

디지털 방송에 있어서의 비디오 압축 기술

정 제 창*, 오 덕길**, 박 재홍**

* 한양대학교 전자전기공학부

** 한국전자통신연구원 방송기술연구부

1. 서론

21세기 멀티미디어 시대를 눈앞에 두고 영상과 음향을 CD, DAT, HDD, Optical Disk 등에 저장하거나 전화선, ISDN, BISDN, 지상파/케이블/위성 등의 채널을 통하여 전송할 필요성이 증가되고 있다. 이에 따라 ISO와 IEC가 협력하여 디지털 동영상 및 음향 압축표준을 제정하는 노력을 계속하여 왔다. 우선 CD와 같은 디지털 매체를 대상으로 1.5Mbps의 동영상 압축표준을 제정하였는데 이것이 바로 MPEG-1(ISO/IEC 11172)이다. MPEG-1의 대상은 비디오 CD 등으로 제한되어 있는데, 화질이 VHS 비디오 테이프 수준에 머무르고 한 장에 74분밖에 기록되지 않는 단점이 있다. 이에 따라 MPEG-1의 주요 응용분야는 비디오 CD, 컴퓨터상의 스트리밍 비디오, 게임 등에 국한되었다.

디지털 방송등 고화질의 응용을 위한 표준으로서 등장한 것이 바로 MPEG-1의 후속인 MPEG-2이다. MPEG-2의 목표는 5 Mbps정도에서 현행 TV 품질을 실현하는 것인데 HDTV(High Definition Television, 고선명 TV) 까지도 수용하고 있다. 이 MPEG-2를 기반으로 하여 1994년 미국 휴즈사에 의해 위성방송이 가장 먼저 디지털화되었고 작년에 미국과 유럽의 일부국가에서 지상파 방송의 디지털화가 시작되었다. 미국은 1998년 - 2006년에 걸쳐 현재의 아날로그 디지털 방송을 디지털로 전환할 계획이며 우리나라도 2001-2010년에 걸쳐 디지털 방송으로 전환할 계획이다.

MPEG-2는 뛰어난 성능과 유연성에 따라 디지털 위성방송(우리나라의 무궁화위성을 통한 DBS도 포함), 고선명 TV, 디지털 비디오디스크(DVD), 주문형 비디오(VOD) 등 많은 분야에서 채용이 결정되어 멀티미디어 혁명을 주도하는 원동력이 되고 있다. 특히 DVD는 한장에 MPEG-2 화질(가변 비트율로서 평균 6 Mbps)의 135분 길이 영화를 담을 수 있어 현재의 비디오 CD나 VHS, 디지털 VCR 등에 큰 영향을 주고, 나아가서 약 4.7GB의 대용량을 실현함으로써 HDD, CD-ROM 등 컴퓨터 보조기억장치 분야에도 파급효과가 매우 클 것으로 보인다.

이러한 배경을 바탕으로 본 논문에서는 MPEG-2 비디오 압축기술에 관하여 간략히 기술하고자 한다.

2. MPEG-2 비디오 개요

MPEG-1 비디오는 CD와 같은 디지털매체를 대상으로 하는 1.5Mbps의 동영상 압축표준이다. MPEG-1은 주로 제한된 파라미터를 갖는 비트열(Constrained Parameter Bitstream)

에서 사용되고 있으며 이는 수평방향 화소수가 768 이하, 수직방향 화소수가 576 이하, 전체 화면이 396 Macroblock 이하, 프레임률이 30 Hz 이하, 복호기버퍼 사이즈가 327,680 bit 이하, 전송률이 1,856,000 bps 이하의 영역이다. MPEG-2 비디오는 선택스상 MPEG-1 비디오를 포함하고 있어 순방향 호환성이 유지되나 역방향 호환성은 유지되지 않는다. 즉, MPEG-1 비디오 비트열은 MPEG-2 비디오 복호기에 의해 복호되지만 MPEG-2 비디오 비트열은 MPEG-1 비디오 복호기에 의해 복호되지 않는다. MPEG-2 오디오가 MPEG-1 오디오와 순방향 역방향 모두 호환성이 유지되는 것과 대비된다.

MPEG-2 비디오는 MPEG-1보다 고비트율의 방송, 통신, 저장미디어에서 고화질의 동영상 전송하거나 저장하는 데 사용된다. 응용분야가 다양해진 만큼 충족시켜야 할 요구조건도 많아졌다. MPEG-2 비디오는 현행 TV나 HDTV를 효율적으로 압축하는 것이 주목적이다. 현행 TV의 화질은 3-9 Mbps에서, 그리고 HDTV 화질은 17~30 Mbps에서 각각 얻어진다. 비트율은 주어진 채널의 용량과 요구화질을 고려하여 선택된다. 예를 들어 무궁화위성 DBS에 있어서는 약 5 Mbps, 미국의 Grand Alliance HDTV 방식에서는 17 Mbps, 전화선을 이용한(ADSL-3 방식) 주문형 비디오나 DVD에 있어서는 4~6 Mbps를 비디오에 할당하고 있다.

MPEG-2 비디오는 일종의 범용(generic)압축 알고리즘으로서, MPEG-1 비디오를 크게 확장 발전시키면서 많은 도구들을 마련하여 응용분야에 따라 이들을 적절히 선택하여 사용하도록 하고 있다. 압축효율의 향상을 위해 MPEG-2 비디오는 MPEG-1 비디오의 각 요소들을 재검토하여 조금씩 개선하거나 새로운 요소를 추가함으로써 전체적으로 상당한 향상을 가져오고 있다. 즉 필드단위의 처리, 움직임추정/보상방식, 양자화, DCT 계수의 주사방식, 가변장 부호화 등 많은 부분들이 개선되었다.

