

論文95-32B-11-3

직교 스플라인 웨이브렛 변환을 이용한 TCVQ 설계에 관한 연구

(A Study on TCVQ Using Orthogonal Spline Wavelet)

柳 充 一 *, 金 仁 謙 *, 金 成 萬 *, 鄭 賢 民 *, 朴 圭 泰 *

(Choong Il Ryu, In Kyeom Kim, Sung Man Kim, Hyun Meen Jung,
and Kyu Tae Park)

요 약

본 논문은 스플라인 웨이브렛 변환과 엔트로피 부호화된 TCVQ(Trellis Coded Vector Quantizer)를 사용한 계층적 대역 분할 부호화기에 대한 구현을 그 내용으로 한다. 영상의 계층적 대역 분할을 통해 원영상을 특정 대역과 방향성을 갖는 다양한 해상도의 대역 성분들로 분리하여 각 대역 성분을 TCVQ에 의해 부호화한다. 대역 분할에 사용된 웨이브렛 기저는 직교 스플라인 함수를 이용하여 설계되었다. 본 논문에서는 Wang[13]등이 제시한 집합 분할 방식에 대한 개선 방안을 제시하였으며 512x512 LENA 영상을 부호화한 것과 제안한 부호화기는 WT-VQ(Wavelet Transformed VQ) 부호화기보다 1dB 정도의 성능 향상을 보였으며, 1.002 bpp에서 37.46dB의 PSNR 성능을 나타내었다.

Abstract

In this paper, the method to incorporate TCVQ(Trellis Coded Vector Quantizer) into the encoding of the wavelet transformed(WT) image followed by a variable length coding(VLC) or an entropy coding(EC) is considered. By WT, an original image is separated into 10 bands with various resolutions and directional components. TCVQ used to compress these WT coefficients is a finite state machine that encodes the input source on the basis of the current input and the current state.

Wavelet basis used in this paper is designed by orthogonal spline function. A modified set partitioning algorithm to Wang's is also presented. A simple modification to Wang's algorithm gives a highly time-efficient result.

Proposed WT-TCVQ encoder shows a very competitive result, giving 37.46dB in PSNR at 1.002bpp when encoding 512x512 LENA.

I. 서 론

ATM과 같은 광대역 통신망을 위한 영상 부호화기는 다해상도 영상의 지원과 영상의 점진적 전송, 셀 손

실로 인한 피해의 최소화 등을 고려해야 한다. 기존의 변화 부호화(Transformed Coding) 방식으로는 이런 특성의 고려가 불가능하며, 따라서 다해상도 부호화 기법이 요구된다. 다해상도 부호화는 셀 손실 등으로 인한 전송 에러의 보상, 가입자의 요구에 부합하는 서비스의 제공, 여러 종류의 영상 서비스간의 호환성(compatibility)과 부호화 방법의 통합성, ATM망의

* 正會員, 延世大學校 電子工學科

(Dept. of Elec. Eng., Yonsei Univ.)

接受日字:1994年8月18日, 수정완료일:1995年10月24日

사용료 부과 등의 관점에서 매우 중요한 개념으로 부각되고 있다¹¹⁾.

웨이브렛 변환은 원 영상을 특정 대역과 방향성을 갖는 다양한 해상도의 대역 성분들로 분리하게 되므로 다중 해상도 영상 부호화에 적합하며 대역별 특성을 고려한 부호화와 전송시 대역별 영상의 우선 순위를 고려할 수 있다. 특별히 웨이브렛 변환은 시간과 주파수에 대한 국부성을 가지고 신호를 표현할 수 있어 비정상 과정을 가지는 영상 신호를 해석함에 유리하고, 이를 이용하여 표현된 영상은 인간의 시각 특성과 비슷하여 최근의 영상 처리 분야에서 각광을 받기 시작했다^{17, 9)}.

Trellis Coded Quantizer(TCQ)^{14, 5)}는 현 상태와 입력 신호에 근거하여 양자화를 수행하는 일종의 finite state machine으로 생각할 수 있으며 트렐리스 검색을 통해 평균 왜곡이 가장 작은 경로를 탐색하여 부호화를 수행하는 과정으로 요약된다. 현재 상태에 따라 적합한 부호책을 결정된 후 선택된 부호책을 사용하여 부호화를 수행하는 일반적인 Finite State Quantizer(FSQ)와의 차이점은 FSQ는 입력 샘플을 각각 개별적으로 부호화 하는데 반해, TCQ는 입력 신호열(input sample sequence) 전체에 대하여 평균 왜곡이 최소가 되도록 부호화 하는데 있다¹³⁾. 따라서 입력 신호는 일정한 길이로 구분되어 신호열 단위로 처리되며 대표 신호의 선택 또한 신호열 단위로 결정된다. 이러한 부호화 기법을 delayed decision coding이라 한다. 현재 상태와 입력에 따라 곧바로 대표 신호를 선택할 경우, 현재 상황에서는 최적의 선택일 수 있으나 이러한 선택이 누적될 경우 평균 왜곡이 최소가 된다고 보장할 수 없다. TCQ는 트렐리스 검색을 통해 부호화된 신호의 평균 왜곡이 최소가 되도록 최적의 경로를 결정한다. TCQ는 부호책의 크기가 커질수록 트렐리스의 상태수가 증가하거나 병렬 천이(parallel transition)하는 가지의 수가 증가하게 된다. 이는 트렐리스의 최적 경로 검색의 복잡도가 증가함을 의미한다. 하지만 Ungerboeck 이 Trellis Coded Modulation(TCM)¹⁶⁾을 제안하면서 사용한 집합 분할과 branch labelling의 개념을 적용함으로써, 기존의 트렐리스를 이용한 부호화 방법에 비해 부호화 복잡성이 상당히 줄어들게 되었다. Trellis Coded VQ는 TCQ를 벡터의 경우로 확장하여 일반화시킨 것이다.

