

영상 압축 기술의 동향 및 알고리즘

金 二 漢
韓國科學技術院

金 聖 大
韓國科學技術院

I. 서 론

머지않은 미래에 가정에서 모든 업무를 수행하고 쇼핑을 비롯한 문화적 욕구를 충족할 수 있는, 컴퓨터, 전화, TV 등 대중매체가 통합된 멀티미디어 시대가 도래할 것이다. 이러한 멀티미디어 시대에는 방대한 정보를 빠른 속도로 주고 받을 수 있어야 하기 때문에 고속의 전송로와 시스템이 요구된다. 그러나 현재의 기술 수준으로는 사용자의 욕구를 충족시킬 수 없으므로 제한된 전송로 하에서 만족할 만한 성능을 얻기 위하여는 데이터 압축이 필연적이다. 단편적인 예로, NTSC신호를 기본으로 하여 720×480 크기의 영상 신호의 경우 초당 250M비트의 데이터량을 갖고 있으므로 현재의 기술로는 전송이 불가능하다. 특히, HDTV와 같이 고화질의 경우 그 양은 기하급수로 증가한다. 그러므로 디지털 영상 압축 기술은 매우 중요하며, 그 응용 분야에 따라 다양한 표준화 그룹이 성립되었다. 표준화 그룹으로는 정지영상에 대하여 JBIG, JPEG과 동영상 압축을 위하여 MPEG 등이 제안되고 있다. 본 고에서는 이들 각 분야에 대하여 표준화 그룹의 동향과 제안 알고리즘에 대하여 살펴볼 것이며, 음성 부호화 기법에 대하여는 간단히 언급할 것이다.

II. 정지영상 압축 기술

컴퓨터를 이용하여 인터넷(internet)에 접속한 경험이 있는 독자라면 월드디지니의 인어 공주에서부터 세계적인 유명 모델 신디 크리프드의 아름다운 모습까지, 만년설 히말라야의 고봉에서부터 광활한 우주의 성운들까지, 그리고 다양한 형태의 문서 등을 컴퓨터 모니터를 통하여 본 경험이 있을 것이다. 이러한 데이터들, 특히 영상은 중대한 정보를 가지고 있기 때문에 여러 분야에 유용하게 사용되고 있다. 그러나 이러한 데이터들은 그 양이 매우 방대하기 때문에 아무런 압축 기술없이 사용

한다면 전송을 위한 많은 시간과 저장 공간을 필요로 한다. 따라서 영상 압축 기술은 필수적이며 반드시 해결해야 할 과제이다.

본 장에서 살펴볼 정지 영상 압축 기술은 의학의 x-선 및 단층 사진, 인쇄 및 신문 사진에서 영상 회의를 위한 자료 화면까지 다양한 분야에 응용될 수 있다. 정지 영상으로는 일반적으로 생각하는 영상(컬러 영상)과 문서(흑백 영상)로 분류할 수 있으며, 이들 각각을 위하여 국제적인 표준화 그룹이 형성되어 있다(컬러 영상을 위한 JPEG과 흑백 영상을 위한 JBIG). 이들 각 그룹으로부터 국제 표준안(International Standard, IS)이 제안되었으며, 다양한 분야에서 이를 이용하고 있다. 한편, 영상을 표시하는 방법으로는 수신한 데이터를 종이에 직접 표시하는 팩시밀리와 같은 하드카피(hardcopy) 방식과 컴퓨터의 모니터에 표현하는 소프트카피(softcopy)로 구분할 수 있는데 본 고에서는 주로 소프트카피의 관점에서 압축 기술을 살펴볼 것이다.

1. JPEG(Joint Photographic Experts Group)

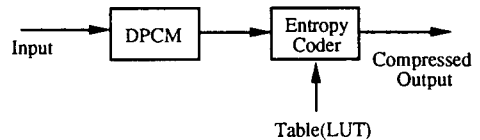
JPEG은 1982년 ISO/CCITT 공동으로 결성된 후, 1987년 7월 10가지의 흑백/컬러 영상에 대한 여러가지 부호화 방식이 제안되어 비교, 검토되었으며, 부호화된 화질의 평가, 하드웨어로 구성되는 부호기/복호기에 대한 성능, 압축률 등을 상호 비교하여 1988년에 최종적으로 ADCT(Adaptive DCT)를 기초로 하는 알고리즘을 채택하였으며, 이러한 부호화 방식을 JPEG방식이라 하였다. 이후 1989년에 JPEG부호화 방식을 채택하여 수정 작업을 거친 후, 1990년 12월에 CD(Committee Draft) 초안이 완성되었으며, 최종적으로 1992년에 국제표준(IS)으로 JPEG알고리즘이 확정되었다.

JPEG은 크게 두 가지의 압축 방식으로 분류할 수 있다. 제1방식은 손실 부호화로 DCT(Discrete Cosine Transform)를 기본으로 하는 방식이며, 제2방식은 무손실 부호화로 DPCM(Differential PCM)을 행하는 방식이다. DCT 방식은 양자화를

포함하기 때문에 일반적으로 완전한 영상을 복원할 수 없는 비가역적 부호화이지만 적은 비트수로도 충분한 복호 영상의 품질을 얻을 수 있으며, JPEG의 기본이 되는 방식이다. 한편, DPCM 방식은 압축률은 작지만 원래의 영상을 완벽하게 재구성할 수 있는 가역 부호화이며, 표준 방식으로 추가된 방식이다.

1) 무손실 부호화

무손실 부호화는 의료 영상과 같이 원래의 정보가 손실되면 중대한 결과를 초래할 수 있는 영상 정보를 압축하는데 사용되며, 압축률은 2:1 정도로 그다지 높지 않다. 그림 1은 무손실 부호화의 처리과정을 나타낸 것으로 그림에서 예측기(predictor)는 인접한 3개의 화소로부터 현재 화소 값을 예측하며, 엔트로피코(entropy) 부호기에서는 예측된 화소 값과 실제의 화소 값을 비교하여 두 값의 차에 대하여 허프만(Huffman) 부호화 또는 산술(arithmetic) 부호화 알고리즘을 적용한다.