MPEG-2 비디오의 범위는 매우 넓지만 응용분야마다 그 중 특정해상도에 특정기능까지만을 사용한다. 따라서 부호기와 복호기를 제작할 때의 편의를 위해 MPEG-2 비디오를 해상도와 기능에 따라 몇 가지로 분류하고 있다. 우선 화면의 해상도는 4개의 레벨로 분류된다. MPEG-1 비디오가 대상으로 하는 것과 같은 작은 화면인 Simple, 현행 TV의 화면 크기인 Main, 유럽 고선명 TV(HDTV)의 화면크기인 High-1440, 미국 고선명 TV를 위한 규격인 High가 그것이다.

또 기능에 따라서는 5개의 프로파일로 나누어진다. 양방향 예측을 이용하는 B 프레임을 제외하여 구현을 용이하게 한 Simple, 많은 기능을 포함하여 대부분의 응용분야에 채택되고 있는 Main, 계층구조를 가지면서 보다 기능이 확장된 SNR Scalable, Spatial Scalable, High 등이 그것이다.

MPEG-2 비디오에서는, 컴퓨터에서 채택되고 있는 순행주사만을 대상으로 하는 MPEG-1 비디오와는 달리, TV에서 사용되고 있는 격행주사 방식의 동화에 대해서도 많은 고려를 하고 있다. 즉, 한 프레임을 두 필드로 나누어 필드구조로 부호화할 수도 있고, 프레임구조로 부호화할 수도 있다. 움직임이 많은 장면은 한 프레임의 두 필드 간에도 큰 차이가 나므로 필드 구조로 부호화하는 것이 효과적이고, 정지화에 가까울수록 두 필드간에 상관도가 높아 프레임구조로 부호화하는 것이 유리하다. 또 프레임구조의 부호화에 있어서도 각 매크로블록(16×16 화소단위)별로 필드단위의 처리가 가능하도록 하여 화면내의 부분적 움직임을 용이하게 처리할 수 있게 하고 있다.

MPEG-2에 새로이 채택된 움직임 추정·보상 방법으로서 듀얼프라임(Dual Prime)이 있다. 이것은 필드단위의 움직임보상을 하되 이에 따라 많아지는 움직임벡터의 양을 효과적으로

줄이는 방식이다. 이 방식은, B 프레임을 생략하여 부호화에 따른 지연시간과 복잡도를 줄이고자 할 때의 보완수단으로서, 여전히 좋은 화질을 유지하는 데 기여한다.

DCT에 있어서는 프레임구조에서도 매크로블록 단위로 프레임모드와 필드모드 중 데이터 발생량이 적은 것을 선택할 수 있다. 따라서 움직임이 많은 경우와 적은 경우 모두 효과적으로 처리할 수 있다. 단, 필드구조에서는 필드 DCT 하나만이 쓰인다.

DCT 계수의 양자화에 있어서 MPEG-1에서는 계수의 크기에 관계없이 양자화시스템이 일정한 선형 양자화가 사용되고 있다. 반면 MPEG-2에서는 계수값이 작을수록 양자화시스템이 작아 세밀하게 양자화하는 비선형양자화가 함께 사용된다. 비선형방식은 복잡도는 증가하지만 평균 양자화잡음을 줄여 양자화기의 성능 향상을 가져온다.

양자화 된 DCT 계수의 가변장 부호화를 위한 주사에 있어서, MPEG-1에서는 지그재그 주사만을 사용하였다. 그러나 MPEG-2에서는 대체 주사(Alternate Scan)방법이 추가로 사용되는데, 이 두가지 중 하나를 화면단위로 선택하여 사용한다. 대체 주사는 DCT의 수직방향 고주파성분을 상대적으로 일찍 주사하여 특히 움직임이 큰 격행주사 영상에 뛰어난 효과를 보인다.

주사된 DCT 계수의 가변장 부호화(VLC: Variable Length Coding)를 위해 MPEG-1에서는 하나의 2차원 VLC 테이블만이 사용되었다. 이에 비해 MPEG-2에서는 화면내 부호화를 위한 테이블을 추가로 사용할 수 있도록 하였다. 화면내 부호화는 화면간 부호화에 비해 DCT 계수값들이 훨씬 커 통계적 특성이 다르므로 이를 반영한 별도의 테이블이 필요했던 것이다. 이 VLC를 사용함으로써 화면내 부호화시의 데이터발생량을 크게 줄일 수 있다.

스케일러빌리티 기능은 MPEG-2에 새로이 도입된 개념으로서, 공간 스케일러빌리티(Spatial Scalability), 시간 스케일러빌리티(Temporal Scalability), SNR 스케일러빌리티(SNR Scalability) 등이 있다.

공간 스케일러빌리티는 우선 화면을 공간해상도가 낮은 기본 계층(예: 현행 TV)과 높은 고위 계층(예: 고선명 TV)으로 나눈다. 기본 계층을 먼저 부호화하고 이어서 기본 계층의 보간성분과 고위계층의 차이성분을 부호화하여, 두 부호화비트열을 함께 보낸다. 이렇게 하면 현행 TV 수신기로도 기본 계층비트열을 복호하여 고선명 TV를 현행 TV 화질로 볼 수 있고, 고선명 TV 수신기는 두 비트열을 모두 복호하여 고선명화면을 재생한다. 시간 스케일러빌리티와 SNR 스케일러빌리티도 공간 스케일러빌리티와 마찬가지로, 기본 계층과 고위 계층으로 나누어 기본 계층의 부호화비트열과, 기본 계층의 확장성분과 고위계층 간의 차이성분의 부호화비트열을 보낸다. 다만 기본 계층과 고위계층의 분류에 있어서 시간 스케일러빌리티는 시간축(화면의 진행방향)으로, SNR 스케일러빌리티는 화소마다의 비트표현상의 해상도에 따라 나누는 점이 다르다.

