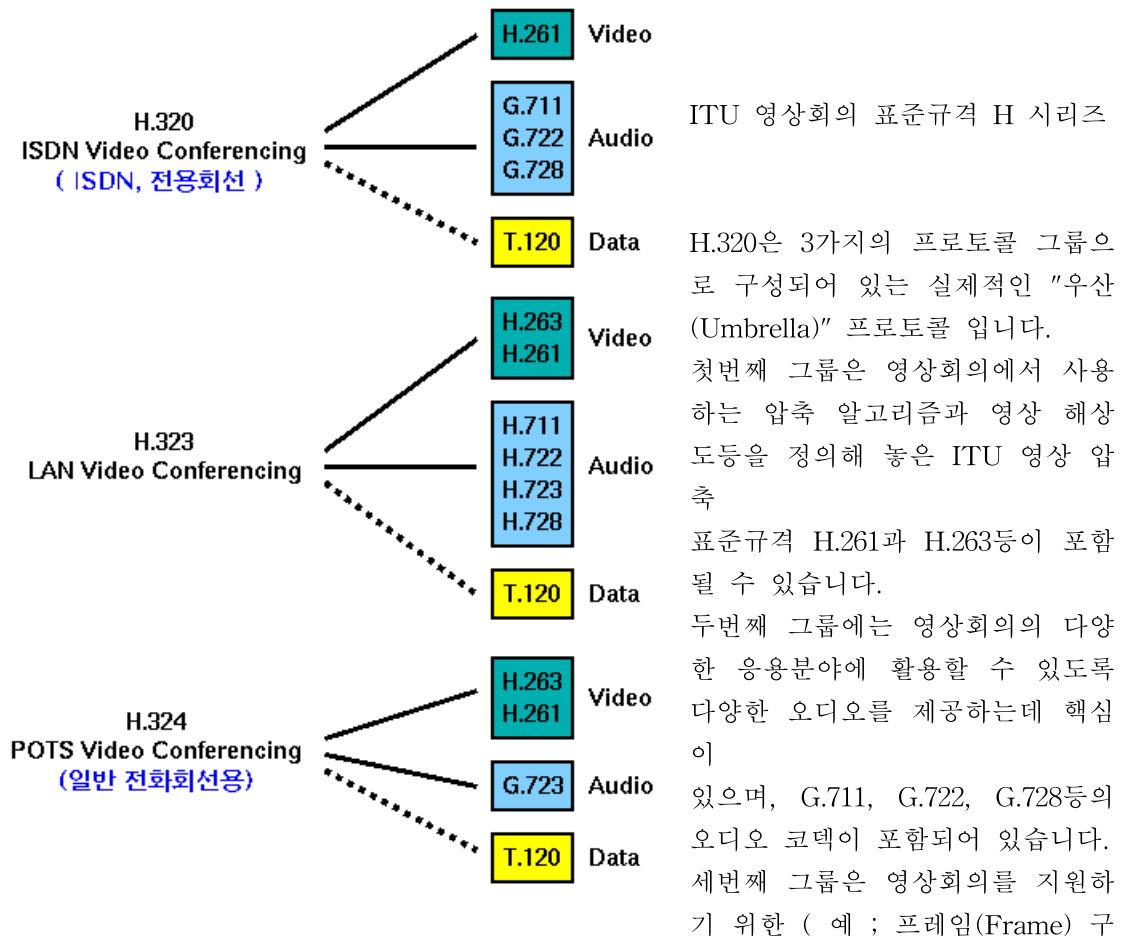


1. H.320 영상의회의 프로토콜의 개요

H.320 영상의회의 프로토콜은 ISDN과 협대역 전송 매체 (T1/E1 전용회선등)를 통한 영상의회의 표준 규격입니다. H.320 프로토콜은 ITU (International Telecommunication Union)에서 지정한 방식으로 멀티미디어 정보통신에서 사용하는 가장 핵심적인 기술입니다. H.320 프로토콜이외에 다른 "H" 프로토콜은 다음과 같습니다.

