

## 목 차

1. 서 론
2. MPEG-1
3. MPEG-2
4. MPEG-4 AVC/H.264
5. VC-1
6. 결 론

## 문양세 · 노용기

(강원대학교 · University of Minnesota)

## 1. 서 론

텔레비전(TV)은 일반인들의 생활에 가장 많은 영향을 미치는 가전제품 중의 하나다. TV는 1940년대에 처음 소개된 이후 컬러 방송과 같은 몇 가지 경우를 제외하곤 괄목할만한 변화가 없었다. 하지만, 최근 IPTV(Internet Protocol TV)가 많은 일반인뿐만 아니라 학문적으로도 큰 관심을 끌고 있다. 그 이유는 IPTV가 전통적인 TV와는 달리 고품질의 프로그램(영화와 드라마 등)을 일반 가정에서 대화적(interactive)이고 개인화(personalized)된 방식으로 시청할 수 있도록 하기 때문이다. 즉, 전통적인 TV가 방송국에서 송출하는 신호를 시청자의 TV에서 받아 화면에 출력하는 획일적이고 단방향성의 시스템이었다면, IPTV는 사용자의 요구를 IPTV 서비스 회사에 보내고 IPTV 회사는 그에 따라 사용자마다 서로 다른 맞춤형 서비스를 제공하는 양방향성의 시스템이다.

이러한 IPTV 서비스를 가능하게 한 가장 중요한 두 가지 요인은 (1) 네트워크 기술의 발전 및 대중화, (2) 멀티미디어 코덱(codec) 및 코딩 기

술의 발전을 들 수 있다. ADSL(Asymmetric Digital Subscriber Line)과 VDSL(Very High Speed DSL) 등의 기술로 인하여 전화선을 통하여 일반 가정에 수십 Mbps까지의 대역폭을 지원할 수 있고, 일부 인터넷 서비스 회사는 광섬유를 일반 가정까지 연결하여 100Mbps를 초과하는 대역폭까지도 지원하고 있다. 또한, 이러한 인터넷 서비스를 이용하는 가정이 급속히 증가하여 인터넷을 통한 잠재적인 IPTV 서비스 시장이 확대되었다<sup>1)</sup>. 이러한 초고속 인터넷도 압축되지 않은 동영상을 전송하기에는 역부족이다. 예를 들어, NTSC(National Television System Committee) TV 화면을 디지털화하여 전송한다면 약 120Mbps의 대역폭이 필요하다<sup>2)</sup>. 고품질 HD(High Definition) 화면의 경우 더욱 높은 대역폭을 요구할 것이다. 따라서, 대용량의 동영상을 효율적으로 압축하는 기술이 필요하며, 1990년대 이후 현재까지 매우 활발히 연구되

1) 2004년 3월 현재, 한국의 초고속 인터넷 보급율은 70%를 넘었다.

2) NTSC 화면은 DVD 화면과 동일하게 720x480의 해상도를 갖는다. 각 픽셀을 12비트로 표현하고 초당 30 프레임으로 구성된다면,  $720 \times 480 \times 12 \times 30 = 120\text{Mbps}$ 를 얻는다.

고 있고, MPEG-1/2/4, VC-1 등의 동영상 압축 기술이 표준으로 제정되었다. 이들 동영상 압축 표준을 이용하면 초당 120Mbps를 요구하던 NTSC 방송 화면은 1.5 ~ 8Mbps, HD 화면의 경우 5 ~ 20Mbps의 대역폭으로 전송 가능하다.

본 고에서는 IPTV 서비스를 제공하기 위한 MPEG-1/2, MPEG-4 AVC/H.264, VC-1 동영상 압축 표준에 대하여 각각 제 2 ~ 5 절에서 개략적으로 설명한다. 이러한 동영상 압축 표준들은 각각 개별적으로 설계되기도, 이전보다 발전된 또는 유사한 압축 기법(tool)을 포함하기도 한다. 동영상 압축 기술의 근본적인 원리는 동영상에 내재되어 있는 중복성(redundancy)을 제거하는 것이다. 이러한 중복성은 공간적(spatial), 시간적(temporal), 코딩(coding) 중복성으로 구분할 수 있다. 공간적 중복성이란 하나의 프레임(또는 화상; frame, picture) 내에서 인접한 픽셀들은 대개 유사한 색상을 가진다는 점이고, 시간적 중복성은 시간적으로 연속되는 프레임은 대개 유사한 이미지로 구성된다는 점이다. 코딩 중복성은 코딩 값들의 분포가 균등하지 않다는 점이다<sup>3)</sup>. 본 고에서는 각 동영상 압축 표준이 이러한 중복성을 어떻게 제거하여 전체적인 동영상 스트림을 어떻게 구성하는지 살펴본다.

## 2. MPEG-1

MPEG(Motion Picture Experts Group)은 원래 ISO/IEC(International Organization for Standardization / International Electrotechnical Commission) 산하의 동영상 압축 표준을 제정하기 위한 전문가 모임의 이름이었으나 나중에 그 표준의 이름으로 굳어졌다. 1988년에 첫 모임을 가졌고 1990년대 들어 점차 많은 전문가가 참여하게 되었다. MPEG-1 표준은 CD-ROM에 저장되고 1.5Mbps 정도의 비트율(bit ratio)을 가지는 동영상을 목표로 진행되었고, 1991년에 세 파트(part)로 최종 고정되었다.

