

論文2002-39SP-6-2

DCT 기반 임베디드 동영상 부호화 및 최적 비트 배분의 기법

(DCT-based Embedded Image Sequence Coding and Bit Allocation Scheme)

鄭 且 根 *

(Cha-Keon Cheong)

요 약

동영상 부호화를 위한 새로운 방법으로 DCT 기반의 임베디드 제로트리 부호화와 연속되는 프레임에서의 최적 비트 배분을 위한 기법을 제안한다. 국제 표준화된 기존 부호화 알고리즘의 구조를 충분히 이용하면서 부호화 효율을 개선시키기 위해, 움직임 검출 및 보상과 DCT 복합 영상 부호화 구조 기반의 임베디드 제로 트리 부호화 기법을 적용한다. 이를 위해 먼저, DCT 변환 계수를 복호 영상에 미치는 중요도에 따라 트리 구조로 재배치한 후, 임베디드 제로트리 부호화 알고리즘을 수정해서 적용한다. 이때, 주어진 전체 비트율에 대해, 연속되는 각 프레임에 최적의 비트가 배분되도록 해서 최상의 복호 영상을 얻기 위한 최적 비트율 제어 기법을 제안한다. 또한, 각 프레임의 비트율 및 부호화 오차를 균일하게 제어해서 각 프레임마다 균일한 품질의 영상이 얻어지도록 한다. 제안 알고리즘의 객관적인 성능을 평가하기 위해 다양한 테스트 영상에 대해, 모의실험 결과를 제시하고 그 성능의 우수성을 입증한다.

Abstract

This paper presents a novel DCT-based embedded zero-tree coding and optimal bit allocation algorithm for image sequence coding. In order to fully utilize the structure of the conventional standard coding algorithm and improve the coding efficiency, motion estimation and compensation(ME/MC)-DCT hybrid coding structure and a modified zero-tree coding algorithm are applied. After the rearrangement DCT coefficients into pyramidal structure according to their significance on the decoded image quality, the modified embedded zero-tree coding is performed on layered coefficients. Moreover, for a given overall bit rates, a new optimal bit control scheme is proposed to achieve the best decoded image quality in the consecutive frames. The rate control scheme can also provide the equal quality of decoded image with the control of bit rate and distortion for each frame. The various simulation results are provided to evaluate the coding performance of the proposed scheme.

Keywords : DCT, 임베디드 제로트리 부호화, 비트배분, 최적 비트율 제어, 복호영상

I. 서 론

실용화가 진전되어 다양한 분야의 응용에 널리 사용

되고 있는 MPEG이나 H.26x와 같은 국제 표준화 동영상 부호화는, 입력영상으로부터 시공간적인 상관성을 제거하기 위해 움직임 추정 및 보상 (Motion Estimation and Compensation: ME/MC)-DCT의 복합 부호화 기법을 적용하고 있다. 그러나, 이들 알고리즘에서는 좁은 범위의 공간적 스케일과 인접된 프레임에서만 시공간적인 상관성을 이용할 뿐이다. 또한, 블록단위로 부

* 正會員, 湖西大學校 電氣情報通信工學部

(Hoseo Univ., The School of Electrical Engineering)

接受日字:2002年4月17日, 수정완료일:2002年9月26日

호화가 수행되므로 출력되는 비트스트림은 순차전송이나 인터넷 등의 응용에 유용한 임베디드 특성을 가질 수 없으며, 한정된 대역폭을 갖는 채널을 통해 효율적인 정보를 전송하기 위한 비트율 제어 알고리즘이 복잡하게 된다. 또한, 대역이 극히 제한된 유무선 통신채널을 통해서 보다 높은 품질의 영상정보를 제공하기 위해서는, 표준화된 부호화 기법의 개선을 포함해서 새로운 영상부호화 및 처리 알고리즘에 관한 지속적인 개발이 요구되고 있다.

영상 부호화 효율을 개선시키기 위해 제안된 다양한 기법 중에서, 웨이브릿 또는 DCT 변환공간에서의 연속적인 양자화에 기반을 둔 임베디드 제로트리에 의한 부호화 기법^[1-3] 제안되어 많은 주목을 받아 왔으며, 관련된 다양한 알고리즘들이 개발되었다.^[4,6-11] 이 기법의 기본적인 원리는, 입력영상의 변환공간으로부터 계층구조의 변환계수로 재구성하고, 인접 계층간의 유사성에 따라 계수들을 분류해서 연속근사 양자화와 엔트로피 부호화 알고리즘을 적용하는 것이다. 이들 기법은 전체 부호화 과정이 간단함에도 불구하고, 뛰어난 비트율-왜곡(Rate-Distortion : R-D)의 부호화 성능이 얻어질뿐만 아니라, 화상정보의 순차전송이나 인터넷 브라우징, 지연이 낮은 화상통신 등의 응용에 유용한 임베디드 비트스트림의 발생이 가능한 특성을 갖는다. 또한, 간단한 비트율 제어 알고리즘으로 한정된 비트 내에서 뛰어난 품질의 영상정보의 발생이 가능하다.

웨이브릿 변환은 입력영상의 유연한 공간-주파수 해상도 성분으로의 분할과, 낮은 주파수 성분으로의 높은 에너지 집중도 등의 특성으로 인해, 주로 정지영상의 부호화나 잡음제거 등 다양한 응용에 이용되어 왔다. 또한, 웨이브릿 변환 기반의 임베디드 제로트리 구조의 영상 부호화는,^[1,2,4-7] 높은 부호화 효율의 특성이 잘 알려져 있다. 그러나, 비록 동영상 부호화를 위한 3D 웨이브릿 변환 기법이 제안되기도 했으나,^[6] 움직임 보상된 차분영상의 경우, 웨이브릿 변환 고유의 특성을 충분히 이용할 수 없어, 부호화 효율의 개선에는 한계가 있다. 따라서, 웨이브릿 기반의 임베디드 제로트리 부호화 기법은 웨이브릿의 공간주파수 변환 특성으로부터 높은 시간적 상관성을 갖는 동영상 부호화에는 그 특성을 충분히 활용할 수 없을뿐만 아니라, 기존의 표준화된 동영상 부호화의 코덱 구조를 전혀 활용할 수 없는 문제가 있다.

