

적응적 웨이블릿 변환을 이용한 저속 비트율 컬러 비디오 코딩에 관한 연구

김 혜 경[†] · 오 해 석^{††}

요 약

본 논문에서는 적응적인 웨이블릿 변환에 기초한 저속 비트율 비디오 코딩 방법의 새로운 알고리즘을 제안한다. 접근 방법은 양자화된 웨이블릿 계수들이 웨이블릿 시프밴드 구조내에서 중복성을 활용하는 메커니즘에 의해서 긴처리 된디블 코딩 절차가 더욱 효과적으로 작용하여 나타날 것이다. 그러므로 본 논문에서는 코딩부분의 최적화 활동에 초점을 맞추어 완전한 중복 블록 움직임 보상된 에리프레임에서 일치를 확보하기 위해 이용하고, 향상된 코사인 윈도우를 적용하였다 또한 웨이블릿 변환은 각 일치한 움직임 보상된 에리 프레임을 전체적인 에너지 컴백션에 도달하도록 적용된다 움직임 벡터의 수평적, 수직적 컴포넌트는 적응적 산술적 코딩을 사용하여 독립적으로 인코딩되는 반면에 의미있는 웨이블릿 계수는 적응적 심슨 코딩을 사용함에 의해서 비트-플레인 순서로 인코딩된다. 재인된 부호기는 28Kbits에서 PSNR이 평균적으로 각각 대략 2.07과 1.38dB에 존재하는 H.263과 ZTE를 초과한다 전체순서 코딩에 대하여도 3DWCVC 방법은 평균적으로 각각 0.35와 0.71dB을 나타내는 H.263과 ZTE 보다 우수한 성능을 보인다

A Study of Very Low Bit-Rate Color Video Coding Using Adaptive Wavelet Transform

Hye-Kyung Kim[†] · Hae-Seok Oh^{††}

ABSTRACT

This paper presents a new method for an efficient coding of very low bit-rate color video based on adaptive wavelet transform. Our approach reveals that the coding process works more efficiently if the quantized wavelet coefficients are preprocessed by a mechanism exploiting the redundancies in the wavelet subband structure. Thus, we focus on optimized activity of coding part, and exhaustive overlapped block motion compensation is utilized to ensure coherency in motion compensated error frames, and raised cosine window is applied. The horizontal and vertical components of motion vectors are encoded separately using adaptive arithmetic coding while significant wavelet coefficients are encoded in bit-plane order by using adaptive arithmetic coding. On average the proposed codec exceeds H.263 and ZTE in peak signal-to-noise ratio by as much as 2.07 and 1.38 dB at 28 kbits, respectively. For entire sequence coding, 3DWCVC method is superior to H.263 and ZTE by 0.35 and 0.71 dB on average, respectively.

[†] 준 회원 숭실대학교 대학원 컴퓨터학과
^{††} 정회원 숭실대학교 컴퓨터학과 교수
논문접수 : 2000년 1월 13일, 심사완료 : 2000년 2월 12일

1. 서 론

미디어 통신시스템에서 저대역의 채널과 함께 작동하는 매우 낮은 비트율의 비디오 압축은 빠르게 확산되는 연구분야가 되었다. 또한 잘 확립된 기법들은 블록기반 움직임 보상(BME, block-based Motion compensation)과 DCT를 기초로 하고 있다. 이러한 기법들 중에서 가장 유망한 한가지는 웨이블릿 변환 체계이며, 이런 변환 코딩기법은 부대역(subband) 코딩 기법들로 결합된다. 웨이블릿 변환은 신호를 10진 연산의 회선에 의해 부대역의 2진수로 구조화된 트리로 분해시키는 역변형이다. 웨이블릿은 상호간에 의존하는 부분들에 대한 역상호관계를 나타내고, 입력 신호를 표현하는 샘플들의 에너지 간결화(compaction)를 수행한다. 또한, 웨이블릿은 공간 주파수 도메인내에 잘 배치되어 영상의 자인적인 특성들에 부합되고, 그들의 다양한 규모의 특징들이 드러난다. 웨이블릿 기반 코딩 체계의 변형 모델은 일반적으로 양자화기에 의해 수행되고, 시각적으로 관련없는 정보를 무시하고, 마지막 단계에서는 엔트로피 코더가 양자화기에 의해 발생된 데이터 스트림내에서 통계적인 중복성을 제거하기 위해 사용된다.

