



특 집 2부 | 멀티미디어 스트리밍 기술의 핵심 1

# 비디오·오디오 압축 기술의 그 깊은 곳

## 멀티미디어 스트리밍

과 몇 년 전만 해도 조금은 생소한 용어였지만, 질이 낮은 음악 샘플을 간단하게 웹에 링크용으로 제공하던 수준의 기술이었고, 일부 상회의 솔루션으로 사용된 것이 전부였다. 인터넷 방송, 인터넷 폰, 화상채팅, 화상메일 용에 일상생활로 파고들고 있다.

이러한 변화는 필자의 십년지기 친구를 구는 군입대 전, 텔레비전, 비디오, 오디오를 즐기고 있었다. 하지만 제대 후 성능 좋은 대신했다. 공중파 TV 시청, 음악, 애니메이션 등 친구의 모든 취미 생활은 보다 폭이것은 현재까지의 멀티미디어 스트리밍 차지한다는 목표를 어느 정도 달성했음을 멀티미디어 스트리밍 기술은 이제 친구의 외출길에도 함께할

스트리밍 기술은 '현재 어떻게 할 것인가'라는 1차적인 문제를 넘어 대용량의 스트리밍 서비스를 어떻게 할 것인가, 스트리밍 서비스의 품질을 높이는 방법은 무엇인가, 복제 방지 문제와 과금 문제를 어떻게 해결해야 할 것인가라는 2차적인 문제로 넘어섰다. 이 글에서는 이런 문제를 해결하기 위한 멀티미디어 스트리밍 기술의 핵심이라 할 수 있는 비디오·오디오 압축 기술에 대해 알아본다.

임익진, 이재욱, 이승재 경희대 실시간 멀티미디어 연구실 limga@oslab.kyunghee.ac.kr

## 기술의 개요

멀티미디어의 기본 단위인 데이터의 기본 단위를 비트라고 한다. 스트리밍 데이터의 전송은 대용량의 멀티미디어 데이터를 전송하는 데 필요한 대역폭을 확보하는 것이 중요하다. 스트리밍 데이터 전송은 실시간 전송과 세션 제어에 HTTP와 자체 프로토콜을 사용하고 있다. 스트리밍 데이터 전송에 관한 네트워크 표준안은 주로 이곳에서 제정되며, RTP/RTCP는 ITU-T의 H.323, H.324 등의 프레임워크를 갖고 있으며, 주로 회의 및 VoIP(Voice over IP)에 MPEG, RTP, SIP 등과 함께 사용된다. 현재의 대세인 MPEG-4에 의해 H.323에 관한 활동은 소강상태에 있다. MPEG의 경우 MPEG-1, MPEG-2에서는 멀티미디어 데이터의 압축만 다뤘는데, MPEG-4로 넘어오면서 통신에 관한 부분도 다루고 있다. 이것은 MPEG-4 자체가 멀티미디어 통신을 전제로 만들어졌기 때문이다. 현재의 시장 기술은 마이크로소프트의 독자적인 WMT 기술을 위시한 스트리밍 기술과 표준 스트리밍 기술을 응용 단계에서 사용자의 편의성을 위해 비표준을 절충하는 기술로 양분된다.

를 줄이기 위한 멀티미디어 압축 기술과 실시간 전송을 지원하는 네트워크 기술로 양분할 수 있다. 하지만 이렇게 양분된 기술은 따로 작동하는 것이 아닌 서로 밀접하게 관련돼 작동한다는 점을 유의해야 한다.

멀티미디어 스트리밍 기술에 관한 표준안은 IETF(Internet Engineering Task Force), ITU(International Telecommunication Union)의 ITU-T(International Telecommunication Union-Telecommunication), MPEG(Moving Picture Experts Group)에서 제정하고 있으며, 각각은 성립 목적이 다른 만큼 다음과 같이 다른 특성을 갖고 있다.

- ◆ IETF : 인터넷과 관련된 표준화 추진
- ◆ ITU-T : 통신에 관한 표준화 추진
- ◆ MPEG : 멀티미디어 전반에 관한 표준화 추진

IETF에서는 스트리밍 전송을 위한 RTP(Realtime Transport Protocol)/RTCP(Realtime Transport Control Protocol)와 스트리밍 제어를 위한 RTSP(RealTime Streaming Protocol)와 회의 참여, 호 관리를 위한 SIP(Session Initiation Protocol), 세션의 기술과 분배를 위해 SDP(Session Description Protocol), SAP(Session Announce Protocol) 등의 규격을 내놓았다. 인터넷에서 멀티미디어 데이터 전송에 관한 네트워크 표준안은 주로 이곳에서 제정되며, RTP/RTCP는 ITU-T의 H.323, H.324 등의 프레임워크를 갖고 있으며, 주로 회의 및 VoIP(Voice over IP)에 MPEG, RTP, SIP 등과 함께 사용된다. 현재의 대세인 MPEG-4에 의해 H.323에 관한 활동은 소강상태에 있다. MPEG의 경우 MPEG-1, MPEG-2에서는 멀티미디어 데이터의 압축만 다뤘는데, MPEG-4로 넘어오면서 통신에 관한 부분도 다루고 있다. 이것은 MPEG-4 자체가 멀티미디어 통신을 전제로 만들어졌기 때문이다.

현재의 시장 기술은 마이크로소프트의 독자적인 WMT 기술을 위시한 스트리밍 기술과 표준 스트리밍 기술을 응용 단계에서 사용자의 편의성을 위해 비표준을 절충하는 기술로 양분된다.