3. MPEG-2 비디오 압축 알고리즘

MPEG-2 비디오는 이산여현변환(DCT; Discrete Cosine Transform), 양자화(Quantization), 가변장부호화(VLC; Variable Length Coding), 움직임 추정/보상(Motion Estimation/Compensation)을 기반으로 한 알고리즘을 사용한다. MPEG-2 비디오의 신택스는 계층적 구조를 가지고 있는데, Sequence Layer 아래에 Random access를 위한 기본 단위인 GOP(Group of Picture) Layer, GOP Layer 아래에 Picture Layer, Picture Layer 아래에 오류가 발생했을 때 오류의 영향을 국한시키는 Slice Layer, Slice Layer 아래에 움직임 추정/

보상의 단위인 Macroblock Layer, Macroblock Layer 아래에 DCT의 단위인 Block Layer가 있다. 따라서 비트열 중간에서 복호화할 경우에는 GOP 단위로 Random Access할 수 있고, 비트열에 오류가 발생할 때에는 Slice 단위로 재동기 및 복구가 수행되므로 다음 Slice부터는 오류의 영향이 미치지 않게 된다.

MPEG-2 비디오는 MPEG-1 비디오를 기반으로 하여 매우 광범위한 응용영역을 지원할 수 있도록 고안되었다. 많은 기능과 압축을 위한 요소들이 새로이 추가되고 기존의 압축요소들이 재검토를 거쳐 상당히 개선되었다. 이 점에서는 최근의 H.263(전화선을 이용한 영상전화를 위한 동영상압축 국제표준)이 기존의 H.261(ISDN을 이용한 영상전화/회의를 위한 동영상압축 국제표준)을 개선하고 있는 것과 비슷하다. 그러나 H.261로부터 H.263으로 진화한 것에 비해 MPEG-1 비디오로부터 MPEG-2 비디오로의 진화는 그 폭과 성능의 향상면에서 훨씬 크다고 할 수 있다. MPEG-1 비디오와 MPEG-2 비디오의 차이가 표 1에 요약되어 있다.

MPEG-2 비디오는 매우 다양한 응용분야를 겨냥하고 있기 때문에, 기능과 해상도에 따라 나누어진 여러 영역 중 어느 영역을 사용할 것인지를 결정하여야 한다. MPEG-2 비디오는 수평방향으로 Profile(Simple, Main, SNR, Spatial, High), 수직방향으로 Level(Low, Main, High-1440, High)을 양측으로 하여 20개의 영역으로 구분하고 있고 이 중 11개 부분이 응용영역으로 권장되고 있다. Profile이란 MPEG-2의 선택스 중 어느 부분을 사용하여 어떤 기능을 포함시키느냐에 관한 것으로 알고리즘의 복잡도를 나타내는 척도이고, Level이란 사용하는 파라미터를 어느 정도의 크기로 제한하느냐에 의해 결정되는 것으로 해상도를 나타내는 척도이다. 예를 들어 Main Profile과 Simple Profile과의 차이는 B(Bidirectional Prediction) 화면과 관계되는 선택스 사용유무에 달려있으며, 방송환경에 많이 사용되는 Main Level과 High Level의 차이는, Main Level에서는 수평화소 720 이하, 수직화소 576 이하, 프레임률 30 Hz 이하 등으로 제한되어 있는 반면, High Level은 수평화소 1920 이하, 수직화소 1152 이하, 프레임률 60 Hz 이하로 되어 있어 전체적으로 최대 10배 안팎의 많은 데이터처리가 요구되나 선택스상으로는 동일하다. 이 중 가장 많이 사용되는 영역으로는 Main Profile@Main Level의 영역으로 기존 NTSC, PAL, SECAM등의 방송을 대체할 수 있어 현재 미국 유럽 등에서 위성방송, 케이블 TV 방송 등에 적용되고 있거나 적용이 검토되고 있고 한국의 디지털 위성방송에도 적용되고 있다. 또한 Main Profile@High Level은 미국을 비롯한 여러 나라의 디지털 HDTV의 표준규격으로 채택되고 있다. 표 2에 MPEG-2 MP@ML의 규격이 요약되어 있다.

1) 부호기/복호기 개요

그림 1에 MPEG-2 비디오 부호기의 전형적인 블록 다이어그램이 각각 나타나 있다. 영상 신호의 압축은 영상신호에 내재하는 다음과 같은 각종 중복성을 제거함으로써 얻어진다.

(1) 색신호간 중복성(Spectral Redundancy) 제거

카메라 등을 통해 들어오는 RGB 영상입력 신호는 색성분간의 신호상관도가 높고 각각 넓은 대역을 차지한다. 이 상관도를 줄여 전체 데이터발생량을 줄이고 인간의 시각특성에 맞도록 하기 위해 YCrCb의 색체계로 변환한다. Y(휘도) 성분은 주파수대역폭도 넓고 인간의 눈이 여기에 민감한 반면, CrCb (색) 성분은 대역폭도 좁고 인간의 눈이 상대적으로 덜 민

감하므로, 수평과 수직방향으로 각각 간축(Decimation)하여 사용할 수 있다.

(2) 공간적 중복성(Spatial Redundancy) 제거

화면내에 인접하는 화소간에는 상관도가 높으므로 Block(8×8 화소) 단위로 변환부호화의 일종인 DCT 및 양자화과정을 통하여 공간적인 중복성을 제거한다. 공간적으로 분포되어 있는 신호의 에너지는 이 변환을 거치면 DC를 비롯한 저주파항에 에너지가 집중되고, 고주파항에는 매우 작은 값만이 존재하여 양자화 후에 0이 되므로 전송이 불필요하여 데이터 압축이 이루어진다. 에너지 집중효율이 높을수록 데이터압축률도 높아지는데 자연화에 대해서는 DCT가 효율적이지만 텍스트, 만화, 컴퓨터 그래픽스에 의한 그림, 애니메이션 등에 대해서는 비효율적이다.

