

웨이브릿 변환을 사용한 초저속 전송 매체용 비디오 코딩

正會員 오 황 석*, 이 흥 규*

Video Coding Based on Wavelet Transform for
Very Low Bitrate ChannelHwang-Seok Oh*, Heung-Kyu Lee* *Regular Members*

요 약

최근 초저속 전송 매체를 위한 비디오 코딩 기법에 많은 관심을 가져왔다. 그러나 초저속 전송로의 제한된 비트율로 인하여 기존의 블럭 기반 변환 알고리즘을 사용한 압축 시스템은 고압축을 할 경우 블럭화 현상을 유발하여 심각한 화질 저하를 초래한다. 본 논문에서는 초저속 전송 매체를 위하여 웨이브릿 변환 영역에서 가변 크기 블럭을 가지는 다중 해상도 움직임 추정 및 보상(VMRME/C), 다중 해상도 벡터 양자화(MRVQ)를 사용하는 비디오 코딩 기법을 제안한다. 실험을 통하여 제안한 비디오 코딩 기법이 기존의 블럭 기반의 동영상 압축 기법(특히 DCT를 기반으로 하는 H.263 기본 모드)보다 PSNR 측면에서 약 0.2-0.6dB 정도 향상됨과 블럭화 현상 제거를 통한 주관적인 화질이 향상됨을 보였다.

ABSTRACT

The video coding for very low bit rate has recently received considerable attention, but conventional block based transform coding schemes suffer from the blocking effect for the constraints of bit rates. In this paper, we present a video coding system using multi-resolution motion estimation/compensation with variable size block(VMRME/C) and multi-resolution vector quantization(MRVQ) in wavelet transform domain for very low bit rate coding. It is shown that the presented scheme has better performance in the peak signal-to-noise ratio(PSNR) by 0.2-0.6 dB as well as subjective quality than that of conventional block based transform video coding techniques(especially, H. 263 which is DCT based video coding).

I. 서 론

최근 초저속 전송 매체를 통한 멀티미디어 서비스에 대한 요구의 증가로 인하여 기존의 공중 전화망이나 이동 통신망 또는 협대역 ISDN 상에서 서비스 제공을 위한 많은 연구가 진행 중이다. ITU-T의 LBC(low bitrate coding) 그룹에서는 GSTN(general swit-

*한국과학기술원 전산학과
論文番號:94290-1018
接受日字:1994年 10月 18日

ched telephone network)상에서 화상 전화기 개발을 위한 알고리즘을 개발, 표준화 및 성능 평가를 하고 있으며[1], ISO/IEC JTC1 SC29 MPEG(moving picture expert group) 4에서는 다양한 응용 서비스를 위한 표준화를 수행하고 있다[2,3].

현재까지는 큰 대역폭을 갖는 전송 매체상에서 다양한 서비스를 제공하기 위해 동영상 데이터 압축에 대한 연구가 많이 있어 왔다. 디지털 기억 장치, 통신, 그리고 방송 등을 위한 MPEG 1, MPEG 2[2,3,13,22]가 표준으로 되었으며, MPEG 4[2,3]에서는 초저속 전송을 위한 영상 압축에 대한 표준화를 연구하고 있다. 이러한 표준안에서의 영상 압축은 영상 내에 존재하는 중복성을 제거하는 것에 기반을 두고 있으며, 크게 프레임 내에서 중복성을 제거하는 코딩과 프레임간의 중복성을 제거하기 위한 코딩으로 나눌 수 있다[5].

기존 표준안은 알고리즘의 규칙성과 계산의 간단성, 그리고 제어의 편리성 및 하드웨어 개발 등을 고려하여 고정된 블록 크기를 갖는 블록 정합 알고리즘, 블록 기반 DCT(discrete cosine transform) 및 양자화를 채택하고 있다. 그러나 움직임이 적은 영상 전화기에 사용되는 영상의 특성을 고려하면 고정된 블록 크기로는 압축률을 높이는 데 한계가 있으며, 블럭화 현상을 유발한다. Zhang[8]은 블럭화 현상을 줄이기 위하여 웨이브릿 변환 영역에서의 영상 코딩을 제안하였다. 웨이브릿 변환 영역에서의 움직임 구조를 이용하여 다중 해상도 움직임 추정 및 보상(MRME/C: multi-resolution motion estimation/compensation) 방법으로 움직임 추정 및 보상을 하였으며, 잔류 오차의 코딩에 다양한 방법을 적용하여 결과를 제시하였다. 실험에서 웨이브릿 변환을 통한 영상의 분해를 먼저 수행한 후 움직임 추정 및 보상한 기법이 공간 3에서의 움직임 추정/보상 기법보다 나은 성능을 보였다. 그러나 Zhang의 기법은 MRME/C시 동일 해상도에서 고정 크기 블럭을 사용하므로 인하여 많은 비트율을 발생 시켰다.

본 논문에서는 초저속 전송 매체용 비디오 코딩에서 요구하는 (1)전송을 제한으로 인한 높은 압축률, (2)고압축에서의 주관적인 화질 향상을 위하여 블럭 단위 변환을 이용한 압축 방법에서 발생하는 블럭화 현상을 제거하고 고정된 블럭 크기로 인한 제약점을 고려하여 높은 화질과 압축률을 얻기 위해서 웨이브

릿 변환 영역에서 가변 크기 블럭, 1/2 화소 다중 해상도 움직임 추정 및 보상(VMRME/C)과 다중 해상도 벡터 양자화(MRVQ)를 이용한 압축 방법을 제안하고 실험 결과를 보인다. 웨이브릿 변환은 블럭 단위 DCT 변환 코딩에서 발생하는 블럭 대비 효과를 크게 줄일 수 있으며, VMRMC는 웨이브릿 변환 영역에서 움직임 구조를 이용한 기법으로 웨이브릿 변환된 영상의 동일 해상도 내에서 움직임의 복잡도에 따라서 가변적인 블럭의 크기를 가진다. 그리고 인트라 프레임 코딩을 위하여 각 웨이브릿 변환 영역의 특성을 반영한 MRVQ를 적용하였다.

본 논문의 구성은 다음과 같다. 제2장에서 제안한 비디오 코딩 기법에 대하여 설명하며, 제3장에서는 실험을 통한 제안한 시스템의 성능을 분석하고, 제4장에서 결론 및 앞으로의 연구 방향을 제시한다.

