

Full-Digital HDTV 에 있어서의 영상 신호 데이터 압축

鄭濟昌, 趙在汶

三星電子(株) 家電 A/V 研究所

I. 서 론

최근들어 디지털 영상신호를 압축하여 저장하거나 전송하는 시스템이 실용화되면서 보다 능률적인 영상부호화 방법에 대한 관심이 고조되고 있다. 응용분야에 따라 국제적으로 표준화 움직임이 활발하여, 영상전하나 영상회의를 위한 부호화 방법은 국제전신전화자문위원회(CCITT)의 H.261으로 표준화 되었고^[1], 정지화의 부호화는 국제표준화기구(ISO)와 국제전기표준회의(IEC), CCITT가 합동으로 JPEG 표준을 마련하였으며^[2], 동화에 대해서는 디지털 저장매체에 적용할 목적으로 1.5Mbps급의 MPEG I이 ISO와 IEC에 의해 마련되었다^[3]. 보다 뛰어난 화질을 요구하는 응용분야를 고려하여 10Mbps 이상까지 확장하여 사용하는 MPEG II 표준이 활발히 진행되고 있으며^[4], CCITT에서는 미래의 광대역 종합정보통신망(B-ISDN)을 이용하는 영상전송을 위하여 ATM 프로토콜에 기초한 부호화 방법을 연구중이고^[5], CCITT와 CCIR이 합동으로 조직한 연구그룹인 CMTT에서는 TV 및 HDTV(high definition television, 고선명 TV) 신호의 디지털전송을 위한 표준을 마련중이다^[6]. 또한 CD-I와 DVI로 대표되는 멀티미디어, 디지털 VCR, 전자 카메라 등에도 영상부호화가 핵심요소로 작용하고 있다.

한편, 미국은 미래의 TV인 HDTV의 표준방식을 정하기 위해 현재 5가지의 방식이 제안되어 ATTC(Advanced Television Test Center)에서 시험중이며, 이 시험 결과를 바탕으로 FCC(Federal Communication Committee)는 93년말에 미국의 HDTV 규격을 결정할 예정이다. 5가지 방식중 일본 NHK가 제안한 narrow-MUSE만이 아날로그 전송방식이고 다른 4개 방식-

ATVA(GI/MIT)가 제안한 두 방식인 Digicipher(ATVA-I)^[7]와 CCDC(ATVA-P)^[8], Zenith/AT&T의 DSC-HDTV^[9], ATRC(Philips, Thomson, DSRC, NBC, CLI)의 ADTV^[10] - 은 데이터 압축과 전송이 모두 디지털인 full-digital 방식이다. 이미 방송이 시작된 일본의 MUSE나 실험방송중인 유럽의 HD-MAC이 subsampling과 TDM(time division multiplexing)에 의해 영상신호의 대역을 압축하고 이를 아날로그 형태로 전송하기 때문에 채널에서의 잡음에 의해 화질의 열화가 커지는 반면, full-digital 방식에 있어서는 수신측에서의 SNR이 어떤 임계치 이상으로 유지되지만 하면 채널 에러가 발생하더라도 에러 정정을 하여 송신측의 화질이 수신측에 거의 완전하게 전달된다. Full-digital HDTV의 신호처리는 크게 신호원 부호화부와 채널 부호화부로 나뉘는데, 신호원 부호화부에서는 주어진 채널(지상방송, 위성방송, 케이블, 광섬유등)의 데이터 전송 속도에 맞추어 영상과 음성의 데이터를 압축하고, 채널 부호화부에서는 에러정정부호(ECC : error correcting code) 및 디지털 변조가 이루어진다. 미국의 경우 FCC에 의해 6MHz의 대역폭을 갖는 기존 NTSC의 터부채널을 이용하여 simulcast 형식으로 HDTV 방송을 하기로 규정되어 있어, 영상신호는 70:1-100:1, 음성신호는 6:1 정도의 데이터 압축이 필요하며, ECC에 의해 약 20%의 데이터량 증가를 가져온다.

본고에서는 full-digital HDTV를 위한 데이터 압축방법과 관련하여, 미국의 4개 full-digital HDTV 제안방식과 CCITT, CCIR, ISO등에서 채택하고 있는 표준 영상압축방법, 그리고 최근 연구되고 있는 여러가지 영상압축 기법들을 비교 검토하고, 나아가 본 연구실에서 개발한 full-digital HDTV용 영상 신호 데이터 압축 방

법을 소개하고자 한다. HDTV와 관련하여, 영상압축 이외의 부분에 대해서는 본 특집에 실린 다른 논문들이 주제별로 자세히 다루고 있다.

II. 영상신호의 데이터 압축

이미 40여년전, 채널의 대역폭과 SNR이 주어졌을 때 이 채널을 통해 임의의 낮은 에러율로 전송할 수 있는 데이터량, 즉 채널 용량을 구하는 식이 Shannon에 의해 발견되었다¹¹⁾. 6MHz의 대역폭을 갖는 현재의 NTSC 지상방송 채널에서는 이 채널용량이 20Mbps 안팎이 되는 바, 대략 영상신호가 17Mbps, 음성신호가 0.5Mbps를 점유하고, 남은 부분을 기타 보조 데이터, 문자방송, 에러 정정 부호등이 차지하게 된다. 따라서 단순 PCM시에 1.2Gbps에 이르는 영상신호의 데이터량을 70:1 이상으로 압축해서 전송해야 하는데, 압축시 영상정보의 손실을 최소화하기 위해서는 영상 신호에 내재하는 중복성(redundancy)-즉, 색 성분간의 spectral redundancy, 화면과 화면간 시간축 방향의 temporal redundancy, 그리고 화면내 공간의 이웃하는 화소간의 spatial redundancy¹²⁾-을 효과적으로 제거해야 한다.

그림 1은 full-digital HDTV 뿐 아니라 여러 응용 분야에 널리 쓰이는 일반적인 영상신호 데이터 압축 방법인 움직임 보상(motion-compensated, MC) 하이브리드 부호화의 블록도이다. 표 1은 미국의 HDTV 제안 방식들 간의 비교표이다. 이하 신호의 흐름을 따라 처리 과정을 기술하기로 한다.

1. Baseband 신호의 형식

HDTV는 현행 TV보다 해상도가 4-5배 정도 되는데, 방식마다 주사선/초당 프레임수/interlace율에 있어서 차이가 있다. 일본 방식은 현행 NTSC(525/30/2:1) 및 PAL, SECAM(625/25/2:1)과의 상호 변환이 용이하도록 하기 위해 1125/30/2:1을 채택하고 있다. 유럽의 HD-MAC은 1250/25/2:1로서 PAL, SECAM과의 상호 변환을 고려하고 있다. 미국에서는 NTSC와의 상호 변환을 고려하여, Digicipher와 ADTV가 채택한 1050/30/2:1(interlaced)과 CCDC와 DSC-HDTV가 채택한 787.5/60/1:1(progressive)의 두 방식이 맞서고 있다. Interlaced 주사방식은 NTSC 보다 주사선 수를 2배로 하여 공간 해상도를 수평 수직 각각 2배로 늘

