

영상 통신 기술

김 종 국*, 박 현 옥**

(*한국과학기술원 정보 및 통신공학과 박사과정

**한국과학기술원 정보 및 통신공학과 조교수)

1. 서 론

오늘날 사회가 고도 정보화되어감에 따라 통신 서비스의 요구가 다양화 되어 음성뿐만 아니라 텍스트, 그래픽, 그리고 영상 데이터와 같은 비음성계의 통신에 대한 관심이 고조되고 있다. 차세대 정보 통신 분야의 꽃이라고 할 수 있는 통합 멀티미디어 시스템은 바로 이러한 데이터들을 처리해야 하는데 특히, 영상정보는 디지털화 했을때 많은 데이터량을 갖기 때문에 디지털 영상부호화기술은 멀티미디어를 지원해 주는 핵심 요소기술이다.

영상 부호화 기술이란 영상신호들이 갖고 있는 공간적(Spatial) 또는 시간적(Temporal) 중복성을 제거함으로써 영상을 표시하는데 필요한 데이터량을 줄이는 것이다. 영상 부호화 기술은 오래전부터 연구되어 왔고 부호화기술이 계속 발전하여 실제 제품화됨에 따라 국제 표준을 제정할 필요성이 절실하게 대두되어 ITU-T(International Telecommunication Union - Telecommunication Standardization Sector)와 ISO(International Organization for Standardization)의 JTC1중심으로 다양한 전문가 그룹을 결성하여 여러가지 표준들을 제정했거나 제정중에 있다.

영상부호화 관련 표준화는 정지영상 부호화의 표준방식인 JPEG(Joint Photographic Experts Group : ISO/IEC JTC1/SC29/WG10)방식 개발을 시작으로 영상회의 및 영상전화용인 ITU-T

의 H. 261이 완성되었고 디지털 기록매체뿐만 아니라 통신 방송용으로 이용될 수 있는 동영상 부호화 표준인 MPEG1,2(Moving Picture Experts Group : ISO/IEC JTC1/SC29/WG11)방식이 완성되었다. MPEG1은 CD-ROM과 같은 매체를 대상으로 한 1.5Mbps급의 동영상 부호화 표준이며 MPEG2는 MPEG1을 기반으로하여 응용분야가 광범위하며 당초 10Mbps까지의 전송률을 갖는 알고리즘으로 출발했으나 10Mbps의 제한이 없어짐으로써 MPEG3단계가 없어지게 되었다. 현재, MPEG에서는 64Kbps이하의 초저속 채널에서도 영상서비스가 가능한 MPEG4에 대한 표준화 노력이 집중되고 있다. 또한, MPEG4그룹은 ITU-T SG15/LBC와 밀접한 협력아래 표준안을 진행중이다.

본고에서는 이러한 영상 통신 서비스를 위하여 이용되고 있는 여러가지 영상 부호화 기법에 관하여 간단히 논하고 현재 제정된 국제 표준안 중에서 JPEG과 MPEG에 대하여 간단히 설명하고자 한다.

2. 영상 부호화 기법

아날로그 영상을 디지털 영상으로 변환하기 위해서 표본화(Sampling)과정과 양자화(Quantization)과정이 요구된다. 일반적인 디지털 칼라 TV영상의 구성은 720×480개의 화소구성에 화소당 16bits가 소요되는데 1초에 30프레임을 전송할 경우 약 170Mbps의 데이터 전송률이 된다. 이러한 방대한

양의 데이터를 기존의 통신로를 이용하여 전송하거나 기록매체에 저장하기 위해선 데이터량의 감축이 필요하다.

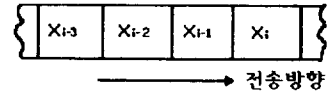
본 절에서는 표준화 및 일반적인 코덱시스템에 널리 이용되고 있는 여러가지 영상 부호화기법의 원리를 간략히 설명하고 또한 초저속 영상통신을 위한 동영상 부호화 기법으로 연구되고 있는 제2세대 부호화 방법인 객체지향 부호화(Object-Oriented Coding), 웨이브릿 기반 부호화(Wavelet-Based Coding) 그리고 프랙탈 부호화(Fractal Coding)에 대해서도 언급하고자 한다.

2.1. 예측 부호화(Predictive Coding)[1,2]

예측부호화는 이미 전송된 화소치로부터 현재의 화소치를 예측하고 현재의 신호값과 예측치와의 차(예측오차)를 부호화하는 방법으로써 인접화소간의 강한 상관성을 가지는 영상신호 특유의 성질을 이용하는 기법이다. 예측부호화는 원신호와 예측치와의 차분치를 전송하는 것이므로 DPCM(Differential Pulse Code Modulation)이라고 부르기도 한다. 차분치의 엔트로피가 원신호의 엔트로피에 비해 훨씬 작기 때문에 차분치에 적당한 부호화당을 행하면 평균 부호길이를 줄일 수 있다. 예측기(Predictor)는 부호화 라인상의 화소만 이용하는 1차원 예측과 종방향으로도 상관성이 있으므로 전송된 라인상의 화소도 이용하는 2차원 예측으로 나눌 수 있으며 몇가지 예를 그림 2.1에 나타냈다.

일반적인 TV화상은 0~255 까지의 값을 가지며 예측오차로서 출현가능한 값은 -255~255까지가 된다. 따라서 부호어로 511개를 준비해야 하므로 예측오차를 양자화하여 전송할 값의 종류를 한정하고 있다. 이렇게 하면 오차가 발생하는데 이 오차를 양자화 오차(Quantization Error)라 한다. 양자화기(Quantizer)는 구현을 쉽게 하기 위해 구간이 일정한 양자화기(Uniform Quantizer)를 사용할 수 있으나 예측오차들이 대체적으로 적은 값을 갖는 통계적 특성을 이용하여 적은 오차에서 세분된 구간을 갖고 큰 오차에서 넓은 구간을 갖는 양자화기(Non-uniform Quantizer)를 사용함으로써 양자화에 의한 오차를 더욱 줄일 수 있다. 양자화를 포함한 예측부호화의 블럭도를 그림 2.2에 나타냈다.

