

MPEG-2 표준 영상 압축 기법의 최근 동향

金在均, 金旻圭

韓國科學技術院 電氣 및 電子工學科

I. 서론

국제 표준화 기구인 ISO-IEC(International Organization for Standardization - International Electrotechnical Commission) 산하에 정보기술 표준화를 위한 합동 기술 위원회(JTC1-Joint Technical Committee 1)가 있다. 여기에 소속된 여러 기술 분과 위원회 중의 하나로서 "coded representation of audio, picture, multimedia and hypermedia"를 위한 SC(sub-committee)-29가 있다. SC-29에 속한 4개의 작업반(WG-working group)중의 하나인 WG-11이 MPEG(moving picture experts group)으로 호칭되고 있다.

MPEG는 "coding of moving pictures and associated audio"를 위한 표준화 연구 작업반으로서, 1988년에 SC-2 기술 분과위원회의 한 연구반으로 시작되었다가 1991년 11월부터 SC-29 소속의 연구반이 되었다. 따라서 MPEG는 현재 ISO-IEC/JTC1/SC29/WG11에 속한다. 1991년 11월에 속칭 MPEG-1 표준안(CD-committee draft)이 완성되었으며, 현재 MPEG-2 표준안의 내용이 활발히 연구 논의 중에 있다. MPEG의 압축기법은 소위 디지털 저장 매체(DSM - digital storage media)를 위한 일반적 부호화기법(generic coding)으로서, 그 응용범위가 컴퓨터 및 가전기기, 통신, 방송 등으로 매우 넓기 때문에, 이에 대한 관심도는 해가 갈수록 더욱 높아지고 있다.

MPEG-1 표준안¹⁾은 최대 비트율 1.5 Mbits/sec의 데이터량으로서 가정용 TV급의 영상과 CD 품질의 오디오 정보 저장을 목표로 하고 있다(별첨부록 참조). MPEG-2 표준안은 비트율 4-10 Mbits/sec의 데이터

량으로서 NTSC, PAL, SECAM 등의 기존 방송 화질 또는 CCIR601의 디지털 TV 화질을 기본 목표로 하면서도 EDTV, HDTV 등의 새로운 TV 형식도 수용할 수 있는 확장성을 갖도록 하고 있다(1992년 7월 제 19차 MPEG meeting).

MPEG 연구반은 현재 CCITT의 ATM video coding 연구반과 공동 작업 관계를 유지하고 있으며, 방송 분야의 CMTT/2 연구반과도 상호 연락체계를 갖추고 있다. 표준안의 진행 일정은 93년 3월의 표준시안(frozen draft) 작성, 93년 11월의 서울 회의에서 표준안(committee draft) 작성을 목표로 하고 있다.

MPEG 부호화 기법은 동영상 압축, 관련 오디오 압축, 그리고 시간 동기 및 정보 다중화 등을 위한 시스템 등의 세가지 분야가 있으나, 본고에서는 동영상 압축 분야만을 고찰하고자 한다. MPEG-2의 오디오 부분은 상대적으로 진행이 늦은 편이다. 먼저 MPEG-2 표준안의 기본 요건들(requirements)을 정리한 다음 MPEG-2 시험 모델(TM-test model)의 주요 내용을 요약하였다.

II. MPEG-2 영상 압축 부호화 표준안의 기본 요건

MPEG 영상 압축 부호화 기법은 두가지 특징이 있다. 첫째는, 여러가지 디지털 기록 매체(DSM)에 공통적으로 적용될 수 있도록 되어있다. 그러므로 여러 매체의 공통적이고 일반적인 부분에 국한된 표준 기법이다(generic coding). 둘째로는 영상(video)의 부호화 표현(coded representation), 즉 비트열의 구문(bitstream syntax)과 이것으로 부터의 복호화 방법만을 명

시하는 표준이다(양립성을 위한 최소 기준). 따라서 부호화기 자체의 설계 및 구현 방법을 명시하는 것이 아니다. 그러므로 적용 분야와 설계자에 따라 부호화기의 기능, 구조, 성능, 복잡도 등이 달라질 수 있다.

표 1은 MPEG-2 기법의 주 응용분야 및 전송율, 그리고 기록 매체 및 송수신의 복잡도를 비교한 것이다. 여러가지 DSM들에의 응용을 위해서 MPEG-2 기법을 개발할 때 고려해야하는 요건들은 다음과 같다^[2].

1. 입출력 영상의 규격

MPEG-2 표준 기법은 격행주사(interlaced scan)의 CCIR601 (4:2:2, 4:2:0, 4:4:4 규격) 영상을 비롯하여 EDTV, HDTV, 그리고 필름 영상과 같은 순행 주사(progressive scan)의 영상 등을 지원하도록 한다. 따라서 MPEG-2는 디지털 위성방송, 기존 TV의 디지털 방송, 케이블 TV, 고선명 그래픽스, 의용 장비, video 극장, EDTV, 그리고 HDTV에 이르기 까지 그 응용분야가 다양하다.

2. 재생 화질

MPEG-2는 MPEG-1에 비해 고화질을 요하는데, CCIR601 영상을 입력으로 할 때 재생영상의 화질은 비트율과 연관되어 다음의 화질을 얻을 수 있어야 한다.

- 3-5 Mbits/s: NTSC/PAL/SECAM 등의 방송 화질
- 8-10 Mbits/s : CCIR601에 근접한 화질

3. 비트율의 유연성

ATM 환경에서의 영상 전송, DSM 등에서의 다중채널 전송, 그리고 DBS등에의 응용을 위하여 고정 비트율(CBR-constant bit rate) 부호화, 가변 비트율(VB-R-variable bit rate) 부호화 모두가 가능해야 한다.

4. 부호화/ 복호화 지연(Coding/ Decoding Delay)

인코더 입력단과 디코더 출력단 사이에서 frame 재배열(reordering), 영상 규격 변환(format conversion), 영상 부호화(source coding), 버퍼 지연(buffer delay), 전송 부호화 및 복호화(transmission coding and decoding) 등에 의해서 지연되는 시간을 말하며, 실시간 통신 분야에서는 150 msec 이하이어야 하며, 그외의 응용분야에서는 500 msec 이하이어야 한다.

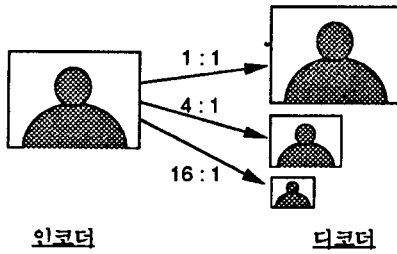
5. 비트열 Scalability

디코더에서 부호화된 비트열 중 일부분을 취하여 인코더와는 다른 공간해상도를 갖는 영상을 직접적으로 재생하거나(resolution scalability), 혹은 다른 시간해상도를 갖는 영상을 직접적으로 재생하는(temporal scalability) 특수재생 기능을 기본적인 scalability라 정의한다(그림 1). 이밖에 다양한 구조 및 복잡도를 갖는 수신기(receiver)로 영상을 재생할 수 있게 하는 complexity scalability와 동일한 해상도를 가지나 다양한 화질의 영상을 재생할 수 있게 하는 coding noise scalability 등도 정의될 수 있다. 여기서 부호화된 비트

