《主 題》

비디오 압축 기술 동향 - MPEG-4 압축기법을 중심으로

김 남 철
(경북대학교 전자·전기공학부)

□차 례□

Ⅰ. 서 론
Ⅱ. MPEG-4 표준화의 중요성
Ⅲ. MPEG-4 고등 영상압축 기법
Ⅳ. 결 론

Ⅰ. 서 론

문자, 그래픽, 오디오, 영상 및 비디오 등의 총합체로서 구성되는 멀티미디어 데이타 중에서 비디오는 그 의미 전달 효과가 매우 크기 때문에 멀티미디어 데이타의 핵심이라 할 수 있다. 그러나, 한편으로 비디오는 단위시간당 발생 또는 재생되는 데이타량이 방대하여 일반적으로 압축 과정 없이는 저장 또는 전송이 어렵기 때문에 멀티미디어 서비스의 확산을 위한 길목에서 병목현상을 일으키는 주범이라고도 할 수 있다. 그 동안 많은 비디오 압축기법이 연구되어 왔으며[1-4], 지금까지 ITU-T(CCITT의 후신)와 ISO/IEC MPEG 그룹에 의하여 여러 가지 비디오 압축에 대한 표준화 작업이 활발히 추진되어 왔다. 이미 제정된 표준에는, 64k-2Mbps 전송률의 비디오전화, 비디오회의 등을 위한 H.261, 1.5 Mbps급 전송률의 디지틀 저장매체를 위한 MPEG-1, 그리고 3M-15Mbps 전송률의 디지틀 TV 및 HDTV 수준의 응용을 위한 MPEG-2 등이 있다. 그리고 극히 최근에 추진되어온 64kbps 이하의 초저속 전송률에서의 응용을 위한 단기 표준으로 H.263이 있는데, 금년 3월 이에 대한 표준안[5, 6]이 작성되었고 금년 안으로 제정될 예정이다. H.263에서는 H.261이나 MPEG-1 및 -2의 비디오 압축 기법인 움직임 보상과 DCT의 혼합기법을 거의 그대로 사용하고 있기 때문에, 일반적으로 H.263의 재생 비디오에는 H.261, MPEG-1 및 -2 시스템에서 고압축을 했을 경우에 발생하는 블록화 현상과 모스키토 현상이 심하게 나타나 품질이 열등한 것으로 알려져 있다. 장기적으로 볼 때, 초저속 전송률에서의 응용에 대한 수요가 확산되도록 하기 위해서는 보다 고등의 압축기법을 개발하여 품질을 개선할 필요가 있으며, 이에 대한 표준화 작업은 MPEG-4[7, 8]라 하여 1998년을 목표로 현재 활발히 추진되고 있다.

국내에서도 비디오 압축 기법 및 표준화에 대한 관심이 점차 높아짐에 따라 그 동안 한국통신학회 및 대한전자공학회의 학술지를 통하여 수차례에 걸쳐서 비디오 압축표준 및 압축기법에 대한 소개가 이루어진 바 있다[9-12]. 또한, MPEG-4에 대한 표준화 동향도 근래에 소개된 바 있으나[13-18], 아직 MPEG-4에 대한 인식이 미흡하고 이의 표준화에 대한 열기가 부족한 상태이다. 본 고에서는 아직 제정되지 않은 MPEG-4의 중요성을 인식시키고 고등의 비디오 압축기법들을 깊이 있게 소개하여 MPEG-4의 표준화 작업에 국내의 많은 산학연 연구그룹들이 함께 참여하는 분위기를 조성하고자 한다.

Ⅱ. MPEG-4 표준화의 중요성

서론에서 이미 언급된 여러 가지 비디오 압축의 표준에 관한 주요 특성들을 비교하면 표 1과 같다. 이들

표 1. 비디오 압축 표준들의 주요 특성 비교

Standards	Bit Rates	Input Format	Features	Applications	Final Date
MPEG-2	3~15 Mbps	CCIR 601(HDTV)	random access(scalability)	TV, HDTV	1993
MPEG-1	around 1.5 Mbps	CIF(SIF)	random access	CD-ROM DAT Winchester disk	1991
H.261	p×64kbps	CIF	integer-pel accuracy	videophone video conference	1990
H.263	below 64kbps	QCIF	very similar to H.261	videophone via PSTN	1995
MPEG-4			advanced technique	mobile net LAN	1998

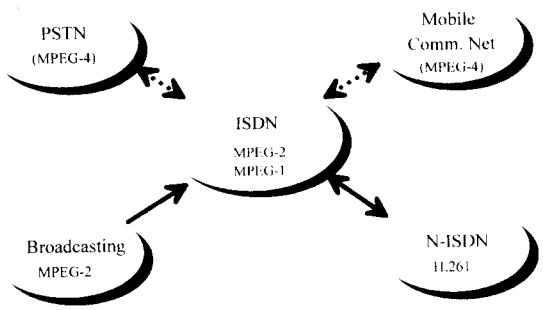


그림 1. ISDN 및 주변 통신망에서의 멀티미디어 서비스를 위한 비디오 압축 표준

중, MPEG-4는 기존의 협대역망인 공중교환망과 이동통신망을 포함하는 64kbps 이하의 초저속 전송로를 통한 비디오전화 등의 멀티미디어 서비스가 가능하도록 하는 것을 목표로 하고 있다. 따라서, 만약 MPEG-4가 성공적으로 제정될 경우, 그림 1에서 볼 수 있듯이 직장에서는 B-ISDN이나 N-ISDN을 통하여 값비싼 고품질의 멀티미디어 서비스를 제공받는 한편, 집에서는 공중교환망을 통하여 차안이나 비행기안에서는 이동통신망을 통하여 값싸고 유용한 멀티미디어 서비스를 받을 수 있게 되며, 또한 각 통신망간의 멀티미디어 서비스의 원활한 소통이 이루어지게 된다. 이로 인하여 진정한 의미의 통신망의 통합화와 멀티미디어 서비스의 대중화가 이루어지게 되며, HDTV 및 VOD 등의 고품질 서비스에 대한 수요도 함께 확산되는 효과도 얻을 수 있다.

이렇게 중요한 MPEG-4의 표준화에 대한 현재 국내의 인식은 H.261, MPEG-1 및 -2 등의 표준화에 대처하지 못한 과거 국내의 능력과 행보 때문에 그다지 높지 않은 편이다. 그러나, 당시의 상황과 현재의 상황은 사뭇 다르다는 점에 주목할 필요가 있다.

첫째로, MPEG-4 표준화의 직접적인 파급효과가 큰다. 표 1에서 H.263, H.261, MPEG-1 및 -2의 비디오 압축 기법은 근본적으로는 움직임 보상과 DCT의 혼합기법이며 정보이론 및 부호화이론에 근거한 것으로, 64kbps 이하의 전송률에서 양호한 재생품질을 유지하는 데에 한계가 있다. 이를 극복하기 위해서 MPEG-4에서는 웨이브렛변환, 프랙탈기반, 객체기반, 모델기반 및 영역기반 기법 등의 고등기법들을 압축기법으로 선호하고 있다. 이러한 고등기법들을 실현하기 위해서는 영상처리, 컴퓨터비전, 컴퓨터 그래픽스, 인공지능 등의 다양한 분야에서의 기술들이 함께 개발되어야 하기 때문에, MPEG-4의 표준화를 위한 노력은 압축기술 뿐만 아니라 자동화, 애니메이션, 멀티미디어, 자동인식 기술 등의 발전에도 크게 기여하게 된다.

둘째로, MPEG-4 표준화에의 참여 기회가 아직 많다. 과거의 표준화 작업들은 기본 압축기법은 이미 학문적으로 완결된 상태에서 시작된 경우가 대부분이었다. 그러나, 현재 MPEG 4에서는 후보 압축기법들이 다양하게 존재하고 있고, 초저속에서 양호한 성능을 보이면서 실시간 하드웨어 구현도 가능한 결정적인 방법이 아직 제안되어 있지 않은 상태이다. 따라서, 1998년에 제정되는 MPEG-4는 H.263보다 약간 개선된 초저속 압축의 중기 표준 정도로 그치고, 이상적으로 지향하고 있는 양호한 성능의 장기 표준의 제정은 그 이후로 연기될 가능성도 큰 것으로 판단된다. 지금이라도 장기적인 안목으로 MPEG-4의 표준화에 참여한다면 아직도 기회는 많다고 볼 수 있다.

