

논문 2005-42CI-3-6

공간 웨이블릿 변환의 복잡도를 줄인 스케일러블 비디오 부호화에 관한 연구

(Scalable Video Coding with Low Complex Wavelet Transform)

박 성 호*, 정 세 윤**, 김 원 하***

(Seong-Ho Park, Se-Yoon Jeong, and Won-Ha Kim)

요 약

인터 프레임 웨이블릿 부호화 기법의 복호화 과정에서 많은 연산량을 차지하는 모듈 중의 하나는 웨이블릿 변환이다. 복호기는 PDA, PC, 휴대폰등과 같이 다양한 단말기 상에서 동작 할 수 있어야 하기 때문에 복호기의 복잡도는 각 프로세서의 계산 능력에 맞게 설계되어야 한다. 따라서 스케일러블 부호화를 위한 코덱 역시 낮은 복잡도로 설계되어야 한다. 본 논문에서는 부호화 성능을 열화시키지 않으면서 공간 웨이블릿 변환의 복잡도를 조절하면서 줄이는 기법을 제안한다. 또한 이 기법은 천천히 변화하는 영상 시퀀스에 대해서는 웨이블릿 변환 시 발생하는 잔상 현상도 줄일 수 있다.

Abstract

In the decoding process of interframe Wavelet coding, the Wavelet transform requires huge computational complexity. Since the decoder may need to be used in various devices such as PDAs, notebooks, or PC, the decoder's complexity should be adapted to the processor's computational power. So, it is natural that the low complexity codec is also required for scalable video coding. In this paper, we develop a method of controlling and lowering the complexity of the spatial Wavelet transform while sustaining the same coding efficiency as the conventional spatial Wavelet transform. In addition, the proposed method may alleviate the ringing effect for slowly changing image sequences.

Keywords: 스케일러블 동영상 압축, 웨이블릿 기반 동영상, 적응적 필터

I. 서 론

방송과 통신 융합이 됨에 따라 현재의 통신환경은 다양한 매체 및 기기 들이 서로 융합되어 있으며 앞으로 이러한 추세는 더욱 심화될 것이다. 이러한 비균일 통

신 환경에서 디지털 융합의 추세에 맞추어 가자면 동영상 전송도 범용 미디어 통신 (Universal Media Access, UMA)이 되도록 해야 한다.

Scalable Video Coding(SVC)의 한 분류인 웨이블릿 기반 스케일러블 동영상 부호화에서 영상 시퀀스는 움직임 보상 시간 필터링(motion compensated temporal filtering, MCTF)과정을 거친 다음 웨이블릿 변환 후 양자화하고 EZW나 EBCOT과 같은 방법으로 엔트로피 부호화한다.^{[1][2][3]}

기존의 웨이블릿 기반 SVC는 MCTF를 통과한 모든 프레임에 대하여 동일한 웨이블릿 필터를 사용하여 공간 웨이블릿 변환을 수행한 반면 본 논문에서는 MCTF를 통과한 결과로 나오는 영상 프레임의 스펙트럼 특성을 분석하여 특성에 맞는 공간 웨이블릿 필터를 사용하

* 학생회원, 경희대학교 전자공학과
(Electronic Engr. Major, KyungHee Univ.)

** 정회원, 전자통신연구원 디지털 방송연구단
대화형미디어 연구팀
(Digital Broadcasting Research Division, Interactive Media Dept., ETRI)

*** 정회원, 경희대학교 전자공학과
(School of Electronics and Information Electronic Engr. Major)

※ 본 연구는 과학 기술부 과학재단 목적기초연구
(R05-2003-000-11653)지원으로 수행되었음.

접수일자: 2005년1월25일, 수정완료일: 2005년5월5일

는 기법을 제안한다. MCTF를 수행한 후 프레임들의 특성에 따른 필터성능을 나타내는 척도로 부호화 이득 (Coding gain)을 적용하였다.^[4] 이를 위해 임의의 필터와 임의의 서브밴드 분할에도 적용할 수 있는 통합 부호화이득을 유도하였다. 유도한 부호화이득 공식을 바탕으로 프레임의 특성을 결정하여 Haar 필터와 9/7 필터를 적응적으로 사용하였다.

이 기법은 기존의 모든 프레임에 대해 동일한 필터를 적용하는 방법에 비해 변환 복잡도를 줄일 수 있으며, 에지성분이 많은 시간축으로 필터링된 프레임을 긴 필터로 변환함에 따라 발생하는 잔상 현상(ringing effect)을 감소시킬 수 있다.^[5]

본 논문의 내용은 다음과 같다. 제 I장에서는 MCTF를 간단히 소개하고 MCTF한 영상들의 특성을 분석한다. 제 II장에서는 통합 부호화이득을 유도하고, 유도된 부호화이득을 바탕으로 MCTF된 영상들의 부호화 성능을 분석한다. 제 III장에서는 MCTF된 영상들의 특징에 따라서 웨이블릿 필터를 선택하는 알고리즘을 제안한다. 제 IV장에서는 제안된 기법의 성능 및 복잡도를 논한다. 마지막 장에서는 본 논문의 전체를 요약하는 결론을 도출한다.

