

영상 정보의 감축 부호화기법

전준현* · 김성대**

* (한국과학기술원 전기 및 전자공학과 대학원)

** (한국과학기술원 전기 및 전자공학과 교수)

1. 서 론

영상신호의 디지털 전송방식은 아날로그방식에 비해서 넓은 대역폭의 전송로가 필요하지만 장거리 전송에서 고도의 전송품질을 유지할 수 있고, 컴퓨터통신과 관련하여 영상정보의 컴퓨터 처리가 용이하다는 장점이 있다. 특히 영상의 디지털전송방식은 CATV, 원거리영상회의, 영상전화 및 TV 방송중계등과 관련하여 매우 중요한 연구과제가 되고 있다. 오늘날 통상 사용되고 있는 아날로그 영상신호의 대역폭은 NTSC, PAL, SECAM 등 방식에 따라 5-8 MHz 정도이다. 그러나 이 신호를 디지털방식으로 전송하기 위하여 8 bit PCM으로 부호화하게 되면 적어도 100Mbps 정도의 전송율을 가진 전송로가 필요하게 된다. 이러한 영상신호를 한정된 전송속도를 가지는 전송로를 통하여 전송하기 위하여는 필수적으로 데이터량을 감축해야 한다.

지금까지 연구된 영상신호의 데이터 감축방법은 예측오차를 이용하는 방식과 변환을 이용하는 방식, 그리고 이 두 방식을 조합한 복합감축방식으로 대별된다. 예측 부호화 방식은 현재 부호화될 화소의 실제값과 이미 부호화된 화소들로부터 얻을 수 있는 예측값과의 차이를 전송하는 방법으로 예측오차의 분산이 원래 영상신호의 분산에 비해 적어진다라는 점을 이용하는 것이다. 변환 부호화 방식은, 원래영상을 적당한 변환방법을 사용하여 변환하게 되면, 변환영역에서 일부 몇개의 변환계수만 큰 값을 가지고, 대부분의 나머지 계수들은 무시할

수 있을 만큼 작은 값을 가지게 된다는 점을 이용한다.

일반적으로 데이터 감축 방식의 선택은 전송로의 전송속도와 수신단에서 원하는 수신영상의 품질등에 달려있다. TV방송중계와 같이 고품질의 수신영상을 원하는 분야를 위해서는 45Mbps 정도의 속도를 가지는 전송로를 대상으로 한 DPCM에 의한 감축방식이 많이 이용되고 있다. 12Mbps 정도의 낮은속도의 전송로를 통하여 CATV 정도품질의 영상을 송수신하기 위한 방법으로는 DPCM에다 운동보상 예측방법을 도입하여 보다 더 예측오차를 줄임으로써 데이터를 감축하는 방식을 사용하기도 한다. 1.5Mbps 이하의 전송로를 대상으로 한 원거리 영상회의, 영상전화시스템등에 관한 연구에서는 변환부호방식 혹은 복합감축방식들이 주로 이용되고 있다.

현재 1.5Mbps급 전송로를 이용할 수 있는 비디오 코덱 시스템들 중에서 대표적인 제품으로는 CLI(Compression Lab Inc)의 VTSL5, NEC 사의 NETEC-X1, McM(GEC McMichael)의 VIDEO-CONFERENCEING-CODEC 등이 있고, 64Kbps급으로는 KDD의 INVITE-64, PICTEL사의 C-2000, MITSUBISHI의 MVC-3000, NEC사의 NETEC-XD, WIDCOM사의 VTC-56등이 있다.

본 논문에서는 지금까지 연구개발되어 있는 방법들 중에서 기본적인면서도 핵심적인 기법들을 간추려서 소개한다. 또한 각 방식들 중에서 기본적인 방식을 선택하여 computer simulation을 하였다. 표-1과 그림1은 사용

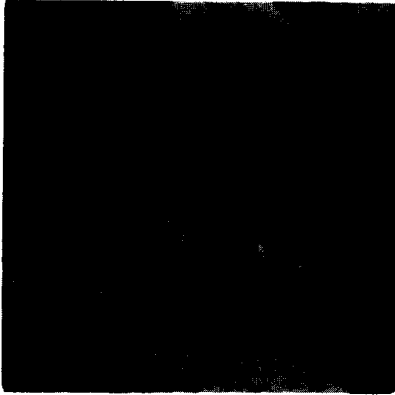
된 영상들의 통계적 특성 및 원영상들을 나타낸다.

2. DPCM을 이용한 부호화 기법

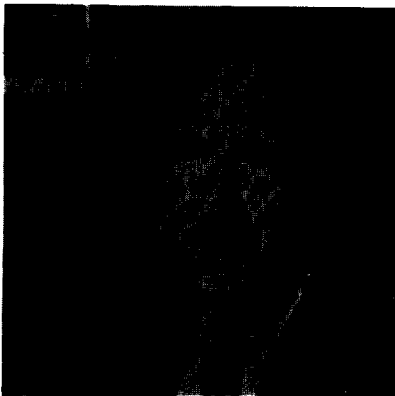
DPCM 방식은 예측부호화 방식으로 간단하면서도 다른 영상신호의 데이터 감축 기법들의 기본이 되는 매우 중요한 방식이다. 이 방식은 현재의 신호값을 과거의 값들로 부터 예측한 값으로 빼어 그차이(예측오차라고 함)를 양자화하여 전송하는 것으로서, 하드웨어의 구현이 비교적 용이하고 커다란 성능의 저하없이 쉽게 데이터 감축율을 $\frac{1}{2}$ ~ $\frac{1}{3}$ 정도 얻을 수 있다.