〈그림 1〉 무손실 부호화의 기본 블럭도

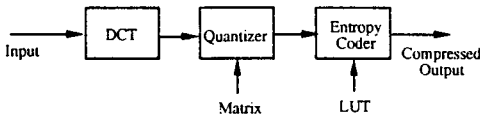
2) 손실 부호화

DCT 방식은 필수 기능인 Baseline 시스템과 선택 기능인 Extended 시스템으로 분류된다. Baseline 시스템은 DCT 방식을 실현하는 모든 부호기/복호기가 갖추지 않으면 안되는 최소한의 기능을 보이고 있으며, 대부분의 응용 분야에서는 충분한 성능을 발휘할 수 있다. Extended 시스템은 넓은 영역의 응용을 위하여 준비된 선택적인 기능이다. 본 고에서는 Baseline 시스템을 기본으로 JPEG에서 채택한 ADCT 방법에 대하여 살펴볼 것이며, JPEG에서 고려한 ADCT의 특징은 아래와 같다.

- (1) No image modeling : 모든 종류의 영상을 동일한 방식으로 처리할 수 있어야 한다.
- (2) Psychovisual thresholding and linear quantization : 인간의 시각 특성을 이용하

여 양자화 효율을 개선한다.

- (3) Adaptivity : 영상의 특성에 적응적인 행렬 및 VLC(Variable Length Code) 표(Look-Up Table, LUT)를 사용한다.
 - (4) Flexibility : 부가적인 정보로 VLC 및 양자화 표를 전송하거나 고정된 값을 사용한다.
 - (5) Comptibility : CCITT 및 CCIR의 표준 영상 전화 및 회의 시스템의 Intra 모드 부호화 방식과 호환성을 가진다.
 - (6) Simple control of picture quality : 영상의 화질 및 압축률은 양자화 행렬을 변화시킴으로써 간단히 수행한다.
 - (7) Sequential and progressive build-up : 순차 및 점진 전송, 재생을 수행한다.
- ADCT 시스템의 전체적인 구조는 그림 2와 같다.



(그림 2) 손실 부호화의 기본 블럭도

가. 변환 부호화(DCT)

DCT 변환을 이용한 부호화 방식은 입력 영상의 공간 방향의 중복성을 제거하는 것으로 원 영상의 모든 화소에 분산된 에너지를 DC를 포함한 낮은 주파수를 갖는 몇 개의 변환 계수에 집중시킨 후 이들의 장점을 적절히 이용한 방식이다. ADCT에 적용될 DCT의 수학적 정의는 다음과 같다.

1) FDCT(Forward DCT)

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

2) IDCT(Inverse DCT)

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

실제 영상에서는 입력 영상을 8×8 블럭으로 분할하고 각 블럭에 대하여 DCT를 수행한다. 이 때

Cr, Cb 등 컬러 성분의 평균치는 -128~+127 값을 가지므로 변화된 계수의 DCT 범위는 -1024~+1023이나, Y 성분의 평균치는 0~255의 범위를 가지므로 컬러 성분과 다르기 때문에 이를 방지하기 위하여 Y 성분에 -128를 더해준 후 DCT를 수행한다. 변환 부호화로 DCT를 선택한 이유는 고속 알고리즘이 제안되어 있으며, 일반적인 영상에 대하여 최적을 성능을 발휘할 수 있기 때문이다.

나. 주관적인 양자화

DCT된 계수의 각 위치에서의 값들은 실제 영상의 특성과 연관되어 있다. 즉, DCT 계수의 각 위치에서의 값은 원 영상의 상관성의 정도를 나타낸다. 따라서 변환 계수들은 위치에서 상대적으로 다른 의미를 가지며, 양자화나 기타 부호화 과정에서 손실이 생기면 계수들의 위치에 따라 복원된 영상의 각기 다른 부분에 영향을 미친다. 한편, 일반적으로 양자화에 의해서 부호화 손실이 발생했을 때 인간의 시각적 특성은 물체의 경계 등과 같이 공간 방향으로 급격한 변화가 있는 부분 보다는 물체의 표면 등과 같이 균일한 부분에서 크게 느껴진다고 알려져 있다. 이러한 인간의 시각적 특성과 DCT 계수들의 특성을 결합하여 부호화에 이용하고 있다.

즉, DC를 포함한 저주파 영역에 대해서는 양자화 간격을 좁히는 반면, 고주파 영역에 대해서 양자화 간격을 넓히므로써 효율적인 압축을 기대할 수 있다. JPEG에서는 양자화 행렬을 권고하였으나, 사용자의 필요에 따라 적합한 양자화 행렬을 만들어 이용할 수 있으며, 양자화 간격은 등 간격(uniform quanization)을 사용하였다.

다. 엔트로피 부호화

이렇게 양자화된 계수 중 F(0,0)를 DC계수라 하며 이 값은 인접한 블럭과의 상관성이 매우 높으므로 이에 대해서는 인접한 블럭의 DC 값과의 DPCM을 수행하고 나머지 AC계수에 대해서는 지그재그 스캔(zig-zag scan)을 통한 가변길이 부호화(variable length coding)와 줄길이 부호화(run-length coding)를 한다. 지그재그 스캔을 하는 이유는 저주파로부터 점차 고주파를 검색하므로써 양자화된 DCT계수에서 '0'이 연속적으로 발

생활 가능성을 도모하기 위한 것으로, 이렇게 발생된 '0'들에 대하여 줄길이 부호화를 함으로써 압축 효율을 증가시킬 수 있다.