- T.120 - 실시간 문서회의용 프로토콜
- H.323 - LAN (근거리 네트워크) 을 위한 영상의회의 프로토콜
- H.324 - 일반 전화회선 및 모뎀 연결을 통한 고품질 영상,오디오 압축

아래의 그림은 영상의회의 표준규격(프로토콜)의 "H"시리즈를 구성하는 프로토콜들 입니다.



ITU 영상의회의 표준규격 H 시리즈

H.320은 3가지의 프로토콜 그룹으로 구성되어 있는 실제적인 "우산(Umbrella)" 프로토콜 입니다. 첫번째 그룹은 영상의회의에서 사용하는 압축 알고리즘과 영상 해상도등을 정의해 놓은 ITU 영상 압축 표준규격 H.261과 H.263등이 포함될 수 있습니다. 두번째 그룹에는 영상의회의의 다양한 응용분야에 활용할 수 있도록 다양한 오디오를 제공하는데 핵심이 있으며, G.711, G.722, G.728등의 오디오 코덱이 포함되어 있습니다. 세번째 그룹은 영상의회의를 지원하기 위한 (예 ; 프레임(Frame) 구

조, 포맷, 다자간 제어(Multipoint Control) 전송 (Transmission)과 제어 (Control) 처리를 담당하고 있는 프로토콜로 구성되어 있습니다.

또한 H.221, H.231 사양을 포함하고 있기도 합니다.

2. H.320의 영상, 음성 코딩 기술

2.1 H.261 영상 압축 표준규격

H.261 은 64Kbps에서 2Mbps까지의 통신 채널을 통하여 사용할 수 있도록 설계된 영상 압축 표준 규격

입니다. 이 표준은 또한 p가 1에서 30까지 의 범위 (예 ; 1BRI ISDN B-채널에 의하여 제공 되는 복수개의

대역폭) 내에 있는 px64로 알려져 있습니다.

H.261은 Intraframe 공간적 엔코딩 (Encoding)과 Interframe 시간적 엔코딩 두가지 방식을 사용하고 있습니다. Intraframe 엔코딩은 DCT ((Discrete Cosine Transform, 분산 코사인 변환 방식) 기술을 사용합니

다. 또한 이것은 8x8 블록내에서 DCT 처리, 양자화 및 Run-Length 엔코딩을 하는 JPEG과 흡사합니다.

Interframe 엔코딩은 이전 프레임(Frame)과 현재 프레임(Frame)간의 차이 를 점검하며, 만약 차이가 일정

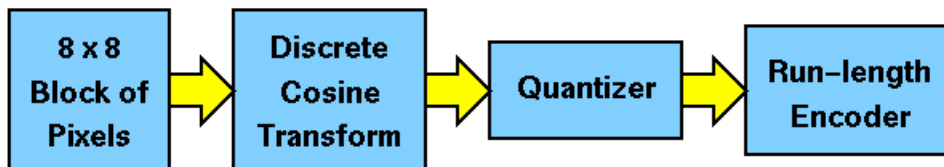
기준 이하라면 아무런 데이터도 보내지 않 습니다.

그 차이가 적당하다면, 블록간의 차이는 Intraframe 엔코딩에서 정한 동일한 방식을 사용하여 엔코딩합니다. H.261은 또한 일정한 Bit Rate를 유지하기 위하여 이들 두가지 엔코딩 방식은 양자화 단계(절차)를 유동적으

로 조정 합니다. Bit Rate를 줄이기 위해서는 양자화 단계를 증가시킵니다.

(발생되 는 정보를 줄임) 증가된 양자화 단계는 떨어지는 화질을 야기시키며, 반대 로 양자화 단계를 줄이면

화질은 향상됩니다. 아래 그림은 H.261에서 사용하는 코딩 절차를 나타냅니다.



H.261 코

딩 절차

H.261은 또한 두가지의 화상 형식을 정의하고 있으며, Common Intermediate Format 또는 CIF 그리고 Quarter CIF 또는 QCIF등으로 구분합니다.