(3) 시간적 중복성(Temporal Redundancy) 제거

시간적으로 인접한 두 화면 간에도 상관도가 높으므로 Macroblock(16×16 화소) 단위로 두 화면 간의 움직임 추정하여 보상함으로써 시간적인 중복성을 제거한다. 움직임추정은 보통 휘도 Macroblock(4개의 휘도 Block)만을 이용하고 움직임보상은 4개의 휘도 Block과 2개(4:2:0의 경우)의 색차 Block 모두에 대해 수행한다. 각 Macroblock은 움직임보상 후 부호화하거나(Inter Mode, 화면간 부호화) 움직임 보상없이 현재 화면을 직접 부호화(Intra Mode, 화면내 부호화)한다. 예를 들어 장면전환이 있는 경우나 움직임이 커서 움직임추정을 위한 탐색범위를 넘어서는 경우에는 움직임보상은 오히려 데이터량이 증가하게 되므로 이럴 때는 화면내 부호화를 선택한다.

(4) 통계적 중복성(Statistical Redundancy) 제거

DCT와 양자화과정을 거친 계수값들은 통계적으로 발생빈도가 크게 다르다. 이 통계적 특성을 이용하면 무손실로 데이터를 더욱 압축할 수 있다. MPEG-2 비디오에서는 가변장부호화(VLC)의 일종인 허프만 부호화를 이용하여 통계적 중복성을 제거한다.

부호기의 채널버퍼는, VLC 출력이 가변(variable)비트율을 갖기 때문에 이를 채널을 통하여 고정(constant)비트율로 전송할 수 있도록 데이터를 일시 저장하기 위한 용도로 사용된다. 복호기측에서는 채널을 통해서 전송된 MPEG-2 비트열이 복호기버퍼에 입력되면 헤더 정보 등을 제외한 비디오데이터들이 VLD(Variable length decoding), 역양자화, 역 DCT 등을 거쳐 Block 단위로 화면이 복원된다.

2) Picture 구성

MPEG-2 비디오에는 움직임보상의 방법에 따라 I, P, B 3종류의 Picture가 있고, random access가 가능한 I picture 주기에 따라 I, P, B를 묶어 하나의 GOP(Group of Picture)를 구성한다.

(1) I(Intra) Picture

화면내의 모든 Macroblock이 Intra Mode로만 구성되기 때문에 시간적 중복성은 여전히 존재한다. 채널 전환시의 화면 복구나 오류의 전파를 막기 위해 일정 간격으로 I Picture를 두어야 한다. I Picture는 전체 화질에 절대적인 영향을 미치므로 P, B Picture에 비해 고화질을 유지할 수 있도록 부호화하여야 한다.

(2) P(Predictive) Picture

현재 프레임에 대해서 이전 프레임의 I Picture 또는 P Picture를 기준으로 하여 순방향(forward) 움직임보상 예측기법을 적용하여 시간적 중복성(temporal redundancy)을 제거한다. GOP의 구조적 특징 때문에 P Picture는 연속되는 P Picture 및 B Picture에 영향을 미치기 때문에 I Picture보다는 다소 떨어지지만 B Picture보다는 나은 화질을 유지하여야 한다.

(3) B(Bidirectional) Picture

현재 프레임에 대해서 이전 프레임의 I/P Picture 그리고 다음 프레임의 I/P Picture로부터 각각 움직임보상된 순방향 예측화면, 역방향 예측화면, 순방향 및 역방향을 보간(Interpolation)한 화면을 사용하여 세 가지 예측신호를 얻어낸 후 이들 예측신호 중 최적의 것을 영상간의 예측신호로 사용하여 시간적 중복성을 제거한다. GOP의 구조상 B-Picture는 다른 Picture에 영향을 주지 않으므로 I, P Picture에 비해 가장 적은 비트를 할당하여 부호화한다. 평균적으로 I, P, B Picture 각각으로부터의 비트발생량은 15:5:1 정도의 비율이 된다.

3) 입출력영상의 규격

MPEG-2 비디오에서는 기본적으로 4:2:0, 4:2:2, 4:4:4 chroma format의 영상을 받아들일 수 있다. 표 3에서 4:4:4 포맷은 색성분에 대해 간축하지 않은 경우이고 4:2:2 포맷은 수직방향으로 2:1로 간축한 경우이며 4:2:0은 수평 수직 방향 모두 2:1로 간축한 경우이다. 색성분의 간축이 화질을 그다지 떨어뜨리지 않고 데이터압축 효과가 있으므로 Main Profile@Main Level을 포함한 거의 대부분의 응용에 있어서 4:2:0이 사용되고 있고, 스튜디오용의 전문가용에서는 4:2:2를 사용하며 4:4:4는 사실상 거의 사용되지 않는다. 다만 최근의 MPEG 회의에서 비트율에 여유가 있는 환경에서는 4:2:0보다 4:2:2를 사용하는 것이 화질에 약간의 향상을 가져온다는 것이 밝혀져 이를 위한 4:2:2 Profile이 표준에 추가되었다.

MPEG-2 비디오 선택스상에서는 sequence extension에 chroma_format을 나타내는 필드가, sequence display extension에는 video-format(NTSC, PAL, SECAM, MAC 등)의 종류를 나타내는 필드가 각각 포함되어 있다.