## 데이터 압축(Codec, Encode/Decode) 기술

멀티미디어는 영상과 소리를 중심으로 문자, 도형, 음성 등의 정보를 디지털 기술로 압축해 결합한 것이다. 음성의 경우 각

각의 정보를 단순히 PCM(Pulse Code Modulation)에 의해 디지털로 표현하면 데이터 양이 방대해진다. 전화음은 64Kbps(8KHz 표본화 주파수에 표본당 8비트), 스테레오 음악을 담은 CD는 1.5Mbps(44.1MHz 표본화 주파수에 표본당 16비트로 2채널), 고선명 HD TV의 경우 스튜디오 규격인 SMPTE 240MHz에 의하면 휘도 성분은 74.25MHz로 표본화되고 두 색 성분은 그 반의 주파수로 표본화된다. 각각 표본당 8비트이므로 발생하는 데이터량은 총 1.2Gbps에 이른다. 멀티미디어 스트리밍에 있어 앞에서 말한 것과 같은 방대한 양의 각종 데이터를 어떻게 압축할 것인가는 핵심 요소 기술이 된다. 세계적으로 과거부터 현재까지 이런 막대한 양의 데이터를 갖는 영상, 음성 등의 정보 유형에 따른 효율적인 데이터 압축 방식의 개발을 위해 치열한 경쟁을 벌였다. 또 한편으로는 국제 표준 데이터 압축 방식의 제정을 위해 국제표준화기구(ISO)나 국제전기통신연합(ITU) 등을 통해 공동의 노력을 기울여왔다.

## 멀티미디어 압축 기술 1, 비디오 코덱

비디오 압축, 즉 부호화와 복원을 위한 비디오 코덱의 국제 표준으로는 H.261, H.263, MPEG-1, MPEG-2, MPEG-4가 있다. 여기서는 이런 표준 코덱에 사용되는 기본적인 압축 원리를 알아보고, 각 표준에 대한 기본 개념과 중요한 요소 등을 살펴본다.

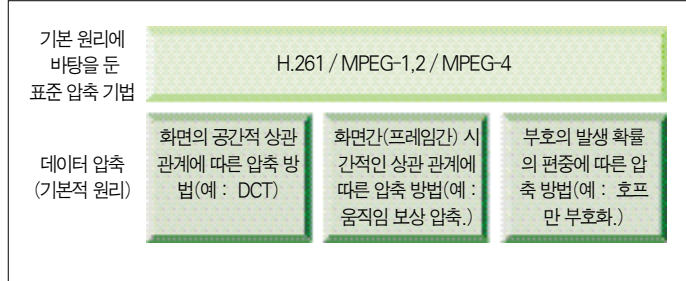
## 표준 코덱에 사용되는 기본 원리

데이터 압축은 데이터에 내재된 중복성을 없애고 중요하고 필요한 것만을 뽑아 데이터의 용량을 줄이는 과정으로 볼 수 있다. 하지만 이렇게 원본의 1/10 이상으로 압축돼 전송되는 작은 데이터는 원본 영상과 차이가 없어야만 한다. 이런 코덱의 압축 기술을 구현하기 위해 하나의 기법만 사용한다는 것은 불가능하다. 영상 데이터의 압축에 사용되는 기술은 DCT(Discrete Cosine Transform), 양자화, 런 길이

(표 1) 멀티미디어 압축 표준

	표준	내용 및 응용분야	표준화 기구
영상	H.261	ISDN을 이용한 영상전화나 영상회의를 위한 영상 표준	ITU-T
	H.263	전화망을 이용한 영상전화에 위한 표준	ITU-T
	MPEG-1	CD에 1.5Mbps로 동영상을 담기 위한 규격	ISO 11172
	MPEG-2	고품질 영상 부호화 위한 압축 표준	ISO 13818 ITU-T
음성	CMIT 721/723	방송국에서 프로그램을 제작하는데 쓰일 소재를 고품질로 전송하기 위한 표준	ITU-R
	G.711	PCM 방식의 음성 부호화	ITU-T
	G.721 G.728	ADPCM 방식의 음성 부호화 저지연 CELP 방식의 음성 부호화	ITU-T ITU-T
음향	G.722	AM 방송 품질	ITU-T
	J.41	FM 방송 품질	ITU-T
	MPEG-1 오디오 MPEG-2 오디오	CD 수준의 MPEG-1 오디오 5.1 채널로 확장한 MPEG-2 오디오	
	AAC(AC-3)	미국의 고선명 TV에 채택됨	업계표준

(그림 1) 동영상 압축의 기본적인 원리와 압축 표준



부호화, 호프만 부호화, 움직임 보상 DPCM(동영상의 경우만 해당) 등 다양하고, 일반적으로 압축기술의 구현은 효과적인 방법을 복합적으로 사용하는 하이브리드 방식을 사용하고 있다.

**화면의 공간적 상관관계에 따른 압축**

한 화면 내에서도 이웃하는 화소끼리는 그 값이 매우 비슷하다. 이것이 화소와 화소 사이에 존재하는 공간적 중복성이다. 이런 공간적 중복성을 없애기 위해 DCT를 일반적으로 많이 사용한다. DCT를 이용한 영상은 공간영역에서 주파수 영역으로 변환할 수 있다. 영상 데이터는 변화가 낮으므로 낮은 주파수, 특히 주파수(DC) 성분이 큰 값을 갖게 되고 높은 주파수 성분은 상대적으로 낮은 값을 갖게 되므로, 대부분의 정보가 낮은 주파수 쪽으로 몰리게 된다. 따라서 양자화 과정을 적절히 거치면 높은 압축율로 우수한 화질을 얻을 수 있다. 이런 압축을 DCT 압축 기술이라고 한다. DCT 압축 기술은 H.261, JPEG, MPEG-1에서 사용되고 있다.

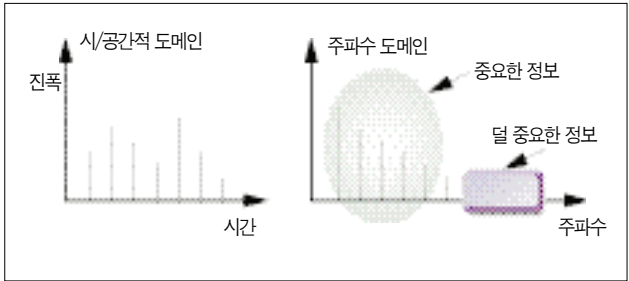
**화면간, 프레임간 시간적 상관관계에 따른 압축**

초당 24장(영화) 혹은 30장(NTSC 컬러 TV)의 화면이 발생할 때 이웃하는 두 장의 화면은 매우 비슷하다. 정지한 경우 완전히 같고 움직임이 있더라도 그 부분을 제외하면 배경은 같다. 이것이 화면과 화면 사이에 존재하는 시간적 중복성이다. 데이터 압축 관점에서 보면 영화나 TV처럼 초당 수 십장의 화면을 취하면 화면간 중복성이 매우 높다. 예를 들어 고정된 장면의 경우 화면 한 장 한 장의 내용이 같으며

