

가변 길이 부호화를 이용한 적응 3차원 변환 부호화 기법

(On the Adaptive 3-dimensional Transform Coding Technique Employing the Variable Length Coding Scheme)

金宗源*, 李新浩**, 李商郁*

(Jong Won Kim, Shin Ho Lee and Sang Uk Lee)

要約

본 논문에서는 연속영상의 부호화를 위해 3차원 DCT(discrete cosine transform) 변환 영역에서 수평·수직·시간 영역에서의 에너지 분포에 적응적인 움직임 영상 부호화 기법을 제안한다. 제안한 방식에서는 3차원 DCT 블록들을 에너지 집중에 따라 분류한 다음, 분류화된 적응 선형주사 기법과 적합한 스텝 크기로 설계된 가변길이 부호(VLC: variable length code) 테이블을 선택하는 기법을 활용하여 부호화를 수행한다. 또한 차신호를 이용하는 프레임간 부호화를 적응적으로 결합하는 구조를 제안하고, 움직이는 영역이 적은 영상에 대해서는 원영상에 3차원 변환 부호화를 적용하는 방식에 비해 상당한 성능 향상을 보임을 제시한다. 다양한 움직임 영상들에 대한 심도 있는 모의실험을 통해서 제안한 3차원 변환 부호화 방식의 성능과 입력 영상 특성의 관계를 분석하였으며, 0.5~1.0 bpp 전송률 대에서는 2차원 운동 보상 변환 부호화 방식에 비해 높은 성능을 대부분의 연속 영상들에 대해서 얻을 수 있음을 입증하였다.

Abstract

In this paper, employing the 3-dimensional discrete cosine transform(DCT) for the utilization of the temporal correlation, an adaptive motion sequence coding technique is proposed. The energy distribution in a 3-D DCT block, due to the nonstationary nature of the image data, varies along the vertical, horizontal and temporal directions. Thus, aiming an adaptive system to local variations, adaptive procedures, such as the 3-D classification, the classified linear scanning technique and the VLC table selection scheme, have been implemented in our approach. Also, a hybrid structure which adaptively combines inter-frame coding is presented, and it is found that the adaptive hybrid frame coding technique shows a significant performance gain for a moving sequence which contains a relatively small moving area. Through an intensive computer simulation, it is demonstrated that, the performance of the proposed 3-D transform coding technique shows a close relation with the temporal variation of the sequence to be coded. And the proposed technique has the advantages of skipping the computationally complex motion compensation procedure and improving the performance over the 2-D motion compensated transform coding technique for rates in the range of 0.5 ~ 1.0 bpp.

*正會員, 서울大學校 制御計測工學科
(Dept. of Control & Instr. Eng., Seoul
National Univ.)

**正會員, 金星社 中央研究所
(Gold Star Central Research Center)
接受日字: 1992年 7月 14日

1. 서론

움직임 영상은 정지영상들의 연속적인 프레임들로 구성되며, 이러한 특성으로 인 해서 수평, 수직 방향의 상관성 뿐만 아니라 시간상의 상관성 활용이 대단히 중요하다. 그러므로 움직임 영상의 부호화에는 움직임을 탐색하여 보상하는 움직임 보상 변환 부호화(motion compensated transform coding: MCTC) 방식이 주로 이용되고 있다.^[1] 그런데 블럭 정합 방식(block matching algorithm: BMA)^[2] 등 움직임 추정 기법들이 요구하는 계산량은 매우 크다.^[3] 또한 잘못 추정된 움직임 정보를 사용하면 움직임 손상(motion artifact)과 같은 문제점이 발생할 우려가 있다. 따라서 움직임 보상을 사용하지 않으면서 시간상의 상관성을 제거할 수 있도록 시간축 상으로도 변환을 수행하는 방식이 3차원 변환 부호화 기법이다.^[4,9] 이 방식은 움직임 보상 기법이 효과를 발휘하는 매우 낮은 전송률에 있어서는 움직임 보상 부호화에 미치지 못하며, 연속적인 프레임들을 동시에 처리해야 하기 때문에 시간상의 지연이 발생되고 프레임 메모리의 크기가 증가하는 단점이 있다. 하지만 움직임 추정 과정에 소요되는 계산량과 하드웨어의 복잡성을 줄일 수 있을 뿐만 아니라 이전 프레임에 대한 의존성이 없으므로 프레임 리프레시(refresh)를 수행할 필요가 없다는 특성을 지닌다.^[10]

Roese, Pratt, Robinson^[4]에 의해 처음 제안된 3차원 변환 부호화는 구역(zonal) 부호화를 이용하는 간단한 방식이었다. 그런데 움직임 영상에서는 공간 세부상(detail), 움직임 등의 공간-시간적인 특성이 국부적으로 달라진다. 따라서 국부적인 특성을 고려하여 분류를 수행하고 각 부류의 특성에 맞게 부호화하는 적응 기법이 제안되어 있다.^[5,9] 그리고 양자화할 계수들의 위치를 옥트리(octree) 구조를 이용하여 효과적으로 전송하려는 시도도 있다.^[6] 한편, 근래에 와서는 DCT 계수의 확률분포에 근거한 불균일 양자기를 이용하는 종래의 기법과는 달리 균일 양자기에 이은 2차원 엔트로피 부호화를 활용하는 방식이 변환 부호화에 널리 이용되고 있다.^[11-13] 그리하여 카메라의 평면이동(panning), 원근이동(zooming)과 같은 영상 전반에 걸친 움직임 보상이나 분할대역(subband) 부호화 기법에 이러한 가변길이 부호화(variable length coding: VLC) 방식의 3차원 변환 부호화를 결합한 기법들이 제안되고 있다.^[7,8]

본 논문에서는 이와 같은 사항들에 초점을 맞춰 3차원 변환 부호화 기법을 수행함에 있어서 다양한 적응 기법들의 효과적인 구현 방법을 연구하였다. 먼저

수평·수직·시간 영역에 대한 에너지 집중도 및 집중 방향에 따라 3차원 DCT 블럭들을 분류하는 새로운 기법을 제안하였다. 그리고 영들의 줄길이(run-length)와 계수 크기에 대한 2차원 VLC를 3차원 블럭에 효과적으로 결합하기 위해서 3차원 DCT 블럭의 분류 기법과 결합된 분류화된 적응 선형 주사(linear scanning) 기법을 사용하였다. 그리고 2차원에 비해 증가된 동적(dynamic) 범위를 지니는 3차원 DCT 계수들을 효율적으로 부호화할 수 있도록 적합한 스텝 크기로 설계된 VLC 테이블을 선택하도록 하였다. 그리고 차신호를 이용하는 프레임간(inter-frame) 부호화를 적응적으로 결합하는 혼성(hybrid) 구조를 함께 제시하였다. 최종적으로는 다양한 움직임 영상에 대한 심도있는 모의실험을 통해서 제안한 적응 3차원 변환 부호화 방식의 성능을 평가하였다. 그 결과 입력 영상의 특성에 따른 제안한 부호화 방식의 성능을 심도 있게 분석하였으며, 대부분의 연속 영상들에 대해 0.5 bpp 이상의 전송률에 있어서는 2차원 움직임 보상 변환 부호화 기법을 능가하는 성능을 지니는 효과적인 부호화 방식을 확인할 수 있었다.