동영상에 예측부호화 기법을 적용할 경우 영상이



$$\begin{aligned} X_i & \dots \text{부호화 대상화소} \\ \hat{X}_i &= X_{i-1} \\ \hat{X}_i &= 2X_{i-1} - X_{i-2} \end{aligned}$$

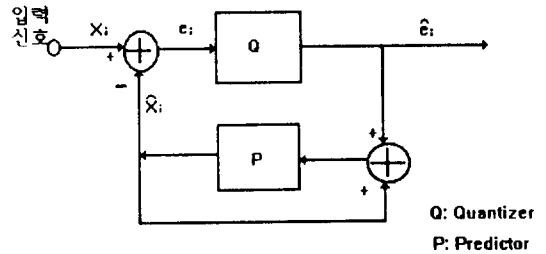
(a) 1차원 예측

$X_{[i-1, j-1]}$	$X_{[i, j-1]}$	$X_{[i+1, j-1]}$
$X_{[i-1, j]}$	$X_{[i, j]}$	$X_{[i+1, j]}$

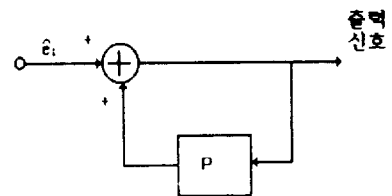
$$\begin{aligned} X_{[i, j]} & \dots \text{부호화 대상화소} \\ \hat{X}_{[i, j]} &= 1/2 [X_{[i-1, j]} + X_{[i, j-1]}] \\ \hat{X}_{[i, j]} &= 1/2 [X_{[i-1, j]} + X_{[i+1, j-1]}] \\ \hat{X}_{[i, j]} &= X_{[i-1, j]} - X_{[i-1, j-1]} + X_{[i, j-1]} \end{aligned}$$

(b) 2차원 예측

그림 2.1 1차원 예측과 2차원 예측



(a) 부호기(encoder)



(b) 복호기(decoder)

그림 2.2 예측 부호화의 블럭도

갖고 있는 시간상의 중복성을 줄이기 위해 전송된 프레임의 화소를 참조화소로서 이용하는 프레임간(Interframe) 예측부호화기법을 이용할 수 있다.

한가지 예를 그림 2.3에 나타내었다.

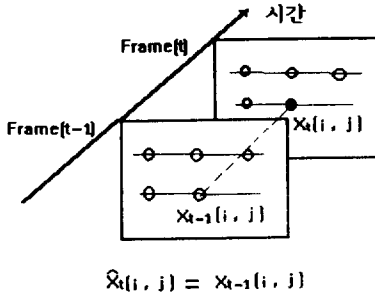


그림 2.3 프레임간 예측

프레임간 예측 부호화는 시간상의 위치 화소를 예측치로 하여 차신호를 구하고 이것을 양자화한 후 전송하는 방식으로써 움직임이 많은 동영상일 경우 예측오차의 증가로 예측 성능이 많이 떨어지는 문제점이 있다. 이러한 문제점을 보완하기 위해 이웃한 영상간의 움직임을 검출하여 보상해주는 움직임 보상 예측부호화(Motion Estimation/Compensation Predictive Coding)기법이 이용되고 있다. 그림 2.4는 그것에 대한 블록도를 나타내고 있다.

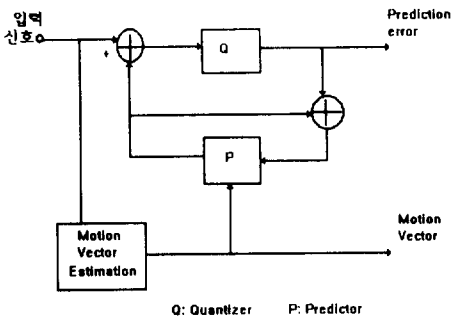


그림 2.4 움직임 보상 예측부호화

움직임 보상 예측부호화는 인접한 영상과 현재의 영상을 비교하여 물체의 움직임 벡터(Motion Vector)를 구하고 이 움직임 벡터를 이용하여 현재의 영상을 예측하는 방식이다. 움직임 벡터 검출방법으로는 블록 정합 알고리즘(Block Matching Algorithm)과 화소 순환 알고리즘(Pel Recursive Algorithm)이 있으며 국제 표준화 그룹(H. 261, MPEG1, MPEG2)에서는 성능은 다소 떨어지나 구현이 간단한 BMA방식을 채택하고 있다. BMA

방식을 이용한 움직임 보상기법은 영상을 여러개의 블록으로 나누어 현재 부호화되는 블록과 가장 오차가 적은 블록을 인접 영상에서 찾아서 그 블록을 현재 부호화되는 블록의 예측으로 이용한다. 움직임을 찾는 방법으로는 전체 탐색(Full Search), 계층 탐색 방법(Hierarchical Search Algorithm), TSS(Three Step Search)등이 있다. BMA방식을 이용한 움직임 보상 부호화는 복호기에 움직임 벡터에 대한 정보를 전송해 주어야 하는데 부호화 효율을 증대시키기 위해 먼저 전송된 움직임 벡터와의 차를 부호화하여 전송하는 방식을 국제 표준화 그룹에서는 채택하고 있다.

2.2. 변환 부호화(Transform Coding)[1]

변환 부호화는 영상 신호를 블록으로 분할하여 블록내의 상관관계를 변환에 의해 제거하고 변환영역에서의 에너지를 집중시키므로 변환 계수들을 적절히 양자화함으로써 부호화 효율을 증가시키는 부호화 기법이다. 그림 2.5는 변환부호화의 간단한 블록도를 나타낸다.

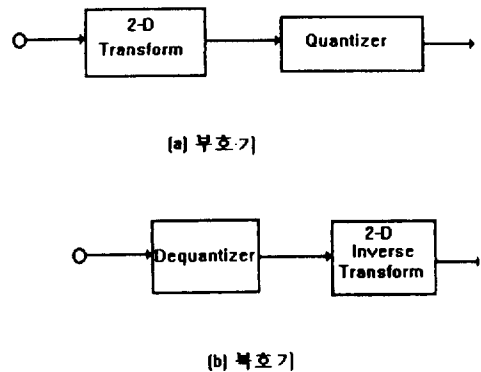


그림 2.5 변환 부호화 블록도

상관관계를 완전히 제거하는 최적의 변환은 KLT(Karhunen Loeve Transform)이나 실시간 H/W구현이 어렵고 전송될 영상신호의 통계적 특성을 미리 알아야 하는 단점등으로 실제 사용되지 않는다. DCT(Discrete Cosine Transform)는 성능이 거의 KLT와 비슷하며 하드웨어 구현이 용이하고 실시간 처리가 가능하며 에너지 집중효율이 높아 압축효율측면에서도 우수하기 때문에 모든 표

준화가 DCT 변환을 이용하고 있다. 일반적으로 부호기의 입력에서 원영상을 8×8블럭으로 분할하고 각 블럭마다 DCT변환을 수행한다.

