

論文2003-40SP-6-14

주파수 특성을 이용한 미세 계위적 동영상 부호화 방법

(A Fine Granular Scalable Video Coding Algorithm using Frequency Weighting)

金昇煥*, 扈堯盛*

(Seung-Hwan Kim and Yo-Sung Ho)

요약

본 논문에서는 최근 비디오 스트리밍 응용을 위해 만들어진 FGS(fine granular scalability) 방식에서 주어진 비트율에서 인간 시각에 좀더 나은 화질을 얻기 위해 주파수 가중 방법을 제안한다. 제안된 방식에서는 각 DCT 계수의 사람 눈에 미치는 중요도를 고려하여 구해진 주파수 가중행렬을 비트평면 부호화를 사용하는 FGS 상향계층 부호화 동작에 적용될 수 있도록 주파수 천이 행렬로 바꾸어서 가중치를 부여한다. 본 논문에서 제안된 주파수 가중 방법을 통해 화질의 향상, 더 미세한 계위 부호화, 시간 영역에서 화질의 균일화를 얻을 수 있었다. 또한, 주파수 가중 방법을 통해 얻은 화질 개선을 설명하기 위해 기존에 사용하던 PSNR(peak signal to noise ratio) 대신에 인간 시각적인 관점에서의 오차만을 고려한 JNDE(just noticeable difference error)를 Weber의 법칙을 근거로 하여 새롭게 제안하였다.

Abstract

In this paper, we propose a progressive scalable video coding algorithm using frequency weighting in the DCT domain. Since the human visual system (HVS) can be modeled as a nonlinear point transformation, called the modulation transfer function (MTF), we can use the frequency weighting matrix to enhance the video image quality. We change this frequency weighting matrix into the frequency shift matrix to apply to the bit-plane coding method for the fine granular scalable (FGS) video coding. We also define a new error metric JNDE (just noticeable difference) to measure the perceptual image quality in terms of human vision.

Keyword : FGS, HVS, JND, Frequency Weighting

I. 서론

일반적으로 비디오 부호화의 목적은 주어진 비트율에서 비디오 화질의 최저화하려는 것이다. 기존의 도시 부호화 시스템에서는 전송 채널의 용량을 미리 알고

있다고 가정하고 채널용량과 같거나 더 적은 비트율 안에서 입력 비디오 신호를 부호화하였다. 또한 복호기에서는 수신된 모든 비트들을 복호할 수 있다는 가정하였다^[1]. 한편, 1990년 중반 이후에 서서히 알려진 인터넷은 급속한 속도로 확산되어 현재는 거의 모든 일반인들에게 알려져 있다. 따라서 인터넷 초기에 전송 대상이었던 텍스트와 정지영상이 현재는 음악이나 동영상과 같은 실시간 멀티미디어 정보가 급격히 늘어나고 있다. 특히, 인터넷 비디오 스트리밍과 같이 네트워크 채널용량이 어느 주어진 시간에 대해서 넓은 범위로

* 正會員, 光州科學技術院 情報通信工學科
(Dept. of Information and Communications, Kwangju Institute of Science and Technology)
接受日:2003年2月2日, 수정완료일:2003年11月22日

에서 변하는 네트워크 비디오 응용에서 부호화기는 시간에 따라 변하는 채널용량을 정확히 예측할 수 없으며, 복호기 또한 다른 운영체제와 자원을 공유하는 경우에 수신된 비트를 모두 복호할 수 없기 때문에 기존의 통신 시스템에서의 요구조건이 다소 변화되었다^[2].

최근에 MPEG-4 표준에서는 스트리밍 비디오를 위해 새로운 프로파일을 제공하였다. 이는 인터넷을 통해 점점 증가하는 스트리밍 비디오에 대한 하나의 부호화 방법을 제시한 것이다. 인터넷 스트리밍을 위해 제안된 FGS(fine granular scalability) 방식은 부호화를 수행하기 전에 미리 전송채널의 용량을 예측하여 최소 채널용량보다 작은 비트율로 기본계층을 부호화하므로 기본계층 데이터의 전송은 항상 보장된다. 또한, 원 영상과 기본계층 영상의 차이 값인 잔여 영상은 네트워크 채널용량이 허용하는 범위에서 향상계층에서 부호화되어 전송된다.

향상계층에서의 FGS 부호화는 먼저 원영상과 기본계층 영상의 차이 값인 잔여 영상을 8×8 블록별로 나누어, 각 블록별로 DCT 변환을 통하여 얻어진 64개의 DCT 계수들을 비트평면 부호화를 이용하여 부호화한다. 따라서 전송채널과 시스템의 복잡성을 고려하여 일부 비트평면까지만 복호할 수 있다. 게다가, 비트평면이 DCT 계수별로 부호화되므로 인간의 시각에 중요한 주파수 성분을 강조하기 위해 각 계수별로 해당하는 가중 값을 고려하여 비트평면을 천이 시킬 수 있다.

