



영상 부호화 기술 분야

김병진*, 박현욱

(*LG전자(주) 영상미디어 연구소 책임연구원, **KAIST 정보및 통신공학과 부교수)

1. 서론

최근 들어 디지털 영상 데이터의 전송, 저장 및 처리에 대한 요구가 급격히 증가 하면서 영상 정보의 다양한 전송 및 저장을 위한 영상 부호화 연구가 국내외 적으로 활발하게 진행되고 있다. 특히 영상의 정보 손실을 최소로 하는 영상 압축 기법 분야는 국제 표준화 기구를 통하여 각 응용 분야에 따라 이미 표준화 되거나 진행되고 있으며, 이 표준화에 따라 영상 압축 시스템 및 영상 복원 시스템의 개발이 행해지고 있다. 여러 표준화와 관련된 응용분야로는 영상전화, 영상회의, 영상감시, TV방송, 비디오텍스, 영상우편 등을 들 수 있으며, 각 응용분야에 따라 영상의 규격, 비트율, 요구 화질등이 서로 다르다.

본고는 이제까지 연구 되어온 영상 압축에 대한 내용 및 적용을 소개 함으로써 영상 압축에 대한 이해를 높이고자 하는데 그 목적이 있으며, 구성은 영상 압축의 필요성, 영상 압축에서 사용하는 정보 이론에 대한 개론, 무손실 영상 압축 기법, 손실 영상 압축 기법에 관하여 설명하고자 한다.

2. 영상 압축의 필요성

아날로그 영상을 샘플링(Sampling) 및 양자화에 의해 디지털 정보로 변환시 발생하는 정보량 즉 비트수는 상상외로 크다. 가령 TV 화질 정도의 저 해상도를 갖는 칼라 비디오 영상(512x512픽셀/칼라, 8비트/픽셀, 3칼라) 1장을 디지털로 변환했을 경우, 발생된 비트량은 이의 약 100,000배의 비트량이 필요하다. 이는 영상 정보의 저장 및 전송을 위해서 다량의 메모리 및 대용량의 채널(Channel)이 필요하다는 것을 의미하는 것으로 실제 응용에서는 영상 압축이 필수적이다. 영상 압축은 영상속에 포함되어 있는 3가지 종류의 리던던시(Redundancy)를 이용하며 이를 정리하면 다음과 같다.

- 공간 리던던시 : 인접 픽셀값 사이의 유사성
- 색 리던던시 : RGB 칼라 사이의 유사성
- 시간 리던던시 : 시간축상 영상 사이의 유사성

영상 압축에 대한 연구는 상기와 같은 리던던시를 제거함으로써, 영상을 표현하는데 필요로 하는 비트수를 줄이는 데 그 목적이 있으며, 여러가지 응용에 따라 다양한 기법이 존재하나 크게 무손실 압축과 손실 압축으로 나눌 수 있다.

무손실 압축이란 압축된 영상을 복원 했을 때 복원된 영상이 원 영상과 동일하도록 하는 압축 방식으로, 원 영상이 갖고 있는 정보를 전혀 잃지 않는 장점을 갖고 있으나 압축률이 작다는 단점을 갖고 있다.

손실 압축은 복원된 영상이 원 영상과 비교 했을 때 얼마간의 화질 저하를 발생 시키는 압축 방식으로, 비록 화질 열화는 존재하지만 무손실 압축에 비해 고율로 압축이 가능하다는 장점이 있다. 일반적으로 높은 압축은 그에 상응하는 화질 열화를 수반한다. 여기서 중요한 점은 이러한 화질 열화가 우리 눈에 보여 지느냐 그렇지 않느냐 하는 점이다. 즉 손실 압축 기법의 성능을 결정하는 요소중 하나는 영상 압축으로 발생된 화질 열화를 일반적 조건 상황에서 사람 눈에 띄지 않게 하도록 하는 것이다.

이와 같은 분류는 원 영상의 성질에 따라 보다 더 세분되어 질 수 있다. 압축하려는 영상이 문서와 같은 2진값을 갖는 바이너리(Binary) 영상일 수 있고 8비트로 샘플링된 연속적 톤(Tone)을 갖는 비디오 영상일 수도 있다. 또는 공간 리던던시를 포함하는(칼라 영상인 경우는 색 리던던시도 포함) 정지 화상일 수도 있고 시간 리던던시까지 포함하는 동 영상일 수도 있다. 결국 압축하고자 하는 영상에 따라 화상에 포함되어 있는 리던던시는 서로 다르며 각 영상에 따라 적절하게 각 영상에 포함되어 있는 리던던시를 제거 함으로써 목표로 하는 압축비를 얻을 수 있다. 본 고에서는 연속적 톤을 갖는 흑백 및 칼라의 정지 영상에 사용되는 무손실 압축 및 손실 압축 기법에 대하여 설명하고 있으며, 본 고에서 언급되어 지는 각 기법은 또한 동 영상에도 적용될 수 있다.