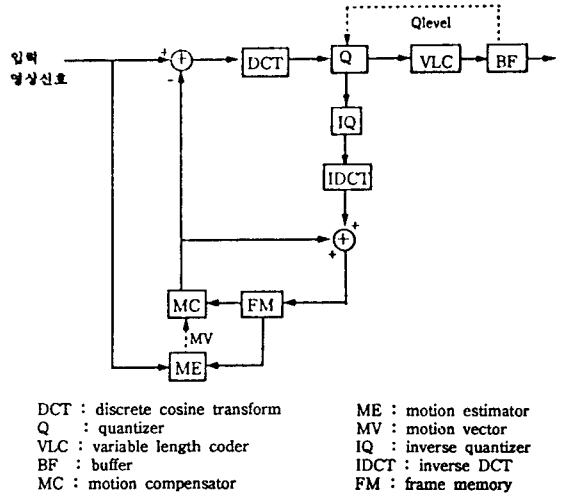


그림 1. 영상신호압축 system block diagram

리고 있고, 시간 해상도는 NTSC와 같다. Progressive 주사방식은 주사선수가 NTSC의 1.5배이므로 공간해상도가 수평 수직 각각 1.5배로 늘었고 시간축 방향으로 초당 프레임수가 2배로 늘어났다. 따라서 전체적으로 progressive방식이 데이터량이 약간 더 많고 특히 수직 방향 움직임이 있는 영상에 대한 해상도가 라인간 flickering 현상이 있는 interlaced 방식보다 뛰어나다. 또 유효 영상 영역이 1280(H)×720(V) 화소이어서 16:9의 화면비와 같아지므로 정방형 화소가 되어, computer graphics와 호환성이 유지되고 신호처리도 용이해진다. Progressive의 단점은 수평주파수(47.2KHz)가 interlaced(31.5KHz)의 1.5배에 달해 CRT 구동부가 기술적으로 한층 어려워지고 시간 해상도가 공간 해상도보다 지나치게 높아 데이터 압축시 시공간에 비트 할당의 균형을 유지시키기 위한 보다 효율적인 알고리즘이 요구된다.

2. 색성분간 중복성의 제거

카메라로부터 출력되는 R, G, B 신호는 색성분간의 상호 상관도가 높아, 이 상관도를 떨어뜨리고 인간의 시각 특성에 맞추도록 하기 위해 Y, U, V의 색체계로 변환한다. 인간의 눈이 가장 민감한 Y 성분은 약 30MHz의 대역을 차지하나 U와 V는 약 15MHz 정도만을 차지하므로 decimation에 의해 샘플의 수를 줄이더라도 정보의 손실이 거의 없다. 보통 Y:U:V를 4:1:1로(U와 V가 수평수직 각각 2:1로 decimation됨) 하는 경우가 많으며, 특히 interlaced 신호에서는 수직방향 deci-

표 1. 미국의 HDTV 제안 방식 비교

	DigiCipher	DSC-HDTV	ADTV	CCDC
Lines per Frame	1050	787 / 788	1050	787 / 788
Frames per Second	29.97	59.94	29.97	59.94
Interlace	2:1	1:1	2:1	1:1
Horizontal Scan Rate	31.469 KHz	47.203 KHz	31.469 KHz	47.203 KHz
Aspect Ratio	16:9	16:9	16:9	16:9
Active Video Pixels	1408H × 960v (luma) 352H × 480v (chroma)	1280H × 720v (luma) 640H × 360v (chroma)	1440H × 960v (luma) 720H × 480v (chroma)	1280H × 720v (luma)
Pixel Aspect Ratio	33:40	1:1	27:32	1:1
Video Bandwidth	21.5 MHz(luma) 5.4 MHz(chroma)	34 MHz(luma) 17 MHz(chroma)	27 MHz(Nyquist Limit)	34 MHz(luma) 34 MHz(chroma)
Colorimetry	SMPTE 240M	SMPTE 240M	SMPTE 240M	SMPTE 240M
Video Compression Algorithm	Motion-compensated DCT coding	Motion-compensated DCT coding	Motion-compensated DCT coding (MPEG-based)	Motion-compensated DCT coding
Block Size	8×8	8×8	8×8	8×8
Video Sampling Frequency	53.65 MHz	75.3 MHz	54 MHz	75.3 MHz
Audio Bandwidth	20 KHz	20 KHz	20 KHz	20 KHz
Audio Sampling Frequency	47.2 MHz	47.2 MHz	48 MHz	47.2 MHz
Dynamic Range	90dB	96dB	96dB	90dB
Number of Audio Channels	4	4	4	4
Video Data Rate	12.59 Mbits /s (16QAM) 17.49 Mbits /s (32QAM)	Automatically varies from 8.6 to 17.1 Mbits /s	14.98 Mbits /s (can be shared with additional audio and /or data)	15.636 Mbits /s
Audio Data Rate	0.5 Mbits /s DOLBY AC-2	0.5 Mbits /s DOLBY AC-2	0.512 Mbits /s (nominal) MUSICAM	0.5 Mbits /s DOLBY AC-2
Control Data	126 Kbits /s	40 Kbits /s (spare)	40 Kbits /s (data)	126 Kbits /s (access control)
Ancillary	126 Kbits /s	413 Kbits /s	512 Kbits /s	126 Kbits /s
Sync	N / A	292 to 544 Kbits /s	N / A	N / A
Total Data	19.51 Mbits /s (16QAM) 23.39 Mbits /s (16QAM)	11.1 to 21.0 Mbits /s	21.00 Mbits /s	19.43 Mbits /s
Error Correction Overhead	6.17 Mbits /s RS(116,106) (16 QAM) RS(155,145) (32 QAM)	1.3 to 2.4 Mbits /s RS(167,147)	23.6% (4.96 Mbits /s RS (depends on Priority))	3.042 Mbits /s RS(154,130)
RF Modulation (Terrestrial)	16 QAM or 32 QAM (3/4 or 4/5 TCM)	2-level and 4-level VSB	Spectrally shaped 32 QAM (9/10 TCM)	16 QAM or 32QAM (TCM)
3dB Bandwidth (Terrestrial)	4.88 MHz	5.38 MHz	5.2 MHz	4.86 MHz
C/N Threshold (Terrestrial)	12.5 dB (16QAM) 16.5 dB (32QAM)	10 dB (2VSB) 16 dB (4VSB)	11.1 dB (HP) 16.1 dB (SP)	19 dB
ATV Interference (C/I)	12.0 dB (16QAM) 16.0 dB (32QAM)	10 dB (2VSB) 16 dB (4VSB)	16.1 dB (HP & SP)	N / A
NTSC Interference (C/I) T	0.0 dB (16QAM) 5.0 dB (32QAM)	-6.0 dB (2VSB) 0.0 dB (4VSB)	-6 dB (HP) -2 dB (SP)	N / A
Channel Equalization (Ghost Canceling)	-2 to +24μs (multiple ghosts)	-2 to +20μs (multiple ghosts)	16μs (may be extended to 40μs)	N / A
RF Modulation (Satellite)	QPSK	MSK	QPSK	N / A
Bandwidth (Satellite)	24 MHz / 2 channels	20 MHz / channels	24 MHz / 2 channels	N / A
C/N Threshold (Terrestrial)	7.5 dB	8 dB	8 dB	N / A

mation에 의해 짝수(혹은 홀수)필드의 색신호를 모두 잃어버리므로, 양 필드의 색성분을 고루 취하기 위한 4:2:0 형식도 쓰이고 있다⁴¹⁾. 또 V신호의 대역폭이 U보다 약간 더 넓어 (NTSC의 경우 I가 Q보다 넓은 대역폭을 갖는다) V와 U의 데이터 할당이 적절히 균형을 이루지 않으면 빨간색 혹은 파란색 계통 어느 한쪽에서 artifact가 두드러질 수 있다.

3. 시간적 중복성의 제거

이웃하는 화면간에 존재하는 상관성을 이용하여 시간적 중복성을 없애기 위해, 움직임 보상 DPCM 방식을 이용한다. 화면을 일정한 크기의 매크로 블록들로 나누어, 각 매크로 블록이 이전 화면의 어느 곳으로 부터 움직여 왔는지를 나타내는 움직임 벡터를 구하고 이 벡터를 이용하여 움직임 보상을 행한다. 움직임 벡터를 구하는 방법으로는 MSE(mean square error)나 MAE(mean absolute error)등의 기준을 사용하는 block matching algorithm(BMA)을 비롯하여 Spatio-temporal gradient 방법, Pel-recursive 방법등이 있다⁴²⁾.