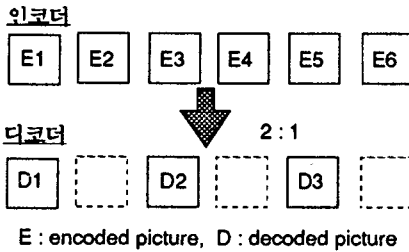
표 1. MPEG-2의 응용 분야

Application Area	Bit rate	Complexity*
ISM : Interactive Storage Media (optical disk, etc)	2 - 10	C >> D
SSM : Serial Storage Media (digital VTR, etc)	5 - 10	C ≅ D
NDB : Networked Database Services (via ATM, etc)	2 - 10	C >> D
ENG : Electronic News Gathering (Satellite ENG)	5 - 10	C << D
IPC : InterPersonal Communications (videoconferencing, videophone, etc)	2 - 10	C ≅ D
RVS : Remote Video Surveillance	2 - 10	C << D
CTV : Cable TV Distribution (on optical networks, copper, ect)	5 - 10	C >> D
STV : Satallite TV Broadcasting	5 - 10	C >> D
TTV : Terrestrial TV Boradcasting	5 - 10	C >> D

*C : Coder, D : Decoder



(a) Resolution scalability



(b) Temporal scalability

그림 1. Scalability의 예

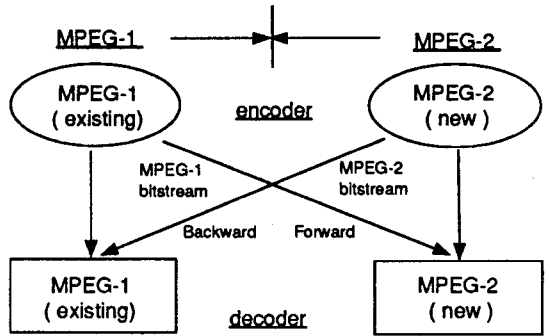


그림 2. MPEG-1과 MPEG-2사이의 호환성 관계

표 2. 인코더와 디코더의 호환성 관계

호환모드	인 코 더	디 코 더
순방향	기존 standard	새로운 standard
역방향	새로운 standard	기존 standard
상 향	저해상도	고해상도
하 향	고해상도	저해상도

열로 부터 전체 영상을 재생한 후 부표본화(subsampling)하는 과정은 포함되지 말아야 한다.

Scalability의 채택은 근본적으로 부호화 효율을 저하시키거나 workstation 등과 같이 다중 해상도(multi-resolution) 기능을 갖는 응용 분야, TV 시스템에서 picture-in-picture, picture-out-picture 등의 응용분야, 그리고 서로 다른 대역폭을 갖는 채널의 접속(예 : satellite→cable system)등의 응용 분야에서 요구되는 기능이다.

6. 호환성(Compatibility)

각각 다른 규격을 갖는 표준화 기법들을 사용하여 영상신호를 부호화하고 복호화 할 수 있을 때 두 표준기법들 사이에는 호환성이 있다고 정의한다. MPEG-2는 다음의 호환성 모드가 고려되어야 한다.

- MPEG-1, H. 261과의 순방향/역방향(forward/backward) 호환성
- EDTV, HDTV, SDTV와의 상향/하향(upward/downward) 호환성

표 2는 호환성을 갖기 위한 송수신단의 관계를 나타낸 것이며, 그림 2는 MPEG-1과 MPEG-2 사이의 호환성 관계를 나타낸 것이다. 호환성은 서로 다른 표준안 사이의 교류나 제조 및 대량 생산성 측면에서 고려해야 하는 기능이다.

그림 3은 호환성을 부여할 수 있는 대표적인 방안들을 나타낸 것이며, 각각은 다음의 특성을 갖는다.

1) Embedded 비트열

CCIR601 영상을 MPEG-1에서 요구되는 SIF 영상과 나머지(CCIR 601-SIF) 영상으로 분리한 후 각각을 MPEG-1 인코더와 MPEG-2 잔여 인코더로 부호화한 후 다중화(multiplexing)하여 전송한다. 전체적으로 MPEG-1 인코더가 MPEG-2 인코더에 포함된(embedded) 형태로 MPEG-1과 MPEG-2 사이에 완전한 호환성(full compatibility)를 부여하나 복잡도가 높은 방안이다.

2) Simulcasting

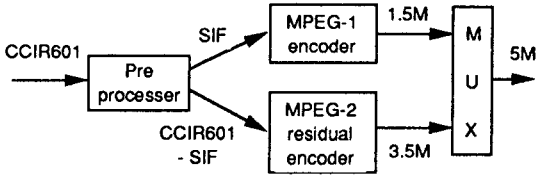
CCIR601 영상과 SIF 영상을 MPEG-2 인코더와 MPEG-1 인코더를 사용하여 각각 독립적으로 부호화한 후 다중화하여 전송한다. 역방향 및 하향 호환성만이 존재하며 완전 호환성이 존재하기 위해서는 MPEG-2 인코더가 syntactic extension 호환적이어야 한다.

3) Syntactic extension

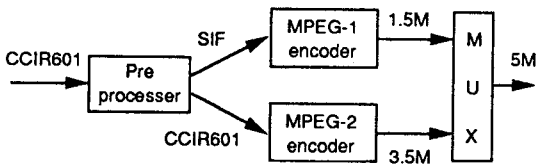
MPEG-2 syntax의 주요 부분이 MPEG-1 syntax로 구성되어 있기 때문에 나머지 부분(extension syntax)이 없어도 부호화 및 복호화가 가능하게 syntax를 구성하는 방식이다. 순방향 및 상향 호환성만을 제공할 수 있는 방안이다.

4) Switchable

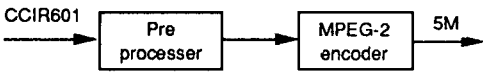
MPEG-1 인코더 및 MPEG-2 인코더를 선택적으로 사용하여 부호화하는 방안으로 경우에 따라서 완전 호환성이 존재할 수 있는 방안이다.



(a) Embedded

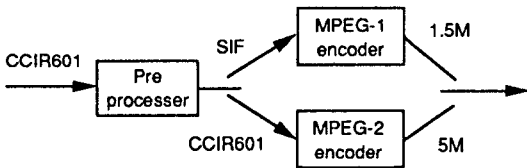


(b) Simulcast



MPEG-2 syntax = MPEG-1 syntax + extension syntax

(c) Syntactic extension



(d) Switchable

그림 3. 여러가지 호환적 방식

7. 재생 기능

화면 재생시 정상 속도에서의 전후 방향으로의 재생(normal forward/backward playback)은 MPEG-2의 기본 기능이며, 이외에 고속 순방향/역방향 재생, 저속 재생, 정지 모드 등이 가능해야 한다.

8. Random Access/ Channel Hopping

상호 작용의 DSM, TV 전송 등에서의 응용을 위하여 DSM에 저장된 부호화된 비트열 중 임의의 위치에서

재생이 가능(random access)해야 하며, TV 채널 선택이 자유로와야 한다(channel hopping). 또한 각 기능의 시간 지연은 각각 0.4초, 0.3-1초 이내이어야 한다.

9. 시스템의 복잡도

MPEG-2 시스템의 복잡도와 성능사이는 서로 절충 관계로 유연성이 필요하며, 표 1과 같이 응용 분야에 따라서 인코더와 디코더의 복잡도가 달라진다. 또한 최소 오류를 요하는 기록 매체에의 응용을 위해서 intra-frame 모드 만을 갖는 codec도 고려해야 한다.

10. 기타 고려 요건

MPEG-2는 표 1에서와 같이 각각 다른 기록 매체 및 전송로를 고려하는 응용 분야에 적용되어야 하기 때문에 다음의 사항들을 추가적으로 고려하여 개발되어야 한다.