일반적으로 DPCM 시스템은 ADC, 예측기, 양자화기, 및 DAC등으로 이루어져 있는데 DPCM의 성능은 주로 예측과 양자화를 얼마나 잘 하느냐에 달려 있다. DPCM은 사용되는 예측방식에 따라서 intraframe coding과 interframe coding 방식이 있다. intraframe coding에서는 같은 frame에 있는 화소들로 부터 현재 화소의 값을 예측하고, interframe coding에서는 과거 frame들 (주로 과거 frame)에 있는 화소들로 부터 예측한다. DPCM을 이용하여 칼라 TV 영상신호를 부호화하는 방법에는 칼라 복합영상신호를 그대로 DPCM하는 composite DPCM 방식과 복합영상신호를 RGB 혹은 YIQ 등으로 분리하여 각 성분에 대해 따로 DPCM을 적용하는 component DPCM방식이 있다. component coding은 composite 방식에 비하여 성분분리등의 추가 하드웨어를 필요로 하기 때문에 composite 방식을 적용하기 힘든 저전송을 부호화에 주로 이용된다.

그림-2는 composite DPCM 방식을 이용한 간단한 45Mbps급 영상신호 부호화 시스템의 block diagram을 나타낸것이다. NTSC 신호의 대역폭이 약 5MHz이고, 예측에 이용될 화소들의 위상들을 고려하여 sampling 주파수는 색부반송파 주파수 f_{sc} (약 3.58MHz)의 3배가



(a) Girl



(b) Cronkite

그림 1. 원영상

표 1 영상데이터의 통계적 특성

	Component	line/pel	평균	표준편차
Girl	Y	256×256	181.4	42.62
Cronkite	Y	256×256	87.3	54.23
MSAM	Y	288×352	64.11	29.31
	R-Y	144×176	17.88	25.94
	B-Y	144×176	0	16.49

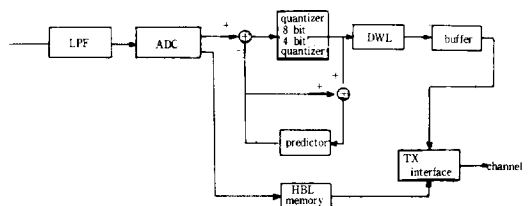


그림 2. Composite DPCM방식

예 측 방 식	예 측 치 x
전 예 측(previous prediction)	c
평균 예 측(average prediction)	$\frac{1}{2}(a+d)$
평면 예 측(planar prediction)	$a+g-i$
사와다예측(sawada prediction)	$\frac{1}{2}(g-e+d+a)$

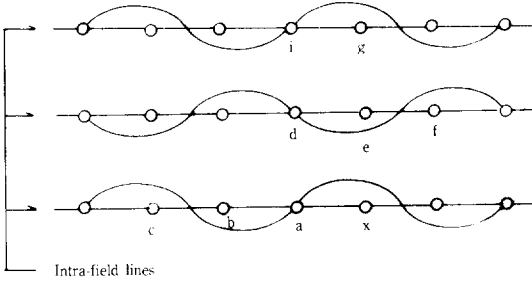


그림 3. Composite DPCM에서 예측 방법

적당하다. 그림-3은 composite DPCM 방식에서 많이 사용되는 예측방법을 나타낸것이다. 예측오차를 전송하기 위하여 양자화 과정을 거치는데 대개의 경우 30 level 정도이며 양자화후 VWC(Variable Word-length Coding)를 추가할 경우에는 70-80 level 정도로하여 보다더 quantization error를 줄이기도 한다¹⁾.

그림 4는 평면 예측(planar prediction)을 사용하여 2차원 DPCM 한 재생 영상으로써 한 화소당 정보량(bit per pel)에 따른 성능을 나타냈으며, Laplacian 특성의 비균등 양자화기를 사용한 후 VWC 하였다. 이때 사용된 영상은 "Girl"이다.

그림 5는 두개의 DPCM loop를 이용하여 영상데이터를 감축하는 CDC(Combinational Difference Coding) 방식의 block diagram을 도시한 것이다.

여기서 intraframe DPCM loop은 공간적으로 인접한 화소들간의 중복성을 줄이는데 사용되고, intraframe DPCM loop는 시간축상의 중복성을 제거하는데 사용된다. 그림에서 보는 바와같이 프레임메모리(frame memory) 와 화소지연소자(pel delay)를 이용하여 현재 부호화하고자하는 화소를 예측한다. 이때 예측치와 실제값 사이에는 차이가 생기게 되는데 대부분의 경우 이 예측오차의 값은 상당히 작게 된다. 이 예측오차를 15 level 양자화기(quantizer)를 이용하여 일차적으로 데이



(a) 1bpp, SNR=25.8dB (b) 2bpp, 2NR=32.1dB

(c) 3bpp, SNR=38.4dB (d) 4bpp, SNR=44.7dB

그림 4. 평면 예측을 사용 2 차원 DPCM 한 재생 영상(Data base: Girl)

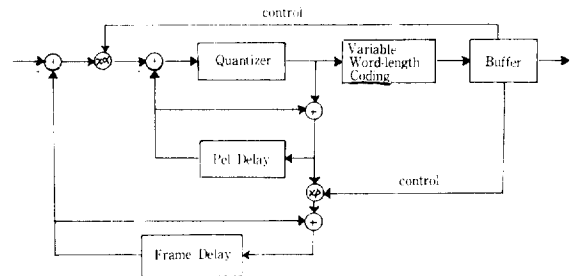


그림 5. CDC 방식

표 2 Entropy

Search methods	Average Entropy
Brute force	3.53
2-D logarithmic search	3.62
Three-step search	3.64
Conjugate direction search	3.75
Menu vector search	3.89