II. MCTF 된 프레임들의 영상 신호 특성

1. MCTF

(Motion Compensated Temporal Filtering)

MCTF는 GOF(Group Of Frames) 내 영상들의 시간적 중복성을 제거하고 영상 신호들의 전반적인 정보를

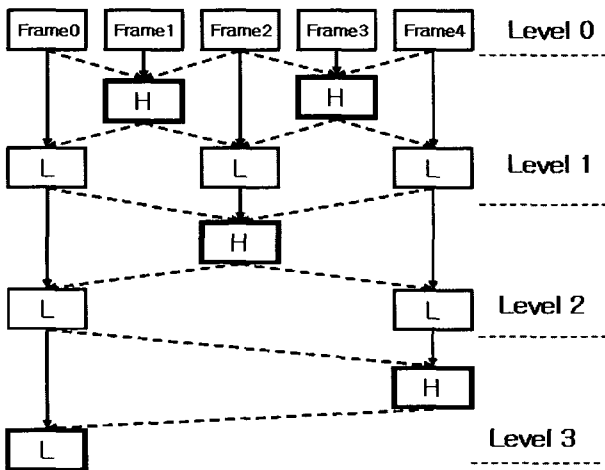


그림 1. 3 레벨 5/3-MCTF
Fig. 1. 3 level 5/3-MCTF.

L프레임이라고 하는 프레임에 모으고, 잔여 영상 신호 (residual image signal)들은 H프레임이라고 하는 프레임에 분산 시킨다. 이러한 MCTF 과정은 반복적으로 L프레임에 대하여 실시하여 한 GOF에서 하나의 L프레임을 생성하고 L프레임을 제외한 나머지 프레임들은 H프레임이 되도록 한다. MCTF를 통과한 L, H프레임들은 공간 웨이블릿 변환 후 부호화되어 전송된다.^{[1][2][3]}

주로 사용되는 MCTF는 Haar MCTF와 5/3 MCTF이다. Haar MCTF는 2개의 프레임으로 각각의 L, H프레임을 만들고 5/3 MCTF는 3개의 프레임으로 H프레임을, 5개의 프레임으로 L프레임을 만든다. Haar MCTF와 5/3 MCTF는 한 GOF 내에서 동시에 사용할 수 있다. 여기서 H프레임은 기존 MPEG의 B, P픽처와 같은 특성을 갖으며, L프레임은 연속적인 5 프레임의 평균적 영상 신호를 포함하고 있어서 기존 MPEG의 I 픽처와 같은 개념이며 거의 모든 에너지가 밀집되어 있다. 그림 1은 3레벨 5/3 MCTF의 과정을 나타낸다. 그림에서 실선은 프레임의 직접적인 흐름을 나타내며 점선은 움직임 보상을 적용하는 것을 나타낸다.

2. 프레임에 따른 영상 신호 분석

그림 2에 실제 'foreman'영상에서 GOF 크기가 5인 한 GOF가 MCTF를 통과한 결과를 나타내었다. 그림에서 보듯이 L프레임은 원래 영상 신호와 비슷하고, H프레임의 경우 참조 프레임과의 차이영상을 나타낸다.그

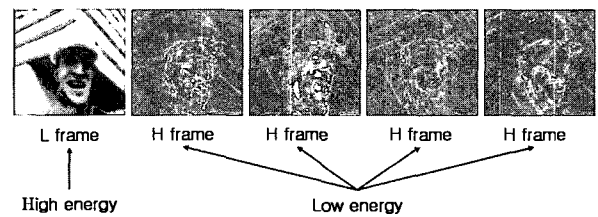


그림 2. MCTF를 통과한 'foreman' 영상의 한 GOF
Fig. 2. Result of GOF after MCTF.

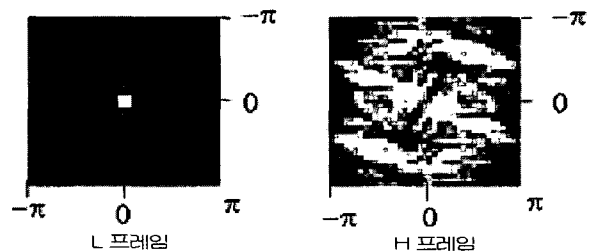


그림 3. L, H프레임의 2차원 PSD 영상
Fig. 3. 2-D PSD image of L and H frame.

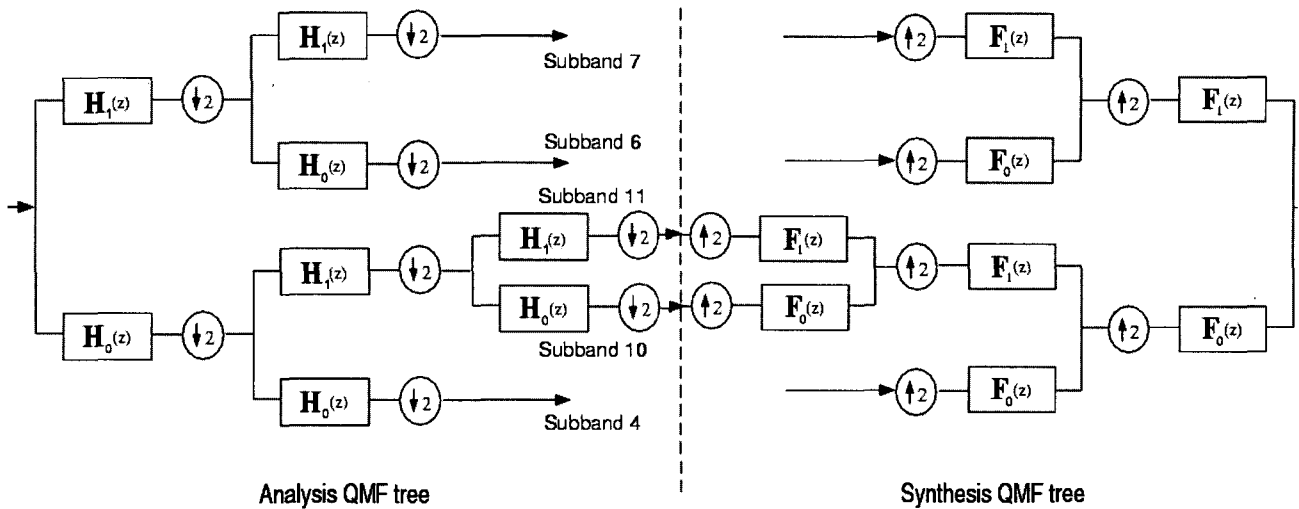


그림 4. QMF의 나무 구조
Fig. 4. Tree-structured connections of QMFs.