II. 적응 3차원 변환 부호화 기법

$L \times M \times N$ 크기 3차원 블럭 $f(x, y, t)$ 에 대한 3차원 DCT는 다음과 같이 정의된다.

$$F(u, v, w) = \sum_{i=0}^{L-1} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} A(u, v, w; j, k, i) \cdot f(x, y, t), \tag{1}$$

$$A(u, v, w; x, y, t) = C(u) \cdot C(v) \cdot C(w) \cdot \frac{2}{L} \cos \left[\frac{u(2x+1)\pi}{2L} \right] \cdot \frac{2}{M} \cos \left[\frac{v(2y+1)\pi}{2M} \right] \cdot \frac{2}{N} \cos \left[\frac{w(2t+1)\pi}{2N} \right], \tag{2}$$

여기서 $C(0) = 1/\sqrt{2}$ 이고 다른 u, v, w 에 대해서 $C(u) = C(v) = C(w) = 1$ 이다. 그리고 3차원 DCT 변환은 1차원 DCT 변환을 수평, 수직, 그리고 시간축으로 각각 행함으로서 가능하다. 본 논문에서는 공간 영역에서의 변환 크기는 일반적으로 이용되는 8×8 크기를 선정하였으며, 시간축상으로는 4 프레임들에 걸친 변환을 적용하였다. 물론 시간축상의 변환의 크기를 증가시키면 변환에 의한 에너지 집중 및 그에 대응하는 부호화 이득을 얻을 수 있다. 하지만 한 블럭에 속하는 화소들간의 균일성을 유지하여 효과적으로 분류가 이루어지도록 한다는 측면이나 프레임 수의 증가에 따른 지연 문제를 고려할 때 $8 \times 8 \times 4$ 크기를 적절한

크기로 판단하였다. 이러한 크기로 3차원 DCT를 수행하면 변환된 dc 계수의 범위는 0 ~ 4096 이고, ac 계수는 -4096 ~ 4096의 범위를 지니므로 DCT 계수들의 동적 범위가 상당히 넓다는 것을 알 수 있다.

3차원 DCT를 이용한 변환 부호화 기법의 블럭선도를 그림 1에 도시하였다. 제안한 기법에서는 입력 영상 프레임들을 먼저 8×8×4 크기의 블럭들로 분할하고 3차원 DCT 변환한다. 이어서 선형 주사, dc 계수의 균일 양자화에 이은 고정길이 부호화, ac 계수들의 균일 양자화와 2차원 VLC를 이용한 부호화 과정을 거친다. 이 과정에서 분류화 기법을 결합하여 선형 주사 방법을 부류에 따라 변화시키며, 부호화된 데이터는 완충장치인 버퍼를 통해서 전송한다. 그런데 영상 프레임들은 영상 내에 존재하는 물체의 형태, 크기, 그리고 움직임에 따라 시시각각으로 특성이 변한다. 따라서 본 논문에서는 변화에 적응성을 지니는 부호화 기법에 목적을 두고 4 프레임 영상 단위마다 분류 기법, 선형 주사 형태, 그리고 VLC 테이블 선택 과정을 적응화하였다. 적응화에 관련된 부분들은 그림 1에 강조하여 표시하였으며, 제안한 적응 3차원 변환 부호화는 적응화에 필요한 통계적인 특성을 얻는 과정을 일차적으로 수행하는 2 패스(pass) 방식으로 구현된다.

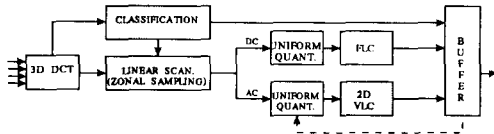


그림 1. 프레임내 적응 3차원 변환 부호화 구조
Fig. 1. Block diagram for the intra-frame 3-D transform coder.

한편, 연속된 영상 프레임들 간에 상관성이 크면 같은 위치상의 화소들은 비슷한 값을 가지므로 영상 내의 물체가 움직이지 않고 카메라도 고정된 경우라면 변화는 거의 없다. 그러므로 움직임 영상은 프레임내의 국부적인 영역에 따라서 변화가 많은 부분과 그렇지 않은 부분으로 나누어 질 수 있으며, 변화가 적은 부분에서는 단순히 차영상을 구하더라도 부호화 이득을 얻을 수 있다. 따라서 이전에 부호화된 마지막 프레임과의 차이에서 얻는 차신호들과 원래 신호를 비교함에 의해서 프레임내/프레임간 부호화를 선택하는 기법이 효과적이며, 이러한 프레임내/프레임간(intra-/inter-frame) 부호화를 적응적으로 수행하는 구조를 그림 2에 제시하였다. 이때 [7]에서

와 같이 카메라의 평면이동이나 원근이동과 같은 전체적인 움직임을 간단한 계산만으로 보상하는 기법을 사용하는 것도 가능하다. 그런데 프레임내/프레임간 적응 부호화를 위해서는 차영상을 얻고 원/차 신호간의 적응적인 선택을 수행하는 부분이 필요하다. 또한 이전에 부호화된 프레임들에 대한 의존성을 지니므로 리프레쉬(refresh)에 대한 배려가 필요하다는 문제점도 있다. 하지만 사용 영상에 따라서 움직이는 영역이 적은 경우에는 상당한 성능 향상을 보인다는 점을 간과할 수는 없다.

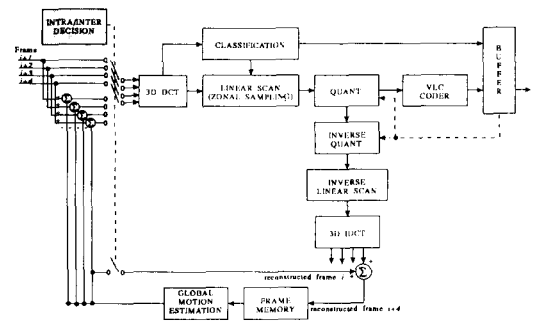


그림 2. 프레임내/프레임간 적응 3차원 변환 부호화 구조
Fig. 2. Block diagram for the adaptive intra-/inter-frame 3-D transform coder.

Ⅲ. 3 차원 DCT 블럭을 위한 적응적인 분류 기법

적응적인 분류 기법을 설명하기 위해서는 그 대상이 되는 3차원 DCT 블럭들의 에너지 분포 특성을 살펴보는 것이 필수적이다. 따라서 그림 3에 제시한 Miss America, Salesman, Calendar & Mobile, 그리고 Tempete 연속 영상의 3차원 DCT 계수들의 분산 분포를 상용로그(log₁₀) 단위로 그림 4에 각각 제시하였다. 움직임을 반영하는 시간상의 에너지 분포는 Miss America, Salesman, Tempete, Calendar & Mobile 영상의 순서로 증가되어 나타남을 알 수 있다. 먼저 Miss America 영상의 경우에는 밝기 변화가 적고 배경이 단순하므로 공간적으로 높은 주파수 대의 에너지가 매우 작게 나타난다. 그리고 움직임에 있어서는 영상의 대부분을 차지하는 인물 전체가 움직이고 있으므로 전체 시간 주파수대에 움직임에 대응하는 에너지들이 비교적 넓게 분포한다. 이에 비해서 Salesman 영상은 영상속에서 상대적으로 작은 인물이 복잡한 배경속에서

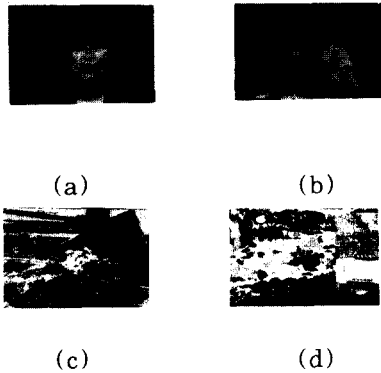


그림 3. 사용한 원 움직임 영상들
(30번째 프레임, 위도 성분)
(a) Miss Ameerica (b)Salesman
(c)Tempete (d)Calender & Mobile

Fig. 3. Original moving sequence images (30th frame).

(a)Miss America, (b)Salesman,
(c)Tempete, (d)Calendar & Mobile.

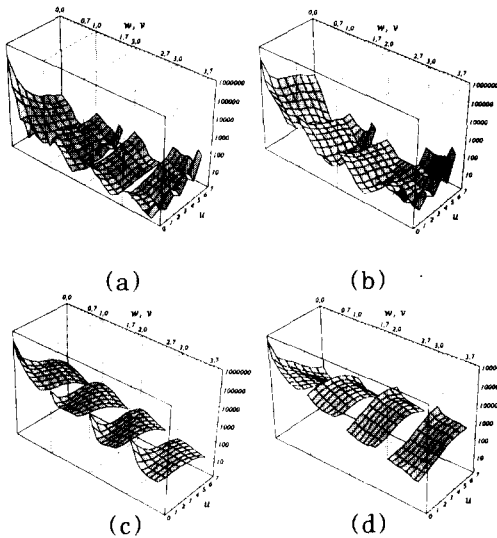


그림 4. 3차원 DCT블럭의 분산 분포
(a) Miss Ameerica (b)Salesman
(c)Tempete (d)Calender & Mobile

Fig. 4. Variance distributions for the 3-D DCT block.