- 압축된 비트열의 편집 기능 부여
- 재부호화 손실(re-coding loss)의 최소화
- 오류 방지(error protection) 및 cell/packet loss resilience
- 실시간 aspect ratio 변경 등

Ⅲ. MPEG-2 동영상 부호화 기법의 개요

MPEG-2의 영상 압축기법은 기존의 MPEG-1의 구조를 기초로 하여 개발되고 있으며, 기본적인 부호화 원리는 다음과 같다.

- 시간 방향의 중복성(temporal redundancy)은 이동 보상형 DPCM(motion compensated DPCM)으로 제거한다.
- 공간 방향의 중복성(spatial redundancy)은 변환 부호화(transform coding)와 양자화(quantization)과정에서 제거한다.
- 통계적 중복성(statistical redundancy)은 엔트로피 부호화(entropy coding)과정에서 제거한다.

그러나 MPEG-2의 목표와 응용환경은 II장의 기본 요건들에 나타난 바와 같이 MPEG-1과 크게 다르기 때문에 각 세부 기법들은 이를 고려하여 개발되고 있다. 현재 MPEG-2 부호화 기법은 TM0^[3], TM1^[4]을 거쳐 TM2^[5]가 완성된 단계이며 1992년 11월에 MPEG-2의 초안(working draft)이 완성될 예정이다. 이제 본 장에서는 TM2를 중심으로 MPEG-2 동영상 압축 기법의

기본원리와 MPEG-1과 크게 달라지는 세부 기법들을 살펴보고자 한다.

1. MPEG-2 부호화기법의 기본원리

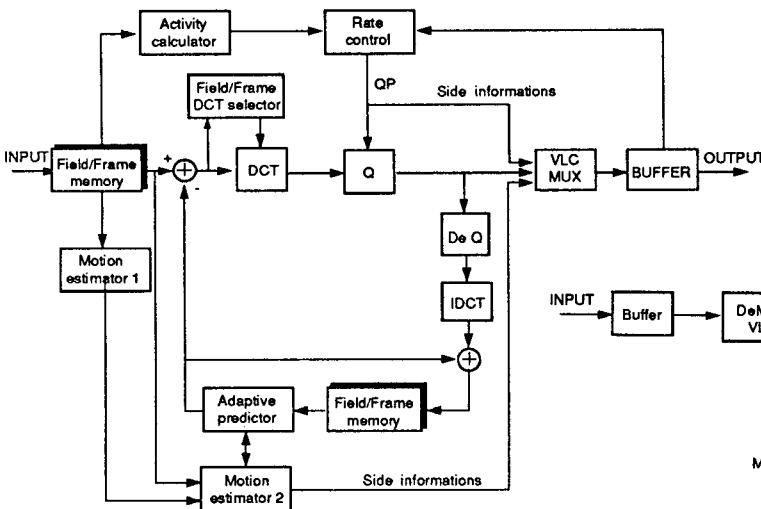
그림 4는 MPEG-2 부호화기의 전체 구성도로서 영상간 예측 부호화와 intraframe 부호화를 수행할 수 있는 구조로 되어있다. 주요 부분은 영상간 예측을 위한 이동 추정기(motion estimator), 예측기(predictor), field/frame memory와 intraframe 부호화를 위한 DCT 변환기와 양자화기(quantizer), 그리고 허프만 부호화를 위한 VLC(variable length coder) 등으로 구성되어 있다. 표 3은 TM2에서 이용하는 주요 부호화기법을 나타낸 것이다. MPEG-2는 각 picture별로 서로 다른 부호화기법을 사용하는데, 각 picture의 부호화 과정 및 특징은 다음과 같다.

1) I-picture

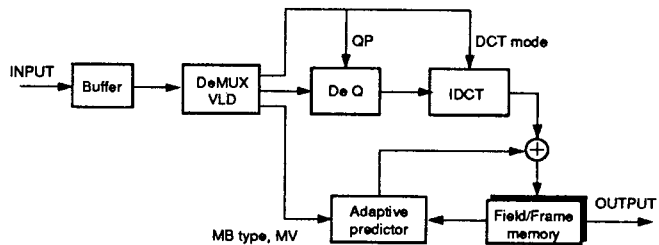
I-picture는 블럭별로 DCT 변환과 양자화 과정을 적용하여 공간방향의 중복성만을 제거한다. GOP(group of pictures)의 구조상 I-picture의 화질이 전체 GOP의 화질에 절대적인 영향을 미치므로 (특히 P-picture들의 부호화 성능에 중요한 영향을 준다) 고품질을 유지하도록 부호화해야 한다. 이를 위해서 양자화 및 허프

표 3. MPEG-2 test model 2의 부호화기법

Basic video format	-Interlaced/progressive format -CCIR 601 4:2:2/4:2:0 format
Pre/Post-processing	CCIR 601 4:2:2 ↔ CCIR 601 4:2:0
Layered structure of video data	sequence, GOP, picture, slice (slave-slice), MB (slave-MB), block (scaled-block, slave-block)
Coding mode	-IPB-picture mode -low delay coding mode
Motion estimation and compensation	-frame/field adaptive ME & MC -ME & MC unit : 16×16 / 16×8 MB -1/2 or 1/4 pels accuracy -+ / -15.5 or more pels / frames
Transform	frame/ field adaptive DCT
Block scanning	zig-zag scanning
Q-step update	M-Quant
Q-weighting matrix	Intra/Non-intra coded MB
Entropy coding	VLC coding
Rate control	feedback/feed-forward control
Compatibility	spatial layering
Scalability	-spatial layering -frequency layering



(a) Encoder



(b) Decoder

그림 4. MPEG-2 부호화기의 전체 구성도

만 엔트로피 부호화시 intraframe 부호화된 영상은 interframe 기법과 구분하여 처리하여 고품질을 유지하도록 한다.

2) P-picture

P-picture는 순방향의 영상간 예측 부호화(forward interframe predictive coding)기법과 DCT 변환기법을 사용하여 데이터를 압축한다. 입력된 현재 영상에 대해서 이전의 I-picture 및 P-picture를 기준으로하여 이동보상형 예측기법을 적용하여 영상간에서 물체의 운동으로 기인한 변화를 보상하여 영상간 중복성을 제거한다. 그 후 영상간의 중복성이 제거된 예측오차에 대해서 DCT 변환과 양자화 과정을 적용하여 나머지 공간방향의 중복성을 제거한다. GOP의 구조상 P-picture는 한 GOP내에서 연속되는 P-picture 및 B-picture들의 화질에 영향을 미치므로 I-picture에 비해서는 적으나 B-picture에 비해서는 많은 비트를 할당하여 부호화한다.

3) B-picture

B-picture는 양방향성 영상간 예측 부호화(bidirectionally interframe predictive coding) 기법과 DCT 변환기법을 사용하여 데이터를 압축한다. 입력된 현재 영상에 대해서 이전의 I-picture 및 P-picture와 다음의 P-picture로 부터 각각 이동보상형 예측된 영상 및 이동보상형 내삽(motion compensated interpolation)된 영상을 사용하여 세가지의 영상간 예측신호를 얻은 후 이들 예측신호 중 최선의 것을 영상간의 예측신호로 선택하여 영상간의 중복성을 제거한다. 그 후 영상간의 중복성이 제거된 예측오차에 대해서 DCT 변환과 양자화 과정을 적용하여 나머지 공간방향의 중복성을 제거한다. GOP의 구조상 B-picture의 화질은 GOP 내의 다른 picture에 영향을 미치지 않으므로 주관적인 화질만을 고려하여 다른 기법에 비해 상대적으로 적은 비트를 할당하여 부호화를 수행한다.