3. 정보 이론 개요

3.1 소스 모델(Source Model)과 엔트로피(Entropy)
정보 소스에 있어서 가장 간단한 형태는 이산 무기억 소스

(Discrete Memoryless Source)로 이는 통계적으로 서로 독립인 소스에 의해 발생되어진 연속적 심볼(Symbol)을 의미하며, 임의의 이산 무 기억 소스 S는 소스 알파벳 $S=\{s_1, s_2, \dots, s_n\}$ 과 각 심볼의 발생 확률 $\{p(s_1), p(s_2), \dots, p(s_n)\}$ 에 의해 특성지어 질 수 있다. 여기서 정보 소스에 있어 임의의 심볼 s_i 에 대한 정보량 $I(s_i)$ 를 정의하면 다음과 같다.

$$I(s_i) = \log \frac{1}{p(s_i)} \quad (1)$$

상기 식에서 로그 함수의 베이스(Base)는 정보량을 표현하는데 필요한 단위를 결정한다. 즉 만약 로그 함수의 베이스가 2라면 정보량은 2진수 즉 비트로 표현됨을 의미한다. 여기서 이산 무 기억 소스의 모든 가능한 소스 심볼의 정보량에 대한 평균값 즉 소스 심볼당 평균 정보량을 구할 수 있는데 이를 정보 이론에서는 엔트로피(Entropy) $H(s)$ 라 부르고 아래와 같이 표현한다.

$$H(s) = \sum_{i=1}^n p(s_i)I(s_i) = - \sum_{i=1}^n p(s_i) \log_2(s_i) \quad (2)$$

여기서 단위는 비트/심볼 이다.

위에서 설명한 이산 무 기억 소스는 실제 응용에 있어서 많은 제약이 따른다. 왜냐하면 일반적으로 실제 응용에 있어서 인접 심볼간은 서로 많은 상관성을 갖고 있기 때문이다. 따라서 이산 기억 소스를 정의할 필요가 있는데 이를 마코프(Markov) 소스라 부르며 마코프 소스는 차수를 갖는다. 즉 m-차 마코프 소스라 함은 임의의 소스 심볼 s_i 의 발생 확률이 해당 심볼 바로 직전에 발생한 m개의 심볼에 의해 영향을 받는다는 것을 의미한다. 이는 조건부 확률로 정의할 수 있으며 아래와 같다.

$$p(s_i | s_{j_1}, s_{j_2}, \dots, s_{j_m}) \quad (3)$$

단, $i, j_p (p=1, \dots, m) = 1, 2, \dots, n$

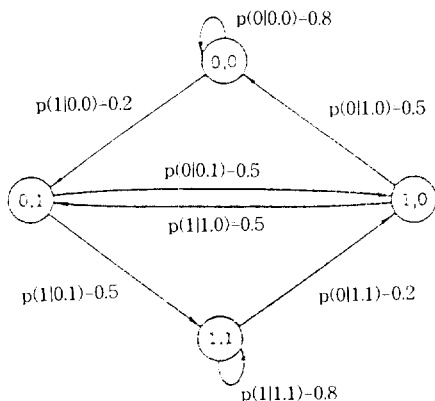


그림 1. 마코프 프로세스 상태도

마코프 소스는 앞서 발생한 m개의 심볼에 대한 상대도(State Diagram)로 표시하면 이해가 편리하다.

일반적으로 m-차 마코프 소스는 n^m 상태(State)를 가지며, 아래 일레로 2진 알파벳을 갖는 2차 마코프 소스를 상태로 표시하였다.

m-차 마코프 소스에 대한 엔트로피를 계산하기 위하여는 먼저 각 상태 $(s_{j_1}, s_{j_2}, \dots, s_{j_m})$ 에 대한 소스 엔트로피 $H(S | s_{j_1}, s_{j_2}, \dots, s_{j_m})$ 를 구해야 하며 구하는 수식은 다음과 같다.

$$H(S | s_{j_1}, s_{j_2}, \dots, s_{j_m}) = - \sum_{i=0}^n p(s_{j_1}, \dots, s_{j_m}) \log p(s_i | s_{j_1}, \dots, s_{j_m}) \quad (4)$$

그런 다음 위에서 구한 조건부 엔트로피는 해당 상태에 대응하는 확률에 의해 조정된 후, 모든 가능한 상태의 합이 m-차 마코프 소스에 대한 엔트로피가 된다. 즉

$$H(S) = \sum_{S^m} p(s_{j_1}, s_{j_2}, \dots, s_{j_m}) H(S | s_{j_1}, s_{j_2}, \dots, s_{j_m}) \quad (5)$$

3.2 가변장 부호(Variable Length Codes)

가변장 부호는 엔트로피 부호라고도 불리며 크게 허프만(Huffman) 부호와 산술 부호로 나뉘어진다. 일반적 응용에서는 허프만 부호가 가장 많이 사용되며, 기본 개념은 발생 확률이 높은 심볼에 대하여는 짧은 길이의 부호어를, 발생 확률이 낮은 심볼에 대하여는 상대적으로 긴 부호어를 할당하여 평균 부호어 길이를 짧게 하는 것이다. 그림 2는 심볼이 5인 허프만 부호화의 작성예를 나타내고 있다.