(a)Miss America, (b)Salesman,
(c)Tempete, (d)Calendar & Mobile.

빠른 움직임을 보이므로 Miss America 영상보다 전체적으로 증가된 에너지 분포를 보인다. 그러나 영상 전체에서 적은 영역에만 움직임이 존재하므로 시간 주파수 대의 에너지는 크게 증가하지는 않음을 관찰할 수 있다. 한편 Tempete 영상에서는 원근이동하는 카메라의 움직임에 따른 시간상의 에너지 분포 상황을 보여 주고 있으며, 복잡한 영상 구성에 의해 전체적인 공간 주파수 영역으로 에너지가 넓게 분포함을 알 수 있다. 그리고 Calendar & Mobile 영상에서는 모형 기차와 카메라의 이동에 따른 에너지들이 각 시간 주파수 대마다 현저하게 나타남을 관찰할 수 있다. 정리하면 3차원 DCT 블럭의 에너지 분포는 입력 영상의 특성에 따른 변화를 반영하고 있으며, 제안한 방식의 성능 향상을 위해서는 이러한 특성에 적응적인 기법이 필요하다는 것을 알 수 있다.

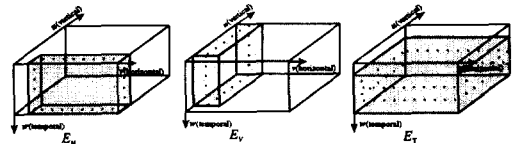


그림 5. 제안한 3차원 DCT 블럭의 분류를 위한 특징 영역들

Fig. 5. Feature regions for the 3-D block classification.

따라서 본 논문에서는 입력 블럭들을 각각의 특성에 맞도록 부호화하여 성능을 향상시키기 위해서 블럭들을 특성에 따라 분류하는 기법을 사용하였다. 3차원 블럭의 특성을 나타내는 요소에는 움직임, 가장자리 방향(edge direction) 등이 있다. 움직임은 시간축 방향으로의 상관 관계를 결정하게 되며, 움직임이 커질수록 프레임간 상관성은 낮아지는 특성을 지닌다. 이를 변환된 3차원 DCT 블럭에서 살펴보면 움직임이 커질수록 시간축 방향 주파수 영역에 고루 에너지가 분포하게 됨을 알 수 있다. 또한 가장자리 방향 및 강도는 수평, 수직 방향에 있어서 에너지의 집중 형태를 반영한다. [15, 16] 이러한 성질들을 이용하여 본 논문에서는 수평 수직 시간 영역에 대한 에너지 집중도 및 집중 방향에 따라 3차원 DCT 블럭들을 분류하는 새로운 기법을 제안한다. [9] 그리하여 이러한 3차원 DCT 블럭들을 효과적으로 분류하기 위한 특징(feature)으로 $L \times L \times M$ 크기 블럭의 각 방향으로의 에너지 집중 정도를 반영하는 수평 수직 시간 영역 에너지 $E_H E_V E_T$ 를 각각 정의하였다. 그림 5에 제시한 영역내에 포함된 계수들의 제곱의 합에 해당하는 이들 특징들은 식 (3.1)에서 식 (3.3)과 같다.

$$E_H = \sum_{u=0}^1 \sum_{v=1}^{L-1} \sum_{w=0}^{M-1} F^2(u, v, w) \tag{3}$$

$$E_V = \sum_{u=1}^{L-1} \sum_{v=0}^1 \sum_{w=0}^{M-1} F^2(u, v, w) \tag{4}$$

$$E_T = \sum_{u=0}^{L-1} \sum_{v=0}^1 \sum_{w=1}^{M-1} F^2(u, v, w) \tag{5}$$

여기서 $F(u, v, w)$ 는 변환된 영역에서의 계수를 나타낸다. 제안한 분류화 기법은 이러 한 E_H, E_V, E_T 특징들로 이루어지는 3차원 특징 공간을 그림 6에 제시한 과정으로 분할함에 의해서 입력 블럭들을 분류한다. 먼저 E_H 및 E_V 에 대해 기준치 T_{Hv} 를 사용하여 공간 영역에서의 에너지 방향 및 크기에 따라 4개의 공간 부류 범주(category)들로 나눈다. 이어서 각 공간 부류 범주에 속한 블럭들을 시간상의 에너지 E_T 를 분리하는 기준치들인 $T_{C0} \sim T_{C3}$ 를 사용하여 각각 2개의 부류들로 분리한다. 즉 본 논문에서 제안한 분류기는 공간 가장자리 형태의 분류에 이어 시간상의 에너지에 따른 분류를 별도로 수행하는 분류기에 해당된다. 전체적으로 보면 입력 블럭들은 4개의 공간 부류 범주들이 각각 2개의 시간 부류들로 나누어지므로 모두 8개의 부류들로 나누어 진다. 먼저 부류 0, 4에는 에너지가 작은 균일한 영역들이 포함되며, 부류 1, 5는 에너지는 크면서 뚜렷한 방향성을 지니지 않는 부류들이다. 부류 2, 6은 수직 가장자리 성분이 강한 부류들이고, 부류 3, 7은 수평 가장자리 성분이 강한 부류들이다. 이를 다시 움직임에 따라 나눠보면 부류 0~3은 대체적으로 움직임이 적은 부류이고 부류 4~7은 움직임의 변화가 심한 부류들이다.

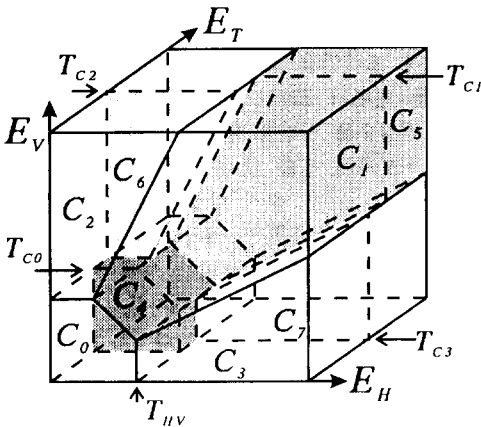


그림 6. 제안한 3차원 분류 기법의 특정 공간 분할
Fig. 6. Scatter diagrams for the proposed classification.

한편 입력 영상들에 따른 특성 변화를 감안하기 위해서는 분류화에 사용되는 기준치들을 입력 영상에 따라 가변시켜 항상 부류마다 적절한 수의 블럭들이 포함되도록 하는 것이 효과적이다. 따라서 본 논문에서는 기준치 T_{Hv} 를 4 프레임의 입력영상 단위마다 모든 블럭에서의 E_H 및 E_V 의 평균에 적절한 상수를 곱한 값을 사용하며, $T_{C0} \sim T_{C3}$ 은 각 공간 부류 범주에 속한 블럭들의 E_T 의 평균치로 설정한다.^[9] 그리고 각각의 부류는 서로 다른 에너지 분포를 가지므로 부류마다 선형 주사 기법을 달리 사용하여 부류의 특성에 따른 부호화를 수행할 수 있다. 또한 변환된 DCT 블럭의 각 계수들이 갖는 시각 특성을 활용하기 위해 인간 시각 특성(human visual system: HVS)에 의한 주파수 가중(frequency weighting)을 주는 기법^[14]을 결합함에 있어서도 부류마다 가중치를 달리 사용할 수 있으므로 보다 효율적이다.^[8]

IV. 3차원 DCT 블럭을 위한 분류화된 적응 선형 주사 기법

영들의 줄길이를 이용하는 2차원 VLC를 적용하기 위해서는 DCT 블럭의 계수들을 1차원으로 배열하여야 한다. 2차원의 경우에는 이를 위해서 계수들의 에너지(variance) 순서에 부합되는 갈깃자(zigzag) 주사가 널리 이용되며, 이러한 주사 기법을 각 시간 주파수대 마다 별도로 수행하여 3차원 블럭을 부호화할 수도 있다.^[6] 하지만 3차원적인 에너지 분포 특성을 활용함에 있어서는 미흡한 점이 있다. 따라서 본 논문에서는 수평·수직·시간 영역으로의 에너지 집중에 근거한 분류기법과 결합하여 분류화된 적응 선형 주사 기법을 제안하였다.