(표 2) 연대별 움직임 보상 압축 기법

연대별 움직임 보상 압축 기법	내 용
1970년대(초기연구단계)	이웃하는 화면간에 움직임 부분과 정지한 부분의 영역 구분, 움직임 영역의 영역 정보와 그 안의 내용만 갱신해 보냄
1980년대 (MPEG이전 H.261 등 널리 쓰이고 있는 블록별 움직임 추정 및 보상방법)	화면을 일정한 크기의 매크로 블록으로 나뉘 단위마다 앞 화면의 어느 곳으로부터 움직여 왔는지 움직임 벡터를 구하고 이를 이용해 움직임 보상, 예측함. 현 매크로 블록과 움직임 보상에 의해 얻어진 이전 화면의 매크로 블록간 차이만을 부호화해 데이터 양을 줄임. DPCM과 호프만 부호를 이용한 무손실 압축이 이용됨. (MPEG 등 동영상 압축기술의 효용이 JPEG 등의 정지영상 압축기술보다 크게 높아짐(약 1/4).

(그림 2) DCT 변환



로 첫 화면만 전송하면 다음 화면들은 '앞 화면과 같다'는 단순 정보만으로 완전하게 전송할 수 있다. 또 움직임이 있는 장면에서도 우선 배경 부분은 정지해 있는 경우가 많고 움직인 부분도 '어떤 부분이 어디로 움직였는지'의 정보를 보냄으로써 데이터 양을 크게 줄일 수 있다. 데이터가 압축되지 않는 경우는 장면 전환이 있어 두 화면간 상관성이 없을 때로, 이 때는 어쩔 수 없이 뒤 화면은 앞 화면의 정보를 사용하지 않고 자체 화면 내에서만 압축한다. 이런 방법으로 압축하는 것이 움직임 보상 압축 방법이다.

**부호화 발생확률의 편중에 따른 압축**

양자화된 움직임 보상 DCT 계수는 통계적으로 어떤 값은 자주 나오고 어떤 값은 희박하게 나타난다. 이것이 앞서 파일 압축의 경우에도 해당되는 통계적 중복성이다. 이것을 없애기 위해서는 무손실 부호화를 이용하는데 MPEG에서 사용하는 호프만 부호화과 파일 압축시 사용하는 LZW 알고리즘이 대표적인 예다.

- ◆ 양자화 : DCT를 통해 얻어진 DCT 계수 값 등을 어떤 상수들로 나눠 유저자리의 비트 수를 줄이는 과정
- ◆ 호프만 부호: 호프만 부호화 방식은 통계적 부호화 방식이다. 빈번히 발생하는 데이터 코드는 적은 수의 비트로 표현하고, 빈번하지 않은 데이터는 상대적으로 많은 비트 수로 표현해 전체 데이터의 크기를 줄이는 방식. 정지 화상과 동영상에 사용한다.

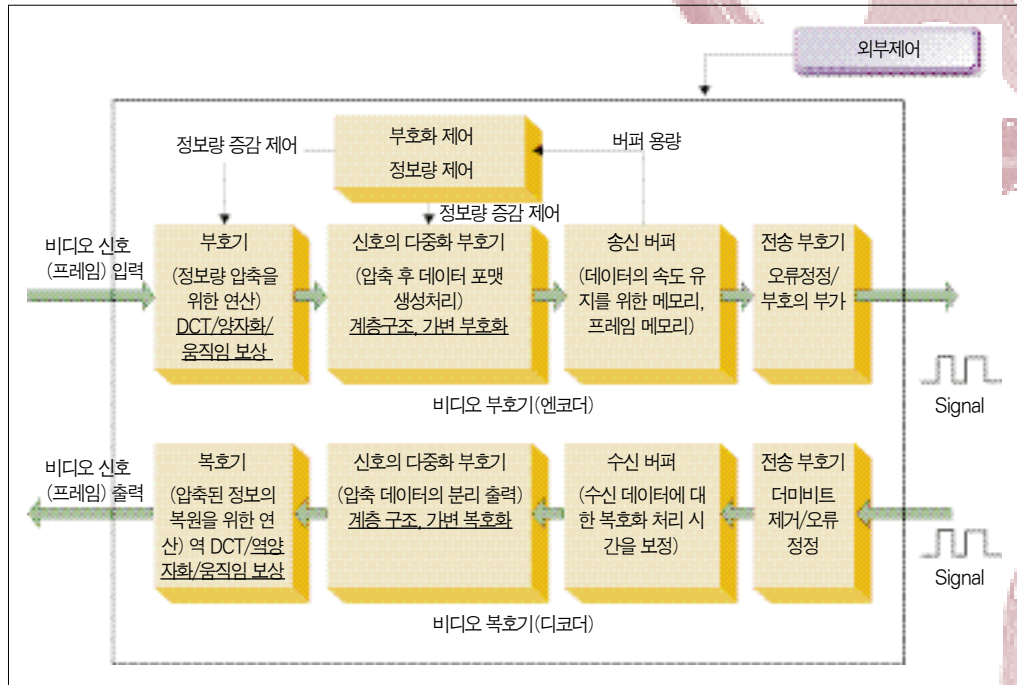
**H.261**

H.261은 ITU-T가 제정한 국제표준으로 ISDN 네트워크에서 이용한 영상전화 및 영상회의를 위한 국제 비디오 압축 표준이다. H.261의 비디오 데이터는 초당 64Kbps 배수로 압축되는데, 이를 흔히 64\*P라고 하며 여기서 P 값은 ISDN 채널의 수에 따라 1~30 사이에서 정해진다. 이 표준은 본래 비디오 폰과 화상회의의 등을 지원하기 위해 개발된 것이다.

**H.261의 기본원리**

H.261에서는 Transform 변환기법으로 DCT를 프레임간 예측기법으로 DPCM(Differential Pulse Code Modulation)을 결합한 '하이

(그림 3) H.261의 하이브리드 방식의 인코더/디코더



(그림 4) H.261로 부호화된 연속된 프레임



브리드 부호화(Hybrid Coding)'를 사용한다. 또한 움직임 보상 기법도 추가할 수 있다.