이상과 같은 변환 기반의 임베디드 제로트리 부호화

기법에서 정지영상의 부호화에는 특별한 부호율 제어 알고리즘이 없어도 임의의 한정된 비트량에서 최적의 화질을 보장할 수 있으나, 동영상의 경우에는 시간적 상관성의 변동, 즉 움직임 추정과 보상의 정도가 프레임마다 다르므로 일률적인 비트배분으로는 최적 품질의 복호영상을 얻을 수 없다. 따라서, 최적의 복호영상을 위한 동영상 부호화에서는 각 프레임의 특성을 반영한 최적 비트배분이 필요하다. 특히, 제한된 채널 대역폭에서의 동영상 통신이나 한정된 메모리의 저장매체에 동영상 정보를 저장하기 위해서는 프레임 단위의 최적 비트배분 알고리즘 개발이 요구되고 있다.

비트배분의 기본 목적은 정해진 부호율로부터 프레임 단위별로 최적인 비트량을 배분해서 부호화 성능을 개선시키도록 하는 것으로, 완전한 일반해를 얻는 것은 용이하지 않다.^[12] 임베디드 웨이브릿 변환 부호화기에 대해, 비트배분의 최적해를 얻기 위한 다양한 기법들이 제안되었다.^[18,19] 즉, Cheng 등은^[18] 임베디드 웨이브릿 부호화기의 동영상 부호화에의 적용을 검토하고, Lagrangian 승수를 사용한 최적 비트배분 기법을 제안하고 있다. [18]에서는 각 프레임간의 R-D의 특성이 지수함수로 모델화 되고, 연속되는 프레임간의 영향이 각 프레임의 분산과 선형 종속적인 관계라는 가정하에 최적해를 구하는 방법을 제시하고 있다. 또한, Yang 등은^[19] 가변블록 움직임 추정과 영역기반의 임베디드 웨이브릿 부호화로 구성된 R-D 최적화 움직임 보상 임베디드 웨이브릿 부호화기의 구조에서 minimax 평가함수를 사용한 프레임 레이트 제어 알고리즘을 제안하고 있다.

본 논문에서는 동영상 부호화를 위한 하나의 새로운 방법으로, DCT 기반의 임베디드 트리구조의 동영상 부호화를 위한 새로운 기법과 연속되는 프레임 단위별로 비트를 최적으로 배분하기 위한 비트율 제어 알고리즘을 제안한다. 먼저, DCT 기반의 임베디드 부호화에서는 기존 표준기법에서 개발된 알고리즘을 최대한 활용할 수 있도록 하기 위해 기존의 MC/DCT 하이브리드 부호화 구조를 그대로 사용한다. 또한, 계층구조의 트리를 구성하기 위해서 한 프레임 전체에 걸쳐 블록별 DCT 계수를 피라미드 구조로 재배열시킨 후, 부호화를 수행한다. 재배열된 DCT 계수를 부호화하기 위해 기존의 SPIHT 알고리즘^[2]을 개선시켜 사용한다. 다음으로, 연속되는 프레임 단위별로 최적 비트배분을 위한 알고리즘에서는 GOP (Group of Picture) 단위로 비트율 제어를 수행하며, 연산 R-D 함수와 프레임간 선형 종속

성을 적용해서 주어진 비트량을 각 프레임에 최적으로 배분되도록 한다. 또한, 프레임 단위별의 R-D 변동을 최소화시키기 위해 각 프레임의 최대 오차가 최소가 되도록 하는 제어방법을 사용한다.

이하, 본 논문의 구성은 다음과 같다. 2절에서는 DCT 계수를 트리구조로 재배열 시켜, 임베디드 제로트리 영상부호화 기법에 의한 부호화에 관해서 기술한다. 기존 웨이브릿 변환 기반의 부호화 알고리즘인 SPIHT 를 수정, DCT 기반의 부호화에 보다 적합하도록 한다. 다음으로 3절에서는 GOP 단위의 각 프레임에 최적의 비트 배분이 이루어지도록 하기 위한 최적 비트배분 알고리즘을 기술한다. 4절에서는 본 제안방법의 객관적인 성능을 평가하고 검증하기 위해 표준 시험 영상을 사용한 모의실험의 결과를 제시하고, 5절에서 결론을 간단히 기술한다.

II. DCT 기반의 트리 구조에 의한 동영상 부호화

1. DCT 계수의 트리 구조로의 재배열

DCT 기반 임베디드 제로트리 영상 부호화의 기본개념은, $N \times N$ 블록의 DCT 계수를 복호영상에 미치는 중요도에 따라 새로운 계수의 집합으로 분류해서 부호화 하는 것이다. 즉, DCT 블록의 계수들을 적당한 트리구조로 취급해서, 기존의 임베디드 제로트리 부호화 기법을 적용시킬 수 있도록 하는 것이다. 이와 같은 개념은 정지영상의 부호화 효율을 개선시키기 위해 Xiong 등에^[3] 의해 처음으로 제안되었다.

이 방법을 간단히 고찰하면 다음과 같다. 8×8 크기의 각 DCT 블록 계수들을 트리 구조로 재배치해서, 기존의 임베디드 제로트리 부호화 알고리즘으로 부호화 한다. 이 때, 프레임 내 각 DCT 블록은 서로 상관성이 없는 독립블록으로 간주하며, 각 DCT 블록마다 DC 값을 별도의 부가정보로서 출력하고 있다. 따라서, DCT 블록간에 존재하는 상관성을 전혀 이용할 수 없으며, 블록마다 출력되는 별도의 부가정보로 인해, 부호화 성능의 개선에 한계가 있다.