4) 부호화모드

(1) Picture 부호화모드

MPEG-2는 앞에서 설명한 바와 같이 GOP 구조단위로 부호화하며 한 개의 GOP마다 1개의 I Picture, N/M-1개의 P Picture, N-1-P개의 B Picture가 있다. 여기서 파라미터 M은 I Picture와 P Picture 사이의 B Picture의 개수+1을 말하는 변수이고, N은 한 GOP의 길이를 말하는 변수이다. 따라서 M=1 이라면 I, P Picture 사이에 B Picture가 없는 IPPP...IPPP 구조의 GOP 구조를 의미하는 것이며 N=9, M=3 이라면 IBBPBBPBBL...와 같은 GOP 구조를 갖게 된다. 여기서 M, N 등의 변수가 클수록 부호화/복호화 지연이 길어지며 Random Access 시간 또한 길어지는 등의 단점이 있기 때문에 M=1 등으로 설정하는 방법 또는 Intra Slice/Intra Column 부호화 방법 등 저지연(Low delay)부호화에 관해서도 연구되어 있

다.

5) 움직임추정 및 보상

MPEG-2에서는 움직임추정(ME : Motion Estimation) 및 보상(MC : Motion Compensation)을 위한 방법으로 Frame MC, Field MC, Dual Prime MC 등의 3가지가 있다. 또한 기본적으로 모든 움직임추정 및 보상은 반화소단위까지 하는 것으로 규정하고 있다.

(1) Frame ME/MC

Frame ME/MC는 MPEG-1에서부터 사용하여 온 움직임 추정방법으로 Top Field와 Bottom Field의 구분없이 Frame 구조로 움직임을 추정하고 보상하는 것이며, 현재 Frame의 부호화하고자 하는 Macroblock에 대해 기준 Frame의 탐색영역내에서 반화소정밀도까지 완전탐색(Full Search)을 수행하여 이 중 가장 작은 MAE(Mean Absolute Error)를 발생시키는 위치를 움직임벡터로 결정한다. 실질적으로 데이터가 화소단위로 주어져 있으므로 화소단위의 1차 완전탐색을 통해 화소단위 움직임벡터를 구하고 그 후 반화소단위의 보간 및 2차 탐색을 통해 반화소단위의 움직임벡터를 구한다. Frame ME의 경우에는 P Picture에서는 1개의 Macroblock당 1개의 움직임벡터(motion vector)를, B picture에서는 1개의 Macroblock당 1개 또는 2개의 움직임벡터를 전송하기 때문에 Field ME/MC에 비하여 움직임벡터 전송에 필요한 비트수가 적다.

(2) Field ME/MC

Frame Structured Picture에 있어서 각 필드별로 움직임의 추정 및 보상을 수행하는 방식으로서, 현재 프레임의 Top Field와 Bottom Field, 기준 프레임의 Top Field와 Bottom Field 사이에서 각각 16×8 sub Macroblock 단위로 Top to Bottom, Top to Top, Bottom to Top, Bottom to Bottom 등의 4가지 움직임벡터를 구한 뒤, 현재 프레임의 Top Field와 Bottom Field 각각에 대해 최소의 움직임보상 에러를 발생시키는 하나씩의 움직임벡터를 선택한다. 따라서 P picture에서는 1개의 Macroblock당 2개의 움직임벡터, B picture에서는 1개의 Macroblock당 2개 또는 4개의 움직임벡터를 전송한다.

MPEG-2에서는 모든 Macroblock에 대해서 Frame/Field Prediction 방법을 다 적용하여 본 뒤 그 중 보다 작은 예측오차를 갖는 Prediction Mode를 사용한다(부호기측). 복호기측에서는 부호기에서 사용한 Prediction 모드가 전송되므로 이에 따라 움직임보상을 수행하여 영상을 복원한다.

(3) Dual Prime ME/MC

이 방법은 Toshiba에서 제안한 움직임 추정/보상 방법으로서 Field ME/MC 방법이 비교적 Macroblock당 발생하는 움직임벡터를 전송하기 위한 비트수가 많은데 반하여, 1개의 Macroblock당 1개의 움직임벡터와 차분 움직임벡터(dmv)만을 전송하는 것으로 비교적 느린 움직임을 갖는 시퀀스에 효과적인 것으로 알려져 있다. 이 방법은 $M = 1$ 인 경우, 즉, IPPPPIP...와 같이 B Picture가 없는 경우에만 사용할 수 있도록 규정하고 있다. 즉, B Picture가 허용되는 경우에는 이를 이용하여 더 좋은 화질을 얻을 수 있으나, 그렇지 않은 경우에는 Dual Prime Prediction을 사용함으로써 가능한 적은 비트발생량으로 화질의 향상

을 가져올 수 있다.

Dual Prime Prediction 방법은 Field Prediction Mode에서 구한 Top to Bottom, Top to Top, Bottom to Top, Bottom to Bottom 4개의 움직임벡터 중 Top to Top과 Bottom to Bottom 움직임벡터는 그대로 Base Motion Vector로 사용하고, Top to Bottom과 Bottom to Top 움직임 벡터는 각각 Scaling($\times 2, \times 2/3$)과 Truncation을 하여 Base Motion Vector를 만든 뒤, 이렇게 만든 4개의 Base Motion Vector 각각에 대하여 수평방향과 수직방향으로 -1, 0, +1 씩의 미세조정(dmv)을 가하여 두개의 16×8 sub MB에 대해 움직임 보상에러가 최소가 되도록 하는 Base Motion Vector와 변위값(dmv)을 보내는 방식이다. Dual Prime ME은 부호기에서의 계산량이 상당히 많은 편으로 1개의 Base Motion Vector당 9개의 예측후보값을 계산해내야 하므로 총 36가지의 후보 중 1개의 Base Motion Vector와 dmV를 계산해야 한다. 한편 복호기측에서는 전송되어온 Base Motion Vector와 dmV값으로부터 2개의 Field Motion Vector값을 계산해내기만 하면 되므로 비교적 간단하게 구현이 가능하다.