타를 감축한후, 양자화된 예측오차를 일종의 엔트로피 부호화방식인 VWC(Variable Word-length Coding) 부호화기에 입력시킨다. 양자화된 예측오차신호가 대부분의 경우 영일 확률이 크므로, 각 화소에 대하여 부호화를 하지 않고 작은 블록, 예를 들면 2 x 4 화소를 한 블

표 3 하드웨어 측면에서 비교

	Motion vector(5, 4)		Worst case	
	search point	sequential step	search point	sequential step
2-D logarithmic search	22	5	21	7
Three-step search	25	3	25	3
Conjugate direction search	14	11	15	12
Menu vector search	26	1	26	1

록으로 하여 VVC를 적용시킨다. 이와같이 하여 부호화된 데이터는, 영상의 움직임 정도에 따라서 그 양이 계속 변하기 때문에 직접 전송로를 통하여 보낼 수가 없다. 그래서 buffer라는 완충장치를 두어 데이터의 전송속도를 일정하게 해준다. buffer의 상태는 영상의 움직임이 많은 화면이 계속될 경우, 전송속도보다 데이터 발생량이 많을 경우, buffer의 용량을 초과할 우려가 있다. 이것은 overflow라고 하는데 이렇게 되면 데이터를 잃어버리게 된다. 그래서 적당한 방법을 사용하여 데이터의 발생량을 조절할 필요가 있다. CDC에서는 이를 위하여 buffer 상태에 따라서 α 의 값을 조절한다. α 는 0과 1사이의 값을 취하는데 데이터량이 많아지면 α 의 값을 작게 하여 예측오차를 줄여서 양자화된 예측오차의 값이 영이될 확률을 높인다. 그와 반대일 경우 α 값을 크게 하여 보다 정확한 영상이 보내어지도록 한다. 그러나 이와같은 α 값의 조절에도 불구하고 계속 buffer의 양이 증가하면 frame subsampling등을 통하여 데이터 발생량을 대폭 줄이는 방법을 사용한다. α 값이 작아져서 1/2정도보다 작아질 경우 blurring 현상이 생기게 되는데 이것을 보상해주기 위하여 β 값을 α 값의 변화에 따라 조절한다.

3. 이동정보를 이용한 영상 부호화 방법

현재까지 제안되어 있는 이동정보를 이용한 영상부호화 방법에는 CRC(Conditional Replenishment Coding) 방법, MCC(Motion Compensated Coding)방법들이 있다. CRC방법은 인접한 두 frame으로 부터 변화가 있는 움직임 영역을 검출한 뒤 그 영역에 해당하는 변화에 관한 정보만 전송하는 것으로 움직임이 적은 영상에서 좋은 성능을 갖는 방법이다. MCC방법에서는 인접한 두 frame 사이의 변화는 물체가 이동함으로 생긴 것이라 가정하여 각 화소별 혹은 각 영역별로 이동정보를

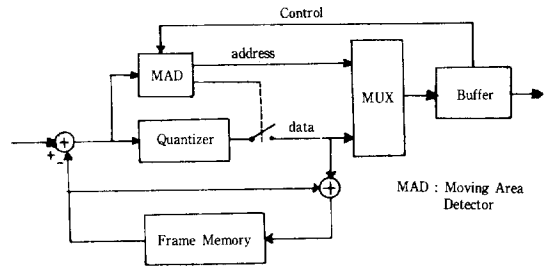


그림 6. CRC방식

추출하여 그 정보를 이용하여 부호화하고자 하는 화소 값을 예측 부호화 한다.

그림 6은 CRC의 encoder부분을 나타낸다. 점선안의 부분은 interframe DPCM과 같은 형태이고, 다만 움직임영역(MAD:Moving Area Detector) 부분만이 추가로 있다는 것만이 다른 점이다. CRC방식은 송·수신단에 프레임메모리가 있는데, 이동검출기는 이 프레임메모리에 들어있는 이전 frame과 입력되고 있는 현재의 frame과의 차이를 계산하여 움직임영역을 검출한다. 움직임영역 검출시 적당한 threshold값을 이용하는데 그 차이가 이 threshold 보다 클 경우 해당 화소를 움직임영역에 포함시킨다. MAD에 의하여 검출된 움직임영역에 포함되는 화소들의 이전화면에서의 값과 현재 화면에서의 값과의 차이를 양자화하여 그 화소의 위치정보와 함께 전송한다. 움직임영역이 작은 경우 이 방법으로 상당량의 데이터를 감축할 수 있지만, 그렇지 못할 경우 위치정보를 전송하는데 너무 많은 데이터가 필요하게 된다. 이를 위하여 clustering이란 방법을 사용하여 움직이는 영역에 포함되는 화소들 중 인접한 화소들을 묶어서 cluster를 만들고 그 cluster의 위치정보를 보내는 방법을 사용한다. 집음의 영향으로 인하여 움직임영역내