본 논문은 TCVC를 이산 웨이브렛 변환을 수행한 영상의 부호화에 적용한 후 여기서 발생한 색인들을 엔트로피 부호화는 문제를 다룬다. 본 논문에서 사용한 웨이브렛 변환을 위한 기저는 직교 스플라인 함수(orthogonal spline function:이하 O-spline)를 이용하여 설계하였다¹²⁻¹⁴⁾. TCVC는 16장의 256계조를 갖는 256x256 영상의 웨이브렛 계수들을 학습 영상으로 하여 설계된다.

본 논문에서는 Wang이 제시한 부호책의 집합 분할 알고리즘에 대한 개선 방안을 제시하였다¹³⁾. 부호책의 크기가 커질 경우 Wang의 알고리즘은 매우 많은 시간을 필요로 한다. 집합 분할은 초기 LBG 부호책을 분할하는 단계에만 사용되는 것이 아니라 TCVC 설계시 empty cell 처리 후 반복해서 사용되기 때문에 집합 분할에 많은 시간이 소요될 경우 TCVC의 설계에 큰 장애가 된다. 본 논문에서는 Wang의 알고리즘에 대한 간단한 수정만으로 집합 분할에 소요되는 시간을 비약적으로 단축시켰다.

II. 웨이브렛 변환

웨이브렛 변환(wavelet transform-이하 WT)은 시간과 주파수에 대해 국부성(locality)을 가지고 신호를 표현할 수 있어 비정상 과정(non-stationary)을 가지는 영상 신호를 해석함에 유리하고, 이를 이용하여 표현된 영상은 인간의 시각 특성과 비슷하여 최근 영상 처리 분야에서 각광을 받기 시작하였다. 지금까지 사용된 이와 유사한 영상 처리 방법은 영상을 피라미드(pyramid) 구조로 표현하는 다해상도 방법 및 대역 분할 부호화(subband coding)방식이 있는데 WT는 이를 하나로 통합하는 이론이라 생각할 수 있다.

이산 WT에서의 기저 함수들은 식 1과 같이 단 하나의 원형 웨이브렛을 천이(translation), 확장(dilation)/수축(contraction)함에 의해서 만들어진다.

$$h_{mn}(t) = a_n^{-m/2} h(a_n^{-m}t - nb_n) \quad m, n \in \mathbb{Z}, a_n > 1, b_n \neq 0 \quad (1)$$

이러한 기저들을 이용한 이산 웨이브렛 변환은 다음과 같다.

$$X_{II}(m, n) = a_n^{-m/2} \int_{-\infty}^{\infty} h(a_n^{-m}t - nb_n)x(t) dt \quad (2)$$

m, n 값들을 어떻게 주느냐에 따라서 확장/수축과 천이된 기저들이 만들어진다.

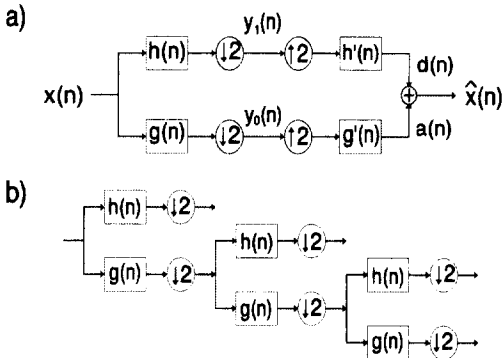
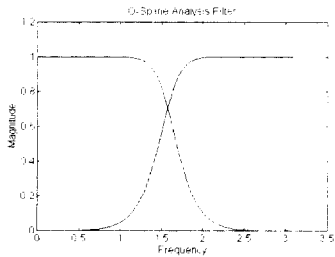


그림 1. 대역 분할 부호화 체계

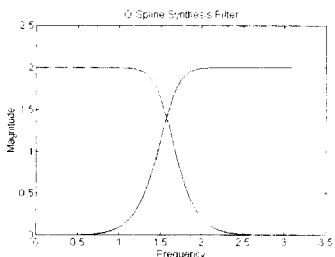
- (a) 영상정보 분해 과정
- (b) 영상 정보 합성 과정

Fig. 1. The coding system of band splitting.

- (a) The procedure of image decomposition.
- (b) The procedure of image synthesis.



(a) O-spline 분해 필터
(a) O-spline analysis filter



(b) O-spline 합성 필터
(b) O-spline synthesis filter

그림 2. O-spline 분해/합성 필터

Fig. 2. O-spline analysis/synthesis filter.

이러한 웨이브렛 변환은 일련의 계층 대역 필터로 필터링하는 것이며, 이에 Nyquist frequency에 따라서 샘플링하는 것으로 해석 가능하다. 따라서 웨이브렛 변환을 다해상도 부호화에 적용할 수가 있는데, 이러한 경우의 대역 분할 구조가 그림 1에 나타나 있다. 그림 1(a)는 원신호를 2개의 대역, $h(n)$ 에 의한 고주파 대역과 $g(n)$ 에 의한 저주파 대역으로 분할하는 구조를 나타내며, 그림1(b)는 저주파 대역을 다시 고주파와 저주파 성분으로 분할하는 과정을 반복하는 octave band splitting 구조를 나타낸다.

본 논문에서 사용된 WT 필터 계수는 B-spline에 기초한 O-spline을 이용하여 얻었다^[10-12]. 그림 2는 O-spline을 이용한 분해 필터 및 합성 필터를 나타낸다.