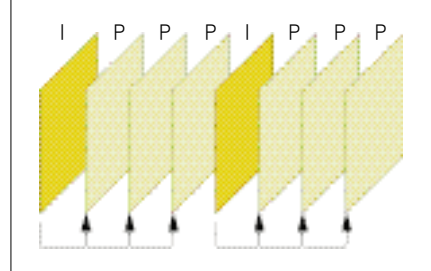
**H.261의 부호화 단계**

영상에서 두 프레임의 차이를 매크로 블록화시키고 8\*8 블록화한 후 DCT 변환을 한다. DCT 변환을 통해 얻은 정보는 양자화와 Run Length Encoding를 거쳐 마지막으로 호프만 부호화(Huffman Coding)으로 부호화된다.

**H.261 코딩 기법**

H.261은 인터 프레임(Inter frame)과 인트라 프레임(Intra frame)

(그림 5) H.261로 부호화된 연속된 프레임



방식을 지원한다. 인트라 프레임 방식은 이전의 화상으로부터 현재의 화상을 예측하는 기법이다. 즉, 이전의 화상과 현재의 화상의 차이만 부호화한다. 이것은 MPEG의 P-프레임과 비슷하다. 인트라 프레임 방식에서는 입력화상이 인접한 다른 화상과는 독립적으로 DCT 부호화된다. 이것은 MPEG의 I-프레임 부호화와 비슷하며, JPEG 코딩과도 비슷하다.

- ◆ 인터 프레임 방식(P-프레임) : P-프레임(forward predicted frames)은 이전 프레임을 참조하는 예측 프레임이다. P-프레임은 연속되는 이미지들의 전체 이미지가 바뀌는 것이 아니라 이미지의 블록들이 옆으로 이동한다는 점에 착안한 것이다. 즉, 움직임이 있는 경우 앞 화면에 있는 물체 자체의 모양에는 큰 변화 없이 옆으로 이동

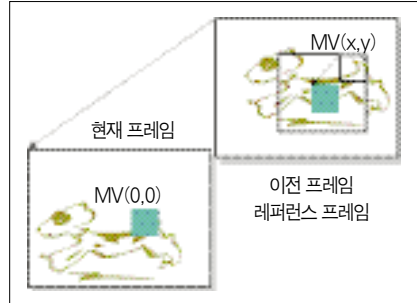
하는 경우가 대부분이므로, 이전의 화면과 현재의 화면의 차이가 매우 적은 것을 이용해 차이 값만을 부호화하는 것이다.

- ◆ 인트라 프레임 방식(I-프레임) : I-프레임(intra frames)은 앞 또는 뒤의 다른 프레임과는 관계 없이 독립적으로 단 하나의 영상으로 부호화된다. 이 부호화 체계는 JPEG 압축과 유사하다. 각 8\*8 블록은 독립적으로 부호화되고, 신호를 별개의 주파수대역으로 분리시키는 DCT를 사용하지 위하여, 먼저 블록은 공간 영역으로부터 주파수 영역으로 변환된다. 대부분의 주파수 정보는 8\*8 블록의 좌측상단 코너에 있다.

**MPEG**

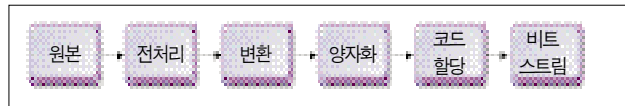
1980년대부터 디지털 AV 서비스의 국제 표준화의 필요성에 따라, 국제적인 표준 방식을 정하는 국제표준화기구(ISO, International Organization for Standardization)와 국제전기표준회(IEC, International Electrotechnical Commission)의 합동기술위원회

(그림 6) 현재 프레임과 이전 프레임의 차이





〈그림 7〉 MPEG의 동영상 압축 절차



(JTC1, Joint Technical Committee 1) 산하의 전문부회(SC29, Sub-Committee 29)에서는 영상신호 압축 부호화에 관한 MPEG(Moving Picture Experts Group) 표준화 작업을 시작했다. 그로부터 지금까지 MPEG은 CD 등 디지털 저장매체에 VHS 테이프 화질의 동영상과 음향을 1.5Mbps 이내로 압축, 저장하기 위한 MPEG-1과 DTV, HDTV 등과 같이 방송용으로 사용하기 위한 고품질 동영상 및 오디오 압축기술(3-10 Mbps)인 MPEG-2를 표준화했다. 또한, MPEG은 64Kbps, 낮은 비트율인 PSTN이나 이동통신 전화에 응용하기 위한 압축으로 멀티미디어 응용제품 전반에 걸쳐 사용 가능한 압축기술인 MPEG-4를 표준으로 내놓았다. 현재 MPEG-4는 계속 연구 중에 있다.

### MPEG 비디오

MPEG 비디오 표준은 다음의 두 가지로 구성된다.

- ◆ 비디오 : JPEG & H.261
- ◆ 오디오 : MUSICAM 테크놀로지

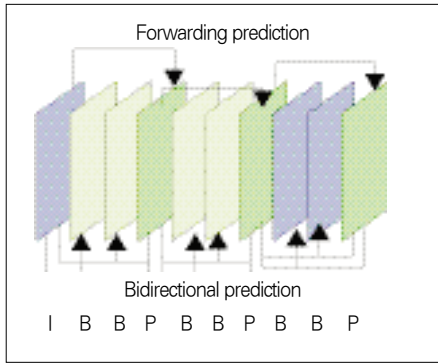
MPEG-1,2의 압축 알고리즘은 JPEG 정지화상 압축 알고리즘과 H.261 동화상 압축 알고리즘 등을 발전, 계승시킨 것이다. MPEG에서는 모든 프레임을 개별 정지화상으로 압축하는 것이 아니라, 인접 프레임 사이에 유사점이 많다는 점을 이용한다. 즉, 동작보상을 하는데 있어 예측과 보간을 이용한다. 그러나 임의접근과 같은 VCR식 제어가 가능해야 하는 등 여러 이유로 인해 MPEG 비디오에서는 자신이 갖고 있는 정보만으로도 복원될 수 있는 프레임이 규칙적으로 삽입돼야 한다. 이러한 프레임은 JPEG과 아주 유사하게 정지화상으로 압축된다. MPEG의 동영상 압축 절차는 다음과 같다.

