

# 웨이브렛 변환계수의 트리구조를 이용한 방송용 HD-VCR의 부호화 기법

正會員 김 용 규\*, 정 현 민\*\*, 이 병 래\*\*\*, 강 현 철\*\*\*\*

## Tree Structured Wavelet Transform Coding Scheme for Digital HD-VCR

Yong Kyu Kim\*, Hyun Meen Jung\*\*, Byeong Rae Lee\*\*\*,  
Hyun Chul Kang\*\*\*\* *Regular Members*

### 요 약

웨이브렛 변환을 이용하여 방송용 HD-VCR(high definition video cassette recorder)의 요구조건들을 충족시키는 부호화 기법을 제안한다. 고정 비트율(constant bit rate)을 얻기 위하여, 앞 화면(frame)의 부호화 결과를 토대로 현재 화면의 양자화 간격을 결정하는 전방 제어 기법(forward rate control)과, 독립부호(IDC:independently decodable code) 와 종속부호(DDC:dependently decodable code)로 구성되는 2단계 부호기(2 level coder) 기법을 제안한다. 에러(error)의 확산을 최소화시키기 위하여, 전체 변환영상을 변환블록(transform block)으로 재구성하고, 각 변환블록을 트리구조(tree structure)로 표현하여 독립적으로 부호화한다. 실험 결과, 제안한 부호화 기법은 기존의 DCT(discrete cosine transform)를 사용한 부호화 기법보다 같은 압축률에서 블록현상(block effect)이 없는 우수한 화질을 나타냈다. 제안한 전방 제어 기법과, 2단계 부호기를 이용하여 매우 정교한 비트량 제어(rate control)를 구현하였다.

### ABSTRACT

A wavelet transform coding method that fulfills the requirements of HD-VCR(high definition video cassette recorder) for studio applications is proposed. A constant bit rate is achieved by a forward rate control technique which determines the quantizer stepsize based on the coding result of the previous frame. We also propose a

\*성경대학교 전산정보학과

\*\*한국통신 연구개발본부 무선통신연구소

\*\*\*한국방송통신대학교 전자계산학과

\*\*\*\*인천대학교 정보통신공학과

論文番號:96365-1118

接受日字:1996年 11月 18日

two-level coder that consists of both the IDC(independently decodable code) and the DDC(dependently decodable code). To minimize error propagation, the transformed coefficients are restructured into transform blocks which are represented by a tree structure. The result shows that the proposed coding scheme produces better picture quality without block effects than that of DCT(discrete cosine transform) based coding schemes at the same compression ratio. The proposed method meets most of the requirements of HD-VCR.

## I. 서 론

고품질의 영상데이터 서비스의 요구가 증대되면서, 기존의 TV 신호보다 해상도가 많이 늘어난 HDTV(high definition television)에 대한 연구가 활발히 진행되었고 아울러 HDTV신호의 기록에 대한 관심도 증가되었다. 따라서 기존의 VCR(video cassette recorder)을 대체할 기록용량(data rate)이 증가된 새로운 기록 시스템의 필요성이 증가되고 있다. 예를 들어 미국방식의 경우는 해상도가  $1920 \times 1050$ 이 되는데, 이 경우 늘어난 정보의 양은 기존 VCR의 5배에 해당하여 HDTV 신호에 적합한 새로운 기록 시스템을 필요로 하게 되었다.

HDTV 신호를 방송용으로 취급하기 위해서는 기록/저장할 수 있는 매체가 여러 가지 제한사항을 갖는다. 이러한 제한사항 중 가장 문제가 되는 것은 전송속도와 기록용량이다. 방송용으로 화질에 중점을 두고 60분 이상의 분량을 기록하기 위해서는 현재의 기록매체들 중 테이프를 이용하는 것이 가장 효율적이라고 볼 수 있으나, 현재 테이프의 기록밀도를 고려할 때, HDTV가 요구하는 1.2 Gbps(HDTV 방식에 따라 약간씩 차이는 있음)의 기록용량과 전송속도를 수용하지는 못한다. 따라서 화질에 최대한 영향을 미치지 않는 범위 내에서 영상 압축을 실시해야만 한다.

영상 압축 기법은 방송용 HD-VCR이 요구하는 기능 중에, 화면(frame)단위의 편집을 위해, 각 화면단위로 일정하게 압축되어야만 하고, 다양한 배속의 재생이 가능하기 위해서는 영상 압축 기법이 기록 시스템과 밀접한 관련을 갖고 있으며 부가기능을 고려한 알고리즘이어야 한다. 또한 화질은 고화질을 유지해야 하고, 여러 번 복사해도 그 화질이 유지될 수 있어야 한다. 따라서 다양한 부가기능과 화면단위의 편집을 위해 낮은 압축률과 고화질을 유지할 수 있는 화면내 부호화(intra frame coding) 기법이 필요하다.

방송용 HD-VCR의 부호화기는 전술한 여러 가지 요구사항과 제한사항들을 만족시키기 위해 여러 가지 기법을 사용하고 있으나, 아직 연구가 진행되어야 하는 실정이다. 대부분의 부호화 추세는 DCT(discrete cosine transform) 기반의 부호화 시스템으로서 화면당 일정한 압축률을 얻기 위한 비트량 제어(rate control) 방법에 많은 연구가 진행되고 있고, 아울러 가변 배속 재생을 위한 부호화 기법도 주된 연구과제로 대두되고 있다. 방송용 디지털 HD-VCR의 경우 고화질을 요구하고 있으며 이에 따르는 압축률은 4:1 이내로 하는 것이 타당한 것으로 알려져 있다[1]. 그러나 영상 데이터를 압축하기 위해 DCT를 근간으로 하는 기존의 표준화된 부호화 기법을 사용할 경우 블록변환(block transform) 기법을 사용하므로, 변환 기법의 특성상 블록현상(block effect)이 발생할 수 있고, 이는 화질에 심각한 결과를 초래한다[2][3][4]. 이러한 블록현상의 근본적인 원인은 변환기법과 영상의 통계적 특성에 기인하기 때문에 비록 4:1과 같이 낮은 압축율로 부호화 한다고 하여도, 영상에 따라서는 블록현상이 나타나기도 한다. 아울러 가변양자화, 또는 허프만 부호화(Huffman code)를 사용할 경우, 비트량 제어에 많은 어려움을 내포할 뿐만 아니라 부가기능(trick play) 또는 화면단위의 편집에 장애요인으로 존재한다.