그림 5는 $N = 15, M = 3$ 일 때, MPEG-2 부호화 기법에서 사용하는 GOP 구조를 나타낸 것이다. 여기서 N은 GOP내의 picture의 갯수를 나타내며, M-1은 두 P(I)-picture 사이에 있는 B-picture의 갯수를 나타낸다.

2. 입력/ 출력 영상의 규격

표 4의 CCIR 601 4:2:0 형식의 영상을 부호화하는 것을 기본으로 하며, CCIR 601 영상의 한 frame은 1/60초의 시간 간격이 있는 두 field(odd 와 even)로 구성되어 있다. 이 밖에 compatibility/scalability를 위한 SIF(MPEG-1의 source input format), SIF-odd/even, SIF-Interlaced, CIF(H. 261의 common intermediate format), HHR(half-horizontal resolution) 형식의 영상도 정의하여 이용한다. 이들 규격의 영상은 CCIR601 4:2:2 영상으로 부터 전처리 과정(filtering과 subsampling)을 통해서 얻는다.

표 4. MPEG-2의 CCIR601 영상 규격

	4:2:2-525	4:2:2-625	4:2:0-525	4:2:0-625
Active line수				
• Luma (Y)	483	576	480	576
• Chroma (Cb, Cr)	483	576	240	288
Line당 화소 수				
• Luma (Y)	720	720	704	704
• Chroma (Cb, Cr)	360	360	352	352
Frame rate(Hz)	30	25	30	25
Aspect ratio	4:3	4:3	4:3	4:3

3. Video Data의 계층 구조

MPEG-2 부호화 기법에서 사용되는 영상 데이터는 비트열 syntax를 고려하여 sequence, group of pictures(GOP), picture, slice, macroblock(MB),

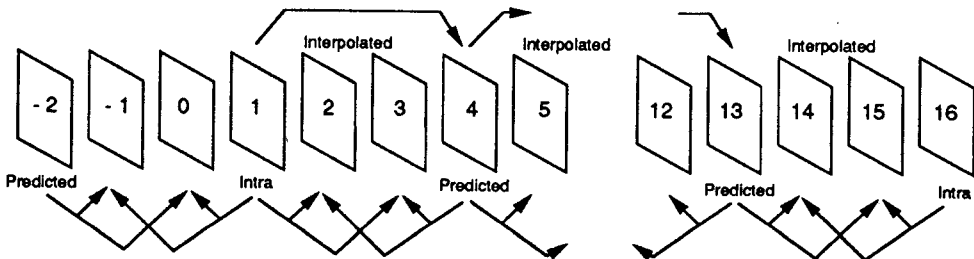


그림 5. Display 순서에서 본 frame-picture coding의 GOP 구조

block 등의 계층 구조로 구분된 후 이용되는데 각 계층 영상의 구조 및 사용 용도는 다음과 같다.

1) Video sequence layer

Video sequence는 부호화에 입력된 영상들의 집합으로 정의되며, 그림 6(a)에서와 같이 연속된 GOP의 배열로 구성된다.

2) GOP layer

그림 6 (b)에서와 같이 1개의 I-picture, $P=(N-1)/M$ 개의 P-picture, 그리고 $(N-1-P)$ 개의 B-picture로 구성되며, B-picture로 시작되어 P-picture에서 끝난다. GOP 매개변수인 N, M 등은 H/W의 복잡도, 부호화 지연, 부호화 성능 등과 밀접한 관계를 갖기 때문에 응용분야에 따라 신중적으로 결정되어야 한다. 이와 같이 video sequence를 GOP별로 부호화 하는 것은 시간축 방향으로의 오류 전파를 막으며, random access 및 고속 재생 등의 특수 재생기능을 가능하게 하기 위함이다.

3) Picture layer

하나의 CCIR601 형식의 영상을 의미하며 원칙적으로 각 picture별로 intraframe /interframe 부호화를 수행한다. 부호화 기법에 따라 picture는 I-picture, P-picture, B-picture로 구분된다.

4) MB slice layer

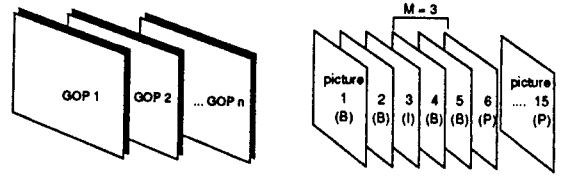
각 picture에서 slice는 그림 6(c)에서와 같은 배열로 구성되며, 원칙적으로 한 slice내의 MB의 수는 가변적으로 정할 수 있다. 매 slice 마다 새로운 양자화기의 계단크기(step size)가 결정될 수 있으며, 각 slice의 위치 정보가 전송되어 디코더와 동기(synchronization)를 가능하게 한다.

5) MB layer

MB slice내에서 각각의 MB는 그림 6(d)에서와 같은 배열로 구성된다. 이때 하나의 MB는 그림 6(e)에서와 같이 16×16 의 Y 신호와 8×8 의 Cr 및 Cb 성분으로 구성되는 4:2:0 MB(1-6 blocks) 8×16 의 Cr 및 Cb 성분으로 구성된 4:2:2 MB (1-8 blocks), 그리고 16×16 의 Cr 및 Cb 성분으로 구성된 4:4:4 MB(1-12 blocks)로 구분된다. 매 MB 마다 slice 내에서 MB의 위치 정보가 전송되며, 원칙적으로 이동보상예측, DCT 변환, 양자화기의 조정은 MB 단위로 수행된다. 또한 각 블럭 단위로 부호화된 데이터에 대한 전송 여부를 MB 단위로 판정하는데, 이때 전송할 데이터가 발생치 않는 MB를 skipped MB라 한다.

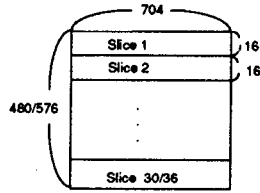
6) Block layer

하나의 MB는 그림 6 (e)에서와 같이 6개(4:2:0

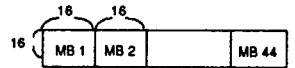


(a) Video sequence layer

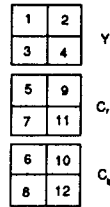
(b) GOP layer



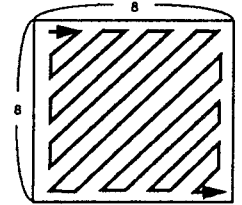
(c) Picture layer



(d) MB slice layer



(e) MB layer



(f) Block layer

(zig-zag scan order)

그림 6. Video data의 계층 구조

MB) 8개, (4:2:2 MB), 또는 12개(4:4:4 MB), 또는 12개 (4:4:4 MB)의 block으로 구성된다. DCT는 block 단위로 수행되며 변환된 DCT 계수는 그림 (f)에서와 같은 순서로 zig-zag 순서로 허프만 부호화되어 전송된다.

4. Picture Sequence의 구조

현재 MPEG-2의 TM에는 다음의 세가지 picture sequence 구조가 제안되어 있다.

1) Frame sequence + frame-picture

하나의 frame에 속한 두 field(odd, even)를 하나의 picture로 간주하고, 하나의 16×16 MB는 두 16×8 크기의 odd sub-MB와 even sub-MB가 결합된 형태이다 (그림 7(a)). 즉 frame 단위로 GOP가 구성되며 부호화 된다.