- ◆ 전처리 : 컬러 스페이스(Color Space) 변환, 필터링(Filtering), 컬러 서브 샘플링(Color Subsampling) 등의 작업
- ◆ 변환 : 영상이 갖고 있는 정보의 중복성을 찾아내는 과정, 파형 방식과 변환 방식이 있음
- ◆ 양자화 : DPCM이나 DCT 과정을 통해 얻은 영상 값을 어떤 실수 값으로 나눠 유효자리의 비트 수를 줄이는 과정

〈표 3〉 변환 과정의 종류와 비교

변환방법	처리방법	쓰임새
파형 방식	영상 신호 자체에서 처리	DPCM(Differential Pulse Code Modulation) ADPCM(Adaptive Differential Pulse Code Modulation)
변환 방식	새로운 영역에서 처리	DCT(Discrete Cosine Transform)

〈그림 8〉 MPEG 비디오의 부호화된 연속된 프레임 패턴



- ◆ 코드 할당 : 빈도수에 따른 처리방법으로 호프만 부호화 방법 등이 사용되는 과정

### MPEG의 프레임 타입 3가지

1 I-프레임(Intra-coded frame) : 데이터 스트림의 어느 위치에도 올 수 있으며, 데이터의 임의 접근을 위해 사용되며, 다른 이미지들의 참조 없이 부호화된다. I-프레임은 정지화상 압축방법(JPEG)을 이용하지만, JPEG과는 달리 MPEG에서는 실시간으로 압축이 이뤄진다. I-프레임의 압축은 MPEG에서는 가장 낮은 압축률을 보인다. I-프레임은 매크로 블럭 내에서 지정된 8\*8 블럭으로 나누고, DCT 기법을 사용한 후, DC 계수는 DPCM 방법으로 부호화하는데, 연속한 블럭 사이의 차이 값을 계산한 후 가변 길이 코딩을 사용해 변환한다.

2 P-프레임(Predictive-coded frame) : 부호화와 복호화를 행할 때 이전의 I-프레임 정보와 이전의 P-프레임의 정보를 사용한다. P-프레임은 연속되는 이미지들의 전체 이미지가 바뀌는 것이 아니라 이미지의 블럭이 옆으로 이동한다는 점에 착안한 것이다. 즉, 움직임이 있는 경우 앞 화면에 있는 물체 자체의 모양에는 큰 변화 없이 옆으로 이동하는 경우가 대부분이므로, 이전의 화면과 현재의 화면의 차이가 매우 적은 것을 이용해 차이 값만 부호화하는 것이다.

3 B-프레임(Bidirectional-coded frame) : 부호화와 복호화를 행할 때 이전, 이후의 I-프레임과 P-프레임 모두를 사용한다. B-프레임을 사용하면 높은 압축률을 얻을 수 있다. B-프레임은 이전의 I-또는 P-프레임과 B-프레임 이후의 I-또는 P-프레임의 차이 값을 가진다.

### H.261과의 차이점

MPEG 비디오의 핵심기술은 이전 프레임과 현재의 프레임의 차이를 이용해 움직임을 추정하고 이를 보상해주는 ME/MC(Motion Estimation/Motion Compensation) 기법과 유효 데이터를 최소화하기 위한 변환 부호화인 DCT의 적용이라고 볼 수 있다. 움직임(동작)보상과 DCT라는 부호화의 기본적인 구조는 H.261과 같지만, 기본적으로 H.261은 통신매체를 대상으로 하고 MPEG1은 저장매체를 위한 것이기 때문에 섬세한 부분에서는 몇몇 부분에서 차이가 난다.

MPEG은 I와 P-프레임 외에 〈그림 8〉과 같이 Bidirectional 프

레이미라는 또 하나의 B-프레임을 추가한다.

### MPEG-2

MPEG-2 표준은 기술적으로는 H.261 표준과 MPEG-1 표준의 연장선상에 있지만, 이러한 기존의 표준들보다 훨씬 다양한 영역에 적용될 수 있으며, 다음과 같은 특징을 갖는다.

- ◆ 높은 화질의 영상을 제공할 수 있기 때문에 고선명 TV 화질까지 포함할 수 있음
- ◆ 각 화면을 프레임 단위나 필드 단위로 처리할 수 있음
- ◆ 순차 주사(Progressive Scanning)의 영상, 격행 주사(Interlaced Scanning)의 영상 처리 가능
- ◆ 화면의 크기나 프레임율이 다른 경우를 동시에 효율적으로 처리할 수 있는 계층적 분해 능력(Scalability)을 가짐

MPEG-2는 뛰어난 성능과 유연성을 가진 특성으로 인해 멀티미디어 스트리밍에 중요한 부분을 차지하고 있다. 또한, MPEG-2는 자체에 오디오와 비디오의 계층화, 부호화 및 복호화 지원, 저장 및 전송 시의 잠음에 의한 에러 대책, MPEG-1 및 H.261 표준과의 순방향 호환성(Forward Compatibility), 랜덤 액세스 및 채널 변경(Channel Hopping), 앞/뒤로 가기/정지/빨리 가기 등 특수 효과, 서라운드 오디오(Surround Audio)를 위한 다 채널 음향(Multi-Channel Audio) 및 다국어 음성 수용, 여러 프로그램의 다중화(Multiplexing), 암호화(Encryption), 편집 기능(Editing), ATM(Asynchronous Transfer Mode) 전송과 같은 가변 데이터율(VBR, Variable Bit Rate) 처리 등을 지원하기 때문에 복잡한 구조이지만 스트리밍에 있어 중요한 부분을 차지하고 있는 압축 기술이다.

### MPEG-4

MPEG-4는 1994년 64Kbps 이하의 전송율에서 고품질의 영상을 제공할 수 있는 동영상 압축 표준을 목표로 출발한 것으로 고성능 멀티미디어 통신 서비스를 고려해 기존 방식뿐만 아니라 새로운 기능을 지원하기 위한 영상의 내용에 근거한 영상 신호의 부호화를 목표로 하고 있는 압축 표준으로 주로 멀티미디어 통신을 목표로 한다.