MPEG-1을 비롯한 동영상 표준은 RGB 색상 공간(color space) 대신 YUV 색상 공간을 사용한다. 하나의 픽셀을 Y, U, V 세 개의 값으로 표현하는데, 여기에서 Y는 밝기(또는 휘도; brightness, luminance), U와 V는 색상(또는 색차; color, chrominance) 값을 갖는다. RGB 픽셀로부터 YUV 픽셀 값을 구하기 위한 공식은 다음과 같다:

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

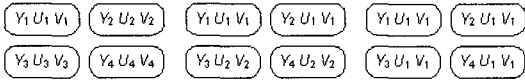
$$U = 0.493 \times (B - Y)$$

$$V = 0.877 \times (R - Y)$$

동영상 표준에서 YUV 색상 공간을 사용하는 이유는 인간의 눈이 밝기보다는 색상 변화에 덜 민감하기 때문에 U와 V 값을 적은 수의 비트로도 표현할 수 있기 때문이다. YUV 색상 공간에 기반하여 하나의 픽셀 색상을 표현하기 위한 포맷은 CCIR(Comité Consultatif International de la Radio) 601 표준에 의하여 4:4:4, 4:2:2, 4:2:0 세 가지 포맷으로 정의되어 있다. 4:4:4 포맷은 모든 픽셀이 독립된 Y, U, V 값을 갖는 포맷이고, 4:2:2와 4:2:0 포맷은 Y 값만이 각 픽셀마다 독립적으로 정해지고, U, V 값은 인접한 둘 또는 네 픽셀이 공유한다. (그림 1)은 각 포맷에 대하여 공유하는 U, V 값을 보인 것이다. (그림 1(b))에서 4:2:2 포맷은 좌우로 인접한 두 픽셀이, (그림 1(c))에서 4:2:0 포맷은 상하좌우로 인접한 네 픽셀이 U, V 값을 공유한다. 따라서, 세 포맷에 있어서 하나의 픽셀을 나타내기 위한 평균 비트 수는 각각 24, 16, 12 비트이다.

하나의 MPEG-1 동영상은 연속된 임의 개수의 화상 그룹(Group of Pictures, GOP)으로 구성되어 있다. 하나의 GOP는 하나 이상의 프레임으로 구성되어 있으며, 각 프레임은 코딩 방법에 따

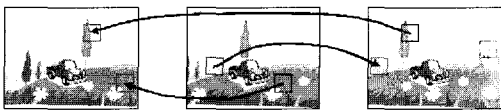
3) 코딩 중복성은 Huffman Coding 기법에 의하여 제거한다. Huffman Coding 기법은 빈도가 높은 코딩 값에 적은 비트를, 빈도가 낮은 코딩 값에 많은 비트를 할당함으로써 평균적인 코딩 비트 수를 감소시킨다.



(a) 4:4:4 포맷 (b) 4:2:2 포맷 (c) 4:2:0 포맷

(그림 1) CCIR 601 픽셀 포맷

라 I-프레임(I-frame, Intraframe), P-프레임(P-frame, Predictive frame), B-프레임(B-frame, Bidirectional frame)으로 구분한다. 이들 프레임 간의 차이점을 설명하기 위하여 (그림 2)의 연속된 동영상 프레임의 예를 이용한다. (그림 2)의 동영상 프레임 내에서 카메라가 오른쪽으로 패닝(panning)하여 배경 화면이 왼쪽으로 움직이고, 앞쪽의 자동차는 동시에 오른쪽으로 움직인다. 프레임 (a)는 GOP가 시작되는 프레임이다. 프레임 (c) 내에서 배경 및 자동차의 상당 부분이 프레임 (a)에도 똑같이 포함되어 있음을 알 수 있다. 따라서, 프레임 (c)는 프레임 (a)로부터 많은 부분을 참조함으로써 표현할 수 있다. 프레임 (b)도 프레임 (c)와 마찬가지로 다른 프레임으로부터 유사한 부분을 참조함으로써 표현할 수 있다. 다만 프레임 (b)가 프레임 (c)와 다른 점은 앞 뒤의 프레임을 동시에 참조한다는 점이다. (그림 2)의 프레임 (a), (b), (c)는 각각 I-프레임, B-프레임, P-프레임에 해당하고, 화살표로 가리키는 블록은 참조하는 블록을 의미한다.

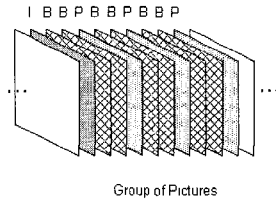


(a) 첫번째 프레임 (b) 두번째 프레임 (c) 세번째 프레임.