6) DCT & Quantization

DCT 변환(Discrete Cosine Transform)은 영상신호 부호화에 매우 효과적인 것으로 알려져, H.261, JPEG, MPEG 등의 국제 표준에 널리 채택되어 왔다. DCT 변환은 영상신호의 공간적인 상관성이 대단히 크다는 사실에 바탕을 둔 것으로, MPEG의 경우 8×8 크기의 block 단위로 수행되는데 8×8 화소에 분산된 에너지를 DC를 포함한 낮은 주파수의 DCT 계수로 집중시킨다. DCT 변환은 해당 Block이 Intra MB인지 Inter MB인지에 따라 각각 영상신호 자체 또는 예측오차를 변환하게 되므로 Intra MB인 경우는 공간적 중복성만을 제거하는 것이 되지만 Inter MB인 경우라면 시간적 중복성이 제거된 영상신호에 또다시 공간적 중복성을 제거하는 것이 된다. 그런데 예측오차 신호는 공간적 중복성이 그다지 크지 않으므로 DCT의 에너지 집중효과도 Inter MB에서는 Intra MB에 비해 다소 떨어진다.

한편, 해당 Macroblock의 두 필드 간의 상관도(정지화일수록 두 필드 간의 상관도가 높고 움직임이 많을수록 상관도가 떨어짐)에 따라 적응적으로 Frame DCT Block 또는 Field DCT Block으로 구분하여 DCT를 수행한다.

양자화(Quantization)는 Block 단위로 얻어진 DCT의 변화계수를 한정된 비트길이를 표현하는 과정으로서, 복호기측에서의 역양자화는 Intra DC 계수와 그 외의 계수로 다음과 같이 나누어져 수행된다.

(1) Intra DC 계수의 역양자화

DCT 변환계수 중 화질에 가장 큰 영향을 미치는 DC 값은 그 정밀도에 따라 8~11비트를 할당하는데, 각각의 양자화 스텝사이즈에 해당하는 8, 4, 2, 1등을 곱하여 DC 계수를 복원한다.

(2) 그 외 계수(Intra AC, Inter DC, Inter AC)의 역양자화

Block내 모든 64개의 DCT 계수에 2를 곱한 뒤 Inter block의 경우에만 해당 계수의 Sign 값(음수일 때 -1, 양수일 때 1)을 더하고, 여기에 Inter/Intra에 따라 그림 4.4와 같이 규정된 가중값 행렬을 곱한 뒤 균일/비균일 양자화기를 구분하는 q_scale_type flag에 따라 선택된 양자화기의 스케일값(Quantizer_scale)을 곱하면 역양자화된 DCT 계수가 얻어진다. 이때

양자화 매트릭스를 사용하는 것은 인간의 눈이 고주파성분의 양자화잡음을 덜 느끼는 점을 이용하여 고주파계수일수록 양자화스텝을 크게 하기 위함이다. 따라서 고주파 DCT 계수들은 크기도 작는데 양자화스텝도 커서 대부분 0이 된다.

7) VLC(Variable length Coding)

VLC(가변장부호화)는 발생확률이 높은 부호어들에 대해서는 부호당 짧은 비트를 할당하고, 발생확률이 낮은 부호들에 대해서는 부호당 긴 비트를 할당하여 부호의 평균길이를 Entropy에 가깝게 하는 수단으로서 Huffman Coding, Arithmetic Coding, Lempel-Ziv 알고리즘 등의 방법이 있다. 영상부호화에 있어서는 이 중 Huffman부호화를 사용하는데, 양자화된 DCT 계수, 움직임벡터의 차신호, 그리고 MB에 관련된 각종 정보가 그 대상이다.

(1) DCT 계수의 VLC

양자화 된 DCT 계수들을 더욱 압축하기 위해 런길이부호화와 허프만부호화를 쓰고 있다. DCT 변환과 양자화과정을 거친 영상신호는 "0"인 계수값이 많으므로 보다 효율적인 부호화를 위해 그림 4.1(f)와 같이 DC로부터 출발하여 지그재그 주사를 하면서 0이 몇 개 반복되고 0이 아닌 값이 나오는지를 (런, 레벨)의 형태로 나타낸다. 8×8 Block의 계수 중 마지막 non-zero 계수까지만 부호화한 뒤 EOB(End of Block) 부호를 사용하여 한 Block의 끝을 나타낸다. 이(런, 레벨) 심벌들은 발생확률이 각각 다르므로 2차원 허프만부호를 써서 더욱 압축하고 있다. 이때 각 블록의 평균값에 해당하는 DC는 화질에 크게 영향을 주므로 보다 충실히 표현할 필요가 있다. 따라서 DC는 AC 계수들에 비해 보다 세밀히 양자화하고, 양자화된 결과에 대해 이전 블록과의 차이를 취해 1차원 허프만부호화한다.

(2) 움직임벡터의 VLC

현 MB와 같은 타입의 바로 전 MB의 움직임벡터와 현재 움직임벡터 간에 DPCM을 수행한 뒤 이 값을 허프만부호화한다. 수직/수평 성분별로 독립적으로 부호화하며 P Picture 경우는 순방향 움직임벡터가 전송되는 반면 B Picture인 경우는 순방향, 역방향 움직임벡터 중 실제 움직임보상에 사용하는 움직임벡터만을 부호화한다.

(3) MB 정보에 대한 VLC

한 슬라이스에서 Macroblock 위치정보(Macroblock Address: MBA), MB의 부호화모드(MB Type), 그리고 MB내에서의 Block들의 부호화패턴(CBP : Coded Block Pattern) 정보에 대해서 허프만부호화를 수행한다.