림 3은 각 프레임 신호의 스펙트럼 분포를 보기위해 전력 스펙트럼 밀도(Power Spectral Density, PSD)를 비교한 것이다. PSD를 보면 L프레임의 경우 저주파에 대부분의 에너지가 모여 있는 반면 H프레임의 경우 전 주파수 대역에 고루 분포되어 있는 것을 알 수 있다.

III. 프레임에 따른 부호화 이득 분석

웨이블릿 변환은 비균일(non-uniform) 서브밴드 분할을 구현한다. 웨이블릿 변환 후 영상 신호의 에너지는 서브밴드 별로 분할되는데 특정 서브밴드에 에너지가 밀집되어 있을 경우 주어진 대역폭에서 부호화 왜곡을 적게 할 수 있다. 따라서 서브밴드 분할된 신호의 에너지 밀집이 클수록 부호화 효율이 증가된다. 밀집 정도를 측정하는 척도로 일반적으로 부호화 이득을 사용한다.^[6]

영상 특성에 따른 웨이블릿 필터의 성능 분석을 위하여 스펙트럼 특성이 다른 영상 신호에 다양한 웨이블릿 필터를 적용하여 부호화 이득을 분석한다. 이를 위해 웨이블릿에 의해서 구현되는 서브밴드의 분할을 수학적으로 모델링 하고, 그 모델을 바탕으로 통합 부호화 이득 공식을 유도한다. 유도된 부호화 이득을 바탕으로 영상신호의 특성에 따른 필터의 성능을 분석한다.

1. 웨이블릿 기반 서브밴드 분할 모델링

웨이블릿 기반 비균일 서브밴드 분할은 QMF

(Quadrature Mirror Filter)가 연속적으로 연결된 트리 구조로 구현 된다.^[7] QMF 트리 구조는 이진 트리 구조로 표현된다. QMF 트리의 수학적 분석을 위해 다음과 같은 변수를 정의한다.

- $S(\text{Subband } i)$: QMF 트리의 노드들의 집합. 뿌리 노드는 1로 시작.
노드 i 의 자식노드는 각각 $2i, (2i + 1)$ 노드로 매긴다. 각 노드의 $2i$ 노드는 lowpass 필터링에, $(2i + 1)$ 노드는 highpass 필터링에 해당된다.
- $L(S)$: QMF 트리 S 의 leaf 노드 집합. 따라서 $|L(S)|$ 는 서브밴드의 수.
- d_i : 노드 i 의 깊이이다. 즉, $d_i = \log_2 i$, 노드 i 의 서브밴드 대역폭은 $\pi/2^{d_i}$.
- $h_i(n), H_i(z)$: 서브밴드 i 를 만드는 분해 필터와 그 응답 함수.
- $f_i(n), F_i(z)$: 서브밴드 i 의 합성 필터와 그 응답 함수.

그림 4에 트리 구조의 QMF의 예를 나타내었다. 그림에서 $H_0(z), H_1(z)$ 는 각각 저주파와 고주파 분해 (analysis) 필터의 전달함수이며 $F_0(z), F_1(z)$ 는 저주파와 고주파의 합성 (synthesis) 함수이다. 서브밴드 분

해/합성 필터는 완벽한 복원을 위해 다음 식을 만족한다.

$$H_0(z) = -F_1(-z), H_1(z) = F_0(-z).$$

$H_i(z)$ 와 $F_i(z)$ 는 다음과 같이 재귀적으로 계산된다^[7].

$$H_i(z) = H_{(i \bmod 2)}(z^{d_i}) \cdot H_{i/2}(z), \quad (1)$$

$$F_i(z) = F_{(i \bmod 2)}(z^{d_i}) \cdot F_{i/2}(z). \quad (2)$$

여기서 $i \geq 2$, $H_1(z) = 1$, $F_1(z) = 1$.

합성과 분해에서의 서브밴드 i 신호는 다음과 같다.

분해에서는,

$$\tilde{x}_i(n) = \sum_{m=-\infty}^{\infty} x(m) h_i(2^{d_i}n - m), \quad (3)$$

합성에서는,

$$\hat{x}_i(n) = \sum_{m=-\infty}^{\infty} \tilde{x}_i(m) f_i(n - 2^{d_i}m). \quad (4)$$

여기서 $x(m)$ 은 분해 쪽에서의 입력신호이고, $\tilde{x}_i(m)$ 은 합성 쪽에서의 i 번째 서브밴드 신호이다.

2. 통합 부호화 이득

앞 절의 모델링을 기반으로 임의의 서브밴드 분할과 임의의 필터에 대해서 적용 가능한 통합 부호화 이득 공식을 유도한다.

일반적으로 서브밴드 부호화(Subband Coding, SBC)의 서브밴드 분할을 QMF 트리의 노드 집합 S 로 표시할 수 있다.