FGS 방식에서는 시간에 따라 변하는 채널용량에 따라 유효 비트율을 적응 적으로 조절할 수 있도록 향상계층에서 비트평면 부호화를 사용하고 있지만, 양자화 테이블을 통해서 각 DCT 계수별로 중요도를 고려하는 기본계층의 부호화에 비하여 훨씬 비효율적이다^[1].

만약, 기본계층에서 충분한 양의 부호화가 이루어진다면 향상계층에서는 거의 잡음성분과 같은 영상이 부호화되기 때문에 모든 계수들에 대해서 중요성을 고려할 필요성이 없겠지만, FGS 부호화는 최소 채널용량 이하로 부호화를 수행하므로 대부분의 비디오 신호는 향상계층에서 부호화되고 여전히 향상계층에는 기존의 영상이 가지고 있는 저주파 성분에 에너지가 집중되는 성질이 남게 된다. 따라서 향상계층에서 DCT 변환을 통하여 얻어진 64개의 계수들을 같은 중요도를 가지고 부호화하는 것은 인간 시각의 관점에서 비효율적이다.

본 논문에서는 향상계층에서 얻어진 64개의 DCT 계수들을 HVS(human visual system) 관점에서 각 DCT

계수의 중요성을 고려하기 위해서 각 DCT 계수별로 인간 시각의 중요도를 나타내는 주파수 가중 행렬을 이용한다. 또한, 주파수 가중 행렬을 통한 DCT 계수별 중요도를 비트평면 부호화를 사용하는 FGS 향상계층 부호화에 적용하기 위하여 주파수 가중 행렬을 수학적인 모델링을 통해 각 DCT 계수가 비트평면 부호화상에서 그 중요도에 따라서 얼마나 상향 천이 되어야 하는지를 나타내는 주파수 천이 행렬로 바꾼다. 여기에서 얻어진 주파수 천이 행렬을 이용하여 각 DCT 계수별 중요도에 따른 비트평면 천이를 통해 인간 시각에 맞는 화질의 향상을 얻는다. 또한 기존의 화질측정 단위인 PSNR (peak signal-to-noise ratio) 대신에, 인간의 시각적 특성에 맞도록 Weber의 법칙을 근거로 하여 인간 시각에 감지되는 오차만을 고려한 새로운 화질측정 단위인 JNDE(just noticeable difference error)를 정의하고 이를 이용하여 주파수 가중을 이용하여 얻어진 화질 향상을 HVS 관점에서 분석한다.

II. FGS 기본 구조

1. FGS 상위계층 부호화

<그림 1>에 보인 것처럼, FGS의 기본계층은 기존의 블록기반 부호화 방법과 동일하게 부호화된다.

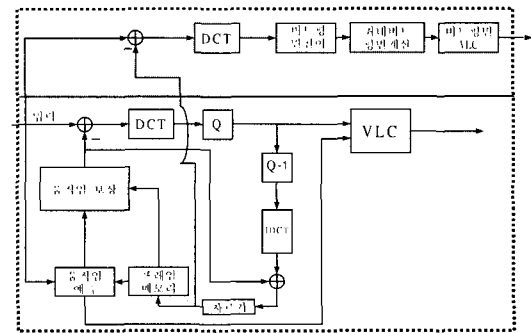


그림 1. FGS 부호기
Fig. 1. Encoder of FGS.

<그림 1>의 사각형 점선으로 표시되어 있는 상위계층에서는 기본계층에서 부호화되지 않은 잔여 영상을 부호화한다. 잔여영상 정보는 8×8 블록별로 나누어져서 DCT 변환되므로, 블록당 64개의 DCT 계수를 가진다. 이러한 한 블록 안에 있는 64개의 계수들은 ZigZag 스캔되고 절대값과 부호를 분리하여 나타내어진다. 다음

은 64개의 나열된 절대값 계수들이다.

10, 0, 6, 0, 0, 3, 0, 2, 2, 0, 0, 2, 0, 0, 1, 0, ..., 0, 0 (절대값)

위에 주어진 십진수의 값에서 최대값 10을 이진수로 나타내면 1010(2)이므로 4비트가 필요하다. 따라서 DCT 계수들이 이진수로 비트평면 부호화되기 위하여 필요한 비트수는 4이다. 다음은 위의 십진수들이 4비트의 이진수로 바뀌어 각 비트 평면을 구성하는 것을 나타낸다^[1].

1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ..., 0, 0 (MSB)
 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ..., 0, 0 (MSB-1)
 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, ..., 0, 0 (MSB-2)
 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, ..., 0, 0 (MSB-3).