2-D DCT은 다음과 같은 수식으로 수행된다.

$$F(u,v) = \frac{1}{4}C(u)C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \cos\left(\frac{(2x+1)u\pi}{16}\right) \cos\left(\frac{(2v+1)v\pi}{16}\right) \quad (1)$$

여기서, x,y = 블럭내의 화소 위치
u,v = DCT 계수의 위치

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}}, & \text{for } u,v=0 \\ 1, & \text{otherwise} \end{cases}$$

그 외의 변환기법으로는 Hadamard, Slant, Harr등이 있다.

2.3 벡터 양자화(Vector Quantization)[3]

벡터 양자화 기법은 일련의 데이터로 구성되는 벡터를 양자화하여 부호화하는 기법으로 각 데이터를 개별적으로 양자화(Scalar Quantization)하여 부호화하는 기법보다 데이터량을 더 줄일 수 있는 잇점이 있다. 또한 벡터양자화는 디코더 구현이 간단하기 때문에 Single-Encoder, Multiple-Decoder 응용분야에 효과적으로 이용될 수 있다. 기본적인 벡터양자화기 블럭도는 그림 2.6과 같다.

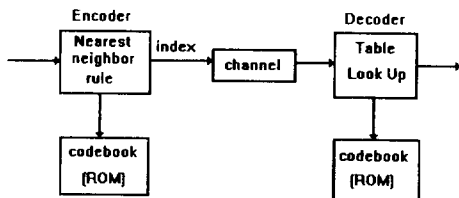


그림 2.6 벡터 양자화 블럭도

벡터 양자화 기법은 일반적으로 그림 2.6과 같이 부호책(Codebook)과 부호기 그리고 복호기로 구성된다. 부호책은 대표되는 벡터들의 집합이며 Linde, Buzo 및 Gray에 의해서 제안된 실험적인 설계방법인 LBG알고리즘에 의해 구한다. 부호기는 입력되는 벡터와 같거나 가장 근사한 대표벡터를

부호책에서 찾아내고 이 벡터를 가리키는 지수(Index)를 전송한다. 복호기는 수신된 지수에 해당되는 대표 벡터를 출력한다. 이와 같이 개념적으로 간단한 벡터 양자화 기법은 많은 영상정보량으로 대표벡터 선정과 갯수가 문제가 된다. 특히 대표 벡터 갯수가 늘어나면 부호책이 커져 부호기에서 대표벡터 탐색이 실시간에 수행되지 못하는 문제가 있다. 따라서 계산 복잡도 및 메모리 요구사항등을 줄일 수 있는 다양한 방식에 관한 연구가 많이 진행되고 있으며 다른 부호화 기법의 결합된 벡터 양자화가 많이 연구되고 있다.

2.4.대역분할 부호화(Subband Coding)[4]

대역분할 부호화의 기본 개념은 QMF(Quadrature Mirror Filter) बैं크를 사용하여 입력신호의 주파수대역을 N개의 Subband신호로 분리한 후 각 Subband신호는 그 주파수 영역에서 인간의 시각 특성에 적합한 부호화 기법을 적용하여 압축을 하는 기법이다. 예를 들면, 신호를 그림 2.7과 같은 응답을 가지는 Analysis Bank로 2채널 대역분할할 수 있다.

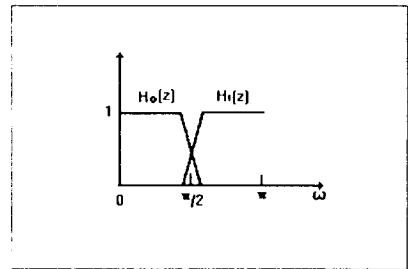


그림 2.7 Analysis 필터의 주파수 응답

분할된 Subband 신호는 1/2로 간축(Decimation)되어 에너지 내용에 따라 각 Subband신호에 적절한 비트수를 할당한다. 이 경우는 $x_1(n)$ 이 $x_0(n)$ 보다 더 작은 에너지를 가지므로 $x_0(n)$ 보다 더 적은 비트수를 할당할 수 있다. 신호의 재구성은 보간(Interpolation)과 Synthesis Bank를 통해 복원된다. 그림 2.8은 일반적인 2채널 대역 분할 부호기의 블럭도이다.

이 방법은 낮은 비트율에서 다른 방법에 비해 전력현상이 현저히 줄며 각 분할된 subband별로 가장 적합한 부호화 방법을 적용하기가 용이하며 하드웨

어 구현시 각 밴드에 대한 병렬처리가 가능하다.

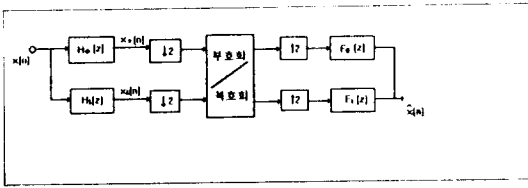


그림 2.8 2채널 대역분할 부호화의 블록도

2.5. 엔트로피 부호화(Entropy Coding)[1,2]

엔트로피 부호화는 가변장 부호화(Variable Length Coding)라 부르기도 하며 Huffman 부호화와 산술 부호화로 나뉘어진다. Huffman 부호화가 가장 많이 사용되며 기본 개념은 발생확률이 높은 신호에 대해서는 짧은 부호어를, 낮은 신호에 대해서는 긴 부호어를 할당하여 평균 부호어 길이를 짧게 하는 것이다. 그림 2.9는 심볼이 9개인 Huffman 부호화의 작성예를 나타내고 있다.

심볼	확률	순서	부호어
0	0.49	0	0
1	0.15	0 0	100
2	0.14	1 0	101
3	0.07	0 1 0	1100
4	0.06	1 1 0	1101
5	0.04	0 1 1	1110
6	0.02	0 1 1 1	11110
7	0.02	1 1 1 1	111110
8	0.01	1 1 1 1 1	111111

그림 2.9 Huffman 부호어 작성 순서

우선 심볼의 출현확률이 낮은 2개를 선택하여 그것에 1 또는 0의 부호를 할당하고 나서 그것을 통합하여 하나의 심볼로 한다. 그것에 따라 심볼수가 8개로 된다. 다음에 8개의 심볼을 다시 출현확률 순에 따라 출현 확률이 낮은 2개를 선택하여 0 또는 1을 할당한다. 그와 같은 처리를 심볼수가 1이 될 때까지 반복한다. 할당된 부호를 역순으로 호출하여 그것에 대응하는 심볼의 부호어(Codeword)로 한다.