이런 일반적인 웨이블릿 기반 코딩 체계는 정지영상이나 비디오에 적용될 수 있고, 최근에는 정지영상이나 비디오에 시간적인 중복성을 이용하는 메카니즘을 추가하는 작업이 제공된다. 비디오 코딩에 대한 대부분의 작업은 2-D 웨이블릿 기반 체계를 확장하는 문제에 대한 해결책을 제공하는 것으로 발표되었고, 세 가지 그룹으로 분류할 수 있다. 첫 번째 그룹은 2-D 웨이블릿 변환의 확장이나 3-D 부대역 코딩(3-D SBC: 3-D subband coding)에 대한 부대역 코딩(subband coding) 체계를 제안한다[1]. 두 번째 그룹은 웨이블릿 도메인내에서 다중해상도의 움직임 보상(MRMC: Multiresolutional Motion Compensation)에 도움을 줄 수 있는 시간적 중복성을 획득하는 것을 시도한다[2] [3]에서 표현된 세 번째 생각은 수정된 블록 매칭 알고리즘을 사용한, 소위 중첩 블록 움직임 보상(OBMC: overlapped block motion compensation)[4]이라 불리는 하나의 체계가 제안되었다. 전통적인 블록 기반 움직임 보상처럼, OBMC는 사전적인 에러 신호내에서 인위적인 블록킹 결과를 제거하는 이점과 함께 사전적 코딩(predictive coding)을 위한 매우 효과적인 기법이다.

3-D SBC와 다중해상도 움직임 보상(MRMC)과는 대조를 이루지만, 혼성의 OBMC/2-D DWT 체계는 공간적인 확장 방법으로 본래부터 양립할 수 없다.

본 논문에서 제안하는 매우 낮은 비트율의 애플리케이션을 위한 고성능의 하이브리드 3-D 웨이블릿 비디오 코딩 체계에서는 블록킹 효과를 제거하기 위하여 OBMC 방법을 적용한다. 또한 움직임 보상에 여러 프레임에 대한 웨이블릿 변환을 따르는 시간 도메인(time domain)의 움직임 평가와 움직임 보상으로 확장된 3-D 서브밴드 웨이블릿 비디오 코딩 체계의 이득은 매우 낮은 계산적인 복잡도와 에러 발생의 예방을 포함한다. 또한 컬러 모델을 적용한 3-D 웨이블릿 비디오 코딩방법은 인터넷 방송 서비스등에서 필수적인 기술로 각광을 받을 것으로 예상되며, 보다 자연스러운 영상 표현을 위하여 필수적으로 연구되어야 할 분야가 되었다.

2. 기존의 연구방향

DCT(discrete cosine transform)를 기반으로 한 영상 코더는 매우 낮은 비트율에서 블록킹(blocking) 효과와 모기 노이즈(mosquito noise)가 발생한다. 서브밴드 코딩 스킴은, 일반적으로 점진적인 영상 전송과 브라우징(browsing)을 위해 사용되고, 블록 기반 DCT에 대한 가능한 대체안을 제공한다. 비록 더욱 향상된 서브밴드 코딩 스킴을 언급하지 않더라도, 전통적인 서브밴드 코딩 스킴은 블록 기반 코더와 비교할 수 있는 객관적인 성능이 이미 산출되었고, 우려할만한 블록킹 산출물을 제거함에 의해 우수한 시각적인 특성을 보여 주었다.

2차원(2-D) 서브밴드 영상 코딩의 확장은 시간 도메인을 포함하여 제안된 3-D(temporal-spatial & spual) 서브밴드 비디오 코딩 알고리즘이 자연스럽게 된다. 3-D 웨이블릿 비디오 코딩 체계의 이득은 그들의 낮은 계산적인 복잡도와 에러 발생의 예방을 포함한다. 전자는 계산적으로 값비싼 시간-도메인 움직임 평가와 움직임 보상이 일관적으로 Harr 웨이블릿이 사용된 시간적인 필터링에 의해 교체된다는 사실 때문이다. 즉, 시간 도메인에서, 두 서브밴드는 각각 두개의 연속적인 프레임의 합과 차로써 획득된다. 후자는 하이브리드 코더를 가진 경우처럼 코더 구조내에 어떤 반복적인 루프가 없다는 사실 때문이다. 이러한 특징은 3-D 서

브랜드 비디오 코더를 무선 통신을 위한 매력적인 도구로 만든다. 그럼에도 불구하고, 시간적인 필터링은 낮은 프레임율에서 높은 움직임 순서에 대한 코딩 이득의 감소에 의해 입증된 것처럼, 일반적으로 비디오 순서들에서 본래부터 가지고 있는 시간적인 중복성을 이용하고 있는 시간-도메인의 블록 기반의 움직임 평가와 움직임 보상 알고리즘만큼 성공적이지 못하였다. 비록 이런 문제가 움직임 적응 구조를 사용함에 의해 경감되어질 수 있더라도, 3-D 서브밴드 비디오 코딩 알고리즘은 주로 높은 비트율의 애플리케이션에 대하여 중앙에 적용할 수 있다.