효율적인 주사 형태가 되기 위해서는 3차원 DCT 블럭의 에너지 분포를 반영하여 선형 주사된 DCT 계수들이 가변 길이 부호화에 적합하여야 한다. 본 논문에서는 이를 위하여 3차원 DCT 블럭의 에너지 분포를 반영하는 척도로 선형 주사 형태 평면(linear scanning shape plane)을 도입하였다. 그림 7(a)에서 제시한 주사 형태 평면은 수직 수평 시간축 상의 지표인 P_v, P_h, P_t 점들에 의해 정의된다. 3차원 블럭을 주사 형태 평면에 평행한 선형 주사 조각(slice)들로 균일하게 분할한 다음, 각각의 조각내에서 주사를 수행하고 인접한 조각에 대해 이를 반복하는 방식으로 주사 순서가 결정된다. 이 과정에서 하나의 선형 주사 조각내에서는 가장 낮은 시간 주파수대 계수들 부터 주사를 시작하고, 높은 시간 주파수대 계수들로 연결한다. 또한 그림 7(a)에 표시한 선형 주사

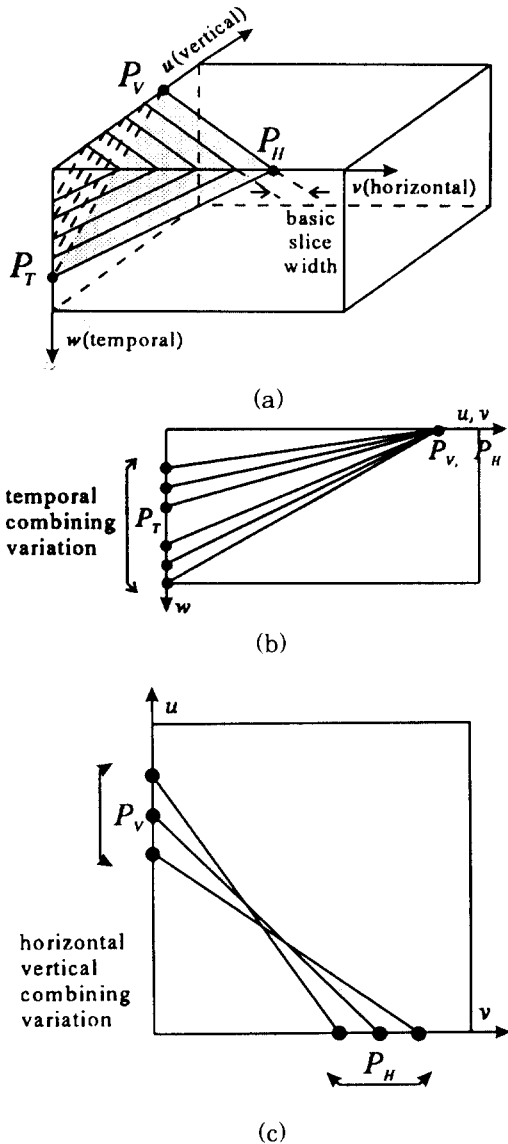


그림 7. 분류화된 적응 선형 주사 결정 기법

- (a) 선형주사 형태 평면의 구성
- (b) 시간축 결합에 따른 적응화
- (c) 가장자리 방향성에 따른 적응화

Fig. 7. Determination methods for the adaptive classified linear scanning shapes.

- (a) Sinear scanning shape plane.
- (b) Adaptation method for temporal direction coupling.
- (c) Adaptation method for the spatial edge direction.

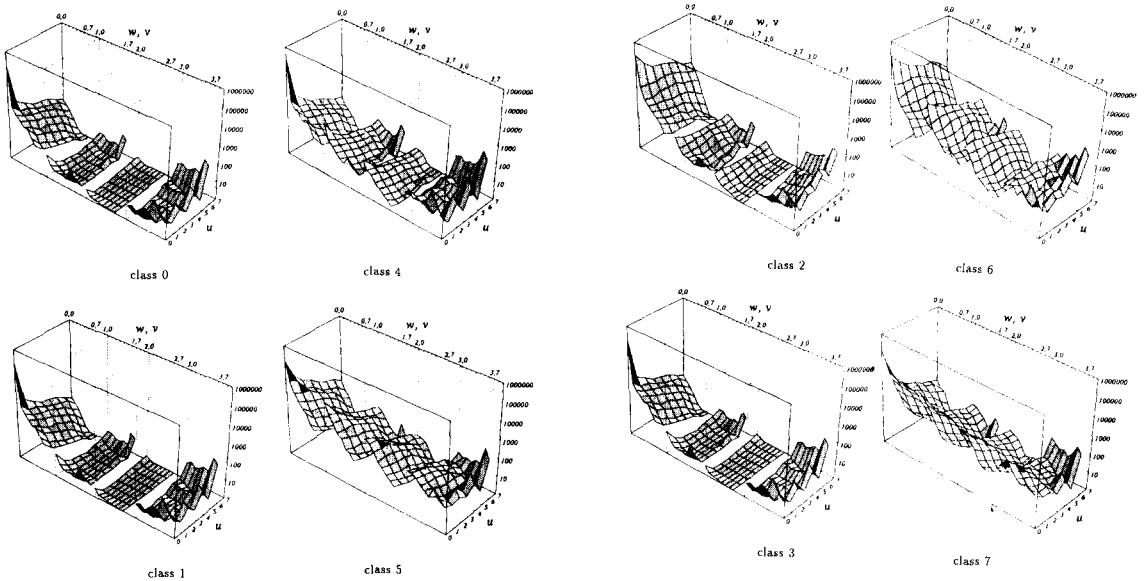
조각의 기준 넓이(basic slice width)는 조각내에 일정 수 이상의 DCT 계수들의 위치를 포함하도록 설정하였으며, 이를 상수 C_s 를 통해 조절하여 선형 주사 조각의 넓이를 변화시키도록 하였다. 종합하면 최종적인 주사 형태는 P_V, P_H, P_T 와 C_s 의 네가지 값들에 의해 유일하게 설정될 수 있는 것이다.

상기한 주사 형태 설정 기법을 활용하면, 공간-시간축 상의 결합 정도는 P_V, P_H 와 P_T 의 비율에 의해 조절할 수 있으며, 그림 7(b)에 이러한 과정을 도식적으로 표현하였다. 이때 이산적인(discrete) 좌표를 지나는 DCT 계수들의 위치를 고려하고 실제적인 DCT 계수들의 에너지 분포를 감안하여 시간적인 결합의 정도를 결정하는 P_V/P_T 또는 P_H/P_T 의 값을 8/8, 8/6.6, 8/5, 8/4, 8/3.5, 8/3, 8/2.7, 8/2.4의 8가지 경우에 한정하였다. 또한 공간상의 에너지 집중에 대해서는 분류 기법과 결합하여 부류마다 수직-수평 공간의 결합 비율인 P_V/P_H 를 적절하게 설정하도록 하였다(그림 7(c) 참조). 그리고나서 시간축 상의 변화에 대한 적응성을 부여하는 P_T 와의 관계를 연결함으로써 사용할 선형 주사 형태들을 결정하였으며, 이를 이용하여 분류화된 선형 주사 형태의 적응 기법을 구현하였다. 단 P_V/P_H 비율과 P_T 와의 관계를 연결함에 있어서 실제적인 DCT 계수들의 에너지 분포 구성에 비해 비현실적인 P_V, P_H, P_T 쌍은 제외시켜야 한다. 그 결과 분류화된 경우를 위해서는 P_V/P_H 의 비율을 8/2, 8/4, 8/6, 8/8, 6/8, 4/8, 2/8의 7가지 경우로 설정하고 P_T 와의 관계를 통해서 40가지의 적응 선형 주사 형태로 확장하였다.