움직임 벡터의 탐색 영역과 정밀도는 오차 신호의 크기를 변화시키며 이는 최종 부호화된 데이터의 양을 변화시킨다. 대부분의 영상에서 수평 방향의 움직임이 수직 방향보다 많으므로 탐색영역도 보통 수평 방향으로 더 넓으며, 탐색영역은 이웃하는 화면간의 최대 변위가 어느 정도인가에 따라 결정된다. DSC-HDTV에서는 넓은 영역으로부터 좁은 영역으로 좁혀 들어오면서 움직임을 추정하는 단계적(hierarchical) 움직임 추정을 사용하므로 탐색영역이 넓지만 정확도가 떨어지는 반면⁴³⁾, 다른 여러 시스템에서 쓰고 있는 full-search 방법은 탐색영역은 좁으나 정확하다.

움직임 벡터의 정밀도는 흔히 화소단위 혹은 더욱 정밀한 반화소 단위까지 하고 있다. 반화소 단위의 움직임 벡터를 추정하기 위해 흔히 쓰이고 있는 방법은 우선 화소단위의 움직임 벡터를 BMA등의 방법에 의해 구하고, 이 벡터 주변의 반화소 위치의 값들을 interpolation에 의해 구한 후, 화소 단위 움직임 벡터를 포함하여 주변 9개의 반화소 단위 벡터 위치에 의해 2차 full-search를 행하여 오차가 가장 작은 위치가 반화소 단위 움직임 벡터로 결정된다. 한편, interlaced 주사방식의 영상 신호에 있어서는 한 프레임 내의 두 필드간에도 움직임이 있을 수 있으므로 필드단위의 움직임 보상이 더 효과적인 경우도 발생한다. 프레임 단위 움직임 보상 시에는 한 매크로 블록이 하나의 움직임 벡터를 필요로 하지만 필드 단위 움직임 보상 시에는 한 매크로 블록이

두개의(필드마다 하나씩) 움직임 벡터를 갖는다. 두 방식을 결합하여 매크로 블록마다 두 방식중 보다 효율적인 움직임 보상 방식을 취하는 적응적인 방법을 쓰면 복잡도는 증가하나 데이터량을 더 감소시킬 수 있다⁴⁴⁾. 최근에는 프레임 단위 움직임 보상을 하면서도 필드간의 움직임을 고려하여, 같은 parity 필드끼리 움직임 보상을 하되(even→even, odd→odd) 움직임 벡터의 연장선상에 가장 가까운 다른 parity 필드의 점을 움직임 보상을 위한 interpolation에 가담시키는, FAMC(field-time adjusted motion compensation) 방법이 제안 되었다⁴⁵⁾. 움직임 처리에 관한 보다 자세한 논의는 본 특집의 다른 논문^[28]을 참조하면 된다.

4. 공간적 중복성의 제거

1) 화면내(intra-frame) 부호화

영상신호는 흔히 상관계수가 0.9 이상인 1차 Gauss-Markov 모델로 근사화 시킨다^{44,45)}. 이때 에너지 집중화와 상관도 제거의 관점에서 이상적인—그러나 신호에 따라 변환 kernel이 달라져 구현이 곤란한—KLT(Karhunen Loeve transform)에 가장 가까운 성능을 내는 변환인 DCT(discrete cosine transform)가 영상 데이터 압축에 널리 쓰이고 있다⁴⁶⁾. DCT 블록사이즈는 에너지 집중과 복잡도간의 타협으로 8×8이 널리 쓰이고 있는데, 그 이상의 크기에서는 복잡도에 비해 효율 증가가 미미하다⁴⁴⁾. 영상을 band pass filter 군에 의해 주파수 성분별로 나눈 후 각 성분을 부호화하는 subband 부호화에 대해서도 많은 연구가 있어 왔으며⁴⁸⁾ 최근에는 subband filter의 특수형인 Wavelet 변환을 영상압축에 이용하려는 시도도 자주 눈에 띄고 있다⁴⁹⁾. 엄밀한 의미에서 변환부호화와 subband 부호화는 동일한 것으로, time-frequency(영상의 경우 space-frequency) 평면에서²⁰⁾ critical sampling을 한 후, 시간 방향과 주파수 방향 어느 쪽으로 격자점들을 묶어 동시에 부호화하느냐에 따라 다르게 불릴 뿐이다. 변환부호화의 경우 대부분 time window의 길이가 짧고(블록 크기와 같은 경우가 많음) 따라서 frequency leakage가 많은 반면, subband 부호화에서는 주파수 영역에서 성분별 분리가 잘 되도록 time window가 길다. Subband 부호화에서는 영상을 대역별로 나누는데 시간을 많이 소모하고, 특히 움직임 보상 DPCM방식에 있어서의 움직임 추정이 블록(혹은 매크로 블록) 단위로 이루어 지므로, 현재 구현되고 있는 거의 모든 시스템이 DCT에 기초한 변환부호화이다. Nonoverlapping 변환부호화의 최대의 단점은 블록의 경계가 눈에 보이기 쉬운 소위 blocking ar-

tifact인데, 이를 줄이고 coding gain을 높이기 위한 노력으로 최근에는 lapped orthogonal transform(LOT)이 제안되었다^[21].

2) 화면간(inter-frame) 부호화

움직임 보상 에러신호는 원화에 비해 일반적으로 에너지가 작고, 화소간의 상관도가 훨씬 떨어져 전술한 1차 Gauss-Markov 모델로 근사화 하기 어렵다. 이웃하는 화소간의 상관계수는 0.5 근처까지 떨어져 DCT를 하더라도 에너지 집중 효율이 높지 않다. Discrete sine transform을 비롯하여^[17], 최근의 MPEG II contest에서^[4] 몇가지 DCT이외의 다른 변환이 시도되었으나 아직 뚜렷하게 효율적인 변환이 발견되지 않고 있으며, 따라서 inter 모드에서도 여전히 DCT를 쓰는 경우가 많다.

3) Intra /inter의 결합

DPCM에 있어서는 수신기를 켜거나 채널을 바꾸거나 채널 에러가 발생하였을 때 화면을 빠른 시간 내에 복구할 수 있도록 하기 위해 모든 블록이 일정한 시간 간격이내에 intra로 부호화 되어야 한다. 일정한 간격의 프레임 단위^[3,4,10] 혹은 slice 단위로^[7,8] intra 부호화하기도 하고, DPCM loop에 leakage factor를 넣기도 한다^[9]. 강제로 intra 부호화되는 부분 이외에는 inter와 intra중 보다 능률적인 쪽을 택한다. 블록마다 inter와 intra의 에너지를 비교하여 작은 쪽을 선택하거나^[3,4] inter와 intra로 각각 부호화한 후 비트 발생량이 적은 쪽을 택한다^[7,8].

5. 양자화

영상신호가 DCT등에 의해 주파수 영역으로 변환된 후, 일반적으로 실수값을 갖는 각 계수들을 한정된 데이터 길이로 표현하기 위해서는 양자화하여야 한다. 이 과정에서 양자화 잡음이 발생하여 lossy coding이 된다. 저주파 성분은 진폭 변화가 크고 고주파 성분은 변화가 작아, 저주파 성분에 많은 비트가 할당된다. 단순히 복구된 영상의 SNR을 높이기 위해서라면 각 주파수 성분별로 통계적 분산을 구하여 양자화 잡음이 주파수에 관계없이 균일하게 유지되도록 비트할당을 하면 되지만, 인간의 시각(human visual system : HVS) 특성이 고주파 성분의 양자화 잡음에는 덜 민감하므로 고주파 성분일 수록 양자화 스텝 사이즈를 키워 양자화 잡음을 더 허용하고 있다^[16]. 이를 위해 DCT 계수들은 먼저 HVS에 따른 weighting matrix에 의해 나뉘진 후, 베퍼와 블록의 복잡성에 따라 결정되는 양자화 스텝 사이즈에 의해 양자화 된다^[3].