서브밴드 도메인(subband domain)의 다중해상도 움직임 평가와 움직임 보상 구조는 [5]내에서 소개되었다. 각 최초의 프레임에 대하여 웨이블릿 변환을 수행한 후에, 계층적인 블록 기반의 움직임 평가와 움직임 보상이 수행되었고, 의미있는 움직임 보상된 웨이블릿 계수의 인코딩이 뒤이어 수행된다. 그러므로, 두개의 연속적인 프레임들 간의 변환하는 움직임은 아마도 두개의 연속적인 웨이블릿 변환된 프레임들간의 변환하는 움직임으로 바뀌어질 수 없을 것이다. 그것은 다중해상도의 서브밴드 도메인의 움직임 평가와 움직임 보상에 대한 성능이 비트율이 감소되는 것처럼 매우 효과적으로 떨어졌음이 관찰되었다.

서브밴드 비디오 코딩 알고리즘에 대한 3가지 타입에서, 시간 도메인(time-domain)의 움직임 평가와 움직임 보상 이후에, 움직임 보상된 여러 프레임은 웨이블릿 도메인에서 인코드된다. 외관상으로, 이런 타입의 비디오 코딩 알고리즘과 전통적인 하이브리드 DCT 비디오 코딩 구조간의 차이는 움직임 보상된 여러 프레임의 인코딩에서 웨이블릿 변환이 DCT를 대신하는 것이다. 그러나, 교체가 만약 중복되지 않은 블록 움직임 보상 방법에 의해 생성된 것이라면 여러 프레임으로 적당하게 실제로 도움이 되지 못하는 것처럼 보인다. 웨이블릿 변환이 한지역에 집중된 블록킹 산출물을 결코 허용하지 않는 전역적 변환은 이해할 수 있고, 그것의 세기는 에너지 컴팩션과 관련하여 심하게 퇴화될 수 있다. 다행하게도, 붙일치는 중복된 블록 움직임 보상(OBMC)기법을 사용하여 크게 완화될 수 있다. 보고된 것처럼, OBMC는 블록킹 효과가 완화되었을 뿐만 아니라, 움직임 보상된 여러 프레임에 대한 전체적인 에너지를 줄였다. 최근에 제안된 매우 낮은 비트율의 웨이블릿 비디오 코딩 알고리즘[6]의 대부분이 이런 형

태의 것이고 OBMC 가 사용되었다.

최근에는, 웨이블릿 변환된 영상의 데이터 구성과 표현에 대한 혁신적인 전략의 이용 때문에 웨이블릿 영상 코딩의 성공이 달성되었다. 지금까지 공개된 웨이블릿 영상의 코더는 네 개가 있다. Shapiro의 EZW(Embedded Zerotropic Wavelet)[3] 코더와 Said와 Pearlman의 SPIHT(set partitioning in hierarchical tree)는 규칙적인 트리 구조와 분할된 집합의 트리(set-partitioned tree) 구조를 무의미한(insignificant) 웨이블릿 계수로 사용한다. Servetto 그리고 다른 곳에서의 웨이블릿 데이터(MRWD)[7]에 대한 형태적인 표현은 서브밴드이내에서 의미있는 계수에 대한 불규칙한 모양의 클러스터를 발견한다. Chai 그리고 다른 곳에서의 의미있는 링크로 연결된 컴포넌트 분석(SLCCA)[8]은 다중해상도의 웨이블릿 데이터(MRWD)를 서브밴드이내에서 의미있는 계수의 클러스터링과 의미없는 계수에 대한 클러스터링의 두가지에 대한 교차-규모(cross-scale)의 의존성을 이용하여 확장한다. 위의 네가지 웨이블릿 영상 코딩 알고리즘 사이에서, SLCCA는 일반적으로 가장 높은 성능을 나타낸다.

Bhutani와 Pearlman은 반복적인 움직임 보상에 의해 획득된 여러 프레임을 인코드하기 위하여 Shapiro의 EZW 알고리즘을 사용하여 제안했다. 그 코더는 MPEG-1 표준과 비교해볼 때 우수한 성능을 보여준다. Kim과 Pearlman은 SPIHT를 3-D 서브밴드 비디오 코딩[4]으로 확장하도록 제안했고, MPEG-2 보다 우수한 결과가 보고되었다.