2.6 웨이브렛 기반 부호화

(Wavelet-based Coding)[4,5]

Wavelet이론은 응용수학, 신호처리 그리고 컴퓨터 비전분야에서 연구되어 온 필터 뱅크(Filter Bank), 다중해상도 신호해석이론과 연관성이 있음이 밝혀졌다. Wavelet이란 h(t)로 정의되는 Mother Wavelet을 변이(Translation) 그리고 확대/축소(Dilation/Contraction)시킴으로써 얻어지는 기저함수(Basis Function)들의 집합을 말하며 어떤 함수 x(t)의 연속 Wavelet변환은 다음과 같이 정의된다.

$$Xw(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)h\left(\frac{t-b}{a}\right)dt \quad (2)$$

여기서 a,b는 실수(a,b ∈ R, a ≠ 0)이다. a = a₀m, b = na₀m₀일때 이산 Wavelet변환이라 하며 정의식은 다음과 같다.

$$Xw(m,n) = a_0^{m/2} \int_{-\infty}^{\infty} h(a_0^{-m}t - nb_0)x(t)dt \quad (3)$$

여기서 m, n은 정수, a₀>1, b₀ ≠ 0이다. 특히, a₀ = 2, b₀ = 1 일때를 Dyadic Wavelet변환이라고 한다.

그림 2.10은 Wavelet변환의 시간-주파수 해상도(Time-Frequency Resolution)를 나타내며 고주파수에서는 시간 해상도가 좋고 저주파수에서는 주파수 해상도가 좋다. 따라서 Nonstationary한 영상 신호처리에 적합하다고 볼 수 있다. 그림 2.11은 이산 웨이브렛 변환에 대한 블록도를 나타내며 Treestructured QMF구조와 같다. 그림 2.12는 웨이브렛 변환의 주파수 특성을 나타냈다.

웨이브렛 기반 부호화는 상기와 같은 웨이브렛이론을 근간으로 밴드마다 인간의 시각 특성을 고려한 기존의 여러 부호화 알고리즘과 접목하여 연구가 되고 있으며 이산 웨이브렛 변환 영역에서 벡터 양자화를 이용하는 방법이 주종을 이루고 있다.

2.7 객체 지향 부호화

(Object-Oriented Coding)[6]

순차적인 동영상을 저속으로 전송하기 위한 Low Bit Rate Coding의 일종으로 일반적으로 움직임이 적은 부분으로 국한된 Video Phone이나 Video Conferencing에 사용한다. 각 영상들은 균일한 움

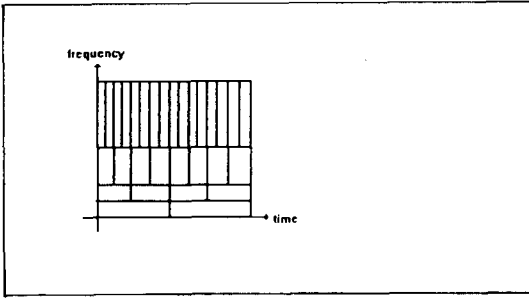


그림 2.10 Wavelet 변환의 시간 주파수 공간도

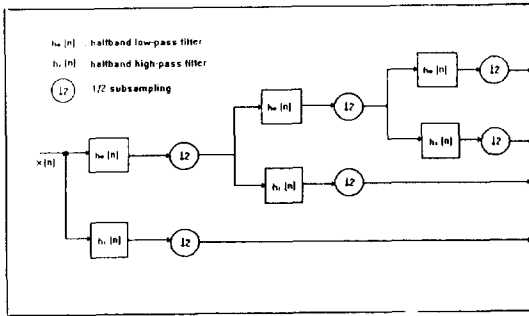


그림 2.11 Wavelet 변환 블럭도

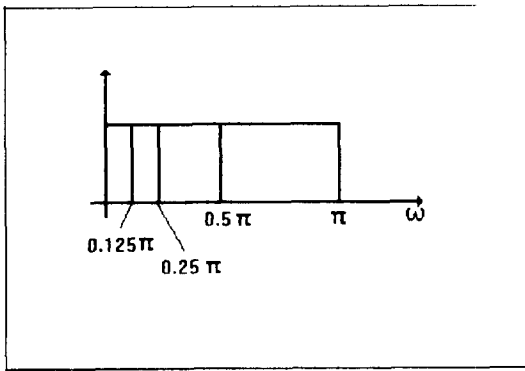


그림 2.12 Wavelet 변환의 주파수 특성

김을 갖는 Moving Object로 분할되어 이러한 각 Object들을 기술하기 위한 움직임(Motion, M_k), 형태(Shape, S_k) 및 색상(Color, C_k)의 세가지 변수(Parameter)로 나타낸다.

Object-Oriented Analysis Synthesis Coder는 영상분석부, 변수의 부호화/복호화부, 영상 합성부, 변수를 위한 메모리의 다섯 부분으로 구성된다. 영상분석부는 순차적인 동영상에서 현재의 영상과 이전의 부호화/복호화과정을 거친 합성영상

으로부터 각 Object로 분할하여 각 Object의 성질에 따라서 해당 변수를 찾아낸다. 즉 이전 프레임에서 움직임으로만 표현될 수 있는 경우(MC-Objects: Motion Compliance Objects)에는 A_{k+1} , M_{k+1} 만을 추출하고, 이전 프레임으로 부터 변화된 부분(MF-Objects: Motion Failure Objects)에서는 M_{k+1} , C_{k+1} 로만 Object를 표시한다. 변수의 부호화/복호화부는 A_{k+1} , M_{k+1} 에 대해서는 이전 Frame으로부터의 예측부호화를 수행하여 시간적 중복성을 압축시키고 S_{k+1} 은 Object의 형태에 따른 부호화를 수행한다. 수신단의 영상 합성부는 수신된 변수(Parameter)로부터 복호화, 이전 Frame으로 부터의 보상을 통하여 현재의 Frame을 재생하고, 송신단의 영상합성부는 송신한 변수로부터 이전 송신 Frame을 보상하여 다음 Frame의 영상 분석부를 위한 기준 화면으로 사용한다. 송수신단의 변수를 위한 메모리는 동일한 변수들을 가지고 있게 된다. Object-Oriented Analysis Synthesis Coding은 위와 같은 방법으로 구성되어 블럭 변환 부호화(Block Transform Coding)시 발생하는 블럭현상, Mosquito현상등의 Degradation대신 Geometric Distortion이 발생되어 인간의 시각에 효율적이다. 또한 이전 Frame과의 변경부분을 적은 부분으로 분할하여 자세히 전송함으로써 화질이 개선된다. 기본적인 블럭도는 그림 2.13과 같다.