이와 같은 단점을 보완하기 위해 WT[5][6][7](wavelet transform)에 대한 연구가 활발하다. WT의 특성상 비정상적(non-stationary)인 영상신호를 처리하기에 유리하고, 블록현상을 제거할 수 있으며, 점진적 전송이 가능하므로 부호화, 복호화시 중요도에 따른 우선 순위의 부여가 가능하며, DCT와는 달리, 화소의 증가에 따라 계산량이 선형적인 증가를 유발하며, 블록변환 기법이 필요 없고, 인간의 시각 특성을 적절하게 반영하는 변환기법이므로 고압축에도 적합하다.

웨이브렛 변환된 영상에서 같은 방향성분을 갖는

대역(band)들 사이에 기하학적 유사성(geometric similarity)이 존재한다. 이러한 유사성(similarity)은 여러 가지의 트리구조로 표현될 수 있다. 최근에 Shapiro [8]는 제로트리(zerotree)를 이용하여 점진적 근사형태의 영상 부호화기(successive approximation image coder)를 개발하였다. 이 기법은 주어진 변환 영상을 여러 번 검색해야 하므로 처리 속도가 늦다는 단점이 있다. 따라서 이 기법은 실시간 처리를 위해 매우 빠르고 복잡한 처리 장치를 필요로 한다. 그 후 단 2회의 검색으로 부호화를 완료할 수 있는 기법이 Kim[9][10]에 의해 제안되었으나, 이 부호화 기법은 5개의 산술 부호화기(arithmetic coder)가 필요하다는 단점이 있다. 따라서 단순한 2가지 부호화 기법 모두 방송용 HD-VCR에 적용하기에는 적합하지 못하다.

본 논문에서는 편집, 고속재생, 고화질 등의 방송용 VCR의 요구조건을 만족시키고 트리구조를 사용하여 부호화 효율을 증대시킨 부호화 기법을 제안한다. 입력영상은 휘도 신호 및 색차 신호를 각각  $1024 \times 1920$ ,  $1024 \times 960$ 의 해상도로 가정하고, 웨이브렛 변환은 4단계(level)까지 반복 분해한다. 전체적인 압축률은 4:1로 가정했다.

## II. 방송용 VCR의 제한사항

### 2.1. 고화질

일반적인 가정용 VCR보다 우수한 화질을 요구하며, 영상압축으로 발생할 수 있는 시각적인 열화가 없어야 한다. 이것은 주관적인 영상 평가에서 영상압축을 거친 영상과 원 영상을 거의 구분할 수 없을 정도의 고화질을 의미하며, 객관적인 영상 평가에서 높은 PSNR을 유지해야 한다. 이러한 고화질을 유지시키기 위하여 일반적으로 4:1 이내의 압축률로 테이프에 기록한다.

### 2.2. 화면단위의 편집

방송용 장비로써 사용될 때 중요한 기능 중에 하나인 편집기능은 화면단위(frame by frame)의 편집을 요구하며, 이것은 동영상 압축의 표준 기법인 MPEG (Moving Pictures Experts Group)과 같이 압축효율을 증대시키기 위해 움직임을 예측하고, 보상하는 방식의 화면간 부호화(inter frame coding)를 사용할 수 없

는 이유가 된다. 즉 모든 영상을 화면내 부호화(intra frame coding)로 처리하여야만 각 화면의 독립적인 첨가, 삭제, 수정이 가능하다.

### 2.3. 다양한 배속의 재생(부가기능, trick play)

테이프를 기록매체로 사용한 VCR은 일반적으로 순차(sequential) 재생만이 가능하므로 부가기능의 구현이 디지탈 VCR이 안고 있는 가장 어려운 문제이다. 정지 재생(still), 느린 재생(slow)은 비교적 용이하게 구현할 수 있으나, 배속 재생(fast forward)의 구현이 핵심연구 분야로 대두되고 있다. DT(dynamic tracking) 헤드나, 심층기록 기법을 이용하지 않을 때 배속 재생 기능의 구현을 위해서는 영상을 작은 부분 단위로 압축하고, 이 압축된 데이터를 테이프의 일정한 위치에 기록해야 한다. 따라서 테이프 형식(format)과 부호화 알고리즘이 밀접한 관계를 갖는다. 이러한 제한사항을 만족시키기 위하여 부호화 알고리즘은 영상의 작은 부분이 독립적으로 부호화 또는 복원될 수 있는 기법을 필요로 하며, 이 영상의 작은 부분을 균일한 비트 수로 압축한다면 용이하게 부가기능을 구현할 수 있으나, 이런 경우 효율적인 압축결과를 얻기 힘들다는 문제점이 있다. 따라서 영상압축 기법을 정할 때 기록매체인 테이프에 영상의 압축결과를 어떻게 배치할 것인가를 함께 고려해야 한다.

### 2.4. 화면단위로 일정한 압축

효율적인 화면단위의 편집이 가능하기 위해서, 또 용이한 부가기능의 구현을 위하여 각 화면 당 균일한 비트량으로 압축되어야 한다. 그러나 MPEG, JPEG (Joint Photographic Experts Group)과 같이 표준화된 영상 압축기법에서는 적응적 양자화 기법과, 허프만 부호를 사용하기 때문에 화면단위로 균일하게 압축하기 힘들다. 일반적으로 비트량 제어 기법은 후방제어(backward control) 기법과, 전방제어(forward control) 기법으로 분류되는데, VCR의 제한사항과 같이 좁은 범위(화면단위)내에서 균일한 비트량 제어를 실시할 때에는 후방제어보다 전방제어가 유리하다. 전방제어를 위해서는 영상이 입력 버퍼에 놓여진 상태에서 간단한 통계적 특성을 이용하여 비트량의 전방제어를 위한 변수 값을 결정해야 한다.

### 2.5. 에러 확산 감소

일반적으로 디지털 VCR 시스템에서 발생하는 대부분의 에러는 테이프와 자기헤드 사이에서 발생한다. 이미 발생한 에러의 확산을 최소화시키기 위하여 ECC(error correction code)를 사용하지만, 어떠한 ECC 기법도 모든 형태의 에러를 정정하지는 못한다. 따라서 영상 압축 기법 자체가 에러의 확산을 최소화시킬 수 있어야 한다.

앞에서 기술한 제한사항들을 만족시키기 위하여 각 화면은 작은 블록단위로 구분되어 각 블록이 독립적으로 복원될 수 있어야 하고, 전체 화면은 일정한 비트량으로 부호화 되어야 한다는 것으로 요약될 수 있다.