심볼	확률	순서	부호어
0	0.40	0	1
1	0.20	0	000
2	0.15	1	001
3	0.15	0	010
4	0.10	1	011

그림 2. 허프만 부호화의 일실시에

우선 심볼의 출현 확률이 작은 것을 2개 선택하여 2개중 발생 확률이 작은 것에 1을 큰 것에 0을 할당하고 나서 2개를 통합하여 하나의 심볼로 만든다. 이때 새로 만들어지 심볼의 발생 확률은 2심볼의 합이 된다. 이와 같이 하여 처음 5개의 심볼은 4개로 된다. 다음에 4개의 심볼을 다시 출현 확률 순에 따라 2개를 선택한 후, 발생 확률이 작은것에 1을 큰것에 0을 할당하고 2심볼을 통합하여 하나의 심볼로 만든다. 이와같은 처리를 심볼수가 2개 남을때 까지 반복한다. 할당된 부호를 역순으로 호출하여 그것에 대응시키면 그것이 심볼의 부호어가 된다.

상기 허프만 부호는 부호화 방법이 간단하고 소스 엔트로피에 비교적 가까운 평균 비트율을 가지나 몇 가지 기본적인 제약을 갖고 있다. 첫 번째 제약은 다음과 같다. 이산 무기억 소스의 임의의 심볼 s_i 에 대한 이상적 이산 부호어 길이는 $-\log_2 p(s_i)$ 이다. 부호어는 반드시 정수값으로 표현되기 때문에 이는 소스 심볼의 확률이 $1/2, 1/4, 1/8, \dots$ 등과 같이 2의 음의 멱급수(Negative Power of Two)를 가져야 함을 의미한다. 따라서 상기에 설명한 바와 같이 이산 무기억 소스의 임의의 심볼이 이상적 이산 부호어 길이와 같아지기 위해서는 심볼의 발생확률이 2의 음의 멱급수로 표현되어야 하나 실제 응용에서는 불가능하다.

또다른 제약으로는 소스의 발생 확률이 변화되었을 때 허프만 부호는 효율적으로 대처하기 어렵다는 점이다. 실제로 사전에 고정된 허프만 부호를 임의의 다른 소스에 적용한 경우 오히려 데이터가 증가하는 경우도 있다. 이와 같은 결점을 해결하기 위해서 허프만 부호는 종종 2 패드(Path) 기법을 사용하기도 한다. 즉 먼저 심볼의 통계적 성질을 구하여 부호어를 발생시킨 후, 이를 근거로 실제 부호화를 행한다. 이의 방법은 압축률은 상승하나 부호화를 2번 행해야 하는 결점이 있다.

이제 산술 부호에 대하여 설명하기로 한다. 상기 설명한 허프만 부호의 부호화 효율을 높이기 위한 방법의 한가지는 소스열을 몇 개의 큰 블록으로 나눈후, 각 블록을 가변장 부호어로 할당하는 것이다. 물론 복호기 측에서는 각각의 부호어에 대응되도록 가변장 블록으로 분리한 후, 원래의 부호로 변환한다. 즉 소스열 블록과 부호어 블록사이에는 일대일 대응관계가 존재한다.

이에비해 산술부호는 비블록 부호(또는 트리(Tree) 부호라고도 부름)로 길이 m 개의 심볼로 구성되는 전체 입력 데이터열에 단 하나의 부호어를 할당하는 기법이다. 산술부호에서 부호어 길이는 대략 이상적 이산 부호어 길이 $-\log_2 p(s_m)$ 에 근접하며, 여기서 $p(s_m)$ 은 소스열 (s_m)의 발생확률을 나타낸다.

산술부호의 기본 개념을 이해하기 위하여 이산 무기억 소스로 발생되어진 2진 심볼열 $[s_m]$ 의 부호화를 생각해 본다. 먼저 심볼열에서 '0'이 나올 확률을 p 라 하면 '1'이 나올 확률은 $q=1-p$ 가 된다. 여기서 왼쪽값은 포함하고 오른쪽값은 포함하지 않는 구간 $(0,1)$ 를 I 라 정의하기로 한다. 한편 길이 m 의 소스열 (s_m)에 있어서 발생할 모든 조합은 2^m 이며, 이의 발생 확률 $p(s_m)$ 의 총 합은 1이 된다. 따라서 서브(Sub)구간 $I_i, i=1, 2, \dots, 2^m$ 은 각 소스열 s_m 에 일의적으로 할당되어질 수 있으며, I_i 의 길이는 $p(s_m)$ 과 같으며 서브구간은 중첩되지 않는다.

실제 부호화 방법을 아래와 같은 에리어스(Elias) 부호 분할을 이용해 설명하기로 한다.