위에 제시된 비트열에서 세로방향으로 각 십진수의 DCT 계수 값을 이진수로 나열하였고, 가로방향으로 64개의 DCT 계수 별로 동일한 자리 수를 가진 비트열을 나열하였다. 그리고 이 동일한 자리 수를 가진 이진수들은 같은 중요도를 가지고 있는 하나의 비트평면을 이루고, 가장 중요한 성분인 MSB 평면부터 가변길이 부호화되어 전송된다.

다음은 각 비트평면 별로 (RUN, EOP)로 나타내어진 가변길이 부호화된 경우를 나타낸다^[5-7].

(0,1) (MSB)
 (2,1) (MSB-1)
 (0,0), (1,0), (2,0), (1,0), (0,0), (2,1) (MSB-2)
 (5,0), (8,1) (MSB-3)

여기서 RUN은 연속된 0의 개수를 나타낸다. 또한 EOP의 경우는 더 이상 1이 나오지 않으면 1 그렇지 않으면 0으로 부호화된다^[1].

2. FGS 주파수 가중
 각 DCT 계수들은 화질에 서로 다른 영향을 미친다. 특히, 저주파 성분이 고주파 성분보다 인간의 시각에 더 민감하므로, 저주파 계수의 정밀도는 고주파 계수의 정밀도보다 더 중요하다. 이러한 관점에서, 저주파 계수들을 위로 천이 시켜 상대적으로 고주파 계수들보다 우선적으로 전송하면, 일부 비트스트림을 처리하는 경우에 화질의 향상을 얻을 수 있다.

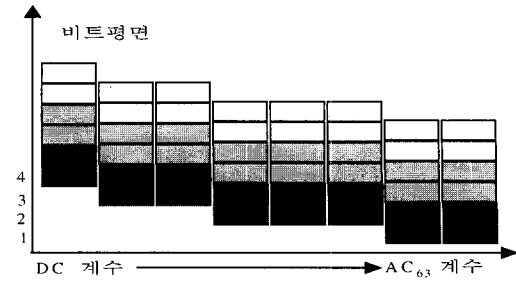


그림 2. 주파수 천이
 Fig. 2. Frequency shifting.

<그림 2>는 저주파 계수부터 차례로 천이 계수를 내림순(4, 3, 2, 1, 0)으로 적용한 예를 나타내고 있다. 비트평면 천이 부호화를 수행하는 경우에, 복호기에서는 부호기에서 천이 한 만큼의 비트 수만큼 다시 하향 천이를 수행하여 원래의 DCT 계수 값에는 변화가 없도록 해야한다.

비트평면 천이는 채널의 용량이 제한되어있거나 자주 변하는 환경에서 중요한 성분을 먼저 전송하여 저 비트율에서 상대적으로 나은 화질을 얻고자 하는 관점에서 매우 중요하다. 이러한 문제는 DCT변환을 이용하여 압축을 수행하는 현재의 부호화 과정을 고려할 때 주파수 영역에서 각 계수별로 중요도를 고려하여 비트평면 천이가 수행되어야한다. 하지만, 지금까지 각 DCT 계수 성분별로 중요도를 고려한 비트평면 천이에 대한 기준이 제시되지 않았다. 따라서 본 논문에서는 인간시각에 맞는 화질향상을 얻기 위해 HVS 관점에서 비트평면 부호화에 적용되도록 각 DCT 계수 성분들을 인간 시각의 관점에서 중요도를 고려하여 비트평면 천이 방법을 제시한다.

III. HVS 기반 주파수 가중

1. 주파수 가중 행렬

HVS는 비선형 주파수 함수인 MTF(modulation transfer function)를 식 (1)과 같이 정의하고 있다^[3].

$$H(f) = a(b + cf) \exp(-cf) \quad (1)$$

여기서 f 는 cycles/degree를 나타내고, a, b, c, d 는 상수 값을 나타낸다. 이러한 상수 값을 조절하여 각 주파수의 가중치를 보여주는 MTF의 모양을 결정한다. 특히, 8×8 블록의 DCT 계수들의 주파수 가중 행렬은

실험을 통하여 제시되었다^[1]. <그림 3>에서 각 위치에 해당하는 가중 값은 실제 DCT 계수들의 곱셈에 대한 가중치를 의미한다.

0.494	1.000	0.702	0.381	0.185	0.084	0.037	0.016
1.000	0.454	0.308	0.170	0.084	0.039	0.017	0.007
0.702	0.308	0.213	0.124	0.064	0.031	0.014	0.006
0.381	0.170	0.124	0.077	0.042	0.021	0.010	0.004
0.185	0.084	0.064	0.042	0.024	0.013	0.006	0.003
0.084	0.039	0.031	0.021	0.013	0.007	0.004	0.002
0.037	0.017	0.014	0.010	0.006	0.004	0.002	0.001
0.016	0.007	0.006	0.004	0.003	0.002	0.001	0.000

그림 3. 주파수 가중 행렬
Fig. 3. Frequency weighting matrix.