여기서 S 의 leaf 노드 $L(S)$ 에 속하는 노드 i , 즉 $i \in L(S)$ 는 서브밴드 i 가 되며 그 서브밴드의 대역폭은 $\pi/2^{d_i}$ 이다. 그림 5는 하나의 서브밴드를 부호화하는 과정을 나타낸다. 서브밴드 신호 $x_i(n)$ 을 양자화 할 때 발생하는 양자화 에러 $q_i(n)$ 은 $x_i(n)$ 과는 독립적이고 평균이 0인 additive white process이다. $e_i(n)$ 은 양자화 에러 $q_i(n)$ 으로 발생한 서브밴드 i 의 복원 에러이다. 서로 다른 서브밴드의 양자화 에러가 독립적이

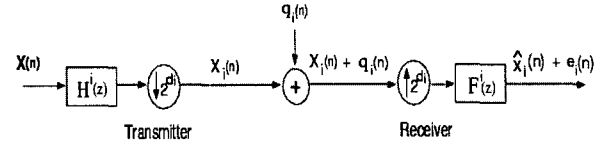


그림 5. 한 서브밴드의 부호화 모델

Fig. 5. Subband model of one channel.

고 평균이 0이기 때문에 복원 에러도 독립적이고 평균이 0이다. 따라서 서브밴드의 복원 신호는 다음과 같은 식으로 표현할 수 있다.

$$\hat{x}(n) = \sum_{i \in L(S)} \{\hat{x}_i(n) + e_i(n)\} = x(n) + e(n)$$

여기서 $x(n)$ 은 원래 신호, $e(n)$ 은 복원 에러이며,

$$E\{|e(n)|^2\} = \sum_{i \in L(S)} E\{|e_i(n)|^2\} \text{이다.}$$

다음으로 임의의 서브밴드 분할 S 에 대한 부호화 이득을 유도한다.

정리 1 : 입력 신호가 wide sense stationary (WSS) 일 때, 통합 부호화 이득 $G(S)$ 는 다음과 같다.

$$G_{SBC}(S) = \frac{\sigma_x^2}{\prod_{i \in L(S)} (\beta_i \sigma_{x_i}^2)^{1/2^{d_i}}} \quad (5)$$

여기서 $\beta_i = \sum_m |f^i(m)|^2$ 로 합성 필터의 에너지이고, σ_x^2 는 입력 신호의 분산, $\sigma_{x_i}^2$ 는 서브밴드 i 신호의 분산이다.

증명 : 그림 5에서처럼 σ_q^2 를 $\frac{\pi}{M_i} = \frac{\pi}{2^{d_i}}$ 의 크기를 갖는 서브밴드 i 의 양자화 에러 분산이며 $f^i(m)$ 는 서브밴드 i 를 얻기 위한 합성 필터이다. WSS 신호는 보간(interpolation) 후에 보간율의 주기(M_i)를 갖는 cyclo-stationary 신호가 되기 때문에 i 번째 서브밴드 채널로부터 복원된 에러 $e_i(n)$ 은 주기 M_i 를 갖는 cyclo-stationary process가 된다^[4].

에러에 대한 자기 상관함수 $R_{e_i}(n, n+m)$ 는 식 4로부터 다음과 같이 구한다.

$$\begin{aligned}
 R_e(n, n+m) &= E\{e_i(n)e_i(n+m)\} \\
 &= \sum_r \sum_s E\{q_i(r)q_i(s)\}f^i(n-rM_i) \\
 &\quad \cdot f^i(m+n-sM_i) \\
 &= \sum_r \sum_s R_{q,q_i}(r-s)f^i(n-rM_i)f^i(m+n-sM_i)
 \end{aligned}$$

그리고 주기 M_i 를 갖는 cyclo-stationary process의 평균 자기 상관함수 $\overline{R}_e(n)$ 는 다음과 같이 구한다.

$$\begin{aligned}
 \overline{R}_e(n) &= \frac{1}{M_i} \sum_{l=0}^{M_i-1} R_e(l, n+l) \\
 &= \frac{1}{M_i} \sum_{l=0}^{M_i-1} \sum_r \sum_s f^i(l-rM_i)f^i(l+n-sM_i) \\
 &\quad \cdot R_{q,q_i}(r-s) \\
 &= \frac{\sigma_q^2}{M_i} \sum_{l=0}^{M_i-1} \sum_r f^i(l-rM_i)f^i(l+n-rM_i).
 \end{aligned}$$

여기서 σ_q^2 는 각 i 서브밴드의 양자화 에러 분산이고, 두 번째 식에서 양자화 에러는 white process이므로, $R_{q,q_i}(r-s) = \sigma_q^2 \cdot \delta(r-s)$ 이다.

$E\{|e_i(n)|^2\} = \overline{R}_e(0)$ 이므로 $E\{|e_i(n)|^2\}$ 는 다음과 같이 구한다.

$$\begin{aligned}
 E\{|e_i(n)|^2\} &= \frac{\sigma_q^2}{M_i} \sum_{l=0}^{M_i-1} \sum_r |f^i(l-rM_i)|^2 \\
 &= \frac{\sigma_q^2}{M_i} \sum_m |f^i(m)|^2.
 \end{aligned}$$

여기서 $\sum_{i \in L(S)} \frac{1}{2^d} = 1$ 이므로, 산술 평균과 기하 평균의 부등식에 따라 에러 분산에 대한 경계 값은 다음과 같이 된다.

$$\begin{aligned}
 E\{|e(n)|^2\} &= \sum_{i \in L(S)} E\{|e_i(n)|^2\} \\
 &\geq \prod_{i \in L(S)} (\sigma_q^2 \sum_m |f^i(m)|^2)^{1/d_i} \quad (6)
 \end{aligned}$$

R_i 를 i 서브밴드에 할당된 비트수라 하면, 각 서브밴

표 1. 서브밴드 가중치

Table 1. Subband weight factor $\beta_i = \sum_m |f_i(m)|^2$.