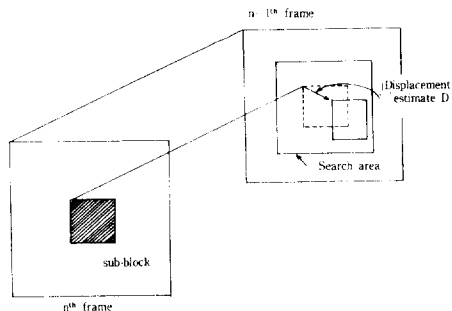


그림-7 BMA에 의한 이동 벡터 검출

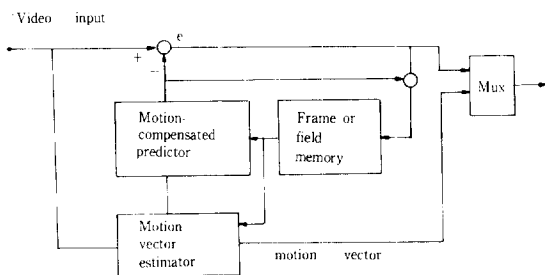


그림 8. BMA를 이용한 MCC방식

고립된 정지화소가 존재하든지 정지영역내에서 움직이는 화소가 존재할 경우 이로 인하여 많은 성능저하가 생기게 되는데, 이러한 경우 정지영역 혹은 움직임 영역 내에 고립된 화소들을 제거하면 위치전송에 필요한 데이터량을 줄일 수 있다. 이러한 방법을 suppression 혹은 bridging이라고 한다. 불균일하게 발생하는 데이터를 일정한 속도로 전송하기 위해서는 buffer가 필요하게 된다. CRC에서는 buffer내에 있는 데이터량을 조절하기 위하여 MAD의 threshold값을 조정하는 방법을 사용한 다.

MCC에서 가장 중요한 부분은 어떻게 이동정보를 검출하느냐이고 그 방법에 따라 BMA(Block Matching Approach)와 PRA(Pel Recursive Approach)가 있다.⁴⁾ 일반적인 물체의 운동은 회전운동, 병진 운동등을 포함 하는 복잡한 형태를 가지지만 인접 frame간의 시간차이가 1/30초 정도 이므로 frame간의 운동량은 극히 적고, 또 이를 병진운동으로 근사화할 수 있다. 이러한 가정에 근거를 두는 BMA방법에서는 현재 frame을 일정한 크기의 block으로 나누고, 현재 coding하고자하는 block과

가장 유사한 block을 과거 framed에서 찾아서 이동 vector를 검출해낸다.⁵⁾ PRA 방식을 이용한 motion vector estimation은 1979년 Netravali와 Robbins에 의해 처음 제안되었다.⁶⁾ 이 방식은 steepest descent algorithm을 이용하여 interframe DPCM coding에서 예측오차를 줄여 나가는 방식으로 각 화소단위로 이동 vector를 추정하는 특징을 갖는다. BMA에 비하여 각 화소에 대한 정확한 motion vector를 찾을 수가 있어서 예측오차를 좀 더 줄일 수 있지만, 많은 계산량으로 인하여 실제 실시간 구현시에 제약이 따른다.

그림7은 BMA에 의한 이동벡터 검출 방법을 도시한 것이고 그림8은 BMA를 이용한 부호화기의 기본적인 block diagram을 나타낸 것이다. $M \times N$ 의 block size에 대하여 frame당 최대운동속도를 p pels/frame 이라 할 경우, 좌우상하 p 화소 만큼 운동이 발생할 수 있으므로 search area의 크기는 $(M+2p) \times (N+2p)$ 가 된다. 따라서 $M \times N$ 의 block이 search area에서 matching될 수 있는 경우의 수는 $(2p+1)^2$ 개가 된다. 두 block사이의 유사성을 나타내는 함수로서 NCCF(Normalized Cross-Correlation Function), NMSE(Normalized Mean Square Error), 그리고 MNAE(Mean Number of bits necessary to binary-code the Absolute Error), MAE (Mean of the Absolute Error)등을 생각할 수 있으나 실제 적용시에 계산량이 비교적 적은 NMSE나 MAE가 많이 사용된다. 이러한 함수들을 이용하여 최적이동변위를 구할때 가능한 모든 예상변위에 대하여 계산하는 brute force search 방식이 있지만 이 방법은 엄청난 계산량 때문에 거의 사용되지 않고, 예상변위를 줄여서 계산량을 감축하는 방법들이 발표되어 있다. 그 대표적인 것으로 2-D logarithmic search, three step search, conjugate direction search, menu vector search 등이 있다. 표2는 data base로 256 x 256 크기의 8비트 화질인 "Cronkite" 16장을 사용하였을 때 각 방법들의 성능을 나타낸 것이며, 표3은 하드웨어 측면에서 고찰을 위해 이동 벡터가(5,4)인 경우에 search 해야할 point 수와 sequential step의 수, 그리고 최악의 경우에 있어서 search point 수와 sequential step의 수를 실었다. 이 두가지 표를 비교하여 보면, Koga의 three-step search가 성능면이나 하드웨어 측면에서 가장 유리하다는 것을 알 수가 있다.⁵⁾