2) Frame sequence + field-picture

하나의 picture는 두 field로 구성되지만, 두 field는

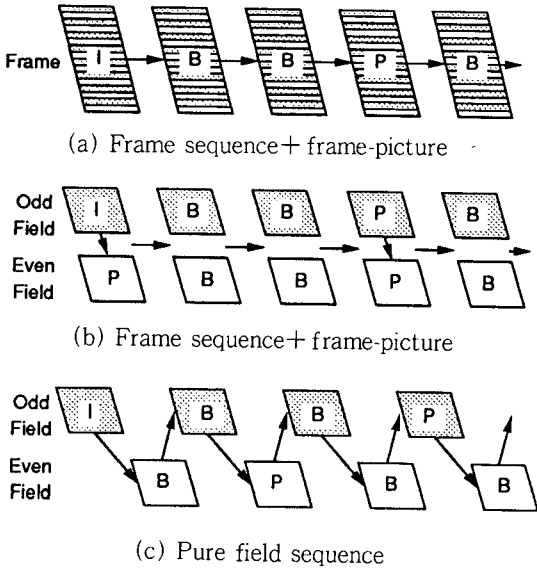


그림 7. MPEG-2의 picture 구조

완전히 독립적으로 부호화된다. 따라서 뒤따르는 even field는 동일한 picture 내의 odd field를 이용한 이동보상 예측이 가능하며, 하나의 16×16 MB는 오직 하나의 field 성분으로 구성된다(그림 7(b)). 즉 frame 단위로 GOP가 구성되며 field 단위로 부호화 된다.

3) Pure field sequence

하나의 picture는 하나의 field만으로 구성된다(그림 7(c)). 즉 field 단위로 GOP가 구성되며 부호화된다. 인코더에서는 위의 세가지 picture 구조 중 하나로 부호화되고, 디코더에서는 세가지 중 어느 것이 전송되어 오더라도 복원이 가능해야 한다. Frame-picture 구조는 성능은 우수하나 부호화/복호화 지연이 길고, memory가 많이 소요된다. 반면 field-picture 구조는 지연과 memory는 감소하나 성능이 다소 떨어진다.

5. 이동 벡터 추정 및 이동보상 예측

복합 부호화기법에서 interframe 예측효율을 높이기 위하여 이동보상 예측기법을 일반적으로 이용하는데, SIF 영상을 기본으로 하는 MPEG-1에 비해서 MPEG-2는 CCIR601 영상을 기본으로 하기 때문에 보다 다양하고 복잡한 이동벡터 추정 및 예측기법을 필요로 한다. 현재 TM2에는 frame MC(motion compensation), field MC, dual field MC, FAMC(field-time adjusted MC), SVMC(single vector MC), 그리고 dual prime MC 등이 제안되고 표준화를 위한 성능 평가 실험이 이루어지고 있다.

1) 이동벡터 추정(motion estimation : ME)
MPEG-2에서 이동보상 예측을 위한 이동벡터(motion vector : MV)는 크게 frame-based MV와 field-based MV로 나눌 수 있다.

추정과정은 integer 화소 단위의 MV를 추정하는 과정(ME1)과 1/2 화소 단위의 MV를 추정하는 과정(ME2)의 2단계로 이루어지며, 블록정합(block matching) 기법을 이용한다. ME1 과정에서는 원영상을 기준 영상으로 사용하고 ME 2과정에서는 local decoder에서 복원된 영상을 기준영상으로 사용한다. 또 ME2에서 1/2 화소 단위 영상은 주위의 integer 단위 화소들을 이용한 bilinear 내삽으로 얻는다.

2) 이동보상 예측(motion compensated prediction : MCP)

그림 8은 P-picture 및 B-picture에 대해서 이동보상을 하는 경우의 관계를 나타낸다. P-picture(M+1번째 picture)에 대해서는 이동벡터 MV(1, M+1)를 이용한 순방향 이동보상형 예측만이 필요하며, B-picture(k번째 picture)에 대해서는 순방향 예측 및 역방향 예측, 그리고 양방향성 보간예측이 요구된다. 표 5는 이들과 예측방식에 대한 예측 관계식을 나타낸다.

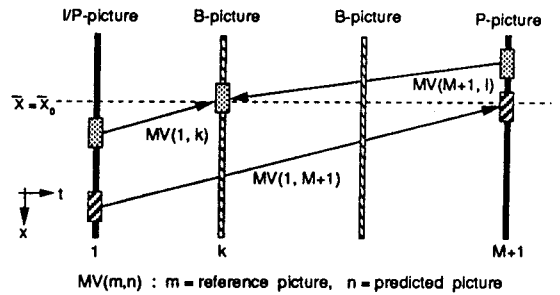


그림 8. 이동보상 예측을 위한 이동벡터의 관계

표 5. P/B-picture에 대한 이동보상 예측 관계식

Picture	예 측	관 계 식
P	순방향	$\tilde{P}_{M+1}(\bar{X}) = P_1(X_0 + MV(1, M+1))$
	역방향	$\tilde{P}_1(\bar{X}_0) = P_M(X_0 + MV(M, 1))$
B	순방향	$\tilde{P}_k(\bar{X}_0) = P_1(X_0 + MV(1, k))$
	양방향	$\tilde{P}_k(\bar{X}_0) = (P_1(X_0 + MV(1, k)) + P_{M+1}(X_0 + MV(M+1, k))) / 2$

이동보상 예측방법은 picture 구조에 따라서 크게 달라지는데 TM2의 주요 MCP 방식은 다음과 같다.

(1) Frame MC

가장 간단한 방식으로서 odd와 even field의 구분없이 frame 구조로 이동 보상 예측된다. 따라서 frame-picture 구조에서만 가능하다. MCP시 1/2 화소 단위의 예측값은 서로 다른 parity의 두 field의 화소들로부터 공간 내삽에 의해 얻는다.

(2) Field MC

각 field 별로 이동보상 예측을 수행하는 방식으로 frame-picture 구조에서는 odd와 even field의 각각 16×8 sub-MB 단위로 예측 신호를 얻는 후, 16×16 크기의 frame 단위 MB에 대한 예측신호로 결합된다. 반면 field-picture 구조에서는 각 field별로 16×16 MB에 대한 예측신호를 얻는다. MCP시 1/2 화소 단위의 예측값은 frame MCP와 달리 하나의 field내의 화소들로부터 공간 내삽에 의해 얻는다.

(3) Dual field MC

그림 9는 dual field prediction의 원리를 나타낸 것으로 같은 parity를 갖는 field로 부터의 예측신호(a1, b2, c1, d2)와 서로 다른 parity를 갖는 field로 부터의 예측신호 (a2, b1, c2, d1)의 산술평균으로 각 field에 대한 예측신호를 구하는 방식이다. 이때 B-picture에서는 순방향 및 역방향 예측 모드만 존재하고 양방향성 보간 모드는 존재하지 않는다. 이외의 모든 사항은 field MC에서와 동일하다.

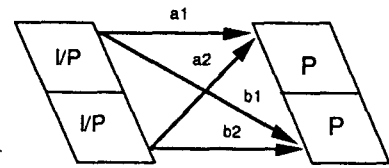
(4) FAMC^{6,7)}

FAMC의 기본 원리는 각 field는 기준 frame 중 같은 parity field 위치로 부터 예측 신호를 얻으며, 두 field의 예측 신호를 얻기 위해서 MB당 오직 하나의 frame 단위 MV를 사용한다는 것이다. 또한 예측 신호의 해상도를 높이기 위하여 그림 10(a)와 같이 서로 다른 parity의 화소를 시간 조절(field-time adjusting)하여 가장적인 평균(weighted average)에 의해 각 field의 예측신호를 얻는다.