QCIF가 176x144의 해상도 (CIF 해상도의 1/4)인 반면에 CIF는 352x288 의 해상도를 지원합니다.

QCIF는 전형적으로 192Kbps 이하의 저 대역폭 이용 시 사용을 합니다. QCIF/CIF 이미지는 YCbCr (Y=Luma, Cb/Cr=Chrominance) 컬러 엔코딩 방식을 사용하여 코딩됩니다.

2.2 오디오 코덱 (G.711, G.722, G.728)

H.320은 영상회의의 활용 분야에 따라 다양한 오디오를 처리할 수 있도록 세가지 오디오 코덱 표준 (G.711, G.722, G.728)을 포함하고 있습니다. G.711은 64Kbps에서 3KHz의 전화급 오디오 품질을 제공하기 위하여

PCM 오디오 엔코딩과 미국, 유럽에서 주로 이용하는 U-law 또는 A-law 방식을 사용합니다.

G.722는 G.711과 흡사하지만 64Kbps 대역폭에서 7KHz의 스테레오 오디오 품질을 지원합니다.

G.728은 단지 16Kbps 대역폭 내에서 거의 전화급에 가까운 3KHz의 오디오 품질을 내기 위하여 Code Excited Linear Prediction (CELP) 방식을 사용합니다.

펄스코드변조 (Pulse Code Modulation, PCM) 방식은 샘플들을 미국 방식인 U-law 또는 유럽 방식인

A-law 양자화 방식을 이용하여 한정된 재구성 세트 중의 하나로 양자화를 시켜주는 과형 코딩 방식입니다.

G.711 표준은 전화 대화 코딩을 위한 표준 방식으로 8bit PCM 을 정의하고 있습니다.

PCM은 적은 Delay 성질을 가지고 있기 때문에 전화 회사에서도 많이 사용하고 있는 방식입니다.

Code Excited Linear Prediction (CELP) 방식은 복잡한 음성 코딩 구조를 가지고 있습니다. CELP는 대화 음성과 음성 트랙의 분석 모델을 비교하며, 원 음성과 모델 사이의 에러를 계산해 냅니다.

그리고는 모델의 특성 매개변수와 에러 내용을 전송합니다. CELP 엔코더는 고품질의 음성을 재생할 수 있습니다.

3. H.320의 프레임(Framing)과 제어(Control) 프로토콜

3.1 H.221 프레임 (Framing) 프로토콜

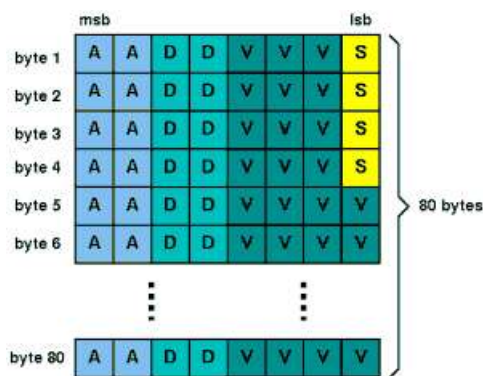
H.221 표준은 64Kbps에서 1,920Mbps 채널을 통한 통신을 위하여 영상/음성 통신 서비스에

서 사용하는 프레임 구조 (Frame Structure)를 정의하고 있습니다. H.221은 길이 80 Byte의 고정 길이를 가진 Frame의

데이터로 구성되어 있습니다. 각 Byte는 H.320에서 각기 다른 프로토콜로 생성된 멀티플렉스된 오디오,

비디오 그리고 인밴드 데이터로 구성됩니다.

아래의 그림은 H.221 Frame의 첫번째 몇 Byte를 보여 줍니다. Bit 위치들은 A, D, V로 표기됩니다.



그리고 S bit는 오디오, 인밴드 데이터, 비디오, 그리고 신호 배열정보들을 포함하고 있습니다.

H.221 프레임 (Framing)의 구조

첫번째 4 Byte의 lsb내에 있는 S bit를 보면, 우리는 이것을 서비스(Service) bit 또는 Frame Alignment Signal (FAS) bit라고 부르며, Frame

의 시작을 나타냅니다.