특집 : 회로 및 신호처리

4. 변환을 이용한 부호화 방식

영상처리에서 많이 이용되는 변환에는 Haar변환, Walsh-Hadamard변환, Slant변환, DCT(Discrete Cosine Transform), DFT(Discrete Fourier Transform) 등이 있다. 변환부호화방식은 원래의 영상전체를 몇 개의 블록으로 분할한 후, 각 영역을 적당한 변환방법을 사용하여 변환한다. 일반적으로 변환을 하게 되면 대부분의 변환계수들은 영에 가까운 작은 값을 가지게 되고, 일부 몇 개의 변환계수들 만이 큰값을 가지게 되는데 이러한 것을 에너지집중현상 이라고 한다. 이때 적절한 방법으로 작은 값을 가지는 계수들을 버리고, 큰값을 가지는 계수들을 부호화함으로써 영상데이터를 감축할 수 있다. 이러한 방법을 사용하는 경우 KLT(Karhunen-Loeve Transform)가 가장 적합한 변환이라고 할 수 있다. 그러나, KLT는 영상의 covariance에 근거한 변환이기 때문에 TV영상과 같이 각 구역마다 각 화면마다 covariance가 다른 경우 매번 covariance를 구해야 한다는 단점이 있다. 영상데이터가 stationary하고 지수형 상관계수를 갖는 일차 Markoff process라고 가정할 경우 DCT는 에너지집중현상면에서 보았을 때 최적변환인 KLT와 가장 유사한 성능을 가지고 있고, 소프트웨어나 하드웨어의 구현이 가능하므로 변환을 이용한 데이터감축기법에서 많이 이용되고 있다.⁷⁸⁾ 이산함수 $f(j)$ 의 1차원 DCT는 다음과 같이 정의된다.

$$F(k) = \frac{2C(k)}{N} \sum_{j=0}^{N-1} f(j) \cdot \cos \frac{[(2j+1)k\pi]}{2N}$$

$$j, k=0, 1, \dots, N-1$$

$$f(j) = \sum_{k=0}^{N-1} C(k) F(k) \cos \frac{[(2j+1)k\pi]}{2N}$$

이때 F는 변환계수(transform coefficient)를 나타낸다.

$$여기서 C(k) = \begin{cases} \frac{1}{\sqrt{2}} & (k=0) \\ 1 & (k \neq 0) \end{cases} \quad K K$$

영상 데이터의 보다 많은 데이터 감축을 위하여 2차원 DCT를 사용하는데 2차원 DCT는 그림9 같은 서로 독립적인 1차원 DCT 2개로 구현 가능하며 변환 sub-block의 크기는 8 x 8 또는 16 x 16을 가장 많이 사용한다.⁸⁾

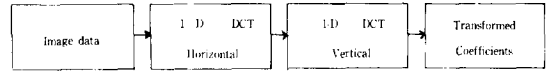


그림 9. 2 차원 DCT의 블록도

8	8	8	7	7	7	5	5	4	4	4	4	4	4	4	4	4
8	8	7	6	5	5	3	3	3	3	3	2	2	2	2	2	2
8	7	6	4	4	4	3	3	2	2	2	2	2	2	2	2	2
7	6	4	3	2	2	2	2	1	1	1	1	0	0	0	0	0
7	5	4	2	2	2	2	1	1	1	0	0	0	0	0	0	0
7	5	4	2	2	2	1	1	0	0	0	0	0	0	0	0	0
5	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0
5	3	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
4	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0

그림 10. 16x16블럭에 대해 한 화소당

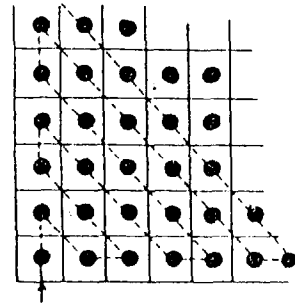


그림 11. Zig-Zag scanning 방법

변환을 이용한 부호화 방식으로는 변환 계수의 block 내에서 통계적으로 높은 에너지를 갖는 낮은 주파수의 계수들이 있는 영역(zone)을 정하여 그 계수들만을 양자화하는 zonal coding 방식과 중요하지 않는 계수들을 버림으로써 데이터 감축을 하는 threshold coding 방식들이 있다. 여기서 zonal coding은 block quantization을 한후 coding 전송할 bit allocation map을 필요로 한다. 그림10은 한화소당 1.5bpp인 bit map 이다. Threshold coding은 threshold보다 작은 값들을 모두 0으로 하고 이를 전송하기 위하여 RLC(Run-Length Coding)을 사용한다. 그리고 threshold 보다 큰 계수들은 모두 똑같은 레벨 수를 갖는 균등양자화기화 VWC(Variable Word-length Coding)을 함께 사용한다.

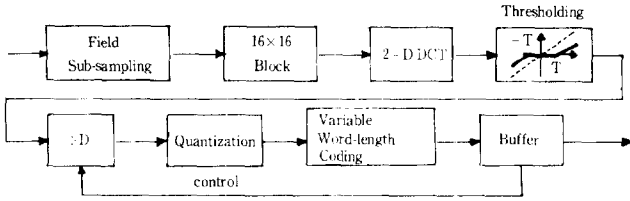


그림12. DCT를 이용한 SAC부호화기법

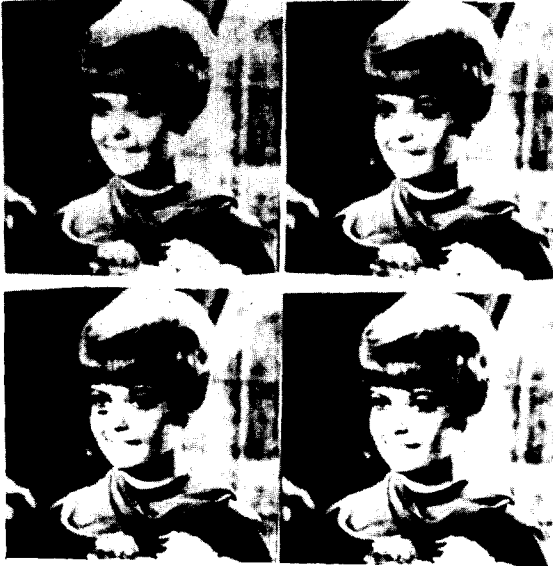


그림13. Zonal coding에 따른 재생 영상 (Data base:Girl)