서브밴드 i	2	3	4	5	6	7	8	9	10
β_i	0.98	1.04	0.97	1.02	1.02	1.08	0.95	1.07	0.98
서브밴드 i	11	12	13	14	15	16	17	18	19
β_i	0.97	1.07	0.99	1.04	1.05	1.00	1.09	0.96	1.01

드의 부호화 비트율은 초당 $1/2^d$ 샘플로 샘플링되기 때문에 $R_i/2^d$ 이며 전체 부호화 비트율은

$$R = \sum_{i \in L(S)} R_i/2^d$$

이다. Shannon의 rate distortion bound에 의하여 양자화 에러의 분산은 $\sigma_q^2 > 2^{-2R} \sigma_x^2$ 을 만족하여야 한다 [13],[14]. 따라서 식 5는 다음과 같이 된다.

$$\begin{aligned}
 E\{|e(n)|^2\} &\geq \prod_{i \in L(S)} (\sigma_q^2 \sum_m |f^i(m)|^2)^{1/d_i} \\
 &\geq -2^{-2R} \prod_{i \in L(S)} (\sum_m |f^i(m)|^2 \sigma_x^2)^{1/2^d} \quad (6)
 \end{aligned}$$

식 6은 복원 에러의 최소 경계 값이다. 또한 식 6은 부호화 이득의 분모 값이 되며 에너지의 밀집(compaction) 정도를 나타낸다. 그리고 신호를 Pulse Code Modulation(PCM)으로 양자화 시켰을 때 발생된 양자화 에러 분산의 최소값이 $2^{-2R} \sigma_x^2$ 인 것을 고려 할 때, 서브밴드 분할 S 에 대한 부호화 이득은 식 5와 같이 된다. □

정규화된 필터에너지 β_i 는 서브밴드 i 의 양자화 잡음이 다른 서브밴드로 이전되는 정도를 나타낸다. 직교 필터인 경우 β_i 는 1이 되어 서브밴드의 양자화 잡음의 이전이 없다. 양자화 잡음의 이전이 적을수록 서브밴드 부호화의 성능이 우수하다. 실제로 주로 사용되는 이중 직교 필터인 경우에도 β_i 는 거의 1에 가까워야 한다. 표 1은 영상 부호화에 널리 사용되는 9/7 이중직교 필터의 β_i 값들이다.

3. 부호화 이득의 분석

다양한 Wavelet 필터로 변화한 L프레임과 H프레임

의 부호화 이득의 분석을 위해 4레벨 MCTF를 통과한 L프레임 H프레임을 추출 하였다. 추출된 프레임에 대해 9/7필터, Haar 필터를 사용하여 서브밴드 분해한 후 부호화 이득을 비교하였다. L프레임은 MCTF의 4 레벨에서 얻었고, H프레임은 1 레벨에서 얻었다. 또한 각 프레임에 대해서 9/7필터와 Haar필터를 적용하여 3레벨 웨이블릿 변환을 하였다.

부호화 이득을 비교하기 위해 다음과 같은 식을 사용한다.

$$G_{comp} = \frac{Gain_{97}}{Gain_{Haar}} = \frac{\prod_{i \in L(S)} (\beta_i^{Haar} \sigma_{x_i^{Haar}}^2)^{1/2^{d_i}}}{\prod_{i \in L(S)} (\beta_i^{97} \sigma_{x_i^{97}}^2)^{1/2^{d_i}}}$$

여기서

- x_i^{Haar} : Haar 필터로 공간 변환 후 생성되는 서브밴드 i 의 신호
- x_i^{97} : 9/7필터로 공간 변환 후 생성되는 서브밴드 i 의 신호
- $L(S) = \{5, 6, 7, 9, 10, 11, 16, 17, 18, 19\}$,
- $\{d_i | i \in L(S)\} = \{2, 2, 2, 3, 3, 3, 4, 4, 4, 4\}$,
- 모든 i 에 대하여 $\beta_i^{Haar} = 1$,
- β_i^{97} 값은 앞 절의 표 1과 같음.

위의 G_{comp} 계산식으로부터 G_{comp} 의 값이 1보다 클 경우 9/7필터의 부호화 이득이 높고 1보다 작을 경우 Haar필터의 부호화 이득이 높다. 표 2는 MPEG에서 일반적으로 사용되는 영상 시퀀스^[8]에 대한 G_{comp} 값들이다.

결과에서 볼 수 있듯이 L프레임의 경우에는 낮은 주파수 영역에 에너지가 많이 모여 있기 때문에 효과적으로 에너지 밀집을 시키는 9/7필터의 사용이 Haar 필터

표 2. 다양한 영상의 L 과 H 프레임에 따른 G_{comp} 값

Table 2. The value of G_{comp} for various image sequences.