본 논문에서는, Xiong 등의^[3] 방법과 개념적으로는 동일하나 구체적인 구현방법에서 조금 다른 알고리즘을 제안해서 사용한다. 입력되는 프레임 전체에 걸쳐, 8×8 크기의 각 DCT 블록의 계수들을 <그림 1>에 나

타낸 것과 같이 4층의 피라미드 구조로 재배열시킨다. 이 구조에서, DCT 계수들 간의 부모-자식의 관계는 웨이브릿 변환에서와 같이 다음으로 정의된다. <그림 1>에 나타난 DCT 계수 i 의 부모는 $\lfloor i/4 \rfloor$ ($1 \leq i \leq 63$)가 되고 계수 j 와 관련한 하위 계층에서의 4 자식의 집합은

$$\{4j, 4j+1, 4j+2, 4j+3\}, 1 \leq j \leq 15 \quad (1)$$

로 정의된다. 이 때, $\lfloor \cdot \rfloor$ 는 정수부분만을 취하는 연산을 나타낸다. 그림 1에서 루트(root)로 취급되는 최상위 계층(숫자 "0"의 성분으로 구성되는 계층)은 DCT 블록의 DC 성분으로 구성된다. 나머지 하위 계층의 계수들은 DCT 블록의 주파수 성분에 대응하는 AC 성분으로 이루어진다. 따라서, 최상위 계층과 그 바로 아래 계층(숫자 "1", "2", "3"으로 구성되는 계층)을 구성하는 전체 계수의 수는 한 프레임 내 8×8 DCT 블록의 수와 같다.

이상의 트리구조는 입력영상 대부분의 에너지가 DC 성분의 최상위 계층과 가장 낮은 AC 성분으로 구성된 그 바로 아래 하위계층에 집중하게 되고, 하위계층으로

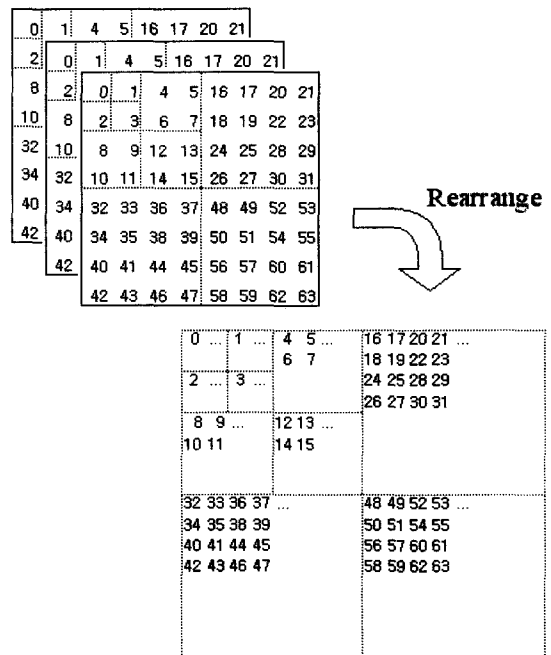


그림 1. DCT 변환 계수의 4층(layer) 피라미드 구조로의 재배열
Fig. 1. The rearrangement of each DCT coefficients into 4 layer pyramidal structure.

내려올수록 급속한 감쇄특성을 갖는다. DCT 계수의 이와 같은 재배열을 통해, DCT 각 블록간의 상관성을 이용할 수 있으며, 급속한 에너지 감쇄특성으로부터 임베디드 제로트리 부호화 기법을 효율적으로 적용할 수 있다.

2. 임베디드 제로트리 영상부호화

2.1절에서 기술한 것과 같이 피라미드 구조로 재배열된 DCT 계수를 양자화하기 위해 SPIHT 알고리즘을^[2] 수정해서 적용한다. 임베디드 제로트리 부호화기에서 비트의 발생을 최소화함과 동시에 임의의 비트량에서도 최상 품질의 복호영상을 얻기 위한 과정을 요약하면 다음과 같다.

- 비트 발생량을 최소화시키기 위해, 중요계수의 집합과 중요하지 않은 계수의 집합을 효율적으로 나타내는 것이 필요하다. 뿐만 아니라, 중요하지 않은 계수들의 부호화를 위한 비트량이 최소값이 되도록 하는 것이 요구된다. 이는 가능한 중요하지 않은 계수들의 그룹을 크게 하는 것과 일치한다.
- 발생되는 비트스트림이 임의의 위치에서 잘려졌을 때에도 최상의 품질을 갖는 복호영상을 얻기 위해서는, 복호영상의 품질에 미치는 영향의 크기에 따라 계수를 분류해서 부호화하는 것이 필요하다. 이는 계층 구조에서 상층에 해당되는 계수일수록 먼저 부호화하고, 남은 비트량에 따라 하위계층의 계수들이 부호화되어야 하는 것을 의미한다.

SPIHT 부호화 기법에서 정의된 것과 같이, 임의의 변환계수의 집합 T 내 계수 $C_{i,j}$ 의 중요도는 다음의 임계값 n 과 T 의 함수로 정의된다.

$$S_n(T) = \begin{cases} 1, & \text{if } \max_{i,j \in T} \{|C_{i,j}|\} \geq 2^n \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

여기서 $|C_{i,j}|$ 는 (i,j) 위치에서의 변환계수의 절대값이다. 재배열된 DCT 계수들에 관한 비트정보를 발생시키기 위해, SPIHT 알고리즘에서와 같이 계수들을 다음과 같은 3가지 종류의 리스트로 분류해서 사용한다.

- List of Insignificant Sets(LIS) : 중요하지 않은 계수들로 구성된 집합의 모음
- List of Insignificant Pixel (LIP) : 중요하지 않은 계수들의 리스트
- List of Significant Pixel (LSP): 중요 계수로 분류된 계수들의 리스트

이상의 임베디드 제로트리 부호화 기법에 관한 개념과 기존 SPIHT 알고리즘을 기반으로 다음과 같은 수정 임베디드 제로트리 부호화 알고리즘을 제안한다.