셋째로, MPEG-4 표준화 결과는 고품질의 응용에도 사용가능하다. 초저속 전송률에서의 압축률은 상대적으로 높아야 하기 때문에 이러한 전송률에서 양호한 성능을 내기 위해서는 고도의 압축기술들이 집약되어야 한다. 따라서 초저속 전송률에서 양호한 품질을 보이는 압축기법이 있다면, 그것은 HDTV 및 VOD등의 고품질의 응용에서도 효율적으로 적용될

수 있다. 그러므로, MPEG-4 표준화의 성공적인 결과는 장차 고품질 고압축의 표준화를 추진하는 데에도 응용될 수 있다.

넷째로, MPEG-4 표준화는 압축 연구그룹간의 협력이 필요하다. 과거의 표준화와는 달리, 현재 후보기법들로 거론되고 있는 고등의 압축기법들이 입력영상과 응용상황에 맞춰 사용자의 선택에 따라 사용될 수 있도록 모두 표준 기법으로 채택될 가능성도 크다. 따라서, 특정 연구그룹이 모든 후보방법들에 대한 표준화 연구를 수행하기 보다는 국내의 압축 연구그룹간의 협력이 절대적으로 요구되고 있다.

끝으로, MPEG-4 표준화는 관련 연구그룹의 지원이 절대적으로 필요하다. 아무리 고등의 압축기법이라도 해도 압축과정만으로는 고압축에 의한 화질열화를 완전히 극복하기 어렵다. 따라서, 송신측에서 전처리 필터처리 기법을 효과적으로 사용하여 압축 효율을 극대화할 수 있게 하고 수신측에서는 영상복원 기법을 사용하여 재생영상의 품질을 향상시키는 방식도 함께 고려되어야 한다. 또한, MPEG-4의 고등의 압축기법들은 비디오의 파형에 관한 정보를 보내기보다는 파형에서 추출한 특징 파라미터들을 전송하기 때문에 채널전송이나 저장시 에러가 발생하거나 고의적인 교란으로 인한 정보보호에 대단히 취약하므로 이의 방지를 위한 채널부호화 및 암호화 기법도 함께 연구되어야 한다. 뿐만 아니라, 실시간 하드웨어 구현이 가능하도록 개발된 기법들의 VLSI설계 연구도 함께 뒷받침되어야 한다. 현재 국내에는 과거와는 달리, 비디오 압축 및 영상처리 등의 연구그룹과 아울러 관련 연구그룹의 기술들이 축적되어 있는 상태이므로 MPEG-4 표준화를 중심으로 지원이 집중된다면 우리나라도 이 분야에서 소기의 성과를 달성할 수 있다고 판단된다.

Ⅲ. MPEG-4 고등 영상압축 기법

3.1 웨이브렛 변환 기법[19-21]

웨이브렛 변환 : 웨이브렛 변환은 Fourier 변환과 같이 기저함수(basis function)들의 집합으로 신호를 분해하여 표현하는 하나의 방법이다. 그러나, Fourier 변환과는 달리 국부적으로 에너지가 집중된 대역통과 신호인 웨이브렛들을 기저함수로 사용한다. 즉, 임의의 연속신호 $f(t)$에 대한 웨이브렛 변환 Wf는 다음과 같이 표현된다.

$$Wf(s, x) = \int f(t) \cdot \psi_{s,x}(t)\, dt$$

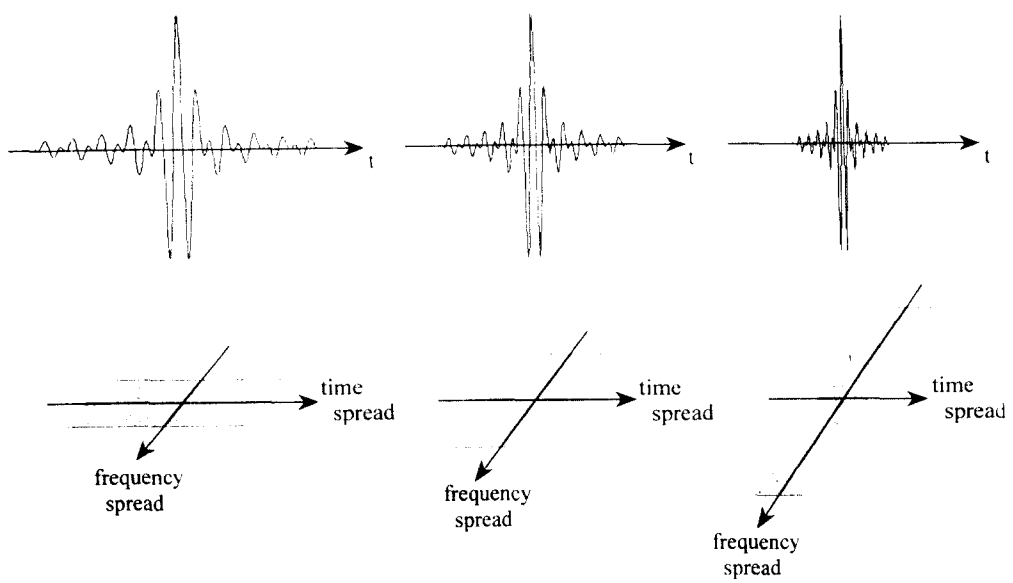


그림 2. 웨이브렛 기저함수들의 예

여기서, 웨이브렛 기저함수 $\psi_{s,x}(t)$들은 모 웨이브렛(mother wavelet)이라 불리는 하나의 원형(prototype) 웨이브렛 $\psi(t)$를 s만큼 확장(축소)시키고 x만큼 이동시켜 얻는다.

$$\psi_{s,x}(t) = \frac{1}{\sqrt{s}}\ \psi(\frac{t-x}{s})$$

웨이브렛 변환에서는 확장된 신호의 주파수를 언급하는 것보다 몇 배로 확장 또는 축소됐는가 하는 스케일(scale)을 언급하며, 일반적으로 스케일은 2배수를 사용하여 $s=2^j$(여기서 j는 정수)인 경우를 흔히 생각한다. 그림 2는 웨이브렛 기저함수들의 예를 보이고 있다. 여기서, 스케일 s가 배로 증가할수록 웨이브렛 함수는 보다 낮은 주파수 대역의 성분을 나타내며 시(공)간 해상도는 떨어지나 주파수 해상도는 좋아지고, 스케일 s가 반으로 감소할수록 보다 높은 주파수 대역의 성분을 나타내며 주파수 해상도는 떨어지나 시(공)간 해상도는 좋아짐을 볼 수 있다. 이러한 시간-주파수 양쪽의 국부화(localization) 특성과 시간-주파수간의 해상도 조절 특성으로 인하여 비정재적 신호 분석 및 처리에 유용하게 쓰이고 있으며, 컴퓨터 비전에서의 다해상도 처리, 음성 및 영상압축에서의 부대역 부호화(subband coding) 등 다양한 분야의 기법들에 대한 단일화된 이론을 제공하여 준다.