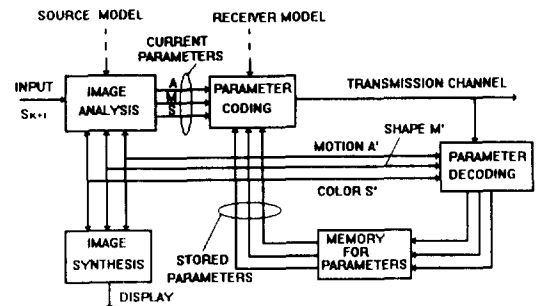


그림 2.13 Object-based 부호화의 블럭도

2.8 프랙탈 부호화(Fractal Coding)[7]

프랙탈은 1970년 중반에 Benoit Mandelbrot에 의해 그 이름이 명명되었다. 프랙탈은 간단하게 결정된 알고리즘에 의해 표현할 수 있는 적은 정보량을 가질지라도 대단히 복잡한 시각 특성을 갖는다.

이들은 자신의 일부 또는 그들 자신 각각의 변환된 복사본에 의해 구성된다는 점에서 중복된 형상을 갖는다. 직선 또는 부드러운 곡선과 같은 종래의 기하학적인 형태보다 프랙탈이 해안선, 구름, 낙엽, 산과 같은 자연 영상을 더 잘 묘사할 수 있다는 인식때문에 프랙탈 영상 부호화 기법이 수년간 연구되어 왔다. 프랙탈의 특징은 다음과 같다.

- i) 미세구조(Fine Structure)를 갖는다.
- ii) 자기 유사성(Self Similarity)를 갖는다.
- iii) 프랙탈의 부분적인 차원(Fractional Dimension)은 Euclidean 차원 보다 더 높다
- iv) 프랙탈은 순환적으로 정의된다.

프랙탈 영상부호화는 어떤 영상에 대해 IFS(Iterated Function System)의 끌개(Attractor)를 추정해야 한다. IFS는 여러개의 축소 의사변환(Contractive Affine Transform)을 모아 놓은 것이다. 어떤 영상에 대해 IFS를 찾는 방법은 두가지가 있다.

첫째, 영상의 자기 유사성에 입각한 영상으로부터 직접 찾는 방법과

둘째, 영상을 더 작은 영상으로 쪼개서 IFS 라이브러리에서 각 분할 영상에 대한 해당 IFS를 찾는 것이다.

프랙탈 영상 부호화가 기반으로 하는 기본적인 정리는 축소의사 변환과 폴라쥬정리이다.

축소의사 변환 정리(Contractive Affine Transform Theorem)

2차원 의사 변환 W는 다음과 같은 공식으로 주어지며,

$$W \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix} \quad (4)$$

변환 W가 임의의 두점 P1,P2에 대해 Distance가 다음 조건을 만족하면 축소(Contractive)라 말한다.

$$d(W(P1), W(P2)) < sd(P1,P2)$$

여기서 $0 < s < 1$

즉 축소 의사 변환을 반복 적용하면 고정된 점에 수렴하게 된다.

폴라쥬 정리(Collage Theorem)

B를 최종 영상이라고 하고 C를 축소인자 $0 < s <$

1를 만족하는 IFS 코드라 했을 때 B와 C(B)사이의 Hausdorff Distance가 E미만이면 그때, B와 Attractor(C)사이의 Hausdorff는 $E/(1-s)$ 미만이다. 여기서 임의의 두 영상에 대한

$$\text{Hausdorff Distance } H(A,B) = \max\{d(A,B), d(B,A)\}$$

이다.

이 정리는 B와 C(B)사이의 Hausdorff Distance가 매우 작도록 IFS code C를 찾을 수 있으면 C의 Attractor은 최종 영상 B에 근접하게 된다는 것이다.

프랙탈 부호화는 실시간 처리가 어려운 단점이 있으나 영상의 화질 및 고압축률에는 유리하기 때문에 멀티미디어 분야에서는 기존의 부호화 방법보다 낫다는 것을 알 수 있다. 마이크로 소프트의 예를 들면 현재 엔카르타(Encart) 멀티미디어 사진에 프랙탈 영상 부호화를 사용하고 있다.

3. 국제표준 부호화 방식

본절에서는 국제표준 부호화 방식인 JPEG표준 방식과 MPEG표준방식에 대해 간단히 서술하기로 한다.

3.1 JPEG(Joint Photographic Experts Group)표준

JPEG은 정지영상을 압축하는 것을 목적으로 ISO/IEC JTC1/SC29/WG10에서 1992년 국제표준으로 제정했다. JPEG의 기능은 다음과 같이 3가지로 구분되며

- ① DCT-Based 부호화 알고리즘
- ② 무손실(Lossless) 부호화 알고리즘
- ③ 계층적(Hierarchical) 부호화 알고리즘

동작모드로는 8×8크기의 블럭영상을 상단 좌측에서 우측으로 위에서 아래 순으로 부호화하는 순차(Sequential)재생과 영상을 점진적으로 개선하여 재구성하는 점진(Progressive)재생 방식으로 나뉘어 표준화가 이루어졌다.

3.1.1 DCT-Based 부호화 알고리즘

그림 3.1은 DCT-Based부호기와 복호기 처리과

정을 나타낸다.

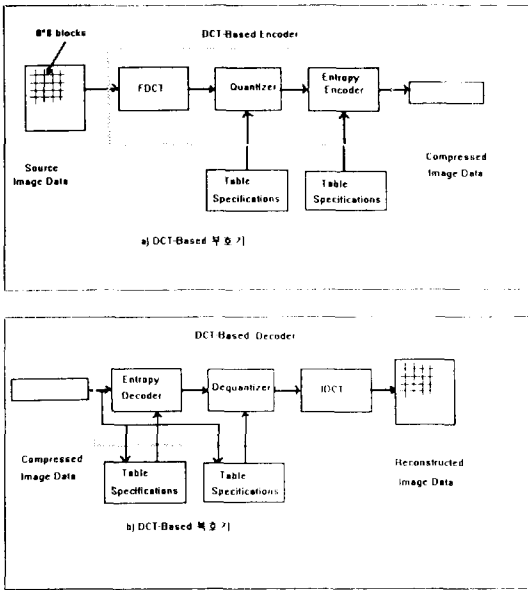


그림 3.1 DCT-Based 부호기/복호기 블록도

입력영상은 8×8의 블록으로 분할되며 그 블록단위로 (1)식의 DCT(Discrete Cosine Transform)을 수행한다. 64개의 DCT계수중 F(0, 0)를 DC계수라 하고 나머지를 AC계수라 한다. 반대로 복호기에서는 IDCT(Inverse DCT)변환에 의해 8x8 블록영상이 복원된다. DCT변환계수는 양자화 테이블을 이용하여 계수 위치마다 다른 스텝크기로 선형 양자화(Linear Quantizer)된다. JPEG에서 권고한 양자화 테이블은 그림 3.2에 주어져 있다.

