

멀티미디어 신호에서 오디오/비디오 신호의 동기화에 관한 연구

김 시 호, 이 승 원, 배 건 성
경북대학교 전자·전기공학부

A Study on the Synchronization of Audio and Video Signals

Si Ho Kim, Seung Won Lee, Keun Sung Bae

School of the Electronic & Electrical Engineering, Kyungpook National University
si5@mir.knu.ac.kr , lsw@mir.knu.ac.kr

요 약

본 연구에서는 MPEG 시스템의 동기화 방식에 대해 조사하고, 이를 바탕으로 특정 어플리케이션 개발에 적용 할 오디오/비디오 동기화 시스템을 구현하고자 한다. 먼저 비디오 신호와 오디오 신호를 각각 MPEG-1 시스템의 비디오 압축 방식과 G.722 광대역 음성 부호화 방식을 이용하여 부호화하고, 부호화된 두 미디어간의 동기화를 위해 시간 정보를 삽입하여 최종적으로 하나의 비트스트림으로 다중화 하고, 다중화된 비트스트림으로부터 오디오/비디오 신호를 각각 분리하여 재생 할 수 있음을 보였다.

I. 서 론

인간의 중요한 정보 전달 수단인 청각 정보와 시각 정보를 동시에 제공하는 오디오/비디오 시스템은 최근에 급속히 발달한 멀티미디어 시스템 중에서 대표적인 것이라 할 수 있다. 오디오와 비디오가 결합된 시스템에서 두 신호의 동기가 맞지 않을 경우, 자연스러운 동영상 신호의 재생을 연출할 수 없기 때문에 멀티미디어 신호의 생성과정에서의 오디오신호와 비디오신호 사이의 동기화는 중요한 문제가 된다. 특히, 실시간 통신을 목적으로 하는 응용의 경우 데이터 저장을 목적으로 하는 응용에 비해 두 신호의 동기화는 더욱 중요하다.

본 연구에서는 여러 가지 동영상 포맷 중에서 현재 널리 사용되는 MPEG 시스템[1]에서의 오디오/비디오 신호의 동기화 방법[2,3]에 대해서 분석하고, 이를 바탕으로 임의의 방식으로 압축된 음성신호의 재생시에 필요한 동기화 기법을 개발하고자 한다. 기본적으로 연구의 목표를 CCTV나 동영상 신호의 저장을 위한 어플리케이션 개발에 두어, MPEG 시스템에 사용되는 고품질 오디오 압축 방식은 배제하고, 적당한 음질을 얻을 수 있는 간단한 알고리즘의 사용을 고려하였다. 따라서, 영상 부호화 방식은 MPEG-1의 압축 알고리즘을 사용하였고, 오디오 부호화 방식은 ITU의 광대역 부호화 표준인 G.722[4,5] 방식을 사용하였다. 그리고 부호화된 오디오 및 비디오 미디어간의 동기를 맞추는 방식은 MPEG-1 시스템의 방식을 적용하였다.

본 논문의 구성은 다음과 같다. 서론에 이어 본론에서는 MPEG-1 시스템[2]의 동기화 방식에 대해 설명하고 오디오 및 비디오 신호의 압축 방식을 각각 설명한다. 그리고 부호화된 오디오/비디오 신호간에 동기를 맞추어 다중화/역다중화 과정을 구현한 시스템에 대해 설명한다. 결론에서는 본 논문에서 구현한 시스템의 실행 결과에 대해 언급하고, 향후 연구 방향을 제시한다.

II. MPEG에서의 동기화 기법

현재 많이 사용되는 동영상 포맷 중에서 MPEG (Moving Picture Experts Group)은 영상과 오디오 압축에 관한 국제 표준으로, 여러 동영상 포맷 중에서 현

재 가장 많이 사용된다. 따라서, 본 연구에서는 MPEG 동영상 포맷에서 사용되는 오디오/비디오 동기 기법에 대해서 설명하고, 이를 바탕으로 MPEG 시스템에 다른 오디오 코덱을 적용하여 동기화 시스템을 구현하는 방법에 대하여 설명한다.