II. 웨이브릿 변환을 이용한 초저속 전송 매체용 비디오 코딩

본 장에서는 제한된 전송 비트율로 인하여 발생하는 문제점을 해결하기 위해 웨이브릿 변환 영역의 동일 해상도 내에서 VMRME/C를 사용하는 비디오 코딩 시스템을 제안한다. 비디오 코딩 시스템의 구조는 그림 1과 같다. 압축 시스템은 블럭화 현상을 제거하기 위한 웨이브릿 변환 부분, 시간적 중복성을 줄이는 움직임 추정 및 보상 부분, 비트율을 줄이기 위한 벡터 양자화 부분, 그리고 코딩 모델의 통계적 중복성을 제거하는 엔트로피 코딩 부분으로 나눌 수 있다. 본 논문에서는 고압축 고품질의 복원 영상을 얻기 위하여 VMRMC, 해상도별 비트 할당을 통한 MRVQ를 제안한다.

본 논문에서 영상의 다중 해상도 변환을 위하여 표 1과 같은 계수를 갖는 Daubechies가 제안한 웨이브릿 9-3 filter[6]를 사용하였다. 이 필터는 대칭적인 성질을 가지고 있으며 이것은 기저 함수의 미분 차수가 증가, 즉 규칙성이 증가된다. 영상 압축 관점에서 보면 영상을 2차원 신호로 볼 경우 신호에 잘 근사하는 기저 함수가 되며, 작은 수의 기저 함수에 많은 에너지를 모을 수 있다. 따라서 엔트로피가 감소하게 되고 에너지가 압축이 용이한 형태로 일정 구역에 몰리게 된다[7]. 일차원 필터를 가로와 세로로 적용하므로

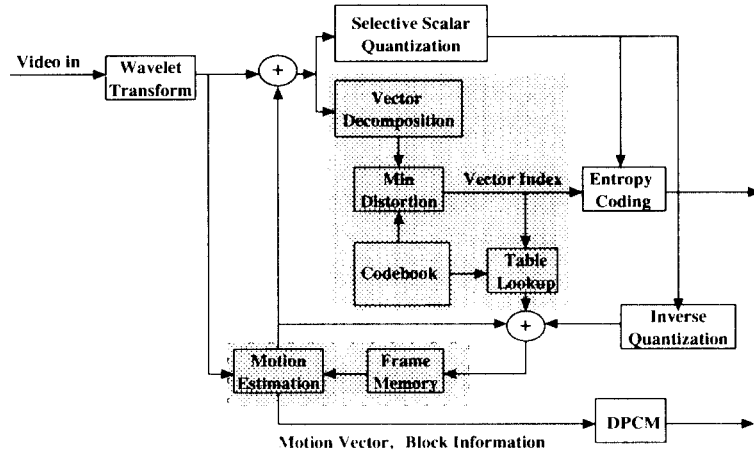


그림 1. 제안된 초저속 전송용 압축 기법의 구조

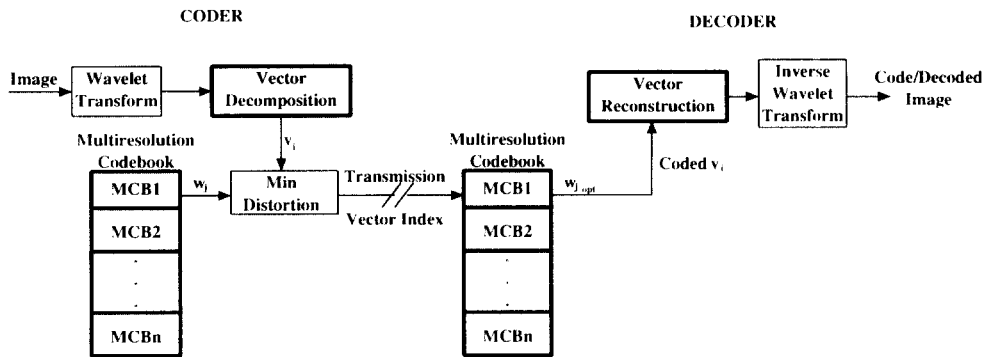


그림 2. 다중 해상도 벡터 양자화 기법

하나의 영상이 그림 3과 같이 4개의 부 영상으로 나누어지며 각 부 영상이 포함하는 주파수 특성이 다르다. 왼쪽 윗부분은 대부분의 에너지가 집약된 저주파 영역이며 오른쪽 윗부분의 영상은 주로 수평 에지 성분을 많이 포함하는 고주파 영역, 왼쪽 아랫부분의 영상은 수직 에지 성분을 많이 포함하는 고주파 영역이 된다. 본 논문에서는 저주파 성분에 2 번의 웨이브릿 변환을 적용하여 그림 3과 같은 다중 해상도 영상을 얻는다.

표 1. Daubechies 9-3 filter의 계수

n	0	±1	±2	±3	±4
$2^{-1/2} \hat{h}_n$	45/64	19/64	-1/8	-3/64	3/128
$2^{-1/2} \hat{h}_n$	1/2	1/4	0	0	0

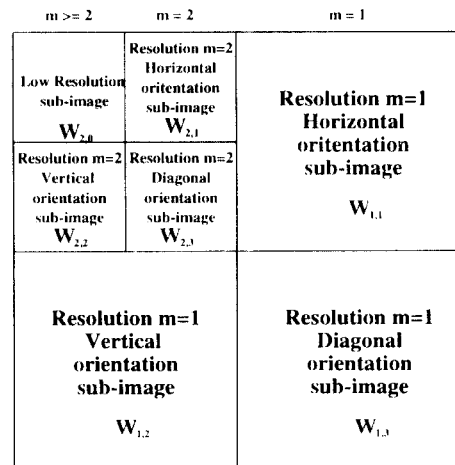


그림 3. 웨이브릿 변환을 통한 영상의 다중 해상도 분해

1. 다중 해상도 벡터 양자화

• 다중 해상도 벡터 양자화(MRVQ)