전체 부호화 시스템의 화질 및 부호화 정보량은 양자화 테이블 값을 변화시킴에 따라 제어할 수 있다. 양자화 테이블값은 양자화 잡음에 대한 인간의 눈 특성을 고려하여 설정할 필요가 있다. 즉, 인간의 눈이 민감한 주파수(저주파)에 대응하는 DCT 계수 값에 대해서는 작은 스텝 크기로, 눈에 둔감한 주파수(고주파수)에 대응하는 DCT계수값은 큰 스텝크기를 할당한다.

DCT와 양자화를 거친 신호는 마지막 단계로 엔트로피(Entropy)부호화 과정을 거친다.

JPEG에서는 엔트로피 부호화로 호프만(Huffman)부호화와 산술(Arithmetic)부호화 두 가지를 지원하고 있다. 산술부호화는 호프만 부호화에

Luminance quantizer table

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	58	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Chrominance quantizer table

17	18	24	47	66	99	99	99
18	21	26	66	99	99	99	99
24	25	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

그림 3.2 JPEG 권고 양자화 테이블

비해 5-10% 압축율이 높지만 구현하기가 복잡하다. 양자화된 DCT계수는 DC계수와 AC계수를 별개로 호프만 부호화하고 있다.

DC계수는 인접블록의 DC계수값과 상관성이 있으므로 현재 블록의 DC계수값과의 차이를 호프만 부호화 테이블로부터 코드를 할당한다. JPEG에서는 휘도성분, 색차성분 각각 독립적으로 DC 호프만 코드 테이블을 권고하고 있다. DCT변환 부호화는 DC를 포함한 낮은 주파수 계수에 에너지를 집중시킨다. 즉 AC계수중 저주파에 해당하는 계수들은 비교적 Nonzero값을 갖는데 비해서 고주파에 해당하는 계수들은 Zero의 값을 가질 확률이 높다. 따라서 DCT변환계수를 호프만 부호화 하기전에 가능한 0이 연속되게 해 줄 경우 부호화 효율을 증대시켜줄 수 있다. 이러한 과정을 위해 양자화한 DCT계수값은 그림 3.3과 같이 Zig-Zag스캔 방식을 채택하고 있다. Zig-Zag스캔된 AC계수값들은 Zero-Runlength의 효과와 Nonzero값에 따라 AC 호프만 부호화 테이블로부터 코드를 할당한다.

그 외에 제어코드로서 각 블록의 최종계수값이 0일 경우에만 전송하는 EOB(End of Block)코드와 Zero-Runlength가 15를 초과했을 때 전송하는 ZRL(Zero Run Length)코드가 있다.

3.1.2 무손실(Lossless)부호화 알고리즘

그림 3.4는 무손실 부호화의 블록도를 나타낸다. 예측기는 인접한 3개의 화소값으로부터 현재의 값을 예측한다. 현재의 화소와 주변화소로 구한 예측치와의 차를 엔트로피 부호화한다. 압축률은 2:1이다. 엔트로피의 부호화는 호프만 또는 산술부호화를 이용한다.

3.1.3 계층적(Hierarchical) 부호화 알고리즘

입력영상의 해상도를 가로, 세로 각각 그림 3.5와 같이 피라미드 모양으로 축소한다. 가장 작은 영상

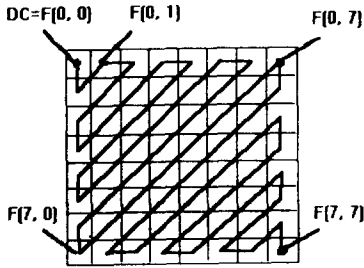


그림 3.3 Zig-Zag 스캔

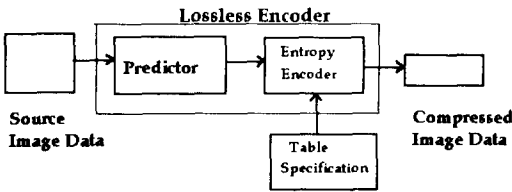


그림 3.4 무손실 부호화 블럭도

을 먼저 부호화하고 다음부터는 전 단계에서 부호화한 것을 보간법에 의해 2배로 확장한 것과 현재 부호화하려는 영상과의 차를 부호화한다. 이 처리를 영상 크기가 입력 영상과 같을 때까지 반복한다.

3.2 MPEG(Moving Picture Experts Group)표준

MPEG은 동영상을 압축하는 것을 목적으로 ISO/IEC JTC1/SC29/WG11에서 현재 MPEG-1, MPEG2를 표준화로 완성하였고 초저속 영상 통신을 위한 MPEG4에 대한 연구가 활발히 진행 중이다.

MPEG1은 1.5Mbps 전송률의 CD-ROM, DAT 와 같은 DSM(Digital Storage Media)을 대상으로 하고 있고 MPEG2는 3Mbps 이상의 전송률로서 CD-ROM, Digital VCR, LDP, CATV 그리고 HDTV를 대상으로 하고 있다. MPEG4는 기존의 전화망(Public Service Telephone Network)이나 무선망을 대상으로 한 64kbps이하의 저속 채널에서도 영상서비스를 가능하게 하자는 데 있다.

MPEG 표준화활동은 MPEG-비디오, MPEG-오디오 및 MPEG-시스템으로 구성되어 있고 본질에서는 최근에 완성된 MPEG2 비디오 표준방식에 대해 간략히 설명한다.