먼저 구간 I 는 2개의 서브구간 $[0,p)$ 와 $[p,1)$ 로 분할된다. 다음 입력되는 열의 심볼값이 '0'이면 서브구간 $[0,p)$ 를 선택하고 '1'이면 $[p,1)$ 을 선택한다. 각 서브구간은 위와같은 방식으로 다시 분할된다. 즉 서브구간 $[0,p)$ 는 $[0, p^2)$ 와 $[p^2,p)$ 로, 서브구간 $[p,1)$ 은 $[p,2p-p^2)$ 과 $[2p-p^2,1)$ 로 나뉘어지며 이의

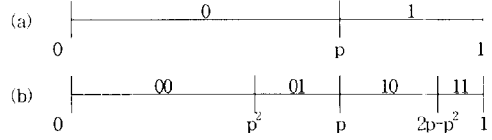


그림 3. 에리어스(Elias) 부호 분할

결과가 상기 그림에 도시되어 있다. 이와같은 구간 분할은 입력 데이터열의 수만큼 행해진다. 한편 입력 데이터열에 따른 구간 선택은 순차적으로 입력되는 심볼값에 따라 순차적으로 결정되어 지며 이를 수식적으로 표현하면 다음과 같다. 즉 $j-1$ 번째의 2진 소스 심볼의 서브구간을 $[L^{(j-1)}, R^{(j-1)})$ 라 하면 (여기서 L 은 왼쪽 끝점을, R 은 오른쪽 끝점을 나타냄), 새로운 서브구간 $[L^{(j)}, R^{(j)})$ 은 다음심볼이 '0'일때

$$L^{(j)} = L^{(j-1)} \tag{6}$$

$$R^{(j)} = L^{(j-1)} + p(R^{(j-1)} - L^{(j-1)}) \tag{7}$$

이 되며, 다음심볼이 '1'일때는

$$L^{(j)} = L^{(j-1)} + p(R^{(j-1)} - L^{(j-1)}) \tag{8}$$

$$R^{(j)} = R^{(j-1)} \tag{9}$$

이 된다.

임의의 입력열에 있어서 위와같은 방법으로 발생된 서브구간은 입력열의 발생확률과 동일한 크기의 폭을 갖는다. 즉 산술부호는 입력 데이터열의 엔트로피와 거의 유사한 값을 갖는것이 가능하다. 실제 응용에 있어서는 정확한 서브구간 분할이 어려움으로 스케일링(Scaling)과 라운딩(Rounding) 기법을 도입하여 이 문제를 해결하고 있다.

3.3 레이트-디스토션(Rate-Distortion)이론과 손실 압축

많은 실제 응용에 있어서, 원래의 화상으로 복원했을 때 발생하는 일정 수준의 화질의 열화는, 응용에 따라 일정 수준이 허용된다. 이와같은 화질 열화는 양자화시의 양자화 간격과 같은 파라메타들의 조정에 의해 발생되는데, 화질 열화의 수준에 대한 결정은 각각의 응용에 따라 서로 다르게 결정된다. 여기서 자주 발생하는 의문점중 하나는 화질 열화의 수준이 결정되었을 때 필요로 하는 최소한의 비트율은 얼마인가 하는 점이다. 이에 대한 해답이 레이트-디스토션(Rate-Distortion) 함수 $R(D)$ 이다. 그림 4는 유한 알파벳을 갖는 이산 소스에 있어서의 전형적인 레이트-디스토션 함수 $R(D)$ 를 보여주고 있다. 임의의 소스를 왜곡없이 압축하기 위한 최소 비트율은 그림에서 $D=0$ 일때의 R 값이며, 이는 소스 엔트로피와 같거나 더 작은 값을 갖는다. 그림에는 또한 저율 압축과 고투율 압축의 경우에 대한

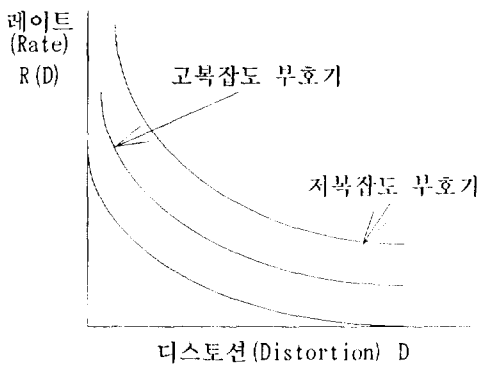


그림 4. 레이트-디스토션 함수의 일예

레이트-디스토션 함수의 일예도 나타내고 있다.

4. 무손실 압축 기법

의료 영상과 같이 영상에 포함되어 있는 정보가 중요한 몇몇 영상 압축 기법의 응용에서는 복원된 영상의 픽셀값이 원영상의 대응되는 픽셀값과 일치하는 무손실 압축을 필요로 한다. 무손실 압축의 경우 복원시 원화상과 동일한 영상을 얻을 수 있지만, 손실 압축에 비해 압축율이 낮다는 단점이 있다.