부호화 알고리즘 :

가. 초기화

- 피라미드 계층구조의 전체 DCT 계수의 집합을 X 라하고, 이를 루트(R)와 그나머지(I)의 두 종류의 집합으로 분리한다 <그림 2(a)> : $R = \text{Roots}, I = X - R$.
- $n = \lceil \log_2(\max_{i,j \in X} \{|C_{i,j}|\}) \rceil$ 를 구한다.
- 좌표 $(i,j) \in R$ 를 LIP에, 집합 I를 LIS에 할당한다.
- LSP를 공집합으로 초기화 한다.

나. 소팅 패스 (Sorting Pass) <그림 2>

- DC성분 (루트층, <그림 2(a)>) LIP내의 모든 (i,j) 의 각 성분에 대한 중요도를 테스트하고 비트를 발생시킨다.
- AC 성분 (계층 2 ~ 계층 0)
 - $S_n(I) = 1$ 이면, 모든 (i,j) 에 대해 집합 I를 부집합 $A(i,j)$ 로 분할한다 <그림 2(b)>.
 - $S_n(A(i,j)) = 1$ 이면, 계층 2를 3개의 성분과 부집합 $B(i,j)$ 로 분리한다 <그림 2(c)>.
 - $S_n(B(i,j)) = 1$ 이면, 계층 1을 3개의 부집합 $D_k(i,j)$ 로 분리한다 <그림 2(d)>.
 - $S_n(D_k(i,j)) = 1$ 이면, $D_k(i,j)$ 를 4개의 성분과 4개의 부집합으로 나누고, 각 성분과 부집합에 대해 중요도 테스트를 수행해서 비트를 발생시킨다. (<그림 2(e)>는 $D_1(i,j)$ 와 $D_3(i,j)$ 가 이 조건을 만족하여 분할된 것을 나타낸 것임).

다. 리파인먼트 패스 (Refinement Pass)

SPIHT 알고리즘과 동일한 것으로, LSP의 각 성분에 대해 보다 정확한 계수값이 되도록 하는 반복과정을 적용한다.

제안기법의 큰 특징은 각 DCT 계수의 중요도에 따라 순서화된 연속 양자화를 취하고, LIS의 사이즈를 가능한 크게 해서 중요하지 않은 계수에서의 비트발생을 최대한 억제하도록 한 점이다. 즉, 재배열된 DCT 계수의 주파수 특성에 따라 LIS에 속하는 집합의 크기를

극대화함으로써 LIP에 속하는 계수의 수를 최소화 되도록 해서 불필요한 비트발생을 최대한 억제하도록 한다. 따라서, 복호영상에 미치는 영향이 큰 계수일수록 먼저 부호화되어 출력될 수 있도록 한다. 또한, 각 DCT 블록의 DC 성분으로 구성된 최상위 계층의 에너지는 다른 계층에 비해 매우 높고, 각 계수들간에 비교적 높은 상관성을 갖고 있다. 따라서, 이 상관성 제거를 위한 별도의 처리과정을 적용할 수 있으나, 알고리즘의 간략화를 위해 본 논문에서는 이를 수행하지 않고 다만 평균값을 구해 이를 하나의 부가정보로서 출력하고 평균치가 제로가 되게 함으로써 불필요한 반복과정을 억제하도록 한다. 마지막으로, 연속 근사 양자화로부터 발생된 비트는 기존의 방법과 동일하게 적응 산술부호화기를 사용해서 엔트로피 부호화 한다.

III. 최적 비트배분 알고리즘

MC-DCT 기반의 동영상 부호화에 임베디드 제로트리 부호화 방법을 사용하는 경우, 연속되는 각 프레임에 최적으로 비트가 배분되도록 하는 것은 복호 동영상의 화질을 개선하는데 있어 매우 중요한 문제이다. 이를 위해 본 논문에서는 하나의 인트라(Intra: I) 프레임과 N-1개의 예측(Predicted: P) 프레임만으로 구성된 N 개의 연속 프레임을 하나의 GOP로 취급하고 이에 대한 최적 비트배분의 문제를 취급한다.

R-D 최적화에 의한 비트배분의 목적은 GOP 내의 각 프레임에서의 가장 큰 부호화 왜곡이 최소가 되도록 하는 것으로, 최적 비트배분을 위한 평가함수는 다음과 같이 주어진다.^[13,15]

$$MIN \{ D_I(R_I) + \sum_{i=2}^N D_i(R_i) \} \quad (3)$$

$$\text{subject to } \sum_{i=1}^N R_i = R_T + \sum_{i=2}^N R_i \leq R_T \quad (4)$$

여기서, R_T 는 GOP내의 전체 목표 비트율, R_i 는 i 번째 프레임의 비트율, R_I 는 Intra 프레임의 비트율, 그리고 $D_i(R_i)$ 는 비트율 R_i 에 대한 i 번째 프레임에 대한 왜곡을 나타내는 함수이다. 인트라(I) 프레임 이후의 모든 프레임은 부호화가 수행된 전 프레임으로부터 예측되어 사용되기 때문에, 각 프레임의 왜곡 $D_i(R_i)$ 는 i 번째 자체 프레임뿐만 아니라 $i-1$ 번째의 전 프레임에도 좌우되는 함수이다.