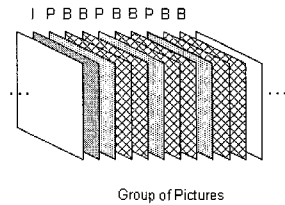
(그림 2) MPEG-1 동영상 프레임

I-프레임과 P-프레임은 다른 프레임으로부터 참조되므로 참조 프레임(reference frame)이라 부른다. B-프레임은 다른 프레임으로부터 참조되지 않는다. MPEG-1 동영상을 중간부터 재생할 경우, 다른 프레임을 참조하는 P-프레임과

B-프레임은 참조되는 프레임이 없다면 재생이 불가능하다. 따라서, 가장 최초로 발견되는 I-프레임으로부터 재생을 시작한다. 만약 I-프레임의 개수가 너무 적다면 중간부터 재생하기 위한 시간이 많이 걸릴 것이고, 만약 너무 많다면 충분한 압축률을 얻기 어려울 것이다<sup>4)</sup>. 대개 초당 2개의 I-프레임을 포함한다. (그림 3)은 MPEG-1 GOP 내의 프레임을 보인 것이다. (그림 3(a))는 프레임을 시간적 순서에 따라 나열한 것이다. 하나의 GOP 내에는 하나의 I-프레임만이 포함되고, P-프레임과 B-프레임의 개수는 고정되어 있지 않다. (그림 3(b))는 GOP 내의 프레임들이 실제로 저장, 전송되는 순서를 보인 것이다. B-프레임을 재생하기 위해서는 P-프레임이 필요하므로, P-프레임이 시간적으로 뒤라 하더라도 먼저 저장, 전송된다.



(a) 프레임의 시간적 순서



(b) 프레임의 전송 순서

(그림 3) MPEG-1 GOP 내의 프레임

하나의 MPEG-1 프레임은 16x16 크기의 이미지 블록들로 분할되어 각각 개별적으로 코딩된다. 이러한 블록을 매크로블록(macroblock)이라

4) I-프레임은 다른 프레임을 참조하지 않으므로 좀더 많은 코딩 데이터를 요구한다. 일부 인코더에서는 I-, P-, B-프레임의 코딩 데이터의 크기 비율을 140:52:36로 가정한다.

부르며, 다른 프레임을 참조하기 위한 단위도 매크로블록이다. I-프레임 내의 모든 매크로블록은 JPEG 표준과 거의 유사한 방법으로 8x8 블록으로 다시 분할되어 각각 DCT(Discrete Cosine Transform) 및 양자화(quantization) 코딩된다<sup>5)</sup>. P-프레임과 B-프레임 내의 각 매크로블록 B에 대하여 다른 참조 프레임으로부터 B와 가장 유사한 같은 크기의 블록 B'을 찾는다. 매크로블록 간의 유사성 척도로는 MMSE(Minimum Mean Square Error) 또는 MAD(Minimum Absolute Difference)가 많이 사용된다. 비교 대상 블록은 참조 프레임 내에서 매크로블록 B와 같은 위치에 있는 블록을 중심으로 300x150 크기의 영역 내의 모든 가능한 16x16 블록이다. 즉, 하나의 매크로블록 B가 참조할 블록 B'을 찾기 위하여 최악의 경우 45,000 번의 블록 비교 연산이 필요하다. 이러한 비교 연산은 하나의 P- 또는 B-프레임 내의 모든 매크로블록에 대하여 수행하여야 하므로, 최대 17,820,000 (= 300\*150\*396) 번의 비교 연산이 일어난다. 초당 30 프레임으로 구성되어 있다고 한다면 1초 간의 동영상을 코딩하기 위하여 534,600,000 번의 블록 비교 연산을 수행하여야 한다. 이러한 과정을 모션 추정(motion estimation)이라 부르고, MPEG 코딩에 있어서 가장 많은 시간이 필요로 한다. 모션 추정을 거쳐 매크로블록 B가 참조할 블록 B'을 발견하면 B로부터 B'로의 벡터를 코딩한다. 이를 모션 벡터(motion vector)라고 부른다<sup>6)</sup>. 실제로, 두 블록 B와 B'이 매우 유사하더라도 동일하지는 않으므로, B와 B'의 대응 픽셀 간의 차이(residual)를 추가적으로 JPEG 표준과 같이 DCT 및 양자화 코딩한다. 만약 매크로블록 B와 충분히 유사한 블록 B'을 찾지 못한다면, I-프레임 매크로블록을 코딩하는 같은 방법으로 B를 코딩한다. 이를 인트라코딩(intracoding)이라 부른다.

<표 1>은 MPEG-1 파라미터를 정리한 것이

다. MPEG-1에서 각 프레임의 가로, 세로 해상도는 크게 잡을 수 있지만, 프레임당 매크로블록의 개수를 초과해서는 안 된다. MPEG-1 동영상에서 가장 많이 사용되는 해상도는 352x288 @ 24 프레임 또는 352x240 @ 30 프레임이다. 특정 MPEG-1 인코더(encoder)에서 생성한 동영상이 표 1의 파라미터 값을 하나라도 만족하지 않는다면 MPEG-1 디코더(decoder)는 그 동영상을 재생할 책임이 없다.

<표 1> MPEG-1 파라미터.

파라미터	값의 범위
가로 해상도	≤ 768 픽셀
세로 해상도	≤ 576 픽셀
한 프레임당 매크로블록 개수	≤ 396
초당 매크로블록 개수	≤ 9900 (= 396 * 25)
초당 프레임 개수	≤ 30
비트율	≤ 1,856,000 bps

### 3. MPEG-2

MPEG-2는 MPEG-1이 고정된 1991년에 작업이 시작되어 1995년에 최종 표준화되었다. MPEG-2는 기본적으로 MPEG-1에 많은 융통성을 부가하여 확장한 표준으로, 모든 MPEG-2 디코더는 MPEG-1 동영상도 재생 가능하다. MPEG-2는 MPEG-1에 비하여 80 Mbps까지 보다 높은 대역폭을 지원하고, 다중 오디오 채널을 포함하며, 인터레이스(interlace) 화면과 다양한 중형비(aspect ratio)의 HD 화면을 지원한다.