MRVQ는 Antonini[6]에 의하여 제안 되었으며, 코딩 부분은 영상을 분해하여 벡터로 만들어 주는 부분, 코드북, 그리고 입력 벡터와 코드북에서 최소의 차이를 갖는 코드벡터를 찾는 부분으로 이루어져 있다. 복호화 부분은 코드북과 코드북의 인덱스를 받아서 해당되는 다중 해상도 코드북에서 코드벡터를 찾아 영상을 재구성하는 부분으로 구성되어 있다. 본 논문에서는 고압축을 얻기 위하여 Antonini와 달리 저주파 부 영상에서도 벡터 양자화를 적용하였다. 또한 MRVQ에서 입력 영상을 벡터로 만들 때 해상도의 특성을 반영하도록 벡터의 차원과 입력 영상의 분해를 특징 추출에 적합하도록 하였다. 즉, 저주파 영역의 영상에 대해서는 벡터의 차원을 줄이고 고주파 영역에서는 임계값을 부여하여 많이 나타나는 에지의 특성을 고려하여 영상을 벡터로 만든다. 벡터화된 영상과 자신이 속한 해상도 코드북을 비교하여 Euclidian distance가 가장 작은 코드벡터를 찾아 그에 해당하는 해상도 구분자와 인덱스를 전송하므로 일종의 분류 벡터 양자화와 비슷한 성격을 띤다. 한편 복원하는 쪽에서는 역과정을 거쳐서 영상을 재 구성할 수 있다. 그러나 전송하는 비트들의 순서로서 해상도 구분이 가능하므로 해상도 구분자에 대한 비트율을 줄일 수 있다. 즉, 벡터 양자화기의 입력 벡터 x 가 해상도 i 에 속하면 코드북 $MCB(i)$ 에 있는 코드벡터 \hat{x}_j 와의 거리 $dist = \|x - \hat{x}_j\|$ 를 가장 작게 하는 인덱스 j 구하여 전송/저장한다.

• 비트 할당

다중 해상도로 변환된 영상에 벡터 양자화를 적용하기 위해서 각 해상도마다 영상의 특징을 잘 반영하도록 부 영상을 벡터로 나누는 것이 필요하다. 초저속 전송용 압축 기법에서는 비트율이 제한 되어 있기 때문에 다중 해상도로 변환된 각 영상에 비트를 어떻게 할당하는가 하는 것이 중요하다. 다중 해상도에서는 눈의 마스크(masking) 효과를 이용하며, 해상도 레벨과 윤곽선 등에 따라 코딩의 타입을 선택할 수 있다. 에너지가 많이 집약된 저주파 영상에서는 벡터의 차원을 낮게하여 코드북의 코드벡터와의 왜곡(distortion)을 줄여 화질 저하를 막고, 고주파 영역에서는

수평/수직 에지 성분의 치중 정도에 따라 에지의 특성을 반영하도록 가로 세로의 길이가 서로 다른 블록으로 분해하여 벡터를 형성한다. 다중 해상도로 분리된 영상은 각 부분이 복원된 영상에 미치는 주관적인 화질의 정도가 다르므로 각 해상도마다의 비트 할당이 서로 다르다. 본 논문에서는 [6]에서와 같이 왜곡 기준(distortion measure)으로 가중치를 부여한 MSE(mean square error)를 사용하였다. [6]에서는 저주파 대역과 고주파 대각선 영역에 대하여 각각 텍스처 코딩과 0비트를 할당하여 코딩을 하였다. 그러나 본 논문에서는 모든 해상도에 대하여 벡터 양자화를 적용하였으며, 동일 해상도는 같은 가중치를 갖는 왜곡 기준을 사용하였다. 즉, 총 MSE 왜곡을 $D^*_T(R_T)$, 총 비트율을 R_T , 그리고 해상도가 m , 방향이 d (horizontal, vertical, diagonal orientation)인 부 영상에서 비트율이 $R_{m,d}$ 라고 하면 평균 왜곡은 식(1)과 같이 정의할 수 있다. $q(x)$ 는 부 영상에 속하는 모든 계수 x 에 대한 양자화 값이다.

$$D_{m,d}(R_{m,d}) = E(|x - q(x)|^c) = d(x, q(x)), c \geq 1 \quad (1)$$

총 비트율이 R_T 라고 했을 때 영상의 총 왜곡은 식(2)로 주어진다.

$$D_T(R_T) = \frac{1}{2^{2M}} D_{M,0}(R_{M,0}) + \sum_{m=1}^M \frac{1}{2^{2m}} \sum_{d=1}^3 D_{m,d}(R_{m,d}) \quad (2)$$

$\frac{1}{2^{2m}} D_{M,0}(R_{M,0})$ 는 저주파수 영역에 대한 왜곡이며, 고주파 영역에서는 해상도 m 과 방향 d 에 따라서 왜곡이 $\sum_{m=1}^M \frac{1}{2^{2m}} \sum_{d=1}^3 D_{m,d}(R_{m,d})$ 와 같이 주어진다. 다중 해상도의 부 영상의 벡터 양자화에 대한 최적의 비트 할당 문제는 다음과 같이 수식화된다.

$$\begin{aligned} \text{Min}_{R_{m,d}, R_{m,d}} [D^*_T(R_T) &= \frac{1}{2^{2M}} D_{M,0} R_{M,0} \times B_{M,0} \\ &+ \sum_{m=1}^M \frac{1}{2^{2m}} \sum_{d=1}^3 D_{m,d}(R_{m,d}) \times B_{m,d}] \quad (3) \end{aligned}$$

$$\text{subject to } R_T = \frac{1}{2^{2m}} R_{M,0} + \sum_{m=1}^M \frac{1}{2^{2m}} \sum_{d=1}^3 R_{m,d} \quad (4)$$

즉, 주어진 총 비트율에서 왜곡을 최소로 하는 각 해상도의 비트율을 구하는 것이다. 식(3)과 (4)를 [6]과 같은 방법으로 lagrangian multiplier 기법을 도입하여 전개하면

$$R_{M,0_m} = R_T + \frac{1}{c} \log_2 \left[\frac{1}{(\alpha_{M,0}(\rho, c) B_{M,0})^{\frac{1}{4^M}-1} \prod_{m=1}^M \prod_{d=1}^3 (\alpha_{m,d}(\rho, c) B_{m,d})^{\frac{1}{4^m}}} \right] \quad (5)$$

$$R_{m,d_m} = R_T + \frac{1}{c} \log_2 \left[\frac{\alpha_{m,d}(\rho, c) B_{m,d}}{(\alpha_{M,0}(\rho, c) B_{M,0})^{\frac{1}{4^M}} \prod_{m=1}^M \prod_{d=1}^3 (\alpha_{m,d}(\rho, c) B_{m,d})^{\frac{1}{4^m}}} \right] \quad (6)$$

와 같이 된다. 각 수식에 사용된 항목들은 [6]에서 사용된 것과 동일하다. 식(5)와 (6)을 기반으로 다해상도 벡터 양자화를 이용하여 [6]에서와 같이 코딩할 때 프레임에 할당된 총 비트수를 식에 의하여 각 부 영상에 할당하고, 이에 따라서 각 부 영상의 코드 벡터 크기와 차원을 결정한다.