각 프레임에서의 R-D 함수가 컨벡스(convex) 함수라 하면, 최적화의 문제는 Lagrangian 승수를 사용해서 그 해를 구할 수 있다.^[15] 즉, 상기 식(3)과 식(4)와 같이 구속된(constrained) 최적화의 문제를 구속되지 않은(unconstrained) 최적화의 문제로 변환시키면, R-D 평가함수는 다음의 식 (5)로 주어진다.

$$J(R_1, R_2, \dots, R_N) = \sum_{i=1}^N D_i(R_i) + \lambda \left(\sum_{i=1}^N R_i - R_T \right) \quad (5)$$

각 프레임의 비트율 R_i 는, 양자화 스텝 사이즈를 조정하는 MPEG이나 H.26x와 같은 부호화기^[14~16]와는 달리, 임베디드 제로트리 부호화에서는 하나의 부호화 제어 파라미터로서 사용된다. 또한, 계산량을 줄이기 위해, 간단한 R-D 모델과 프레임간 선형중속성을 가정해

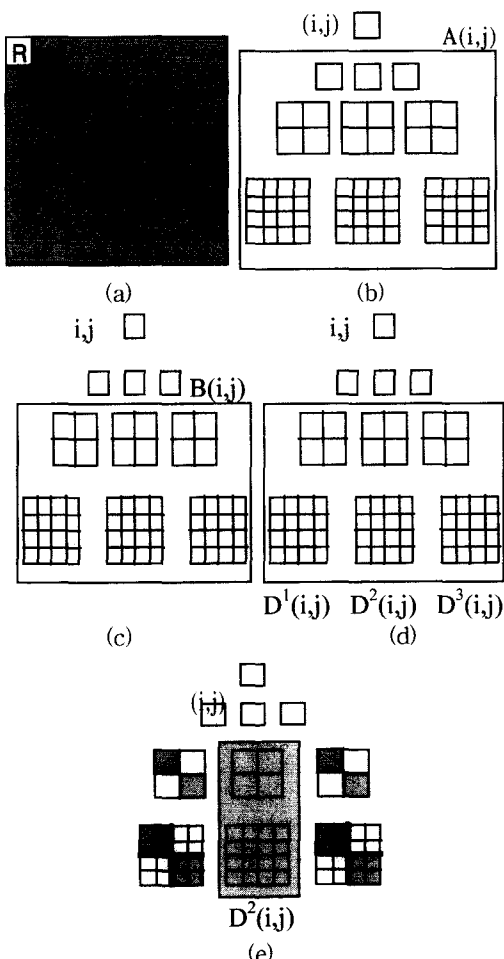


그림 2. 소팅 패스의 과정에서 수행되는 계층간 계수들의 분류와 분할의 방법
Fig. 2. The set partitioning scheme in the procedure of sorting pass.

서 사용한다. Minimax 기법^[10]을 사용해서 각 프레임에 대한 R-D 변동의 범위를 줄여, 각 프레임에서 균일한 화질이 얻어지도록 한다.

단일 프레임에 대한 연산 R-D 함수는 다음과 같이 지수함수로 모델링이 가능하다.

$$D_i(R_i) = \sigma_i^2 2^{-\alpha_i R_i} \quad (6)$$

여기서, σ_i^2 는 부호화하기 전 DCT 계수의 분산에 해당되는 부호화 왜곡을 나타내고, α_i 는 훈련 데이터를 사용해서 추정되는 부호화 파라미터이다. 이와 같은 지수함수의 R-D 평가함수 모델은 비트율 R_i 가 매우 낮은 경우 R-D 곡선의 모델링에 적합하지 않지만, 대부분의 비트율에서 이와 같은 지수함수의 모델은 R-D 특성을 나타내는 근사모델로 충분하다.^[12]

연속되는 P-프레임의 부호화에서, 프레임 f_i 의 예측 오차는 전 프레임 f_{i-1} 이 어떻게 부호화되었는가에 크게 좌우되므로, 각 프레임마다 독립인 R-D 함수를 가정해서 사용하는 것은 최적 비트배분의 목적을 달성할 수 없다. 실제로, 프레임 f_i 내의 각 (R_i, D_i) 점에서, 프레임 f_{i+1} 에서는 다른 R-D 함수가 된다. 이와 같은 문제를 R-D 함수의 프레임간 종속성이라 한다.^[15,17]

이 문제에 대한 일반해를 구하는 것은 간단한 문제가 아니다. Ortega 등은^[12] 각 R-D 함수에서의 데이터 점의 수가 매우 적은 경우에 한해 실행 가능한 해결 방안을 제시하고 있다. 그러나, 이 기법은 각 프레임의 차분에 대한 R-D 함수에 많은 데이터점을 갖는 임베디드 제로트리 부호화에는 적합하지 않다. 따라서, 보다 효율적인 방법으로^[18] 프레임 f_i 가 프레임 f_{i+1} 에 미치는 영향을 프레임 f_i 의 부호화 오차에 대한 프레임 f_{i+1} 의 분산의 선형종속으로 제한하는 것이다. 이는 프레임간 종속성의 문제를 해결하기 위한 하나의 좋은 방법이라 할 수 있다.

본 논문에서 제안하는 최적 비트배분의 기법은 다음의 2종류 모드로 구성된다. 첫 번째 모드는 [18, 19]에서와 같이 하나의 GOP 상에서 수행하고, 인트라(I) 프레임과 인터(P) 프레임을 구분해서 비트배분을 구하는 것이다. 일반적으로 보다 좋은 화질의 인트라(I) 프레임은 연속되는 인터(P) 프레임의 화질을 개선시킬 수 있으므로 인트라 프레임에 보다 많은 비트가 배분되도록 비트량을 조절하는 것이 필요하다. 따라서, 각 GOP에

서 I 프레임은 P 프레임과는 다른 방법으로 부호화 되도록 한다. 즉, I 프레임의 비트배분은 이후 연속되는 모든 P 프레임이 동일한 왜곡과 비트를 갖는 것으로 가정하고, 다음의 식을 사용해서 결정되도록 한다.

$$\begin{aligned} D_{tot} &= \omega D_I + (1-\omega)(N-1)D_P \\ R &= R_I(D_I) + (N-1)R_P(D_P) \leq R_{tot} \end{aligned} \quad (7)$$

여기서 ω 는 가중치로서 $0 < \omega < 1$ 의 값을 갖고, D_P 와 R_P 는 각각 P 프레임에 대한 왜곡과 비트율을 나타낸다.