다해상도 표현 : 각 스케일을 중심으로 웨이브렛 변환을 다루기 위하여, 편의상 웨이브렛 변환을 아래와 같이 재 표현하기로 하자.

$$W_j f(x) = Wf(s, x)\,|_{s=2^j}$$

또한, 연속신호 $f(x)$가 유한한 해상도를 갖고 있다고 가정하면, $W_1 f(x)$는 한 단계 낮은 해상도에서의 상세정보를 의미하는 것으로 생각할 수 있다. 이때, 한 단계 낮은 해상도에서의 나머지 부분인 개략정보를 $A_1 f(x)$라 하면 이들 사이의 관계는 아래와 같이 표현할 수 있다.

$$\begin{aligned} f(x) &= A_0 f(x) \\ &= A_1 f(x) + W_1 f(x) \end{aligned}$$

이것을 임의의 해상도 상으로 확장하여 생각하면 다음과 같이 다해상도 상의 신호들을 얻게 된다.

$$\begin{aligned} f \rightarrow\ & A_1 f \rightarrow A_2 f \rightarrow \cdots \rightarrow A_j f \\ & \rightarrow W_1 f \rightarrow W_2 f \rightarrow \cdots \rightarrow W_j f \end{aligned}$$

그리고 이들 사이의 관계는 다음과 같이 표현할 수 있다.

$$A_{j-1} f(x) = A_j f(x) + W_j f(x)$$

또한, 위 식들의 관계로부터 다음과 같이 각 연산자들 간의 관계식도 얻을 수 있다.

$$\begin{aligned} A_j &= A_{j-1} \circ A_1 \\ W_j &= A_{j-1} \circ W_1 \end{aligned}$$

이러한 관계들을 블록도로 표현하면 그림 3과 같이 트리 구조의 분해 및 합성 과정으로 나타낼 수 있다.

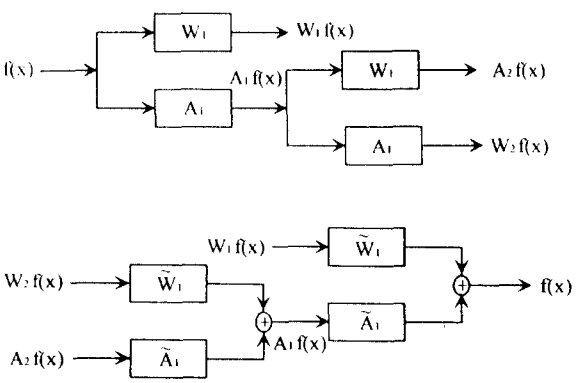


그림 3. 연속 신호의 웨이브렛 변환에 의한 다해상도 분해 및 합성

이산 웨이브렛 변환 (DWT) : 연속신호 $f(x)$가 유한한 해상도를 갖는다는 것은 유한한 대역폭을 갖는 것을 의미하며, 이러한 신호는 샘플링 이론에 의하여 디지틀 신호 $f(n)$으로 변환하여도 다시 복원할 수 있다. 또한, 해상도 또는 대역폭이 반으로 줄어든 $A_1 f(x)$ 및 $W_1 f(x)$의 샘플링 주파수는 $f(n)$에 대한 것보다 반으로 줄어들어도 된다. 그러므로 그림 3에 대응되는 이산신호 및 시스템들을 생각할 때, 연속 연산자 A_1는 디지틀 LPF 필터 $h(n)$과 downsampler의 조합으로 대치될 수 있으며, W_1는 디지틀 HPF 필터 $g(n)$과 downsampler의 조합으로 대치될 수 있다. 이러한 내용을 블록도로 표현하면 그림 4와 같이 디지틀 필터에 의한 트리 구조의 분해 및 합성 과정으로 나타낼 수 있다. 여기서 보면 DWT 변환 및 역변환은

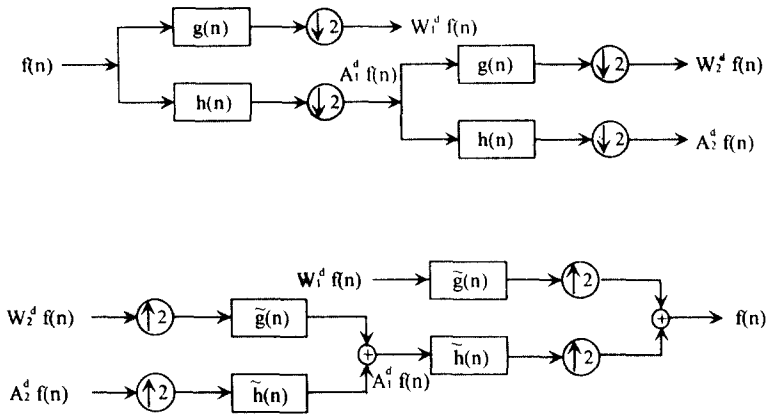


그림 4. DWT에 의한 다해상도 분해 및 합성

결국 트리구조의 부대역 분해 및 합성 과정임을 알 수 있으며, 이러한 사실은 이미 정량적으로도 입증된 바 있다[21].

어떠한 LPF 필터 $h(n)$과 HPF 필터 $g(n)$를 사용하느냐는 것은 몇 가지 제한조건 내에서의 선택의 문제이다. 이론적으로 가장 간단한 한 예는 각각 $\pi/2$ (quadrature frequency)의 차단 주파수를 갖는 ideal LPF와 ideal HPF로 선택하는 것이나, 이 필터들은 실현불가능하다. $\pi/2$를 중심으로 roll-off하는 실현가능한 필터들 중, 그림 5와 같이 $h(n)$과 $g(n)$의 주파수 특성이 서로 거울상(mirror image)의 형태가 되면 엘리아싱(aliasing)이 없는 복원이 가능한데, 이러한 특성을 갖는 필터 쌍을 QMF(quadrature mirror filters)라 부른다.

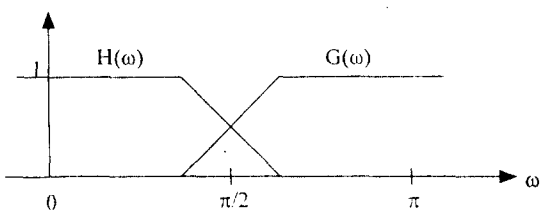


그림 5. QMF 필터의 주파수 특성

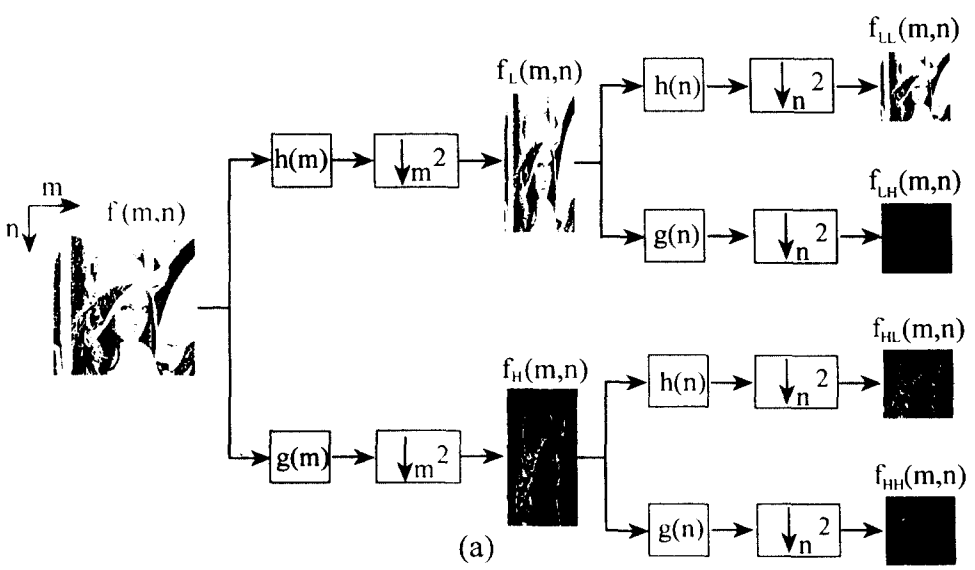
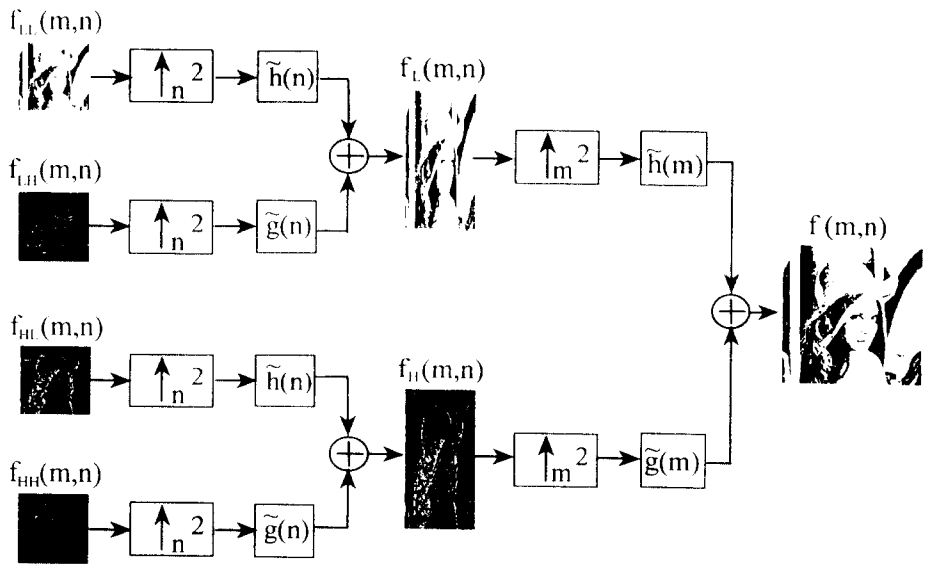