	Mobile	City	Foreman	Soccer
L프레임	1.5680	1.1956	1.8977	1.2928
H프레임	1.0651	0.9840	1.0814	1.0441
	Crew	Football	Harbour	Bus
L프레임	1.3232	2.2070	1.664	2.0175
H프레임	1.226	1.4652	1.1303	1.3039

를 사용하는 것보다 우수한 성능을 나타낸다. Mobile, City, Foreman과 같이 움직임이 느리고 움직임 벡터의 크기가 작아 움직임 예측이 잘 된다. 이러한 영상 시퀀스에서 얻어지는 H프레임에는 그림 2에서 보듯이 잔여 영상만을 포함하는 Inter block들을 대부분 포함하고 있다. 따라서 그림 3에서 보듯이 에너지 스펙트럼이 넓게 퍼져있어서 영상들은 에너지 밀집의 효과가 크지 않기 때문에 9/7필터와 Haar 필터의 성능에 차이가 크지 않음을 알 수 있다. 반면에 움직임이 빠르고 불규칙한 영상 신호들은 움직임 예측이 어렵다. 이러한 영상 시퀀스에서 얻어지는 H프레임에는 Intra-block으로 부호화될 영역이 많이 포함되어 있어서 영상 신호의 스펙트럼이 넓게 퍼져 있지 않다. 이러한 영상 시퀀스에 속하는 Football, Bus등에는 9/7필터를 적용하는 것이 더 효율적이다.

결론적으로 움직임 예측이 잘되는 영상 시퀀스에서 만들어지는 H프레임에는 Haar 필터를 적용시켜도 부호화 효율의 저하시키지 않으면서도 9/7필터보다 필터의 길이가 짧기 때문에 변환의 복잡도를 줄일 수 있다.

4. 제안하는 프레임 적응적 공간 웨이블릿 변환

앞서 분석한 것과 같이 저주파 대역에 대부분의 에너지를 포함하는 원영상과 유사한 L프레임은 저 주파수의 서브밴드에 에너지 밀집(compaction)을 효과적으로 하는 웨이블릿 필터가 필요하다. 그러나 움직임 예측이 잘되는 영상의 경우 H프레임의 잔여 영상 신호의 에너지 스펙트럼이 넓게 퍼져 있어 에너지 밀집의 효과가 크지 않다.

본 논문에서는 위와 같은 분석을 바탕으로 L과 H프

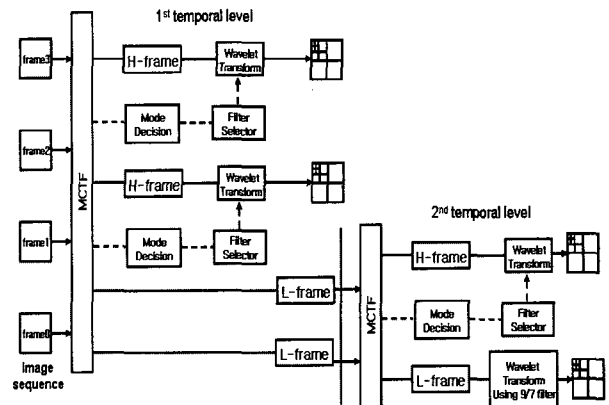


그림 6. 프레임의 특성에 따른 공간 웨이블릿 기법
Fig. 6. Frame adaptive spatial Wavelet transform.

레이들의 특성에 맞는 공간 웨이블릿 필터를 사용하는 기법을 제안한다. 즉, L프레임에 서브밴드 분해도가 높은 긴 웨이블릿 필터를 사용하고 영상의 움직임 예측이 잘 되는 H프레임에는 Haar 필터와 같은 짧은 웨이블릿 필터를 사용한다. 그림 6에 제안 기법의 블록도를 나타내었다.

실제 부호화 과정에서 부호화 이득을 계산한 후 프레임 특성을 판별하기 어렵기 때문에 움직임 예측의 정도를 판별하여 H프레임의 모드를 결정하고 웨이블릿 필터 매니저가 필터를 선택한다.

제안하는 기법은 복원된 영상 시퀀스의 화질에는 크게 영향을 미치지 않으면서 변환 복잡도를 줄이고 에지 성분이 많은 H프레임을 긴 필터로 변환함에 따라 발생하는 잔상 현상(ringing effect)을 감소시킬 수 있다.^[5]

IV. 실험

실험에 사용된 S/W는 MPEG SVC의 MSRA와 MC-EZBC이다^[9].

웨이블릿 변환의 복잡도에 대한 측정을 위해 기존의 방법과 제안된 방법의 웨이블릿 변환곱셈의 수를 계산하는 식은 다음과 같다.

$$C_{conventional} = GOF_{size} \cdot \sum_{l=0}^L \frac{MN}{4^l} \cdot (L_{low} + L_{high})$$

$$C_{proposed} = F_L \sum_{l=0}^L \frac{MN}{4^l} \cdot (L_{low} + L_{high}) + (GOF_{size} - F_L) \cdot \sum_{l=0}^L \frac{MN}{4^l} \cdot (2 + 2)$$

여기서

- M, N 은 프레임의 크기,
- L_{low}, L_{high} 는 lowpass, highpass 필터의 길이,
- GOF_{size} 는 GOF의 프레임 수
- F_L 은 GOF 내의 L프레임 수,
- L 은 웨이블릿 분할 레벨이다.

다음은 위의 식을 근거로 제안된 방법과 기존 방법에 대한 복잡도 개선율(CI)에 대한 식을 나타낸다.

$$CI = \frac{C_{conventional}}{C_{proposed}} \times 100(\%)$$

제안한 기법을 MC-EZBC S/W에 적용하였다, GOF

크기는 16 이고, 웨이블릿 분할 레벨은 3으로 하였다. 표 3, 4에 GOF 크기 16, 9/7필터의 사용 웨이블릿 분할 레벨 3일 경우 기존의 방법과 제안한 방법의 복잡도와 개선율을 나타내었고, 표 5와 6에 기존 방법과 제안한 방법의 부호화 성능을 PSNR값으로 비교하였다. MC-EZBC S/W는 Haar MCTF 기반으로 인접한 두 L프레임 간에 모션을 찾을 수 없는 픽셀의 수가 60%이상이면 장면 전환이 발생 한 것으로 간주한다.^[10] 이 경우에는 MCTF를 수행하지 않고 두 프레임 모두 인트라 프레임으로 처리한다. 실험에서는 L프레임과 인트라 프레임에는 9/7필터를 사용하고 그 외의 다른 H프레임에는 Haar 필터를 사용하였다. 그림 7은 MC-EZBC의 MCTF 수행과정의 예를 나타낸다.

