 Stream을 위한 다수의 PAD Thread를 생성하게 되고 생성된 Stream Thread는 StreamList객체에서 관리 되는 구조이다.



[그림 9] PAD 모듈을 장착한 RTSP서버의 Interface 구조

그림 9는 구현한 PAD 모듈을 장착한 RTSP 서버의 Interface 구조를 나타낸다. 그림 9와 같은 구조로 장착하여 동시에 여러 개의 PAD 세션의 전송 서비스를 구현할 수 있다.

4. 결론

본 연구에서는 MPEG 오디오 데이터를 RTP 기반의 실시간 전송 포맷으로 전환하여 전송하는 모듈을 설계, 구현하였다. 본 모듈은 MPEG 비디오 영역을 RTP 기반으로 전송하는 시스템과 MPEG 시스템의 전송에 의한 주문형 오디오 및 비디오 시스템(AOD & VOD, Audio & Video On Demand System), 멀티캐스트를 이용한 다자간 화상회의 시스템(Multiple Conference System) 등에 응용할 수 있다.

향후 연구 계획으로는 MPEG-4 기반의 데이터 전송부분과 실시간 전송에서 손실 되는 패킷이 발생할 경우 QoS를 보상 할 수 있는 전송 방식의 개량 부분에 대한 연구가 진행될 예정이다.

본 연구는 2002년도 TIT 학술연구조성비 사업으로 이루어졌습니다.

[참고문헌]

- [1] IETF RFC 1889 RTP: "A Transport Protocol for Real-Time Application", January 1996
- [2] IETF RFC 1890 RTCP: "RTP Profile for Audio and Video Conferences with Minimal Control", January 1996
- [3] 정제창 역, "그림으로 보는 최신 MPEG,", 교보문고, pp.197~220,1995
- [4] IETF RFC 2250: "RTP Payload Format for MPEG1 / MPEG2 Video", RFC 2250, January 1998
- [5] Ross Finlayson, "A More Loss-Tolerant RTP Payload Format for MP3 Audio", IETF Draft, 2000.12
- [6] JMF & JMStudio, Sun Microsystems Co.
<http://www.sun.com/software/communitysource/jmf>
- [7] W.Richard Stevens, TCP/IP Illustrated, Volume 1, Addison Wesley, pp.148~151, 1994.
- [8] Live.com Streaming Media, Live.com,
<http://live.sourceforge.net/>