양자화 스텝 사이즈는, 양자화기 출력이 통계적 특성에 따라 가변장부호(variable length code : VLC)에 의해 더욱 압축되는 경우에는, 균일(uniform)해도 그다지 성능이 떨어지지 않으나^[3, 7, 8, 10], 만일 VLC가 뒤따르지 않을 경우에는 양자화기 스텝 사이즈를 입력신호 진폭에 따라 변화되도록 하는, 즉 발생빈도가 높은 작은 진폭에서 보다 세밀하게 양자화하는 최적(Lloyd-Max) 양자화기가 요구된다^[14]. CCIR 권고 723과 DSC-HDTV에서는 VLC를 쓰는 데도 불구하고 최적 양자화기를 사용하고 있다^[6, 9]. 또한 DSC-HDTV에서는 DCT 계수별로 따로 양자화기를 선택한 후 선택패턴을 베퍼 양자화하고 있다.

6. 가변장 부호화(VLC)

양자화된 DCT 계수는 최종적으로, 통계적 특성에 따라 가변장 부호를 사용하는 entropy 부호화를 통하여 더욱 데이터 압축이 이루어지는데, 이 과정은 양자화 과정과는 달리 손실이 없는 lossless coding이다. Entropy 부호로는 Huffman code, arithmetic code, universal code 등이 있는데^[22], 양자화된 DCT 계수는 주로 Huffman code를 사용하여 VLC 된다. 이때, DCT 계수중 DC 계수와 AC계수를 구분하여 서로 다른 방법으로 부호화 한다. 보통 각 블록의 DC값은 주변 블록의 DC값과 많은 상관성이 있으므로 이전 블록의 DC값과 차이를 구하여 그 차이값을 부호화하고, 첫번째 블록의 DC는 DC값의 가변범위의 중간값인 128과의 차이를 구하여 부호화 한다. 이렇게 구해진 DC의 차이값들은 일차원 VLC를 통하여 부호화하는데 그 한 예가 그림 2에 나타나 있다^[3].

또한 AC는 DCT 영역에서 DC계수 부근의 AC계수 값이 0이 아닐 확률이 높고, DC에서 멀어질수록 0이 발생할 확률이 높다는 점을 이용하여, 보다 더 효과적인 데이터 압축을 위해 계수들을 재정렬하는데, 주로 그림 3에 나타난 것과 같이 zig-zag scan을 통하여 1차원으로 정렬한다. 여기서 0이 연속적으로 나타나는 갯수(zero-run)와 0이 아닌 계수들의 값(level)을 (run, level)의 2차원으로 표현한다. 예를 들어 zig-zag scan이 되어 30, 2, 0, 0, -8, 0, 0, 0, 9, .. 와 같이 정렬된 DCT 계수는 run-level 부호화를 통하여 (0, 30), (0, 2), (2, -8), (3, 9), .. 와 같이 표현된다. 그리고 zig-zag scan된 계수들이 어떤 위치 이후에 계속해서 끝까지 0이 발생할 경우는 블록의 끝을 나타내는 EOB(end of block) 부호를 추가한다.

DC difference	Size	VLC
0	0	100
1	1	00
2~3	2	101
4~7	3	110
8~5	4	1110
16~31	5	11110
32~63	6	111110
64~127	7	1111110
128~255	8	

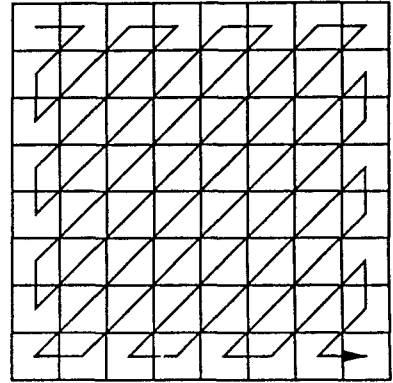


그림 3. DCT영역에서의 zig-zag scan 방법

DC difference	Size	Additional Code
-255 ~ -128	8	00000000 ~ 01111111
-127 ~ -64	7	0000000 ~ 0111111
-63 ~ -32	6	000000 ~ 011111
-31 ~ -16	5	00000 ~ 01111
-15 ~ -8	4	0000 ~ 0111
-7 ~ -4	3	000 ~ 011
-3 ~ -2	2	00 ~ 01
-1	1	0
0	0	
1	1	1
2~3	2	10 ~ 11
4~7	3	100 ~ 111
8~15	4	1000 ~ 1111
16~31	5	10000 ~ 11111
32~63	6	100000 ~ 111111
64~127	7	1000000 ~ 1111111
128~255	8	10000000 ~ 11111111

그림 2. 1-D VLC table for DC difference

이와같이 (run, level) 부호화된 데이터는 2-D Huffman 부호표를 이용하여 VLC되는데, 이러한 예는 그림 4에 나타낸 것과 같다. 여기서는 level에 있어서 양수와 음수 값이 나타날 확률이 거의 같으므로 level을 절대값으로 취하고 sign은 별도로 붙인다. 그런데 zero-run의 값이나 level의 값이 매우 클 경우에는 부호어의 길이가 너무 길어져서 구현하기가 힘들어지므로, 최대 부호길이를 제한하기 위해 escape sequence를 도

입한다. 즉, run이나 level이 일정한 값 이상으로 클 경우는 escape + run + level + sign과 같이 고정길이의 표현함으로써 부호어가 너무 길어지는 것을 방지한다.

7. Rate Control

HDTV 전송채널의 대역폭은 고정되어 있는 반면, 영상 데이터들은 최종적으로 VLC로 부호화 되므로 발생하는 데이터의 양은 시간에 따라서 달라진다. 따라서 주어진 전송율에 맞추어 발생하는 데이터량을 조절하는 rate control(또는 버퍼 control)이 필요하다. Rate control은 주로 버퍼의 충만도에 따라 양자화기의 step size를 가변함으로써 데이터의 발생량을 조절한다. 즉 발생한 비트 수가 기준치 이상이면 버퍼에 채워지는 데이터량이 증가하므로 양자화 step size를 증가시켜 다음에 발생할 비트수를 감소시키고, 기준치 이하로 데이터가 발생하면 그 반대로 수행하여 전체적으로 버퍼의 상태가 일정치를 유지할 수 있도록 조절한다. 이때 양자화 step size에 따라서 양자화 에러가 발생하게 되고 이는 곧 화질에 직접적인 영향을 미치게 되므로 버퍼 control 기법은 영상의 화질을 좌우하는 매우 중요한 요인중의 하나라고 할 수 있다.

Rate control 방법은 영상 신호 압축 방식에 따라 다소 차이가 있는데, MPEG에서와 같이 target 비트를 설정해 놓고, 실제 발생한 데이터로써 가상적인 버퍼(virtual buffer)의 충만도를 계산하고 이에 따라 양자화 step size를 결정하고 계속하여 target 비트를 update 시켜 나가는 방법이 있고^{3, 4)} Digicipher, CCDC, DSC-HDTV등에서 사용하는 것과 같이 매 slice마다 실

R \ A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26...	31	32....	41	42.....	255																			
0	2	4	5	7	8	8	10	8	12	12	11	13	13	13	14	14	14	14	14	14	14	14	14	14	14	14	14	14	15	15	15																			
1	3	6	8	10	12	13	13	15	15	15	15	15	15	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16																			
2	4	7	10	12	12																																													
3	5	8	12	13																																														
4	5	10	12																																															
5	6	10	13																																															
6	6	12	16																																															
7	6	12																																																
8	7	12																																																
9	7	13																																																
10	8	13																																																
11	8	16																																																
12	8	16																																																
13	8	16																																																
14	10	16																																																
15	10	16																																																
16	10	16																																																
17	12																																																	
18	12																																																	
19	12																																																	
20	12																																																	
21	12																																																	
22	13																																																	
23	13																																																	
24	13																																																	
25	13																																																	
26	13																																																	
27	16																																																	
28	16																																																	
29	16																																																	
30	16																																																	
31	16																																																	
32 ↓ 63																										6+6+8=20																								

EOB=2 bits

그림 4. VLC code length table

제 버퍼(physical buffer)의 상태에 따라 결정되는 step size를 feedback 받아서 다음 slice를 부호화해 나가는 방법이 있다^[7, 8, 9].