3) 점진 전송

JPEG은 영상을 표현하는 데 있어 순차(sequential)적으로 하는 것과 점진(progressive)적으로 하는 것으로 나뉘어져 있다. 이들은 팩시밀리와 같이 종이 위에 기록하는 하드카피와 비디오텍스(videotex)와 같이 모니터 상에 영상을 표시하는 소프트카피 방식에 있어 효과적으로 사용된다. 하드카피 방식은 사용자가 영상이 전달되는 과정에 관여하지 않고 단순히 전달이 끝난 후에 관심이 있으므로 하드웨어 구현상 순차재생이 적합하다. 그러나 이와 같은 순차재생은 사용자가 데이터 베이스를 검색하는 데 있어서는 적합하지 않다. 즉, 찾고자 하는 내용의 대략적인 윤곽만 보기도 사용자의 요구를 충족할 수 있으므로 좀더 효율적인 방법이 요구되며, 이를 위한 것이 점진전송 방법이다. 점진전송 알고리즘을 간단히 설명하면 아래와 같다.

- (1) 먼저 원 영상의 해상도를 가로, 세로 각각 2의 배수로 축소한다.
- (2) 축소된 영상에 손실 또는 무손실 부호화 방법을 적용한다.
- (3) 축소된 영상을 보간법을 이용하여 2배로 확장한다.
- (4) 확장된 영상과 원 영상과의 차를 다시 손실 또는 무손실 부호화한다.
- (5) 위의 과정을 반복함으로써 점진 전송의 범위를 확대할 수 있다.

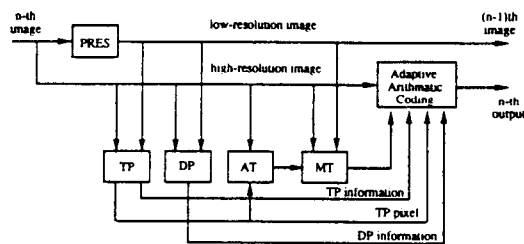
JPEG에서는 위와 같은 전반적인 알고리즘만을 소개하였을 뿐 세부적인 내용에 대해서는 사용자에게 일임하였으나 축소와 보간을 위하여 간단하면서 효율적인 방법을 권고하고 있다. JPEG의 점진 전송 방법은 인터넷의 mosaic나 netscape 등을 통하여 정보를 얻고자 할 때 명확하게 확인할 수 있다.

2. JBIG(Joint Bi-level Image Experts Group)

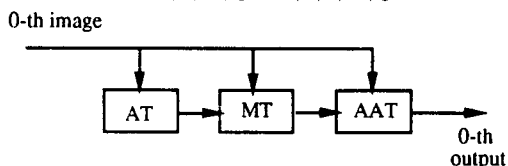
새로운 영상 통신 서비스로서 데이터 베이스 검

색, 오디오 그래픽 회의 등으로 대표되는 소프트카피 통신이 주목을 끌고 있다. 소프트카피 통신이란 이용자의 모니터에 정보를 표시하는 것인데 정보의 재생, 소거, 편집 등의 가공이 용이하다. 앞서 컬러 영상에 대한 압축 기술을 기술한 바 있으므로 본 절에서는 흑백 영상을 위한 압축 기술을 살펴볼 것이다.

지금까지 흑백 영상 표준 부호화 방식으로는 MH(Modified Huffman), MR(Modified Relative edge address designate), MMR(Modified MR) 등이 있으며, 하드카피 통신인 팩시밀리 등에 널리 사용되고 있다. JBIG은 소프트카피의 점진 전송을 목적으로 한 것으로 그림 3과 같이 5개의 블럭으로 구성되어 있다.



(a) 최저 해상도 이외의 계층



(b) JBIG 부호기의 구성

(그림 3) JBIG 부호기의 구성

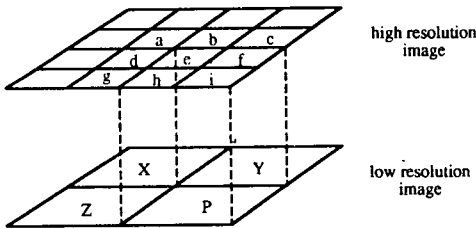
1) PRES(Progressive REduction Standard)

PRES 알고리즘은 흑백 영상의 점진적인 전송을 위하여 영상을 축소하는 과정으로, 그림 4와 같은 방법에 의해서 고해상도 영상으로부터 저해상도 영상을 얻을 수 있다. 즉, 저해상도 영상의 'P' 화소의 색상을 결정하기 위하여 이미 구해진 인접한 화소 'X,Y,Z'와 고해상도 영상의 'a,b,c,d,e,f,g,h,i,' 화소들을 이용하며, 'P' 화소를 결정하는 기본 연산식은 아래와 같다.

$$SUM = 4 \times e + 2 \times (b + d + f + h) + (a + c + g + i) -$$

$$X-3 \times (Y+Z)$$

위 식에서 SUM의 크기가 5보다 크면 'p' 값은 '1' (흑)이고, 그렇지 않으면 'P'의 값은 '0'(백)으로 한다. 또한 위 식은 예측기(TP, DP) 등에서 예측을 위하여 사용되고 있다. 그러나 JBIG에서는 이를 좀더 빠르고 효과적으로 수행하기 위하여 발생할 수 있는 모든 경우의 수(인접한 화소들의 갯수, 2^{12})에 대하여 테이블을 만들어 놓았으며, 알고리즘 적용시 이를 참조한다.



〈그림 4〉 PRES 알고리즘

2) TP(Typical Prediction)

TP는 부호화하려는 영상(고해상도 영상)과 부호화된 영상(저해상도 영상)과의 관계를 이용하는 것으로 예측 가능한 화소에 대해서는 전송하지 않으려는 데 목적이 있다. 그림 4와 같이 고해상도 영상에서의 4개의 화소는 저해상도 영상에서 1개의 화소와 정합되며, 따라서 이를 포함하는 저해상도 영상의 화소를 중심으로 3×3 화소들이 모두 흑 또는 백이면 고해상도 영상에서의 화소의 값은 이미 부호화된 화소와 같은 값을 가질 확률이 매우 높으므로 이런 경우에는 부호화 대상에서 이를 제외한다. 이들은 저해상도 영상에서는 1라인을 단위로, 고해상도 영상에서는 2라인을 단위로 하여 진행시키며, 이에 따라 데이터의 압축을 기대할 수 있다.