III. 트렐리스 부호화 벡터 양자화기

VQ(Vector Quantization)는 각 샘플을 독립적으로 부호화 하는 대신 일련의 샘플들을 블록으로 나눈 입력 벡터를 대표 벡터에 사상시키는 방법으로, 대표 벡터들의 집합을 부호책(codebook) 이라 한다. 부호책의 각 벡터들은 색인(index)되어 있어서 부호화될 입력 벡터와의 왜곡을 최소로 하는 대표 벡터의 색인의 전송된다. 신호의 복원을 위해서 복호화단에서도 부호화단과 동일한 부호책을 가지고 있어야 한다. 벡터 양자화기의 설계는 일종의 clustering 알고리즘과 유사한데, 다양한 설계 방법이 연구되어 왔다. 본 논문에서는 그 중에서 가장 일반적으로 사용되고 있는 LBG 알고리즘을 이용하여 TCVCQ를 설계하기 위한 기본 부호책을 제작한다. LBG 알고리즘은 비균일 스킨라 양자화기의 설계를 위한 Lloyd 알고리즘^[15]을 VQ의 부호책 설계를 위해 일반화시킨 것으로 주어진 신호에 대해 국부적으로 최적(locally optimal)인 부호책을 작성할 수 있다.

TCVCQ는 트렐리스 부호화 양자화기(TCQ : Trellis Coded Quantizer)를 입력 소스가 벡터인 경우로 일반화한 것이다. TCVCQ 는 현재 상태에 근거하여 사용할 부호책을 결정하고, 또한 현재 입력에 근거하여 출력 대표 벡터를 주어진 부호책에서 결정하는 finite state machine으로 생각할 수 있다. TCVCQ는 트렐리스의 검색 깊이(search depth)만큼의 부호화 지연(coding delay)를 두어 다양한 경로를 탐색한 후 최

소의 평균 왜곡을 가지는 경로를 선택하며 이러한 최적의 경로를 탐색하기 위해서 Viterbi 알고리즘을 사용한다. 이러한 TCVQ의 성능 한계는 알파벳 제한 비트율 왜곡 이론(alphabet constrained rate distortion theory)에 의하여 주어진다^[17].

비트율 R , 벡터의 크기를 n , 탐색 깊이를 L 이라 하면 크기 2^{nR+K} 의 부호책이 주어질 때 이를 Ungerboeck의 집합 분할 방법에 의해 K 개의 부집합으로 분할하고 이렇게 분할된 부집합을 트렐리스의 각 가지에 할당한다. 따라서 TCVQ에서는 사실상 일반 VQ에 비하여 동일한 비트율에서 더 큰 크기의 부호책을 사용하게 되며 이에 따른 성능 향상을 기대할 수 있다. 부호화 과정은 다음과 같다. 길이 L 의 입력 벡터열 X' 이 주어질 때 가지에 할당된 k 번째 부집합내에서 X' 의 i 번째 벡터 X'_i 와 가장 근사한 대표 벡터 $\hat{X}_{k,i}$ 를 찾고 그 때의 왜곡 $d(X'_i, \hat{X}_{k,i})$ 를 i 번째 트렐리스 단의 branch metric으로 삼는다. 이렇게 주어진 branch metric을 이용하여 최소의 평균 왜곡을 갖는 경로를 Viterbi 알고리즘을 이용하여 검색하고 이렇게 결정된 L 개의 $\hat{X}_{k,i}$ 의 색인 i 와 그때의 가지 번호를 전송하고 다음의 L 개의 벡터에 대한 부호화를 반복한다.

수신단에서는 수신된 L 개의 채널 부호어로부터 가지 번호에 의해 경로를 결정하고 색인 i 에 의해 그 가지에 할당된 부집합 내에서 대표 벡터를 결정하여 복호화를 수행한다. 본 논문에서는 $R = 1$, 검색 깊이 $L = 512$ 인 TCVQ를 설계한다. TCVQ 설계 과정은 크게 집합 분할(set partitioning), branch labelling, 최적화 과정으로 요약된다.

TCM에서 사용된 변조 신호의 집합 분할은, 분할이 진행됨에 따라 각 집합 내의 채널 부호들간의 최소 유클리디안 거리가 최대가 되도록 집합을 구성해 나간다. TCVQ의 초기 LBG VQ 부호책 $\{c_i | i = 0 \dots M-1\}$ 가 주어졌을 때의 집합 분할 알고리즘은 다음과 같다.

- STEP 1) d_i 의 단조 증가 방향으로 정렬된 테이블 $\{c_{i1}, c_{i2}, d_i = \|c_{i1} - c_{i2}\| | i = 0, \dots, M(M-1)/2\}$ 을 구성한다.
- STEP 2) $c_{01} \in D_0, c_{02} \in D_1$ 로 할당하고 0번째 항은 제거한다.

STEP 3) 테이블을 처음부터 순차적으로 검색하여 j 보다 작은 모든 i 에 대해서는 어느 부호어도 부집합에 속해 있지 않고 $(c_{i1} \in D_0, c_{i2} \in D_1) \text{ OR } (c_{i2} \in D_0, c_{i1} \in D_1)$ 를 만족하는 j 를 찾는다.

STEP 4) 만일 c_{i1}, c_{i2} 가 모두 부집합에 할당되어 있다면 j 번째 항은 테이블에서 제거하고 STEP 6으로 간다.

STEP 5) 만일 어느 한 부호어가 부집합에 할당되어 있다면 다른 부호어는 반대편 부집합에 할당한다.

STEP 6) 부집합 D_0 (또는 D_1)의 크기가 $M/2$ 에 달하면 할당되지 않은 나머지 부호어 모두 D_1 (또는 D_0)에 할당한다. 그렇지 않을 경우 STEP 3으로 간다.