### Ⅲ. 테이프 형식

방송용 HD-VCR의 부호화기는 용이한 부가기능의 구현을 위하여 각 화면마다 또한 각 SYNC 블록마다 일정한 비트량으로 압축되어야 하며, 이 요구사항은 일반적으로 부호화기의 압축성능에 따른 화질과 상반되는 요소로 작용한다. 따라서 보다 정교하고, 효율적인 입력영상의 분석을 통해서 정확한 출력 비트량을 예측하여야 한다. 복호화기(decoder)의 경우에는 테이프의 기록/재생시 ECC의 정정능력을 초과한 에러가 발생했을 때, 출력 버퍼에 에러 숨기기(error concealment) 작업을 요구하는 신호를 필요로 한다. 에러의 확산을 최소화시키기 위해서는 가능한 한 영상을 작은 단위로 구분하여 독립적으로 부호화(encoding)/복호화(decoding)를 실시할 수 있어야 하고, 높은 부호화 효율을 위해서는 보다 광범위한 단위마다 일정한 비트량으로 부호화하여야 한다.

웨이브렛 변환을 거친 영상 신호를 부호화하기 위한 최소의 단위를 변환블록(transform block)이라고 하고 그림 1에서 웨이브렛 변환된 결과와 변환블록의 관계를 나타냈다. 변환블록은 하나의 DC 계수와 이 계수를 뿌리(root)로 하는 모든 노드들로 구성된다. 본 논문에서는 입력영상을 4단계까지 분해하여 부호화하였으므로, 하나의 변환블록은 256(16×16)개의 변환계수들로 이루어진다. 변환계수를 부호화하는 기본단위가 변환블록인 반면 VCR 테이프에 기록/재생되는 기본단위를 SYNC 블록이라고 정의하고, 하

나의 SYNC 블록은 4개의 변환블록(2개의 Y 변환블록과 Cb, Cr 변환블록 각 1개)으로 구성된다. 앞에서 기술한 디지털 VCR의 제한 사항 중에서 에러의 확산을 줄이고, 부가기능을 쉽게 구현하기 위하여 가능한 최소의 크기로 SYNC 블록을 설정하였다. 본 논문과 같이 영상을 4 단계까지 분해하였을 때 독립적인 기록/재생의 기본 단위는 원 영상에서 16×32의 크기에 해당된다. 따라서 테이프의 기록/재생시 발생하는 에러는 각 SYNC 블록으로 국한시킨다.

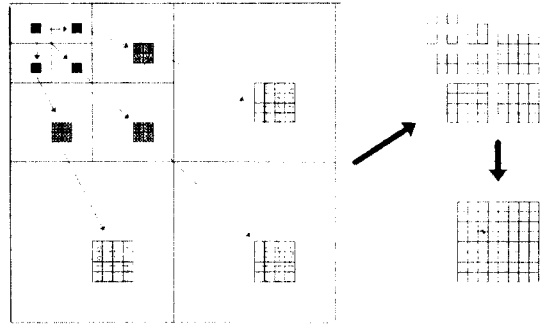


그림 1. 변환영상과 변환 블록의 관계

Fig. 1 The relationship between WT frame and transformed block

표 1. 테이프 형식

Table 1. The tape format

Items	Data
Compression ratio	4:1
Image Size	1024×1920(Y)
	1024×960(Cb)
	1024×960(Cr)
Rotations	5400 rpm
1 Frame	24 tracks / frame
	4 channels / frame
	3840 SYNC's / frame
1 Channel	6 tracks / channel
1 Track	160 SYNC's / track
1 SYNC Block	4 transform blocks / SYNC
	256 bytes / SYNC
1 Transform Block	16×16 pels / block

부가기능의 용이한 구현과 에러의 확산을 최소화시키기 위하여 테이프의 기록/재생의 기본단위인 SYNC 블록은 4:1 압축을 가정하여 256 바이트로 정했다. 본 논문에서 정의한 테이프 형식을 표 1에 나타내었다. 이 테이프 형식을 VCR 테이프에 기록하면 그림 2와 같이 각 SYNC 블록이 배치된다. 각 SYNC 블록에는 영상의 휘도신호  $16 \times 32$ , 색차신호  $16 \times 16$  크기의 영역에 해당하는 압축결과를 수용한다.

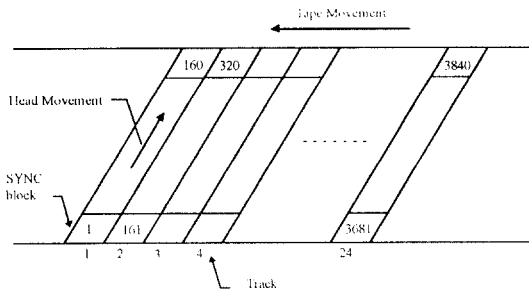


그림 2. 테이프에서 트랙과 SYNC 블록의 배치  
Fig. 2 The arrangement of track and SYNC block on a tape

그림 3은 VCR 테이프의 트랙에 부호화 결과가 SYNC 블록의 형식으로 배치된 구조를 나타냈다. 각 SYNC 블록은 평균 256 byte를 수용할 수 있으므로 부호화된 결과가 1번, 2번, 3번, 4번, 7번 SYNC 블록과 같이 256 바이트에 미달하는 SYNC 블록과 5번, 6번과 같이 256 바이트를 초과하는 SYNC 블록이 발생한다. 이러한 경우 효율적인 압축능력과 화질을 유지하기 위하여 256 바이트를 초과하는 정보는 같은 화면내의 인접 SYNC 블록에 분산 배치한다. 그러므로 5번 SYNC 블록과 같이 압축결과가 256 바이트를 초과하는 정보는 1번 SYNC 블록에 분산 배치하고, 6번 SYNC 블록의 경우에는 2번, 3번, 4번, 7번 SYNC 블록에 배치한다.

그림 3과 같은 방법으로 압축결과를 각 SYNC 블록에 배치하며, SYNC 블록은 그림 4와 같은 형태로 구성된다. 각 SYNC 블록은 영상에서 해당되는 고유한 위치의 부호화 결과를 저장하여야 하므로 각 SYNC 블록의 내부는 우선적으로 해당하는 고유 위치의 영상 압축 결과를 저장하며, 이 부분은 부가

능 구현시 우선적으로 복원된다. 만약 해당 SYNC 블록의 정보량이 256 바이트에 미달되었을 경우 나머지 부분은 다른 SYNC 블록의 초과된 정보가 기록된다.