• 코드북의 생성

해상도별로 최적의 비트를 할당한 후 각 해상도의 특성을 잘 반영하는 코드북의 구성이 벡터 양자화의 성능에 큰 역할을 한다. 본 논문에서는 코드북 생성을 위해 Learning Vector Quantization Program Package 를 이용하였다[9]. 코드북의 크기가 i , 초기 코드북을 m_i 라 했을 때 학습을 위한 훈련 벡터 x 를 식(7)의 입력 벡터와 가장 근사한 코드북을 찾는 기준에 적용하여, 이 식에 의하여 구해진 벡터 m_i 를 식(8)과 같이 계속 수정해 간다.

$$c = \min_i \|x - m_i\| \quad (7)$$

$$m_i(t+1) = m_i(t) + \alpha(t)[x(t) - m_i(t)] \quad (8)$$

여기서 $0 < \alpha(t) < 1$ 이며, $\alpha(t)$ 는 시간에 대하여 상수이거나 단조 증가하는 함수이다[9].

2. 가변 크기 블럭 움직임 추정(VMRME)

제안된 움직임 추정 방법은 블럭 정합 알고리즘에 기반을 두고 정확한 움직임과 최소의 비트를 할당하기 위해서 웨이브릿 변환을 이용한 다중 해상도 분해된 영상의 동일 해상도 내에서 가변 크기의 블럭, 1/2 화소 MRME/C, 그리고 계층 내 및 계층 간의 예측을 특징으로 한다.

비디오 코딩에서 실시간 구현의 용이성, 알고리즘의 복잡도, 필요한 계산량을 등을 고려하여 움직임 추정은 블럭 정합 알고리즘을 사용한다. 그러나 고정된 크기의 블럭 단위마다 움직임 벡터를 추정하므로 영상의 국부적인 변화에 따른 움직임을 효과적으로 반영하기 어렵다. 블럭 정합 알고리즘에서 효과적인 움직임을 반영하기 위해서 블럭의 크기를 작게 하면 정확한 움직임을 추정할 수 있지만 움직임 정보의 전송 부담이 커진다. 이러한 문제점을 해결하기 위해서 움직임 추정의 성능과 움직임 정보의 전송 부담간의 타협으로 가변 블럭 크기 모델을 이용한다. 움직임이 같은 영역 내에서는 큰 블럭을, 여러가지 움직임이 있는 복잡한 영역에서는 작은 블럭 크기를 할당하므로 고정된 블럭 크기에서 발생하는 블럭화 현상을 어느 정도 줄일 수 있으며, 작은 움직임 벡터 수에서 나은 화질을 얻을 수 있다. 반면에 가변 크기의 블럭으로 나눈 정보를 가지고 있어야 하므로 비트 손실을 가져오지만 가변 블럭의 정보를 이진 트리를 이용하여 트리의 각 노드에 한 비트를 할당하므로 움직임 벡터 수의 감소에 비하여 비트 증가량은 미약하다. 블럭의 분해는 먼저 가장 큰 블럭의 크기와 가장 작은 블럭의 크기를 결정된 후 최소한의 블럭 분해가 발생하도록 가로/세로 또는 세로/가로로 나눈다. 주어진 블럭을 나누는 것은 다음의 알고리즘에 의하여 수행된다.

블럭 분해 알고리즘

- block size: $m \times n$
- $I_c(x, y), I_r(x, y)$: luminance of pel (x, y) at current and reference frame
- $|\cdot|$: absolute operator
- matching criteria: MAE(mean absolute error)
- SAD: sum of absolute difference

$$SAD(x, y) = \sum_{i=1}^m \sum_{j=1}^n |I_c(i, j) - I_r(i-x, j-y)|$$

$$SAD_1(x_1, y_1) = \sum_{i=1}^{m/2} \sum_{j=1}^n |I_c(i, j) - I_r(i-x_1, j-y_1)|$$

$$SAD_2(x_2, y_2) = \sum_{i=m/2}^m \sum_{j=1}^n |I_c(i, j) - I_r(i-x_2, j-y_2)|$$

or

$$SAD_1(x_1, y_1) = \sum_{i=1}^m \sum_{j=1}^{n/2} |I_c(i, j) - I_r(i-x_1, j-y_1)|$$

$$SAD_2(x_2, y_2) = \sum_{i=1}^m \sum_{j=n/2}^n |I_c(i, j) - I_r(i-x_2, j-y_2)|$$

IF ($SAD(x, y) < \alpha(SAD_1(x_1, y_1) + SAD_2(x_2, y_2))$)

THEN no block partition

ELSE block partition

블럭 분해 정보는 이진 트리 데이터 구조로 표현되며, 현재 블럭이 나누어 지면 0을, 이상 분해되지 않으면 1을 할당한다. 블럭 정보의 저장은 이진 트리의 깊이 우선 탐색에 의하여 방문되는 순서에 의하여 이루어진다. 한편, 알고리즘에서 각 부 영상별 웨이브릿 변환 계수의 중요도에 따라서 α 를 다르게 설정하므로 복원된 영상에 영향을 미친다. 즉 저해상도에서는 α 를 작게 하므로 작은 블럭으로 나누어 화질 향상에 기여하고 고해상도에서는 크게 하므로써 움직임 벡터의 수를 줄인다. 이렇게하여 웨이브릿 변환의 특성을 이용하여 적은 비트율로써 영상을 표현할 수 있다.