8) Video Buffering Verifier(VBV)

VBV는 CBR(Constant Bit Rate), 즉 고정된 전송률의 비트열에서 만족시켜야 하는 요구조건을 규정하기 위한 것으로, 부호기측 출력에 연결된 가상적인 복호기이다. 가변 비트율(VBR: Variable Bit Rate)의 경우에는 MPEG-2: Part 1 시스템 부분에 정의되어 있는 STD(System Target Decoder)에 따른다. 가상적인 복호기에서는 일정한 전송률로 비트열이 들어오게 되며, 이 데이터가 복호기 버퍼에 쌓이게 된다. 복호기에 입력된 데이터는 하나의 Picture 단위로 정해진 복호시간에 순간적으로 빠져나가게 된다. 이러한 가정을 전제로 한 VBV의 버퍼상태는 그림 4.5와 같다. 그림 4.5에서, VBV 버퍼에 채워지는 데이터량은 시간

에 대해 일정한 기울기를 갖고 증가하고 있으며, 각 복호시간에 버퍼를 빠져나가는 양은 해당 Picture의 부호화시 발생되었던 데이터량과 같다. VBV상의 버퍼 상태 변화량은 부호기 내에 있는 버퍼의 버퍼상태 변화량을 거울을 통해 보는 것과 같은 대칭형태이다. Low Delay Mode일 경우에는 VBV에서 Overflow는 발생하지 않으나 Underflow는 발생할 수 있다. 이 경우 직전에 복호가 완료된 화면을 계속 보여주고 정상적인 VBV 동작이 일어날 때까지 기다린다. 이러한 경우를 'Skipped Picture'라고 한다.

Low Delay Mode가 아닐 경우에는 Skipped Picture가 허용되지 않고 따라서 Overflow나 Underflow가 생기지 않도록 부호화하여야 한다. VBV의 버퍼상태는 버퍼의 최대 크기인 B_{max} 와 0 사이의 값을 갖도록 한다. 실제 구현할 복호기는 VBV상에서 Underflow나 Overflow가 생기지 않도록 설계하여야 한다. VBV의 초기화는 vbv_delay라는 파라미터를 통해 이루어지는데, 최초 VBV 버퍼가 비어있는 상태에서 시작하여 sequence_header()부터 picture_start_code까지 입력한 후, picture_header()에 들어있는 vbv_delay 시간 동안 데이터를 채워넣는다. 실제 복호기에서도 이와 같이 동작하여 VBV 버퍼의 움직임과 동일하게 움직이게 된다. 그 후 계속적으로 VBV 버퍼상태를 감시하는데, 한 Picture 데이터를 빼내기 전과 후에 버퍼상태를 추출하여 이 값이 항상 VBV의 크기인 B_{max} 와 0 사이에 있는지 감시한다. 주기적으로 감시하는 시간은 Frame Picture인지 Field Picture 인지와 repeat_first_field상의 값 등에 따라 달라지게 된다.

4. 요약 및 결론

본 논문에서는 최근 전세계적으로 실현되고 있는 디지털 방송의 기반이 되는 표준인 MPEG-2의 비디오 압축방식에 관하여 기술하였다. 디지털 방송은 여기에 오디오 압축, 다중화, 채널부호화, 변조 등이 결합되어 이루어지므로 다양한 신호처리 및 통신 기술에 관한 연구가 선행되어야 한다. 또한 이를 구현하기 위한 VLSI 기술, 네트워크 기술, 소프트웨어 기술 등도 함께 연구되어야 한다. 디지털 방송은 궁극적으로 멀티미디어 정보화 사회의 근간이 되는데 최근 디지털 방송에 각종 부가 서비스 기능을 넣기 위한 표준화가 북미와 유럽을 중심으로 진행되고 있다. 또한 차세대 표준인 MPEG-4가 표준화 최종 단계까지 와 있다.

디지털 방송의 성공적 도입을 위한 학계, 연구계, 산업계, 정부의 효율적 역할 분담과 협력이 절실히 요구된다.

5. 참고문헌

1. ISO/IEC 13818-2: Information Technology - Generic Coding of Moving Pictures and Associated Audio, International Standard, 1996
2. ISO/IEC 11172-2: Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbps, International Standard, 1993
3. ITU-R Rec. H.261: Video Codec for Audio Visual Services at p x 64 kbps, March 1993
4. 정 제창 역, "그림으로 보는 최신 MPEG", 멀티미디어 통신 연구회 편, 교보문고, 1996

표 1. MPEG-1과 MPEG-2의 규격비교표

사 양	MPEG-2(MP@ML)	MPEG-1
화소수	720×480	360×240
전송률	15Mbps 이하	1.5Mbps 이하
화면처리모드	Frame 모드/Field 모드	Frame 모드
움직임 벡터의 예측 방법	Frame간 예측/Field간 예측	Frame간 예측
IDCT mismatch	Q^{-1} 후 DCT 계수의 합이 짝수인 경우[7][7] 계수의 마지막(LSB) 비트를 반전	Q^{-1} 후 발생한 각각의 DCT계수가 짝수일 경우 절대값을 1 감소시킨다.
Escape sequence syntax	VLC할 수 없는 (Run, Level)값은 Escape code(6bit)+FLC(18bit)를 사용	VLC할 수 없는 (Run, Level)값은 Escape code(6bit)+FLC(14bit) 또는 FLC(22bit)를 사용.
Escape sequence usage	VLC할 수 있는 (Run, Level)값은 escape format을 사용하는 것이 허용되지 않음	VLC할 수 있는 (Run, Level)값은 escape format을 사용하는 것이 허용됨.
Chroma samples horizontal position	색차신호의 수평방향 샘플위치는 휘도신호의 샘플위치와 일치함.	색차신호의 수평방향 샘플위치는 휘도신호의 샘플들 중간에 위치함.
Slices	slices는 반드시 MB상의 같은 row상에서 시작되고 끝나야 한다.	slices는 반드시 MB상의 같은 row상에서 시작되고 끝날 필요는 없다.
D-picture	D-picture는 사용하지 않는다.	picture coding type의 하나인 D-picture가 정의되어 있다.
Full-pel motion vector	full_pel_forward_vector와 full_pel_backward_vector flag가 반드시 "0"이어야 한다. 즉 half-pel 움직임벡터만을 사용한다.	full_pel_forward_vector와 full_pel_backward_vector flag를 1로 setting하면 full-pel 움직임벡터를 사용하는 것이 가능.
Aspect Ratio information	sequence header에 있는 4bit aspect_ratio_information은 display aspect ratio이고 pel aspect ratio는 이것과 frame size, display size로부터 구해짐	4bit의 pel_aspect_ratio 값이 sequence header에 명시됨
Forward_f_code & Backward_f_code	움직임벡터를 디코딩하는 데 필요한 f_code는 picture header에 있는 forward/backward_horizontal/vertical_f_code 임	움직임벡터를 디코딩하는 데 필요한 f_code는 picture header에 있는 forward_f_code와 backward_f_code임.
constrained_parameter_flag & Maximum horizontal_size	constrained_parameter_flag 개념의 syntax는 profile과 level 개념으로 바뀌었다. MPEG-2 MP@ML에서는 horizontal_size는 720 pixel을 넘지 못한다.	constrained_parameter_flag가 1로 설정되어 있으면 constraints들을 만족하고 있고 horizontal_size도 768 미만임.