인트라 프레임의 비트를 할당한 후, 연속되는 P 프레임의 비트배분은 다음과 같은 프레임간 선형 종속성의 함수를 사용해서 수행한다.

$$D_i(R_i) = D_{i,org} + \beta_i D_{i-1}(R_{i-1}) \quad (8)$$

여기서 $D_{i,org}$ 는 움직임 검출과 보상에서 입력영상을 그대로 사용했을 때의 최적화된 최소왜곡을 나타낸다. β_i 는 프레임간 종속성의 정도를 나타내는 파라미터이다. 움직임 보상된 i 번째 P 프레임의 비트율 R_i 는 Lagrangian 승수 λ 를 사용해서 구속되지 않은 평가함수 $J(D)$ 를 식(9)와 같이 고쳐서 할당한다.

$$J(D) = D_{tot} + \lambda \left(\sum_{i=1}^N R_i - R_{tot} \right) \quad (9)$$

이 때 [12]의 방법과 같이, 프레임간 종속성 파라미터 β_i 는 다양한 훈련 데이터를 사용해서 추정한다. 실험 결과, β_i 는 0.6 ~ 0.9의 범위의 값이 되는 것을 확인했다. 또한, β_i 의 값은 움직임이 적을 때는 1에 가까운 값으로 수렴한다. 따라서, β_i 는 움직임추정 블록별로 움직임의 정도에 따라 선형으로 근사화된 값을 사용한다. 각 프레임에서의 최적 비트율은 모든 프레임이 동일한 왜곡이라는 가정 하에 프레임간 선형 종속성과 지수함수의 R-D 함수를 사용해서 구한다. 따라서 식(8)을 식(6)에 대입하고, 이를 식(9)에 적용시킴으로서 GOP 내 연속되는 각 P 프레임에 대한 비트율 R_i 는 다음의 식으로 추정할 수 있다.

$$R_i = \begin{cases} \frac{1}{\alpha_i} \log_2 \frac{\sigma_i^2}{D_i}, & i=1 \\ \frac{1}{\alpha_i} \log_2 \frac{\beta_i D_{i-1} + \sigma_i^2}{D_i}, & i=2,3,\dots,N \end{cases} \quad (10)$$

이상과 같은 알고리즘을 기반으로 최적 비트배분을 위한 과정은 다음의 간단한 반복 2분법에 의해서 수행된다.

1. 다음과 같이 D_{g1} 과 D_{g2} 의 초기치를 선정한다.

$$R(D_{g1}) < R_i < R(D_{g2})$$

2. $D_g = (D_{g1} + D_{g2})/2$ 로 두고,

만약 $R(D_g) < R_i$ 이면 D_{g1} 을 D_g 로 대체한다.

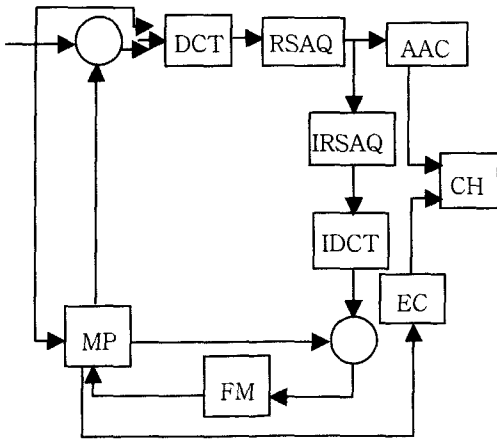
그렇지 않으면 D_{g2} 를 D_g 로 대체한다.

3. $|R(D_g) - R_i| < \delta$ 가 될 때까지 2의 과정을 반복한다.

여기서 δ 는 양수의 임의의 매우 작은 값으로 선정한다.

IV. 모의실험의 결과

<그림 3>은 부호화를 위한 전체 블록도를 나타낸 것이다. 이 그림으로부터 DCT 기반 임베디드 트리구조의 부호화는 양자화 과정이 재배열과 연속근사 양자화로 바뀐 것 외에는 기존 표준 알고리즘과 동일하다는 것을 알 수 있다. 따라서, 알고리즘의 최적화가 이루어진 기존 알고리즘의 하드웨어 구조나 부호화 과정을 그대로 이용할 수 있다.



RSAQ: Rearrangement and Successive Approximation Quantization,

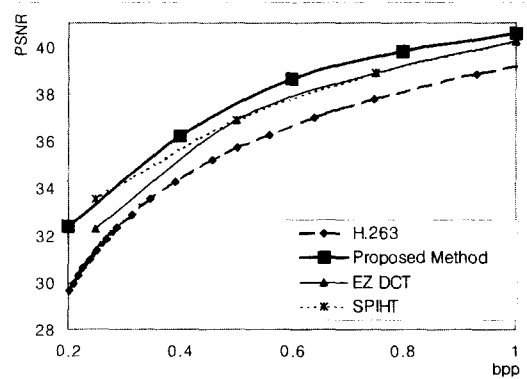
AAC: Adaptive Arithmetic Quantization, IRSAQ: Inverse RSAQ,

MP: Motion Prediction, FM: Frame Memory, EC: Entropy Coding

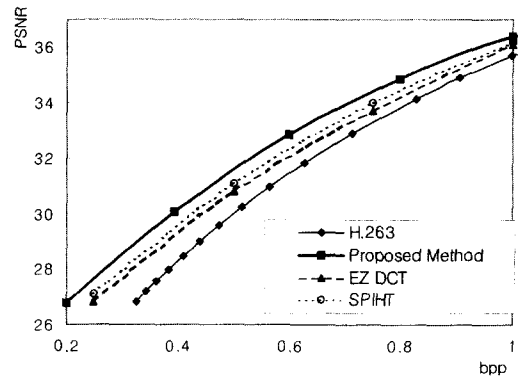
그림 3. MC-DCT 기반 임베디드 제로트리 부호화를 위한 부호화기

Fig. 3. The structure of encoder for the embedded zerotree coding based on the MC-DCT.