전체 부호책을 원하는 개수의 부집합으로 분할하기 위해서는 이진 트리의 경우와 마찬가지로 분할된 부집합에 대해 다시 동일한 알고리즘을 적용하여 분할을 반복하면 된다. 집합 분할 작업은 상당한 시간을 요할 수도 있는데 예를 들어 부호책의 크기가 1024인 경우, 이들 부호어간에 산출되는 거리의 경우의 수는 $1024 \times (1024-1)/2 = 523,776$ 이 된다. 이 경우 테이블을 소팅하는 과정보다는 알고리즘을 진행해 나가면서 테이블의 항을 제거하고 테이블을 갱신하는 과정에서 많은 시간을 요하게 된다. Wang과 Moayeri^[13]의 경우는 STEP 4에서 $i \geq j$ 인 모든 i 번째 항에 대하여 $c_{i1} \in c_{i-1}, c_{i2} \in c_{2-1}, d_i \in d_{i-1}$ 를 수행함으로써 테이블을 갱신한다. 이같은 방식으로는 1024개의 부호어를 4개의 부집합으로 분할하는데 SPARC2 워크스테이션에서 약 24시간이 걸린다. 하지만 알고리즘을 진행함에 따라 c_{i1}, c_{i2} 두 부호어 모두 부집합에 할당되어 있는 항들이 테이블의 앞부분에 몰리게 됨에 착안하여 STEP 4의 테이블 갱신 과정을 $i \leq j$ 인 모든 i 번째 항에 대하여 $c_{i1} \in c_{i-1}, c_{i2} \in c_{2-1}, d_i \in d_{i-1}$ 작업을 행함으로써 테이블의 갱신 시간을 비약적으로 단축시킬 수 있다. 실제로 위의 예와 동일한 집합 분할을 할 때 약 2~3분의 시간이면 가능하다.

TCVQ 설계의 다음 과정으로 집합 분할 과정에서 주어진 부집합들을 트렐리스 가지에 할당하는 branch labelling 과정이 필요하다. 본 논문에서는 4 상태 트렐리스를 사용하여 그림3과 같이 4개의 부집합을 가지에 할당하였다. 이러한 방식으로 부집합을 할당하였을

때 이것이 적절한 선택일 때는 트렐리스의 상태에서 나오는 가지에 할당된 부집합들의 합집합은 입력 벡터의 dynamic range를 포괄하는 부호책이 될 수 있으며, 따라서 입력 벡터와 전혀 동떨어진 대표 벡터가 선택될 가능성은 배제된다는 추정에 의해서 정당하다 할 수 있다.

지금까지 설명한 바와 같이 LBG 알고리즘에 의해 주어진 초기 부호책에 대하여 집합 분할, branch labelling이 이루어진 후 입력 벡터에 대한 부호화가 진행된다. 그런데 이 시점에서 주어진 LBG 부호책의 대표 벡터는 비정형화된 전역 탐색 VQ에 기반을 둔 것이기 때문에 학습을 통해 트렐리스 구조에 적합한 새로운 centroid를 계산함으로써 부호화의 최적화가 가능하다. 이러한 최적화 과정은 새로운 대표 벡터에 의한 전체 왜곡의 개선이 일정 임계치를 넘지 못할 때까지 반복된다.

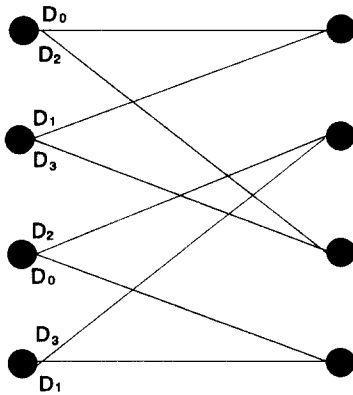


그림 3. TCVQ 설계에 사용한 4 상태 트렐리스도
Fig. 3. The diagram of 4-state trellis using TCVQ.

IV. WT-TCVQ 영상 부호화기의 설계

영상의 대역 분할 부호화의 장점은 해당 대역의 통계적 특성에 최적인 부호화, 그리고 인간의 시각 특성에 따른 적응적인 부호화를 할 수 있다는 것이다. 또한 특정 대역의 양자화 오차 및 손실은 해당 대역에 한정되며, 잡음 스펙트럼의 조정(noise spectrum shaping)이 가능하다^[16]. 이러한 경험적인 요인들 이외에, 정보 이론 관점에서의 분석을 통한 최근의 연구에 의하면^[19-21], 대역 분할을 거친 Gaussian 소스의 대

역 성분들은 원신호에 비하여 유한차 엔트로피와 엔트로피 레이트 사이의 차가 감소한다. 이는 간혹된 대역 성분은 원소스에 비하여 평균적으로 더 작은 메모리를 갖게 되며, Gaussian 과정의 경우 더 평탄한 스펙트럼을 갖게 됨을 의미한다. 대역 분할 부호화의 부호화 이득은 이러한 엔트로피의 감소에 기인한다.

본 논문에서는 그림 4와 같이 주어진 영상을 10개의 대역으로 분할한다. 이러한 각 대역별 영상은 상이한 통계적 성질을 가지고 있으므로, 정보의 손실과 정보량을 최소화하기 위하여, 해당 대역에 최적인 양자화기를 설계하여야 한다. 계층적 대역 분할은 원영상을 특정 대역과 방향성을 갖는 대역 성분들로 분리하므로, 각 대역별로 가지는 16장의 대역 영상들을 학습 영상으로 사용하여 독립적인 TCVQ 부호책들을 설계하였다. TCVQ의 검색 길이는 $L = 512$ 이고, 그림 3과 같이 4개의 상태를 갖는 트렐리스를 사용하였다. 일반적으로 상태수가 많은 트렐리스를 사용할수록 성능은 우수해지나 부호화의 복잡도가 증가하게 되는 단점이 있다.