MPEG은 비디오와 오디오 신호를 압축하기 위한 국제표준으로 널리 알려져 있으며, 크게 비디오 압축, 오디오 압축, 그리고 시스템 파트로 나누어진다. 그 중에서 MPEG 시스템 파트는 비디오와 오디오 파트에서 만든 스트림(elementary stream)을 저장 또는 전송하기 위하여 패킷화하는 과정으로 시스템 코딩에 관한 전반적인 내용을 기술한다. 이 시스템 코딩 과정은 크게 두 가지로 구별할 수 있는데, 하나는 저장 매체에 저장하기 위한 프로그램 스트림(PS:Program Stream)을 만드는 과정이고 또 하나는 네트워크에서 전송 또는 방송 등을 위한 트랜스포트 스트림(TS:Transport Stream)을 만드는 과정이다. 시스템 인코딩 과정은 단순히 각각의 압축된 오디오, 비디오 스트림들을 묶어 결합시키는 기능뿐만 아니라 스트림을 디코딩하는 과정에서 시스템 디코더 내부의 버퍼제어 및 각 디코딩된 스트림들의 동기를 맞추어 재생하기 위한 변수들이 삽입된다. 따라서, MPEG에서 두 미디어간의 동기를 맞추는 역할은 이 시스템 파트에서 담당하게 된다.

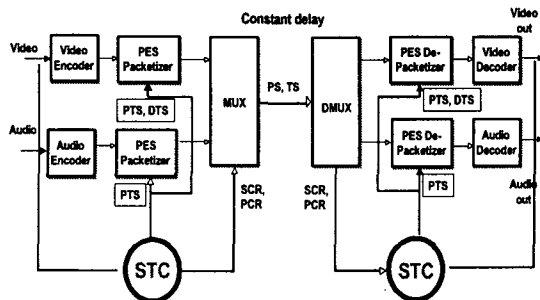


그림 1. MPEG 시스템의 시간 모델

그림 1은 MPEG 시스템에서 비디오와 오디오 스트림 간의 동기를 맞추기 위한 시간 모델을 나타내고 있다. MPEG 시스템에서는 시스템 디코딩 과정에서 두 미디어간의 동기를 맞추기 위하여 특정 시계(STC: System Time Clock)를 가지고 시간 정보를 코딩한다. 즉, 시스템 디코더에서는 STC를 기준으로 해서 비디오 또는 오디오가 재생되어야 하는 시점을 비디오 정보 또는 오디오 정보에 함께 실어서 보낸다. 여기서, 각 신호가 재생되는 시점을 PTS(Presentation Time Stamp)라고 하는데, 각 신호가 패킷화되는 과정에서 헤더에 삽입되어 전송된다. 이때, 각 스트림의 모든 프레

임에 PTS정보가 삽입되는 것은 아니며, PTS 정보가 없는 프레임에 대해서는 보간을 통해서 적당한 재생 시점을 찾게 된다. 또한 인코더에서는 PTS의 생성에 기준 시계로 사용되는 STC에 대한 정보를 PCR(Program Clock Reference) 또는 SCR(System Clock Reference)의 형태로 코딩하여 지속적으로 전송하는데, 표준안에서는 전송 주기를 PCR과 SCR에 대해서 각각 0.1 sec, 0.7 sec 미만으로 정하고 있다. 또한 PTS와 함께 DTS(Decoding Time Stamp)도 시스템 인코딩 시에 첨가되는데, 이는 비디오 또는 오디오 신호가 디코딩되어지는 시점을 나타낸다. 오디오 신호일 경우에는 디코딩되어야 할 시점과 재생되어야 할 시점이 같지만, 비디오 같은 경우에는 디코딩 시점과 재생 시점이 다르기 때문에 두 가지를 구별하여 전송하게 된다.

한편, 디코더에서는 전송된 시간정보(PCR, SCR)를 참고로 하여 자신이 가지고 있는 STC를 이 시간정보에 맞추고, 오디오/비디오 신호의 PTS를 이 STC와 비교하여 재생시점을 정하게 된다. 이때, STC에 사용되는 시간은 시스템 클럭을 사용하며 42 비트로 표현된다. MPEG-1과 MPEG-2의 차이점은 MPEG-2에서는 27MHz로, MPEG-1에서는 90kHz로 시스템 클럭이 코딩된다는 점이다.