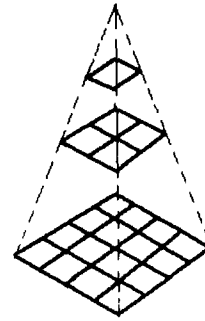


그림 3.5 Hierarchical multi-resolution encoding

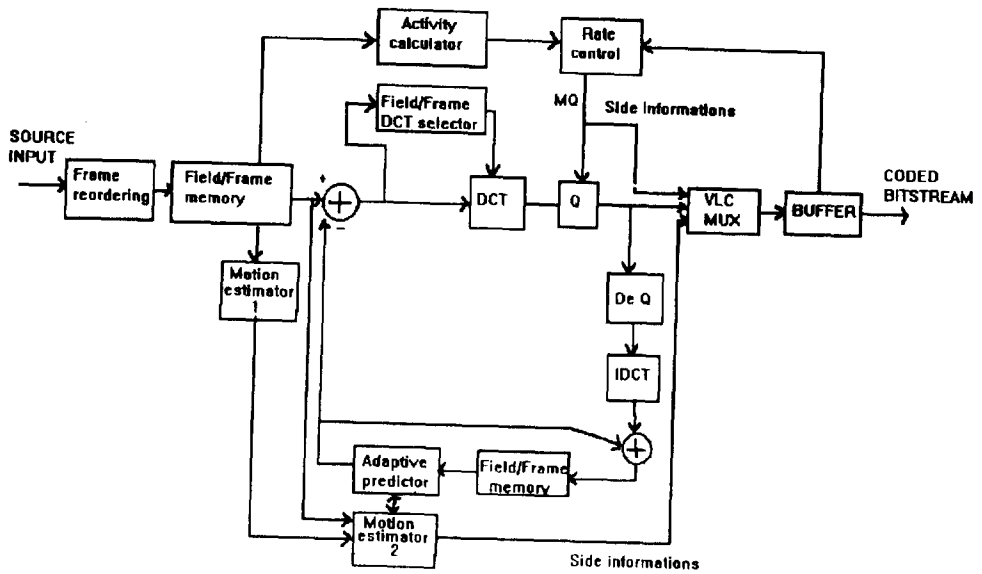
3.2.1 MPEG2 비디오 알고리즘

3.2.1.1 기본원리

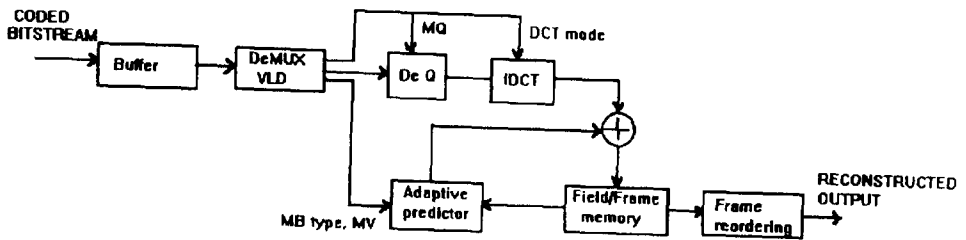
그림 3.6은 MPEG2 부호기/복호기의 블럭도를 나타내고 있다.

입출력 영상의 규격은 CCIR-601 Format을 기본으로 한다. MPEG2 비디오 압축 알고리즘은 시간 중복성(Temporal Redundancy)을 줄이기 위해 Macroblock(16×16화소)단위로 두 화면간의 움직임 추정 및 보상을 이용하고 공간 중복성(Spatial Redundancy)을 줄이기 위해서 Block(8×8화소)단위로 DCT 압축 기법을 이용하며 통계적 중복성(Statistical Redundancy)를 줄이기 위해 가변장부호화(VLC)의 일종인 호프만 부호화를 이용한다.

부호기의 채널 버퍼는 호프만 부호화의 가변 비트율을 고정 비트율로 전송할 수 있도록 데이터를 일시저장하기 위함이다. MPEG은 I,P,B 3종류의 영상 타입이 있으며 각 타입마다 다른 부호화기법을 이용하고 있다. I(Intra)영상은 랜덤 액세스를 위한 기준이 되며 전체 화질에 절대적인 영향을 미치므로 압축률이 높지 않다. 블럭별로 DCT변환과 양자화 과정을 적용하여 공간방향의 중복성만 제거한다. P(Predictive)영상은 이전 영상(I 또는 P)을 기준영상으로 하여 순방향 움직임 보상 예측 기법을 이용하여 시간 중복성을 제거한다. I 영상보다 적으나 B 영상보다는 많은 비트를 할당하여 부호화한다. B(Bidirectional)영상은 이전 영상(I 또는 P)과 다음 영상(I 또는 P)으로부터 양방향성 움직임보상예측기법을 이용하여 시간 중복성을 제거한다. 다른 영상에 비해 상대적으로 적은 비트를 할당하여 부호화한다. MPEG에서는 I 영상을 주기로



아) MPEG2 부호기



바) MPEG2 복호기

그림 3.6 MPEG2 부호기/복호기 블럭도

하여 I,P,B를 묶어 하나의 GOP(Group of Picture)를 구성하고 있다. MPEG표준은 복호기가 아니라 복호화과정을 정의하고 있으므로 복호기를 구현하는 방법은 한 가지로 제한하지 않고 있다.

3.2.1.2 비디오 데이터의 계층구조

MPEG2 부호화 기법에서 사용되는 영상 데이터는 비트열 Syntax를 고려하여 계층적인 구조를 갖는데 Video Sequence Layer아래에 Random Access를 위한 기본단위인 GOP(Group of Picture) Layer, Picture Layer, 오류의 영향을 국한시키는 Slice Layer, 움직임 보상의 단위인 Macroblock Layer, DCT단위인 Block Layer순으로 구성되어 있다. 그림 3.7은 각 계층의 영상 데이터 구조를 나

타내고 있다.

3.2.1.3 움직임 추정 및 보상

(Motion Estimation/ Compensation)

영상간 예측 효율을 높이기 위해 움직임 보상 예측 기법을 일반적으로 이용하는데 SIF(Source Input Format)영상을 기본으로 하는 MPEG1에 비해 CCIR-601 Format영상을 기본으로 하는 MPEG2는 보다 다양하고 복잡한 움직임 추정 및 예측 기법을 필요로 한다. MPEG2에서는 움직임 추정 및 보상 방법으로 Frame MC, Field MC, Dual Prime MC의 3가지가 있다.

3.2.1.4 DCT변환 부호화 및 양자화

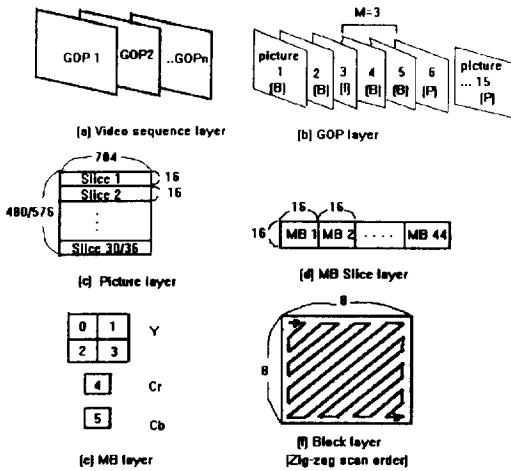


그림 3.7 Video data의 계층구조

DCT변환은 8x8화소의 Block단위로 수행되며 Intra MB일 경우 공간적 중복성만 제거하지만 Inter MB일 경우는 시간적 중복성이 제거된 것에 또 다시 공간적인 중복성을 제거한다.