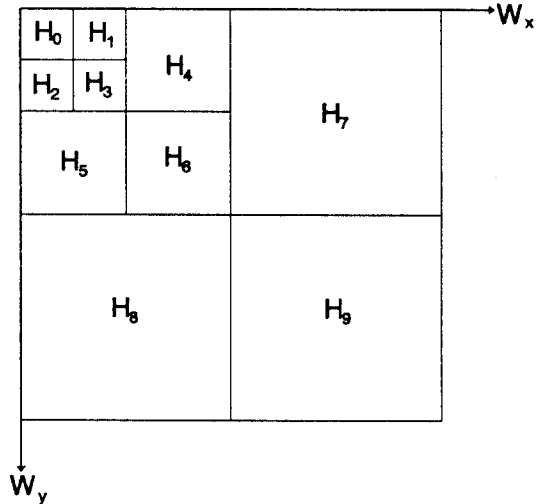


그림 4. WT을 이용한 대역 분할
Fig. 4. Band splitting using WT.

다해상도 부호책의 각 대역별 벡터의 크기와 학습 샘플, 학습 벡터의 개수가 표 1에 나타나 있다. 웨이블릿 분해된 영상의 기본 성질은 대역별로 수직, 수평, 대각선 성분으로 나타나게 된다. 따라서 각 대역별 부호책의 블럭 크기는 이런 특성을 수용하도록 설계하였

다. 저해상도 대역 성분들은 상대적으로 높은 비트율로 부호화해야 하므로 블록 크기를 작게 하여 부호책의 크기를 적당히 제한할 필요가 있다.

표 1. 다해상도 부호책

Table 1. The codebook of each dimensional image.

대역	블록 크기	학습 샘플	학습 벡터
H0	1×1	32×32×16	32×32×16
H1	1×2	32×32×16	32×32×16/2
H2	2×1	32×32×16	32×32×16/2
H3	1×2	32×32×16	32×32×16/2
H4	1×2	64×64×16	64×64×16/2
H5	2×1	64×64×16	64×64×16/2
H6	2×1	64×64×16	64×64×16/2
H7	2×2	128×128×16	128×128×16/4
H8	2×2	128×128×16	128×128×16/4
H9	2×2	128×128×16	128×128×16/4

가장 저주파 대역 성분인 H0의 대역은 가장 중요한 저주파 성분이 들어 있으므로 이 부분은 블록 크기 1×1, 즉 TCQ를 사용하였고, 대역 H1은 주로 수직 방향 성분의 경계들이 살아 있으므로 이를 잘 표현하기 위해서 1×2 크기로 잡아 주었다. 마찬가지로 대역 H2도 수평 방향 경계 성분들이 살아 있으므로 이를 잘 표현하기 위해서 2×1로 설계하였다. H7~H9의 고주파 부분의 대역 성분은 그 중요도가 상대적으로 덜하므로^[12] 부호화율을 높이는 방향으로 벡터의 크기를 잡는다. 따라서 이 부분은 2×2로 설계하였다.

TCVQ의 출력 채널 부호어는 트렐리스의 가지 번호를 나타내는 비트 s_0 와 그 가지에 할당된 부집합 내의 부호어 색인을 나타내는 비트 s_1 으로 구성된다. 이들을 엔트로피 부호화하는 방법에는 가지의 발생 확률은 동일하다고 가정하여 s_0 는 엔트로피 부호화 없이 전송하고 s_1 은 엔트로피 부호화를 거쳐 전송하는 방법(TCVQ1)과 하나의 상태에서 나오는 가지들에 할당된 부집합들을 super set으로 묶어 이 super set내의 색인을 엔트로피 부호화하여 전송하는 방법(TCVQ2)이 있다. 전자에서 있어서의 출력 엔트로피는 $P(n|m)$ 을 m 번째 부집합이 주어졌을 때 그 부집합 내의 n 번째 부호어가 선택될 확률이라 할 때, 다음과 같이 주어

진다.

$$H(n|m) = 1 - \sum_{m,n} \log_2 P(n|m) \quad (3)$$

식 3은 항상 1보다 크므로 이보다 더 낮은 비트율의 부호화기가 요구될 경우 벡터의 경우로의 확장이 필연적이다. 하지만 이러한 경우 부호화기의 복잡도가 상대적으로 증가하게 되며 부호화 성능 또한 감소된다.

Marcellin은 하나의 상태에서 나오는 가지에 할당된 부집합들을 묶어 super set이라 명명하고 이들의 색인만을 채널 부호어로 삼았다^[14]. 각각의 부집합은 공통된 부호어를 가지고 있지 않으므로 super set 내의 부호어로부터 출력 가지 번호와 차기 상태를 결정할 수 있다. 따라서 가지 정보와 색인 정보를 총괄하여 하나의 super set 색인으로 표현할 수가 있다. 이와 같은 엔트로피 부호화 기법을 이용할 경우의 출력 엔트로피는 다음과 같다.

$$H(k|s) = - \sum_{k,s} P(k,s) \log_2 P(k|s) \quad (4)$$

윗 식에서 $P(k|s)$ 는 super set s 가 주어졌을 때 super set내의 k 번째 부호어가 선택될 확률을 의미한다.

웨이브렛 변환을 거친 각 대역 성분들의 부호화 방법 및 비트 할당을 해당 대역의 통계적 특성에 따라 최적화 함으로써 부호화 효율을 극대화 할 수 있다. 각 대역 성분들의 최적 부호화를 위하여 앞에서 살펴본 바와 같이 대역 성분별로 독립적인 부호책을 설계하였으며, 다음과 같은 방법으로 최적 비트 할당이 이루어진다.