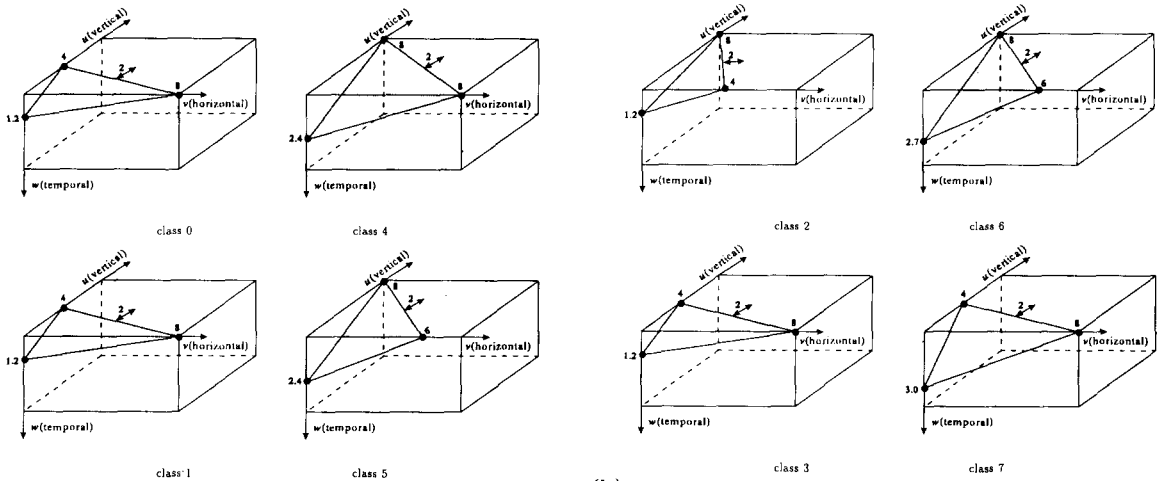
실제적인 구현 측면에서 본 논문에서 사용한 선형 주사 형태의 적응화는 영상에서 얻는 3차원 DCT 블록의 분산 분포와 미리 결정된 P_V, P_H, P_T, C_s 쌍에 의한 선형 주사 형태들과의 정합 순위에 따라 결정된다. 즉 일종의 패턴 정합 방식에 의해서 주어진 DCT 분산 분포에 최적으로 부합되는 선형 주사 형태를 선택하고, 주사 형태 선택정보를 전송하는 방식으로 구현된다. 또한 그림 8(a)에는 Salesman 영상에 대해서 부류별 DCT 분산 분포를 제시하였다. 이에 대해 적응화된 선형 주사 형태를 그림 8(b)에 각각 제시하였으며, 각 부류의 선형 주사 형태들이 공간상의 가장자리 방향성과 시간상의 에너지 분포에 부합됨을 관찰할 수 있다.

V. DCT 계수들의 양자화와 적응적인 VLC 테이블 선택 기법

본 논문에서는 변환 부호화의 문제점인 블록 현상



(a)



(b)

그림 8. 분류화된 분산 분포 및 선형주사 형태(Salesman 3~6번째 프레임들)
 (a) 분류화된 DCT 블록의 분산 분포들, (b)분류화된 선형 주사 형태들

Fig. 8. Classified variance distributions and the corresponding classified linear scanning shapes(Salesman sequence).

- (a) Classified variance distributions for the 8 classes.
- (b) Classified linear scanning shapes for the 8 classes.

을 감안하여 dc 계수는 10 비트로 균일 양자화하고 고정길이 부호로 전송하는 방식을 사용하였다. 또한 넓은 동적 범위를 지니는 3차원 ac 계수들은 확률분

포에 최적화된 불균일 양자기의 사용이 효과적으로 판단된다. 하지만 불균일 양자기를 사용하면서 버퍼 제어를 수행하려면 다양한 전송률에 대한 불균일 양

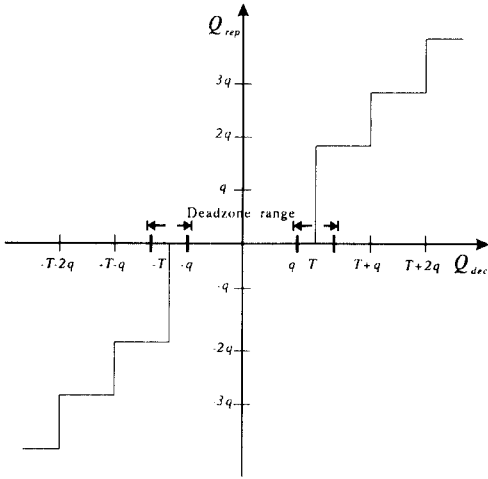


그림 9. 데드존이 있는 균일 양자기의 양자화 특성
 Fig. 9. Quantization characteristic of the uniform quantizer with deadzone.

자기들의 설계 및 복잡한 버퍼 제어 방식과 같은 복잡한 처리 과정이 필요하다.^[14] 따라서 본 논문에서는 ac 계수들을 그림 9와 같은 가변 기준치 데드존(variable threshold deadzone)을 사용하는 균일 양자기를 사용하여 양자화 하였다.^[12] 이어서 양자화된 계수들은 영들의 줄길이나 계수 크기에 대한 2차원 VLC로 부호화한 다음, 버퍼를 통해 일정한 전송률로 전송하였다. 버퍼 제어는 적당한 수의 블록들을 단위로 하여 버퍼의 상태에 따라 양자기의 스텝 크기 q 를 조절하는 간단한 방식을 사용하였다. 그리하여 버퍼 넘침(overflow)이나 버퍼 고갈(underflow)이 없는 양자화를 수행할 수 있었지만, 복잡한 버퍼 제어 기법을 사용하면 보다 향상된 양자화를 수행할 수 있다고 믿어진다.

이와 같은 양자화 과정에서 넓은 범위를 지니는 3차원 ac 계수들을 효율적으로 양자화하기 위해서는 부호화가 이루어지는 상태에 부합하는 스텝 크기로 설계된 VLC 테이블을 선택해야 한다. 그런데 이를 효과적으로 구현하기 위해서는 전송률과 버퍼 수준이 주어졌을 때 부호화 과정에 사용되는 스텝 크기의 평균치를 정확하게 예측하는 것이 필요하다. 그리하여 실제적인 양자화 과정을 일차적으로 수행하는 방식을 배제하고 이를 위한 효율적인 방법을 모색한 결과 본 논문에서는 [17]에서 사용한 해석적인 기법을 변경하여 사용하였다. 먼저 정지 영상 부호화의 표준 기법인 JPEG^[13]에 대한 분석 논문인 [17]에서는 전체적인 JPEG 부호화 과정이 전송량과 왜곡치의 관계에 의해서 분석되어 있다. 따라서 양자화할 DCT

계수들의 확률 분포, 사용하는 양자기의 스텝 크기 집합과 선형 주사 형태, 그리고 VLC 테이블 등 부호화 관련 변수들이 주어진 경우에는 예상되는 부호화 전송량 및 각 계수들의 예상 왜곡치(distortion)들을 수식적으로 계산할 수 있다. 특히 라플라시안 분포를 가지는 경우에 대해서는 부호화 전송량 또는 계수들의 왜곡치를 전술한 부호화 변수들의 관계식으로 표현할 수 있으므로 수치적인 적분(numeric integration) 과정이 필요 없게 된다. 이때 확률 분포와 DCT 계수들로부터 얻는 통계적인 성질들을 연결하는 매체로는 각 계수들의 분산치를 사용하면 된다. 따라서 본 논문에서는 이러한 관계식들 중에서 부호화 전송량에 관한 부분을 3차원 변환 부호화의 경우로 변형한 식 (4)를 이용하여, 표준적인 스텝 크기로 설계된 VLC 테이블과 선형 주사 형태에 대해 양자기 스텝 크기 q 를 사용할 때 각 블록에서 발생되는 전송량을 추정한다.

$$R_{block} = C_{dc} + \sum_{n,l,m} C_{ac}[l,m] \cdot M[n,q,m] \cdot L[n,q,l] + C_{eob} \quad (6)$$

