

영상신호의 압축 기술 및 전개 방향



박 구 만

삼성전자(주) 신호처리연구소
선임연구원

1. 머리말

영상압축기법은 지난 10여년간 JPEG, H.261, MPEG-1, MPEG-2, MPEG-4 등으로 이어지는 표준화에 의해 실용기술로서 기반이 확대되어 주변 생활에 적용되기 시작하여 일반 소비자들까지도 기술의 혜택을 누리는 정도가 되었다.

지난 10여년간 영상/음향 신호처리기술, 통신 기술, 반도체 기술, 소프트웨어 기술 등의 꾸준한 발전 및 전송로 또는 저장매체의 용량 증대에 힘입어 광역 통신 서비스가 주위에서 서서히 실현되어 가고 있다.

이는 멀티미디어 서비스의 제공을 의미하고 그 핵심인 영상-음향(Audio-visual) 정보의 접근이 손쉬워진다는 것을 뜻한다.

특히 정보전달 효과가 가장 큰 것은 영상이다.

고품위 영상·음향서비스가 가능하기 위해서는 영상·음향(특히 영상)대역 압축기술이 뒷받침되어야 한다. 통신망에서 영상이란 디지털 영상을

의미한다. 비록 디지털 영상이 다양한 서비스와 조작성이 가능해도 디지털로 표현했을 때 방대한 비트(bit) 소모량 때문에 전송 또는 저장에 어려움이 많아진다.

따라서 영상신호가 가지고 있는 공간적, 시간적, 확률적인 중복성(Redundancy)을 제거하는 신호원 압축을 하여 필요 비트량을 낮추어야 한다.

본 기고에서는 디지털 영상신호의 특성과 기존의 압축부호화 기술을 간단히 소개하고, 향후 전개될 압축부호화 기술 및 부호화에 관련된 중요한 국제표준안들의 흐름을 소개하였다.

또한 이들 흐름에 따른 응용분야도 소개하였다.

2. 디지털 영상 및 압축 부호화 기술

1) 압축의 필요성

영상신호 압축의 필요성은 다음의 예들을 통해서 이해할 수 있다. 36mm 컬러 사진필름을 12 μ m 마다 표본화하여 디지털 영상을 얻으면 영상의 크기는 대략 가로 3000개, 세로 3000개의 화소(picture element)들로 구성된다.

각 화소마다 24비트로 표현하면 영상 하나를 위해서 205Mbits가 필요하다. 요사이 PC통신에서 많이 쓰는 1400bps(bits per second) 모뎀으로 전송을 하면 250분 정도의 긴 시간이 걸린다.

또한 우리나라의 위성방송 방식으로도 채택된 MPEG-2 표준안을 예로 들기로 한다. 한 화면의 유효크기가 가로 720화소, 세로 480 화소이고 초당 29.97개의 화면이 발생하는 방송신호는 스튜디오에서 약 166Mbps정도가 된다.

이를 9Mbps정도로 전송하려면 약 18:1 정도

의 압축을 해야 한다. 이때 수신된 영상의 품질은 압축하기 전의 영상과 거의 구분이 되지 않아야 한다. 여기서 음향신호도 압축하여 전송한다.

음향신호는 48KHz의 표본화 주파수를 사용하고 표본당 16 비트이며 2개 채널을 사용한다고 가정한다면 초당 발생하는 비트량은 약 1.5Mbps 정도 발생한다. 만약 256Kbps로 전송한다면 약 6:1 정도의 압축을 해야 한다. 음향신호는 왜곡이 생겼을 때 우리의 귀에 고통을 주기 때문에 압축률을 높이지 않는다.

여기서 영상신호와 음향신호의 발생량을 비교하면 약 100배 정도의 차이가 난다. 따라서 데이터 압축의 비중은 영상이 클 수 밖에 없다.

2) 중복성 제거

앞에서도 언급한 바와 같이 3가지 중복성을 제거하여 영상신호를 압축 부호화할 수 있다.

가. 공간적 중복성

화면을 이루는 최소의 화면 구성요소를 화소라고 부른다. 수평과 수직방향으로 가까이 있는 화소일수록 비슷한 값을 가진다. 화소들간에 DPCM(Differential pulse code modulation)과 같은 간단한 기법만 사용해도 압축 효과를 얻을 수 있다.

기존의 부호화 방식에서 가장 널리 사용되는 DCT(Discrete cosine transform)는 영상신호를 변화시켜 계수들로 표현해 준다. 계수들간에는 중복성이 없으며, 자연적이고 일반적인 영상은 소수 몇개의 계수에 영상 전체 성질이 포함되므로 이 둘만을 택하면 압축이 가능해진다.