$D_k(R_k)$ 와 R_k 를 각각 k 번째 대역의 왜곡 및 비트율이라 한다면, 비트 할당의 목적은 복원 영상의 왜곡을 최소화 할 수 있도록 주어진 비트수 R_{max} 를 각 대역 성분의 부호화를 위해 적절히 배분하는 것이므로 총 비트수 R 이 주어진 비트수 R_{max} 를 넘지 않는 범위 내에서 전체 왜곡 D 가 최소가 되도록 R_k 를 결정하는 것으로 요약된다. 즉 다음 식을 만족하는 R_k 의 해를 구하는 것이다.

$$\min R_k \left\{ D = \sum_k D_k(R_k) \mid \frac{1}{N} \sum_k N_k R_k \leq R_{max} \right\} \quad (5)$$

여기서 N 은 전체 영상의 화소수이며 N_k 는 k 번째 대역 영상의 화소수를 의미한다.

각 대역 성분에 대한 distortion rate 함수

Dk(Rk)가 연속이고 convex하다고 가정하면 식 (5)를 다음과 같은 Lagrange multiplier 방법을 통하여 표현할 수 있다.

$$\min_{R_k} J = \min_{R_k} \left\{ \sum_k D_k(R_k) + \lambda \sum_k N_k R_k \right\} \quad (6)$$

J는 총 비용(overall cost) 함수이며 λ는 Lagrange multiplier이다. 주어진 총 비트율에서 최소의 총 왜곡은 식 6의 1차 편미분치를 0으로 하여 얻어진다. 이러한 미분방정식의 해를 구하여 (Rk,Dk)점을 찾는 알고리즘은 많은 사람들에 의해 제안되었는데, 수식적인 표현이 불가능한 경우를 포함하여 대역 분할 부호화에 있어서 어떠한 제한 조건도 가하지 않은 비트 할당 알고리즘으로서 Westerink등이 제안한 최적비트할당 알고리즘^[21]과 Riskin이 제안한 G-BFOS 알고리즘^[20] 등이 있다. 본 논문에서는 Westerink의 최적비트할당 알고리즘을 사용하였다.

V. 실험 결과 및 결과 고찰

본 논문에서 제안한 알고리즘의 성능을 분석하기 위하여 다양한 통계적 성질을 갖는 256x256 해상도 및 256의 계조도를 갖는 16장의 학습 영상을 사용하여 각 대역 성분에 대한 부호책을 설계하였다. 웨이브렛 변환은 원 영상을 특정 대역과 반향성을 갖는 대역 성분들로 분리하므로 각 대역 성분들을 학습 영상으로 사용하여 설계된 부호책들은 해당 대역의 통계적 특성에 잘 부합된다. O-스플라인 웨이브렛 변환을 거친 학습 영상들은 대역 성분별로 모아지고 이를 이용하여 TCVC부호책을 설계하였다.

그림 5은 이러한 10개의 대역의 부호책으로부터 참고문헌^[21]의 최적 비트 할당 알고리즘을 이용하여, 엔트로피 부호화된 TCVC 시스템의 operational DRF(Distortion-Rate Function)이다. 그림에서 VQ라고 명명한 곡선은 각 대역에 대하여 TCVC부호책을 사용하는 대신에 표 1에서 제시된 바에 의한 학습 벡터에 따라 일반 VQ를 적용했을 때의 성능 곡선이다. 그림 5에서 보듯이 super set의 개념을 적용하여 엔트로피 부호화를 수행한 TCVC2의 성능이 가장 우수함을 볼 수 있으나 비트율이 높아질수록 성능상의 이득이 점차 감소하여 1 bpp 이상에서는 TCVC1과 차이가 없음을 알 수 있다. 전 비트율에 걸쳐

TCVC는 VQ를 적용한 웨이브렛 부호화에 비하여 성능이 우수하여 PSNR로 환산시 최고 1dB까지 성능 향상을 얻을 수 있음을 알 수 있다.

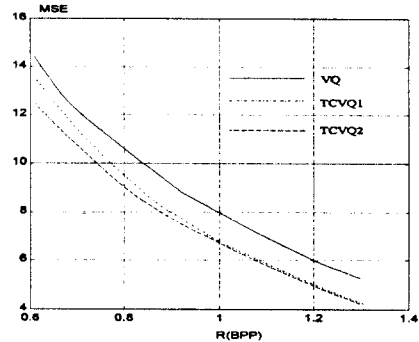


그림 5. WT-TCVC 시스템의 성능 곡선
Fig. 5. The Curve of WT-TCVC system's performance.

표 2. TCVC2에 의한 부호화 결과
Table 2. The coded results by TCVC2 method.

대역	LENA		PEPPER	
	bpp	MSE	bpp	MSE
H0	6.907	0.219	6.943	0.253
H1	2.881	1.197	2.838	0.952
H2	2.526	0.562	2.695	0.857
H3	2.769	0.397	2.706	0.225
H4	2.071	1.279	2.000	0.936
H5	1.807	0.754	1.919	0.981
H6	1.955	0.391	1.888	0.289
H7	0.683	3.316	0.688	3.569
H8	0.608	2.022	0.699	3.366
H9	0.318	1.558	0.442	4.377
총	1.002	37.457	1.057	36.143

제안한 영상 부호화기의 성능을 평가하기 위하여 16장의 학습 영상에서 제외된 256x256 LENA 영상과 학습 영상에 포함된 256x256 PEPPER 영상에 대하여 부호화를 수행하였다. LENA영상을 학습영상에 포함시킨 경우의 실험 결과는 1.057bpp에서 36.44dB를 얻어 그 결과가 LENA를 제외한 부호책의 결과와 거의 차이가 나지 않았다. 부호화는 먼저 주어진 영상을 그림 4와 같이 웨이브렛 변환으로 3단계, 10개 대역으로

로 분할한 뒤 각 대역별로 독립적으로 설계된 부호책을 이용하여 부호화하였다. 그 결과로 LENA 영상에 대해서는 1.078 bpp에서 36.38 dB를, PEPPER 영상에 대해서는 1.05 bpp에서 35.11 dB의 PSNR 결과를 얻을 수 있었다.