MPEG에서는 기본적으로 화면을 I(intra), B(bidirectional interpolative), P(predictive) 모드로 나누어 각각 서로 다른 target 비트를 설정한다. 각 모드에 대해 slice당 target 비트를 계산하고 실제로 한 slice에서 발생된 비트와의 차이에 해당하는 만큼으로 각 모드에 대하여 가상적인 버퍼로 update하고 이 가상 버퍼의 상태에 따라 다음에 부호화 될 slice의 양자화 step size를 결정하여 부호화 한다. 이렇게 한 frame의 부호화가 끝나고 난 뒤 frame당 발생한 비트량과 남아 있는 화면의 모드를 감안하여 다음 frame의 target 비트를 계산하고 계속적으로 부호화해 나간다. 화면내에서 I 모드와 P 모드가 공존하므로 MPEG에서의 rate control과는 달리 slice 단위로 실제 버퍼의 총만도에 따라 양자화 step size를 결정한다. 즉 한 slice에서 발생한 데이터를 버퍼에 채워 넣고 그동안 채널을 통하여

데이터가 전송되어 빠져나가고 난 후의 버퍼상태에 비례하여 다음 slice의 양자화 step size를 결정한다. 이 방법은 MPEG에서의 방법을 사용했을때 I, B, P각 화면단위로 화질의 변화가 심하게 나타나는 것과 달리 비교적 안정된 화질을 유지해 나갈 수 있다는 장점이 있다.

III. 음성신호의 데이터 압축

HDTV에 있어서는 고해상도 영상과 함께 CD 음질의 4채널 음성신호를 전송해 주어야 하는데, CD(16비트, 44.1KHz sampling)나 DAT(16비트, 48KHz sampling)처럼 단순한 PCM 방식으로 보내면 약 3Mbps를 차지하여 낭비요인이 된다. 따라서 이를 6:1로 압축하여 4채널 합하여 0.5 Mbps를 차지하도록 하고 있다. 음성신호는 영상신호에 비해 redundancy가 적으므로 영

상과 같은 높은 압축률을 얻을 수 없다. 음성부호화의 핵심은 신호를 transform이나 subband analysis에 의해 주파수 성분을 나눈 후 noise masking과 같은 인간의 청각특성을 활용하여 각 주파수 대역별로 귀로 느껴지지 않을 정도의 양자화 잡음을 발생하도록 양자화하는 것이다. ADTV의 MUSICAM 방식에서는 시간영역에서 window의 길이가 길고 주파수 영역에서 이웃하는 밴드끼리만 겹치도록 하여 subband filter bank를 설계하고 있고¹⁰⁾, Digicipher와 DSC-HDTV에서 채택하고 있는 Dolby사의 AC-2 방식과 CCDC에서는 반대로 시간영역에서 이웃하는 window끼리만 겹치고 주파수 영역에서 filter의 주파수 응답이 길어(즉, leakage가 많아) 시간영역에서 aliasing을 제거하는 소위 TDAC(time domain aliasing cancellation) 방법²³⁾을 채택하고 있다. 음성 부호화에 관한 보다 자세한 논의는 본 특집의 다른 논문²⁹⁾에서 다루고 있다.

IV. 채널 부호화

ECC(error correcting code)로는 burst 에러에 강한 Reed-Solomon 부호와 random 에러에 강한 길쌈(convolucional) 부호를 결합하여 사용하는 concatenated ECC 방식이 우수한 것으로 알려져 있다²⁴⁾. 또 ECC 출력신호의 변조는 지상 방송처럼 대역폭이 좁고 송신 전력을 키울 수 있는 환경에서는 32QAM등과 같이 대역효율(bandwidth efficiency)이 높고 전력 효율(power efficiency)이 낮은 방식을 택하지만, 반대로 위성 방송처럼 대역폭이 넓고 송신 전력이 낮은 환경에서는 QPSK등과 같이 대역효율이 낮고 전력 효율이 높은 방식을 택한다³¹⁾. 그런데 ECC와 변복조부를 독립적으로 최적화하더라도 최적의 성능을 내지는 않기 때문에 최근에는 길쌈부호와 디지털 변조를 결합하여, 출력신호 symbol들이 길쌈부호에 의해 고유하게 형성되는 trellis를 따라 변하는 TCM(trellis coded modulation) 방식²⁵⁾이 CCITT의 고속 모뎀에 채용되고 있고, HDTV 전송 수단으로도 유력하다. TCM은 symbol간의 Euclidian distance를 이용함으로써, 비트간의 hamming distance를 이용하는 단순 길쌈 부호보다 개선된 성능을 얻으며, TCM을 쓰지 않고 단순 변조하는 것에 비해 3dB 이상의 이득을 얻고 있다. 보다 자세한 논의를 본 특집의 다른 논문³⁰⁾에서 다루고 있다.

V. HDTV 영상 데이터 압축을 위한 새로운 시도

미국에서 제안된 여러가지의 full-digital HDTV 방식들은 이상에서 설명한 것과 같은 기본적인 데이터 압축 방식을 근간으로 하여 이루어져 있다. 여기서는 최근 본 연구실에 의해 제안되어 우수한 화질을 보여주고 있는 full-digital HDTV용 영상데이터 압축 방식에 대하여 설명하고자 한다. 기본적으로 1050/30/2:1을 사용하고 있으며, 유효 영상 영역의 화소수는 1408(수평)×960(수직)이다. DCT 단위인 블록은 8×8, 움직임 추정 단위인 매크로 블록은 32(수평)×16(수직)이고 한 slice는 수평방향으로 44개의 매크로 블록이 모여서 이루어져 있다. 강제 intra 모드는 11 매크로 블록마다 한번씩 생기며, 움직임 벡터의 탐색 영역은 -32~31(수평), -16~15(수직)이다. 그림 1의 기본적인 데이터 압축 방식에 부가 보완된 부분에 대한 자세한 설명은 다음과 같다²⁷⁾.

1. 적응 Prefilter

부호화기에 입력되는 영상 신호가 복잡하거나 고주파 성분이 많으면 부호화되었을 때의 데이터량이 많아져서 버퍼가 overflow될 염려가 있으므로 영상 신호의 대역을 적절한 방법으로 제한할 필요가 있다. 이때 주로 low pass filter의 성질을 가지 prefilter를 통하여 대역을 제한하는데 그 예로서는

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \text{ 이나 } \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \text{ 등이 있다.}$$

그러나 복잡하지 않은 영상 신호가 계속해서 입력되어 버퍼 속의 데이터가 줄어들었을 때, 즉 양자화 step size가 작을 때는 이 prefilter가 오히려 화질을 열화시키는 요인이 될 수 있으므로, 본 방식에서는 버퍼의 상태에 따라 low pass filter(LPF)와 all pass filter(APF) 사이에서 적응적으로 변하는 prefilter를 설계하였다. 즉,

$$\text{prefilter} = \alpha \times \text{LPF} + (1-\alpha) \times \text{APF}$$

여기서 α 는 양자화기의 스텝사이즈(qlevel)에 의해 결정되는 parameter로서 예를 들면 그림 5와 같다.