나. 시간적 중복성

NTSC TV신호는 초당 29.97개의 화면으로 구성되어 있다. 대개의 경우, 앞뒤의 가까운 화면일수록 비슷한 장면을 담는다. 현재 부호화하려는 화면의 내용 중에서 이미 앞의 화면에서 부

호화한 것들과 중복되는 것들이 있으면 이들은 제외하고 중복되지 않는 내용만을 부호화할 수 있다.

앞뒤 화면에서 특정 부분(블럭)간의 움직임 추정(Motion Estimation)을 한 후, 앞 화면의 내용으로 현재 화면의 부분을 대체시킨다.

이때 대체시키는 것을 움직임 보상(Compensation)이라고 하고 시간적 중복성의 가장 중요한 요소가 된다. 대체시킨 블럭과 원래 블럭간에는 오차가 있고 이 오차값만을 부호화한다. 이것은 시간축으로 DPCM을 하는 것과 같은 의미이다.

이로써 높은 압축효과를 얻는다. 뒤에 언급할 기존의 국제표준화 방식은 공간적으로는 DCT를 사용하고 시간적으로 DPCM을 이용하는 혼합형식(Hybrid)이다.

다. 확률적 중복성

발생 확률이 제각기 다른 심볼로 구성된 유한 집합이 있고, 심볼들의 발생 확률을 더하면 1이 되는 집합이 있다. 이 때 발생 확률이 높은 심볼은 짧은 비트열로 표현하고 발생 확률이 낮은 심볼은 상대적으로 긴 비트열로 표현하는 가변길이 부호(variable length code)를 사용하면 모든 심볼에 대해 단순히 같은 길이의 이진부호로 표현했을 때 보다 평균 비트 소모량을 줄일 수 있다.

발생 확률에 따라 부호의 길이를 다르게 사용하는 방법을 엔트로피 부호화라고 하며, 확률적 중복성을 제거하는 것을 의미한다. 대표적인 가변길이 부호인 허프만부호(Huffman code)가 가장 많이 사용된다. 지금까지 언급한 세가지 중복성 제거 방법이 압축기술의 근간을 이루고 있다.

3) 무손실압축 및 손실 압축(lossless compression and lossy compression)

사용하는 부호화 방식에 따라 압축후 복원한 영상이 원래의 영상과 동일하거나 또는 훼손될 수도 있다.

무손실 압축이란 원래의 영상과 압축했던 영상이 모든 화소가 일치되는 것을 의미한다. 앞에서 언급한 엔트로피 부호화가 이에 해당한다. 원래 영상의 품질, 디지털 상태의 데이터를 그대로 유지하면서 저장 또는 전송에 필요한 필요 비트수를 줄이는 것이다. 보통 PC에서의 압축프로그램이 단적인 예이다. 스튜디오 등에서 방송원본 필름이나 비디오를 처리할 때 쓰일 수 있다.

DCT부호화의 경우 DCT계수들중에서 저주파 범위에 있는 일부 계수들만을 선택하고 나머지를 버려도 영상을 복원했을 때 원래 영상의 에너지 대부분을 포함 할 수 있다.

하지만 복원후 원래 영상과 비교하면 값이 다르게 된다. 단지 눈으로 비교할 때 왜곡이 별로 나타나지 않을 뿐이다. 이를 손실 압축이라고 한다.

영상부호화에서 우리가 추구하는 것은 수치적으로는 손실이 있지만 시각적으로는 손실을 발견하기 어려운 시각적 무손실 압축 방법이다.

부호화했던 결과 영상을 평가하는 방법은 객관적 방법과 주관적 방법의 두가지가 있다. 객관적 방법은 원래영상과 결과영상의 차이값을 계산하여 결과영상의 충실도를 PSNR(Peak to peak signal to noise ratio)와 같은 수치로 표현하는 것이다.

반면에 주관적 평가 방법은 사람의 눈으로써 결과영상의 만족도를 표현하는 것이다. 비록 수치는 낮아도 눈으로 보았을 때 만족스러우면 옳은 결과를 얻은 것이다.

4) 기존 압축기술의 종류

개별적인 화소가 모인 파형을 대상으로 통계적

특성을 찾고 또한 특정 모델을 가정하여 부호화한다. 단일 화소, 화소의 블럭 또는 이들의 변환 영역(DCT)값들이 부호화 하기 위한 재료가 된다.

예를 들면 DPCM, DCT, 부대역(subband) 및 웨이블릿(wavelet) 변환, 벡터양자화 등이다. 움직임 추정 및 보상 기술도 화소에 기초하고 있다.