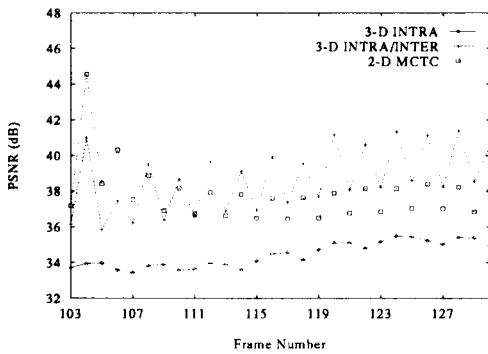
여기서 C_{ac} 와 C_{eob} 는 dc 계수와 eob(end of block)의 부호 길이를 각각 의미하며, 는 VLC 테이블에 의한 영들의 줄길이 l 와 계수 크기 m 에 대한 ac 계수들의 부호 길이 집합이다. 그리고 선형 주사된 계수 인덱스 n 에 대해서 $L[n,q,l]$ 은 영들의 줄길이가 l 일 확률, $M[n,q,m]$ 은 계수 크기가 m 일 확률이며, 구체적인 식은 [17]로 부터 쉽게 유도할 수 있다. 정리하면, VLC 테이블 선택 과정은 주어진 스텝 크기에 대한 블록당 예상 전송량 R_{block} 을 전체 영상에 대하여 환산하고, 이를 버퍼의 수준 및 부호화 전송량에 비교함에 의해 수행된다. 즉 비교를 통해서 적합한 양자기 스텝 크기를 효과적으로 결정하여, 예상된 스텝 크기에 부합하는 VLC 테이블을 선택하도록 하는 것이다.

한편, 변환에 의한 에너지 집중에 의해서, 특히 움직임이 적은 단순한 영상 부분들에서는, 높은 공간 시간 주파수 계수들이 아주 미약한 크기만을 가지고 있어 부호화시 거의 영으로 전송되는 경우가 발생된다. 따라서 이러한 계수들의 위치를 별도로 전송하여 양자화 과정에서 배제하는 방식이 구역 표본화(zonal sampling) 기법이다. 그런데 구역 표본화를 사용하려면 구역 설정을 위한 기준치를 결정하고 구역 정보를 부가적으로 전송해야 한다. 그리고 대부분의 미약한 에너지를 가지는 계수들은 선형 주사 순서에 있어 뒷 부분에 위치하므로 부호화 과정에 영향을 미치지 못하는 경우가 많다. 이와 같은 이유들로 인하여 구

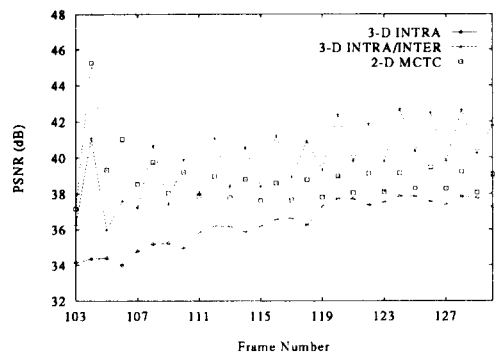
역 표본화 기법은 VLC를 사용하는 2차원 기법들에서 널리 사용되지 않고 있다. 하지만 3차원의 경우에는 움직임이 적어서 상당수의 계수들의 에너지가 미약하면 부호화되는 계수들 간의 영들의 수가 상당히 길어지고, 짧은 영들의 줄길이에 대해 적은 비트의 부호를 할당하는 VLC의 효율이 상당히 저하된다. 따라서 3차원의 경우에는 구역 표본화 기법의 효율성이 상대적으로 크다고 판단되며, 전술한 VLC 선택 기법에 활용한 해석적 기법을 활용하면 구역 표본화를 효과적으로 구현할 수 있음은 주목할 만하다. 즉 테드존을 사용하는 균일 양자기의 경우로 변형된 예상 왜곡치식을 사용하면 3차원 DCT 계수마다의 오차를 예상할 수 있으므로, 부호화를 수행하지 않을 구역을 적응적으로 표본화할 수 있다. 이때 선형 주사 형태와의 관계를 살펴서 별도의 구역 표본화의 필요성을 확인하는 과정을 추가하면 보다 효과적인 것으로 믿어지며, 앞으로 계속된 연구가 필요하다고 보인다.

VI. 모의실험 및 결과

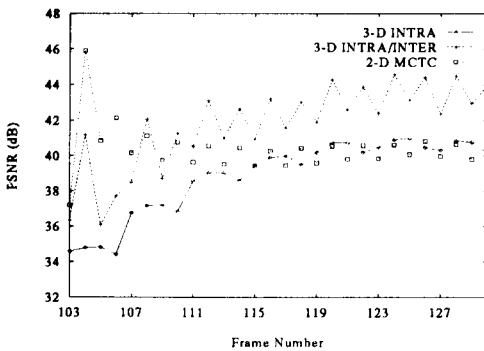
본 논문에서 제안한 적응 3차원 변환 부호화 기법을 Miss America, Salesman, Calendar & Mobile, 그리고 Tempete 연속 영상을 포함한 다양한 연속 영상들에 대하여 모의실험 하였다. 이때 비교를 위해서 CCITT RM7에 의한 2차원 움직임 보상 변환 부호화도 함께 수행하였다. 사용한 영상들은 모두 칼라 영상들이며, Y:U:V 성분의 비율이 4:1:1인 CIF(common intermediate format) 형식으로 구성된 영상을 각각 30 프레임씩 사용하였다. 그리고 본 논문에서 적용화된 3차원 변환 부호화의 모의 실험을 수행한 구체적인 과정은 다음과 같다. 먼저 적응적인 VLC 테이블의 선택은 부호화 영상을 포함한 10 종류의 연속 영상들에 대해서 8, 16, 32의 균일한 스텝 크기로 설계한 VLC 부호 테이블들 중에서 선택하도록 하였다. 그리고 현 단계에서 선형 주사



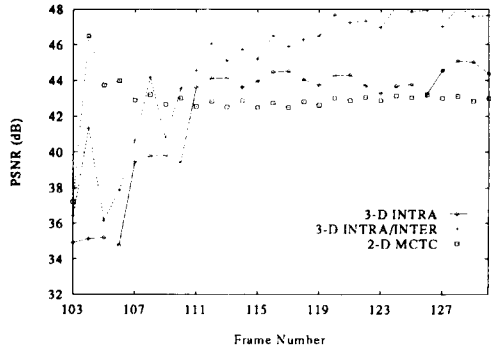
(a)



(b)



(c)



(d)

그림 10. Salesman 영상에 대한 모의 실험 결과.

(a) 0.2 *bpp* (b) 0.3 *bpp* (c) 0.5 *bpp* (a) 1.0 *bpp*

Fig. 10. 3D DCT simulation results for the Salesman sequence.

(a) 0.2 *bpp*, (b) 0.3 *bpp*, (c) 0.5 *bpp*, (a) 1.0 *bpp*.