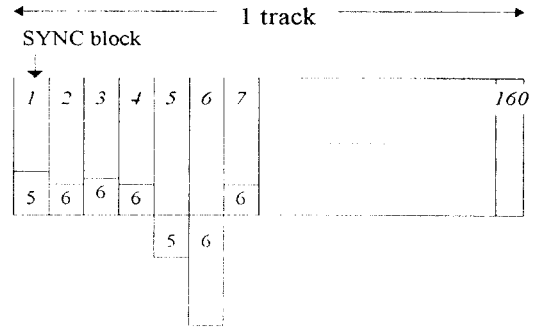


그림 3. 각 SYNC 블록의 비트 배열의 예  
Fig. 3 The bit arrangement of SYNC block

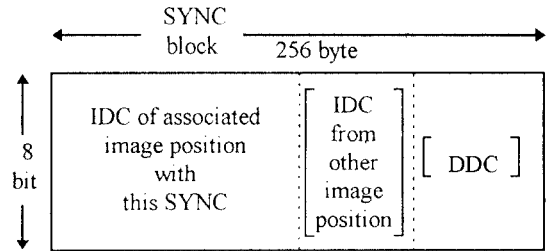


그림 4. SYNC 블록의 구성  
Fig. 4 The organization of a SYNC block

#### IV. 제안된 부호화 기법

WT는 영상을 피라미드(pyramid) 구조로 표현하는 기존의 방식인 다해상도 분석(multi resolution analysis)과 대역분할 부호화(subband coding)를 하나로 통합한 이론이므로 변환된 각 부영상들은 저주파 대역부터 고주파 대역으로의 계층적 정보를 갖고 있고, 이를 이용하여 보다 적절한 부호화를 행할 수 있다.

웨이브렛 변환의 특징은 주어진 영상신호를 ①계층적(hierarchical)인 형태로 분해하여 표현한다는 것과, ②계층적 구조에서 각 계층간에 서로 기하학적인 유사성(self-similarity)이 존재한다는 것과, ③분해된 계층구조는 영상과 같이 2차원 신호에서 각 대역이

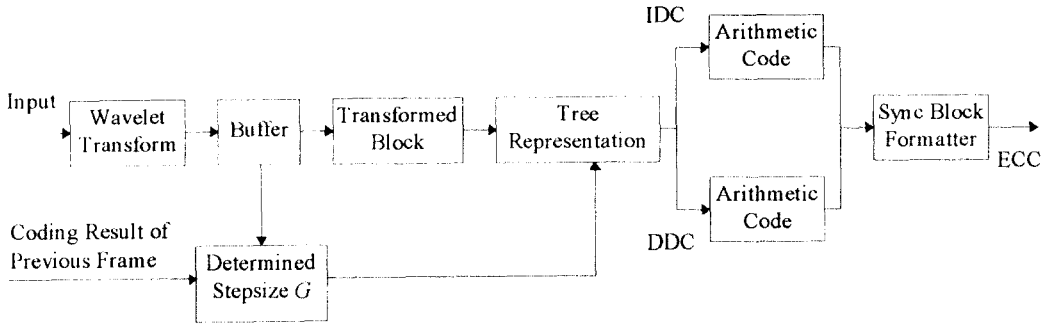


그림 5. 부호화기의 전체 구성도  
Fig. 5 The encoder block diagram

고유한 방향성분을 갖게 된다는 것이다. 따라서 효율적인 웨이브렛 변환을 이용한 부호화 기법은 이러한 특징들을 이용하는 것이 바람직하며, 본 논문에서는 이러한 특징들을 이용한 트리(tree) 형태의 부호화 기법을 제안한다.

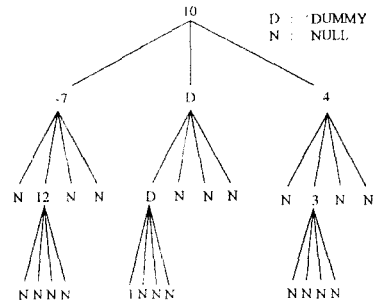
제안된 영상 부호화 기법을 그림 5에 블록도 형태로 나타냈다. 변환된 입력영상은 변환블록으로 재구성되고 각 블록은 제안된 트리구조로 표현된다. 이 트리를 효율적으로 부호화 시키기 위해 적응 모델(adaptive model)의 산술 부호화기(arithmetic coder) [11]를 사용했다. 제안된 기법은 단 한번의 검색을 사용하며, 양자화 단계값  $G$ 를 변화시켜 다양한 압축률이 얻어진다. 산술 부호화기에 의해 압축된 데이터는 테이프에 저장되기 위하여 SYNC 블록의 형태로 재구성된다. 제안된 부호화 기법은 프레임 단위의 부호화 또는 필드 단위의 부호화 모두 가능하지만 본 논문에서는 프레임 단위를 중심으로 설명한다.

4.1. 트리구조의 표현

변환된 영상은 에러의 확산(propagation)을 최소화시키기 위하여 가능한 한 작은 크기의 변환블록으로 재구성했다. 그림 6은 웨이브렛 변환을 3단계까지 반복 분해한 경우에 대하여 부호화 과정을 설명하고 있다. 그림 6(a)는 하나의 DC 성분과 이에 해당하는 AC 성분으로 구성된 변환블록을 나타내고 있다. 이 변환블록은 그림 6(b) 형태의 제안된 트리구조로 표현된다. 제안된 트리구조에서 3개의 하위노드를 갖는 DC 성분의 계수를 뿌리(root)로 간주한다. 반면에 가장 높은 주파수 성분을 나타내는 대역은 하위노드를

10	7	0	1	0	0	0	0
0	4	0	0	0	0	0	0
0	0	0	3	0	0	0	0
0	0	0	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
0	0	0	0	0	0	0	0

(a) 변환블록의 예( $G = 1$ 인 경우)  
(a) Example of a transformed block(for  $G = 1$ )



(b) 변환블록의 트리구조의 표현  
(b) Tree representation of above block

10'	-7'	4'	12'	3'	1'
•010	-11D•00	N•100NN	DN•NN	N•1NN	NNNN
					•NNN
					NNNN

(c) 위 트리구조를 부호화한 결과  
(c) Encoding data stream of given tree

그림 6. 트리구조의 표현과 부호화 실시의 예  
Fig. 6 The simple example of the tree representation and encoding

가지 못하고, 그 이외의 노드는 모두 4개의 하위노드를 갖는다. 이 트리구조는 웨이브렛 변환계수의 계층적 특성을 적절히 이용하기 위한 것이다.