화면간의 움직임 추정 및 보상은 특정 크기의 화소 블럭단위로 이루어진다. 움직임 보상은 동화상 압축의 가장 중요한 수단이고, MPEG (Moving Picture Experts Group)등의 국제표준안은 블럭단위의 DPCM을 사용한다. 움직임도 평행이동만을 가정하였다.

3. 새로운 영상 압축 기술 및 국제 표준안의 전개방향

1) 새로운 영상 압축 기술

기존의 예측과 변환기법을 사용한 전통적인 방법으로부터 벗어나 새로운 시도가 지난 10여년 전부터 본격화되었다.

새로운 압축기술의 우수성이 입증되기도 하지만 아직은 특정 모델을 가정한 경우이므로 일반적인 경우로의 확대가 필요하다. 또한 압축 기술이 상업적으로 관심을 받으면서 국제 표준화가 비슷한 시기에 시작되었기 때문에 성능이 입증되고 실용화 개발이 잘 된 기존의 압축 기술이 국제표준안에서 채택되었다.

하지만 초저속 부호화기, 즉 고압축 부호화기에서 기존 방식의 한계가 인정되기 때문에 새로운 압축기술은 필수적이다. 새로운 압축기술을 간단히 소개하기로 한다.

가. 물체중심 분석·합성 부호화 (object oriented analysis synthesis coding)기법

비디오의 각 화면을 분석해서 움직임이 없는

정적객체(배경)와 3차원(또는 2차원)의 움직임 객체들로 구분한 후, 움직임 객체들의 형태, 움직임 및 칼라 정보에 관한 파라미터들을 추출하고 전송하는 기법이다.

여기서 형태는 객체 영역의 개략적인 윤곽 정보를 말하며, 칼라는 객체 영역내부의 질감을 말한다. 모델기반기법에서 움직임 객체는 기지의 3차원 객체라고 가정하지만, 객체기반 기법에서 움직임 객체들은 미지의 3차원(또는 2차원)객체들이라고 가정한다.

나. 분할기반 부호화(segment-based coding) 기법

영상을 일정한 크기의 블록들로 나누어 부호화하는 것을 블록 부호화라고 하며, 영상내에서 동질성을 갖는 영역들로 분할한 후 분할된 각 영역의 윤곽 정보(모양, 위치) 및 내부의 질감 정보를 부호화하는 것을 분할기반 부호화라고 한다.

분할 기반 비디오 압축기법에서는 영상을 정적영역(배경영역)과 움직임 영역들로 분할한 후, 움직임 영역들의 윤곽, 움직임 및 질감에 관한 파라미터들을 추출하고 전송한다.

모델기반 및 객체기반 기법과는 달리 움직임 영역에 관한 사전지식이 전혀 요구되지 않는다.

다. 모델기반 기술

화면내의 객체를 의미론적(semantic) 표현을 하여 파라미터로 저장해 놓는다 영상이 입력되면 영상을 파라미터로 표현하여 이를 전송한다.

얼굴표정이 가장 대표적인 부호화 대상이다. 얼굴, 어깨 등의 형상에 관한 3차원 모델과 지식을 송신측과 수신측에서 함께 준비해 놓고, 송신측에서는 그 모델과 지식에 근거하여 입력영상을 분석 인식하여 움직임과 모양 또는 내용 변화만이 송신 파라미터로서 전송된다.

수신측에서는 인식결과에 따라 변형된 동영상을 합성해 낸다.

이때, 분석에는 주로 영상처리 및 컴퓨터 비전 기술이 사용되며 합성에는 컴퓨터 그래픽스 기술이 주로 사용된다.

라. 프랙탈(fractal) 기술

프랙탈 영상압축기법은 자기유사성(self-similarity)이라는 특성을 갖는 프랙탈 기하 이론에 근거한 압축 기법으로, 영상내의 각 블록들을 축소변환 파라미터들로 표현함으로써 영상을 부호화한다.

2) 부호화 형식의 새로운 전개

전통적으로 압축 부호화는 제한된 초당 전송률을 맞추기 위해 버퍼를 이용하거나 또는 미리 부호화하기 전에 영상마다 표현 비트량을 할당시키는 방법으로 고정 비트율 전송을 가능하게 했다.

화면내에서 물체의 움직임이 크면 그만큼 많은 비트량을 소모한다.

따라서 출력 비트율이 심하게 변동한다. 고정 비트율을 유지하기 위해서는 영상의 품질에 영향을 주는 압축부호화 파라미터(예로서 양자화 단계값)를 조절하여서 달성하였다.