형태 결정 과정에서 구역 표본화의 구현은 생략하였다. 또한 분류 기법의 적용은 휘도 성분에만 국한하였으며, 이는 휘도 성분에 비해서 중요성이 떨어지는 색도 성분들에 대해서는 분류 기법의 추가적인 적용이 불필요하다는 분석에 따른 것이다. 그리하여 휘도 성분만을 $8 \times 8 \times 4$ 블록별로 8개의 가장자리 및 시간 에너지 부류들로 나누어 3 비트의 분류 정보를 전송하고 분류화된 적응 선형 주사 기법을 적용하였다. 그리고 나머지 색도 성분들은 시간축상의 에너지 분포에 적응화된 방향성 없는 선형 주사 형태를 사용하였다. 이와 같은 과정으로 프레임내 3차원 변환 부호화를 수행하였으며, 프레임간 부호화의 결합은 원/차신호의 적응적인 선택을 ac 분산 비교 방식에 의해 구현하였다. 즉 $16 \times 16 \times 4$ 크기의 휘도 신호 블록 단위에 대해 차신호를 사용하는 경우의 ac 분산이 원래 신호를 사용하는 경우보다 1/4 이하인 경우에만 프레임간 부호화가 이루어지도록 하였다. 이때 색도 신호에 대해서는 휘도 신호에 의해 이루어진 선택을 동일하게 적용하는 방식을 사용하였다. 그리고 프레임간 부호화를 결합하는 경우에는 움직임 보상을 사용하는 경우와 마찬가지로 부호화를 시작할 때 필요로 하는 기준이 되는 이전 프레임은 원 영상을 그대로 사용하였다. 프레임내 3차원 변환 부호화, 프레임내/프레임간 3차원 변환 부호화, 그리고 움직임 보상을 이용하는 2차원 변환 부호화의 세가지 방식으로 수행한 모의실험 결과를 먼저 Salesman 영상을 사용한 경우에 대해 그림 10(a)에서 그림 10(d)까지 제시하였다. 이때 휘도 성분인 Y 성분에 대한 결과만을 제시하였으며, 각각의 그림들은 0.2, 0.3, 0.5, 그리고 1.0 bpp의 부호화 전송률에 대한 결과에 해당한다. 결과에 의하면 3차원 변환 부호화의 성능은 부호화 전송률이 증가함에 따라 급격하게 향상됨을 알 수 있다. 즉 낮은 전송률에서는 움직임 보상을 사용하는 경우에 비해서 낮은 성능을 보이지만 전송률이 증가하면서 점차로 격차를 좁혀 나가 1.0 bpp에서는 성능이 역전되는 추세를 보인다. 그리고 복잡하지만 고정된 배경 속에서 비교적 적은 크기를 차지하는 인물이 빠르게 움직이는 Salesman 영상이 차영상에 이용하는 프레임간 부호화 기법의 결합이 상당한 성능 향상을 보여 3가지 기법 중에서 가장 우수한 성능을 보임을 확인할 수 있다. 한편 그림 11에 모의 실험에 의해 1.0 bpp로 복원된 Salesman 영상들을 부호화 방식마다 각각 제시하였으며, 주관적인 성능 평가도 PSNR(peak signal to noise ratio)로 측정된 객관적 성능 평가와 일치함을 알 수 있다. 그리고 프레임내 3차원 변환 부호화에 의해 복원된 영상과 차신호

를 결합한 방식을 비교해 보았을 때 얼굴 부분 등에서 차신호를 이용하는 방식이 원영상에 보다 충실함을 알 수 있다.



(a) (b) (c)

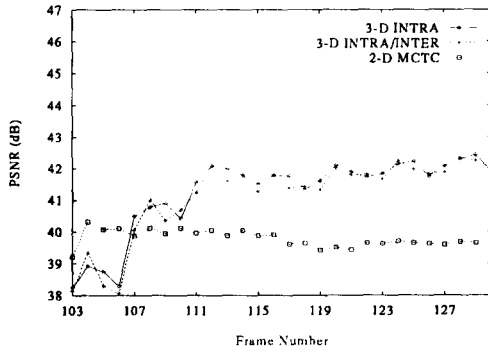
그림 11. 모의 실험에 의해 복원된 Salesman 영상들(30번째프레임, 휘도성분)

- (a) 프레임내 적응 3차원 변환 부호화 (1.0bpp)
- (b) 프레임내/프레임간 적응 3차원 변환 부호화(1.0bpp)
- (c) 움직임 보상을 사용한 2차원 부호화 (1.0bpp)

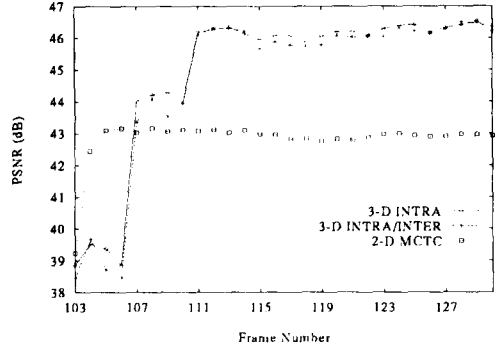
Fig. 11. Encoded Salesman sequence (30th frame).

- (a) Intra-frame 3-D transform coder an 1.0 *bpp*.
- (b) Intra-/Inter-frame 3-D transform coder at 1.0*bpp*.
- (c) 2-D transform coder with motion compensation at 1.0 *bpp*.

또한 시간상의 변화에 따른 부호화 성능 변화를 관찰하기 위해서 Miss America, Tempete, 그리고 Calendar & Mobile 영상들에 대한 결과를 0.2, 1.0 bpp에 대해 그림 12에서 그림 14까지 제시하였다. 그 결과 Miss America 영상은 항상 40 dB 이상의 매우 높은 PSNR 성능을 보이는 반면에 Calendar & Mobile 영상은 20 dB 부근에서 시작하여 30 dB에 미치지 못하는 성능을 보임을 확인할 수 있다. 그리고 대부분의 경우에 3차원 변환 부호화 방식이 움직임 보상을 사용하는 기법보다 높은 성능을 보이고 있다. 이러한 결과는 Miss America 영상의 경우에는 3차원 변환 후의 에너지 분포가 매우 단순하여 3차원 변환 부호화가 효과적으로 수행되었기 때문으로 여겨진다. 그리고 Tempete 영상의 경우에는 영상 구성이 매우 복잡하며, 카메라의 원근이동을 ब्ल록 결합 방식 움직임 추정에 의해 보상하는 것에 한계가 있기 때문으로 생각된다. 그런데 움직임을 보상



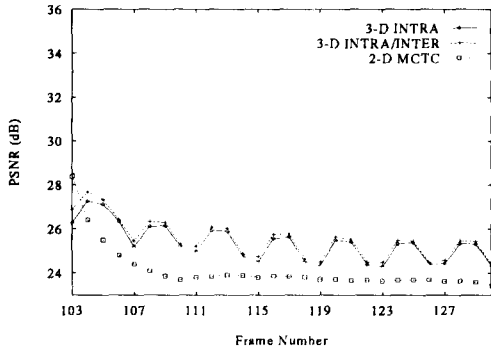
(a)



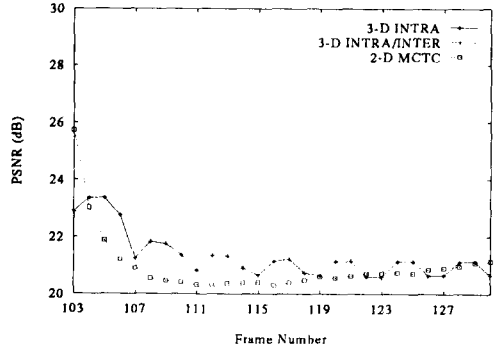
(b)

그림 12. Miss America 영상에 대한 모의 실험 결과
(a) 0.2 *bpp* (b) 1.0 *bpp*

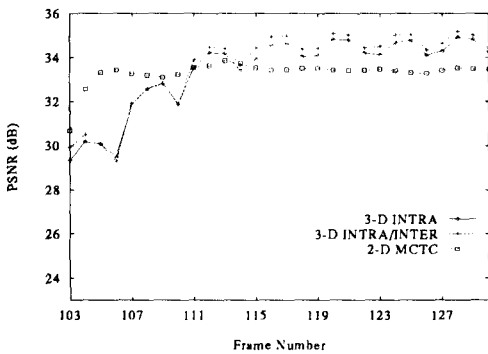
Fig. 12. 3D DCT simulation results for the Miss America sequence.
(a) 0.2 *bpp*, (b) 1.0 *bpp*.



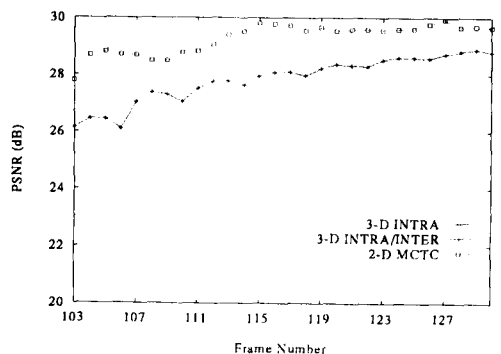
(a)



(a)



(b)



(b)

그림 13. Tempete 영상에 대한 모의 실험 결과
(a) 0.2 *bpp* (b) 1.0 *bpp*

Fig. 13. 3D DCT simulation results for the Tempete sequence.
(a) 0.2 *bpp*, (b) 1.0 *bpp*.