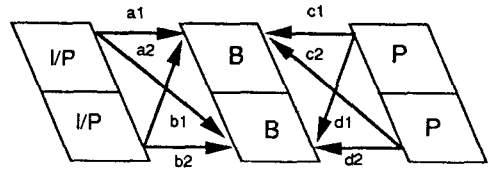
(5) SVMC⁸⁾

SVMC는 field MC, dual field MC, FAMC의 장점을 이용한 방식으로 다음의 네가지 예측 모드가 있다 (그림 10).

- Simplified FAMC
- Same parity field MC
- Near field MC
- Modified dual field MC

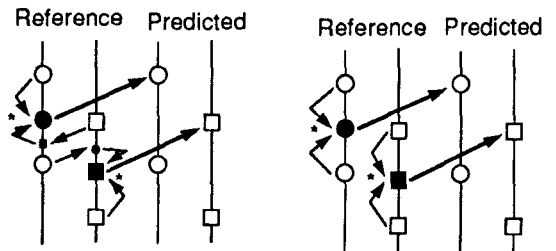


(a) P-picture



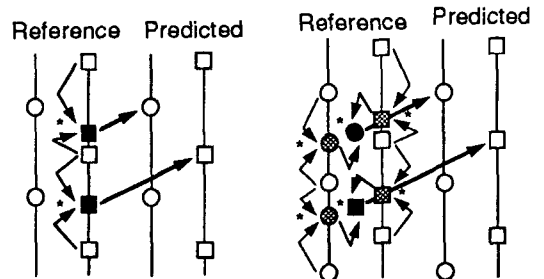
(b) B-picture

그림 9. Dual field MCP의 예 (frame-picture인 경우)



(a) FAMC

(b) Same parity field MC



(c) Near field MC

(d) Modified dual field MC

그림 10. SVMC 예측 모드 (*: averaging)

SVMC는 어떤 예측모드를 사용하든지 MB당 오로지 하나의 MV만을 사용하기 때문에 MV양이 상당히 줄어든다. 네가지 예측 모드중 simplified FAMC와 modified dual field MC에서는 B-picture의 양방향 보간 예측 모드는 사용하지 않는다.

(6) Dual prime MC⁹⁾

TM1에서 정의한 dual field MC를 개선한 방식으로 서 dual field MC가 MB당 내개의 MV가 필요한 반면 dual prime MC는 MB당 한개의 MV와 작은값(+/- 1 이내)을 갖은 하나의 부가적인 차벡터(differential vector)만을 사용한다.

MPEG-2 표준안에서는 위의 여러가지 예측 모드중 다음의 5가지 후보중 하나로 선택될 가능성이 높으며 이에 대한 연구가 활발히 진행 중이다.¹⁰⁾

- Frame + field + dual prime
- Frame + field + sVMC
- Frame + field + sVMC (1/2 화소 정확도)
- Frame + field + dual prime + SVMC
- Frame + field + dual prime + (SVMC - dual)

6. DCT 변환 부호화

DCT 변환은 영상신호의 공간방향으로의 상관성이 큰 것에 바탕을 둔 것으로 영상의 모든 화소에 분산된 에너지를 DC를 포함한 낮은 주파수(sequency)를 갖은 몇개의 변환 계수에 집중시킨 후 이들의 장점을 적절히 이용하는 방식이다.

DCT 변환은 8×8 block 단위로 수행되며, intra-frame 부호화 방식에서 공간방향의 중복성을 제거할 때와 interframe 부호화 방식에서 영상간 예측오차(영상간 예측오차는 영상간의 중복성은 제거되었으나 공간방향의 상관성은 존재한다)에서 공간방향의 상관성을

제거할 때 사용된다. MPEG-2에서는 격행주사 규격 영상의 DCT 변환 효율을 증대시키기 위해서 그림 11과 같이 block을 2개의 구조로 구분한다. 즉 odd field와 even field에서의 영상이 비슷할 때 (field간 운동체의 이동이 작거나 두 field에서의 예측성능이 비슷할 때)는 frame-block을 구성하고, 그 반대의 경우에는 field-block을 구성한 후 DCT 변환을 수행한다.

7. 양자화기

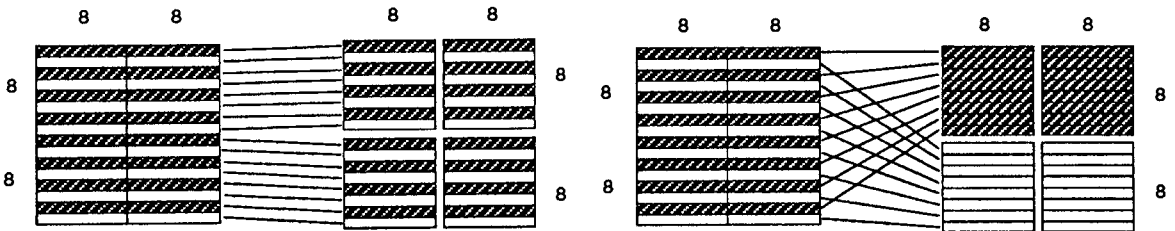
MPEG-2에서는 인간의 시각적인 특성과 부호화 효율을 동시에 고려하여 8×8 block 단위로 양자화를 수행한다. 즉 전체 화질에 중요한 영향을 미치는 순서에 의해서 I-picture > P-picture > B-picture의 순서로 데이터 비트를 할당할 수 있도록 양자화를 수행한다. Intraframe 부호화될 MB와 interframe 부호화될 MB에 대한 양자화 과정은 다음과 같다.

1) Intra-MB

I-, P-, B-picture에서 intra-MB의 변환 계수에 대해서 dead zone이 없는 균등 양자화기(uniform quantizer)를 사용한다. 화질에 가장 큰 영향을 미치는 DC 계수는 양자화 계단 크기를 항상 '8'로 하며 AC 계수는 시각특성을 이용한 그림 12(a)의 weighting matrix를 이용하여 양자화한다. 또 AC 계수는 부호화 효율을 높이기 위하여 양자화 레벨 이동(quantizing level shift)를 이용한 양자화를 수행한다.

2) Nonintra(inter)-MB

P-와 B-picture에서 nonintra-MB의 변환계수는 dead zone을 갖는 균등 양자화기에 의해서 DC와 AC 구분없이 양자화된다. 이때 MPEG-1과 달리 부호화 효율을 높이기 위하여 인간 특성을 이용한 그림 12(b)의 weighting matrix를 사용한다.



(a) Frame DCT block

(b) Frame DCT block

그림 11. Frame /field DCT block

8 16 19 22 26 27 29 34	16 17 18 19 20 21 22 23
16 16 22 24 27 29 34 37	17 18 19 20 21 22 23 24
19 22 26 27 29 34 34 38	18 19 20 21 22 23 24 25
22 22 26 27 29 34 37 40	19 20 21 22 23 24 26 27
22 26 27 29 32 35 40 48	20 21 22 23 25 26 27 28
26 27 29 32 35 40 48 58	21 22 23 24 26 27 28 30
26 27 29 34 38 46 56 69	22 23 24 26 27 28 30 31
27 29 35 38 46 56 69 83	23 24 25 27 28 30 31 33

(a) Intra (b) Non-intra

그림 12. Quantization weighting matrix

8. 부호화 모드

MPEG 부호화 기법에서 부호화 모드는 picture 및 MB 단위로 결정한다. 결정된 부호화 모드에 따라 각 MB는 적응적으로 부호화되어 전체적인 데이터량 및 화질을 조정하는데, 부호화 모드를 결정하는 과정은 다음과 같다.