- (a) 0.5bpp, SNR=48.6dB
- (b) 1.0bpp, SNR=51.5dB
- (c) 1.5bpp, SNR=53.4dB
- (d) 2.0bpp, SNR=55.3dB

VWC에는 여러가지가 있을 수 있지만 source entropy에 가장 가까운 code를 만들 수 있는 방법인 Huffman code를 사용한다. 이 Huffman code를 zero AC계수의 run length를 부호화하는 데 이용되면 더 많은 데이터 감축을 얻을 수 있으며, 이러한 RLC를 할 때 0을 잇달아서 내보내기 위해 block을 scanning하는 방법도 여러가지 있다. 그 중에서 그림11의 zig-zag scanning⁸⁾을 가장 많이 사용한다.

그림 12는 DCT 변환을 이용하여 영상데이터를 감축하는 화상적응 부호화기(SAC:Scene Adaptive Coder)의 block diagram을 나타낸 것이다.⁹⁾ 먼저 작은 블록으로 나뉘어진 입력영상을 2차원 DCT한 후 thresholding 한 후 평균 밝기를 나타내는 F(0,0)는 9비트 PCM을 하고 그 이외의 계수들에 대해서는, VWC와 RLC를 하여 전송한다. 또 buffer 조절은 thresholding 할 때 사용

하는 threshold나 정규화 계수인 D를 변화시킴으로 이루어진다.

변환계수의 통계특성에 따라 부호화 방법을 달리하는 적응식(adaptive) 변환부호화 방식의 대표적인 것은 적응식 block quantization 이다.⁷⁾ 이것은 변환블록을 통계적 특성에 따라 몇 개로 분류(classification) 한 후 비트 할당(bit allocation)을 달리하여 전송하는 방식이다. 즉, activity가 큰 블록에는 많은 비트를 할당하며 activity가 작은 블록에 대해서는 적은 비트가 할당된다. Activity로는 보통 변환계수들의 에너지를 사용한다.

그림 13은 위의 여러가지 변환 방식들 중에서 블록 양자화(block quantization)를 사용한 zonal coding 방식에 따른 재생 영상들로써 16x 16 2차원 DCT를 사용하였으며 Laplacian 특성의 균등 양자기를 사용하였다. 이때 사용영상은 "Girl"을 사용하였으며, 1bpp이하의 낮은 bit rate에서도 좋은 성능을 나타냄을 알 수가 있다.

5. Vector Quantization을 이용한 부호화 방법

최근들어 VQ(Vector Quantization)을 이용한 영상데이터 감축기법에 대한 연구가 활발하게 이루어지고 있다. 이 방법은 개별신호를 각각 부호화하는 것보다 신호들의 조합을 단위로 부호화하는 것이 좋다는 rate-distortion 이론에 근거를 둔 것이다. 그림14는 VQ를 이용

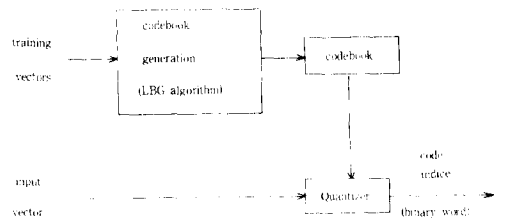


그림14. 일반적인 VQ의 개념도

특집 : 회로 및 신호처리

하여 부호화하는 방법에 대한 개념도를 나타낸 것이다. 신호들의 조합인 한개 vector가 입력되면 codebook에 저장되어 있는 reference vector들 중에서 입력 vector와 가장 유사한 reference vector를 선정하여 그 vector의 index를 전송한다는 것이다.¹⁰⁾ VQ의 성능은 codebook에 따라서 크게 좌우된다. codebook에 저장되어 있는 reference vector를 설계할때 입력될 vector와 그 vector에 해당될 reference vector와의 평균거리가 최소가 되도록 LBG algorithm을 사용한다. 이 때 유사한 성능을 가지는 vector들을 이용하는데 이것을 training vector라고 한다.

VQ를 실제 부호화에 적용하기 위해서는 여러가지 문제점이 있다. 좋은 성능을 위하여 vector의 차수를 높일수록 좋은데 이러한 경우 reference vector의 수가 많아지게 되고 그 결과 입력 vector에 해당하는 reference vector를 찾아내는데 필요한 계산량이 많아진다. 그리고 codebook의 설계시 사용한 training vector와 입력 vector와의 통계특성이 다를 경우 성능저하가 생긴다는 것이다.



그림 15. 일반적인 VQ 방식에 따른 재생영상

(Data base: Girl, Cronleite)

- (a) 0.5bpp, SNR=29.0dB
- (b) 0.5bpp, SNR=31.0dB
- (c) 0.5bpp, SNR=32.3dB
- (d) 0.5bpp, SNR=33.8dB

그림 15는 통계 특성이 다른 여러가지 영상을 training vector로 사용하여 256과 512개의 vector를 갖는 codebook

(0.5bpp, 0.56bpp)들을 각각 설계한후 4 x 4 2차원 VQ를 한 재생 영상들로서 이때 사용된 영상은 "Girl"과 "Cronkite"이며 블럭간의 경계 현상이 심한 것을 알 수가 있다.