5) MPEG 표준은 JPEG 표준과 달리 8x8 블록을 코딩하기 위한 디폴트 양자화 테이블(quantization table)을 정의하고 있다. 양자화 테이블은 양자화 계수들을 포함하는 8x8 크기의 테이블이고, 계수 값이 클수록 양자화에 의한 압축률은 높아지나, 화질은 저하된다.

6) 대개의 경우, 인접한 모션 벡터는 유사한 크기와 방향을 가진다. 따라서, 각 모션 벡터를 개별적으로 코딩하는 것이 아니라, 인접한 모션 벡터 간의 차이만을 코딩한다. 이때, 모션 벡터 차이는 거의 0에 가까운 경우가 많으므로 코딩을 위해 할당할 비트 수를 줄일 수 있다. 이러한 코딩 기법을 예측 코딩(predictive coding)이라 한다.

MPEG-1에 비하여 확장된 좀더 세부적인 MPEG-2 표준 내용으로는 다음과 같은 것들이 있다. MPEG-1에서는 모션 추정을 위하여 픽셀 단위로 매크로블록을 비교하였으나, MPEG-2에서는 선형 보간법(linear interpolation)에 기반하여 반 픽셀 단위로 매크로블록을 비교한다. 픽셀 색상 표현을 MPEG-1에서는 4:2:0 포맷을 썼으나, MPEG-2에서는 4:4:4 및 4:2:2 포맷으로도 표현할 수 있다. MPEG-1에서는 DCT 변환 후에 계수를 최대 8 비트를 이용하여 표현하였으나, MPEG-2에서는 최대 10 비트를 이용할 수 있다. MPEG-2에서는 에러가 발생하여 매크로블록이 소실된 경우 다른 매크로블록으로 대체하여 그 에러를 은폐(concealment)하기 위한 기법도 도입되었다. MPEG-2는 16:9 종횡비의 동영상을 지원함과 동시에 이러한 동영상이 4:3 종횡비의 화면에서 출력되는 경우를 대비하여 어떤 영역을 출력할 것인지를 나타내는 Pan & Scan 정보를 함께 코딩한다.

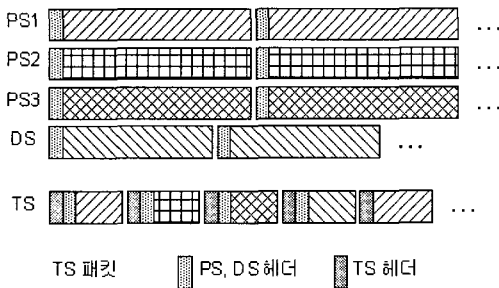
MPEG-2 표준 내용 중에는 MPEG-2로 코딩된 동영상과 오디오를 전송하기 위한 시스템 단계(systems layer)가 정의되어 있다. 시스템 단계는 다시 프로그램 스트림(Program Stream, PS)과 전송 스트림(Transport Stream, TS) 방식으로 구성된다. PS는 DVD와 같이 안정된 환경에서의 전송을 위한 방식이며, TS는 ATSC(Advanced Television Systems Committee) 방송 신호와 같이 에러가 발생할 여지가 있는 환경에서 다중 채널을 전송하기 위한 방식이다. TS 방식에서는 에러 처리 등의 내용이 포함되어 PS 방식에 비해 더 복잡하며, TS 방식을 통해 여러 채널 신호가 동시에 전송되므로 신호 수신자 측에서는 원하는 채널을 효율적으로 찾아 추출해 내는 것이 중요하다. MPEG-2 시스템 단계는 MPEG-4 표준에서도 그대로 사용되고 있다.

MPEG-2 시스템 단계에서는 동영상 및 오디오 인코더를 통해 출력된 스트림과 데이터 스트

림(data stream, DS)을 기본 스트림(elementary stream)이라 부른다. DS는 연관된 프로그램과 관련된 데이터를 저장하는 스트림으로, DVD 자막(caption)이나 서브타이틀(subtitle)을 예로 들 수 있다. 기본 스트림은 가변 길이의 연속된 짧은 스트림으로 분할되고, 각각의 앞에 헤더(header)를 붙인다. 이를 패킷화된 기본 스트림(Packetized Elementary Stream, PES)이라 부른다. PES는 PS 및 TS 방식에서의 공통 최소 단위이다. PS 방식에 의한 하나의 PS는 연속된 임의 개수의 팩(pack)으로 구성된다. 하나의 팩은 헤더와 하나 이상의 PES로 구성되고, 2KB의 길이를 가진다. 이 길이는 하드디스크 또는 DVD에서 한번에 읽어 들이는 데이터의 양과 같도록 의도적으로 정해진 것이다. 팩 헤더에는 SCR(System Clock Reference) 데이터가 포함되어 있는데, 팩 내에 포함된 PES를 언제 재생할 것인지 디코더 클럭과 동기화를 이루기 위한 것이다. 하나의 프로그램은 최대 16 개의 동영상, 32 개의 오디오, 16 개의 데이터 스트림을 포함할 수 있다.