이때 화질은 움직임이 많을수록 떨어진다. 고정비트율은 화질의 변화를 보인다.

반면에 전송채널에서 일정한 범위내에서 전송 비트율을 변동시킬 수 있으면 화질을 항상 일정하게 유지할 수 있다. 지난 수년간 이 분야가 중요한 연구 주제로 여겨져 왔다.

가변비트를 전송방식이 BISDN(broadband Integrated services digital network)의 핵심인 ATM(asynchronous transfer mode)망에서 채용하고자 하는 방안이다. 또 다른 중요한 흐름은 초저속 부호화라고 할 수 있다.

즉, 고압축률의 실현이다. 현재 진행중인 MPEG-4의 표준화는 이를 달성하기 위한 것이다. 64Kbps 이내에서의 비디오 통신에 적합한 부

호화 방식을 표준화하는 것이다.

혼합형(Hybrid) DCT/DPCM을 사용할 지 아니면 객체기반 또는 모델기반 부호화 방식을 사용할 지는 결정되지 않았다.

초저속이기 때문에 상업적인 성공은 비디오를 부호화하는 능력에 달려있다.

기존의 기술에 의한 표준화 방식들은 64Kbps 이상에서 좋은 비디오 품질에 도달하기 때문에 이대로는 충분하지 않다.

새로운 부호화 방식의 개발 또는 충분한 개선이 필요하다. 이와 같은 초저속 부호화기는 기존의 PSTM(public switched telephone network) 또는 미래의 저대역 전송로에서 비디오 정보를 전달하기 위함이다.

3) 국제표준안의 흐름

정지영상 부호화를 목적으로 한 JPEG(Joint Photographic Experts Group), 비디오 전화기 또는 화상 회의에 사용할 목적으로 마련된 H, 261, 디지털 저장매체에 기록을 목적으로 한 MPEG-1, 저장, 방송 등 용도가 크게 넓어진 MPEG-2 등의 표준안은 1970년대의 제안된 DCT와 움직임 보상기술을 이용한 결실이라고 할 수 있다.

표준화는 초기에는 특정 응용을 목적으로 하는 경향이 강했지만 MPEG-2와 MPEG-4에 이르면서 표준이 특정 응용으로 부터 독립적이면서 용도가 넓어지는 일반적(generic) 목표를 가지게 되었다.

표준은 도구모음의 역할을 하고 다른 응용간의 용이한 데이터(bitstream)교환을 위해 데이터의 표현수단(syntax)만을 제공만하고 어떻게 부호화해야 할 지는 응용하고자 하는 사용자에게 맡긴다.

따라서 표준은 복호화 규칙만을 정의한다. 이에 따라 표준의 상업성이 더욱 증대되고 있다.

MPEG-2의 경우에는 각각의 응용이 모든 표현을 실현할 수 없으므로 프로파일(profile)과 단계(level)를 정의하여 사용목적에 맞는 프로파일과 단계를 선택하게 하였다.

프로파일은 기능에 관련된 의미로서, 각 프로파일마다 약간씩 차이가 나거나 또는 정교한 정도에서 차이가 나는 알고리즘을 의미하고, 레벨은 영상의 크기, 초당 화면수, 초당 전송 비트량과 같은 해상도를 의미한다.

MPEG-4에서는 이것이 더욱 발전되어 표준화 방향이 2가지의 커다란 줄기로 가지게 되었다.

첫째는 각각의 응용분야에서 공통의, 혹은 상호 필요로 하는 정보매체(2차원 또는 3차원 음향영상 내용물)를 효율적으로 공유할 수 있도록 해주는 MSDL(MPEG-4 Syntactic Description Language)의 개발이고 둘째는 MSDL에 적합하게 응용될 수 있으면서 MPEG-4에서 추구하는 각종의 기능(functionally)들을 가능한한 많이 만족시켜 줄 수 있는 음향/영상 부호화 기술의 개발이라고 볼 수 있다.

여기서 MPEG-4의 응용분야는 영상-음향 데이터베이스 접근, 유무선 영상-음향 통신 및 교환(audio-visual communications and messaging), 원격감시 및 제어(remote monitoring and control)등이다.

5. 맺음말

본 기고에서는 멀티미디어 서비스의 핵심인 디지털 영상신호의 특성과 압축원리 및 종류를 간단히 설명하였으며, 기존의 압축부호화 기술 및 새로운 영상 압축 기술의 간단한 소개와 함께 국제 표준화의 전개 방향을 통해 전반적인 압축부호화의 방향을 파악하였다.