3) DP(Deterministic Prediction)

DP는 저해상도 영상을 얻는 축소 방법에 기인하여 이미 부호화된 주위 화소를 가지고 부호화 화소의 레벨을 절대적으로 예측할 수 있는 지를 판단

하고, 예측할 수 있는 경우는 이 화소의 부호화를 생략하는 방식이다. 이들 과정 또한 테이블로 정의 되어 있어 연산을 용이하게 하였다. DP에 의해 부호화 화소 수가 감소하기 때문에 부호화 효율의 향상이 가능하다.

4) MT(Model Template)

MT는 부호화 송신이 끝난 화소들이 취하는 패턴에 기초하여 마코프 모델을 구축해서 그 상태를 인식한 후, 부호화를 행하는 것이다. 최저 해상도 이외의 계층에 사용되는 템플레이트에는 부호화가 끝난 근방 6화소와 저해상도 영상 4화소를 참조 영상으로 하고, 축소 방식과 연관되는 x, y좌표의 위상을 위한 2비트를 포함한 2^{12} 상태의 마코프 모델을 보고 부호화를 생각한다. 반면, 최저해상도 영상에서는 근방 10개의 화소를 참조한다. 즉, MT는 템플레이트의 화소 값들을 배열하여 배열된 것과 일치하는 것을 찾아내는 것이다.

III. 동영상 압축 기술

멀티미디어라는 단어가 떠오르면 영상 분야를 접촉하여 본 독자라면 대부분 MPEG(Moving Picture Experts Group)을 생각하게 될 것이다. 영상 압축 기술하면 MPEG을 떠올릴 만큼 MPEG은 일상적인 언어가 되어버렸다. 실제로 MPEG은 모든 압축 기술을 포함할 만큼 매우 광범위하게 연구되었으며, 현재 대부분의 압축 기술이 표준화되어 있다. MPEG은 움직임이 있는 연속 영상을 대상으로 한 것으로 JPEG, JBIG과는 구분이 된다.

정지영상에 비하여 동영상은 상대적으로 많은 정보를 가지고 있으므로 이를 저장하거나 통신로를 통해 전송하기 위해서는 그에 적합한 데이터의 압축 기술이 요구된다. 이러한 요구를 충족시키기 위해 MPEG에서는 다양한 전송률하에서 동영상의 압축 기술 표준안을 제정하기 위한 작업이 진행중에 있다.

MPEG-1은 15Mbps 전송률의 일반 디지털 저

장 매체를 대상으로 하며, MPEG-2는 런던회의 (92년 11월)에서 전송률을 3-10Mbps에서 3-15Mbps로 상향 조정하고 그 대상을 원래의 CD-ROM, Digital VTR, LDP, CATV에 HDTV를 추가로 포함시키는 대신에, HDTV를 목표로 한 40Mbps의 MPEG-3계획을 취소시켰다. 1994년에 통과된 DIS에서는 1920x1152해상도에서 100Mbps까지의 영상을 처리할 수 있도록 확장하였다. 한편, 현재 공중전화망, 비디오폰, 비디오텍스 등을 대상으로 한 수십 kbps 이하의 MPEG-4가 1997년에 CD, 1998년에 DIS제정을 목표로 표준화가 진행중이다. 일반적으로 MPEG이라 불리는 방법은 MPEG-2를 의미함을 상기시키는 바이다. 본 고에서는 이들 중 MPEG-2와 MPEG-4로 구분하여 제1세대 부호화와 제2세대 부호화에 대하여 고찰할 것이다.

1. MPEG-2

앞서 언급했듯이 MPEG-2는 MPEG이라는 압축 알고리즘의 대명사가 되었으며, 그 응용 범위는 디지털 저장 매체에서 부터 HDTV까지 폭 넓게 적용되고 있다. 이들 적용을 위하여 MPEG-2에서는 Profile과 Level을 정의하였다. Profile이란 특정한 이용을 목적으로 MPEG-2 표준에 포함되어 있는 여러가지 기능들 중에서 필요한 부분만을 모아서 만든 특정 기능들의 집합을 말한다. 특히

기능과 가격면을 고려하여 MPEG-2 표준방식이 도입될 초창기에 가능한 넓은 응용범위에서 사용될 수 있도록 MPEG-2 Main Profile을 정의하고 있다. 나중이라도 필요에 따라 기존에 존재하는 Profile을 확장 또는 축소하여 새로운 Profile을 정의할 수는 있지만, 가능한 Profile의 종류를 제한하여 국제 표준화의 여러 잇점을 최대한으로 유지하도록 권고하고 있다. 현재 MPEG-2모임에서는 Main Profile 외에도 Simple Profile, SNR profile, Spatial Profile 및 High Profile을 정의하고 있다.

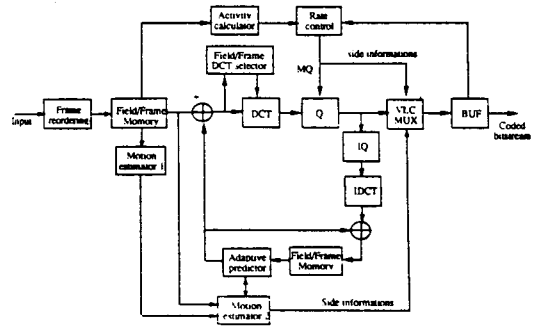
MPEG-2의 특정 Profile에서 관련되는 변수들에 대한 제한 조건을 두어 몇 가지 다른 Level들을 정의한다. 예를들면, 수행할 수 있는 기능은 비슷하더라도 입력으로 수용할 수 있는 영상의 크기를 제한함으로써 서로 다른 응용 분야에 보다 경제적으로 이용할 수 있게 한다. MPEG-2에서는 기능과 가격면을 고려하여 Main Level을 정의하고 있으며, Low Level과 High Level 등도 정의해가고 있다. 아래의 표 1은 MPEG-2에서의 Profile과 Level의 관계를 잘 설명하고 있다.