표 2. MPEG-2 MP@ML의 주요사항

프레임포맷	720×480(주파수 29.97Hz) 720×576(주파수 25Hz)	
부호화전송률	최대 15 Mbps	
Y, Cr, Cb 표본화 비	4 : 2 : 0	
화면의 종류	I(Intra), P(Predictive), B(Bidirectional) Picture	
부호화단위	프레임(Frame)구조/필드(Field)구조	
움직임 벡터 예측	프레임구조	프레임(16×16)/필드(16×8)/dual prime(M=1인 경우에만 존재)
	필드구조	필드(16×16)/필드(16×8)/dual prime(M=1인 경우에만 존재)
움직임벡터 탐색범위	-128 화소 ~ +127.5 화소, half-pixel단위	
버퍼의 크기	1.75 Mbit(1,835,008 bit)	
호환성	MPEG-1에 대해 상위 호환성(forward compatibility)	
DCT DC 정밀도	8 bit/9 bit/10 bit/11 bit	
VLC table	MPEG-1과 동일테이블/신규테이블	
DCT계수의 scan 방법	MPEG-1과 동일스캔(zig-zag)/alternate 스캔	
VBR(variable bit rate)	대응함	

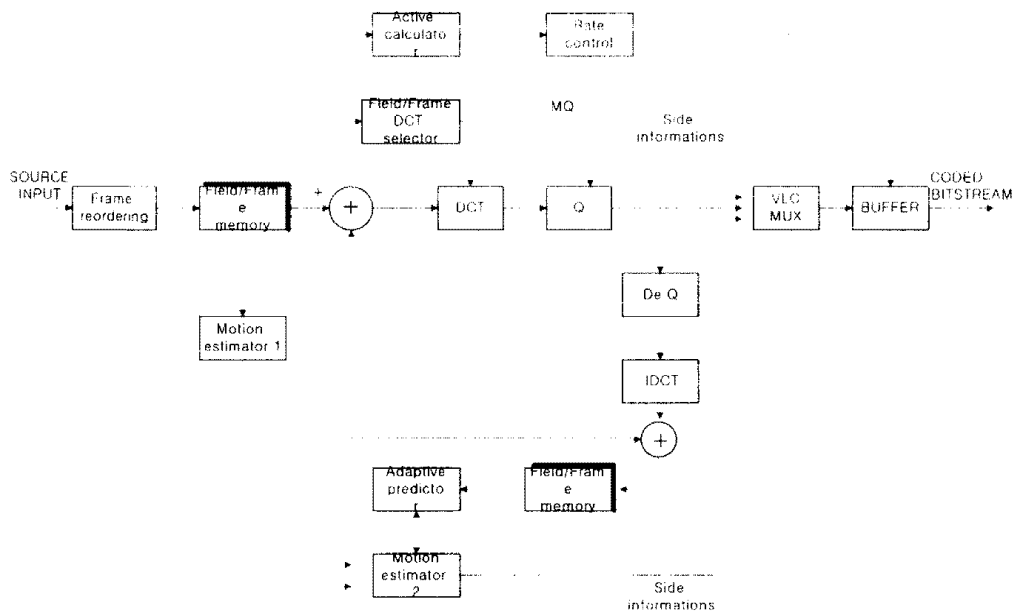


그림 1. MPEG-2 비디오 부호기

표 3. MPEG-2 입출력 영상규격

	4:4:4-625	4:4:4-525	4:2:2-625	4:2:2-525	4:2:0-624	4:2:0-525
유효라인수						
휘도신호	576	480/483	576	480/483	576	480/483
색차신호	576	483	576	480	288	240
라인당 유효화소						
휘도신호	720	720	720	720	720	720
색차신호	720	720	360	360	360	360
프레임률	25	30	25	30	25	30

8	16	19	22	26	27	29	34	16	16	16	16	16	16	16	16
16	16	22	24	27	29	34	37	16	16	16	16	16	16	16	16
19	22	26	27	29	34	34	38	16	16	16	16	16	16	16	16
22	22	26	27	29	34	37	40	16	16	16	16	16	16	16	16
22	26	27	29	32	35	40	48	16	16	16	16	16	16	16	16
26	27	29	32	35	40	48	58	16	16	16	16	16	16	16	16
26	27	29	34	38	46	56	69	16	16	16	16	16	16	16	16
27	29	35	38	46	56	69	83	16	16	16	16	16	16	16	16

그림 2. Intra Block과 non-Intra Block의 가중값행렬