일반적으로 연속된 두 프레임 사이에 움직임을 추정하는 경우 움직임의 최소 크기를 화소 단위로 한다. 그러나 더욱 정확한 움직임을 찾기 위해서 화소와 화소 사이의 값을 보간법으로 구하여 움직임을 추정할 수 있다. 특히 이웃한 화소 간의 값의 차이가 클 때 1/2 또는 1/4 화소까지의 움직임을 정확히 추정하므로 복원된 영상의 화질을 높일 수 있다. 그림 4의 (a)는 1/2 화소 단위까지 움직임을 찾는 것을 보여준다. 블럭 정합 알고리즘을 이용하여 1/2 화소 단위까지 움직임을 찾기 위해서 먼저 화소 단위로 움직임 벡터를 구한 후 이 움직임 벡터가 가리키는 화소를 중심으로 1/2 화소 만큼 이동한 8개의 1/2 화소(그림에서 빗금친 사각형)에 대하여 움직임 벡터를 구하기 위한 거리 함수의 기준에 가장 적합한 화소를 선택하므로 1/2 화소까지의 움직임을 추정할 수 있다[10]. 이때 실제 화소가 아닌 화소의 값은 그림 4의 (b)와 같이 인접한 화소 값의 보간으로 구하여 사용한다. 그러므로 인접한 화소 간 값의 차이가 클 경우 이들 사이에 있는 1/2 화소는 실제 값을 가진 화소들의 평균이 됨으로 움직임 벡터를 찾고자 하는 블럭의 값과 정합이 잘 될 수 있다.

발생하는 블럭화 현상을 줄이고 움직임 벡터 수와 계산량, 그리고 정확한 움직임 추정을 위해서 웨이브릿 변환 및 분해를 통하여 다중 해상도로 구성된 영상에서 가변 크기 블럭의 MRME를 한다. 먼저 해상도가 가장 낮은 곳에서 전체적인 움직임을 찾고 상관성이 깊은 다음 계층에서 상위 계층에서 찾은 움직임

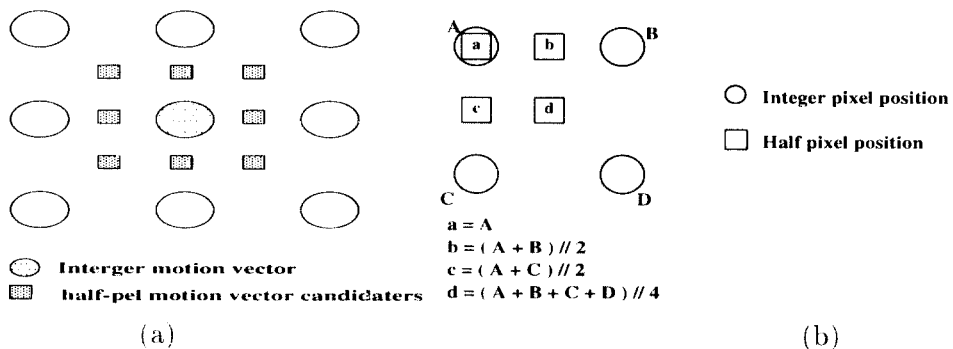


그림 4. (a) 1/2 화소 움직임 추정[10], (b) 1/2 화소 움직임을 위한 보간법[10]

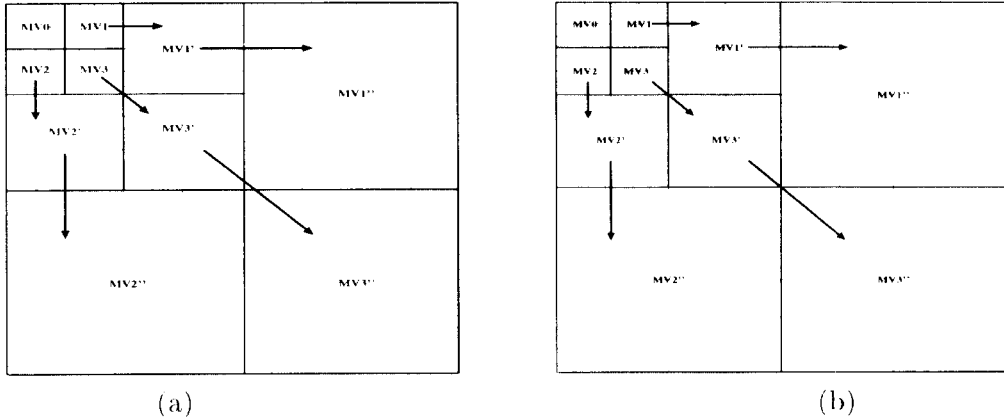


그림 5. (a)다중 해상도 움직임 추정 시 고정 블록 크기와 움직임 벡터의 예측, (b)가변 크기 블록의 다중 해상도 움직임 벡터의 예측

벡터를 이용하여 초기 벡터 $MV_{m-1, i} = 2 \cdot MV_{m, i}$ 와 같이 예측한 후 세부적으로 움직임을 추정한다. 이때 움직임 추정은 1/2 화소 단위까지 하며, 상위 계층에서 정확한 움직임을 찾기 위해서 가변 크기 블록으로 움직임을 찾는다. 상위 계층의 가변 블록 정보는 다음 계층에서 사용된다. 그림 5의 (a)는 Zhang이 제안한 고정 크기 블록 MRME 기법을 보여주며, (b)는 제안한 VMRME/C 기법을 보여준다. 이러한 움직임 추정을 통하여 구해진 움직임 벡터와 가변 블록 정보는 통계적 중복성을 줄이기 위해 엔트로피 코딩을 하여 전송된다.

움직임 보상 후의 잔류 오차는 양자화 후 run-length encoding과 허프만 코딩을 통하여 압축이 된다. 잔류 오차의 값이 0을 중심으로 정규 분포를 따르며, dead-zone을 갖는 양자화를 통하여 비트율을 줄일 수 있다.

III. 성능 평가 및 분석

제안한 비디오 코딩 시스템은 전송매체의 대역폭이 10kbps 이하인 초저속 전송망을 대상으로 하며, 입력 영상은 128 lines/frame, 160 pels/line, 5 frames/sec로 밝기 성분만 고려한다. 실험을 위해 표준 영상 Miss America, Clare, 그리고 Alex와 카메라로 영상 전화기용 영상 즉 사람의 흉상을 중심으로 하는 영상을 초

당 5 프레임씩 받아서 사용하였다. 실험 영상 중 초당 5 프레임 추출한 Miss America 영상과 Hwang-seok 비디오 영상을 구현된 시스템의 성능을 평가하기 위한 데이터로 사용하였으며, Clare, Alex, 그리고 카메라를 이용하여 받은 나머지 영상들을 프레임 내부에서 코딩 시 해상도별로 코딩 블록을 만들 때 학습을 위한 데이터로 사용하였다. 그림 6은 제안된 시스템의 성능 평가를 위해 사용된 영상의 첫번째 프레임이며, 각각 2단계 웨이브릿 변환 및 분해를 통하여 형성된 다중 해상도 분해된 영상을 보여준다.