MPEG-2 알고리즘은 근본적으로 연속 영상을 다루고 있기 때문에 제한된 전송로나 저장 매체에서 실시간으로 처리하기 위해서는 JPEG과는 달리 다양한 압축 기술이 첨가되어야 한다. 그림 5는 MPEG-2 부호기의 구조도이며, 표 2는 MPEG-2 Main Profile MP@ML의 주요 사양이다. 특

<표 1> MPEG-2 Profile 및 Level

Profile \ Level	Simple	Main	SNR Scalable	Spatial Scalable	High
High [1920×1080×30 or 1920×1152×25]		MP@HL (US ATV)			HP@HL
High-1440 [1440×1080×30 or 1440×1152×25]		MP@H1440		SSP@H1440 (EC HDTV)	HP@H1440
Main [720×480×29.97 or 720×576×25]	SP@ML (digital CATV)	MP@ML (digital DBS, video disc)	SNP@ML		HP@ML
Low [352×288×29.97]		MP@LL	SNP@LL		

히 MPEG-2는 디지털 저장 매체에도 효과적으로 사용되어야 하기 때문에 무작위 접근(random access)이 필요하므로 이에 적합한 알고리즘이 요구된다. 대부분의 압축 기법은 영상 내, 또는 영상과 영상사이의 중복성을 찾아내어 이를 제거함으로써 얻을 수 있다. 영상 내의 중복성을 제거하는 기법은 앞서 기술한 JPEG의 ADCT 알고리즘과 크게 다르지 않으므로 본 절에서는 영상과 영상사이의 중복성을 제거하는 방법에 대하여 고찰할 것이다.



〈그림 5〉 MPEG-2의 영상 부호기의 구조

〈표 2〉 MPEG-2 MP@ML의 주요 사양

프레임 포맷	720×480(주파수 29.97Hz) 720×576(주파수 25Hz)
부호화 전송률	최대 15Mbps
Y, Cr, Cb 표본화 비	4 : 2 : 0
화면의 종류	I(Intra), P(Predictive), B(Bidirectional)
부호화 단위	프레임 구조/필드 구조
움직임 벡터 예측	프레임(16×16)/필드(16×8)
움직임 벡터 탐색 범위	-128화소~+127.5 화소, half-pixel 단위
버퍼 크기	1.75 Mbit(1,835,008bit)
호환성	MPEG-1과에 대해 상위 호환성(forward compatibility)
DCT DC 정밀도	8bit/9bit/10bit
VLC 표	MPEG-1과 동일 테이블/신규 테이블
DCT 계수의 스캔 법	MPEG-1과 동일 스캔(zig-zag scan)/신규 스캔
VBR(variable bit rate)	대응함.

1) 압축 알고리즘

저장된 비디오의 무작위 접근을 위하여 MPEG-2에서는 Intra 영상(I-영상)을 사용하고 있으며 이에 대한 압축 기술은 JPEG의 그것과 대동소이하다, 여기에 고압축률을 얻기 위하여 영상과 영상 사이의 상관성을 이용하여 압축을 도모하였고 이를 위하여 연속 영상에서 물체의 움직임을 추정하였으며, 효과적으로 적용하기 위하여 예측된(Predicted) 영상(P-영상)과 양방향 보간된(Bidirectional Interpolated) 영상(B-영상) 등 2가지 영상 타입을 이용하고 있다. I-영상은 무작위 접근을 위한 액세스 점을 제공하나 압축률은 높지 않다. 반면, P-영상은 이전 영상(I 또는 P)을 기

준 영상으로 이용하여 움직임 보상을 함으로써 압축률을 향상시킬 수 있으며, B-영상은 가장 높은 압축을 제공하나 예측을 위해서 이전 영상 및 다음 영상이 기준 영상으로 필요하기 때문에 부호화 순서와 모니터에 도시(display)하는 순서가 서로 다르므로 그에 적합한 알고리즘이 필요하다. 앞서 기술했듯이 I-영상의 압축 기술은 JPEG의 그것과 큰 차이가 없으므로 P, B-영상의 압축 기술만을 논할 것이며, 이에 대하여 좀더 자세히 기술하면 다음과 같다.

가. 움직임 추정

연속 영상으로부터 물체의 움직임을 추정하는 방법으로는 크게 상관성을 이용한 방법(correla-

tion-based method)과 영상의 시공간 방향의 경사를 근거로 한 방법 (gradient-based method)으로 분류할 수 있다. 상관성을 이용한 방법은 영상의 밝기 분포로부터 특징(feature)을 추출한 뒤, 연속 영상에서 그 특징 간의 유사도를 측정하여 움직임을 추정하는 방법으로, 특징으로는 일정한 크기의 영역(region), 선(line), 혹은 물체(object) 등이 사용되며, 이외에 좀더 복잡한 특징을 얻기 위하여 영상을 변환(transform)하기도 한다. 반면, 경사를 근거로 한 방법은 물체의 움직임으로 발생하는 밝기 분포의 시간적 변화와 공간적 구조의 상호 관계로부터 움직임을 추정하는 방법으로 아래와 같은 운동제한식(motion constraint equation)을 이용한다.

$$s_t + s_x v_x + s_y v_y = 0$$

여기서 (s_t, s_x, s_y)는 각각 시공간 방향의 경사이고, (v_x, v_y)는 각 방향의 움직임을 나타낸다.

MPEG-2의 신택스에서는 움직임 벡터를 어떻게 표현하느냐는 명문화하였으나, 어떻게 움직임 벡터를 추정하느냐는 표준화하지 않았다. 그러나 MPEG-2가 블록 단위의 계산(DCT를 기반으로 한 연산)을 기반으로 하고 있기 때문에 상관성을 이용한 방법 중 일정한 영역을 특징으로 사용하는 블록 정합 기법(Block Matching Algorithm, BMA)이 일반적으로 사용되고 있다.