먼저, 본 논문에서 제안한 DCT 기반 트리구조에 의한 정지영상 부호화의 부호화 효율을 평가하기 위한 모의실험을 수행했다. 512×512 크기의 시험 영상 "Lena"와 "Barbara"에 대한 모의 실험의 결과를 나타낸 것이 <그림 4(a), (b)>이다. DCT 블록의 사이즈는 8×8의 크기이고, DCT 계수를 4층의 트리 구조로 재배열한 후 부호화 한 것이다. 객관적인 평가를 위해, SPIHT, EZDCT, H.263+ 인트라(I) 부호화의 결과를 함께 나타낸다. 이 때 SPIHT와 EZDCT의 계층분할도 4층으로 제한했다. 이 결과로부터, 제안 방법은 EZDCT 보다 약 1~1.5dB 정도 개선되고, SPIHT 보다 조금 더 좋은 부호화 효율이 얻어지는 것을 확인할 수 있다.



(a)



(b)

그림 4. 정지영상에 대한 부호화의 모의실험 결과: (a) "Lena", (b) "Barbara"

Fig. 4. The simulation results for still image coding with (a) "Lena" and (b) "Barbara".

다음으로 동영상 부호화의 효율을 평가하기 위해, QCIF 사이즈의 표준 테스트 영상 "Carphone"과 "Foreman" 동영상을 사용했다. 제안방법의 객관적인 성능은 H.263 (TM5)과 비교해서 평가했다. 즉, 각 프레임의 PSNR(Peak Signal-to-Noise Ratio)과 비트 사용량을 비교했다. 모의실험에서 사용한 GOP의 크기는 하나의 인트라 프레임과 9개의 예측 프레임으로 구성된 $N=10$ 프레임으로 하고, 목표로 하는 전체 비트량은 256kbps로 가정했다.

<그림 5(a), (b)>는 이들 시퀀스에 대한 PSNR의 결과를 나타낸 것이다. 이 결과, 제안방법은 H.263에 비해 최대 4.5dB 정도 개선되는 것을 확인했다. 뿐만 아니라, 각 프레임에서의 왜곡의 변동이 크게 감소되어 있고, 프레임 스킵(skip)이 일어나지 않는다는 것을 결과로부터 확인할 수 있다.

<그림 6(a), (b)>는 두 시퀀스에 대한 발생된 비트량의 결과를 나타낸 것이다. 이 그림에서 임펄스 함수와 같은 파형은 인트라(I) 프레임의 비트를 나타내는 것이

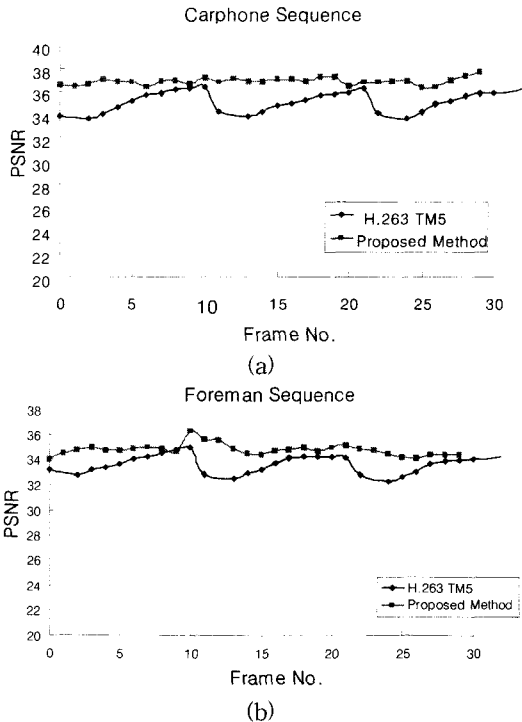


그림 5. 동영상 부호화에 대한 모의실험 결과 (PSNR에 의한 성능 비교) (a) "Carphone" 동영상, (b) "Foreman" 동영상.

Fig. 5. The simulation results for the image sequence coding using (a) "Carphone" and (b) "Foreman" (PSNR performance comparison).

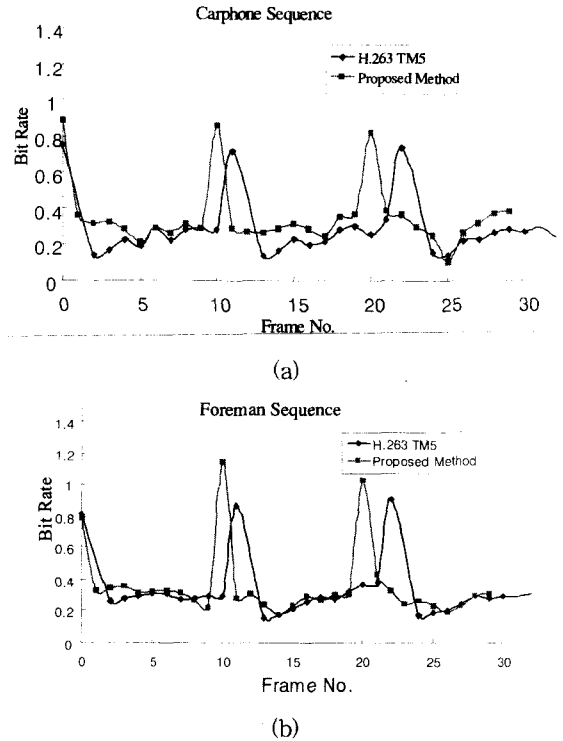


그림 6. 연속 각 프레임에 대한 비트 배분의 결과, (a) "Carphone" 동영상, (b) "Foreman" 동영상

Fig. 6. The bit allocation results with respect to each consecutive frame using (a) "Carphone" and (b) "Foreman" image sequence.

고, P 프레임에 대한 비트량이 H.263에 비해 평탄한 것을 알 수 있다. 또한, 실험에서 제안방법은 각 프레임에 걸쳐 블록 잡음이 줄어들고, 프레임에 걸쳐 보다 일정한 품질의 영상이 얻어지는 것을 확인할 수 있었다.