### MPEG-4의 특징

MPEG-4 표준은 제작자, 서비스 공급자, 사용자의 요구를 만족시키기 위한 방법을 제공하려는 데 초점을 두고 있다. 제작자에게는 제작상의 최대 효율을 올리기 위해 디지털 TV나 애니메이션 그래픽, 웹 페이지와 같은 내용을 다루는 데 있어 훨씬 재사용성이 높고 유연한 작업을 할 수 있도록 했고, 서비스 제공자들은 MPEG-4를 이용해 서비스 제공 지역에 따라 적합하게 해석돼 서비스가 제공될 수 있도록 투명하게 정보를 활용할 수 있도록 했다. 사용자에게는 제작자가 설

정해 놓은 범위 내에서 최대한 상호 작용할 수 있도록 했다.

MPEG-4는 표준화를 위한 4가지 목표를 달성했다. 첫째, 소리와 영상 정보는 소리/영상 객체(AVO, Audio/Visual Objects)라 하고, 이들이 동기화될 수 있도록 해 마이크, 카메라 컴퓨터 등의 서로 다른 매체를 통해 만들어진 정보를 하나로 모을 수 있도록 했다.

둘째, 장면에서 소리/영상 객체가 서로 결합돼 하나의 객체를 이룰 수 있도록 했다. 셋째, 이들을 서로 다중화 및 동기화를 통해 네트워크 채널들로 전송될 수 있도록 했다. 마지막으로 장면들을 사용자가 이들 소리/영상 객체를 이용해 상호 작용할 수 있도록 했다.

MPEG-4는 또한 멀티미디어 전송 통합 프레임워크(DMIF, Delivery Multimedia Integration Framework)를 응용 프로그램과 네트워크 사이에 뒤 네트워크를 통한 원격 응용 프로그램에 명령을 내릴 수 있도록 편의성을 제공한다. 이 DMIF 구조를 구현해 단순한 인터페이스를 가진 응용 프로그램이 자세한 네트워크에 대한 고려를 하지 않아도 되게 한 것이다.

이러한 다양한 기능을 하나의 알고리즘으로 구현하는 일은 불가능하기 때문에 MPEG-4에서는 많은 압축 요소를 표준 메뉴 형식으로 수용해 응용에 따라 선택할 수 있도록 했다. 즉, 압축에 필요한 도구들을 정하고, 이 도구를 결합해 여러 압축 알고리즘을 만들며, 하나 이상의 알고리즘과 프로파일들, MSDDL(MPEG-4 Syntactic Description Language)이라는 언어를 새로 만들어 정의한다. 따라서 MPEG-4 응용 프로그램간의 데이터 송수신은 우선 상태가 어떤 프로파일, 알고리즘, 도구의 복호기를 갖고 있는지 확인해 복호 가능한 모드로 교신하고 필요한 경우, 복호에 필요한 프로그램을 먼저 다운로드한 후에 내용물을 전송한다.

〈표 4〉 MPEG-4의 주 적용 대상

대상분야	응용 서비스의 예
영상 전화 (실시간 영상/음향 통신)	개인간 통신 다자간 영상회의
멀티미디어 (영상/음향 프로그램 재생 및 검색)	대화형 멀티미디어 데이터베이스 멀티미디어 비디오텍스트 멀티미디어 주석 멀티미디어 발표(슬라이드 쇼)
원격 감시 (영상/음향 데이터의 획득 및 감시)	가정, 빌딩 학교의 모니터링 교통 상황 모니터링 현장 전문가의 영상입력 이동차량 및 로봱(piloted and pilotless)
비디오 저장과 전송	멀티미디어 전자우편 영상 전화 응답기
교육	자습용/교육용 비디오 수업에의 응용 대화형 훈련
오락	오락 비디오, 여행자 오락



**MPEG-4의 응용**

MPEG-4는 실시간 영상, 음향 통신이 중요 기술이 되는 영상 전화, 영상 회의 등의 분야에 가장 쉽게 적용될 수 있다. 또한 영상, 음향 프로그램 재생, 대화형 멀티미디어 데이터베이스, 멀티미디어 비디오 텍스트, 슬라이드 쇼 등에도 응용될 수 있다. 그 밖에 가정, 학교 등에서 모니터링을 위해 사용될 수 있으며, 멀티미디어 전자우편, 교육용 비디오, 대화형 훈련, 오락 등 매우 많은 응용 범위를 가질 수 있다. (표 4)는 MPEG-4의 주 적용 분야를 보여준다.

**멀티미디어 압축 기술 2, 오디오 코덱**

스트리밍 기술을 두 가지로 분류했을 때 비디오 스트리밍 부분이 상당히 중요한 비중을 차지한다. 하지만 비디오 스트리밍이 사람의 시각에 중점을 뒀다고 하면, 사람의 청각에 중점을 둔 오디오 스트리밍 또한 비디오 스트리밍만큼 중요하다.

일반적으로 오디오는 아날로그 형태로 된 사운드를 표본화(Sampling)와 양자화(Quantizing) 과정을 거쳐 디지털 사운드 형태로 저장한 것을 말한다. 이렇게 변환된 디지털 사운드 데이터는 매우 크기 때문에 일반적으로 압축해서 저장한다.

아날로그 사운드는 기본적으로 입력된 값을 그대로 저장하는 방법의 PCM 방식으로 표본화돼 디지털 사운드로 변환된다. 이 방식은 압축과정을 거치지 않기 때문에 용량이 매우 크며, 높은 음질을 요구하는 CD나 DAT(Digital Audio Tape) 등에서 사용된다. PCM 방식은 큰 용량으로 인해 일반적으로 사용되지 않고, 압축을 통해 저장매체에 저장 또는 스트리밍에 이용하게 된다. 압축 방식에 따라서 A-Law,  $\mu$ -Law, G.721, G.722, MPEG 오디오 등 다양한 형태를 갖고 있다.