BMA는 영상을 고정된 크기의 영역으로 분할한 뒤, 각 영역 내의 이동벡터는 동일하다고 가정하여 영역 단위의 움직임을 추정하는 방법으로 유사도를 측정하는 비용함수(cost function)로 MAE (Mean Absolute Error), MSE (Mean Square Error) 및 NCCF (Normalized Cross Correlation Function) 등이 사용된다. BMA를 사용하여 움직임을 추정할 때, 일정한 범위에 걸쳐 탐색을 해야 하는데 물체의 움직임이 크다면 계산량이 증가하기 때문에 이의 효율적인 계산을 위하여 MPEG-2에서는 telescopic BMA와 HBMA (Hierarchical BMA)가 연구되었으나, 하드웨어의 구현이 용이한 전영역 탐색(full search) BMA가 사용되고 있다.

나. 움직임 보상

앞서 기술한 기법으로 움직임을 추정된 뒤, 이를 효과적으로 사용하기 위하여 동영상 압축 기술에서는 움직임 보상을 하고 있다. 움직임 보상은 연속 영상 사이의 시간 중복성을 이용하는 기술로 높은 압축율을 기대할 수 있다. 즉, 단순히 현재 영상과 이전 영상과의 DPCM을 하는 것이 아니라 현재 영상과 움직임 보상된 이전 영상과의 DPCM을 함으로써 성능을 향상시킬 수 있다. MPEG-2에서는 이를 위하여 P, B-영상을 두고 있는데 P-영상의 경우는 ITU-T 표준인 H.261과 같은 화상 전화를 위한 압축 알고리즘이 기본이 된다. 반면, B-영상은 이전 영상과 다음 영상 모두를 이용하여 움직임 보상을 하는 것으로 움직임 보상 보간(motion compensated interpolation, MCI) 기법이 사용되었다. MCI는 MPEG-2 압축 기술의 핵심기술이며, 무작위 접근을 향상시키고 동시에 영상 화질을 상당히 개선시키면서 오류 영향을 감소시킨다.

다. 움직임 표현

움직임 정보에 의한 부호화 이득과 움직임 정보를 부호화하는데 필요한 부담을 적절히 절충해야 한다. MPEG-2에서는 움직임 보상 단위로 16×16 블록을 사용한다. 이 블록을 매크로 블록이라 하는데, 각각의 매크로 블록에 관련된 움직임 정보는 이전 블록에 나타나는 움직임 정보에 대해 차등 부호화된다. 차등 움직임 벡터는 화상에 따라 다르게 선택될 수 있고 시간 해상도, 공간 해상도 및 특정 순서에서의 움직임 특성에 따라 선택된다. 최대 범위는 가장 움직임이 많은 상황을 수용할 수 있도록 충분히 크게 선택해야 한다. 차등 움직임 벡터는 움직임 벡터 필드의 높은 공간 상관도를 이용하여 더 높은 부호화 효율을 얻기 위해서 가변길이 부호화된다.

2) Rate control

MPEG-2 부호화에서는 I, P, B 등 다른 종류의 영상 구조를 가지고 있어, 각 영상에서 발생하는 데이터량이 차이가 심하며, 움직임 정도와 같은 영상의 특성에 따라 가변적 데이터량이 발생한다. 따라서 각 영상에 데이터 발생 목표량을 할당하고,

영상의 특성이 변화함에 따라 데이터 발생 목표량을 변화시켜 주어야 한다. 한 화면 내에서도 사람의 시각 특성을 이용하여 결함을 잘 인식하지 못할 부분에는 적은 데이터를 할당시키고, 인식하기 쉬운 부분에는 많은 데이터량을 할당하여 전체적으로 좋은 화질을 느낄 수 있도록 한다. 같은 전송률에서도 화질의 차이를 보일 수 있는 이유는 rate control 방법이 다르기 때문이다. MPEG-2는 기본적으로 복호화 과정에 대한 표준이므로 rate control 알고리즘을 규정하지는 않는다. 단지 MPEG-2의 테스트 모델(test model)에서 사용한 rate control은 부호화할 다음 화면에 대한 데이터량을 할당하고, 가상적인 버퍼를 사용하여 각 매크로 블럭에 대한 양자화 간격의 계산을 위한 기준 값을 추출하여 이를 DCT 계수 양자화에 이용한다.

2. MPEG-4

MPEG-4 활동은 1992년 7월 브라질 리오데자네이로 회의에서 초저속 동영상 부호화를 연구하는 제안이 있는 이후, 1993년 7월 뉴욕 회의에서의 MPEG-4세미나로부터 시작되었다. 그러나 아직 초기 단계에 머물러 있고 활동의 결과도 미미하여 표준안이 성립되기까지는 많은 기간이 필요할 것이다. 표 3은 MPEG-4작업의 일정이다. MPEG-4가 사용될 망의 전송률이 극히 제한적인 관계로 영상의 화질 자체가 만족할 만한 수준에 이르지 못할 가능성이 많다. 따라서 이에 부수되는 오디오의 음질에 따라 전체적인 비디오와 오디오의 복합적 성능이 크게 좌우될 가능성이 있다.

현재 MPEG-4에서는 내용별(Content-based) 부호화, 모델기반(Model-based) 부호화 그리고 프랙탈(Fractal) 부호화 등의 제2세대 부호화 방법이 적극적으로 연구되고 있다.