또한, 제안된 부호화기가 학습 영상에 사용된 256x256 영상외의 일반적 영상의 부호화에도 효과적으로 사용될 수 있도록 설계되었는가를 확인하기 위하여 학습 영상에 포함되지 않은 512x512 LENA와 학습 영상에 포함되기는 했지만 그 크기가 512x512로서 부호화기 설계에 사용된 것과는 다른 통계적 특성을 가지는 PEPPER에 대하여 부호화를 수행하였는데, 표 2는 비트율이 약 1 bpp일 때의 성능을 측정하기 위하여 LENA와 PEPPER영상에 대하여 IV절에서 설명한 최적 비트 할당에 의해 구해진 비트율로 각 대역을 부호화 했을 때의 실제 비트율과 왜곡을 나타낸 것이다. 웨이브렛의 대부분의 정보는 H0대역에 집중되어 있고 이 영역의 왜곡은 전체 영상의 품질에 큰 영향을 미치는 반면, 비트율의 증가는 적은 화소수로 인하여 타 대역에 비해 전체 비트율 증가에 미치는 영향이 상대적으로 작으므로 높은 비트율로 부호화를 행하게 된다. 또한 diagonal 정보는 적은 비트율로 부호화되어도 전체 영상에 큰 영향을 미치지 않는다. 이러한 대역 특성에 따른 부호화 결과가 표 2에 잘 나타나 있다.

이상의 같은 부호화 결과로서 512x512 PEPPER 영상에 대하여 1.06 bpp에서 약 36.14 dB의 PSNR을 얻을 수 있었으며, 512x512 LENA 영상에 대해서는 1.002 bpp에서 37.46 dB의 PSNR을 얻었다. 이러한 결과는 부호화기가 다양한 크기와 다양한 통계적 특성을 가지는 영상의 부호화에 효과적으로 사용될 수 있도록 설계되었음을 의미한다.

표 3은 LENA영상에 대하여, 제안된 부호화기의 성능을 여러 가지 WT-VQ 부호화기의 성능과 비교한 표이다. 두번째 열과 세번째 열은 Antonini가 제안한 웨이브렛 필터^[8]를 변환단에 사용한 것이다. Antonini의 QMF필터의 설계 조건과 이산 웨이브렛 변환의 설계 조건을 결합시켜 웨이브렛 기저를 설계하였다. VQ는 일반적인 LBG VQ에 의한 부호화기를 뜻하며 ECVQ(Entropy Constrained VQ)는 VQ를 비트율-왜곡 이론(rate-distortion theory)의 관점 하에서 일반화시킨 것으로 VQ의 설계시 최소 왜곡만을 중시한 것이 아니라, 부호화기의 출력 엔트로피까지 고려하

여 설계한 VQ를 말한다^[19]. 표 3에서 처음 3가지 부호화기는 256 계조도의 512x512 영상 16장을 가지고 설계된 부호화기들이다. 제안한 부호화기는 약 1 bpp의 비트율에서 타 부호화기에 비해 0.84 dB에서 2.25 dB 정도 더 높은 PSNR을 나타내었다. O-스플라인을 WT와 ECVQ를 결합한 부호화기^[12]와 비교해 보면 0.011 bpp가 더 낮은 비트율에서 오히려 0.84 dB 더 높은 결과를 얻을 수 있음을 알 수 있다. 여기서 제안한 부호화기에 성능이 가장 근접한 O-스플라인 + ECVQ는 설계시부터 512x512 영상을 이용한 것임을 감안할 때 TCVC가 여타의 VQ에 비하여 우수한 성능을 보임을 알 수 있다.

복원된 영상은 일반적으로 VQ의 단점으로 지적되는 블록 효과가 나타나지 않았으며 경계 성분도 선명하게 표현된다. 이는 VQ와 웨이브렛을 이용한 계층적 대역 분할 부호화가 효과적으로 결합되었음을 의미한다.

표 3. LENA 영상에 대한 성능 비교

Table 3. The performance comparison of LENA image.

비교 방법	O-spline + VQ	O-spline + ECVQ	O-spline + TCVC2
bpp	1.009	1.01311	1.002
PSNR [dB]	34.23	36.6201	37.46

VI. 결 론

본 논문에서는 ATM과 같은 광대역 통신망의 특성이 여러 가지 해상도의 영상 서비스 제공, 점진적 전송 등의 특성을 만족시키기 위하여 O-스플라인 웨이브렛 변환과 TCVC를 이용한 영상의 계층적 대역 분할 부호화기를 제안하였다. 제안한 부호화기는 다양한 통계적 특성을 가지는 16장의 256x256 크기의 256 계조도 영상을 학습 영상으로 사용하여 다해상도 부호책을 설계하였다. O-스플라인 웨이브렛 변환을 이용하여 계층적 대역 분할된 학습 영상들은 대역 성분별로 모아서 TCVC 알고리즘에 의하여 각 대역 성분별로 독립적인 부호책을 설계하였다. 학습 영상에 포함된 256x256 PEPPER의 부호화 결과는 1.05 bpp에서 35.11 dB의 PSNR 결과를 보였으며 학습 영상에 제외한 256x256 LENA 영상의 경우 1.078 bpp에서

36.38 dB의 PSNR을 얻을 수 있었다. 제안한 부호화기가 다양한 크기와 다양한 통계적 특성을 가지는 영상에 효과적으로 사용될 수 있도록 설계되었는가를 확인하기 위하여 512x512 PEPPER와 512x512 LENA 영상에 대하여 부호화를 수행한 결과 PEPPER는 1.06 bpp에서 36.14 dB를, LENA는 1.002 bpp에서 37.46 dB의 우수한 결과를 얻을 수 있었다.