주어진 양자화 단계(quantizer stepsize)  $G$ 에 따라 트리의 각 노드는 *POSITIVE*, *NEGATIVE*, *NULL*, *DUMMY*의 4가지 형태의 노드로 다음 식에 의해 분류된다.

$$T(i, j) = \begin{cases} \text{POSITIVE} & \text{if } |C_{ij}| + G/2 \geq G \text{ and } C_{ij} > 0 \\ \text{NEGATIVE} & \text{if } |C_{ij}| + G/2 \geq G \text{ and } C_{ij} < 0 \\ \text{DUMMY} & \text{if } |C_{ij}| + G/2 < G \text{ and } \exists m \in D_{c_q}, |m| \geq G \\ \text{NULL} & \text{if } |C_{ij}| + G/2 < G \text{ and } \forall m \in D_{c_q}, |m| < G \end{cases} \quad (1)$$

여기서  $D_{c_q}$ 는 변환계수  $C_{ij}$ 의 모든 하위노드의 집합

반올림된 계수값의 크기가 양자화 단계  $G$  보다 크면 그 계수의 부호에 따라서 *POSITIVE* 또는 *NEGATIVE*의 속성을 갖는다. *DUMMY*는 반올림된 현재 계수의 값이  $G$  보다 작지만 하위노드(descendent nodes)의 계수값 중에  $G$  보다 큰 것이 있다는 것을 의미하며 그림 6(b)에서  $D$ 로 표시된 노드가 여기에 해당된다. *NULL*은 반올림된 현재 계수와 모든 하위노드의 크기가  $G$  보다 작다는 것을 의미하고 그림 6(b)에서  $N$ 으로 해당 노드를 표시했다. 상위노드의 속성이 *NULL*이 아닌 경우에만 현재노드가 부호화되어 전송되고, *NULL*인 경우에는 전송하지 않는다. 트리구조의 모든 하위노드는 상위노드의 변환계수가 전송된 후 전송된다. 그림 6(c)와 같이 각 변환계수는 노드의 형태를 나타내는 정보와, 크기를 나타내는 이진 심분(binary symbol)이 결합하여 변환계수값을 표현한다.

제안된 부호화 기법에서는 균일 양자화기(uniform quantizer)를 사용하여 각 변환계수를 양자화 한다.  $G$  값을 변화시키 다양한 압축률을 얻을 수 있고 부호화시 발생하는 부호화 오차는 모두 양자화 오차에서 기인한다. 일반적으로, 칼라(color)영상의 부호화 기법은 3가지 색성분을 각각 독립적으로 취급하기 때문에 각 색성분들 사이에서 불균형의 양자화 오차가 존재하는 문제가 발생하기 때문에 본 논문에서는 각 색성분에 대하여 동일한 양자화 단계  $G$ 를 적용하므로 색

성분들 사이에 비교적 균등한 양자화 오차를 유도하였다.

#### 4.2. 변환계수 크기 표현

일반적으로 수를 표현할 때는 수의 부호(sign, +/-)를 나타내는 비트와, 크기를 나타내는 비트가 필요하다. 그러나 임의의 수를 최소한의 비트 수로 나타낼 때 MSB는 항상 '1'이 된다. 어떤 수가 2진수로 표현되었을 때 그 수의 시작과 끝을 알 수 있다면, 그 수의 부호(sign) 비트와 MSB를 공유할 수 있기 때문에 1비트를 절약할 수 있다. 따라서 주어진 변환계수  $C_{ij}$ 는 다음과 같이 표현된다.

$$C_{ij} \approx C_{ij} = s(2^H + \sum_{q=0}^{H-1} b_q 2^q)G \quad (2)$$

$$s = \begin{cases} 1 & \text{for } T(i, j) = \text{POSITIVE} \\ -1 & \text{for } T(i, j) = \text{NEGATIVE} \end{cases}$$

$$b_q = \begin{cases} 0 & \text{for } |C_{ij} + G/2| < (2^q + \sum_{k=q+1}^H b_k 2^k)G \\ 1 & \text{for } |C_{ij} + G/2| \geq (2^q + \sum_{k=q+1}^H b_k 2^k)G \end{cases}$$

$$H = \left\lceil \log_2 \left( \frac{|C_{ij} + G/2|}{G} \right) \right\rceil$$

여기서  $\lceil x \rceil$ 는  $x$ 를 넘지 않는 최대 정수

변환블록의 변환계수들이 앞에서 정의한 순서대로 전송될 때, 각 변환계수는 반드시 노드의 형태를 나타내는 정보로 시작하고 크기를 나타내는 정보가 첨부되므로, 변환계수의 크기를 나타내는 정보의 시작과 끝은 노드의 형태 정보를 통해서 알 수 있다. 예를 들어  $G=1$ ,  $C_{ij}=9$ 인 경우 노드의 형태 정보는 +: *POSITIVE*가 되어 '+001'으로 표현된다. 또한  $G=1$ ,  $C_{ij}=-1$ 인 경우는 '-'로 표현된다. 그림 6(b)의 트리나 그림 6(c)와 같은 부호열(encoding data stream)로 표현되어 산술 부호화기의 입력으로 주어진다.

#### 4.3. 산술부호화

제안된 형태의 트리구조와 변환계수의 크기를 기록하기 위해서 적절한 방법으로 부호화하여 압축해야 한다. 일반적으로 부호화 성능과 구현의 용이성을 놓고 고려할 때, 허프만 부호화 기법을 사용하는 방법과, 산술 부호화 기법을 사용하는 방법이 있다. 허

프만 부호의 장점은 평균 부호길이가 작고, 순간적인 디코딩(instantaneous decoding)이 가능하다는 것이다. 그러나 허프만 부호는 전체 트리구조를 표현하는데 사용되는 각 심볼의 확률을 미리 알고 있거나, 적절한 확률 모형을 설정해야 하는 단점이 있다. 다양한 입력 영상의 트리구조를 부호화할때, 모든 입력 영상에 대해 사전에 심볼 확률을 미리 구하는 것은 불가능하므로, 적절한 확률 모형을 설정하는 것이 바람직하다. 그러나 입력영상의 실질적인 심볼들의 확률들이 이미 설정된 확률 모형과 차이가 있을 때는 압축효율이 크게 저하된다. 또한 어느 하나의 심볼에 큰 확률( $\approx 1$ )이 할당될 때는 부호화 성능이 떨어지며, 일반적으로 허프만 부호를 사용하여 1bpp 이하로 압축할 수 없다는 문제점을 안고 있다. 이와 같은 단점들을 극복하기 위하여 적응모델(adaptive model)을 사용한 산술 부호화 기법을 이용하였다. 산술 부호화기를 사용할 때, 표현된 트리구조가 가지고 있는 엔트로피에 가장 근접한 압축 결과를 얻을 수 있고, 적응 모델을 사용하므로, 트리구조에 대한 확률 모형을 설정할 필요도 없고, 부호화를 위해 각 심볼의 확률을 미리 구할 필요가 없다는 장점이 있다.

제안된 트리구조 표현과 이에 따른 변환계수의 크기 표현기법으로 영상을 압축할 때는 트리구조의 표현을 위하여 4가지 심볼(POSITIVE, NEGATIVE, DUMMY, NULL)과 변환계수의 크기를 2진수로 표현하기 위하여 2가지 심볼('0', '1')이 필요하다. 따라서 총 6 종류의 심볼의 산술 부호화기를 사용했다.