코딩된 영상 화질은 아래와 같이 정의 되는 PSNR에 의하여 평가되며,

$$PSNR = 10 \cdot \log_{10} \frac{255^2}{MSE} \quad (dB), \quad MSE: \text{mean square error} \quad (9)$$

움직임 추정의 성능은 예측 이득(predictive gain: PG)으로 평가한다.

$$PG(dB) = 10 \cdot \log_{10} \frac{255^2}{\sigma_{DEP}^2} \quad (10)$$

예측 이득은 화소의 최대값 제공에 대한 잔류 오차의 분산비로 예측 이득이 클수록 압축이 용이하다. 또한 압축율은



그림 6. 시스템 성능 분석에 사용된 실험 영상

$$\text{compression ratio} = \frac{\text{number of bits for original image}}{\text{number of bits for compressed image}} \quad (11)$$

로 정의된다.

프레임 내부로 위한 MRVQ의 성능을 평가하기 위해서 일반적 많이 알려져 있는 블럭을 기반으로 DCT 및 스양자화를 하는 JPEG baseline 코딩과 웨이브릿을 한 후 해상도 별로 스칼라 양자화 및 VLC(variable length coder)와 성능을 비교하였다. 그림 7은 Hwang-seok과 Miss America에 대한 압축률과 c의 원영상에 대한 PSNR을 보여준다.

Hwang-seok 양서는 제안한 MRVQ가 압축률이 25 이하일 경우 약 2dB 이상 PSNR의 향상을 보이며, 그 이상의 고축률일 경우는 더욱 큰 화질의 차이(1.5-2dB)가 있을 수 있다. 그러나 Miss America 영상에서는 양율이 낮을 때는 JPEG baseline과 MRVQ가 비슷하지만 약간 뒤지지만 압축률이 25 이상의 경우 MRVQ우수함을 볼 수 있다. 이와 같은

성능을 나타내는 이유는 Miss America 영상은 화소의 평균이 66, 표준 편차가 30이며 Hwang-Seok 영상에서는 화소 평균이 165, 표준 편차가 60으로 JPEG baseline으로 압축을 할 경우 Miss America의 경우 DC계수에 모든 에너지가 집중되고 AC계수들은 대부분 0 또는 0에 가까운 값이다. 그러므로 AC 계수가 양자화되어도 큰 화질 저하가 발생하지 않으며, 반면에 Hwang-seok 영상에서는 0이 아닌 AC 계수가 많이 발생하며 고압축이 될 경우 양자화에 의하여 왜곡이 크게 된다. 그래서 복원된 영상 화질에 큰 영향을 미친다. 고압축 측면에서 보면 제안한 MRVQ가 기존의 블럭 단위 변환 알고리즘보다 나은 성능을 지닐 수 있다.

움직임 추정 및 예측의 성능을 평가하기 위해서 Zhang이 제안한 동일 해상도 내에서 고정 크기 블럭을 이용한 MRME/C 방법과 비교하였다. 또한 움직임 추정시 가변 블럭의 효과와 1/2 화소 움직임 추정 및 예측의 효과를 알아보기 위해서 각각을 분리하여 비교하였다. 그림 8과 9는 움직임 추정시 예측 이득을 보여준다.

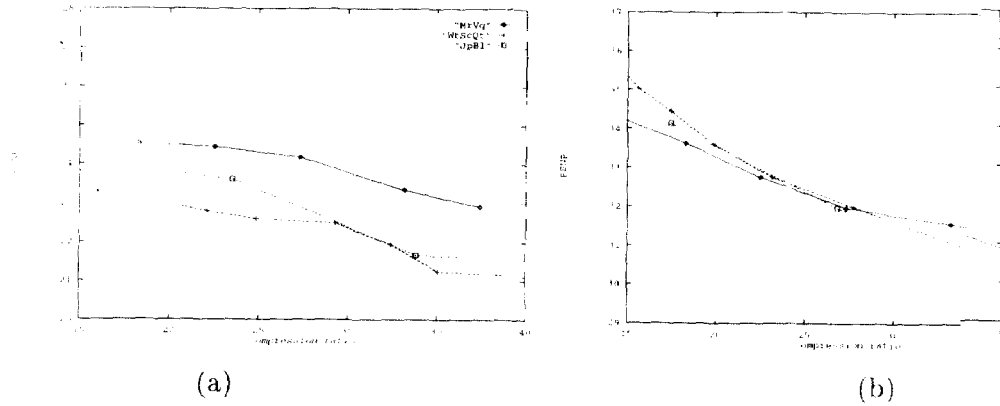


그림 7. Hwang-seok 영상(a)과 Miss America 영상(b)의 첫번째 프레임에 대한 JPEG baseline (JpBl), 웨이환 및 스칼라 양자화(WtScQt), 그리고 MRVQ(MrVq)의 압축률과 PSNR

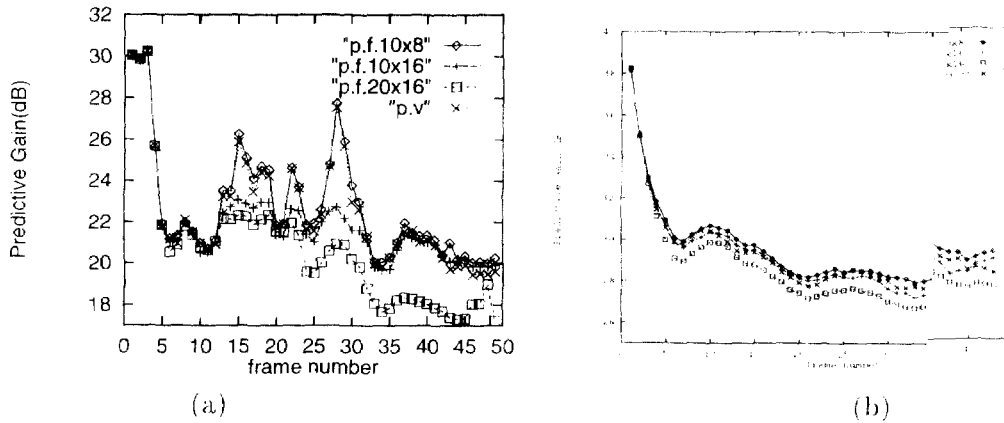


그림 8. 다중 해상도 피라미드 구조에서 Hwang-seok 비디오 영상(a)과 Miss America 비디오 영상(b)의 임 추정시 예측 이득(P: 계층간의 예측, h: 1/2 화소 움직임 추정, f: 고정 크기 블럭, v: 가변 크기 블럭, mxn: 블럭 크기)