2. 주파수 가중을 이용한 비트평면 천이

비트평면 천이는 천이 값에 대해 2의 지수 배로 그 계수 값을 증가시키는 의미가 있다^[2]. 즉, 어떤 계수를 n 비트 더 보내기 위해 요구되는 가중 값은 2ⁿ이 된다. 따라서 가중 행렬이 비트평면 부호화와 결합하기 위해 먼저 각 가중 값을 2에 대한 지수 형태로 나타내야 한다.

$$C'(i, j, k) = 2^{n_{fu}(i)} \cdot C(i, j, k) \quad (2)$$

$$2^{n_{fu}(i)} = C'(i, j, k) / C(i, j, k) \quad (3)$$

$C(i, j, k)$ 는 k 번째 매크로블록의 j 번째 블록내 i 번째 DCT 계수, $C'(i, j, k)$ 는 주파수 가중을 고려한 DCT 계수, $n_{fu}(i)$ 는 주파수 i에 대한 비트평면 천이 계수를 나타낸다. <그림 3>에서 각 DCT 계수의 가중 값은 곱셈에 대한 가중 값이므로 주파수 가중되지 않은 DCT 계수 값인 $C(i, j, k)$ 에 주파수 가중 행렬에서 나타내는 해당 계수의 가중 값을 곱한 경우를 각 DCT 계수별로 주파수 가중을 적용한 경우라 볼 수 있다.

비트평면 부호화에서는 정수단위로 비트평면 천이를 수행하기 때문에 실제로 주파수 가중 행렬에서 주어진 가중 값들이 가장 적은 오차를 갖는 2에 대한 지수의 형태로 나타내져야 한다. 이를 위해 다음과 같은 일련의 과정을 수행한다. 먼저 <그림 3>에서 보여주는 가중 값이 실제 2에 대한 지수의 값들로 변환되기 위해 <그림 4>에서 나타내는 바와 같이 가장 중요한 주파수 성분을 얼마만큼 천이 할 것인가를 나타내주는 최대 주파수 천이계수 $\max(n_{fu}(i))$ 를 결정 해야한다. 최대 천이 주파수 계수를 결정하는 문제는 기본계층의 부호화 율

을 등을 비롯한 여러 가지 많은 고려사항이 요구되어진다. 일단 최대 주파수 천이계수가 결정이 되면 <그림 3>에 있는 가중 값들을 2에 대한 지수 형태로 변환하기 위해 <그림 3>에 있는 가중 값에 $2^{\max(n_{fu}(i)})$ 를 곱하여 크기를 조정한다. <그림 4>의 경우는 최고 주파수 천이가 4인 경우이므로 <그림 3>의 가중 값들이 2⁴ 배만큼 증가되었다.

7.907	16.00	11.23	6.102	2.969	1.358	0.598	0.256
16.00	7.278	4.936	2.729	1.352	0.627	0.278	0.120
11.23	4.936	3.422	1.990	1.032	0.497	0.227	0.100
6.102	2.729	1.990	1.233	0.680	0.344	0.208	0.075
2.969	1.352	1.302	0.680	0.393	0.212	0.107	0.512
1.358	0.627	0.497	0.344	0.212	0.120	0.064	0.032
0.598	0.278	0.227	0.164	0.107	0.064	0.035	0.017
0.256	0.120	0.100	0.075	0.051	0.032	0.017	0.009

그림 4. 크기가 조정된 가중 행렬
Fig. 4. Scaled weighting matrix.

<그림 4>에 주어진 크기가 조정된 주파수 가중 행렬로부터 정수단위의 비트평면 천이 계수를 얻기 위해, 각 크기가 조정된 가중 값이 올림 되어 정수로 바뀌고, 천이 계수를 구하기 위해 밑이 2인 로그를 취한다. 그리고 로그값이 다시 내림 되어 정수 단위의 천이 계수를 얻는다. 이러한 일련의 과정들을 식 (4)에서 잘 보여주고 있다.

$$n_{fu}(i) = \lfloor \log_2 \lceil C'(i, j, k) / C(i, j, k) \rceil \rfloor \quad (4)$$

식 (4)에서 나타내는 $C'(i, j, k) / C(i, j, k)$ 의 값은 <그림 3>에서 보여주는 가중 값과 동일하다. 각 DCT 계수는 <그림 5>에 보인 주파수 가중 값에 따라 천이된 뒤에 부호화된다.

3	4	3	2	1	0	0	0
4	3	2	1	1	0	0	0
3	2	2	1	1	0	0	0
2	2	1	1	0	0	0	0
1	1	1	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

그림 5. 주파수 천이 행렬
Fig. 5. Frequency shifting matrix.