#### 4.4. 비트량 제어

제안된 트리표현과 변환계수의 크기 표현 방법을 통해 각 변환블록은 하나의 부호열로 표현되며, 이 부호열은 산술 부호화를 통해 압축된다. VCR의 부가 기능을 용이하게 구현하고, 에러의 확산을 최소화하기 위해 웨이브렛 변환의 성질을 고려하여 영상을 가장 작은 단위로 부호화한 반면, 부호화 효율을 극대화시키기 위하여 VCR 편집의 최대 단위인 화면마다 균일한 비트율로 압축하는 전방 제어방식의 비트량 제어 기법을 사용하였다. 이러한 부호화 방식의 압축률과 화질은 주어진 양자화 단계  $G$ 에 의해서 결정되므로, 매 화면마다 균일한 비트량으로 압축하기 위해서는  $G$ 의 값을 매우 정교하게 조정해야한다. 그러나

아무리 정교한 전방제어 방식으로  $G$  값을 추정하더라도 정확한 비트량을 예측하는 것은 불가능하다. 산술부호(arithmetic code)를 사용하면서 각 화면 당 일정한 비트량을 유지하기 위하여, 독립부호(IDC: independently decodable code)와 종속부호(DDC: dependently decodable code)[2][12]로 구성된 2단계(level) 부호화 기법의 사용을 제안한다. 종속부호는 Cr, Cb의 신호성분 중에서 가장 고주파 대역의 계수들의 최하위 비트(LSB: least significant bit)로 구성된 부호열이고, 독립부호는 종속부호를 제외한 모든 부호열을 의미한다. 독립부호의 비트량이 목표량보다 적다면 종속부호가 나머지 목표량까지 기록된다. 대부분의 중요한 정보가 독립부호에 담기게되므로 종속부호 없이 복원된 영상도 우수한 화질을 유지할 수 있다. 고속재생시(high speed playback)에는 독립부호만을 복원하여 영상을 구성한다. 각 SYNC 블록에는 그 블록 고유의 독립부호가 있기에 다른 SYNC 블록과 무관하게 독립적으로 복호화가 가능하며, 종속부호는 잉여 비트를 갖고 있는 SYNC 블록에 저장된다. 이와 같은 2단계 부호화를 통해서 보다 정교한 비트량 제어가 가능하다. 따라서 양자화 단계  $G$ 로 부호화할 때 독립부호와 종속부호가 갖는 비트량을 각각  $bit\_IDC(G)$ 와  $bit\_DDC(G)$ 라고 할 때 양자화 단계  $G$ 가 다음 조건을 만족시키면 된다.

$$bit\_IDC(G) < \text{목표 비트량} < bit\_IDC(G) + bit\_DDC(G) \quad (3)$$

독립부호가 갖는 비트량이 목표 비트량 보다 작고, 독립부호와 종속부호가 갖는 비트량의 합이 목표 비트량 보다 크거나 같은 조건을 만족하는 양자화 단계  $G$ 를 결정하면 된다. 만약 2단계 부호화 기법을 사용하지 않을 경우, 전체 부호화된 비트량이 항상 목표 비트량과 같은 조건을 만족하는 양자화 단계  $G$ 를 찾아야하지만, 실제로 그러한 조건을 만족시키는 것은 불가능하다. 따라서 2단계 부호화 기법을 이용하여  $G$  값의 선택 범위가 넓게 하였다.

#### 4.5. 양자화 단계의 결정

본 논문에서는 양자화 단계를 결정하기 위한 통계값으로 유효(significant) 변환계수의 개수를 이용했다. 유효 변환계수는 임계값(양자화 단계  $G$ )보다 큰



변환계수의 개수를 의미한다. 실험적으로, 같은 압축률에서 통계적 성질이 상이한 여러 영상들에 대해 유효 변환계수의 개수는 거의 변화가 없었다. 표 2는 5가지의 MPEG 표준영상을 4:1 압축하여 구해진 유효 변환계수의 수를 나타냈다. 표 2에서 보는 바와 같이 4:1로 압축하기 위하여 양자화 단계  $G$ 값이 광범위하게 변화했으나 유효 변환계수의 개수는 비슷한 숫자를 유지하고 있다.

표 2. 압축율 4:1에서 여러 영상에 대한 유효계수의 수  
Table 2. The No. of significant coefficients of various images at compression ration 4:1

Images Items	football	flower garden	mobile	jopple	tennis
$G$	0.5845	1.01	1.61	0.656	0.555
No. of significant coefficients	77,080	76,579	78,854	74,533	77,725

최초 영상의 경우, 변환영상의 누적분포함수(cumulative distribution function)를 이용하여, 입력 영상의 크기(size)와 압축률에 따라 미리 정해진 유효 변환계수의 개수와 가장 근접한 누적분포함수를 만족하는 값을  $G$ 의 값으로 설정한다. 이전 화면의 통계적 특성이 현재 부호화할 화면과 매우 유사하므로 계속 진행되는 영상은 앞 영상의 부호화 결과를 이용하여  $G$  값을 결정한다. 이전 화면에 적용된 유효 변환계수의 개수, 부호화된 최종 비트 수와 종속부호에서 마지막으로 전송한 계수의 위치를 다음 화면에 알려준다. 만약 이전 화면의 부호화 결과, 요구하는 비트량이 목표 비트량에 미달되었을 때는 잉여 비트 수에 비례하여 현재 영상의 유효 변환계수의 개수를 늘려 설정하고 이에 따라  $G$  값을 결정하며, 만약 이전 화면의 압축결과가 목표 비트량을 초과하는 경우에는 이전 화면에서 마지막으로 전송된 계수의 위치에 따라 유효 변환계수의 수를 감소시키고 이에 따라  $G$  값을 결정한다. 칼라 영상의 경우 3 가지 색성분에 모두 같은  $G$  값을 사용하여 각 색성분이 거의 비슷한 화질을 유지하도록 하였다. 이러한 기법을 사용하여 각 색성분에 대한 비트 배분 문제를 해결하였다.

이와 같이 이전 화면의 부호화 결과를 토대로 현재 화면의 부호화 변수를 결정하는 경우에는 장면이 바

뀔 때 문제가 발생할 수 있고, 또한 최초 화면의 부호화 변수의 결정이 문제가 된다. 그러나 제안한 기법은 단순히 이전 화면의 부호화 변수를 수정없이 그대로 사용하는 것이 아니고, 이전 화면의 부호화 결과를 토대로 새롭게 부호화 변수를 결정하기 때문에 장면이 바뀌어도 비교적 안정되게 최적의 부호화 변수를 선택한다.