TS 방식에 의한 하나의 TS는 다중 채널을 지원하기 위하여 하나 이상의 PS 및 이들에 대한 설명을 포함하는 기술 데이터(descriptive data)로 구성된다. TS는 연속된 임의 개수의 TS 패킷(TS packet)으로 분할되고, 하나의 TS 패킷은 4 바이트 헤더와 함께 항상 188 바이트 길이를 갖는다. 각 TS 패킷 헤더에는 PID(Program Identification) 코드가 들어 있고, 해당 TS 패킷 내에 포함된 스트림이 어떤 프로그램에 속하는지 찾는 데에 필요하다. 하나의 TS 패킷 내에는 한 종류의 스트림만이 포함될 수 있다. 예를 들어, 한 TS 패킷 내에 같은 프로그램이라도 동영상 스트림과 오디오 스트림이 함께 포함될 수 없다. TS 내에는 각 TS 패킷이 어떤 프로그램의 어떤 스트림인지를 나타내는 두 가지 테이블을 포함하고 있다. PAT(Program Allocation Table)는 PID = 0인 TS 패킷 내에 포함되어 있

고, 각 프로그램에 대하여 그에 대한 정보 (이름 등) 및 해당 PMT(Program Map Table)가 포함된 TS 패킷의 PID 정보를 갖고 있다. PMT 내에는 어떤 PID를 가지는 TS 패킷에 어떤 (동영상, 오디오, 데이터) 스트림이 포함되어 있는지 나열하고 있다. TS 수신자는 PAT와 PMT 두 테이블과 PID를 이용하여 원하는 채널 스트림을 추출할 수 있다. (그림 4)는 세 개의 PS와 하나의 DS로 구성된 TS의 예를 보인 것이다.



(그림 4) 세 개의 PS와 하나의 DS로 구성된 TS의 예.

MPEG-2 동영상 표준은 매우 다양한 응용에서 사용될 수 있고, 각 응용마다 압축 기법과 파라미터를 서로 다르게 설정할 수 있다. 응용에 따라 MPEG-2 표준을 일부만 구현해도 무방한 경우도 많다. 예를 들어, 비디오 컨퍼런싱(video conferencing) 응용의 경우 HD 화질의 동영상까지 지원하도록 할 필요는 없다. 따라서, MPEG-2 표준에서는 실제 응용에서 동시에 많이 사용되는 압축 기법의 집합과 파라미터의 집합을 정의하고, 이를 각각 프로파일(profile)과 레벨(level)이라고 부른다. <표 2>와 <표 3>은 각각 MPEG-2 프로파일과 레벨의 일부를 보인 것이다. 이러한 프로파일과 레벨의 모든 조합이 유효하지는 않다. 예를 들어, SP 프로파일과 HL 레벨이 함께 사용될 응용은 거의 없을 것이다. <표 4>는 유효한 프로파일@레벨 조합의 일부를 보인 것이다. 예를 들어, DVD 동영상을 위한 조합은 MP@ML이며, HDTV 방송 신호는 MP@HL이다.

<표 2> MPEG-2 프로파일의 일부

프로파일 이름 (코드)	프레임 타입	픽셀 색상포맷
Simple Profile (SP)	P, I	4:2:0
Main Profile (MP)	P, I, B	4:2:0
High Profile (HP)	P, I, B	4:2:2

<표 3> MPEG-2 레벨의 일부

레벨 이름 (코드)	최대 프레임 해상도	최대 프레임율	최대 비트율 (Mbps)
Low Level (LL)	352x288	30	4
Main Level (ML)	720x576	30	15
High Level (HL)	1920x1152	30	80

<표 4> MPEG-2 프로파일/레벨 유효 조합의 일부

조합 코드	해상도	최대 프레임율	픽셀 포맷	비트율 (Mbps)	응용
SP@LL	176x144	15	4:2:0	0.096	무선 전화
MP@LL	352x288	30	4:2:0	4	셋톱 박스
MP@ML	720x480	30	4:2:0	15	DVD
MP@HL	1920x1080	30	4:2:0	80	HDTV

#### 4. MPEG-4 AVC/H.264

MPEG-4 표준은 MPEG-2에 비하여 보다 다양한 응용을 지원하고 높은 압축률을 얻기 위하여 작업이 진행되었다. MPEG-4 표준화 과정 동안 라이선스 비용에 대한 사용자들의 거부와 MPEG-2 코딩 기술의 눈부신 발전으로 인하여, 한때 MPEG-4 표준화 자체에 대한 실효성이 의심되기도 하였다. 유사한 시기에 ITU-T (International Telecommunication Union - Telecommunication Standardization Sector)에서는 비디오 컨퍼런싱을 위한 낮은 비트율과 높은 압축률의 동영상 압축 표준을 개발 중이었다. ISO/IEC와 ITU-T의 전문가 그룹이 연합하여 JVT(Joint Video Team)를 구성하였고, 이 팀에서 MPEG-4 AVC(Advanced Video Coding) 표준을 제정하였다. 이 표준은 MPEG-4 Part 10에 해당하며, 같은 표준을 ITU-T에서는 H.264라 부른다. 최근 들어 MPEG-4 표준이라고 하면 대

개의 경우 MPEG-4 AVC/H.264를 의미한다.