Frame 시리즈의 수신측에서는 각 Framed의 시작으로 동기를 맞추기 위하여 이들 Bit를 이용하고 있습니다.

일정하게 S bit들을 점검하는 것은 배열 이상 Frame등 bit-Slip을 막을 수가 있습니다

각 H.221 Frame의 첫번째 12 Byte는 S bit를 포함하고 있습니다.

그림에서 볼 수 있는 바와 같이 대략적으로 채널의 1/4이 오디오를 위한 채널이며, 1/4이 인밴드 데이터를 위한 채널, 그리고 대략 1/2 정도의 채널 (Frame 동기를 위한 bit들보다 적음)은 비디오 전송을 위하여 할당합니다.

3.2 H.231 멀티포인트 제어 프로토콜

H.231은 2명이상의 사용자들이 영상회의를 동시에 할 수 있는 다자간 회의 제어용 표준 규격을 정의하고 있습니다. 또한 데이터 암호화 (Encryption), 영상회의 내에서 그래픽의 사용 방식등을 정의하고 있습니다.

다자간 회의 (Multipoint Videoconferencing)은 MCU (Multipoint Control Unit)라는 특수한 브릿지형의 장비를

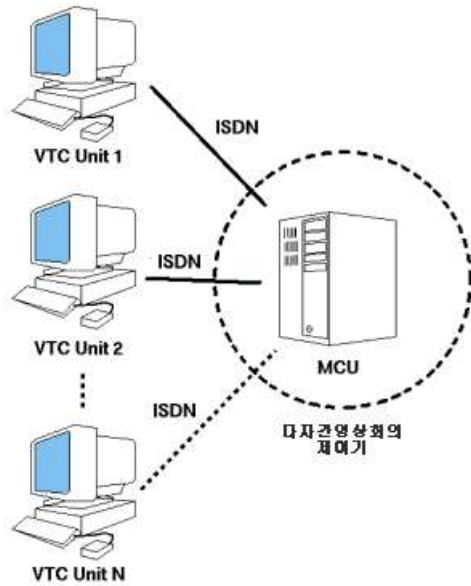
사용하여 이루어 집니다. 두명이상의 사용자가 동시에 다자간 회의를 하기를 원한다면, 사전

에 MCU로 모두를 연결해야만 합니다.

MCU는 H.242, H.243 프로토콜을 사용하여 각 참여자들의 영상회의단말기들간에 정보를 교환시켜 줍니다.

이 통신 동안에 MCU는 각 영상회의단말기의 영상 포맷(CIF 또는 QCIF), 오디오 성능, 데이터 전송 속도등에 관한 정보들을 모으게 됩니다.

일단 MCU가 각 영상회의단말기들의 모든 정보들을 데이터베이스화 한다면, 모든 사람들이 참여할 수 있도록 가장 낮은 공통 사양에서 회의 기준을 지정하게 됩니다. 전형적인 MCU의 그림은 아래와 같습니다.



다자간회의제어장치 (MCU)의 전형적인 구성 방식

MCU는 일반적으로 사무실 내 또는 MCU 서비스 제공자의 네트워크 상에 위치합니다. 요즘은 다른 나라에서는 MCU 전문 서비스 회사가 있어서, MCU를 개별적으로 설치하지 않고 단지 MCU 서비스 회사로 접속하여 다자간회의서비스

를 제공받을 수 있는 시스템이 보편화되어 있습니다.

H.263 비디오 부호화

1. 서론

디지털 비디오 응용에 대한 관심이 증가됨에 따라 H.261, MPEG-1 그리고 MPEG-2와 같은 표준 비디오 압축 부호화 방식들이 제정되었다. 그러나 이 방식들은 전화망이나 무선망과 같이 좁은 대역폭을 갖는 환경에서는 좋은 성능을 제공하지 못한다. 따라서 ITU-T에서 Video Coding for Low Bit Rate Communications라는 명칭하에 H.263 표준[1]을 제정했다. H.263은 원래 64Kbits/s이하의 저 전송률 영상통신을 위한 영상 부호화 표준으로 제시되었다. 그러나 현재는 저 전송률뿐만 아니라 폭 넓은 전송률에도 적용함으로써 경우에 따라서는 H.261과 MPEG 부호화 방식과도 경쟁할 수 있는 기법으로 인정되고 있다.