이와 같이 VQ를 이용하여 영상을 부호화할 때 생기는 여러가지 문제점들을, 즉 많은 계산량과 memory, 경계 현상 및 불안정성 등을 줄이기 위해 제안된 VQ를 이용한 부호화방법으로 MSVQ(Mean Separated VQ), GSVQ(Gain-Shape VQ), MVQ(Multistage VQ) 등이 있다.¹²⁾ 그림-16은 이들에 대한 간단한 block diagram을 나타낸 것이다.

MSVQ는 입력 vector로 부터 평균을 추출하고 입력 vector와 평균과의 차이를 VQ해서 보내는 방법이다. 그리고 GSVQ는 입력 vector의 크기에 해당하는 feature (예를들면 energy)에 대하여 정규화한 후 VQ하는 방식이다. MVQ방식은 2개이상의 codebook을 직렬로 연결하여 전체의 계산량과 필요한 기억소자의 양을 줄인 것이다.

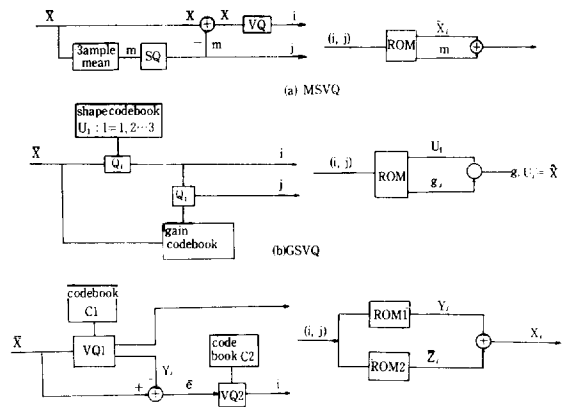


그림 16. 여러가지 VQ를 이용한 감축기법

6. 복합 감축 방식

영상테이타를 좀 더 효율적으로 감축하기 위하여 2절에서 5절까지 살펴본 방법들을 적절히 결합한 여러가지 감축 기법들이 이용되고 있다. 복합 감축 기법은 특

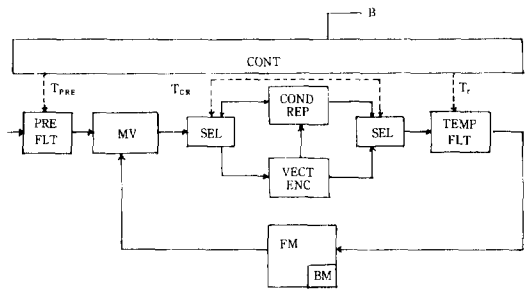


그림 17. BMA-VQ 방식

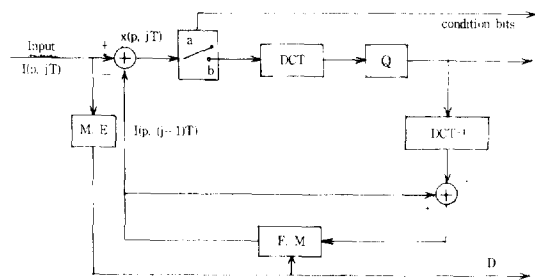
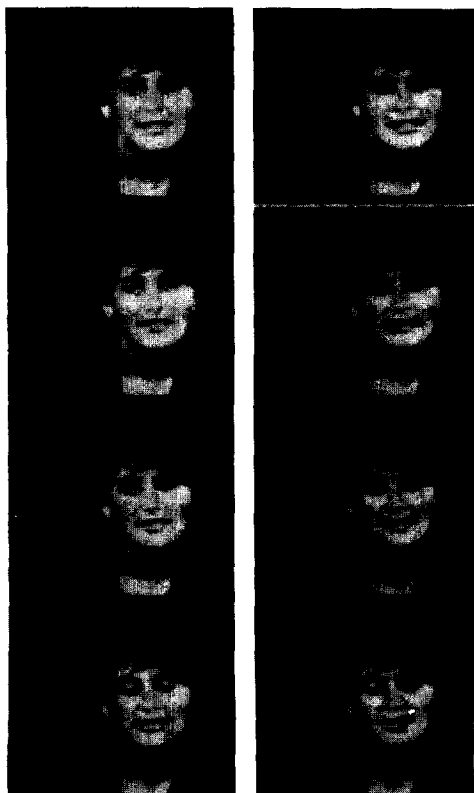


그림 18. 조건부 보충 방식을 이용한 이동 보상형 복합 부호화기



(a) 원영상 (b) 재생영상

그림 20. 그림 18의 방식에 따라 모의실험한 재생영상(Data base:MSAM)

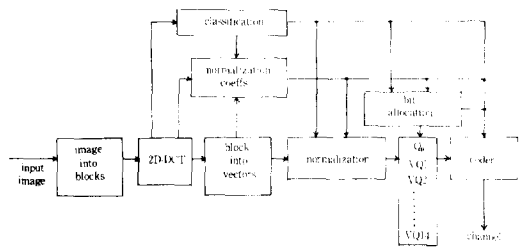


그림 19. DCT를 이용한 adaptive VQ의 송신단 구성도

이 영상전화 혹은 원거리 영상회의시스템을 위한 64Kbps 급 video codec에 많이 이용되고 있는데, BMA-DCT, BMA-VQ, 혹은 DCT-VQ 방법들이 주류를 이루고 있다.