부호기에서는 천이한 만큼 다시 하향 천이를 통하여 값 자체의 변화가 없도록 한다. 이러한 의미에서 주파

수 가중 행렬을 적용하는 것은 DCT 계수 값을 변화시키는 것이 아니라 얼마나 빨리 전송하는가의 문제임을 알 수 있다. 또한 최대 천이계수가 4인 경우에 주파수 가중 행렬에 의해 계산된 DC의 천이계수는 3이지만, DC값은 DCT 변환이 수행된 각 블록의 평균 명암 값을 의미하고, 일반적으로 가장 큰 신호 에너지를 가지고 있기 때문에 최고 천이계수 4를 사용했다.

IV. 화질평가 척도

본 논문에서는 적절한 화질 평가를 위해서 원영상과 복원영상 사이의 오차 값에만 의존하는 PSNR(peak signal-to-noise ratio) 대신에 인간 시각으로 구별되는 차이점을 고려하기 위하여 식 (5)와 같이 표현된 Weber의 법칙에 근거하는 새로운 화질평가 척도인 JNDE(just noticeable difference error)를 정의한다.

$$\frac{\Delta I}{I} = \alpha \tag{5}$$

여기서 I 는 초기에 주어진 자극의 값, ΔI 는 초기 자극에 대한 차이값. α 는 인간이 감지할 수 있는 초기 값과 변화된 값의 최소 비율로 보통 0.02 정도이다.

본 논문에서는 Weber의 법칙을 원영상과 복원된 영상과의 관계에 적용한다. 즉, 원영상을 초기에 주어진 자극(I)으로 간주하고, 원영상과 복원된 영상과의 오차 (ΔI)를 변화된 자극의 값으로 간주하면 주어진 오차 값에 대한 인간 시각의 민감도는 원영상의 화소 값에 의존한다. 원영상이 어두운 색을 가진 경우는 아주 밝은 흰색인 경우보다 똑같은 오차 값에 더 민감하다.

식 (6)은 원영상의 화소 값 p_k 에 오차 D 가 주어졌을 때, 인간 시각에 오차로 간주되는 범위를 나타낸다.

$$\frac{\Delta I}{I} = \frac{D}{p_k} \geq \alpha \tag{6}$$

식 (7)은 식 (6)에서 주어진 오차값 D 에 대해 오차로 간주되는 원영상의 화소 값 p_k 의 최대 경계값(p_{th-D})이 D/α 임을 나타낸다.

$$p_k \leq \frac{D}{\alpha} = p_{th-D} \tag{7}$$

원영상의 화소 값 p_k 를 [0,255] 사이에서 균일하게 분포된 이산확률 변수로 가정하면, 원영상의 화소 값이

p_{th-D} 보다 작을 확률은 식 (8)과 같다. 여기에서 균일한 이산확률 변수로 가정하는 경우는 원영상의 화소 값의 분포를 가장 일반적으로 고려한 경우이며, 만약에 우리가 원영상의 화소 값의 분포를 나타내는 히스토그램을 알고있다면 이용하여 각각의 영상에 알맞게 확률 값을 구할 수가 있다.

$$P(p_k \leq p_{th-D}) = \sum_{k=0}^{p_{th-D}} \frac{1}{2^n} = \frac{(D/\alpha)+1}{2^n} \tag{8}$$

이때, 복원된 영상과 원영상과의 오차는 원영상의 화소 값에 관계없이 발생하므로 오차 D 와 원영상의 화소 값이 서로 독립적이다. 따라서 오차 D 가 주어졌을 때, 인간 시각에 오차로서 간주될 확률 P.N.E.D는 식 (9)와 같이 나타낼 수 있다.

$$P_{N.E.D} = P(p_k \leq p_{th-D}) = P(p_k \leq p_{th-D}) = \frac{(D/\alpha)+1}{2^n} \tag{9}$$

따라서 JNDE 관점에서 오차 값이 D 인 오차의 개수는 실제 발생한 오차의 개수에 그 오차값 D 가 오차로 간주될 수 있는 확률 값을 곱하여 구한다. 이를 통하여 식 (10)에서와 같이 인간시각 관점에서의 오차만을 고려할 수가 있다.

$$\text{오차 } D \text{의 개수 } JNDE = P_{N.E.D} \cdot (\text{오차 } D \text{의 개수}) \tag{10}$$

표 1. 각 오차에 대한 화소 개수
Table 1. Pixel number in each error.

W	N	W-N	D	JND(W)	JND(N)	JND(W-N)
13,341	12,150	1,191	0	13,341	12,150	1,190
23,197	21,010	1,377	1	4,626	4,347	274
17,211	17,222	-11	2	6,790	6,795	-5
12,386	12,890	-504	3	7,306	7,603	-298
8,898	9,443	-545	4	6,986	7,414	-428
6,457	6,885	-428	5	6,330	6,751	-421
4,747	5,055	-300	6	4,747	5,055	-308
3,529	3,743	-214	7	3,529	3,743	-214

<표 1>에서는 주파수 가중을 사용하여 부호화를 수행한 경우와 주파수 가중을 사용하지 않고 부호화를 수행한 두 복원된 영상에 대해 인간시각 관점에서의 화질을 비교하기 위한 실험이다. 위에 제시된 두 복원된 영상은 같은 PSNR 값을 지니고 있으나 원영상과 두 복원영상간에 나타나는 화소 값에서의 오차 D 의 분포는 현저히 다르다. 여기서 D 는 복원된 영상과 원영

상 사이의 동일 위치에서 화소 값 차이를 나타낸다.