보통 prefilter에 주로 사용되는 LPF는 $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ 인데

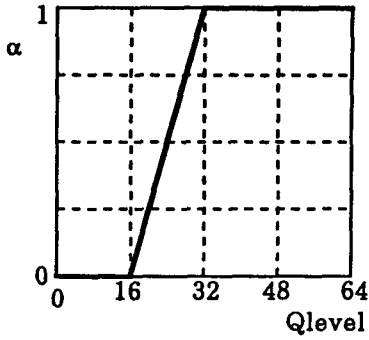


그림 5. Qlevel에 따른 α 의 변화

$$\frac{1}{256} \times \begin{bmatrix} 4 & 1 & -3 & 7 & -3 & 1 & 4 \\ 1 & -8 & -8 & 11 & -8 & -8 & 1 \\ -3 & -8 & 5 & 42 & 5 & -8 & -3 \\ 7 & 11 & 42 & 92 & 42 & 11 & 7 \\ -3 & -8 & 5 & 42 & 5 & -8 & -3 \\ 1 & -8 & -8 & 11 & -8 & -8 & 1 \\ 4 & 1 & -3 & 7 & -3 & 1 & 4 \end{bmatrix}$$

그림 7. 7×7 diagonal filter의 계수

이 filter의 2차원 주파수 특성은 그림 6과 같다. 그런데 인간의 시각 특성을 대각선 방향보다 수평, 수직방향으로 더 민감하므로 위와 같은 특성을 갖는 LPF로 prefilter를 구성하는 것보다, 대각선 방향으로 주파수 대역이 다소 좁더라도 수평, 수직 방향으로 보다더 넓은 대역폭을 갖는 LPF가 요구된다. 본 방식에서는 이에 적합한 7×7 diagonal filter로 LPF를 설계하였는데, 그 filter의 계수는 그림 7과 같으며 이 filter의 주파수 특성은 그림 8과 같다. 그림 8에서 보는 바와 같이, 3×3 LPF보다 수평, 수직 방향으로 더 높은 주파수까지 포함하고 있음을 알 수 있으며, 이 filter를 사용하여 부호화 해 본 결과 시각적으로 훨씬 더 화질이 뛰어난 결과를 얻을 수 있었다.

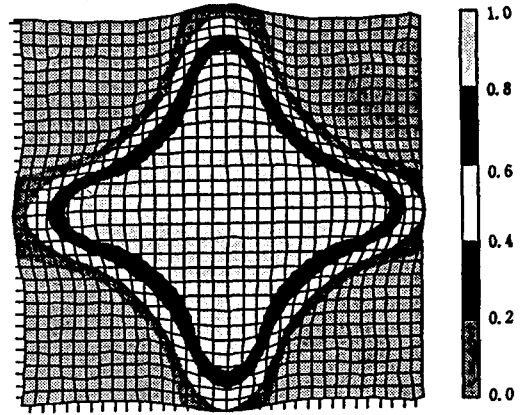


그림 8. Diagonal filter의 주파수 특성

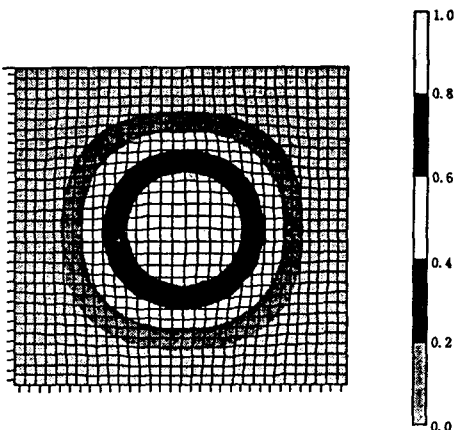


그림 6. $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ filter의 주파수 특성

2. 반화소 움직임 추정

움직임 보상을 위한 반화소 단위(실제로는 주로 반화소 단위로 움직임 벡터를 결정함) 움직임 벡터를 찾기 위해서, 진술한 바와 같이 기존 방법에서는 화소단위 움직임 벡터를 찾은 후에도 interpolation과 9-point search를 위한 시간이 추가로 더 걸려 실시간 hardware를 구현하는 데에 어려움이 있다. 본 방식에서는 이러한 interpolation과 9-point search 없이, 훨씬 간단하면서도 정확도가 뛰어난 반화소 단위의 움직임 벡터를 구할 수 있는 방법을 제시한다. 정확한 움직임 벡터 부근에서 x, y축으로 조금씩 움직여 가면 움직임 보상 에러들이 움직임 거리에 대해 증가하는데, x축 방향으로 움직였을때의 움직임 보상 평균 에러 E(x)를 다음과 같은 함수로 modelling할 수 있다.

$$E(x) = a |x - b|^7 + c, (a > 0, |b| < 1)$$

여기서 b 는 움직임 보상 에러가 최소가 되는 x 의 값 즉, 최적의 움직임 벡터에 해당하는 값이다. $\gamma=2$ 즉, 2차함수로 modelling하여 반화소 움직임 추정하는 경우도 있으나⁹⁾, 본 방식에서 보다 더 정확한 반화소 단위 움직임 벡터를 찾기 위한 최적의 γ 를 simulation을 통해 구한 결과 MSE를 사용할 때는 1.5, MAE를 사용할 때는 1.0으로 나타났다. $\gamma=1$ 인 경우 그림 9에서와 같이 $x = -1, 0, +1$ 일 때의 평균 에러값 P_{-1}, P_0, P_1 을 이용하여 b 를 구하면 다음과 같다.

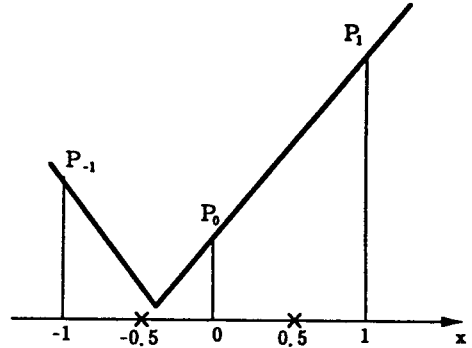


그림 9. 선형근사화한 motion compensation error

$$\frac{P_1 - P_0}{P_{-1} - P_0} = \frac{|1-b| - |b|}{|1+b| - |b|}$$

여기서 $-1/4 \leq b \leq 1/4$ 일 때는 반화소 단위의 수정이 없는 것으로 생각할 수 있고 $b > 1/4$ 이거나 $b < -1/4$ 일 때는 각 방향으로 반화소 단위의 수정이 있는 것으로 판단할 수 있다. 즉, 화소단위의 움직임 벡터가 MV_x 이고 반화소 단위의 움직임 벡터가 MV_x' 이라 할 때 다음과 같이 반화소 단위의 움직임 벡터를 구할 수 있다.

$$\begin{aligned} MV_x' &= MV_x - 0.5 && \text{if } 2(P_{-1} - P_0) < P_1 - P_0 \\ MV_x' &= MV_x + 0.5 && \text{if } (P_{-1} - P_0) < 2(P_1 - P_0) \\ MV_x' &= MV_x && \text{otherwise} \end{aligned}$$

이상과 같이 추가적인 반화소 단위의 탐색 절차없이 간단한 비교만으로 반화소 단위의 움직임 벡터를 구할 수 있으며, y 축으로도 마찬가지로 방법이 적용될 수 있다. 제안된 방법의 성능을 실제 영상을 이용하여 시험한 결과, "flower garden" 영상의 경우 움직임 보상후의 PSNR이 화소단위 움직임 벡터를 썼을 때보다 0.8dB 증가하였으며, interpolation과 9-point search에 의한 반화소 단위 움직임 벡터를 썼을 때보다 불과 0.08dB 낮았다¹²⁶⁾.

3. Inter/ Intera & Field/ Frame 모드의 결정

흔히 블록마다 inter와 intra의 에너지를 비교하여 작은 쪽을 선택하는데^{13, 4)}, 이 방법의 단점은 에너지가 실제 부호화 했을 때 발생하는 데이터량과 정확하게 비례하지 않는다는 점이다. 본 방식에서는 보다 더 효율적인 데이터 압축을 위해 각 블록을 inter /intra 2가지 모드에 대해 DCT, 양자화 VLC한 다음 발생하는 데이터의 양을 비교하여 작은 쪽을 그 블록의 모드로 결정하

였다^{7, 8)}.