바이너리(Binary) 영상에 대하여 몇몇 무손실 압축 방식이 알려져 있으며, 연속적 톤(Tone)을 갖는 영상에 있어서는 런랭쓰(Runlength)부호화,비트플레인(Bit Plane) 부호화,계층적 변환 부호화 등 여러가지 방식들이 제안되고 있다. 최근에는 쿼-코더(Q-Coder)와 같은 적응적 산술 부호와 같이 간단하면서도 효율적인 방식도 제안되었다.

본 장에서는 무손실 압축 기법에서 가장 일반적인 방식인 비트 플레인(Bit Plane) 부호와 예측 부호 그리고 손실과 무손실 부호를 동시에 사용하여 무손실 압축을 행하는 방법에 대하여 설명하고자 한다.

4.1 비트 플레인(Bit Plane) 부호

각 픽셀값이 k 비트로 표현되는 $N \times N$ 크기의 영상을 고려해 보자. 2진수로 표현되어 있는 각 픽셀값에서 어느 한 비트만을 선택한다면,비트 플레인이라고 불리어지는 $N \times N$ 크기의 2진 영상을 얻는것이 가능하다. 일례로 각 픽셀값중 최상위 비트만을 취한다면, 최상위 비트 플레인을 나타내는 $N \times N$ 2진 영상을 얻을 수 있다. 다른 비트들에 대해서도 이와같은 방식을 반복한다면,원 영상은 k 개의 $N \times N$ 비트 플레인 으로 분해되어질 수 있다.

이와같이 원 영상을 각 비트 플레인으로 분해하여 무손실 2진 압축 기법을 이용하여 부호화 하면 동일 비트 위치의 값은 서로 상관성이 높기 때문에 큰 압축효율을 기대할 수 있다. 또한 최상위 비트 플레인은 원 영상의 중요한 정보를 많

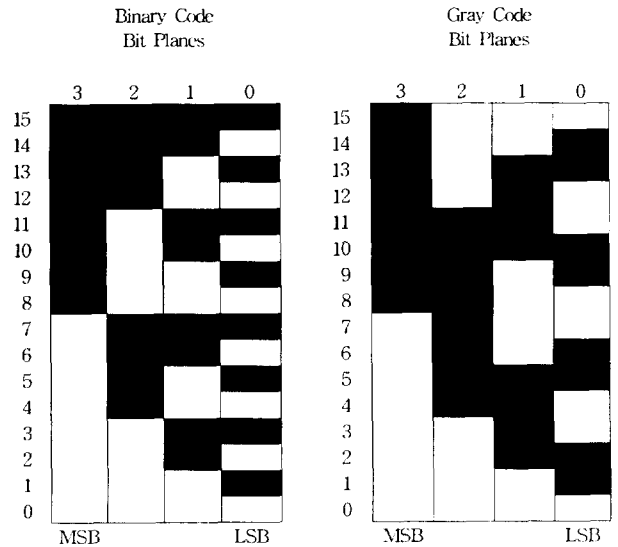


그림 5. 바이너리(Binary) 부호와 그레이(Gray) 부호

이 포함하고 있을 뿐 아니라 높은 상관 관계로 인하여 고효율로 압축이 가능하며,일반적으로 비트 자리수가 높을수록 더 많은 정보를 포함하고 있으며 압축 효율도 더 높다. 이와같은 성질 때문에 비트 플레인 부호는 영상 복원시 중요도가 높은 순서 즉 최상위 비트 플레인 부터 순차적으로 영상을 복원하므로, 영상의 화질이 순차적으로 상승하므로 영상의 고속 탐색과 같은 응용에 특히 유용하다. 그림 5는 바이너리(Binary) 부호와 그레이(Gray) 부호를 나타내었다.

4.2 예측(Predictive) 부호

일반적인 화상의 경우 인접 픽셀값 사이에는 큰 상관성이 존재한다. 따라서 임의의 한 픽셀값에 대한 많은 정보를 인접 픽셀값으로 부터 얻는것이 가능하다. 예측 부호 기법에서는 위와같은 성질을 이용하여 부호화를 행하는 데, 부호화의 기본 개념은 원하는 픽셀값의 예측을 주위의 픽셀값을 이용하여 행하는 것이다.

먼저 가장 간단한 예측기로서 디피시엠(DPCM : Differential Pulse Coded Modulation) 예측기를 살펴보자. 그림 6은 이미 입력된 픽셀값들의 구조를 보여주고 있으며, 이 픽셀들은 앞으로 입력될 픽셀값을 예측하는데 사용된다.