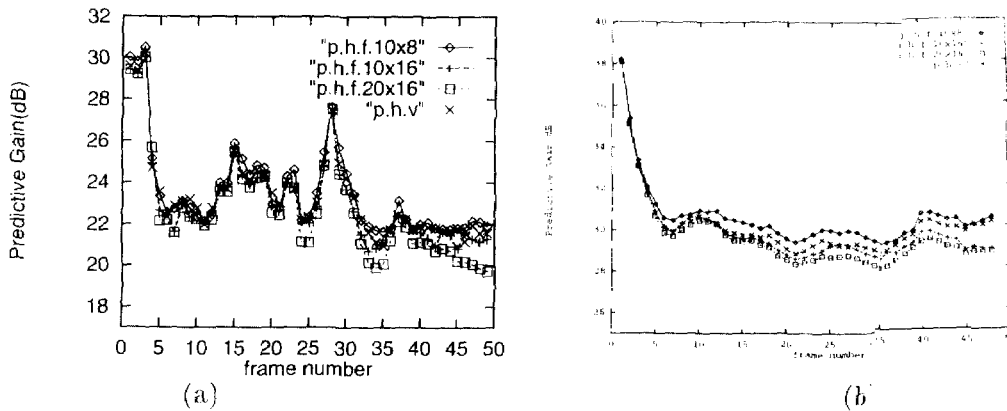


그림 9. 다중 해상도 피라미드 구조에서 Hwang-seok 비디오 영상(a)과 Miss America 비디오 영상(b) 1/2 화소 움직임 추정시 예측 이득(P: 계층간의 예측, h: 1/2 화소 움직임 추정, f: 고정 크기 블럭, v: 가변 크기 블럭, mxn: 블럭 크기)

그리고 표 2, 3, 4는 50 프레임까지 움직임 추정을 한 후 한 프레임당 평균 블럭 수와 예측 이득을 보여준다. 정수 화소 단위에서는 VMRME/C가 MRME/C 보다 같은 수의 움직임 정보를 사용하면서 Hwang-seok 영상에서는 약 0.5 dB가, Miss America 영상에서는 0.2 dB가 향상됨을 알 수 있다.

표 2. Hwang-seok 비디오 영상에 대한 정수 화소 움직임 추정 시의 평균 블럭 수와 평균 예측 이득(dB)

움직임 추정 블럭의 크기(X×Y)	10×8	10×16	20×16	가변 크기
평균 블럭수	112	56	28	51
평균 예측 이득	22.7392	22.0684	20.9180	22.5877

표 3. Miss America 비디오 영상에 대한 정수 화소 움직임 추정 시의 평균 블럭 수와 평균 예측 이득

움직임 추정 블럭의 크기(X×Y)	10×8	10×16	20×16	가변 크기
평균 블럭수	112	56	28	47
평균 예측 이득	29.5213	29.1083	28.4605	29.3223

표 4. Hwang-seok 비디오 영상과 Miss America 비디오 영상에 대한 정수 화소와 1/2 화소 움직임 추정 시의 평균 예측 이득 비교(정수 화소 단위: I, 1/2화소 단위: H)

움직임 추정 블럭의 크기(X×Y)	10×8		10×16		20×16		가변 크기	
	I	H	I	H	I	H	I	H
Hwang-seok	22.73	23.13	22.06	23.13	20.91	22.77	22.58	23.41
Miss America	29.52	30.43	29.10	29.74	28.46	29.45	29.32	30.03

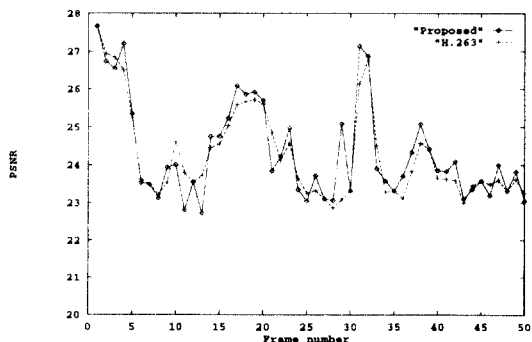
표 4에서는 정수 화소 단위에서 움직임 추정 결과와 1/2 화소 단위에서 움직임 추정 결과를 보여준다. 위 표로부터 가변 크기 블럭의 효과가 1/2 화소 움직임 추정보다 정수 화소 움직임 단위일 경우 크게 나타남을 볼 수 있다. 즉 가변 블럭 크기 움직임 추정과 1/2 화소 단위 움직임 추정 및 예측을 할 경우 상방의 효과가 서로 상쇄되어 전체적으로는 성능 향상을 가져 오지만 각각의 성능의 선형함은 되지 않는다. 그림 10에서는 9600bps 비트율에서 H.263 기본 모드 코딩 결과와 제안한 시스템의 결과를 비교 하였다. 각 실험 영상에서 제안한 시스템의 화질이 H.263의 기본 모드 결과보다 보다 평균적으로 0.2-0.6 정도 우수함을 보여준다.

표 5. 비트율 9600bps에서 실험영상에 대한 제안한 시스템과 H.263 기본 모드의 평균 PSNR

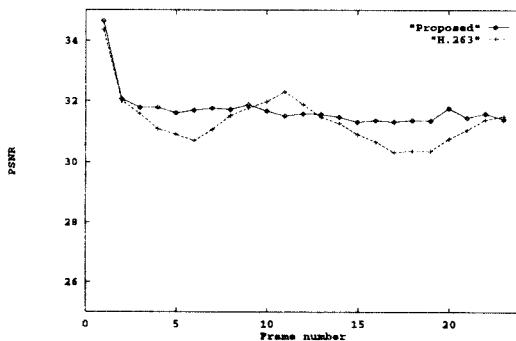
실험 영상	Hwang-Seok	Miss America
Proposed	24.5658	31.8256
H.263	24.3738	31.0578

IV. 결론 및 향후 연구 방향

본 논문에서는 기존의 블럭을 기반으로 하는 변환을 이용한 비디오 압축 시스템의 블럭화 현상을 제거하고, 연속된 프레임간의 정확한 움직임 추정과 정보량을 줄이기 위해서 신호 표현과 인간 시각 특성에



(a)Hwang-seok 영상에 대한 PSNR.