또한 영상 신호가 interlaced로 scan 되어 있으므로 field 또는 frame 두가지 모드가 가능한데, field /frame 모드를 결정하기 위해서는 매크로 블록내의 블록을 2블록씩 묶어서 field 또는 frame 2가지 경우에 대해 각각 부호화하여 DCT-양자화-VLC한 다음 발생하는 데이터량이 적은 모드로 선택한다.

4. 비선형 버퍼 Control

버퍼는 VLC에서 발생하는 가변적인 데이터량과 일정한 전송 데이터량과의 차이를 조절하기 위해 존재하며, 버퍼 상태에 따라 발생 데이터량을 조절한다. 보통은 버퍼에 쌓인 데이터량과 비례하여 양자화 스텝(qlevel)이 결정된다¹¹⁾. 그러나 이 경우, 화면의 변화가 심할 때 버퍼에 입출력되는 데이터량도 변화가 많아지고 따라서 qlevel이 심하게 진동하므로, 양자화 에러의 변화 즉, 화질의 변화가 많아져 전체적으로 화질이 나쁘게 느껴진다. 본 방식에서는 그림 10에 나타난 바와같이 qlevel과 버퍼 상태를 비선형의 관계가 되게 함으로써 버퍼상태의 급격한 변화에도 화질의 변화가 심하지 않도록 하여 전체적으로 화질이 좋은 상태를 유지할 수 있게 하였다. 또한 qlevel이 매우 큰 부분 즉 60, 61, 62 일 경우에는 inter /intra, DC /AC 구분없이 양자화 step size를 64, 128, 256으로 하였고 qlevel이 63이면 DCT 계수를 모두 0으로 만들어 버퍼의 overflow를 방지하였다.

5. Dynamic Escape Sequence

VLC를 할 때, 부호의 최대길이를 제한하기 위해 도입하는 escape sequence에서 부호 길이를 escape 부호

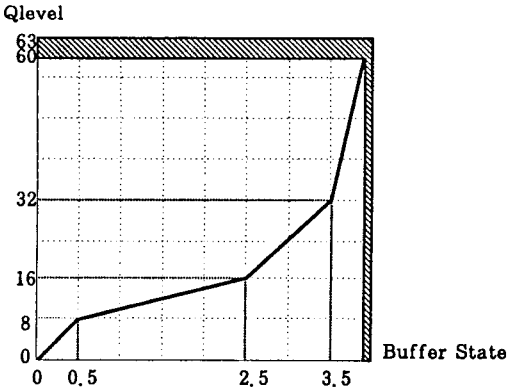


그림 10. Buffer state와 Qlevel의 관계를 나타내는 graph

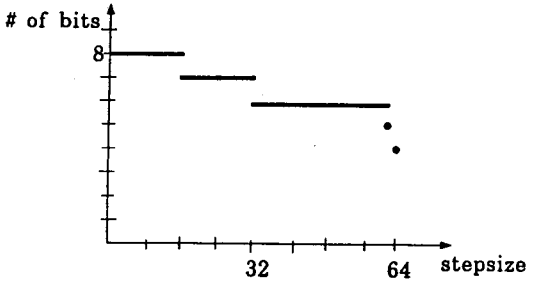
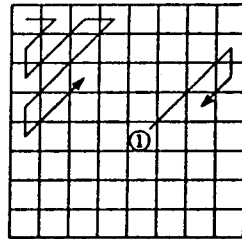


그림 11. 각 stepsize에 대해 level을 표현하기 위해 필요한 bit 수

에 6비트, run length에 6비트, level과 sign에 9비트를 할당하여 총 21비트의 escape sequence로 구성한다고 가정하면 부호어의 길이는 21비트 이하로 한정된다. 여기서 level은 양자화 step size와 관계되어 있어, step size가 커질수록 level의 최대값이 줄어들게 되고, 이를 이용하여 level에 소요되는 비트수를 그림 11과 같이 감소시킬 수 있다. Run length를 표현함에 있어서도 escape sequence가 발생하기 직전에 0이 아닌 계수가 발생한 위치를 알고 있기 때문에 그림 12와 같이 run length의 최대 길이를 예측할 수 있다. 이러한 점을 이용하여 run length에서의 소요 비트수를 감소시킬 수 있다.

6. Forward Analysis

버퍼의 상태에 따라 qllevel을 제어할 때에 주로 slice 단위로 수행하므로 한 slice내의 qllevel(squant)은 일정하다. 그러나 인간의 시각 특성은 고주파 성분이 많고 복잡한 화면에서보다 단순하고 변화가 없는 화면에서 양자화 에러를 더 민감하게 느낀다. 따라서 한 slice내에서도 매크로블록의 복잡도에 따라 qllevel를 재조정하여(mquant), 시각적으로 양자화 에러를 덜 느끼게 할 필요가 있다. 흔히 매크로블록 내의 화소값들의 분산을 복잡도로 이용하는 데, 이 방법은 단순 edge(qllevel이 낮아야 함)와 복잡한 부분(qllevel이 높아야함)과의 구분이 어려워, 경우에 따라서는 화질에 역효과가 나타난다. 본 방식에서는 DCT 계수들로부터 화면의 특징을 추출하여 4가지로 구분하였다. 즉, 단순한 영역, 복잡한 영역, edge가 있는 영역, 그밖의 영역으로 구분하여 각



Maximum Run=23
소요 bit 수=5 bits

그림 12. 어떤 위치에서 다음 runlength를 표현하기 위해 필요한 bit 수

각에 대해 qllevel을 약간 가변함으로써 시각적으로 화질이 개선되도록 하였다. 이를 위하여 우선 DCT 영역을 그림 13과 같이 4가지의 부영역으로 나누었다.

Forward analysis는 매크로블록 단위로 수행되므로 DCT 영역의 각 부영역은 매크로블록 내의 8개 블록에 대해 평균값을 취하였다. 여기서 화면의 특징을 추출하는 rule은 다음과 같이 차례로 우선순위를 가지도록 정하였다. 즉,

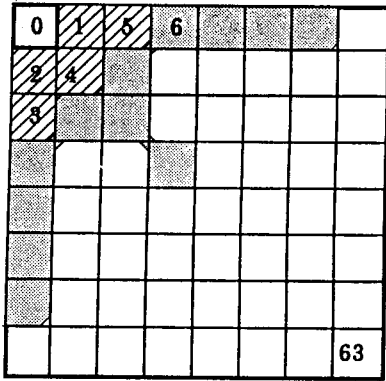
단순 영역 : if (E+H < 100 & L > 10)

Edge 영역 : if (L + E > 3×H)

복잡한 영역 : if (L + E < 3×H & E+H > 150)

그밖의 영역 : otherwise

이렇게 판별된 각 영역에 대해 단순 영역과 edge 영역의 qllevel은 약간 낮춤으로써 양자화 에러를 줄이고, 복잡한 영역에서는 양자화 에러에 둔감하므로 qllevel을 약간 올림으로써 양자화 에러가 다소 증가하도록 하여 전체적으로 발생하는 데이터량이 균형을 맞추도록 하였다. 그 밖의 영역에서는 qllevel의 변동이 없다.



- ⊙ : DC
- ▨ : L(low frequency)
- ⊠ : E(edge)
- : H(high frequency)

그림 13. DCT 영역을 4개의 부영역으로 나눈 그림

7. 실험결과

여러가지의 영상신호 데이터에 대하여 computer 모의실험을 통하여 본연구실에서 제안한 방식의 성능을 test해 본 결과, 70:1의 높은 압축률에도 불구하고 대부분의 영상에서 원화와 거의 비슷한 화질이 얻어졌다. 움직임이 많고 복잡한 영상에서는 원화와 미세한 차이를 느낄수 있으나, 비교적 움직임이 적거나 단순한 영상인 원화와 차이를 거의 느낄수 없었다. 또한 가까운 거리에서 자세히 관찰하면 약간의 artifact를 발견할 수 있으나, 정상적인 시청거리(HDTV의 경우에는 화면 높이의 3배)에서는 원화와 거의 같은 화질의 영상을 감상할 수 있다(주관적 평가). 압축, 복원한 영상의 화질을 나타내는 객관적 평가 수치인 peak-to-peak signal to noise ratio(PSNR)은 28-38 dB 정도를 유지하였다. 화질과 시스템의 복잡성을 놓고 볼 때 미국의 4개 디지털 HDTV 제안 방식에 근접한 것으로 평가되며, 모의 실험의 결과를 바탕으로 1993년 대전 EXPO 전시를 위해 현재 encoder와 decoder 하드웨어를 설계 제작중이다.