예측기에서 사용되는 픽셀의 수를 m 이라 할때 m 을 예측기의 차수라 하며,이는 예측기의 성능을 단적으로 보여준다. 일

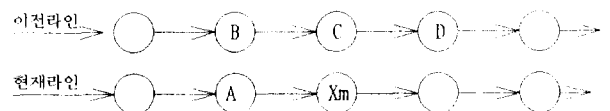


그림 6. 디피시엠(DPCM) 예측기 구조

반적으로 차수가 높을수록 성능이 좋으나, 4차 이상의 경우 차수의 증가에 비해 성능의 향상은 그다지 눈에 띄지 않는다고 하는 실험결과가 텔레비전 영상등을 대상으로한 실험에서 보고 되었다. 그림 6의 경우 예측의 일례로 다음과 같은 3차 예측기를 사용하여 예측하는 것이 가능하다.

$$\hat{x}_m = 0.75A - 0.50B + 0.75C \quad (10)$$

한편 위와같이 예측된 픽셀값은 원 화상의 픽셀값과 예측 오차가 발생하게 되는데, 이는 앞에서 설명한 허프만 부호와 같은 가변장 부호(엔트로피 부호라고도 함)을 사용하여 부호화 함으로써 최종적으로 무손실 압축이 가능하게 된다.

5. 손실 압축 기법

손실 압축 기법은 높은 효율의 압축비를 얻기 위하여 압축 시 화질 열화를 허용하는 기법으로 기본 구성은 다음과 같다.

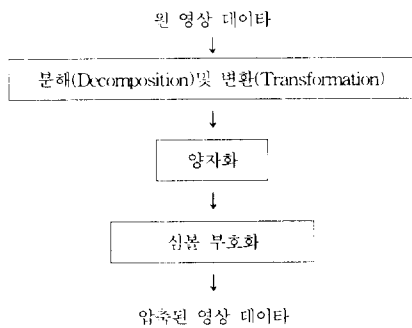


그림 7. 손실 압축의 일반적 구조

손실 압축시 먼저 영상을 분해 또는 변환을 행하는 이유는 다이내믹(Dynamic) 범위를 줄이고 영상에 포함되어 있는 리던던시 정보를 제거하는 함으로써, 보다 효율적으로 부호화 될 수 있도록 하기 위함이며, 이 단계는 무손실 기법으로 가역적이다. 손실 압축 부호화에서 왜곡이 생기는 단계는 양자화 단계로 앞서 설명한 레이트-디스토션 함수 $R(D)$ 에 의해 양자화가 행해지며, 따라서 화질 및 데이터량의 크기도 대부분 이 단계에서 결정된다. 이제 손실 압축 기법중 가장 많이 사용되는 몇가지 기법에 대하여 보다 자세하게 알아 보도록 한다.

5.1 변환(Transform) 부호

일반적인 변환 부호 기법은 먼저 $N \times N$ 크기의 원 영상을 보다 작은 $n \times n$ 크기의 서브 블록으로 분할한 후, 각각의 서브 영상에 대하여 유니터리(Unitary) 변환을 행한다. 변환의 목적은 원 신호 사이의 상관성을 제거하는데 있으며, 변환 결과는 변환 계수중 일부 계수에 에너지가 집중되는 현상을 보인다. 이와 같은 방법에 의해 에너지가 거의 없는 변환 계수는 양자화시 제거

되며, 결과적으로 압축을 얻는 것이 가능하다. 또한 사람의 특성을 이용하여 눈으로 보기에는 왜곡이 없는 압축을 얻는 것도 가능하다. 변환 부호의 기본 구성을 그림 8에 도시하였다.

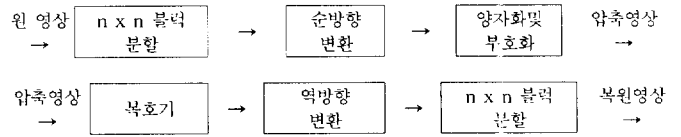


그림 8. 변환 부호의 블록도

영상 압축을 목적으로 사용되는 변환은 다음과 같은 몇 가지 특성을 가져야 한다.

1. 영상 데이터의 상관성을 제거할 것
2. 영상과 무관한 기저함수(Basis Function)를 가질 것
3. 구현이 용이할 것(처리속도가 빠를 것)

변환 부호에는 좌표축 회전에 의한 변환이나 기본 함수 분해에 의한 변환이 있으나 본 고에서는 주로 많이 사용하는 영상 변환에 대하여만 설명하도록 한다.

에너지 집중도의 관점에서 볼 때 최적의 변환 방식은 KLT(Karhunen Loeve Transform)이나 불행하게도 KLT의 기저함수(Basis Function)은 부호화 하고자 하는 영상에 의존하는 한편 계산이 복잡하여 특별한 응용외에는 거의 사용되어지지 않는다.

또다른 변환 방식으로 DFT(Discrete Fourier Transform)이 있으며 주파수 해석이나 필터링에 많이 사용되어 지고 있다. $n \times n$ 크기의 블록 f 에 대한 순방향과 역방향의 2-D DFT는 아래와 같이 정의된다.

$$F(u, v) = \frac{1}{n} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} f(j, k) \exp\left\{-\frac{2\pi i(uj + vk)}{n}\right\} \quad (11)$$

$$f(j, k) = \frac{1}{n} \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} F(u, v) \exp\left\{\frac{2\pi i(uj + vk)}{n}\right\} \quad (12)$$

여기서 $i = \sqrt{-1}$ 을 의미한다. 일반적으로 DFT에 의해 발생되는 변환계수는 복소수(Complex)이며 이는 많은 연산을 필요로 하나 빠른 연산법이 알려져 있다.