**MPEG-1 오디오**

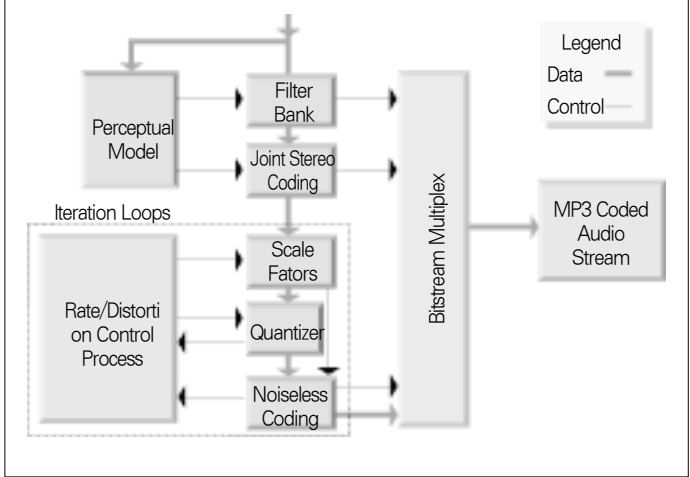
MPEG1 오디오는 높은 품질에 효율적인 스테레오 부호화를 위한 ISO/IEC에서 권고한 표준방식으로 MPEG내 동영상 압축과 병행해 표준화됐다. MPEG-1 오디오의 정식 표준번호는 ISO/IEC IS 1117 2-3으로 1992년에 국제표준이 되었고, Layer-1, 2, 3 이라는 세 가지 모드를 갖고 있다.

레이어-1은 네델란드 필립스가 DDC(Digital Compact Cassette)를 위해 개발한 PASC 압축 알고리즘에 사용됐으며, 레이어-2와 레이어-3으로 갈수록 보다 발전되어 더 큰 압축으로도 우수한 음질을

(표 5) MPEG 오디오의 압축 모드

음질	주파수 대역	모드	비트율	압축률
전화 수준	2.5KHz	Mono	8Kbps	96 : 1
단파 방송 수준	4.5KHz	Mono	16Kbps	48 : 1
AM 방송 수준	7.5KHz	Mono	32Kbps	24 : 1
FM 방송 수준	11KHz	Stereo	56~64Kbps	26~24 : 1
TAPE 수준	15KHz	Stereo	96Kbps	16 : 1
CD 수준	15KHz 이상	Stereo	112~128Kbps	14~12 : 1

(그림 9) MPEG 오디오의 블록 다이어그램



낼 수 있게 됐다. Layer-3의 경우에는 디지털 오디오 방송을 위해 개 발됐으며, 현재 가장 널리 사용되고 있다.

- ◆ Layer-1 : 압축률 1:4, 스테레오 신호의 경우 384Kbps에 해당
- ◆ Layer-2 : 압축률 1:6~1:8, 스테레오 신호의 경우 256~192Kbps에 해당
- ◆ Layer-3 : 압축률 1:10~1:12, 스테레오 신호의 경우 128~112Kbps에 해당

MPEG-1 오디오에서 사용하는 압축 알고리즘은 32 밴드에 기초한 서브밴드 코딩(대역분할부호화)과 MDCT(Modified Discrete Cosine Transform : 변형 이산 여현변환)를 사용한다. 큰 소리와 상대적으로 낮은 소리가 동시에 들릴 때는 낮은 소리는 큰 소리에 묻혀 잘 들리지 않게 되는 청각적인 인식 효과(마스킹 효과, Masking Effect)를 이용해 높은 효율의 압축을 이뤄낸다. 이러한 압축 방식을 이용해 (표 5)와 같은 다양한 압축 모드로 원하는 형태로 압축해 사용할 수 있다.

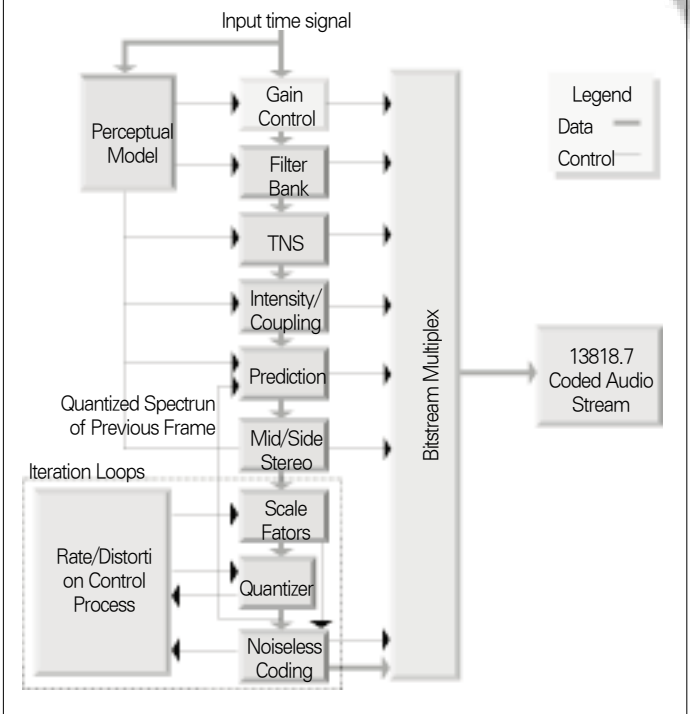
**MPEG-2 AAC(Advanced Audio Coding)**

MPEG-2 AAC 국제 표준은 차세대 디지털 오디오 부호화 기술로 인정 받고 있는 기술로서 디지털 텔레비전 방송, 디지털 오디오 방송, 인터넷 오디오, 다운로드 음악 등에 활용할 수 있는 기술이다. MP3(MPEG-1 Audio Layer-3)보다 우수한 압축율과 오디오 품질을 갖는 것으로 평가 돼, 차세대 디지털 오디오 플레이어의 핵심 기술로 예측되고 있다.

MPEG-2 AAC는 다 채널 고품질화를 목표로 표준화가 진행돼 1994년 11월에 IS1383-3으로 승인돼 표준화가 완결됐다. AAC에는 MPEG-1 오디오를 바탕으로 압축 효율을 높이기 위한 고분해된 필터 뱅크, 강력한 오디오 지각 모델, 역 방향 적응 예측, 조인트 채널 부호화 및 허프만 부호화 등 효율적인 부호화 방법을 결합한 방법을 이용했다.