그림 6. 2차원 DWT에 의한 다해상도 분해 및 합성



(b)

그림 6. (계속)

영상신호에 적용되는 2차원 DWT는 수직방향으로 1차원 변환하고 그 결과를 다시 수평방향으로 1차원 변환하는 것으로 수행할 수 있으며, 이러한 방식의 다해상도 분해 및 합성 과정에 대한 블록도를 나타내면 그림 6과 같다. 저주파 대역 LL에 다시 2차원 DWT을 가하여 전체를 7대역으로 분해한 결과는 그림 7과 같다.

웨이브렛 변환 부호화는 다해상도 처리가 가능하고 점진적 전송과 가변율 전송이 용이하다는 장점을 가지고 있다. 2차원 DWT를 비디오 압축에 적용할 경우, 여러 가지 형태의 변형이 가능하지만 간단한 예로는 MPEG 계열의 압축 시스템에서 DCT블록 대신에 DWT블록으로 치환해 넣는 것이다.

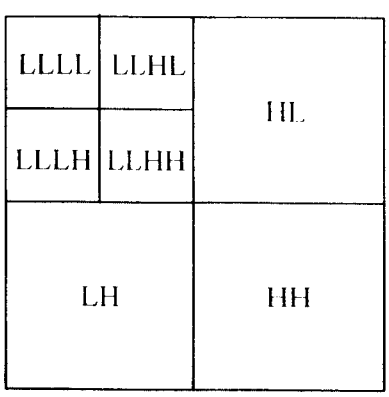


그림 7. 2차원 DWT에 의한 7대역 분해의 결과

3.2 프랙탈기반 기법[22, 23]

프랙탈 영상 압축기법은 자기유사성(self-similarity)이라는 특성을 갖는 프랙탈 기하이론에 근거한 압축기법으로, 영상내의 각 블록들을 축소변환(contractive transformation) 파라미터들로 표현함으로써 영상을 부호화 한다. 반복 축소변환에 의한 프랙탈 영상 압축이론은 Barnsley가 최초로 제안하였고, Jacquin은 이 이론에 근거하여 블록기반 프랙탈 압축기법을 제안하였으며 현재 Jacquin의 방법은 프랙탈 압축 기법의 근간이 되고 있다.

축소변환과 콜라주 정리 : 완전 측도 공간(complete metric space) (X, d)에서 정의되는 변환 $\tau : X \rightarrow X$에 대해서, 임의의 두 원소 x, y 사이의 거리보다 변환 후 변환된 두 원소 $\tau(x)$, $\tau(y)$ 사이의 거리가 가까워질 때, 이를 축소변환이라 한다. 즉, 축소변환은 다음과 같이 표현된다.

$$d(\tau(x), \tau(y)) \le s\, d(x, y), \qquad 0 \le s \le 1$$

여기서 d는 Hausdorff 거리 측도를 말하고 s는 축소변환 τ의 축소인자를 나타낸다. 이와 같이 정의되는 축소변환 τ는 변환과 무관하게 요지부동한 유일한 하나의 고정점(fixed point)을 갖는다.

$$x_a = \tau(x_a)$$

그리고 임의의 원소 $z \in X$에 대하여 축소변환 τ를 반복적으로 적용할 경우 고정점 x_a에 수렴하게 된다.

$$\lim_{n \to \infty} \tau^n(z) = x_a \qquad \forall z \in X$$

또한, 축소변환 τ는 다음과 같은 콜라주 정리(collage theorem)를 만족한다.

$$d(x, x_a) \le \frac{1}{1-s}\ d(x, \tau(x)), \qquad \forall z \in X$$

프랙탈 부호화 : 프랙탈 부호화기에서는 콜라주 정리에 근거하여 주어진 전송률에서 원 영상 x와 재생영상 x_a와의 거리(또는 MSE) $d(x, x_a)$를 최소화하기 위해서, 원 영상 x와 콜라주 영상 $\tau(x)$와의 거리 $d(x, \tau(x))$를 최소화하도록 하는 축소변환 τ를 찾는 일을 한다. 그런데, 실제 영상 내에는 일반적으로 성질이 다른 여러 가지 형태의 프랙탈 물체들이 공존하므로, 이를 효과적으로 부호화하기 위하여 축소변환 τ는 다음과 같이 N개의 축소변환 τ_i들의 콜라주로 이루어진 것으로 간주한다.

$$\tau = \bigcup_{i=0}^{N} \tau_i$$

여기서, 각 τ_i들은 그림 8에서와 같이 편의상 $M \times M$ 크기(예 : 8×8)의 블록 r_i들을 서로 겹침이 없이 발생시키는 것으로 간주한다.

그리고, 이 때의 τ_i는 다음과 같은 affine 변환의 일종으로 구성된다고 가정한다.

$$\tau_i(p) = \alpha \cdot I(S(p)) + t$$

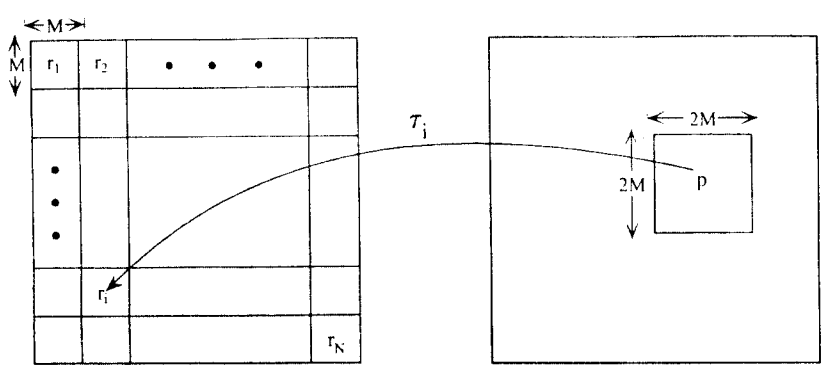


그림 8. 치역 블록의 구성과 정의역 블록의 탐색

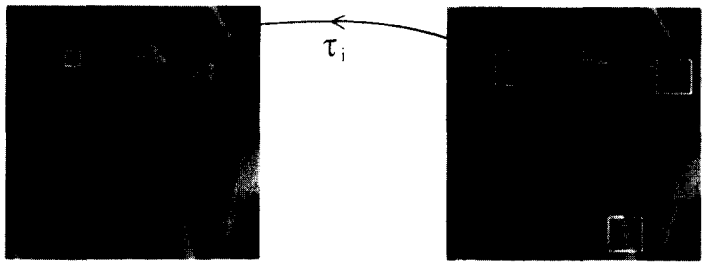


그림 9. 프랙탈 부호화에서의 탐색 과정

여기서 p는 정의역(또는 pool) 블록(크기 $2M\times2M$)을 나타내며, S는 $1/2\times1/2$로 블록을 축소하는 것을 나타내고, I는 회전, 거울상 등의 isometry 변환, α는 대비 조정 인자, t는 밝기 조정 인자를 나타낸다. 따라서, 실제 부호화기에서는 주어진 치역블록 r_i에 대해서 어떤 정의역 블록 p로부터 축소되고 어떤 I 변환과 어떤 대비 조정 인자 α 및 밝기 조정 인자 t에 의해 생성된 $\tau_i(p)$와의 거리 $d(r_i, \tau_i(p))$가 최소가 되는지를 BMA의 전영역 탐색(full search)방식으로 찾아 이때의 p의 위치, I의 종류, α 및 t의 크기에 관한 정보(τ_i)를 수신 측으로 전송한다. 그림 9는 실제 영상에 대하여 주어진 하나의 치역 블록에 대하여 MSE를 최소로 하는 정의역 블록을 찾는 과정을 나타낸 것이다.