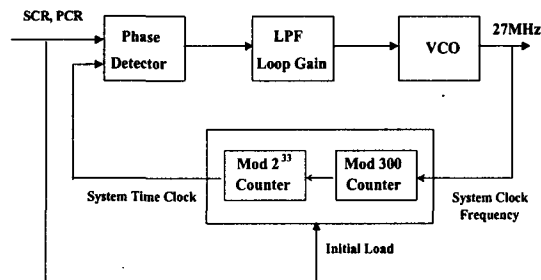


그림 2. 디코더에서 시스템 시계의 복원

그림 2는 실제 디코더에서 시스템 클럭을 복원하는 과정을 나타내고 있다. 전송된 SCR 또는 PCR이 디코더의 STC와 다를 경우, 전송된 시계 정보 PCR에 STC를 맞추기 위해 디코더에서는 PLL(Phase Lock Loops)을 사용한다. 수신된 PCR과 디코더의 STC에 차이가 생길 경우, VCO(Voltage Controlled Oscillator)의 값에 변화가 생기고 이 값에 의해 STC 클럭은 조절되게 된다.

III. 동기화 시스템의 구현

본 연구에서 동기화 시스템의 구현에 사용된 비디오 코덱은 MPEG-1 비디오 코덱[6]이며, 오디오 코덱

으로는 ITU-T의 표준안인 G.722 방식을 사용하였다. 또한 두 미디어간의 동기화 부분을 담당하는 시스템 인코딩은 MPEG-1 시스템을 사용하였다. MPEG-1 시스템에서의 미디어 다중화 방식은 MPEG-2에서 PS에 해당하는 것으로, 각 미디어의 디코딩된 스트림을 패킷화하는 과정에서 PTS 또는 DTS 정보가 첨가되고, 두 미디어를 섞어 팩(pack)으로 만드는 과정에서 SCR 정보가 삽입된다.

MPEG-1 동영상 압축 방식은 JPEG의 화면내 부호화 기법과 H.261의 화면간 부호화 기법을 결합하여 개선한 방법으로 움직임이 적은 영상을 기본으로 하며, 1.5 Mbps의 낮은 전송률에 기반을 두고 있다.

G.722 부호화 방식은 광대역 음성 음질 수준의 부호화기의 필요성에 의한 CCITT의 연구 결과로 1988년에 채택된 방식이다. G.722 부호화 방식은 16 kHz로 샘플링된 신호를 이용하여 48, 56, 64 kbps 세 가지 전송속도를 가지도록 오디오 신호를 압축하며, 과형부호화 방식을 기반으로 하기 때문에 일반적인 협대역 음성 부호화 방식에 비해 다양한 종류의 오디오 신호를 원신호에 보다 충실하게 합성할 수 있는 방식이다. 또한 특정 모델을 선택하여 특징을 추출하는 과정이 없기 때문에 연산량도 그다지 크지 않다는 장점이 있다. 이러한 점들을 고려하여 본 연구의 시작 단계에서 목표한 어플리케이션 개발에 적합한 방식으로 선택하였다.

원래의 G.722 방식은 샘플 단위로 동작하지만 본 연구에서는 일정한 길이(20 msec)의 프레임 단위로 동작시키고, 각 프레임마다의 부호화된 신호에 비디오 신호와의 동기를 위한 신호를 첨가하였다. 그림 1은 헤더 정보가 붙은 부호화된 오디오 신호를 나타내는데, 헤더에 기록된 정보는 시스템 인코딩 과정에서 PTS 또는 DTS 생성에 필요한 정보를 제공한다.