MPEG-4 표준은 MPEG-2 표준에서의 여러 압축 기법들을 전수받음과 동시에 다양한 새로운 압축 기법을 소개하였다<sup>7)</sup>. 이러한 새로운 기법들 중의 몇 가지를 설명한다. MPEG-1 및 MPEG-2 표준에서 매크로블록은 16x16 크기로 고정되어 있다. 하지만, MPEG-4에서는 매크로블록을 분할하여 각 서브매크로블록(sub-macroblock)에 대하여 별도의 모션 벡터를 지정할 수 있다. 이러한 기법이 유용한 경우는 매크로블록의 일부는 배경에 해당하고 나머지는 움직이는 물체에 해당되는 경우이다. 이때, 두 부분을 분할하여 각각 같은 크기의 가장 유사한 블록을 찾는 것이 유리할 것이다. MPEG-4에서 가능한 서브매크로블록의 크기는 16x8, 8x16, 8x8이며, 서브매크로블록은 다시 8x4, 4x8, 4x4 크기로 분할할 수 있다. 또한, MPEG-2에서는 각 매크로블록에 대해 모션 벡터를 구하기 위하여 반 픽셀 단위 위치의 블록과 비교하였으나, MPEG-4에서는 1/4 픽셀 단위로 비교한다. MPEG-1과 MPEG-2에서는 프레임 단위로 P-프레임과 B-프레임을 정의하였으나, MPEG-4에서는 슬라이스(slice) 단위로 P-슬라이스(Predictive-slice), B-슬라이스(Bipredictive-slice)를 정의한다. 여기에서, 슬라이스는 하나의 프레임 내에 연속되는 매크로블록의 집합으로, 임의 개수의 매크로블록을 포함할 수 있고, 여러 처리 및 코딩 파라미터 설정을 위한 단위이다. 하나의 P-슬라이스 또는 B-슬라이스 내의 매크로블록은 최대 15개까지의 참조 프레임을 참조할 수 있다. 이러한 기법은 주기적인 화면으로 구성된 동영상을 코딩할 때에 상당한 압축 효율을 거둘 수 있다. MPEG-1과 MPEG-2에서 B-프레임은 참조 프레임으로 사용되지 않았으나, MPEG-4에서는 슬라이스 단위로 다른 블록을 참조하므로 이러한 제약이 없어졌다. 예를 들어, 하나의 매크로블록이 B-슬라이스를 포함하는 프레임 내의 블록을 참조할 수도 있다.

MPEG-1과 MPEG-2에서는 매크로블록을 코딩할 때에 8x8 크기의 블록으로 분할하여 각각에 대해 DCT 변환 및 양자화 코딩을 수행한다. MPEG-4에서는 앞에서 설명한 바와 같이 매크로블록이 최대 4x4 크기로 분할될 수 있으므로, DCT 및 양자화 코딩 단위도 4x4 크기의 블록이 된다. MPEG-4에서는 4x4 블록에 대하여 DCT 및 양자화 코딩을 수행하는 대신 거의 같은 효과를 얻을 수 있는 정수 변환(integer transform)을 수행한다. DCT 변환의 단점은 높은 정밀도의 실수에 대하여 수행하여야 하므로 계산 비용이 크다는 점이다. 정수 변환은 이러한 단점을 없앴과 동시에 DCT 및 양자화를 한 단계로 합쳤다는 점에서 더욱 계산 비용을 줄일 수 있다. DCT 변환의 또다른 단점은 양자화를 통하여 압축률을 높였을 때, 복원된 프레임 내에 인접한 8x8 또는 4x4 블록들 간에 색상의 불연속이 생긴다는 점이다. 이러한 현상을 블로킹(blocking)이라 부른다. MPEG-4에서는 블로킹 현상을 줄이기 위하여 디블로킹 필터(deblocking filter)를 도입하였다. 디블로킹 필터는 인접한 4x4 블록 간의 가로, 세로 경계의 색상을 검사하여, 화면이 부드럽게 출력되도록 주변의 픽셀 색상을 참조하여 각 경계에 대해 최대 네 개의 픽셀 값을 변경할 수 있다. 이 밖에 MPEG-4 표준에서는 무손실 코딩(lossless coding)을 위하여 CAVLC(Content-based Adaptive Variable Length Coding) 및 CABAC(Context-based Adaptive Binary Arithmetic Coding) 기법을 채용하였다.

MPEG-2에서와 같이 MPEG-4에서도 프로파일과 레벨을 정의하고 있다. <표 5>는 Baseline, Main, Extended 세 가지 프로파일 내에 포함된 기법들을 나타낸 것이고, 표 6은 다섯 가지 레벨에 대하여 설정된 파라미터를 정리한 것이다.

7) MPEG-4 표준은 MPEG-2 표준과 호환성을 가지지 않는다. 따라서, MPEG-4 동영상은 MPEG-2 디코더에 의하여 재생되지 않는다.