프랙탈 복호화 : 프랙탈 복호화기에서는 수신된 축소변환 τ_i에 관한 정보를 가지고 임의의 초기영상 z에 대하여 축소변환 τ를 반복(실제로는 10회 정도) 적용함으로써 재생영상 x_a를 얻는다.

$$x_a = \lim_{n\to\infty} \tau^n(z)$$

그림 10은 원 영상과는 전혀 무관하게 선택된 Pepper영상을 초기영상으로 하여 수신된 축소변환 τ_i에 관한 정보에 근거하여 축소변환 τ를 10회 반복 적용하였을 때 각 단계에서 얻은 영상들을 보여주고 있다.

일반적으로 프랙탈 압축 기법은 고압축율에서도 블록화 현상이 심하지 않으며, 영상의 에지나 질감의 시각적 재생 화질이 우수한 것으로 알려져 있다. 프랙탈 압축 기법에서는 부호화에서의 전영역 탐색의 부담이 대단히 크므로 이를 줄이기 위한 노력이 계속되고 있으며, 복호화에서는 반복없이 재생하는 방법들

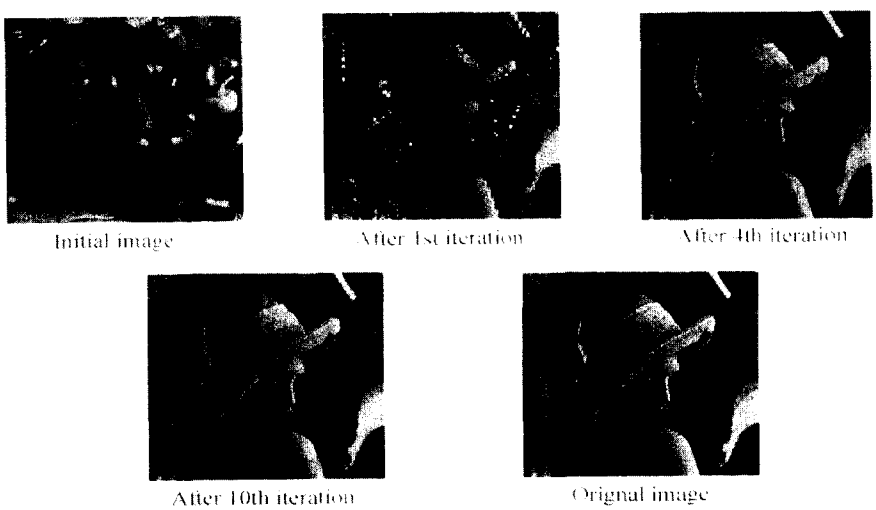


그림 10. 프랙탈 복호화에서의 각 영상들

이 연구되고 있다. 프랙탈 기법을 비디오 압축에 적용할 경우, MPEG 계열의 압축 시스템에서 DCT블록 대신에 프랙탈기반 부호화기로 대치하는 방식은 그다지 성능이 좋지 않은 것으로 알려져 있으므로, 보다 효율적인 비디오 압축에의 적용 방법에 관한 연구도 계속되고 있다. 프랙탈 기법은 부호화의 실시간 처리 측면에서는 매우 불리하지만, 실시간 복호화만 요구되는 멀티미디어 온 디맨드와 같은 응용에서는 상대적으로 유리한 기법이라고 볼 수 있다.

3.3 모델기반 기법[24, 25]

모델기반 압축 기법에서는, 대상 물체(주로 얼굴)의 형상에 관한 3차원 모델과 지식을 송신측과 수신측에서 함께 준비해 놓고, 송신측에서는 그 모델과 지식에 근거하여 입력영상을 분석 인식하고, 수신측에서는 인식결과에 따라 변형된 동영상을 합성해낸다. 이때, 분석에는 주로 영상처리 및 컴퓨터 비전 기술이 사용되며, 합성에는 컴퓨터 그래픽스 기술이 주로 사용된다. 이러한 모델기반 비디오 합성 및 압축 기법은 1980년대 초부터 활발히 연구되기 시작하여, 80년대 후반에 들어서면서 스웨덴의 Forchheimer 그룹 및 일본의 Harashima 그룹 등에 의하여 본격적인 연구가 시작되어 현재에는 응용단계로 들어서고 있다.

모델기반 부호화 시스템의 구성 : 모델기반 부호화 시스템의 전체적인 구성은 그림 11과 같다. 송신측에서는 초기에 머리의 형태, 기하학적인 비율, 표면의 색 및 질감 등의 기본적인 속성들이 추출되어 전송된다. 이러한 속성 정보에 근거하여 송신측 및 수신측에서는 그림 11의 가운데 영상과 같은 입력 영상에 정합된 철선 프레임(wire frame) 모델을 갖게 된다. 다음, 매 프레임마다 얼굴 특징들이 추출되어 정합된 후, 머리의 움직임과 얼굴의 표정 변화에 따라 움직임 파라미터와 표정 변형 파라미터들이 추정되어 전송된다. 수신측에서는 전송되어온 분석정보에 따라 모델이 변형되고, 표정합성 및 질감매핑(texture mapping)에 의해 합성 영상이 생성된다.

모델의 구성 및 정합 : 얼굴모델은 일반적으로 영상합성이 신속하고 용이하게 수행될 수 있는 그림 12와 같은 철선 프레임을 사용한다. 이러한 얼굴모델은 표준적인 얼굴을 근사시켜 철선 프레임 정점들의 3차원 위치좌표를 축적해 놓음으로써 형성된다.

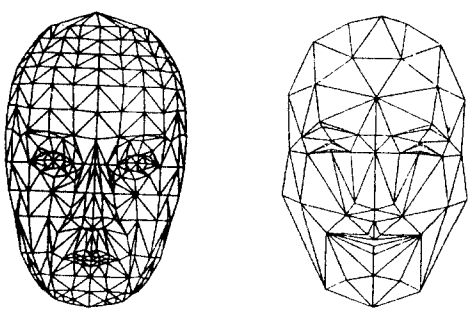


그림 12. 얼굴의 철선 프레임 모델

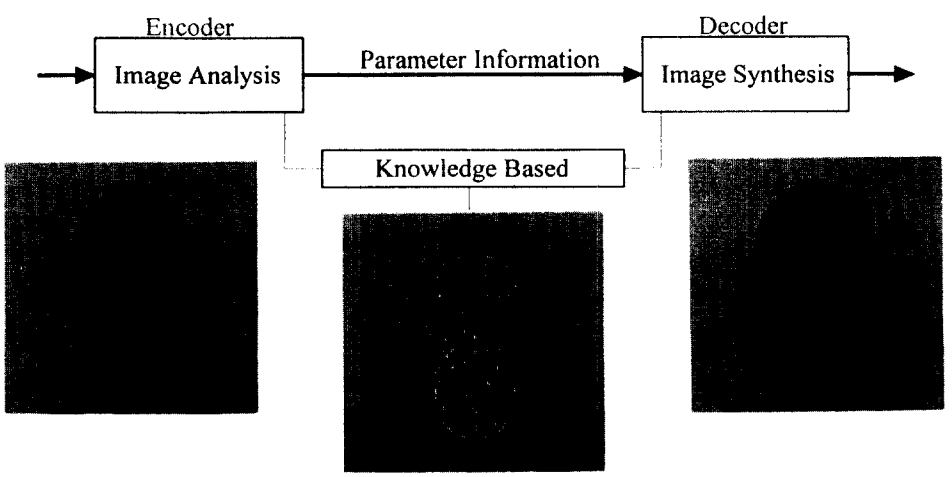


그림 11. 모델기반 부호화 시스템의 개념도

얼굴영상의 분석 : 얼굴에서의 변화는 표정변화와 머리의 3차원적인 변화로 나눌 수 있다. 얼굴표정을 분석하기 위한 연구로 심리학 분야에서 FACS(face action coding system)가 연구된 바 있고, 모델기반 부호화에서도 대부분 FACS에 근거해서 얼굴 윤곽이나 얼굴 표정 변화에 중요한 특징점들의 변화로 얼굴표정을 분석한다. 머리의 3차원적인 변화는 주어진 몇 개의 특징점들의 2차원 이동정보를 이용하여 3차원 움직임 파라미터를 추정한다.