H.263의 부호화 기법은 H.261 기법을 기반으로 하지만 H.261의 비트열 구조의 일부분이 선택사항이 되었고 약간의 추가적인 연산에 의해 저 전송률에서 보다 좋은 화질과 향상된 에러 복원을 제공할 수 있다. 따라서 ITU-T H.324[2], H.320[3], H.310[4]과 같은 표준 비디오 단말 등에 적용되고 있다. 또한 H.263은 비제한적인 움직임 벡터(unrestricted motion vector), 조건부 산술부호화(syntax-based arithmetic coding), 고급예측(advanced prediction), 그리고 순방향과 양방향 예측을 사용하는PB-프레임(predictive and bi-directional predictive frame)과 같은 4개의 선택적 표준기법들을 사용함으로써 보다 향상된 압축 성능을 제공한다.

2. H.263 표준 부호화 기법

H.263은 H.261 또는 MPEG처럼 움직임 보상을 이용한 예측 부호화와 DCT 변환에 의하여 시공간적으로 높은 상관성을 갖는 영상신호 성분들간의 상관성을 제거한 후, 양자화에 의하여 이 신호 성분들을 몇 개의 대표 값으로 표현함으로써 정보압축을 수행한다. 그리고 가변 길이 부호화(VLC: Variable Length Coding)를 사용하여 통계적 상관성을 제거함으로써 보다 높은 압축 효율을 얻는다. 그림 1은 H.263 비디오 부호기의 구성을 보여준다. H.263 부호화는 H.261의 경우와 마찬가지로 입력 영상을 화면, GOB(Group Of Block), MB(MacroBlock),블록과 같은 계층적인 구조로 분할한 후 처리한다.

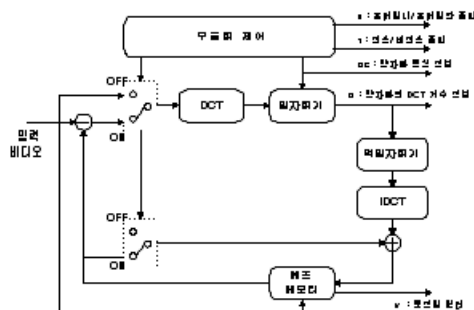


그림 1. H.263 비디오 부호기의 구성도

부호화기에 입력되는 각 영상은 $K \times 16$ 라인으로 구성되는 GOB들로 분리된다. 여기서 K 는 입력 영상 규격에 관계된 변수로 sub-QCIF, QCIF 그리고 CIF에 대해서는 $K=1$, 4CIF에 대해서는 $K=2$, 16CIF에 대해서는 $K=4$ 이다. 따라서 한

영상 당 GOB의 수는 입력 영상 규격이 sub-QCIF이면 6개, QCIF이면 9개, CIF와 4CIF 그리고 16CIF이면 18개이다. 각 GOB는 16*16인 휘도성분과 8*8인 색차성분들(CB, CR)로 구성되는 MB들로 구성되며, 각 MB들은 다시 4개의 휘도성분 블록과 2개의 색차성분 블록들로 구성된다.

예측 부호화 (Predictive Coding)

예측 부호화는 어떤 화소값을 시공간적으로 인접한 화소값으로 표현하는 것이다. H.263에서는 예측 부호화를 수행하는 화면간(Inter) 부호화 모드와 예측 부호화를 수행하지 않는 화면 내(Intra) 부호화 모드가 있다. 화면간 부호화의 경우는 그림 1에서 스위치들이 ON되고 화면 내 부호화의 경우는 스위치들이 OFF된다. 부호화 모드의 결정은 화면단위 또는 MB단위로 결정할 수 있는데, I-프레임(intra frame)인 경우는 화면단위에서 부호화 모드가 화면 내 모드로 결정되므로 이 프레임은 예측 부호화를 수행하지 않는다.