<표 1>에서 W는 주어진 오차 D에 대하여 주파수 가중을 적용한 경우의 오차 개수를 나타내고, N은 주파수 가중되지 않은 경우의 오차 개수를 나타낸다, 그리고 $\alpha=0.02$ 를 적용하였다. 또한, 각 오차 D에서 주파수 가중한 경우와 주파수 가중되지 않은 경우의 오차 개수를 뺀 값을 W-N로 나타내었다.

JNDE를 통하여 인간 시각에 오차로 간주되는 화소의 개수를 다시 계산한 결과를 각각 JND(W), JND(N), JND(W-N)으로 표시하였다. <표 1>에서와 같이, JNDE를 적용하지 않았을 때 오차 1인 경우에 주파수 가중의 경우가 주파수 가중되지 않은 경우보다 1377개나 더 많은 오차 개수가 JNDE를 적용해서 계산했을 때 274개로 그 차이가 현저히 줄었다. 즉, 주파수 가중을 적용한 경우에 인간 시각에 덜 민감한 오차 D가 작은 부분에 많은 오차가 집중되어 상대적으로 주파수 가중하지 않은 경우보다 향상된 화질을 얻을 수 있다.

V. 실험 결과 및 분석

본 논문에서 제안한 미세 계위적 동영상 부호화 방법의 성능을 검사하기 위해, CIF(352 화소 × 288 줄 × 297 화면) 형식의 Foreman 비디오 시퀀스를 사용하였다. 실험을 통하여 주파수 가중을 통하여 얻어진 세 가지 사항에 대하여 비교 설명하고자한다. 주파수 가중을 통하여 기존의 방법에 비하여 얻어진 이점은 더 미세한 향상계층제공, 주어진 비트율에 대한 화질향상, 시 영역에서 화질의 균일화를 얻을 수 있었다.

부호화된 비트평면	FW0	FW1	FW2	FW4
기본계층 비트율	373	373	373	373
기본계층+향상계층1개	523	522	513	460
기본계층+향상계층2개	1499	1164	1050	754
기본계층+향상계층3개	3645	2880	1914	1321
기본계층+향상계층4개	7061	5696	3850	2312

그림 6. 향상계층의 증가에 요구되는 비트율
Fig. 6. Required bit-rate for increasing the number of enhancement layers.

<그림 6>은 주파수 가중을 사용하지 않은 경우(FW0)와 주파수 가중을 최대 천이계수(FW(1, 2, 4))를 변화시키면서 향상계층을 증가시켰을 때 부호화된 비트율을 나타낸다. <그림 6>에서보면 주파수 가중을

사용하지 않은 경우에 요구되어지는 비트율이 370에서 2400 kbits/sec 사이일 때 2개의 향상계층까지 조절되지만, 최대주파수 4를 사용한 FW4 경우에는 4개의 향상계층까지 더 미세하게 조절 될 수 있다.

FGS 경우 기본계층의 부호화 율에 따라서 향상계층에서 주파수 가중한 경우의 효율이 다르게 나타난다. 다음은 기본계층의 비트율을 증가시키면서 시간 영역에서의 화질 및 공간영역에서의 화질을 비교하였다. 참고로 주파수 가중을 한 경우에는 가중을 하지 않은 경우보다 PSNR 면에서는 더 낮은 화질을 얻었지만 인간 시각에는 훨씬 나은 화질을 얻을 수 있었다.

비트율(kbits/s)	평균 PSNR	PSNR 분산
비주파수가중(1047)	33.18	2.43
주파수가중(904)	32.10	0.42

그림 7. PSNR 변화
Fig. 7. PSNR 분산

<그림 7>에서는 기본계층을 320 kbits/sec로 부호화하였다. 주파수 가중을 하지 않은 경우는 기본계층+향상계층을 1047 kbits /sec로 부호화하였고 주파수 가중을 사용한 경우는 904 kbits/sec로 부호화하였다. <그림 7>에서와 같이 주파수 가중을 사용한 경우는 297화면 동안 PSNR의 분산이 0.42로 나타났고 주파수 가중을 사용하지 않은 경우는 2.43으로 나타났다.