### V. 실험 및 결과

제안한 기법을 검증하기 위하여 실제 부호열을 생성하여 성능을 평가하였다. 위상의 선형특성을 만족하며 비교적 적은 탭(tap) 수를 갖는 Antonini의 9-7탭의[5] 쌍직교(biorthogonal) 필터로 WT를 구현하였다. 각 필터는 계수들의 계수의 합이 1이 되도록 조정하여 실험하였다. 웨이브렛 변환은 대칭확장기법(symmetric signal extension)의[13][14] 경계조건을 이용하여 4단계까지 반복 수행했다. 변환계수는 주어진 양자화단계  $G$ 로 균일양자화를 실시했다. 엔트로피 부호화(entropy encode)로 적응형(adaptive model)의 산술부호화기를 이용했다. 부호화 기법의 일반성을 위하여 양자화단계(quantizer stepsize)  $G$ 를 나타내는 2 byte의 부가정보를 사용한다. 화질을 평가하기 위하여 원영상  $f(i, j)$ 와 복원영상  $f'(i, j)$  사이의 PSNR(peak signal to noise ratio)을 다음 식으로 측정하였다.

$$PSNR = 10 \log_{10} \frac{255^2}{E\{|f(i, j) - f'(i, j)|^2\}}$$

비트량 제어 기법의 타당성은 목표 비트량에 대한 실제 부호화된 비트량의 백분율을 이용하여 측정했다.

장면 변화에 따른 비트량 제어 성능과, 부호화 성능을 검증하기 위하여 3 종류의 영상 "Train", "Woman in Flowers" 및 "Models"을 서로 연결하여 실험영상으로 사용했다. 따라서 실험영상에 두번의 장면 바뀐(scene change)이 있다. 모든 색성분의 PSNR과 목표 비트량에 대한 실제 비트량의 백분율을 각각 그림 7과 그림 8에 나타내었다. 3가지 실험 영상 모두 40 dB 이상의 PSNR을 나타냈고, 시각적으로 원영상과 거의 구분할 수 없을 정도의 화질을 유지하였다. 목표 비트량에 대한 실제 부호화된 비트량은 2번의 장면

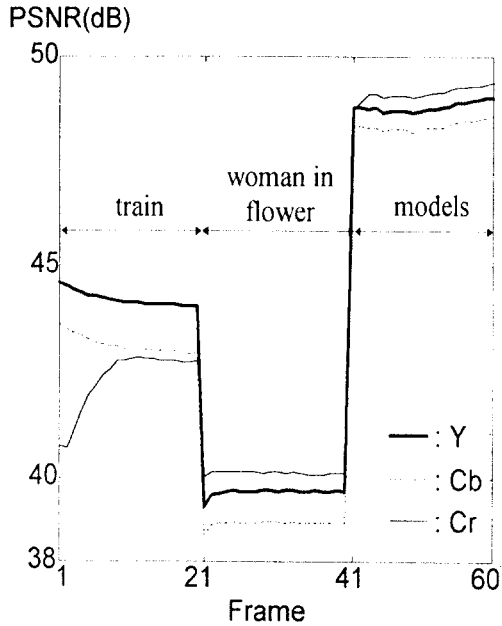
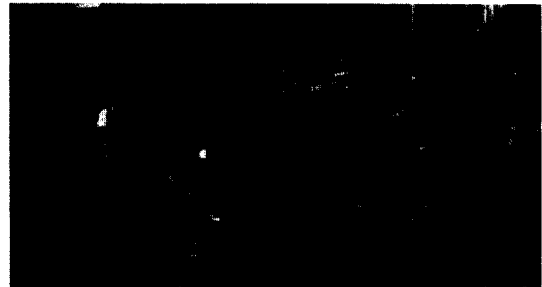


그림 7. 실험영상의 부호화 결과  
Fig. 7 The experiments results in PSNR

변화를 포함하여 평균 99.897%를 유지했다. 제안된 기법이 장면 바뀔 때 잘 대처하며 높은 화질을 유지하고 있으며 매우 정교한 비트량 제어가 이루어지고 있다는 것을 알 수 있다. 장면 바뀔 때 발생했을 때 적절한 양자화단계  $G$ 가 일반적으로 3화면이 지나는 동안에 결정되었다. 본 논문에서 제안된 부호화 기법은 칼라(color) 영상을 처리하는 경우, 각 색성분들 사이에서 발생할 수 있는 PSNR의 불균형 문제를 개선하기 위하여 각 색성분에 대하여 동일한 양자화 단계  $G$ 를 적용하였으므로, 각 색성분에 대한 부가적인 비트



a) train



b) woman in flower



c) models

a)~c) 실험에 사용된 원영상  
a)~c) The original image of test sequence

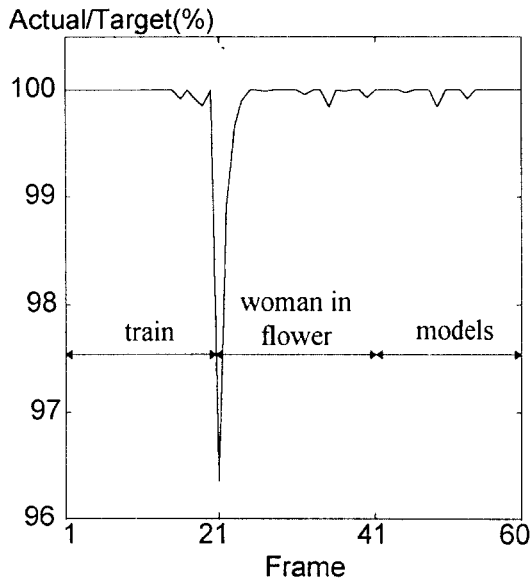


그림 8. 실험영상의 비트 제어 결과  
Fig. 8 The experiments results of the proposed rate control algorithm



d) train



e) woman in flower



f) models

(d)~(f) 실험을 통해 얻어진 결과영상  
 (d)~(f) The result image of test sequence

그림 9. 실험에 사용된 원영상과 결과 영상  
 Fig. 9 The original and result image of test sequences

배분 없이 그림 7과 같이 3가지 색성분들의 PSNR이 비교적 균등한 결과를 얻었다. 또한 각 색성분을 거의 동일한 PSNR로 부호화할 수 있었으므로, 일반적인 경우 휘도 신호에 색차신호 보다 많은 비트를 할당할 수 있었다. 실험에 사용된 3가지 표준영상에 대한 원영상과 결과영상을 그림 9에 나타내었다.