(b)Miss America 영상에 대한 PSNR

그림 10. 비트율 9600bps에서 제안한 시스템과 H.263 기본 모드 코딩 결과

적합한 웨이브릿 변환을 이용하여 초저속 전송율에 적합한 압축 기법을 제안하였다.

제안한 압축 시스템의 프레임 내의 중복성 제거를 위한 MRVQ는 JPEG 기본 알고리즘과 웨이브릿 변환 다중 해상도 피라미드에서 스칼라 양자화를 통한 엔트로피 코딩보다 초저속 전송을 위한 고 압축률일 때 성능이 나음을 보였으며, 프레임 간의 중복성 제거를 위한 코딩에서 VMRME/C는 Zhang이 제안한 MRME/C 보다 실험 영상에서 성능이 나음을 보였다. 그리고 전송율 9600bps 제한 조건에서 화질 변화를 H.263 기본 모드와 비교하였다.

앞으로 더욱 연구되어야 할 부분은 초저속 전송율을 갖는 전송로에서 코딩된 비트열의 동기화 및 제어 정보를 위한 효율적 구문 정의, 영상에 관계없이 일정한 화질을 유지하기 위한 비트율 제어 부분에 많은 연구가 있어야 한다.

참 고 문 헌

1. R.Scharphorst and R. for Q2/15, *ITU-T Recommendation H.263*, Oct. 1995.
2. KAIST, *ISO/IEC MPEG and CCITT ATM Video Coding Algorithms*. KAIST, July 1992.
3. KAIST, *ISO/IEC MPEG and CCITT ATM Video Coding Algorithms*. KAIST, July 1993.
4. M. Rabbani and P. W. Jones, *Digital Image Compression Techniques*, SPIE-Optical Engineering Press, 1991.
5. K. Gyu-Taek, "A study on motion estimation algorithm combined pel-recursive method and matching method," Master's thesis, Dept. of CS, KAIST, 1993.
6. M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. on Image Processing*, vol. 1, pp. 205-220, Apr. 1992.
7. S. W. Kim, "A multiresolution based moving image compression system," Master's thesis, Dept. of CS, KAIST, 1993.
8. Y.-Q. Zhang and S. Zafar, "Motion-compensated wavelet transform coding for color video compression," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 2, pp. 285-296, Sept. 1992.
9. T. Kohonen, J. Kangas, J. Laaksonen, and K. Torkkola, *LVQ-PAK: The Learning Vector Quantization Program Package ver2.1*. Oct. 1992.
10. S. S. COST 211 ter, *Simulation model for very low bit rate image coding(SIM3)* Mar. 1993.
11. M. Chan, Y. Yu, and A. G. Constantinides, "Variable size block matching motion compensation with application to video coding," in *Proc. IEE*, vol. 137, pp. 205-212, Aug. 1990.
12. I. Distein, K. Tose, and A. Heiman, "Variable block-size transform image coder," *IEEE Trans. on Communications*, vol. 38, pp. 2073-2078, Nov. 1990.
13. D. L. Gall, "MPEG: A video compression standard for multimedia application," *Communication of the ACM*, vol. 34, pp. 46-58, Apr. 1991.
14. G. K. Wallace, "The JPEG still picture compression standard," *Communication of the ACM*, vol. 34, pp. 30-44, Apr. 1991.
15. M. Liou, "Overview of the px64 kbit/s video coding standard," *Communication of the ACM*, vol. 34, pp. 59-63, Apr. 1991.
16. Y.-S. Jehng, L.-G. Chen, and T.-D. Chiueh, "A motion estimator for low bit-rate video codec," *IEEE Trans. on Consumer Electronics*, vol. 38, pp. 60-69, May 1992.
17. A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. KAP, 1992.
18. B.-C. Kim, "A study on the corona effect of hierarchical motion estimation," Master's thesis, Dept. of CS, KAIST, 1992.
19. L.-W. Lee, J.-F. Wang, J.-Y. Lee, and J.-D. Shie, "Dynamic search-window adjustment and interlaced search for block-matching algorithm," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 3, pp. 85-87, Feb. 1993.
20. J. W. Kim and S. U. Lee, "Variable block size motion estimation techniques for the motion sequence coding," *KISS*, vol. 30, pp. 357-368, Apr.

1993.

- 21. A.C.P.Loui and .Liou, "High-resolution still-image transmission based on CCITT H.261 codec," *IEEM.LE Trans. on Circuits and Systems for Video Technology*, vol. 3, pp. 164-169, Apr. 1993.
- 22. I. J. MPEG93/457, *MPEG 2 Test Model 5 draft*. Apr. 1993.



吳 皇 錫(Hwang-Seok Oh) 정회원
 1969년 11월 17일생
 1992년 2월: 경북대학교 자연과학
 대학 전자계산학과
 졸업(학사)
 1994년 2월: 한국과학기술원 전산
 학과 졸업(공학석사)
 1994년 3월~현재: 한국과학기술원
 전산학과 박사과정

주관심분야: Video coding, Parallel processing for
 video coding, Multimedia service system



李 興 奎(Heung-Kyu Lee) 정회원
 1955년 10월 3일생
 1978년 2월: 서울대학교 전자공학
 과 졸업(학사)
 1981년 2월: 한국과학원 전산학과
 졸업(이학석사)
 1984년 8월: 한국과학기술원 전산
 학과 졸업(공학박사)

1985년 3월~1986년 8월: U. of Michigan, U.S.A (Re-
 search Scientist)

1986년 9월~현재: 한국과학기술원 전산학과 교수(부
 교수)

1991년~현재: 공업진흥청 JTC1/SC29(전문위원)

1994년~현재: ITU-T SG9(VOD 연구그룹 의장)

주관심분야: Real-time processing, Fault-tolerant sys-
 tem, Multimedia service system, Real-
 time operating system