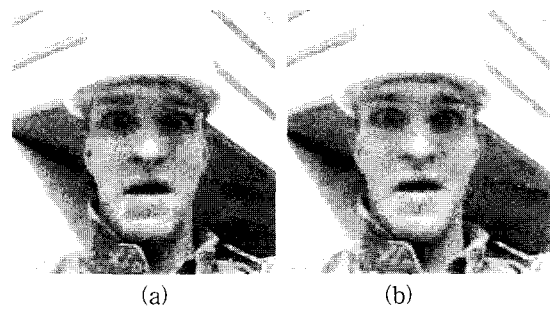


그림 8. (a) 주파수 가중 하지 않은 경우, (b) 주파수 가중한 경우
Fig. 8. (a) No frequency weighting, (b) frequency weighting.

<그림 8>에서는 Foreman Sequence의 93번째 화면으로서 <그림 7>에 주어진 비트율로 부호화되었다. <그림 8>과 같이 주파수 가중을 한 경우(31.24 dB)가 주파수 가중을 하지 않은 경우(31.87 dB)보다 PSNR면

에서는 작지만 훨씬 더 나은 화질을 보이고 있다.

<그림 9>에서는 기본계층을 374 kbits/sec로 부호화 하였다. 주파수 가중을 하지 않은 경우는 기본계층+향상계층을 1499 kbits/sec로 부호화하였고 주파수 가중을 사용한 경우는 1321 kbits/sec로 부호화하였다. <그림 9>에서 보여주는 바와 같이 주파수 가중을 사용한 경우는 297화면 동안 PSNR의 분산이 1.38로 나타났고 주파수 가중을 사용하지 않은 경우는 4.40으로 나타났다.

비트율(kbits/s)	평균 PSNR	PSNR 분산
비주파수가중(1499)	36.30	4.40
주파수가중(1321)	43.45	1.38

그림 9. PSNR 변화
Fig. 9. PSNR 분산

<그림 10>에서는 <그림 9>에 명시된 비트율로 부호화된 화면이다. 주파수 가중을 한 경우(32.68 dB)가 주파수 가중을 하지 않은 경우(32.14 dB)보다 PSNR면에서는 비슷하지만 훨씬 더 나은 화질을 보이고 있다.

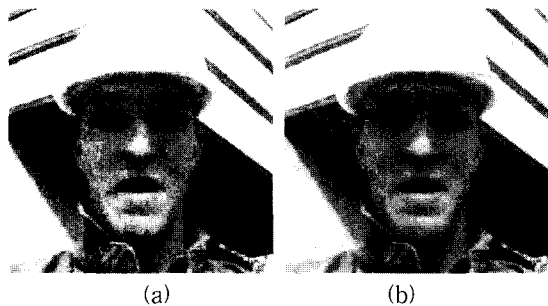


그림 10. (a) 주파수 가중 하지 않은 경우, (b) 주파수 가중한 경우
Fig. 10. (a) No frequency weighting, (b) frequency weighting.

<그림 11>에서는 기본계층을 415 kbits/sec로 부호화 하였다. 주파수 가중을 하지 않은 경우는 기본계층+향상계층을 1496 kbits/sec로 부호화하였고 주파수 가중을 사용한 경우는 1349 kbits/sec로 부호화하였다. <그림 11>에서 보이는 바와 같이 주파수 가중을 사용한 경우는 297화면 동안 PSNR의 분산이 1.45로 나타났고 주파수 가중을 사용하지 않은 경우는 3.12로 나타났다.

<그림 12>에서는 <그림 11>에 주어진 비트율로 부호화되었다. 주파수 가중을 사용한 경우(31.32 dB)에

주파수 가중을 하지 않은 경우(31.92 dB)보다 PSNR면에서는 비슷하지만 더 나은 화질을 보이고 있다.

비트율(kbits/s)	평균 PSNR	PSNR 분산
비주파수가중(1496)	34.18	3.12
주파수가중(1349)	32.93	1.45

그림 11. PSNR 변화
Fig. 11. PSNR 분산

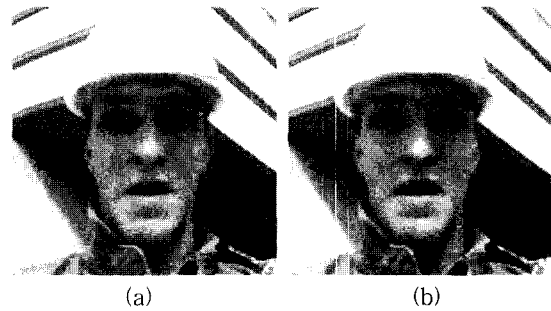


그림 12. (a) 주파수 가중하지 않은 경우, (b) 주파수 가중한 경우
Fig. 12. (a) No frequency weighting, (b) frequency weighting.

<그림 13>에서는 이전에 사용한 기본계층 부호화율보다 상대적으로 많은 781.6 kbits/sec로 기본계층을 부호화하였다. 주파수 가중을 하지 않은 경우 기본계층+향상계층을 2824 kbits/sec로 부호화하였고 주파수 가중을 사용한 경우는 2272 kbits/sec로 부호화하였다.