그림 14. Calendar & Movile 영상에 대한 모의 실험 결과
(a) 0.2 *bpp* (b) 1.0 *bpp*

Fig. 14. 3D DCT simulation results for the Calendar & Movile sequence.
(a) 0.2 *bpp*, (b) 1.0 *bpp*.

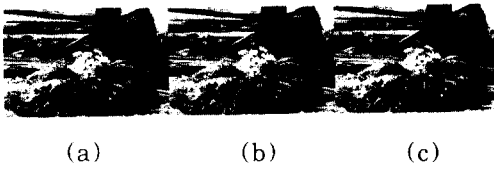


그림 15. 모의 실험에 의해 복원된 Salesman 영상들(30번째프레임, 휘도성분)
 (a) 프레임내 적응 3차원 변환 부호화 (1.0bpps)
 (b) 프레임내/프레임간 적응 3차원 변환 부호화(1.0bpps)
 (c) 움직임 보상을 사용한 2차원 부호화 (1.0bpps)

Fig. 15. Encoded Salesman sequence (30th frame).
 (a) Intra-frame 3-D transform coder an 1.0 bpps,
 (b) Intra-/Inter-frame 3-D transform coder at 1.0bpps,
 (c) 2-D transform coder with motion compensation at 1.0 bpps.

하는 기법의 성능이 Calendar & Mobile 영상의 경우에는, 위의 다른 영상들과는 달리, 약간 높다는 사실을 지적하는 것이 흥미 있다. 이와 같은 현상은 Calendar & Mobile 영상에 대해서는 3차원 DCT 블록의 변화도가 가장 심하며, 평면이동을 위주로 구성된 움직임의 보상이 비교적 용이하다는 사실에 연유한다고 믿어진다. 한편, 그림 15에 Tempete 영상의 경우에 대해서 0.2, 0.5 그리고 1.0 bpps의 전송률로 부호화한 결과를 각각 제시하였다. 제시된 영상들을 살펴보면 낮은 전송률일수록 변환 부호화의 단점인 블록 현상이 다소 나타나며, 전송률이 증가하면 사라지는 것을 확인할 수 있다.

Ⅶ. 결론

본 논문에서는 연속영상의 부호화를 위해 3차원 변환된 영역에서 수평·수직·시간 영역에서의 에너지 분포에 적응적인 움직임 영상 부호화 기법을 제안하고 있다. 모의 실험을 통해서 입력 영상의 특성에 따

른 부호화 방식의 성능을 심도 있게 분석하였으며, 0.5~1.0 bpps 정도의 전송률 대에서는 운동 보상을 사용하는 방식에 비해 향상된 성능을 대부분의 연속 영상들에 대해서 얻을 수 있음을 입증하였다. 또한 차신호를 이용하는 프레임간 부호화를 결합하는 구조를 제안하고 원영상에 적응 3차원 변환 부호화를 적용하는 방식에 비해 특정한 영상들에 있어서는 상당한 성능 향상을 보임을 제시하였다. 결론적으로 본 논문에서 제안한 3차원 변환 부호화 기법은 기존의 움직임 보상 부호화 기법에 비해 복잡한 움직임 보상 과정을 제외할 수 있는 특성을 지니고 있으면서 일정 수준 이상의 전송률 대에서는 움직임 보상 방식을 능가하는 매우 효율적인 부호화 방식임을 알 수 있다. 또한 프레임내 적응 3차원 변환 부호화는 이전 프레임에 대한 의존성이 없으므로 영상의 선택적인 재생에 매우 유리한 특성을 지니고 있다. 따라서 압축률은 비교적 낮지만 높은 화질을 구현하는 기법으로 [8]에서와 같이 디지털 VCR에 이상적으로 결합될 수 있는 특성을 지니고 있다고 판단된다. 또한 3차원 변환을 위해서 DCT 대신에 LOT(lapped orthogonal transform)를 사용하는 경우에는 추가적인 성능 향상을 가져올 것으로 기대된다. [10]

參 考 文 獻

[1] N. G. Musmann, P. Pirsh and H. J. Grallat, "Advances in picture coding," *Proc. of the IEEE*, vol. 73, no. 4, pp. 523-548, April 1985.
 [2] J. R. Jain and A. K. Jain, "Displacement measurement and it's applications in interframe image coding," *IEEE Trans. Commun.*, vol. COM-29, no. 12, pp. 1799-1808, Dec. 1981.
 [3] E. Buboiss and J. Konrad, "Review of techniques for motion estimation and motion compensation," *HDTV'90 Colloquium*, pp. 3B.3.1-3B.3.19, 1990.
 [4] J. A. Roesse, W. K. Pratt and G. S. Robinson, "Interframe cosine transform image coding," *IEEE Trans. Commun.*, vol. COM-25, no. 11, pp. 1329-1339, Nov. 1977.
 [5] M. Gotze and G. Ocylok, "An adaptive interframe coding system for images,"

- in *Proc. Internat. Conf. on Acoust., Speech, and Signal Proc.*, Paris, France, May 1982, pp. 448-451.
- [6] A. Baskurt and R. Goutte, "3-dimensional image compression by discrete cosine transform," in *Proc. Fourth European Signal Proc. Conf.*, Grenoble, France, Sept. 1988, pp. 79-82.
- [7] T. Akiyama, T. Takahashi and K. Takahashi, "Adaptive three-dimensional transform coding of moving pictures," in *Proc. PCS*, Mar. 1990, pp. 8.2.
- [8] K. Onishi, T. Itow, H. Nishikawa, et. al, "An experimental home-use digital VCR with three dimensional DCT and superimposed error correction coding," *IEEE Trans. Consumer Electronics*, vol. CE-37, no. 3, pp. 252-260, Aug. 1991.
- [9] 이신호, 김인철, 김종원, 이상욱, "적용 3차원 변환 부호화 기법," 전자 공학회 논문지 B, 제29권 B편 제12호, 1992년 12월.
- [10] Y. Q. Zhang and S. Zafar, "Motion-compensated wavelet transform coding for color video compression," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. CSVT-2, no. 3, pp. 285-296, Sep. 1992.
- [11] CCITT, "Draft revision of recommendation H. 261: Video codec for audiovisual services at $p \times 64$ Kbits/s", Documentation 572, Working Party XV/1, Specialist Group on Coding for Visual Telephony, July 1990.
- [12] CCITT, "Description of reference model 7(RM7)," Documentation 446, Working Party XV/4, Specialist Group on Coding for Visual Telephony, Jan. 1989.
- [13] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consumer Electronics*, vol. CE-38, no. 1, pp. X VIII-X X X IV, Feb. 1992.
- [14] K. H. Tzou, "An intrafield DCT-based HDTV coding for ATM networks," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. CSVT1, no. 2, pp. 184-196, June 1991.
- [15] D. S. Kim and S. U. Lee, "Image Vector Quantizer Based on a Classification in the DCT Domain," *IEEE Trans. Commun.*, vol. COM-39, no. 4, pp. 549-556, April 1991.
- [16] Y. S. Ho and A. Gersho, "Classified transform coding of images using vector quantization," in *Proc. Internat. Conf. on Acoust., Speech, and Signal Proc.*, Glasgow, Scotland, May 1989, pp. 1890-1893.
- [17] A. C. Hung and T. H. Meng, "Optimal quantizer step sizes for transform coders," in *Proc. Internat. Conf. on Acoust., Speech, and Signal Proc.*, Tronto, Canada, May 1991, pp. 2621-2624.
- [18] H. S. Malvar and D. H. Staelin, "The LOT: Transform coding without blocking effect," *IEEE Trans. Acoust., Speech, and Signal Proc.*, vol. ASSP-37, no. 4, pp. 553-559, Apr. 1989.

 著 者 紹 介

金 宗 源(正會員) 第 29卷 第 12號 參照.
 현재 서울대학교 제어계측공학과
 박사과정

李 商 郁(正會員) 第 25卷 第 3號 參照.
 현재 서울대학교 제어계측공학과
 부교수

李 新 浩(正會員) 第 29卷 第 12號 參照.
 현재 금성사 중앙 연구소
 연구원