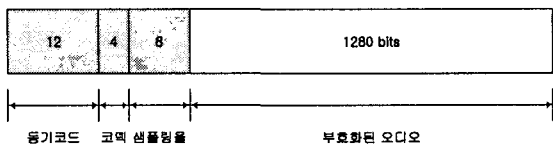


그림 3. 동기 정보가 첨가된 오디오 신호

그림 3의 첫 번째 블록은 프레임 단위로 처리되는 오디오 신호의 프레임 사이의 식별을 위해 할당된 동기 코드이다. 새로운 프레임을 알리기 위한 정보를 12 비트의 연속된 1 값을 이용하여 저장하였다. 이 정보를 이용하여 시스템 디코더는 12 개의 연속된 1의 값이 나오면 20 msec 길이의 새로운 프레임으로 인식하고 이를 처리한다. 두 번째 영역은 오디오 코덱을 구분하기 위한 식별자로서, MPEG-1에서 정의되어 있는 세 가지

종류[7](Layer1, 2, 3) 이외의 방식임을 나타내고 있으며, 구현한 시스템에서는 4 비트를 할당하여 나타내었다. 그리고 오디오 코덱에서 사용한 16 kHz의 샘플링율을 나타내기 위해 8 비트를 할당하였다. 이들 세 가지 동기를 위한 정보에 이어 20 msec 단위로 처리된 오디오 신호 자체의 정보를 덧붙여서 오디오 비트스트림을 구성하였다.

그림 4에 본 연구에서 구현한 단순화된 오디오/비디오 동기 시스템의 구조도를 나타내었다. 이를 이용하여 MPEG-1 비디오 신호와 G.722 오디오 신호에 동기화를 위해 시간 정보를 삽입하여 최종적으로 하나의 비트스트림으로 다중화 하고, 다중화된 비트스트림으로부터 오디오/비디오 신호를 각각 분리하여 재생할 수 있음을 확인하였다.

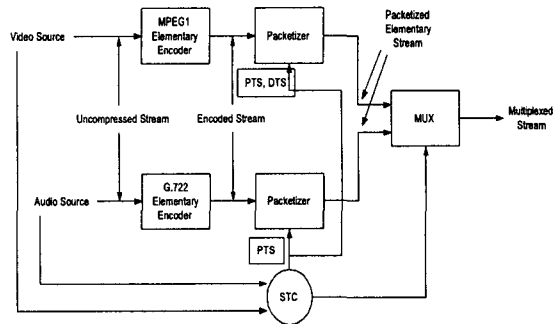


그림 4. 구현한 동기화 시스템의 구조도

IV. 결론

본 연구에서는 CCTV나 적당한 음질을 갖는 동영상 신호의 저장 시스템 등을 위한 어플리케이션 개발의 기초연구로서 오디오/비디오 신호의 동기화 시스템의 구현에 관하여 연구하였다. 이를 위해 가장 보편적으로 사용되는 동영상 압축 방식인 MPEG-1 시스템에서의 동기화 방식에 대한 조사가 선행되었으며, 이를 바탕으로 MPEG 비디오 압축 방식을 기본으로 하여 G.722 광대역 음성 부호화 방식을 적용한 동기화 시스템을 구현하였다. 앞서 언급한 것처럼, 본 연구를 바탕으로 하여 임의의 압축 알고리즘을 적용하여 얻은 오디오 및 비디오 신호의 동기를 맞추어 전송/저장 및 재생 할 수 있는 시스템 구현을 계속해서 연구할 계획이다.

참고 문헌

[1]Barry G. Haskell, Atul Puri, and Arun N. Netravali, *Digital Video: An Introduction to*

MPEG-2, Chapman & Hall.

- [2]"Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s : System," ISO/IEC 11172-1: (November 1991).
- [3]"Generic Coding of Moving Pictures and Associated Audio Information: Systems," ISO/IEC 13818-1 (April 1995).
- [4]"7 kHz AUDIO - CODING WITHIN 64 KBIT/S," ITU-T Recommendation G.722, Nov. 1988.
- [5]Xavier Maitre, "7 kHz Audio Coding within 64 kbits/s," IEEE Journal on Selected Areas in Commun., pp. 283-298, Feb. 1988.
- [6]"Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s : Video," ISO/IEC 11172-2: (November 1991).
- [7]"Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s : Audio," ISO/IEC 11172-3: (November 1991).