한편 P-프레임(predictive frame)인 경우에는 MB단위에서 부호화 모드를 결정할 수 있는데, 대부분의 MB들이 화면간 예측을 사용하지만 장면전환과 같이 화면간의 상관도가 낮은 부분에서는 화면 내 부호화를 사용하는 것이 보통이다.

화면간 예측에서는 그림 1에서 입력 비디오신호와 예측메모리에 저장된 예측신호와의 차이값이 취해지는데, 화면 내의 동일 위치뿐만 아니라 수평 및 수직방향으로 움직임 벡터를 이용한 움직임 보상을 수행할 수 있다. H.261에서는 움직임 보상을 위하여 MB 당 [-15,15] 화소 간격의 범위 내에서 화소단위의 정밀도를 갖는 하나의 움직임 벡터를 사용하지만, H.263에서는 [-16,15.5] 화소 간격의 범위 내에서 1/2 화소단위의 정밀도를 갖는 움직임 벡터를 사용한다. 또한 선택적 표준기법인 고급예측에서는 한 MB 당 4개의 움직임 벡터를 수용할 수 있다. PB-프레임을 사용할 경우에는 B-프레임의 예측을 위해 MB 당 추가적으로 델타 벡터(delta vector)라는 움직임 벡터를 사용한다. 한편 비제한적인 움직임 벡터를 사용하는 경우는 [-31.5, 31.5] 화소 간격의 범위 내에서 1/2 화소단위의 정밀도를 갖는 움직임 벡터를 사용한다. 하지만 부호기에서 움직임 정보의 사용 여부는 설계자의 자유이며, 움직임 벡터를 구하는 방법 역시 설계자의 자유이다.

DCT 변환과 양자화

MB 내에 있는 블록에 대하여 88 DCT(Discrete Cosine Transform)를 수행하여 블록 내에 있는 영상신호의 에너지를 저주파쪽으로 편중 시킨다. H.263에서 사용하는 DCT 및 IDCT(Inverse Discrete Cosine Transform)는 H.261과 동일하다. DCT에 의하여 저주파 성분에 모인 88 블록 내의 영상신호의 주파수 성분들을 양자화에 의하여 몇 개의 대표 값으로 표현함으로써 정보를 압축한다. H.263은 역 양자화 처리에 대해서만 정의되어 있고 양자화 처리에 대해서는 설계자가 결정할 수 있게 하였다. 즉 양자화기의 결정준위(decision level)에 대해서는 정의되지 않고, 재생준위(reconstruction level)에 대해서만 정의된다. 다만 양자화기의 종류는 화면 내 부호화를 수행하는 블록의 DCT 계수 중 DC 계수(INTRADC)에 적용되는 양자화기와 나머지 DCT 계수들에 대하여 화면 내 부호화에 적용되는 31 종류의 양

자화기, 그리고 화면간 부호화에 적용되는 31 종류의 양자화기가 사용되는 것으로 표준화되어 있다.

3. H.263의 선택적 부호화 기법

H.263에서는 앞에서 설명한 H.263의 표준기법뿐만 아니라 비제한적 움직임 벡터, 조건부 산술부호화, 고급예측 그리고 PB-프레임 모드와 같은 4가지의 선택적인 부호화 방식을 채용할 수 있다. 이 방식들은 단독적으로 사용할 수도 있고 서로 결합하여 사용할 수도 있는데, 이 방식들의 사용 여부는 화면계층의 헤더에 있는 PTYPE으로부터 알 수 있다.

이 방식들은 선택 사항이므로 이 방식들을 사용하여 영상통신을 하고자 할 경우는 상대 단말들이 이 방식들을 지원할 수 있어야 한다. 따라서 영상통신을 시작하기 전에 H.245[5]와 같은 단말들 간의 협상 절차에 의해서 이 방식들의 사용 여부를 결정한다.