〈표 5〉 MPEG-4 프로파일에서 포함하는 압축 기법

Tools	Baseline	Main	Extended
I- and P-slices	✓	✓	✓
B-slices	✗	✓	✓
Multiple Reference Frames	✓	✓	✓
Flexible Macroblock Ordering (FMO)	✓	✗	✓
Arbitrary Slice Ordering (ASO)	✓	✗	✓
Deblocking Filter	✓	✓	✓
CAVLC Entropy Coding	✓	✓	✓
CABAC Entropy Coding	✗	✓	✗
4:2:0 Chroma Format	✓	✓	✓
Interlaced Coding (PicAFF, MBAFF)	✗	✓	✓

〈표 6〉 MPEG-4 레벨 파라미터

Level Number	Max Macroblocks Per Second	Max Macroblocks Per Frame	Max Bit Rate (Mbps)
1	1485	99	0.064
2	11880	396	2
3	40500	1620	10
4	245760	8192	20
5	589824	22080	135

### 5. VC-1

VC-1은 본래 Microsoft의 Windows Media Player에서 재생하는 스트리밍 동영상을 위한 Windows Media 9이라는 이름의 코덱의 일부였다. Microsoft는 2003년 이 코덱을 SMPTE (Society of Motion Picture and Television Engineers)에 제출하였고, SMPTE는 2006년 초에 Windows Media 9의 주요 내용인 WMV (Windows Media Video) 9을 기반으로 421M이라는 동영상 코덱을 발표하였다. SMPTE 421M은 VC-1의 또다른 이름이다. 이러한 과정에서 VC-1 표준이 Microsoft 소유로 알려지기도 했으나, 현재 15개 정도의 회사가 VC-1에 대한 특허권을 공유하고 있다. VC-1 표준의 기술적인 사양을 자세히 살펴보면 놀라울 정도로 MPEG-4 AVC/H.264 표준과 많은 기법을 공유하고 있으

며, 화질 및 압축률 등의 성능도 MPEG-4 AVC/H.264 표준과 비교될 정도이다. 실제로 MPEG-4 AVC/H.264 표준 작업 팀에 Microsoft 소속 연구원들이 활발히 참여하였다. MPEG-4 AVC/H.264 표준이 최대의 성능을 거두기 위하여 모든 가능한 (복잡한) 기법을 포함한 것에 비하여, VC-1 표준은 커다란 성능 상의 저하가 없는 한도 내에서 덜 복잡한 기법을 채택한 것으로 보인다. Microsoft는 VC-1이 MPEG-4 AVC/H.264에 비하여 같은 압축률로 거의 동일한 화질을 얻을 수 있는 반면, 디코더의 복잡도 면에서는 2배 정도 효율적이라고 주장하고 있다. VC-1은 MPEG-2, MPEG-4 AVC/H.264와 함께 차세대 동영상 미디어인 Blu-ray Disk와 HD DVD에서 지원하는 동영상 압축 표준으로 선정되었다. 미디어 제작자는 세 가지 표준 중에서 하나를 선택하여 타이틀을 제작할 수 있지만, 타이틀 플레이어는 세 가지 표준을 모두 지원하여야 한다.

VC-1과 MPEG-4 AVC/H.264 압축 기법에 있어서의 중요한 차이점은 다음과 같다. MPEG-4에서는 DCT 및 양자화 코딩의 단위로 4x4 크기의 블록을 사용한다. VC-1에서는 4x4 크기 외에도 8x8, 8x4, 4x8 크기의 블록을 사용한다. VC-1에서도 정수 변환을 수행하지만, 세부적인 방법은 다르다. MPEG-4에서는 모션 추정을 위하여 최대 4x4 크기로 매크로블록을 분할할 수 있지만, VC-1에서는 8x8 크기까지 분할할 수 있다. 비교되는 블록 위치의 단위는 MPEG-4에서와 같이 1/4 픽셀이다. 하지만, 1/4 픽셀 위치의 색상 값을 구하기 위한 방법은 약간 다르다. MPEG-4에서는 P- 및 B-프레임 내에서 참조하는 프레임의 개수가 최대 15개까지 가능하나, VC-1에서는 이러한 복수 개의 참조 프레임을 허용하지 않는다. MPEG-4에서는 복수 개의 참조 프레임에 대한 모션 추정을 수행하기 위하여 인코더의 프로세싱 비용이 크게 올라갔고, 디코더에서도 모든 참조 프레임을 관리하기 위한 저장 비용이 많이



증가하였다. 무손실 엔트로피 코딩을 위하여 MPEG-4에서는 CAVLC와 CABAC 두 가지 기법을 포함하고 있으나, VC-1에서는 CAVLC와 유사한 적응 가변길이 코딩 (adaptive variable-length coding) 기법만을 포함하고 있다. 실제로, CABAC는 매우 복잡한 기법으로 많은 프로세싱 비용을 요구하여, 인코더뿐만 아니라 디코더의 성능을 떨어뜨리는 주요 요인 중의 하나로 꼽히고 있다. VC-1에서도 MPEG-4와 유사한 디블로킹 필터를 도입하였고, Microsoft의 주장으로는 VC-1의 디블로킹 필터가 MPEG-4에 비하여 단순하면서도 텍스처(texture) 등을 잘 나타낸다고 한다.

VC-1에서도 MPEG-2/4 표준에서와 같이 프로파일 및 레벨을 정의하고 있다. <표 7>은 VC-1에서 정의한 프로파일과 레벨 조합의 일부를 보이고 있다.