1) Picture별 부호화 모드

MPEG-2는 GOP 단위로 부호화하는데 하나의 GOP에는 I(=1)개의 I-picture와 P(=(N-1)/M)개의 P-picture, 그리고 B(=(N-1-P))개의 B-picture로 구성된다. 즉, 각 picture별 부호화 모드는 GOP 구조에 따라서 자동적으로 결정된다. N=15, M=3으로 GOP의 변수를 결정할 경우 그림 6에서와 같이 picture별 부호화 모드가 결정된다. MPEG-2 TM에서는 CCIR 601-525 영상은 N=15, M=3이며, CCIR 601-625 영상은 N=12, M=3으로 한다.

한편 부호화 지연을 줄이기 위하여 frame-picture에서 M=1, field-picture에서 M=1 또는 M=2으로 하는 방안과 I-picture를 없애고 매 P-picture마다 intra slice를 두는 방안 등의 '저지연 부호화 모드'에 대한 표준화 연구도 병행되고 있다.

2) MB별 부호화 모드

각 MB에 대한 부호화 모드는 picture별로 다음의 순서에 의해서 결정된다.

(1) I-picture

- a. Compatible prediction
- b. Field/frame DCT coding
- c. Modified quantizer

(2) P-picture

- a. Prediction mode and MC/no MC
- b. Compatible prediction
- c. Intra/Inter

- d. Modified quantizer
- e. Field/frame DCT coding
- f. Coded/not Coded

(3) B-picture

- a. Interpolative/Forward/Backward and prediction mode
- b. Compatible prediction
- c. Intra/Inter
- d. Modified quantizer
- e. Field/frame DCT coding
- f. Coded/not Coded

결정된 부호화 모드에 대한 정보는 매 MB별로 수신 단으로 전송해야 하는데 compatible 예측, 그리고 field/frame DCT coding에 대한 정보는 각각 별도의 flag 신호를 사용하여 전송하며, 나머지 정보는 MB type 신호에 실어서 허프만 부호화하여 전송한다.

9. 허프만 엔트로피 부호화(Huffman Entropy Coding)

허프만 엔트로피 부호화는 존재 확률 빈도가 높은 부호들에 대해서는 부호당 적은 비트를 할당하고 존재 확률 빈도가 낮은 부호들에 대해서는 부호당 많은 비트를 할당하여 원 신호에서의 부호들을 각 부호의 길이가 다른 가변길이를 갖는 uniquely decodable code로 부호화함으로써 평균 전송 데이터율을 감축하는 것을 목적으로 한다. MPEG-2에서는 양자화된 DCT 변환계수, 이동벡터의 차신호, 그리고 MB에 관련된 정보를 허프만 부호화하여 전송한다.

1) DCT 변환계수의 허프만 부호화

DCT와 양자화 과정을 거친 영상신호는 그림 6(f)와 같은 zig-zag scanning 순서에 의해서 허프만 부호화된다. 마지막 non-zero 계수까지만 부호화한 후 EOB(end of block) code를 사용하여 block의 종료를 나타낸다. 이때 intraframe 부호화된 MB의 DC(intra DC 계수)와 기타 성분을 구분하여 다음과 같은 모델에 근거하여 부호화 된다.

◦ Intra DC 계수: 1차원 예측 부호화(predictive coding)

◦ 기타 계수: 연속 부호화(sequential coding)

2) 이동 벡터의 허프만 부호화

바로 직전에 전송된 동일한 type의 이동벡터를 예측기로 사용하는 DPCM을 수행한 후 이동벡터의 차를 허프만 부호화 한다. 이때 slice에서 최초의 MB에서와 바로 이전 MB가 intra MB인 경우 예측된 이동벡터는

'0'의 값을 갖는다. 성분별로 독립적으로 부호화하며, B-picture인 경우 실제 사용되는 이동벡터만을 부호화한다.

3) MB 정보에 대한 허프만 부호화

MB slice 내에서 MB의 위치정보 및 MB의 부호화 모드, 그리고 MB내에서의 block들의 부호화 정보에 대해서 허프만 부호화를 수행한다.

10. 전송율 조정(Rate Control)

엔트로피 부호화기를 거친 데이터와 기타 control 신호들은 multiplexer에서 적절한 형태로 조합되어 수신단으로 전송된다. 그러나 부호화기에서 발생하는 데이터량은 부호화기에 입력되는 영상의 특성에 따라 picture간 및 picture내의 MB별로 불규칙하게 발생한다. 이와 같이 불규칙하게 발생하는 데이터를 고화질을 유지하며 일정한 전송속도를 갖는 전송로를 통해서 전송하기 위해서는 그림 13과 같은 전송율 조절이 필요하다. TM1에서는 MB 단위로 전송 버퍼(buffer)와 부호화될 신호의 복잡도(activity 등)를 기준으로 전송율을 조정하는 feedback과 feed-forward 방식을 결합한 방식을 사용한다. 전체적으로 다음과 같은 3단계의 과정으로 구성되어 있다.

3) Step 3(adaptive quantization:feedback control)

각 MB에서 실제 양자화 과정에서 이용하는 M-quant(modified quantizer)를 계산하는 과정으로 영상내의 activity를 기준하여 step2 과정에서 구한 기준 양자화 계단으로 부터 M-quant를 계산한다.

IV. 결 언

이상에서 MPEG-2 표준 부호화 기법의 위치, 표준안의 기본요건 및 표준 영상압축 기법의 개요를 고찰하였다.

MPEG 표준 기법은 일반적인 부호화 기법(generic coding)으로서 그 적용대상의 폭이 매우 넓다. 여기에는 최근에 새로운 정보 산업 분야로 크게 거론되고 있는 멀티미디어 입출력 기기분야와 컴팩트 디스크, 디지털 VTR 등 가전 산업 분야가 있다. 또 디지털 위성 TV 방송, CATV, 디지털 오디오 방송, 디지털 EDTV 및 HDTV등 각종 고품질 방송 분배 분야가 있으며, B-ISDN과 관련된 영상 database 서비스, 원격 감시, 개인통신등의 통신 산업 분야가 있다.

그러나 이 표준기법은 최소한도로 필요한 기본적인 부호화 기법만을 명시한 것이므로 각 적용대상에 따라서 변형 보완되어야 할 내용이 많다. 따라서 표준안이 완성된 후에도 이것을 응용하기 위해서는 추가적인 응용 연구가 많이 필요하게 된다. 그러므로 우리나라의 학계와 산업계에서도 표준안이 연구 정리되는 과정에서 국내외의 관련 표준화 연구 활동에 적극 참여함으로써 새로운 정보 산업의 발전 단계에 늦지 않도록 해야 할 것이다.

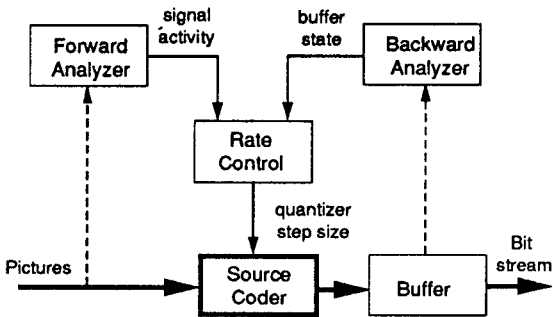


그림 13. Rate control

1) Step 1(target bit allocation)

임의의 한 frame을 부호화 할 때 발생할 비트량(bits)을 사전에 추정해 내는 과정으로 GOP내에서 picture들 간의 관계를 이용한다.

2) Step 2(rate control : feedback control)

가상버퍼(virtual buffer) 개념에 의해서 각 MB에서 전송율을 조정하기 위해서 사용할 기준 양자화 계단(reference quantization step)을 계산하는 과정이다.