VI. 결론

HDTV는 2000년대 전자산업을 선도할 것으로 예측되는 주요 상품이기에 때문에 국가간 혹은 기업간에 치열


한 개발 경쟁을 벌이고 있다. 특히 미국을 중심으로 개발 진행중인 디지털 HDTV의 핵심 기술중의 하나인 디지털 영상 데이터 압축기술은 HDTV 뿐 아니라 화상전화, 디지털 VCR, 멀티 미디어, 멀티 채널 NTSC 방송(지상, 위성, 케이블 방송 모두 포함) 등 다양한 분야에 쓰이며, 국제적으로도 각 응용 분야마다 효율적인 부호화 방법이 쏟아져 나오고 있다. 비록 때 늦은 감이 있지만 우리나라도 최근 이 경쟁에 뛰어들어 연구가 활발히 진행되고 있는데, 보다 적극적인 관심과 투자에 의해 자체 기술을 축적하고 우수한 특허를 많이 발굴해 내어 장차의 HDTV 경쟁에 대비해야 할 것이다.

감사의 글

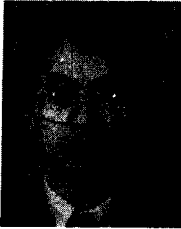
본고의 바탕이 된 한국형 HDTV방식 개발을 위해 지난 2년간 영상 데이터 압축에 관해 함께 연구하고 토론 하였던 영상 연구실 HDTV 팀원들 모두에게 감사하며, 아울러 정성껏 티이핑해준 이미자씨에게도 고마움을 표한다.

參 考 文 獻

- [1] "Video codec for audio visual services at p x 64 Kbits/s," CCITT Recommendation H. 261, 1990.
- [2] "JPEG digital compression and coding of continuous-tone still images," ISO-IEC/JTC1 Committee Draft 10918, 1991.
- [3] Motion Picture Expert Group, "MPEG Video Simulation Model Three (SM3)," Doc. ISO-IEC/JTC1/SC2/WG8/MPEG90, July 1990.
- [4] Motion Picture Expert Group, "MPEG II, Test Model 1, Draft Revision 1," Doc. ISO-IEC/JTC1/SC2/WG11/MPEG92/160, May 1992.
- [5] CCITT/SG15/WP1, "Report of the fourth meeting of the experts group for ATM video coding in Stockholm," 1992.
- [6] "Transmission of component-coded digital television signals for contribution-quality applications at the third hierarchical level of

- CCITT Recommendation G. 702," CCIR-CMTT/2, Document CMTT/13, 1992.
- [7] General Instrument Corporation, "DigiCipher™ HDTV System Description," 1991.
- [8] MIT, "Channel Compatible DigiCipher HDTV System," May 1992.
- [9] Zenith/AT&T, "Digital Spectrum Compatible HDTV: Technical Details," Sep. 1991.
- [10] Advanced Television Research Consortium, "Advanced Digital High Definition Television: System Description," Jan. 1992.
- [11] C. E. Shannon, "A Mathematical Theory of Communication," Bell Syst. Tech. J., vol. 27, pp. 379-423(Part 1), pp. 623-657(Part 2), 1948.
- [12] M. Rabbanni and P. Jones, *Digital Image Compression Techniques*, SPIE Optical Engineering Press, 1991.
- [13] J. K. Aggarwal and N. Nandhakumar, On the computation of motion from sequences of images-A review, *Proc. IEEE*, vol. 76, no. 8, pp. 917-935, Aug. 1988.
- [14] A. K. Jain, Image data compression: A review, *Proc. IEEE*, vol. 69, no. 3, pp. 349-389, Mar. 1981.
- [15] Jae S. Lim, *Two-dimensional Signal and Image Processing*, Prentice-Hall, 1990.
- [16] K. R. Rao and P. Yip, *Discrete Cosine Transform*, Academic Press, 1990.
- [17] P. Yip and K. R. Rao, "On the computation and the effectiveness of discrete sine transform," *Comput. Electr. Eng.*, pp. 7:45-55, 1980.
- [18] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 5, pp. 1278-1288, Oct. 1986.
- [19] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 11, no. 7, pp. 674-693, July 1989.
- [20] J. Jeong and W. J. Williams, "Kernel design for reduced interference time-frequency distributions," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 40, no. 2, pp. 402-412, Feb. 1992.
- [21] H. S. Malvar and D. H. Staelin, "The LOT: transform coding without blocking effects," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 4, pp. 553-559, Apr. 1989.
- [22] T. J. Lynch, *Data Compression: Techniques and Applications*, Lifetime Publication, London, 1985.
- [23] J. P. Princen and A. B. Bradley, "Analysis/synthesis filter bank design on time domain aliasing cancellation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 5, pp. 1153-1161, Oct. 1986.
- [24] G. D. Forney, *Concatenated Codes*, MIT Press, Cambridge MA, 1966.
- [25] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets. Part 1: Introduction; Part II: State of the art," *IEEE Commun. Mag.*, vol. 25, no. 2, pp. 5-21, Feb. 1987.
- [26] J. Jeong and W. Ahn, "Subpixel-accuracy motion estimation using a model for motion-compensated prediction errors," to appear in Proc. Signal Processing Conference, Korea, 1992.
- [27] HDTV 지상방송용 신호처리기 및 관련 반도체 연구개발-삼성전자종합연구소편 2차년도 보고서, 상공부, 체신부, 과기처, 1992.
- [28] 최윤식, "완전 디지털 HDTV 방식의 움직임 처리 방식," 전자공학회지, 1992년 12월호.
- [29] 이병욱, "HDTV용 고음질 디지털 오디오 데이터 압축기술," 전자공학회지, 1992년 12월호.
- [30] 박종석, "완전 디지털 HDTV의 전송방식," 전자공학회지, 1992년 12월호.
- [31] 이상훈, "HDTV를 위한 전송방식," 전자공학회지, 1992년 12월호.
- [32] 김재균, "MPEG II의 최근 기술 동향," 전자공학회지, 1992년 12월호. 

筆 者 紹 介

**鄭 濟 昌**

1957年 9月 10日生

1980年 2月 서울대학교 전자공학과(학사)

1982年 2月 한국과학기술원 전기 및 전자공학과(석사)

1990年 7月 미국 미시간(앤아버)대학 전기공학과(박사)

1982年 2月~1986年 8月 한국방송공사 기술연구소 연구원, 뉴미디어연구개발

1990年 7月~1991年 1月 미국 미시간(앤아버)대학 Research Associate

1991年 2月~현재 삼성전자 A/V연구소 영상연구실 수석연구원, 디지털 HDTV 연구개발 담당

주관심분야 : 영상/음성 데이터 압축, Wavelet변환, Digital Spectral Analysis,
 디지털신호처리, 영상신호처리

**趙 在 汶**

1961年 8月 17日生

1984年 2月 서울대학교 전기공학과(학사)

1986年 2月 한국과학기술원 전기 및 전자공학과(석사)

1991年 8月 한국과학기술원 전기 및 전자공학과(박사)

1986年 3月~1987年 2月 미국 UC Irvine대학 연구원

1991年 9月~현재 삼성전자 A/V연구소 영상연구실 선임연구원, 디지털 HDTV 연구개발담당

주관심분야 : 디지털 영상 신호처리, 영상데이터 압축, Medical Imaging