<표 7> VC-1 프로파일과 레벨 조합의 일부

프로파일	레벨	최대 비트율 (Mbps)	해상도 @ 프레임율 (Hz)
Simple	Low	0.096	176x144 @ 15
	Medium	0.384	352x288 @ 15
Main	Low	2	320x240 @ 24
	Medium	10	720x480 @ 30
	High	20	1920 x 1080 @ 30
Advanced	L0	2	352x288 @ 30
	L1	10	720x480 @ 30
	L2	20	1280x720 @ 30
	L3	45	1920x1080 @ 30
	L4	135	1920x1080 @ 60

## 6. 결론

앞에서 설명한 모든 동영상 압축 표준은 모두 비대칭(asymmetric) 시스템으로 설계되었다. 즉, 각 표준에서는 동영상 압축을 해제하고 화면에 출력하기 위한 디코더에 대해서만 정의되어 있으며, 표준에 따르도록 동영상을 압축하기 위한

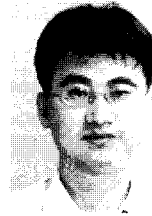
인코더는 정의되어 있지 않다. 압축된 결과가 동영상 압축 표준에 정의된 디코더에 의하여 재생이 가능하다면, 어떠한 알고리즘의 인코더를 사용하더라도 무방하다. 인코더의 복잡도는 디코더에 비하여 훨씬 높으며 인코딩 방식에 따라 같은 동영상을 같은 화질로 압축하더라도 압축 비율이 달라질 수 있다. 실제로, MPEG-2 표준이 제정된 초기에 비하여 최근의 MPEG-2 인코더는 성능이 두 배 이상 향상되었다. 이러한 비대칭 시스템은 IPTV나 DVD 플레이어와 같이 하나의 복잡한 인코더가 생성한 압축 동영상을 매우 많은 수의 비교적 간단한 디코더가 재생하는 환경에 적합하다.

## 참고문헌

- [1] ISO/IEC, Information Technology ? Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1,5 Mbit/s ? Part 2: Video, International Standard, ISO/IEC 11172-2, 1993.
- [2] ISO/IEC, Information Technology ? Generic Coding of Moving Pictures and Associated Audio Information: Systems, International Standard, ISO/IEC 13818-1, 2000.
- [3] ISO/IEC, Information Technology ? Generic Coding of Moving Pictures and Associated Audio Information: Video, International Standard, ISO/IEC 13818-2, 2000.
- [4] ITU-T, Advanced Video Coding for Generic Audiovisual Services, Recommendation H.264, Mar. 2005.
- [5] D. Kumar, P. Shastry, and A. Basu, "Overview of the H.264/AVC," Texas Instruments Developer Conference,

- Bangalore, India, Presentation, Nov. 2005.
- [6] C. Poynton, Digital Video and HDTV Algorithms and Interfaces, Morgan Kaufmann, Jan. 2003.
- [7] W. Simpson, Video Over IP: A Practical Guide to Technology and Applications, Focal Press, Sept. 2005.
- [8] G. J. Sullivan, P. Topiwala, and A. Luthra, "The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions," In Prof. SPIE Conf. on Applications of Digital Image Processing, Aug. 2004.
- [9] P. Symes, Digital Video Compression, McGraw-Hill, Oct. 2003.
- [10] J. W. Weber and T. Newberry, IPTV Crash Course, McGraw-Hill, Nov. 2006.
- [11] 오쿠보 사카에, 카도노 신야, 키쿠치 요시히로, 스즈키 테루히코, 개정판 H.264/AVC 교과서, 인프레스, 2005년 11월.
- [12] 카메라야마 와타루, 하나무라 츠요시, 디지털 방송 교과서 (상) MPEG-1/2/4/H.264/AVC, 인프레스, 2004년 9월.

## 저자약력



**문양세**

1991년 2월 한국과학기술원 전산학과 학사  
 1993년 2월 한국과학기술원 전산학과 석사  
 2001년 8월 한국과학기술원 전자전산학과 전산학전공 박사  
 1993년 2월~1997년 2월 현대전자산업(주) 주임연구원  
 2001년 9월~2002년 2월 (주)현대시스콤 선임연구원  
 2002년 2월~2005년 2월 (주)인프라벨리 기술위원(이사)  
 2005년 3월~현재 강원대학교 컴퓨터학과 조교수  
 2005년 3월~현재 한국과학기술원 AITrc 연구원  
 관심분야 : 데이터 마이닝, 스트림 데이터, 멀티미디어 시스템, 데이터베이스 및 응용, 이동/무선 통신 서비스 및 시스템  
 이 메 일 : ysmoon@kangwon.ac.kr



**노웅기**

1991년 2월 한국과학기술원 전산학과 학사  
 1993년 2월 한국과학기술원 전산학과 석사(멀티미디어 전공)  
 2001년 2월 한국과학기술원 전산학과 박사(데이터 마이닝 전공)  
 2001년 2월~2003년 9월 (주)티맥스소프트 책임연구원 (미들웨어 개발)  
 2003년 10월~2005년 3월 (주)티맥스테이터 수석연구원 (DBMS 개발)  
 2005년 4월~2006년 5월 한국과학기술원 전산학과 초빙교수  
 2006년 6월~현재 미국 University of Minnesota 방문연구원  
 멀티미디어 시스템, 멀티미디어 코덱, 멀티미디어 내용기반 검색, 데이터 마이닝/데이터 웨어하우징, 정보 검색  
 이 메 일 : lohwh@cs.umn.edu, wklh2@gmail.com