비제한적 움직임 벡터 모드

H.263의 표준 부호화에서는 움직임 벡터가 부호화된 이전 영상 내에 있는 화소들을 참조해야 하고 범위도 $[-16,15.5]$ 로 제한된다. 그러나 비제한적 움직임 벡터 모드에서는 부호화된 이전 영상의 경계에 있는 화소들을 외부로 확장함으로써 외부에 있는 화소들을 참조할 수 있고, 탐색범위도 $[-31.5,31.5]$ 로 확장된다. 따라서 이 모드는 영상의 경계를 따라 움직임이 있거나, 카메라의 움직임과 같은 전역 움직임에서 두드러진 이득을 얻을 수 있다. 특히 작은 영상 규격(sub-QCIF 또는 QCIF)에서 이득을 얻는다. 비제한적 움직임 벡터 모드에서 움직임 벡터가 이전 영상의 외부에 있는 화소를 참조할 경우 이 외부 화소는 영상 경계에 있는 화소로 대치되는데, 이 경계 화소는 영상 내부의 화소만을 참조하여 얻어진 화소단위의 움직임 벡터가 가리키는 화소이다.

조건부 산술부호화 모드

이 모드에서는 블록 계층의 DCT 계수에 대한 가변길이 부호화(VLC) 방식으로 허프만 부호표를 사용하는 대신에 비트열 데이터 구조(구문구조)에 근거한 산술부호화를 사용한다. 이 방식을 사용하면 보다 적은 비트를 사용하면서도 재생된 영상의 화질을 동일하게 유지할 수 있다. 이와 같은 이득은 부호화되는 영상과 적용되는 비트율 등에 따라 차이가 있지만 보통 화면간 부호화의 경우는 평균적으로 약 3-4%, 화면 내 부호화의 경우는 평균적으로 약 10% 정도의 이득을 나타낸다.

고급 예측 모드

H.263의 표준 부호화에서는 하나의 MB에 대하여 하나의 움직임 벡터가 사용되지만, 이 모드에서는 하나의 MB에 대하여 4개의 움직임 벡터가 사용될 수 있다. 움직임 보상은 그림 2에 나타난 바와 같이 현재 블록의 움직임 벡터에 의해서 보상된 신호성분과 주위 블록에 해당되는 움직임 벡터들을 이용하여 움직임 보상된 신호성분들을 가중 합(weighted sum)함으로써 처리된다. 즉, 복원되는 현재 블록의 화소값 $P(x,y)$ 은 다음 식과 같이 현재 블록에 대한 움직임 벡터(MV)에 의해 움직임 보상된 신호 $q(i,j)$ 와 상하에 있는 블록에 해당되는 벡터들(MV_{above}, MV_{below})에 의해서 움직임 보상된 신호 $r(i,j)$ 그리고 좌우에 있는 블록에 해당되는 벡터들(MV_{left}, MV_{right})에 의해서 움직임 보상된 신호 $s(i,j)$ 에 미리 결정된 가중치 $H_k(x,y)$ 를 부여한 후 8로 나눈 값이다.

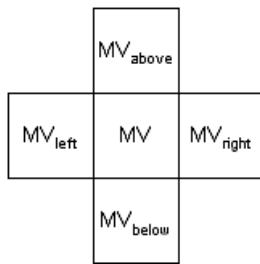


그림 2. 움직임 보상에 이용되는 움직임 벡터들

$$P(x,y) = (q(i,j)H_0(x,y) + r(i,j)H_1(x,y) + s(i,j)H_2(x,y) + 4)/8$$

여기서 (i,j) 는 1/2 화소단위로 움직임 보상된 화소의 위치이고, (x,y) 는 현재 블록의 화소 위치이다.

PB-프레임 모드

PB-프레임은 그림 3에 나타난 바와 같이 복호화된 이전 P-프레임으로부터 예측되는 P-프레임과 복호화된 이전 P-프레임과 현재 P-프레임으로부터 양방향 예측되는 B-프레임으로 구성되는데, 이 두 프레임은 마치 하나의 프레임처럼 부호화된다.