부 록 : Press Release for MPEG-1 CD11172

The ISO/IEC WG11 of JTC1/SC29(MPEG) met in Kurihama, Japan, in the period of November 18th to 26th. In the first week of the meeting a Committee Draft(CD) of the MPEG standard was finalised. The techniques developed by the MPEG Committee will enable many applications requiring digitally compressed video

and sound. The storage media targeted by MPEG include CD-ROM, DAT, and computer disks and it is expected that MPEG-based technologies will eventually be used in a variety of communication channels such as ISDN and local area networks and even in broadcasting applications.

At the rate of 1.2 Mbits per second, good quality pictures have been demonstrated at 24, 25 and 30 frames per second, and at a spatial resolution of 360 samples per line. This resolution is consistent with the resolution of consumer grade television. To code stereo sound of compact disc quality, a rate of approximately 0.2 Mbits per second is required, resulting in a total rate of 1.4 Mbits per second. Such a rate could permit numerous applications including video and associated audio on compact disc.

The Committee Draft "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s" consists of three parts: System, Video and Audio. The System part (11172-1) deals with synchronisation and multiplexing of audio-visual information, while the Video (11172-2) and Audio part (11172-3) address the video and the audio compression techniques respectively.

System

The MPEG-System committee completed and approved for release the technical specification for combining a plurality of coded audio and video streams into a single data stream. The specification provides fully synchronised audio and video and facilitates the storage in and the possible further transmission of the combined information through a variety of digital media.

The systems coding includes necessary and sufficient information in the bit stream to provide the system-level functions of synchronisation of decoded audio and video, initial and continuous management of coded data buffers to prevent overflow and underflow, random access start-up, and absolute time identification. The coding layer specifies a multiplex data format the allows

multiplexing of multiple simultaneous audio and video streams as well as privately defined data streams.

The basic principle of MPEG System coding is the use of time stamps which specify the decoding and display time of audio and video and the time of reception of the multiplexed coded data at a decoder, all in terms of a single 90 KHz system clock. This method allows a great deal of flexibility in such areas as decoder design, the number of streams, multiplex packet lengths, video picture rates, audio sample rates, coded data rates, digital storage medium or network performance. It also provides flexibility in selecting which entity is the master time base, while guaranteeing that synchronisation and buffer management are maintained. Variable data rate operation is supported. A reference model of a decoder system is specified which provides limits for the ranges of parameters available to encoders and provides requirements for decoders.

Some optional sets of constraints provide a framework for common industry acceptance of certain key parameters for use by decoder designers and information providers. While the MPEG Systems specification is included in the current work item of MPEG, it is designed for compatibility with future extensions to audio, video and hypermedia coding and a wide variety of bitrates.

Video

Dozens of algorithmic approaches were carefully reviewed over a period of 3 years refining and perfecting this video compression algorithm. While the MPEG compression algorithm is optimised for bitrates of about 1.5 Mbit/s, it can perform very effectively over a wide range of bitrates and picture resolution. The video standard does not recommend a particular way of encoding pictures and much flexibility is given to implementers of the standard to use the MPEG syntax to optimise the visual quality and access options. The color resolution is given particularly high attention so as to support computer applications, games and ani-

mation.

The compression techniques developed in MPEG rely on the discrete cosine transform(DCT) for spatial redundancy reduction and motion compensated inter-frame coding to take into account the high temporal correlation of video signals by using information from both the past and the future. The statistics of the resulting information also can be exploited to further reduce the bitrate through the use of special codes known as Huffman codes. While the discrete cosine transform has been widely used for many years, the techniques developed within MPEG also exploit the characteristics of the human visual system to optimise the perceived image quality. Eventual coding impairments are concentrated in frequencies and regions of the picture where they are perceptually minimal.

Audio

The audio coding experts of MPEG finalised an audio coding algorithm after having reviewed and tested many approaches over the last three years. The resulting digital audio bitrate reduction technique supports several bitrates covering a range from intermediate to compact disc quality. This latter quality can be obtained at a total bitrate of 256 Kbit/s for a stereophonic program.

Depending on the applications, three layers of the coding system with increasing complexity and performance can be used. In all three layers the time domain input audio signal is converted into a frequency representation. In Layers I and II a filterbank creates 32 subband representations of the input audio stream which are then quantised and coded under the control of a psychoacoustic model from which a blockwise adaptive bit allocation is derived. With respect to Layer I, Layer II introduces further compression by redundancy and irrelevance removal on the scalefactors and more precise quantisation. In Layer III additional frequency resolution is provided by the use of a hybrid filterbank. Every subband is thereby further split into higher resolution frequency lines

by a linear transform that operates on 18 subband samples in each subband. The frequency lines are again quantised and coded under the control of psychoacoustic model. In Layer III, nonuniform quantisation, adaptive segmentation and entropy coding of the quantised values are employed for a better coding efficiency.


The range of bitrates(total rate for both channels)provided by the standard is between 64 kbit/s and 448 Kbit/s. The standard also supports coding starting at 32 kbit/s for a single channel. In all layers a joint stereo mode that exploits stereophonic irrelevance or stereophonic redundancy can be used as an option to improve the subjective quality.

參 考 文 獻

- [1] ISO-IEC /JTC1 /SC20 /WG11, MPEG92 /-, "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s," CD 11172, revision 2, Mar. 1992.
- [2] ISO-IEC /JTC1 /SC29 /WG11, MPEG92 /229 (revised), "Information on requirements for MPEG-2 Video," Jul. 1992.
- [3] ISO-IEC /JTC1 /SC29 /WG11, MPEG92 /086, "MPEG-2 Test Model 0," Mar. 1992.
- [4] ISO-IEC /JTC1 /SC29 /WG11, MPEG92 /160, "MPEG-2 Test Model 1," May 1992.
- [5] ISO-IEC /JTC1 /SC29 /WG11, MPEG92 /245, "MPEG-2 Test Model 2," Jul. 1992.
- [6] ISO-IEC /JTC1 /SC29 /WG11, MPEG92 /100 (Matsushita), "Field-time adjusted MC for frame-base coding," Mar. 1992.
- [7] ISO-IEC /JTC1 /SC29 /WG11, MPEG92 /249 (Matsushita), "Simplification of FAMC," Jun. 1992.
- [8] ISO-IEC /JTC1 /SC29 /WG11, MPEG92 /246 (KDD), "Simulation results of TM1 core experiment," Jul. 1992.
- [9] ISO-IEC /JTC1 /SC29 /WG11, MPEG92 /259

(Toshiba Corporation), "Unification of FAMC and dual field prediction," Jul. 1992.

[10] ISO-IEC/JTC1/SC29/WG11, MPEG92/- ,

"Reprot of the Frame/Field prediction Ad-hoc group at the Tarrytown meeting," Sep. 1992. 

筆者紹介



金 在 均

1938年 9月 17日生

1962年 한국항공대학 응용전자과 졸업

1967年 서울대학교 대학원 전자공학과(석사)

1972年 미국 남가주대학교 대학원(공학박사)

1972年 ~ 1973年

미국 NASA GSFC 연구원

1984年 ~ 1985年

과학기술처 전기전자 연구조정관

1973年 ~ 현재

한국과학기술원 전기 및 전자공학과 교수



金 旻 圭

1961年 6月 10日生

1987年 고려대학교 전자공학과 졸업

1989年 한국과학기술원 전기 및 전자공학과(석사)

1992年 12月 현재 한국과학기술원 전기 및 전자공학과(박사과정)