제안한 기법을 MSRA에도 적용하여 보았다. MSRA S/W는 OBMC(Overlapped Block Motion Compensation)기반의 5/3MCTF를 사용하여 양방향 움직임 보상을 한다.^[11] 따라서 인접한 프레임들은 항상 인트라 프레임으로 처리되어 L프레임과 잔여 영상 신호만을 포함하는 H프레임만 발생한다. 실험을 위해 L프레임에는 9/7필터를 H프레임에 Haar 필터를 적용하였다. MSRA S/W의 경우 공간 웨이블릿 변환의 복잡도는 L프레임에 9/7필터를 나머지 H 프레임에 Haar 필터를 적용하였기 때문에 영상의 내용에 관계없이 프레임 사이즈, 프레임율이 동일하면 같다. 표 7은 MSRA의 웨이블릿 복잡도와 개선율을 나타낸다. 표 8, 9에 기존 방법과 제안한 방법의 부호화 성능을 PSNR값으로 비교하였다.

실험결과에서 보듯이 기존의 방법과 제안된 방법의 부호화 성능의 차이는 PSNR로 0.03dB 이내로 주관적인 화질에 있어서는 차이가 발생되지 않았다. 또한 기존방법

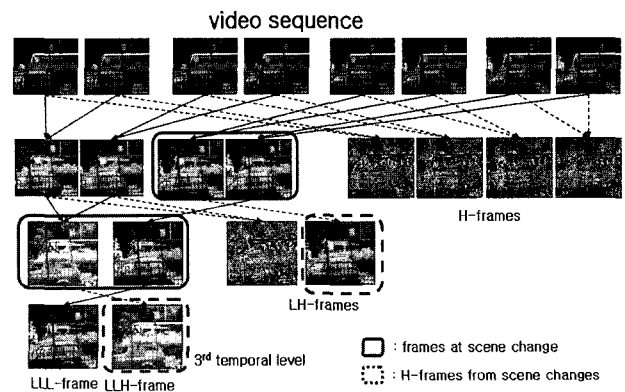


그림 7. MC-EZBC의 MCTF 예
Fig. 7. Example of intra mode coding(MC-EZBC).

표 3. City 영상의 웨이블릿 변환 복잡도 (MC-EZBC)
Table 3. Wavelet transform complexity of city sequence. (MC-EZBC)

size	frame /s	Conventional		Proposed		CI (%)
		# of multiplication	# of using filter(9/7)	# of multiplication	# of using filter(9/7)	
4CIF	30	2908200000	300	959713920	32	33.33
CIF	30	727056000	300	239928480	32	33.33

표 4. Crew 영상의 웨이블릿 변환 복잡도 (MC-EZBC)
Table 4. Wavelet transform complexity of crew sequence. (MC-EZBC)

size	frame /s	Conventional		Proposed		CI (%)
		# of multiplication	# of using filter(9/7)	# of multiplication	# of using filter(9/7)	
4CIF	30	2326579200	240	690703200	15	29.68
CIF	30	581644800	240	172675800	15	29.75

표 5. Mobile 영상의 PSNR값 (MC-EZBC)
Table 5. PSNR value of mobile sequence (MC-EZBC).

size	f/s	bps	Conventional			Proposed		
			Y	U	V	Y	U	V
CIF	30	1024	31.45	36.91	36.21	31.59	37.10	36.30
CIF	15	512	20.52	32.14	31.16	20.52	32.34	31.32
QCIF	15	256	20.01	29.27	28.33	20.00	29.53	28.57
QCIF	15	128	20.34	27.65	26.33	20.31	27.78	26.46

표 6. City 영상의 PSNR값 (MC-EZBC)
Table 6. PSNR value of city sequence(MC-EZBC).

size	f/s	bps	Conventional			Proposed		
			Y	U	V	Y	U	V
4CIF	30	3000	36.18	44.35	46.16	36.14	44.61	46.50
4CIF	30	1500	33.63	42.78	44.68	33.62	43.03	44.99
CIF	30	750	31.58	42.75	44.35	31.51	43.07	44.54
CIF	30	384	29.81	40.88	42.30	29.81	40.85	42.25

과 제안된 방법의 성능의 우수성은 영상 시퀀스의 종류에 따라서 차이가 나기 때문에 어떤 방법이 우수하다고 평가 할 수 없었다. 그러나 제안된 기법은 공간 웨이블릿 변환의 소요 연산량은 평균 1/3로 개선하였다. 따라서 제안된 기법은 부호화 성능을 열화시키지 않으면서도 공간

표 7. 웨이블릿 변환 복잡도 (MSRA)
Table 7. Wavelet transform complexity(MSRA).

size	frame /s	Conventional		Proposed		CI (%)
		# of multiplication	# of using filter (9/7)	# of multiplication	# of using filter (9/7)	
4CIF	30	2585100000	240	767448000	15	29.68
CIF	30	646272000	240	191862000	15	29.68

표 8. City 영상의 PSNR값 (MSRA)
Table 8. PSNR value of city sequence(MSRA).