본의 실험에서 기존에 DCT를 근간으로 이루어진 부호화 기법[15] 보다 약 3 dB 향상된 결과를 얻었으

며, 아울러 블록현상(block effect)이 완전히 제거되었다. 2단계 부호화 기법을 이용하여 화면당 일정한 비트량으로 부호화하였고, 고속재생시(high speed playback) 독립부호를 사용하여 테이프의 각 SYNC 블록이 독립적으로 복원될 수 있으므로 제안된 기법은 방송용 HD-VCR에 적합했다.

## VI. 결 론

WT를 이용하여 방송용 HD-VCR의 여러 가지 요구조건을 만족시키는 부호화 기법을 제안하였다. 제안한 부호화 기법은 기존의 DCT를 이용한 부호화 기법보다 우수한 화질을 나타내었고, 또한 DCT의 문제점인 블록현상(block effect)을 완전히 제거하였다. 이전 화면의 부호화 결과를 토대로 현재 화면의 부호화 변수를 결정하는 전방제어기법(forward rate control)은 비교적 간단하면서도 매우 효율적임이 판명되었다. VCR의 고속재생시 독립부호만을 재생하는 2단계 부호화 기법을 이용하여 방송용 HD-VCR의 부가기능의 구현이 가능했다.

WT는 DCT나 DFT처럼 변환 커널(kernel)이 하나로 고정되어 있지 않기 때문에, 입력되는 영상신호에 따라 적응적으로 최적의 커널을 구하여 부호화에 이용하는 것이 가능하므로 이에 대한 연구가 요구된다. 또한 초압축 부호화(very low bit rate coding) 분야에도 WT를 이용한 부호화 기법의 연구가 기대된다.

## 참 고 문 헌

1. P. Kauff, S. Rauthenberg, R. Ritter, M. Charatishvili, and M. Hahn, "An improved coding scheme for studio recording of interlaced and progressive HDTV signals", International workshop on HDTV'92 Proceedings, pp.711-719, Nov. 1992.
2. M. S. Hong, Y. Kim, et al., "Adaptive bit allocation coder for DVCR trick play," *IEEE Conf. on Cons. Elec.*, pp. 110-111, 1992.
3. S. I. Kim, Y. C. Park, D. H. Youn, W. K. Kim, and B. K. Yoo, "Bit rate reduction algorithm for digital VCR," *IEEE Trans. on Consumer Electronics*, vol. 37, No. 3, pp. 267-274, 1991.

4. D. L. McLaren, and A. Gersho, "Multiresolution high-definition television compression algorithm for digital video tape recording," *Optical Engineering*, vol. 32, No. 7, pp. 1549-1558, Jul., 1993.
5. M. Antonini, et al, "Image coding using wavelet transform", *IEEE Trans. Image Proc.*, vol. 1, No. 2, pp. 205-220, Apr., 1992.
6. S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation", *IEEE Trans. PAMI.*, vol. 11, No. 7, pp. 674-693, 1989.
7. O. Rioul and M. Vetterli, "Wavelets and Signal Processing", *IEEE SP magazine*, vol. 8, No. 4, pp. 14-38, Oct., 1991.
8. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients." *IEEE Trans. Signal Processing*, vol. 41, pp. 3445-3462, Dec., 1993.
9. Y. Kim, H.G. Kim, and K.T. Park, "New tree structure with conditional height difference for wavelet transform image coding," *IEE Electronics Letters*, vol. 31, No. 2, pp. 90-92, Jan., 1995.
10. Y. Kim, I. Choi, I. Lee, T. Yun, K. T. Park, "Wavelet transform image compression using human visual characteristics and a tree structure with a height attribute," *Optical Engineering*, vol. 35, No. 1, pp. 204-212, Jan., 1996.
11. I. H. Witten, R. M. Neal, and J.G. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol. 30, pp. 30-44, Jun., 1987.
12. H. M. Jung, Y. Kim, S. Rhee, H. Y. Song, K. T. Park, "HD-VCR Codec for Studio Application using Quadtree Structured Binary Symbols in Wavelet Transform Domain", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, No. 5, pp. 506-513, 1996.
13. Stephen A. Martucci, "Signal extension and noncausal filtering for subband coding of images", *SPIE, vol. 1605 Visual Communications and Image Processing'91*, pp. 137-148, Sep., 1991.
14. H. J. Barnard, J. H. Webber, and J. Biemond, "Efficient signal extension for subband/wavelet

decomposition of arbitrary length signals", *SPIE Vol. 2094*, pp. 966-975, 1993.

15. J. H. Kim, et al, "Two-Layered DCT Based Coding Scheme for A Digital HD-VCR," *IEEE Conf. on Cons. Elec.*, pp. 24-25, 1994.



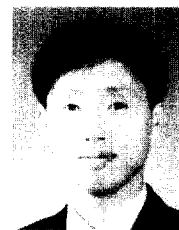
김 용 규(Young Kyu Kim) 정회원  
 1985년 2월: 연세대학교 전자공학과 졸업(공학사)  
 1987년 2월: 연세대학교 본대학원 전자공학과(공학석사)  
 1995년 8월: 연세대학교 본대학원 전자공학과(공학박사)

1987년 3월~1992년 2월: 삼성전자  
 1996년 3월~현재: 성결대학교 전산정보학과  
 ※주관심분야: 신호처리, 영상부호화, 멀티미디어



정 현 민(Hyun Meen Jung) 정회원  
 1984년 2월: 연세대학교 전자공학과 졸업(공학사)  
 1986년 2월: 연세대학교 본대학원 전자공학과(공학석사)  
 1996년 2월: 연세대학교 본대학원 전자공학과(공학박사)

1986년 2월~1992년 2월: 한국통신 연구개발본부  
 1996년 2월~현재: 한국통신 무선통신연구소 무선망 설계연구팀장  
 ※주관심분야: 영상신호처리, 영상부호화, 무선망 설계업



이 병 래(Byeong Rae Lee) 정회원  
 1985년 2월: 연세대학교 전자공학과(공학사)  
 1987년 2월: 연세대학교 대학원 전자공학과(공학석사)  
 1992년 2월: 연세대학교 대학원 전자공학과(공학박사)  
 1992년 11월~현재: 한국방송통신

대학교 전자계산학과 부교수  
 ※주관심분야: 신경회로망, 패턴인식, 영상처리 등

강 현 철(Hyun Chul Kang)

정회원

1982년 2월: 연세대학교 공과대학 전자공학과 졸업  
(공학사)

1984년 2월: 연세대학교 대학원 전자공학과 졸업(공학석사)

1989년 8월: 연세대학교 대학원 전자공학과 졸업(공학박사)

1994년 9월~1995년 8월: 미네소타대학 객원교수

1990년 3월~현재: 인천대학교 정보통신공학과 부교수

※주관심분야: 영상인식, 영상압축, 영상통신