얼굴모델의 변형 : 얼굴모델을 전송된 파라미터를 이용해서 변형시키는 과정은 머리의 3차원 운동 파라미터에 따라 얼굴모델을 스케일링, 이동, 회전시킨 후 얼굴표정 파라미터에 따라 얼굴모델을 변형시킨다. 얼굴모델을 변형한 예로써 그림 13에 3차원 변형된 모델과 입모양이 변형된 모델들을 나타내었다.

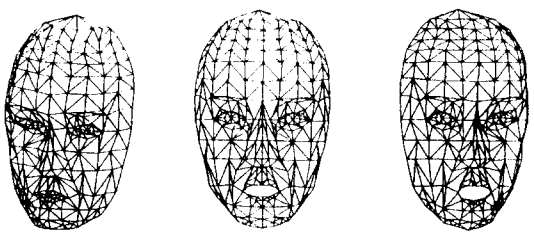


그림 13. 얼굴모델의 변화 예

얼굴영상 합성 : 질감매핑은 물체 고유의 질감을 잘 표현하기 위해서 개발된 방법으로 스크린 영역의 좌표에 해당하는 질감영역에서의 질감을 스크린 영역에 매핑하는 컴퓨터 그래픽스의 한 기법이다. 모델기반 부호화에서는 대부분 삼각형 구조의 얼굴모델을 사용한다. 삼각형은 세 점이 하나의 평면을 정의하기 때문에 선형변환인 affine 변환만으로 임의의 스크린 영역의 좌표에 해당하는 질감영역의 좌표를 간단히 구할 수 있다. 이러한 이유로, 얼굴영상 합성을 하는 시스템의 영상합성법으로 affine 변환을 이용한 질감매핑이 주로 사용된다. 그림 14는 3차원 운동 파라미터와 얼굴표정 파라미터에 얼굴 모델을 변형하고 질감매핑으로 합성한 결과이다.

모델기반 부호화는 초저속 전송률에서도 부자연스러운 부호화 잡음 없는 출력영상을 생성시킬 수 있기 때문에 비디오 전화, 지능형 인간-기계 정합장치(man-machine interface) 등의 광범위한 분야에 응용될 수 있는 압축 기법이다. 그러나, 입력 영상에서 자동적으로 정확하게 특징점과 머리의 3차원 움직임을 추정하는 것이 대단히 어렵기 때문에 실시간 응용을 위해서는 앞으로도 많은 연구가 필요하다.

3.4 객체기반 기법[26, 27]

객체기반 압축 기법에서는, 그림 15에서와 같이 비디오의 각 프레임 영상을 영상분석을 통하여 움직임이 없는 정적 객체(배경)와 3D(또는 2D)의 움직임 객체들로 구분한 후, 움직임 객체들의 형태, 움직임 및 칼라 정보에 관한 파라미터들을 추출하고 전송하는 기법이다. 여기서 형태라 함은 객체 영역의 개략적인 윤곽 정보를 말하며, 칼라라 함은 객체 영역 내부의 질감을 말한다. 모델기반 기법에서 움직임 객체는 미지의 3D 객체라고 가정하지만, 객체기반 기법에서 움직임 객체들은 미지의 3D(또는 2D) 객체들이라고 가정한다. 이러한 객체기반 기법은 독일의 Hannover 대학을 중심으로 활발히 연구되고 있다.

영상 분석 : 객체기반 기법에서는 영상분석을 통하여 이전 영상의 움직임, 형태 및 칼라 파라미터들로부

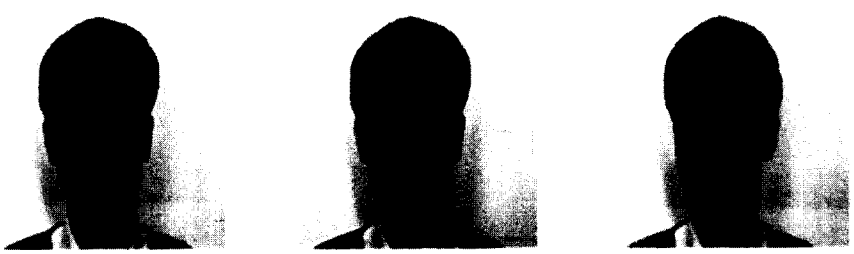


그림 14. 합성된 얼굴영상

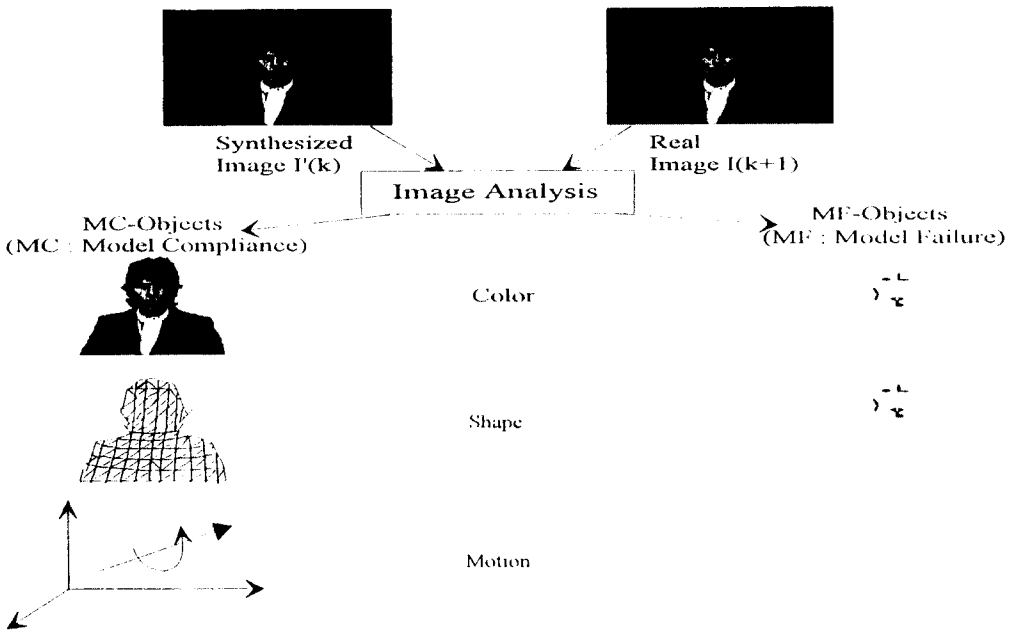
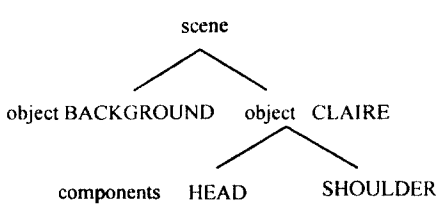


그림 15. 객체기반 부호화기의 원리

터 현재 영상의 움직임, 형태, 칼라 파라미터들을 추정한다. 먼저, 영상 합성, 변화 검출 및 3D 움직임 추정 등의 과정을 거쳐 현재 영상에서 배경과 움직임 객체들을 검출한다. 이 때, 흔히 움직임 객체는 그림 16에서와 같이 하나이며 이들은 다시 몇 개의 객체 성분들로 구성되는 것으로 간주된다. 다음, 각 움직임 객체에 대해서 움직임 및 형태 정보를 전송하고, 움직임 보상을 수행하여 움직임 보상으로 보완될 수 없는 객체 내부의 영역들을 검출한다. 이들은 MF(model failure) 영역이라 하여 이들의 형태 및 칼라 파라미터들을 전송한다. 영상 합성시 사용되는 모델 객체는 그림 17에서와 같이 추출된 움직임 객체의 2D 마스크인 실루엣(silhouette) 내에서 타원형의 깊이를 갖는 3D 철선 프레임에 질감매핑을 하여 얻게 된다.