이에 비하여 DCT(Discrete Cosine Transform)은 성능면에서 앞서 설명한 KLT와 거의 비슷한 에너지 집중도를 가질뿐 아니라 계산량도 DFT에 비하여 적기 때문에 현재 가장 많이 사용되어지고 있다. 이의 순방향과 역방향에 대한 변환식은 다음과 같다.

$$F(u, v) = \frac{4C(u)C(v)}{n^2} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} f(j, k) \cos\left[\frac{(2j+1)u\pi}{2n}\right] \cos\left[\frac{(2k+1)v\pi}{2n}\right] \quad (13)$$

$$f(j, k) = \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} C(u)C(v)F(u, v)$$

$$\cos\left[\frac{(2j-1)u\pi}{2n}\right] \cos\left[\frac{(2k+1)v\pi}{2n}\right] \quad (14)$$

여기서 j,k = 블럭내의 픽셀 위치
u,v = DCT 계수 위치

$$C(w) = \frac{1}{\sqrt{2}} \text{ for } w=0, C(w)=1 \text{ for } w=1,2,\dots,n-1$$

DFT와 유사하게 DCT도 빠른 연산법이 존재한다. 한편 DCT가 DFT에 비하여 더 높은 압축 효율을 갖는 이유는 스푸리어스(Spurious) 주파수 성분이 덜 발생하기 때문이다. 이외에 WHT(Walsh-Hadamard Transform), 슬랜트(Slant) 변환등이 있다.

5.2 벡터 양자화(Vector Quantization) 부호

벡터 양자화 기법은 일련의 데이터로 구성되는 벡터를 양자화하여 부호화하는 기법이다. 벡터 양자화는 스칼라 양자화보다 압축비가 더 높으며, 복호기도 비교적 간단하다. 기본적인 벡터 양자화기의 블럭도를 그림 9에 나타내었다.

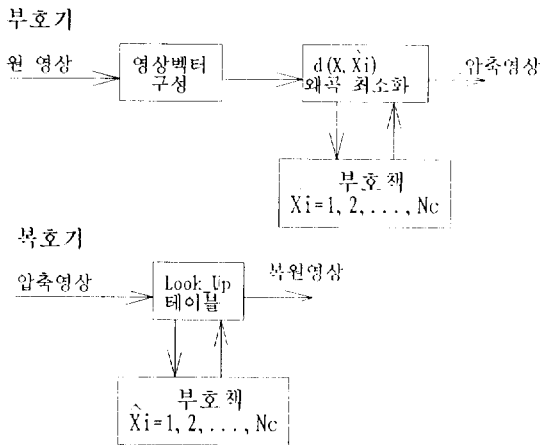


그림 9. 벡터 양자화기의 블럭도

원 영상은 먼저 n-차원의 영상 벡터로 분해되어지는데 영상 벡터를 구하는 방법은 응용에 따라 서로 다르게 구해진다. 다음에 이와같이 얻어진 영상 벡터 X와 부호책(Codebook)내에 있는 부호벡터 $\hat{X}_i, i=1, \dots, N_c$ 를 순차적으로 비교하여 왜곡이 최소가 되는 부호벡터를 찾는다. 최소 왜곡 부호벡터가 선택되면, 그것의 지표(Index) k가 $\log_2 N_c$ 비트를 사용하여 전송된다. 복호기에서는 부호기와 동일한 부호책으로부터 지표 k에 해당되는 부호벡터를 찾아서 원 영상으로 복원한다.

벡터 양자화에서의 효율은 부호책을 얼마나 잘 만드는가에 크게 의존하는데 Linde-Buzo-Gray(LBG) 구성방식이 가장 많이 알려져 있다. 벡터 양자화에서는 부호책의 크기선정이 중요한데, 효율을 높이기 위해 대표 벡터의 수를 늘리면 부호책이 커져 부호기에서 대표벡터를 선정하기위한 탐색시간이 길어져 실시간 수행이 어렵게 된다.

따라서 계산 복잡도 및 메모리에 요구 사항을 줄일수 있는 다양한 방식에 대한 연구가 많이 진행되고 있으며 다른 부호화 기법과 결합한 벡터 양자화가 많이 연구되고 있다.

5.3 서브밴드(Subband) 부호

서브밴드 부호에 있어 원 영상은 먼저 서로 다른 범위의 필터로 필터링되어지며 이에따라 서로 다른 영역의 주파수 스펙트럼을 포함하는 영상들이 얻어 지는데 이를 서브밴드라 부른다. 각각의 서브밴드는 원 영상에 비해 축소된 대역폭을 갖기 때문에 다운샘플링(Downsampling) 되어질수 있다. 상기와 같이 필터링 및 서브샘플링(Subsampling)과 같은 일련의 과정을 분해단(Analysis Stage)라 한다. 이와 같이 얻어진 서브밴드들은 다음에 하나 혹은 하나이상의 부호기를 사용하여 부호화된다. 서로다른 비트율 또는 서로 다른 부호화 기법조차도 각각의 서브밴드에 적용할 수 있으며 이는 여러가지 응용을 가능하게 해 준다. 여기서 강조하고 싶은 사항은 원 영상을 서브밴드로 나누는것 자체는 아무런 압축이 발생하지 않는다는 사실이다. 중요한것은 원 영상을 부호화 하는것 보다는 서브밴드를 부호화 하는것이 더 효율적이라는 사실이다. 아래에 1차, 2밴드의 서브밴드 부호화의 블럭도를 도시하였다.