또한 격행주사(Interlaced Scan)규격 영상의 DCT변환 효율을 증대시키기 위해 Odd Field와 Even Field의 상관성에 따라 적응적으로 Frame DCT 혹은 Field DCT Block으로 구분하여 DCT를 수행한다.

양자화기는 8x8 화소 block 단위로 양자화를 수행하며 I, P, B 영상 순으로 비트를 할당하도록 양자화한다.

3.2.1.5 VLC(Variable Length Coding)

MPEG2에서는 양자화된 DCT 변환계수, 이동벡터의 차신호 그리고 MB에 관련된 정보를 호프만 부호화하여 전송한다. 양자화된 DCT계수는 Intra DC 계수일 경우, 인접 블록의 DC계수간의 차이값을 호프만 부호화를 하고 그 외의 계수값은 (Run, Level)심볼에 따라 호프만 부호화한다.

이동벡터는 바로 전의 MB의 이동벡터와 현재의 MB의 이동벡터의 차이값을 호프만 부호화한다. 한 Slice에서 MB의 위치정보, 타입 그리고 MB 내에서의 Block들의 부호화 패턴(Coded Block Pattern)정보에 대해서 호프만 부호화하여 전송한다.

3.2.1.6 전송율 조정(Rate Control)

MPEG 부호화는 I, P, B 영상에 따라 부호화되는 데이터량이 매우 차이가 나기 때문에 고화질을 유지하며 일정한 전송속도를 갖는 전송로를 통해서 전송하기 위해서는 전송을 조절이 필요하다. 따라서 MB 단위로 전송 버퍼와 입력 신호의 복잡도(Activity)를 기준으로 전송률을 조정하는 Feedback과 Feedforward 방식을 결합한 방식을 사용한다.

4. 결 론

본고에서는 디지털 영상 부호화 기법으로 일반적으로 많이 이용되고 있는 예측 부호화, 변환 부호화, 벡터 양자화, 대역 분할 부호화, 엔트로피 부호화, 또한 제2세대 부호화 기법으로 MPEG4에서 연구되고 있는 객체지향, Wavelet, 프랙탈 부호화에 대해서 그리고 국제 표준 알고리즘인 JPEG, MPEG 표준안에 대해 간단히 소개하였다.

영상 데이터는 많은 정보량에 의한 메모리 용량이나 통신속도의 제한 때문에 오늘날에는 그것을 해결하기 위한 다양한 디지털 영상 부호화 기술이 발전하게 되었고 호환성을 보장하기 위한 여러가지 영상부호화 기술의 표준화가 결실을 맺어감에 따라 고도 정보화 사회의 다양한 통신 서비스가 가능하게 될 것이다.

그동안 국내의 기술진은 국제 표준화 활동에 대한 대응이 충분하지 못한 아쉬움이 있다. 표준안은 기본적인 부호화 기법을 명시한 것이므로 응용에 따라 변형해야 할 사항이 많다. 따라서 응용에 따른 추가적인 연구가 필요하며 또한 핵심 ASIC 기술 개발에 한층 노력을 해야한다. 그리고 초저속 영상 통신을 목적으로 하고 있는 MPEG4에 대해서도 적극적으로 참여하여 새로운 정보산업의 발전에 뒤떨어지지 않도록 제2세대 부호화 기법에 대해서도 적극 연구가 수행되어야 할 것이다.

참 고 문 헌

- [1] A. K. Jain, "Fundamental of digital image processing", Prentice Hall, 1989.
- [2] R. C. Gonzalez, R. E. Woods, "Digital image processing", Addison-Wesley, 1992.

[3] A. Gersho, R. M. Gray, "Vector quantization and Signal compression", KAP, 1992.

[4] P. P. Vaidyanathan, "Multirate systems and filter banks", Prentice Hall, 1993.

[5] M. Vetterli, "Wavelets and filter banks: theory and design", IEEE Trans. signal processing, vol. 40, no. 9, pp. 2207-2232, sept. 1992.

[6] H. G. Musmann, M. Hötter and J. Ostermann, "Object-oriented analysis-synthesis coding of moving images", Signal processing: Image communication, vol. 1, no. 2, oct. 1989, pp. 117-138.

[7] Guojun Lu, "Fractal image compression", Signal processing:Image communication, vol. 5 no. 4, oct. 1993, pp. 327-343.

[8] W. B. Pennebaker, J. L. Mitchell, "Still image data compression standard", Van Nostrand Reinhold, 1993.

[9] ISO /IEC 13818-1,2,3(1994), "Information technology-coding of moving picture and associated audio", MPEG-2 Draft International Standard.

[10] 남재열 외 2명, "영상부호화 기술 동향", 한국통신학회지, 제11권 제8호, 1994년 8월, pp. 23-35.

[11] 권주환 외 2명, "MPEG Video 표준 해설", 한국통신학회지, 제11권 제8호, 1994년 8월, pp. 37-49.



김종국(金鍾國)

1961년 3월 26일생. 1983년 2월 경북대 공대 전자공학과 졸업. 1985년 2월 동 대학원 전자공학과 졸업(석사). 1984년 12월~현재 삼성전자(주) 선임연구원. 1992년 3월~현재 한국과학기술원(서울) 정보및 통신공학과 박사과정. 주관심 분야: Wavelet 변환, 영상 데이터압축, 영상통신.



박현욱(朴玄旭)

1959년 7월 5일생. 1981년 2월 서울대 공대 전기공학과 졸업. 1983년 2월 한국과학기술원 전기 및 전자공학과 졸업(석사). 1988년 2월 한국과학기술원 전기및 전자공학과 졸업(공학박사). 1989년 7월~92년 3월 미국 워싱턴대학교 연구원. 1992년 5월~93년 7월 삼성전자 정보컴퓨터연구소 수석연구원. 1993년 8월~현재 한국 과학기술원(서울) 정보및 통신공학과 조교수. 주관심 분야: 영상처리, 영상통신, 영상처리시스템, 의료영상시스템.