TCVQ는 그 복잡도가 일반 VQ에 비하여 그리 높지 않으면서도 높은 압축 효율을 나타내며, 웨이브렛 변환과 결합시 높은 부호화 효율 외에 영상의 다해상도 표현과 점진적 전송이 용이한 장점이 있으므로 DCT 부호화의 뒤를 잇는 차세대 부호화 방식으로서의 활용이 기대된다.

참 고 문 헌

- [1] 이 광기, 김 인겸, 박 규태, "ATM망에서의 가변비트율 영상부호화와 트래픽 제어," 한국통신학회지, 제 9권 8호, pp. 581-591, 1992
- [2] 이 광기, 김 인겸, 정 준용, 류 충일, 박 규태, "ECVQ를 이용한 영상의 계층적 대역분할 부호화," 전자공학회 논문지, 제 31권 4호, pp. 215-223, 1994
- [3] A. Gersho and R. M. Gray, Vector Quantization and Signal Compression, Kluwer Academic Publishers, 1992.
- [4] R. M. Gray, "Time-invariant trellis encoding of ergodic discrete-time sources with a fidelity criterion," IEEE Trans. Inform. Theory, Jan. 1977.
- [5] A. J. Viterbi and J. K. Omura, "Trellis encoding of memoryless discrete-time sources with a fidelity criterion," IEEE Trans. Inform. Theory, May 1974.
- [6] G. Ungerboeck, "Channel coding with multilevel/phase signals," IEEE Trans. Inform. Theory, Jan. 1982.
- [7] O. Rioul and M. Vetterli, "Wavelets and signal processing," IEEE Signal Processing magazine, Vol 8, No. 4, pp. 14-38, Oct 1991.
- [8] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," IEEE Trans. Image Processing, IP-1, No. 2, pp. 205-220, 1992.
- [9] S. G. Mallat, "A theory for multi-resolution signal decomposition: the wavelet representation," IEEE Trans. PAMI., PAMI-11, No. 7, pp. 674-693, 1989.
- [10] Michael Unser, Akram Aldroubi and Murray Eden-"A family of polynomial spline wavelet transforms," Signal Processing 30, pp. 141-162, 1993.
- [11] Michael Unser, Akram Aldroubi and Murray Eden-"Fast B-spline transforms for continuous image representation and interpolation," IEEE Trans. PAMI, Vol. 13, No. 3, March 1991, pp. 277-285.
- [12] 정 준용, 김 인겸, 류 충일, 이 광기, "스플라인 웨이브렛 변환을 이용한 영상의 다해상도 부호화에 관한 연구," 전자공학회 추계학술대회 논문지, 제 16권 2호, pp. 698-701, 1993년
- [13] H. S. Wang and N. Moayeri, "Trellis coded vector quantization," IEEE Trans. Commun., Aug. 1992.
- [14] M. W. Marcellin, "Transform coding of images using trellis-coded quantization," ICASSP, Apr. 1990, pp. 2241-2244.
- [15] S. P. Lloyd, "Least squares quantization in PCM," IEEE Trans. Inform. Theory, IT-28, No. 2, pp. 129-137, 1982.
- [16] N. S. Jayant and P. Noll, Digital Coding of Waveforms: Principles and Applications to Speech and Video, Prentice-Hall, N.Y., 1984.
- [17] W. A. Pearlman and A. Chekima, "Source coding bounds using quantization reproduction levels," IEEE Trans. Inform. Theory, May, 1984.
- [18] T. R. Fischer, M. W. Marcellin, and M. Wang, "Trellis coded vector quantization," IEEE Trans. Inform. Theory, Nov. 1986.
- [19] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," IEEE Trans. ASSP, Jan.

1989.

[20] E. A. Riskin, "Optimal bit allocation via the generalized BFOS algorithm," IEEE Trans. Inform. Theory, Mar. 1991.

[21] P. H. Westerink, J. Biemond, D. E. Boekee, "An optimal bit allocation Algorithm for sub-band coding," IEEE ICASSP, pp. 757-760, 1988.

저 자 소 개



金仁謙(正會員)

1967년 3월 10일생. 1990년 2월 연세대학교 전자공학과 졸업(공학사). 1992년 8월 연세대학교 전자공학과 졸업(공학석사). 1992년 9월-현재 연세대학교 전자공학과 박사과정

柳充一(正會員)

1970년 2월 25일생. 1992년 2월 연세대학교 전자공학과 졸업(공학사). 1994년 8월 연세대학교 전자공학과 졸업(공학석사). 1994년 8월-현재 대우전자 중앙연구소



金成萬(正會員)

1960년 11월 5일생. 1984년 2월 연세대학교 전자공학과 졸업(공학사). 1986년 2월 연세대학교 전자공학과 졸업(공학석사). 1986년 9월-현재 연세대학교 전자공학과 박사과정 재학중. 1990년 3월-현

鄭賢民(正會員)

1960년 11월 5일생. 1984년 2월 연세대학교 전자공학과 졸업(공학사). 1986년 2월 연세대학교 전자공학과 졸업(공학석사). 1986년 3월-1992년 3월 한국통신 연구개발원. 1992년 3월-현재 연세대학교 전자공학과 박사과정 재학중

재 한국통신 재직중

朴圭泰(正會員) 第30卷 B編 第10號 參照

현재 연세대학교 전자공학과 교수