1) 내용별 부호화

내용별 부호화란 영상의 화질을 영상 내의 내용(content)에 따라 다르게 표시하자는 것이다. 즉, 영상 회의와 같이 같은 화면이라도 회의 발언자의 화질을 향상시킴으로써 같은 전송률 하에서 소기의 목적을 얻을 수 있다. 그러므로 이 방법의 핵심

〈표 3〉 MPEG-4의 작업 일정

Nov. 94	Call for Proposal(Tools, Algorithms, Syntaxes) PPD version 1
Mar. 95	PPD version 2 Definition of the first evaluation process
July 95	Finalize definition of the first evaluation process
Oct. 95	Deadline for submissions
Nov. 95	Evaluations begin Call for proposal(for the second evaluation process)
Jan. 96	Verification Models version 1 First definition of the second evaluation process
Mar. 96	Collaboration on VMs→Version 2 Contributions on tools welcome Finalize definition of the second evaluation process
July 96	VMs improvement process continues
Oct. 96	Deadline for submissions
Nov. 96	Final evaluations begin First Working Draft
July 97	Final version of the Working Draft
Nov. 97	CD
Mar. 98	DIS
Nov. 98	IS

기술은 영상을 내용별로 분할해야 하는 것이다.

제1세대 부호화 기법에서는 블럭 단위로 연산을 하기 때문에 낮은 비트율에서는 블럭화 현상(blocking effect)과 물체의 경계 근처에서 흐려지는 현상(blurring)은 시각적으로 거슬린다. 그러므로 인간의 시각이 물체의 경계에 민감하다는 점을 감안하여 영상을 물체 위주로 분할하여 부호화하는 방법이 내용별 부호화이다.

이와 같이 영상을 내용별로 분할하기 위해서는 기존의 블럭 단위의 연산으로는 불가능하다. 따라서 기존의 많은 연구에서는 영상을 효율적으로 분할하기 위한 기법이 연구되었으며, 연구 결과는 크게 영상의 밝기 값을 이용하는 것과 영상 내의 물체의 움직임 정보를 이용하는 것으로 분류할 수 있다.

영상의 밝기 값을 이용하여 영상을 분할하는 방법에는 영역기반 분할, 경계기반 분할 그리고 히스토그램을 이용한 분할이 있으며, 최근에는 morphology 필터를 이용한 분할 기법이 관심을 받으며 많은 연구가 진행되고 있다.

한편 물체의 움직임을 이용하여 내용별로 분할하는 기법은 MPEG-4의 표준안으로 지금까지 가장 관심이 집중되었던 방식이다. 그 기본 개념은 먼저 영상을 움직임이 있는 물체(object)들로 분할한 뒤, 각 물체를 움직임(motion parameter), 윤곽선(contour) 그리고 색 변수(color parameter)로 표현하는 것으로 물체의 움직임을 추정할 때 사용하는 소스 모델(source model)에 따라 그 정확도에서 차이가 있다. 즉, 물체의 움직임을 고려할 때 물체가 단순히 2차원 움직임을 하였는가, 아니면 복잡한 3차원 움직임을 하였는가. 그리고 물체의 성질을 강체(rigid body)로 볼 것인지 아니면 비강체(flexible body)로 볼 것인지에 따라 달라진다. 대부분의 방법에서는 물체를 강체로 보고 3차원 움직임까지 고려하여 6, 8개의 매개 변수를 사용한다.

이 방법은 일정 크기의 블록 또는 매크로 블록 단위로 처리하는 H.261, MPEG-2 등에 비해 물체의 경계가 잘 보존되고 물체 위주의 예측 성능이 뛰어난 장점이 있다. 그러나 영상을 내용별로 분할하는 과정에서 많은 오차가 발생할 수 있으므로 물체 표현이 용이하지 않다. 현재 연구되고 있는 MPEG-4의 제2세대 부호화 방법 가운데 초저속에서 가장 우수한 결과를 내고 있으며, 장면 전환시의 영상 내 부호화(intra coding)에 소요되는 비트를 고려한 비트율 조절 및 화질 저하 방지에 대한 연구가 진행되고 있다.

2) 모델기반 부호화

모델기반 부호화는 정보 이론에 근거한 기존의 부호화 방법과는 달리 영상 분석과 컴퓨터 그래픽을 접목시킨 부호화 방법이다. 즉, 이 방법은 물체의 워핑(warping)과정을 기반으로 한 것으로 열만전 관심을 끌었던 국내 영화 '구미호'에서 사람이 구미호로 변하는 과정에서 사용한 방법과 유사하다. 그러므로 이 방법에선 물체의 표현을 위한

모델을 우선적으로 만들어야 한다. 특히 인간 얼굴의 외형적인 특징 및 해부학적인 해석까지 도입한 모델을 비롯하여 수많은 삼각형 조각으로 3차원 얼굴 모델을 구현한 철선 구조 모델(wire-frame model)이 일반적으로 사용되고 있다. 이 방식은 부호기/복호기 양측이 동일한 3차원 얼굴 모델을 갖고 있음을 전제로 한다.

입력 영상이 들어오면 먼저 일반적인 철선 구조 모델을 대상 영상에 정합시키며, 얼굴의 특징점인 눈, 코, 입 그리고 귀 등의 위치 정보를 이용하여 모델을 조절한다. 이 방법의 장점은 영상을 보내지 않고 단지 모델의 변수만 보내면 된다는 것이다. 그러나 이 방법의 문제점은 모델링되지 않은 물체 즉, 다른 동물이나 구름 등을 어떻게 처리할 것인가 하는 것이다. 다시 말하면, 표현하고자 하는 모든 종류에 대한 모델이 필요하다는 것이다. 그리고, 또 하나의 문제점은 영상 분석이 완벽할 수 없으므로 이에 따른 오류로 인한 영향이 심각할 수 있다.

3) 프랙탈 부호화

프랙탈 부호화는 수학기론에서 도입된 것으로 영상을 아주 작은 단위로 나누어 보면 각 단위는 유사한 형태로 표현될 수 있다는 가정하에서 공통으로 쓸 수 있는 기본 형태를 찾는 것으로 이 방법은 수행된다. 일정한 크기의 블록 영상을 기본 형태로 하여 이것을 확대, 회전 등의 기하학적 변환 및 밝기 값의 scaling 등을 통하여 전체 영상 내에서 유사성을 조사하여 기본 형태와 변환 정도로 전 영상을 표현할 수 있다. 프레임 내의 유사성만을 이용한 프레임 내 처리에선 압축률이 높을 경우, DCT보다 우수함이 입증되어 있다. 그러나 프레임 간 처리를 위한 프레임 간 유사성에 대한 연구가 더 필요하다.