그림 17은 KDD에서 개발한 INVITE-64의 block diagram을 나타낸 것으로 BMA-VQ방식을 채택하고 있다.¹²⁾ 먼저 BMA로서 입력신호를 예측하고 예측오차를 VQ 한다. 여기에서 CRC를 추가한것은 보다 감축효율을 높이기 위한 것이다.

그림 18은 BMA-DCT 방법을 사용한 영상감축시스템의 block diagram 이다.¹⁴⁾ INVITE-64에서와 마찬가지로 입력신호의 예측에서 BMA를 사용하고, 그 예측오차를 DCT하여 부호화하는 방식이다. 부호화 방식으로 앞절에서 논의한 VWC를 많이 사용한다. 그림19는 DCT-VQ 방식의 구성도이다.¹⁵⁾ 먼저 2차원 DCT를 수행하고 그 결과에 따라서 block을 몇개의 class로 나누어 각 class마다 coding을 달리하는 adaptive VQ 방식이다.

위 방식들 중에서 그림 18의 MCT(Motion-Compensation + Transform) 방식은 현재 과학기술원에서 개발 중인¹⁶⁾¹⁷⁾ 64Kbps급 codec system에 응용되고 있으며, 모의 실험 결과 그림20와 같은 재생 영상을 얻을 수가 있었다. 사용된 영상은 "Miss America"(MSAM)로써, 칼라 영상을 사용하였으며, 112(line) x 128(pel) 크기로 windowing한 후 초당 15장씩 전송하도록 하였다. 이때

특집 : 회로 및 신호처리

설계 전송되는 pel당 평균 비트수는 0.126bpp정도가 되며, 평균 SNR은 Y에 대해 약 33-34dB, R Y와 B-Y에 대해서는 30-32dB, 25dB 정도로 나타났다.¹⁷⁾

7. 맺음말

현재까지 발표되어 있는 영상신호의 대역폭 감축기법들 중에서 기본적인 방법들에 대하여 살펴 보았다. 점차 구체화되고있는 종합정보통신망(ISDN)에서의 주요 정보원으로 음성, 데이터 및 영상정보를 들수 있는데, 그 중에서도 영상정보는 정보의 전달효과가 클 뿐 아니라 전송해야 하는 데이터량도 또한 막대하기 때문에 ISDN의 표준화 과정에서 매우 주요한 위치를 차지하고 있다. 현재 단계에서 ISDN에서의 기본 access는 64-Kbps급으로 될 전망이기 때문에 영상정보를 수용하기 위해서는, 영상데이터를 대폭 감축하는 것이 필요하다.

외국에서는 이를 위한 시스템들이 이미 개발되어 비싸긴하지만 실용화 단계에 있고, 뿐만아니라 CCITT에서도 이들에 대한 표준화 작업이 한창 진행중에 있다.

앞으로 국내통신의 선진화를 이루기 위해서는 ISDN의 광역화를 조기에 성취해야 하고, 이를 촉진하기 위해서는 국내에서의 영상정보를 위한 video codec의 개발 연구가 적극적으로 활성화 되어야 하겠다.

참 고 문 헌

- 1) 박근수, 45Mb/s TV 코덱시스템의 성능개선에 관한 연구, 한국과학기술원 석사학위논문, 1984.
- 2) Yasuda, et al, "Transmitting 4MHz TV signals by combi-national difference coding," IEEE, COM-25, pp 508-516, 1977.
- 3) B.G.Haskell, "Interframe coding of 525-line, monochrome television at 1.5 Mbit/s," IEEE, COM-25, Nov, 1977.
- 4) 장주옥, "CRC 와 MCC를 이용한 1.5Mbps 비데오코덱에 관한 연구," 한국과학기술원 석사학위논문, 1985.
- 5) 정주홍, "이동벡터 검출을 위한 block matching 방법에 관한 연구", 한국과학기술원 석사학위논문, 1986.
- 6) Netravali, and Robbins, "Motion-compensated Television Coding : Part I" BSTJ V-58 pp. 629-668, 1979.
- 7) W.K.Pratt, "Digital Image Processing," Wiley-interscience pp 662-709, 1977.
- 8) 전준현, "DCT 고속 알고리즘의 하드웨어 구현을 위한 성능 개선 및 설계에 관한 연구", 석사학위논문, 한국과학기술원, 1985.
- 9) W.H.Chen, W.K.Pratt, "scene Adaptive Codec", IEEE COM-32, No-3, March, 1984.
- 10) R.M.Gray "Vector quantization," IEEE ASSP Magazine vol-1, pp 4-29, April 1984.
- 11) Y.Linde, A.Buzo, M.M.Gray, "An algorithm for vector quantization ," IEEE COM-28, pp 84-95, Jan. 1980.
- 12) 박종윤, "선형예측 벡터 양자화를 이용한 영상부호화 방법", 한국과학기술원 석사학위논문, 1987.
- 13) H.Yamaguchi, M.Wada, and H.Yamamoto, "A 64 kbit/s Integrated Visual communication System New Communication Medium for the ISDN," IEEE SAC-4, No.8, pp 1202-1208, Nov. 1986.
- 14) CCITT SGXV, Report of fifth meeting in Tokyo, July 1986.
- 15) T.Saito et al, "Adaptive DCT image coding using gain/s-hape vector quantizer," ICASSP 86, pp. 129-132.
- 16) "A study on the development of iamge data cpmression fechnique", 연구보고서, 한국과학기술원 1986.
- 17) "A study on the development of video image coding tec-hnique", 연구보고서, 한국과학기술원 1987.