MPEG-1 오디오와 비교할 때 MPEG-2 AAC는 다음과 같은 특징이 포함돼 있다. 첫째, 다중 채널화했다는 점인데, MPEG-1 오디오의 스테레오 기능이 MPEG-1.2 AAC에서는 6채널까지 확장돼 영화

(그림 10) MPEG-2 AAC의 블록 다이어그램



관과 같은 입체음향을 그대로 만끽할 수 있게 되었다. 6채널의 성분들(즉, 스피커의 위치를 살펴보면 C(Center), L(Left), R(Right), LS (Left Surround), RS(Right Surround)의 5개 광대역신호(20KHz)와 저주파 성분(120KHz)만을 별도로 제공하는 LFE(Low Frequency Enhancement) 신호로 이뤄져 있다. 통상 이를 5.1 채널이라고 부르는데, L, R, LS, RS의 4채널로 이뤄진 기존의 돌비서라운드 입체음향에 비해 더욱 입체감이 향상됐다. 또한 다국어 지원(Multi lingual) 기능이 강화됐다. 아시아, 유럽 등 여러 지역의 위성방송에서도 볼 수 있듯이 멀티미디어 및 정보통신기술의 발전에 따라 지구촌이 점차 하나가 되고 있다. 이에 대응해 MPEG-1.2 AAC는 5.1 채널 외에 별도로 7개 국어까지의 부가 음성을 보낼 수 있는 기능이 들어있다.

둘째, AAC는 MPEG1 오디오에서 사용된 샘플링 주파수의 반에 해당하는 16KHz, 22.05KHz, 24KHz의 샘플링 주파수를 사용할 수 있도록 하고 있다. 이는 한정된 비트율에서 멀티채널 및 다국어 지원의 많은 데이터를 효과적으로 압축하기 위해서는 입력신호의 대역이 좁을 경우 표본화 주파수를 줄이는 것이 유리하기 때문이다.

셋째, MPEG-1 오디오의 데이터 구조가 고정적인 반면 AAC의 데이터는 유동적인 구조를 갖는다. MPEG-1 오디오는 곡 전체의 정보를 담은 헤더 뒤에 데이터가 프레임이라는 단위로 저장되는데 이 프레임의 크기가 고정돼 있어 압축률이 높은 부분에서도 쓸데 없는 용량을 차지한다. 이에 반해 AAC의 프레임 구조는 가변구조로 압축률에 따라 크기가 변하므로 전체 파일의 용량이 훨씬 줄어든다. 실제로 MP3 파일과 비교하면 최대 30%까지 용량을 줄일 수 있다.

AAC의 또 다른 장점은 음질에 있다. MPEG-1 오디오에 비해 AAC는 TNS와 프레딕션이라는 두 가지 기법을 통해 음질을 향상한다. TNS(Temporal Noise Shaping)는 양자화 보정 기술로 아날로그의 연속적인 음악 데이터를 0과 1의 디지털 데이터로 만들 때 생기는 오차를 지능적으로 줄여 잡음을 감소시키고 원음에 가깝게 만든다. 프레딕션(Prediction)이란 TNS로 보정된 수치를 기억하는 것으로 앞에서 보정된 정보를 기억해 다음에 같은 데이터가 나타날 때 기억된 데이터를 사용한다. 만일 양자화 단계에서 같은 음의 보정 수치가 다르다면 다른 소리로 들리게 되므로 이를 같게 만드는 기술이다.

이와 같이 MPEG-2 AAC는 MPEG-1 오디오보다 압축율과 음질 면에서 우위에 있지만 아직 해결해야 할 문제가 적지 않다. 그 중 한 가지가 인코딩 시간이다. 원음을 AAC 포맷으로 만드는 인코딩 시간이 MPEG-1 오디오에 비해 최대 열 배 이상 걸린다. 물론 인코딩 소프트웨어의 성능 향상으로 이 문제는 차츰 개선되겠지만 인코딩 시간이 느려 AAC의 데이터를 만드는 데 많은 시간이 소요된다는 치명적인 약점을 갖고 있다.

**마치면서**

WMT내에서는 필터라고 불리우는 컴포넌트(일종의 DLL)로 스트리밍 기반 기술을 제공하고 있다. 하지만 스트리밍 기술의 모든 것을 WMT에서 제공해 주지는 않을뿐더러, MS의 특성상 표준안을 수정한 것이 대부분으로 플랫폼이 어긋날 경우 호환성에서 문제를 갖고 있는 것들도 있다. 또한 인터페이스가 한정된 상태에서의 개발은 자칫 개발자를 나락의 구렁텅이로 몰고 갈 위험성도 있다. 하지만 그렇다고 WMT가 가지는 모든 것을 항상 포기하고 스스로 모든 것을 만들어야 한다는 말 또한 어불성설이다. 여기서 개발자에게 필요한 덕목은 WMT 기술 특성과 표준화된 요소 기술을 동시에 볼 수 있는 안목이 요구된다 할 수 있겠다. 그리고 이 글을 통해 그러한 시야를 갖고자 하는 독자에게 작은 도움이라도 됐으면 하는 바람이다. **황**

정리 : 강경수 elegy@sbmedia.co.kr

**참고 자료**

- 1 IETF 홈페이지, <http://www.ietf.org>
- 2 ITU-T 홈페이지, <http://www.itu.int/ITU-T/>
- 3 Mpeg.org 홈페이지, <http://www.mpeg.org>
- 4 Mpeg 위원회 홈페이지, <http://www.cselt.it/mpeg/>
- 5 스타더스트 홈페이지, <http://www.stardust.com>
- 6 Multimedia Systems, Standards and Networks, Atul puri & Tsuhan Chen, Marcel Dekker
- 7 Multimedia Communications/Application Networks, Protocols and Standards), Fred Halsall, Addison-Wesley
- 8 IP Telephony Packet-based multimedia communication system, Olivier Hersent, David Gurlle & Jean-Pierre Petit, Addison-Wesley
- 9 Packet Video Communications over Networks, Z.S. Bojkovic & Kamisetty R, Rao, Prentice Hall