IV. 음성 부호화 기법

음성의 디지털 부호화 기술은 크게 세 가지로 분류할 수 있다. 첫째, 음성 파형을 샘플링하여 양자

화하는 파형 부호화(waveform coding)와 음성의 주기와 성도의 계수 등 음성의 특징만 추출하여 전송해서 수신측에서 음성을 재생하는 분석, 합성에 의한 보코딩(vocoding) 방식, 그리고 파형 부호화 방식과 분석, 합성 방식의 이점만 사용하는 혼합 부호화(hybrid coding) 방식이다.

파형 부호화 방식은 전송 속도가 16k~64kbps로서 비교적 높지만 음질이 우수하여 일반 음성 통신에 많이 사용되고 있다. 파형 부호화 방식 중 현재 가장 많이 사용되는 것은 PCM(Pulse Code Modulation) 방식이나 전송 속도는 64kbps로 대역폭의 사용면에서 볼 때 아날로그 방식보다 비경제적이다. 따라서 음성의 대역폭의 축소를 위하여 음성 신호의 중복성을 이용한 예측 부호화 방식이 제안되었으며, 대표적인 예로 ADPCM(Adaptive Differential PCM)과 ADM(Adaptive Delta Modulation)을 들 수 있다.

혼합 부호화 방식은 전송 속도가 4.8k~16kbps로 비교적 낮기 때문에 모뎀을 사용해서 기존의 아날로그 회선으로 음성은 전송할 수 있는 이점이 있지만 음질은 일반적으로 파형 부호화 방식보다 떨어진다. 혼합 부호화 방식은 크게 시간 영역의 혼합 부호화와 주파수 영역의 혼합 부호화 방식으로 나눌 수 있다. 전자의 대표적인 예로는 RELP(Residual Excited Linear Prediction), APC(Adaptive Predictive Coder) 및 최근 각광을 받고 있는 CELP(Code Excited Linear Prediction), MPLPC(Multi-Pulse LPC)와 RPE-LPC(Regular Pulse Excited LPC) 등을 들 수 있다. 그리고 후자의 대표적인 예로는 SBC(Sub-Band Coding)와 ATC(Adaptive Transform Coding)가 있다. 이들 혼합 부호화 방식들은 파형 부호화에 비해 전송 속도가 훨씬 낮지만 시스템이 복잡하다는 것이 큰 단점이다.

한편, 보코딩 방식은 음성 파형을 직접 양자화하지 않고 음성 파형을 분석하여 유/무성음 구별, 기존 주기, 성도의 계수 등 음성의 특징만을 추출하여 전송하기 때문에 전송 속도가 50~4.8kbps로서 매우 낮지만 부호기가 복잡하고 음질에 아직도 문제점이 있는 것이 단점이다. 혼합 부호화나 보코딩

방식은 현재 일반 상용 음성 통신보다는 군 통신 등 특수 통신에 사용되고 있으며, 많은 연구 결과 음질이 점차 개선되고 있다.

V. 결 론

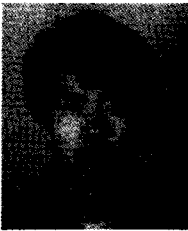
본 고에서는 전반적인 영상 압축 기술에 대하여 살펴보았다. 정지 영상의 압축 기술로 컬러 영상을 위한 JPEG과 흑백 영상을 위한 JBIG이 표준화되어 여러 분야에 효과적으로 사용되고 있으며, 특히 인터넷의 WWW를 통하여 다양한 형태의 응용을 볼 수 있다. 한편, 동영상 표준화 그룹 MPEG에서는 전송률이 높은 알고리즘에 대하여는 어느 정도 표준화가 이루어져 있으나(MPEG-1, MPEG-2), 수십 kbps 정도의 낮은 전송률에서는 이제 시작하는 단계이다. 따라서 우리도 나름대로의 알고리즘을 정립하여 표준화에 기여하는 것이 바람직하다고 여겨진다.

참 고 문 헌

- [1] CCITT/ISO/IEC, "International Standard JBIG Coded Representation of Picture & Audio Information Progressive Bi-level Image Compression Standard", Sep. 1990.
- [2] JPEG IS, "Information Technology-Digital Compression and Coding of Continuous-Tone Still Images," ISO/IEC 10918, 1993.
- [3] MPEG-1 IS, "Information Technology-Coding of moving pictures and associated audio for digital storage media up to 1.5Mbit/s," ISO/IEC 11172-1,2,3, 1993.
- [4] MPEG-2 DIS, "Information Technology-Coding of moving pictures and associated audio," ISO/IEC 13818-1,2,3, 1994.

- [5] ISO/IEC JTC1/SC29/WG11, "Coding of Moving Pictures and associated audio," MPEG 94, Nov. 1994.

저 자 소 개



金 二 漢

1966年 1月 29日生

1988年 忠南大學教 電子工學科(工學士)

1990年 韓國科學技術院 電氣 및 電子工學科(工學碩士)

1995年 韓國科學技術院 電氣 및 電子工學科(工學博士)

1995年~現在

韓國科學技術院 電氣 및 電子工學科 研究員

주관심 분야 : 영상처리, 영상통신, 멀티미디어 등



金 聖 大

1953年 12月 26日生

1977年 서울大學教 電子工學科(工學士)

1979年 韓國科學技術院 電氣 및 電子工學科(工學碩士)

1983年 프랑스 INPT ENSEEIHT(工學博士)

1984年~現在

韓國科學技術院 電氣 및 電子工學科 教授

주관심 분야 : 영상처리, 영상통신, 컴퓨터 비전, VLSI 구현 등