움직임 추정 : 객체기반 기법의 효율성은 효과적인 움직임 추정 기법에 달려 있다. 움직임 추정은 3차원 또는 2차원 움직임 모델에 근거하여 수행된다. 움직임 객체내의 화소(x, y)가 화소(x', y')으로 이동한다고 가정할 때, 3차원 움직임 모델은 다음과 같이 8개의 파라미터로 표현되는 모델이 흔히 사용된다.

$$x' = \frac{a_1 x + a_2 y + a_3}{a_7 x + a_8 y + 1}$$



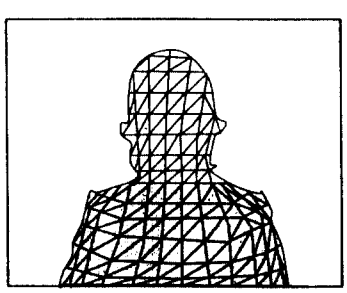
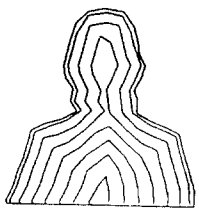
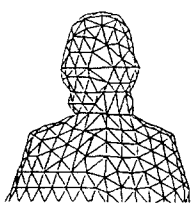


그림 16. 객체들의 구성

object silhouette



contour lines



mesh of triangles

model object

그림 17. 객체 실루엣으로 부터의 모델 객체의 생성

$$y' = \frac{a_4 x + a_5 y + a_6}{a_7 x + a_8 y + 1}$$

그리고, 2차원 움직임 모델은 다음과 같은 6개의 파라미터로 표현되는 affine 모델이 흔히 사용된다.

$$x' = a_1 x + a_2 y + a_3$$

$$y' = a_4 x + a_5 y + a_6$$

이들 파라미터의 추정은 일반적으로 FD(frame difference)나 DFD(displaced FD)의 에너지를 최소화하는 방식에 근거하여 순환적으로 수행된다.

객체기반 기법에서는 기존의 블록기반 움직임 보상/DCT 기법보다 움직임 정보를 효율적으로 부호화하기 때문에 초저속 전송률에서 어느 정도 양호한 품질의 영상을 얻을 수 있다. 그러나, 배경이 복잡하여 배경과 움직임 객체와의 분리가 용이하지 않거나 움직임 객체의 수 및 그 성분들의 수가 사전에 가정한 수보다 많을 때에는 부호화 효율이 급격히 떨어지게 되는 단점이 있다.

3.5 영역기반 기법[28, 29]

그림 18에서 영상을 일정한 크기의 블록들로 나누어 부호화하는 것을 블록기반 부호화라고 하며, 영상을 동질성을 갖는 영역들로 분할한 후 분할된 각 영역의 윤곽 정보 및 내부의 질감 정보를 부호화하는 것을 분할기반 또는 영역기반 부호화라고 한다. 영역기반 비디오 압축 기법에서는 영상을 정적 영역(배경)과 움직임 영역들로 분할한 후, 움직임 영역들의 윤곽, 움직임 및 질감에 관한 파라미터들을 추출하고 전송한다. 이 때, 모델기반 및 객체기반 기법과는 달리, 움직임 영역에 관한 사전 지식이 전혀 요구되지 않는다. 이러한 기법은 스위스의 Kunt 그룹을 중심으로 활발히 연구되고 있다.

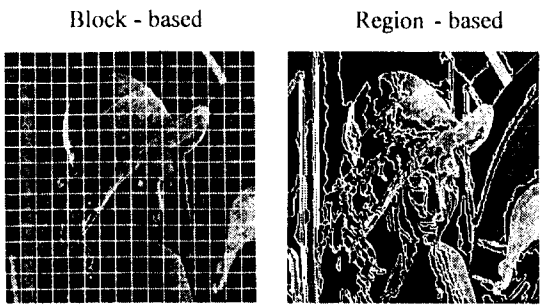


그림 18. 블록기반 기법 및 영역기반 기법

그림 19는 비디오의 한 프레임 영상을 화소의 동질성에 근거하여 자세히 분할한 후, 다시 움직임 정보에 근거하여 병합 정리한 결과를 나타낸다. 여기서, 최종 분할된 각 영역들은 배경을 제외하고는 모두 움직임 영역을 나타낸다. 각 영역마다 영역의 크기와 움직임 모델의 적합도에 따라 움직임 벡터, 2D 또는 3D 움직임 파라미터를 추정하여 윤곽정보와 함께 전송한다.

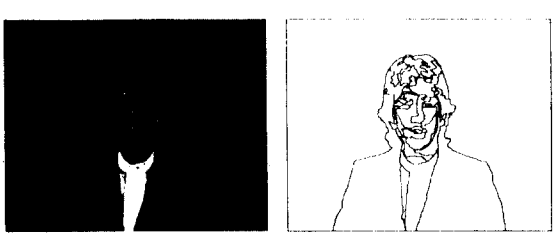
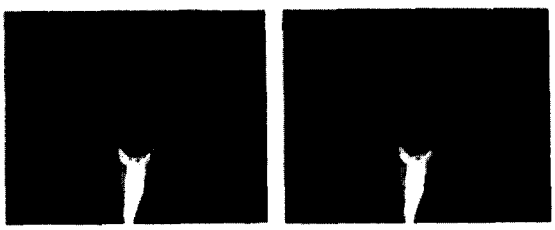


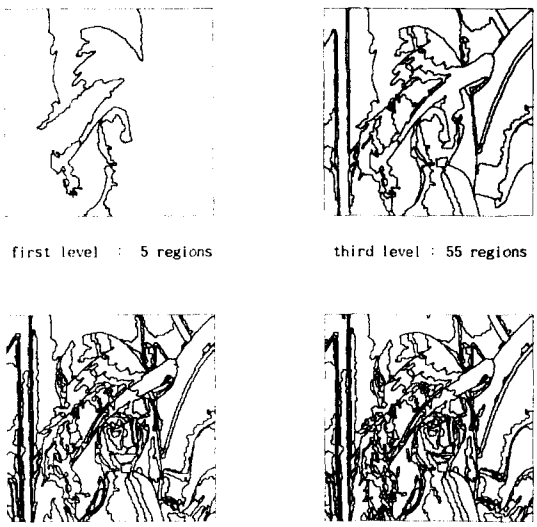
그림 19. 분할된 움직임 영역들

그리고, 움직임 보상후 보상이 되지 않는 영역에 대해서는 질감 정보도 추가로 전송한다. 그림 20은 블록기반 BMA와 영역기반 움직임 추정의 성능을 비교한 것이다. 여기서 두 방법들의 움직임에 관한 오버 헤드는 동일하나, 화질은 두드러지게 영역기반이 우수한 것을 볼 수 있다.



motion compensation based on BMA　　region-based motion compensation

그림 20. 블록기반 BMA 및 영역기반 움직임 추정의 성능 비교



first level : 5 regions　　third level : 55 regions

fifth level : 205 regions　　sixth level : 308 regions

그림 21. 모폴로지에 근거한 다단계 영상 분할

영상분할 : 영역기반 기법의 효율성은 효과적인 영상 분할 기법에 달려 있다. 특히, 영상 분할의 정확도 및 고속 처리가 크게 요구된다. 영역기반 기법에서 흔히 사용되는 순차적 분할기법으로는 CLRG(centroid linkage region growing), MST(minimum spanning tree), RAG(region adjacency graph) 등의 알고리듬이 있다. 이들 순차적인 알고리듬들은 일반적으로 처리속도가 느리지만, 최근에는 priority que 등을 사용하여 이들을 고속화하는 연구도 활발히 수행되고 있다. 한편, 고속 병렬처리를 위하여 신경회로망을 이용한 분할 기법도 연구되고 있으며, 유럽에서는 RACE MORPHECO 프로젝트를 통하여 모폴로지(morphology) 필터에 의한 고속 분할 기법의 개발에 열을 올리고 있다. 그림 21은 모폴로지 필터에 의한 다단계 영상분할의 결과들을 나타낸 것이다.