V. 결 론

본 논문은 동영상 부호화의 성능을 개선하기 위한 방법으로 DCT 기반 임베디드 영상 부호화기와 연속 프레임에의 최적 비트배분을 위한 비트율 제어 알고리즘을 제안하고 모의실험 결과로 이의 성능을 검증했다. DCT 기반의 임베디드 제로트리 부호화 알고리즘은 기존의 웨이브릿 기반의 임베디드 부호화 알고리즘의 개념을 확장해서 적용한 것으로, DCT 기반 트리구조의 특성을 충분히 활용하기 위해 수정된 부호화 방법을 적용했다. 또한, 연속되는 각 프레임에 최적의 비트를 배분하기 위해, 선형 R-D 함수와 프레임간 종속성을

사용해서 GOP 단위의 비트율 제어 알고리즘을 제안했다. 모의실험 결과, 제안 알고리즘의 부호화 성능은 기존 방법에 비해 0~4dB 정도 PSNR 면에서 개선되는 것을 확인하고 있다. 또한, 한정된 비트량에서도 프레임 스kip이 발생하지 않으며, 각 프레임에서 비교적 균일한 품질의 영상이 얻어지는 것을 검증하고 있다.

이상과 같은 제안방법의 부호화 과정은 각 프레임의 DCT 변환계수의 트리구조로의 재배열과 연속근사 양자화를 적용하기 때문에 별도의 프레임 메모리가 요구되는 것 이외에는 기존의 방법에 비해 계산량이 크게 증가하는 것은 없다. 그러나, 비트율 제어를 위한 파라미터의 실시간 추정이나, 각 프레임의 특성을 보다 정확하게 반영한 프레임간 종속성 계수의 추정 등의 분야에서 보다 진전된 연구결과가 요구된다.

참 고 문 헌

- [1] J. M. Shapiro, "Embedded image coding using zerotree of wavelet coefficients," *IEEE Trans. Signal Processing*, Vol. 41, No. 12, pp. 3445~3462, Dec. 1993.
- [2] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuit & System Video Technology*, Vol. 6, No. 3, pp. 243~250, June 1999.
- [3] Z. Xiong, O. Guleryuz, and M. T. Orchard, "A DCT based embedded image coder," *IEEE Trans. Signal Processing Letters*, Vol. 3, pp. 289~290, Nov. 1996.
- [4] S. A. Martucci, I. Sodagar, T. Chiang, and Y. Q. Zhang, "A zerotree wavelet video coder," *IEEE Trans. Circuit & Syst. Video Technol.*, Vol. 7, No. 1, pp. 109~118, Feb. 1997.
- [5] Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space-frequency quantization for wavelet image coding," *IEEE Trans. Image Processing*, Vol. 6, No. 5, pp. 677~693, May 1997.
- [6] J. Y. Tham, S. Ranganath, and A. A. Kassim, "Highly scalable wavelet based video codec for very low bit rate environment," *IEEE J. Selected Areas Commun.*, Vol. 16, No. 1, pp. 12~27, Jan. 1998.
- [7] J. Roh and D. Miller, "A new set partitioning method for wavelet based image coding," In *Proc. IEEE Int. Conf. Image Processing*, pp. 102~106, Oct. 1998.
- [8] J. E. Fowler, "Video coding using perceptually weighted vector zerotrees and adaptive vector quantization," In *Proc. IEEE Int. Conf. Image Processing*, pp. 117~121, Oct. 1998.
- [9] J. Li and S. Lei, "An embedded still image coder with rate-distortion optimization," *IEEE Trans. Image Processing*, Vol. 8, No. 7, pp. 913~923, July 1999.
- [10] A. Islam and W. A. Pearlman, "An embedded and efficient low-complexity hierarchical image coder," In *Proc. Visual Commun and Image Processing (VCIP)*, 1999.
- [11] C. K. Cheong et al., "DCT-based significance tree image sequence coding for progressive transmission," *ICIP 2000*, Vol. II, pp. 859~862, Sept. 2000.
- [12] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," *IEEE Signal Processing Mag.*, pp. 23~50, Nov. 1998.
- [13] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoustic, Speech, and Signal Proc.*, Vol. 36, No. 9, pp. 1445~1453, Sept. 1988.
- [14] G. Keesman et al, "Bit-rate control for MPEG encoders," *Signal Processing: Image Commun.*, pp. 545~560, June 1995.
- [15] K. Ramchandran et al, "Bit allocation for dependent quantization with application to multiresolution and MPEG video coder," *IEEE Trans. On Image Proc.*, Vol. 3, No. 5, pp. 533~545, Sept. 1996.
- [16] W. Ding and B. Liu, "Rate control of MPEG video coding and recording by rate-quantization modeling," *IEEE Trans. Circuit and Sys. Video Tech.*, Vol. 6, No. 1, pp. 12~20, Feb. 1996.

- [17] T. Wiegand et al, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H. 263 standard," *IEEE Trans. Circuit and Sys. Video Tech.*, Vol. 6, No. 2, pp. 182~190, April 1996.
- [18] P. Cheng et al, "Rate control for an embedded wavelet video coder," *IEEE Trans. Circuit and Sys. for Video Technology*, Vol. 7, No. 4, pp. 696~702, Aug. 1997.
- [19] Y. Yang and S. S. Hemami, "Minmax frame rate control using a rate-distortion optimized wavelet coder," *ICIP99*, pp. 551~555, 1999.

 저 자 소 개



鄭 且 根(正會員)

1982년 2월 : 경북대학교 전자공학과 졸업. 1984년 2월 : 서울대학교 대학원 전기공학과 공학석사. 1993년 2월 : 일본 동경대학 전기공학과 공학박사. 1984년 1월~1997년 8월 : LG 종합기술원 책임연구원. 1995년 4월~1996년 3월 : 일본 방송통신기구 초빙연구원. 2002년 1월~2002년 4월 : 동경대학 초빙교수. 1997년 9월~현재 : 호서대학교 정보제어공학과. <주관구분야 : 디지털 신호처리, 디지털 영상통신, 영상처리 및 부호화, Image Sensor 등>