<그림 13>에서 보여주는 바와 같이 주파수 가중을 사용한 경우는 297화면 동안 PSNR의 Variance가 2.57로 나타났고 주파수 가중을 사용하지 않은 경우는 5.95로 나타났다. 따라서 주파수 가중의 경우가 그렇지 않은 경우보다 균일한 화질을 제공함을 알 수 있다.

비트율(kbits/s)	평균 PSNR	PSNR 분산
비주파수가중(2824)	39.64	5.95
주파수가중(2272)	37.00	2.57

그림 13. PSNR 변화
Fig. 13. PSNR 분산

VI. 결 론

본 논문에서는 인간 시각의 특성을 고려하여 각 DCT 계수의 중요성에 따라 비트평면 천이를 통한 주

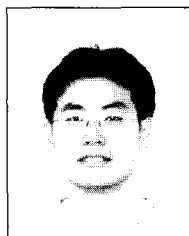
과수 가중을 사용하여 인간 시각에 향상된 화질을 얻고자 하였다. 또한 기존의 화질측정 단위인 PSNR 대신에, 인간의 시각적 특성에 맞는 새로운 화질측정 단위 JNDE를 정의하고, 이를 이용하여 주파수 가중을 이용한 화질 향상을 HVS 관점에서 분석하였다. 주파수 가중을 적용했을 때 적은 비트로 부호화했지만 오히려 더 향상된 화질을 얻을 수가 있었고, 시간 영역에서의 균일한 화질, 더 미세한 향상계층을 제공하였다. 또한, HVS 관점에서의 주파수 가중은 64개의 계수들에 대한 중요도를 모두 고려해서 얻은 결과이므로 사실상 HVS 관점에서는 최적화된 부호화 방법으로 볼 수 있다. JNDE를 적용하여 이미지의 화질을 비교했을 때 시각에 감지되지 않은 부분을 제거했기 때문에 PSNR보다 더 효과적으로 인간 시각에 맞게 화질을 비교할 수 있었다. 따라서 JNDE는 HVS 관점에서의 화질 평가의 기준으로 사용될 수 있다.

참 고 문 헌

[1] W. Li, "Overview of Fine Granularity Scalability in MPEG-4 Video Standard," IEEE Trans. Circuit and System for Video Tech

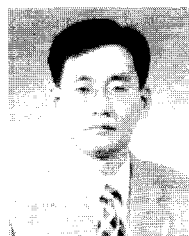
nology, vol. 11, no. 3, pp. 301-317, March 2001.
 [2] H. Radah, M. Van der Schaar and Y. Chen, "The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming Over IP," IEEE Trans. Multimedia, vol. 3, no. 1, pp. 53-68, March 2001.
 [3] K. Rao and P. Yip, Discrete Cosine Transform, Academic Press. INC, New York, 1990.
 [4] K. Rao and P. Yip, "Human Visusal Weighted Progressive Image Transmission," IEEE Trans. Comm. vol. 38, no. 7, pp. 1040-1044, July 1990.
 [5] W. Li, Bit-plane coding of DCT coefficients for fine granularity scalability, ISO/IEC JTC1/SC29/WG11, MPEG98/M3989, Oct. 1998.
 [6] S.-C. S. Cheung and A. Zakhor, Matching persuit coding for fine granular video scalability, ISO/IEC JTC1/SC29/WG11, MPEG98 /M3991, Oct. 1998.
 [7] W. Li, Fine granularity scalability using bit-plane coding of DCT coefficients, ISO/IEC JTC1/SC29/WG11, MPEG98/M4204, Dec. 1, 1998.

저 자 소 개



金昇煥(正會員)
 2001년 2월 : 아주대학교 전자공학과 졸업(학사). 2003년 2월 : 광주과학기술원 정보통신공학과 졸업(석사). 2003년 3월~현재 : 광주과학기술원 정보통신공학과 박사과정. 소속기관 : 광주과학기술원 정

보통신공학과. <주관심분야 : 영상신호 처리 및 압축, 멀티미디어 통신, 비디오 계위 부호>



扈堯盛(正會員)
 1981년 2월 : 서울대학교 전자공학과 졸업(학사). 1983년 2월 : 서울대학교 전자공학과 졸업(석사). 1983년 3월~1995년 9월 : 한국전자통신연구소 선임연구원. 1989년 12월 : University of California,

Santa Barbara Dept. of Electrical and Computer Eng. 졸업(박사). 1990년 1월~1993년 5월 : 미국 Philips 연구소 Senior Research Member. 1995년 9월~현재 : 광주과학기술원 정보통신공학과 교수. 소속기관 : 광주과학기술원 정보통신공학과. <주관심분야 : 디지털 신호처리, 영상신호처리 및 압축, 초저속 영상통신, 디지털TV와 고선명 TV 방식, 멀티미디어 통신, 디지털 워터마킹>