영역기반 기법에서는 기존의 블록기반 기법보다 움직임 정보를 효율적으로 부호화할 수 있고, 객체기반 기법에 비해서는 사전지식이 전혀 필요없어 사용환경에 거의 제한을 받지 않는다는 점이 유리하다. 그러나, 영역들의 윤곽에 관한 오버 헤드가 상대적으로 크기 때문에 윤곽 정보의 근사화 및 윤곽 정보의 축소 방안이 계속 연구되어야 한다. 또한, 영상분할의 정확도 및 고속처리에 관한 연구도 함께 수행되어야 한다.

Ⅳ. 결 론

본 고에서는 MPEG-4 표준화의 중요성에 관하여 살펴 보았고, MPEG-4의 비디오 압축 알고리듬으로 현재 인급되고 있는 5가지 고등 압축기법의 기본적인 특성을 소개하였다. MPEG-4의 직접적인 응용 측면에서는 비디오 품질이 상대적으로 낮고 시장규모도 그다지 크지 않은 것으로 생각할 수 있겠으나, 본 고의 고찰을 통하여 그 표준화의 파급효과 및 영향력은 매우 광범위한 것임을 알 수 있었다. 또한, 웨이브렛 변환, 프랙탈기반, 객체기반, 모델기반 및 영역기반 기법의 5가지 고등 압축기법들은 어떠한 특성과 장단점을 갖고 있는지 알 수 있었다. 앞으로 이들 방법의 장점과 특성을 살려서 복합적으로 결합함으로써 더욱 효과적인 압축기법을 도출하기 위한 연구도 계속되어야 하겠다. 끝으로, 국내 산학연의 여러 관련 그룹들이 MPEG-4 표준화에 관심을 갖고 연구와 협력을 계속한다면 좋은 결실을 맺을 수 있으리라 기대된다.

참 고 문 헌

1. A. K. Katsaggelos, Proc. SPIE Visual Communications and Image Processing '94, Chicago, 1994.
2. CIPIC, Proc. Int'l Picture Coding Symposium PCS'94,

Sacramento, 1994.
3. IEEE, Proc Int'l Conf. on Image Processing '94, Austin, 1994.
4. L. T. Wu, Proc. SPIE Visual Communications and Image Processing '95, Taipei, 1995.
5. ITU-T Study Group 15, Working Party 15/1, Report of the Rapporteur for Very Low Bitrates Visual Telephony, Sep. 1993.
6. ITU-T, Draft Recommendation H.263 (Video Coding for Narrow Telecommunication Channels at 64 kbps), Geneva, Mar. 1995.
7. ISO/IEC JTC/SC29/WG11, MPEG-4 Requirements Document, MPEG94, July 1994.
8. ISO/IEC JTC1/SC29/WG11, MPEG-4 Functionalities, 1994.
9. 정동석, 영상 부호화 기술과 국제표준화 동향, 한국통신학회지, 제11권 제8호, 1994년 8월.
10. 박종국, HDTV 특집, 전자공학회지, 제19권 제12호, 1992년 12월.
11. 김이한, 김성대, "영상 압축 기술의 동향 및 알고리즘", 컴퓨터와 멀티미디어 기술 특집, 전자공학회지, 제22권 제4호, pp. 64-75, 1995년 4월.
12. 남재열, 하영호, "디지틀 영상 표준화 및 서비스 개요", 디지틀 영상 서비스 기술 특집, 전자공학회지, 제22권 제7호, pp. 29-42, 1995년 7월.
13. 문주희, "MPEG-4 표준화 현황 및 영상 부호화 기술 동향", 전자공학회지, 제22권 제1호, pp. 56-65, 1995년 4월.
14. 김용한, 이상미, 안치득, "MPEG 4 표준화 동향 및 전망", 영상 부호화 기술과 국제표준화 동향 특집, 한국통신학회지, 제11권 제8호, pp. 50-59, 1994년 8월.
15. 김성한, 이홍규, "초저속 전송을 위한 wavelet변환 기반의 동화상 압축 기술", 영상 부호화 기술과 국제표준화 동향 특집, 한국통신학회지, 제11권 제8호, pp. 60-77, 1994년 8월.
16. 서정태, 강현우, 윤대희, "프랙탈 영상 압축영상", 컴퓨터와 멀티미디어 기술 특집, 전자공학회지, 제22권 제4호, pp. 13-21, 1994년 10월.
17. 김남철, 홍원학, 김상현, "3차원 모델기반 동영상 전송", 한국통신학회지, 제10권 제4호, pp. 17-24, 1993년 4월.
18. 이태원, 고성제, 김성대, 김남철, 조남익, 최병욱, 백준기, 초저속 동영상 압축 및 처리 기법 연구, 특정연구 1차 중간보고서, 한국과학재단, 1995년 6월.
19. O. Rioul and M. Vetterli, "Wavelets and signal processing", IEEE Signal Processing Magazine, pp. 14-38, Oct. 1991.
20. I. Daubechies, "Orthonormal bases of compactly supported wavelets", Comm. on Pure and Applied Math., vol. XLI, pp. 909-996, 1988.
21. S. G. Mallat, "A theory for multiresolution signal decomposition : the wavelet representation", IEEE Trans. Patt. Analysis., Machine Intell., vol. 11, pp. 674-675, July 1989.
22. M. F. Barnsley, Fractals Everywhere, Academic Press, Boston, 1988.
23. A. E. Jacquin, "Fractal image coding : A review", Proc. IEEE, vol. 81, No. 10, Oct. 1993.
24. R. Forchheimer and T. Kronander, "Image coding-From waveforms to animation", IEEE Trans. Acoust. Speech, Signal Processing, vol. 37, pp. 2008-2023, Dec. 1989.
25. K. Aizawa, H. Harashima, and T. Saito, "Model-based analysis synthesis image coding(MBASIC) system for a person's face", Signal Processing : Image Comm. vol. 1, pp. 139-152, Oct. 1989.
26. M. Hotter, "Object-oriented analysis-synthesis coding based on moving two-dimensional objects", Signal Processing : Image Comm., vol. 2, pp. 409-428, Dec. 1990.
27. J. Ostermann and P. Gerken, "Object-oriented analysis-synthesis coding based on source models of moving 2D and 3D-objects", ISO/IEC JTC1/SC29/WG11 MPEG93/710, July 1993.
28. P. Cicconi and H. Nicolas, "Efficient region-based motion estimation and symmetry oriented segmentation for image sequence coding", IEEE Trans. Circuits and System for Video Tech., vol. 4, pp. 357-364, June 1994.
29. C. Gu and M. Kunt, "Very low bit-rate video coding using multi-criterion segmentation", Proc. IEEE Int'l Conf. on Image Processing '94, Austin, 1994.

김 남 철

• 1955년 7월 29일생
• 1978년 2월 : 서울대학교 전자공학과 졸업(공학사)
• 1980년 2월 : 한국과학원 전기 및 전자공학과 졸업 (공학석사)
• 1984년 2월 : 한국과학기술원 전기 및 전자공학과 졸업(공학박사)
• 1991년 8월~1992년 8월 : 미국 Syracuse 전기및전산학과 객원교수
• 1990년 1월~1994년 12월 : 생산기술연구원 HDTV 신호처리 분과위원
• 1994년 1월~현재 : 대한전자공학회 및 한국통신학회 편집위원
• 1984년 3월~현재 : 경북대학교 전자 전기공학부 교수
• 주관심분야 : 영상처리, 영상압축, 컴퓨터 비전