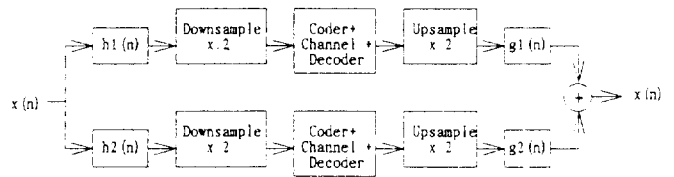


그림 10. 서브밴드 부호화의 블럭도(1차/2밴드)

5.4 웨이브렛(Wavelet) 부호

웨이브렛이란 h(t)로 정의되는 Mother 웨이브렛을 변위(Translation) 및 확대(Dilation)/축소(Contraction) 시킴으로써 얻어지는 기저함수들의 집합을 의미하며, 임의의 함수 x(t)의 연속 웨이브렛 변환은 아래와 같이 정의된다.

$$Xw(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) h\left(\frac{t-b}{a}\right) dt \quad (15)$$

여기서 a,b는 실수(a, b ∈ R, a ≠ 0)이다.

a = a₀^m, b = nb₀ 일때 이산 웨이브 변환이라 하며, 정의식은 아래와 같다.

$$Xw(m, n) = a_0^{\frac{m}{2}} \int_{-\infty}^{\infty} h(a_0^{-m}t - nb_0)x(t) dt \quad (16)$$

여기서 m,n은 정수이며 a₀>1, b₀≠0 이다.

웨이브렛 변환의 시간-주파수 해상도의 특징은, 고주파에서는 시간 해상도가 좋고 저주파에서는 주파수 해상도가 좋다.

따라서 영상의 성질이 Nonstationary 경우에 적합하다고 할 수 있다. 웨이브렛 기반 부호화는 상기와 같은 웨이브렛 이론을 근간으로 여러 밴드마다 사람의 시각 특성을 고려한 기존의 여러 부호화 알고리즘과 함께 연구되고 있으며, 이산 웨이브렛 변환 영역에서 벡터 양자화를 이용하는 방법이 주종을 이루고 있다.

6. 결 론

본 고에서는 이제까지 연구되어온 디지털 영상 부호화 기법중 특히 영상 압축 기법들에 관하여 주요 개념 위주로 설명하였다. 설명되어진 여러가지 압축 기법들은 국제 표준 기법인 JPEG이나 MPEG등의 표준화에서도 그대로 사용되어지고 있으므로 이에 관련하여 도움이 되리라 생각된다.

영상의 정보손실을 최소로 하는 영상압축 기법분야는 이미

국제 표준화 기구등을 통하여 각 응용분야에 따라 표준화가 진행되고 있거나 완료되어져 있다. 모쪼록 국내에서도 이와 같은 표준화 대열에 적극 동참하여 새로운 정보산업의 발전에 앞서감은 물론 지적 재산권등을 적극 확보함으로써 경쟁력을 강화시켜야 할 것이다.

참 고 문 헌

- [1] M. Ajid Rannani, Paul W. Jones "Digital Image Compression Techniques," SPIE OPTICAL ENGINEERING PRESS, 1991
- [2] A.K. Jain, "Fundamental of Digital Image Processing," Prentice Hall, 1989
- [3] ISO/IEC 13818-1, 2, 3, "Information technology coding of moving picture and associated audio," MPEG-2 Draft International Standard

저 자 소 개



김병진(金炳辰)

1962년 7월 27일생. 1985년 2월 고려대 공대 전기공학과 졸업. 1984년 12월-현재 LG전자(주) 영상미디어 연구소 책임연구원. 1993년 9월-현재 한국과학기술원 정보및 통신공학과 석사과정. 관심분야 : 광기록 재생 기술



박현욱(朴玄旭)

1959년 7월 5일생. 1981년 2월 서울대 공대 전기공학과 졸업. 1983년 2월 한국과학기술원 전기및 전자공학과 졸업(석사). 1988년 2월 한국과학기술원 전기및 전자공학과 졸업(공학박). 1989년 7월-92년 3월 미국 워싱턴대학교 연구원. 1992년 5월-93년 7월 삼성전자 정보컴퓨터연구소 수석연구원. 1993년 8월-현재 한국과학기술원(서울) 정보및 통신공학과 부교수. 주관심분야 : 영상처리, 영상통신, 영상처리시스템, 의료영상시스템.