size	f/s	bps	Conventional			Proposed		
			Y	U	V	Y	U	V
4CIF	30	1024	29.66	39.04	40.08	29.48	39.06	40.11
CIF	30	512	31.78	39.33	40.84	31.73	39.09	40.80
CIF	30	256	30.32	41.99	43.99	30.10	41.89	43.66
QCIF	15	64	35.22	43.01	44.78	35.21	43.01	44.79

표 9. Mobile 영상의 PSNR값 (MSRA)
Table 9. PSNR value of mobile sequence(MSRA).

size	f/s	bps	Conventional			Proposed		
			Y	U	V	Y	U	V
CIF	30	1024	28.97	33.47	32.97	28.84	33.46	32.96
CIF	15	512	27.53	32.09	31.46	27.43	32.09	31.46
CIF	15	384	25.1	29.33	28.73	25.06	29.34	28.73
QCIF	15	192	25.57	29.76	29.36	25.57	29.77	29.36

Wavelet의 복잡도를 1/3정도로 감소시킨다.

V. 결론

본 논문에서는 시간 축으로 필터링 된 프레임 특성에 따라 공간 웨이블릿 필터를 다르게 적용하여 공간 변환하는 프레임 적응적 인터프레임 부호화 방식을 제안하였다. 본 논문에서는 부호화 이득을 기반으로 시간 축으로 필터링 된 프레임 특성을 알아내는 방법을 제시하였다, 또한 실제 시스템에서는 부호화 이득을 측정하기가 어려움으로 MCTF 과정에서 연결되지 않은 화소들을 세워서 프레임 특성을 알아내는 방법을 제안하였다. 제안한 기법은 부호화 성능을 유지하면서 공간 웨이블릿 변환의 복잡도를 평균 1/3정도로 줄였다. 따라서 부호화 과정에서 연산의 많은 부분을 차지하는 역 웨이블릿 변환에 제안된 기법이 적용되면 연산량을 크게 줄일 수 있어 하드웨어로 구현이 용이하다.

참 고 문 헌

- [1] J. R. Ohm, "Complexity and Delay Analysis of MCTF Interframe Wavelet Structures", ISO/IEC JTC1/SC29/WG11 M8520, Klagenfurt, July 2002.
- [2] S.-J. Choi and J. W. Woods, "Motion compensated 3-D subband coding of video," IEEE Trans. Image Proc., vol. 8, no. 2, February 1999.
- [3] M. Schaar and D. Turaga, "Wavelet coding for video streaming using new unconstrained motion compensated temporal filtering", ISO/IEC JTC1/SC29/WG11 M8675, Klagenfurt, July 2002.
- [4] Bong-Keun Choi, Won-Ha Kim, "Dynamic Algorithm for Constructing the Optimal Subband Decomposition", IEICE, Information and Systems vol. E86-D, March 2003.
- [5] G. Strang and Truong Nguyen, "Wavelets and Filter Banks", Wellesy-Cambridge Press, 1997.
- [6] V.P. Sathe and P.P. Vaidyanathan, "Effects of multirate systems on statistical properties of random signal," IEEE Trans. Acoust., Speech Signal Process., vol.41, no.1, pp.131-146, Jan. 1989.
- [7] P. P. Vaidyanathan, Multirate Systems and Filter Banks, Prentice-Hall, 1993.
- [8] M. van der Schaar, J. Ridge, "Description of Core Experiments in SVC," ISO/IEC JTC 1/SC 29/WG 11/N6521 July 2004
- [9] Software package "MPEG-CVS",
<http://mpeg.nist.gov>
- [10] J. R. Ohm, "Three-dimensional subband coding with motion compensation," IEEE Trans. Image Proc., vol. 3, no. 5, September 1994.
- [11] Jizheng Xu, Ruiqin Xiong, Bo Feng, Gary Sullivan, Ming-Chieh Lee, Feng Wu and Shipeng Li, "3D Sub-band Video Coding using Barbell lifting," ISO/IEC JCT1/SC29/WG11 M10569/S05, March 2004.
- [12] M. Antonini, M.Barlaud, P. Mathiu and I. Daubechies, "Image coding using Wavelet transform", IEEE trans. Image Processing, vol. 11, pp674-693, July 1989.
- [13] S.-Ta Hsiang and J. W. Woods, "Embedded video coding using invertible motion compensated 3-D subband/wavelet filter bank," Signal Processing: Image Communication vol. 16, pp. 705-724, May 2001.

저 자 소개



박 성 호(학생회원)
 2004년 명지대학교 전자공학과
 (공학사)
 2004년~현재 경희대학교 일반대
 학원 전자공학과 재학
 (공학석사)

<주관심분야 : 비디오 압축, 스케일러블 비디오
 코딩>



정 세 윤(정회원)
 1995년 인하대학교 전자공학과
 (공학사)
 1997년 인하대학교 전산공학과
 정보공학전공(공학석사)
 1997년~현재 한국전자통신연구원
 선임연구원

<주관심분야 : 스케일러블비디오코딩, 멀티미디어
 통신, 디지털 대화형 방송>



김 원 하(정회원)
 1985년 2월 연세대학교 전자공학과(공학사)
 1988년 5월 Univ. of Wisconsin-Madison, 전기공학(석사)
 1997년 5월 Univ. of Wisconsin-Madison, 전기공학(박사)
 1996년 1월~1996년 7월 (미) Motorola, intern
 1997년 8월~2000년 2월 (미) Los Alamos National Lab. 연구원
 2000년 3월~2003년 8월 명지대학교 정보통신공학과 조교수

2003년 8월~현재 경희대학교 전자정보대학 부교수
 <주관심분야 : 멀티미디어신호처리, 영상처리>