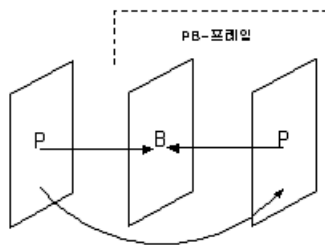


그림 3. PB-프레임 모드에서의 예측

PB-프레임 모드에서는 하나의 MB은 12개의 블록으로 구성되는데, 앞에 있는 6개의 블록은

P-블록이고, 후위에 있는 6개의 블록은 B-블록이다. PB-프레임이 화면간 부호화 형식일 때 P-프레임은 MB단위로 화면 내 부호화가 가능하지만 B-프레임은 항상 화면간 부호화만을 수행한다. 따라서 화면 내 부호화되는 MB에 대해서도 B-프레임에 있는 MB에서 사용되는 움직임 벡터 정보가 전송된다.

B-프레임의 양방향 예측을 위해서는 이전 P-프레임에 대한 순방향 움직임 벡터(MVF)와 PB-프레임에 속한 P-프레임에 대한 역방향 움직임 벡터(MVB)가 필요하다. 이 움직임 벡터들은 P-블록에 대한 움직임 벡터인 MV와 B-프레임의 예측을 위해 MB 당 추가적으로 전송되는 델타 벡터(delta vector) MVD를 이용하여 구해진다. 그림 4는 순방향 움직임 벡터(MVF)와 역방향 움직임 벡터(MVB)의 예를 보여준다.

B-프레임에 속한 블록에 대한 움직임 보상은 역방향 움직임 벡터가 가리키는 위치가 현재 MB 영역에 해당하는 부분은 양방향 예측에 의해서 움직임 보상을 수행하고, 이 외의 영역은 순방향 움직임 벡터만을 사용하는 순방향 예측을 사용한다.

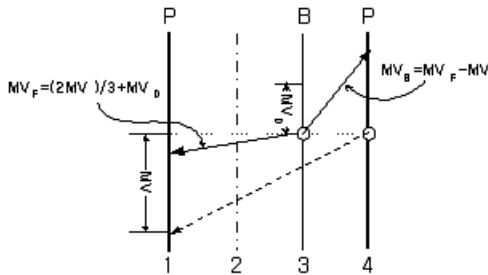


그림 4. PB-프레임 모드에서의 움직임 벡터

4. 결론

지금까지 비디오 부호화 표준인 H.263에서 사용되는 표준 부호화 기법과 4종류의 선택적 부호화 기법들에 대해서 살펴보았다. H.263은 기존 표준 비디오 부호화 방식에 비하여 저 전송률에서 보다 향상된 압축 성능을 제공하므로 향후 IMT-2000등에서 영상통신 서비스를 제공할 수 있는 부호화 방식으로 고려된다. 뿐만 아니라 현재 표준화가 진행 중인 H.263+[6], H.263++, H.26L 등은 H.263에 여러 가지 발전된 선택적 표준기법들을 첨가하여 보다 향상된 압축 성능, 계층적 부호화 등과 같은 다양한 기능을 제공할 뿐만 아니라 채널 전송 오류에 대한 강인성 증대와 QoS (Quality of Service)가 보장되지 않는 망에서의 효율적인 영상 전송을 가능하도록 연구되고 있다.

[참고 문헌]

[1] ITU-T Recommendation H.263 (March 1996): Video Coding for Low Bit Rate Communication.

[2] ITU-T recommendation H.324 (1995): Terminal for low bitrate multimedia communication.

[3] ITU-T recommendation H.320 (1993): Narrow-band ISDN visual telephone systems and terminal equipment.

[4] Draft ITU-T recommendation H.310 (January 1996): Broadband audiovisual communication systems and terminals.

[5] ITU-T Recommendation H.245 (1998): Control Protocol for Multimedia Communication.

[6] Draft ITU-T Recommendation H.263 Version 2 (Sept. 1998): Video Coding for Low Bit Rate Communication.