최근에는, Vass 그리고 다른 곳에서 SLCCA 데이터 구조와 표현 전략은 낮은 계산적인 복잡도의 비디오 코딩을 위하여 적용된다. Sarnoff Corp의 제로트라 엔트로피(ZTE) 비디오 코더에서, H263과 같은 종류의 시간-도메인에서 블록 기반 움직임 평가와 움직임 보상을 수행한 후에, EZW와는 다른 알고리즘이 움직임 보상된 여러 프레임의 표현과 인코딩을 위해 제안되었다.

본 논문에서는, 고성능의 하이브리드 3-D 웨이블릿 비디오 코딩 알고리즘이 매우 낮은 비트율의 애플리케이션을 위하여 개발되었다. 본 논문에서 사용한 H263 권고안을 기반으로 한 정제된 시간-도메인의 움직임 평가는 시간적인 중복성을 줄이기 위해 사용되었고, 완전한 중복 블록 움직임 보상 기법은 움직임 보상된 여러 프레임내에서 일치를 보충하기 위해 이용된다. 또한, 웨이블릿 변환은 전역적인 에너지 컴팩션을 달성

하기 위해 각각의 균일한 움직임이 보상된 에러 프레임으로 적용된다. 웨이블릿 분해에 대한 서브밴드이내의 클러스터링 특성이 형태적인 조건부 팽창 연산에 의해 획득된 연결 컴포넌트로써 의미있는 웨이블릿 계수를 구성하고 표현하는 것에 의해 활용되는 반면에, 의미있는 웨이블릿 계수에 대한 교차 규모(cross-scale)의 의존성은 다른 등급에서 클러스터간의 의미있는 연결에 의해 활용된다. 마지막으로, 움직임 벡터는 직접 인코딩되고, 의미있는 웨이블릿 계수는 비트-플레인 순서로 인코딩되고, 두가지 모두 적응적인 산술적 코더를 사용하여 공간에 변화기 쉬운 순서로 인코딩된다.

3. 제안 알고리즘 소개

3.1 영상 코딩

3.1.1 적응적인 웨이블릿 변환

영상 압축에서, 웨이블릿 변환의 목적은 픽셀간의 상관관계를 감소시키거나 제거하는 것이다. 2가 원소의 웨이블릿 변환을 사용하여, 시그널은 소위 옥타브 밴드에서 분해된다. 서브밴드에 대한 중간 주파수가 낮아질수록, 대역폭은 더욱 좁아지고 주파수의 해상도는 더욱 높아진다. 만약 주파수 구성요소의 합계가 더 낮은 주파수에 대한 것보다 공간적 주파수가 높은 것 더 적다면, 이것은 영상 부호화에 대하여 최적이다. 물론, 이것은 모든 영상들에 대한 경우가 아니다. 이미지 "Barbara" 처럼 특색있는 결을 가진 영상들은 중간 지점에 대해 높은 부분을 갖고 높은 공간적 주파수를 가진다. 만약에 그런 영상들에 대하여 웨이블릿 변환에 대한 이점을 사용하기를 원한다면, 그것은 주파수 도메인내에서 부가적인 분해를 만드는데 적당하고, 더 나아가서 웨이블릿 계수들의 전체 상관관계가 기대된다. 남겨진 과제는 분해 결정에 대한 좋은 표준을 발견하는 것이다.

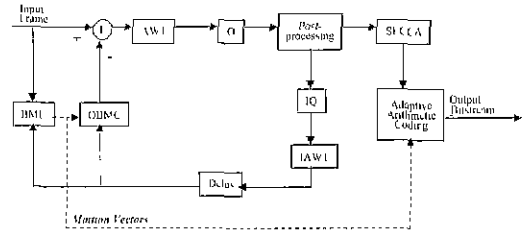
이러한 공헌에서, 서브밴드 내부에 있는 상관관계를 평가하는 표준함수가 제안된다. 그것에 대하여, 자동 상관 함수는 다음과 같이 소개된다.

$$R_{xx}[l, k] = \sum_m \sum_n d_{m,n} \cdot |d_{m+l, n-k}| \quad (1)$$

비율은 다음과 같다.

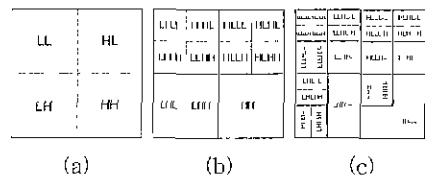
$$q_x = \frac{R_{xx}[1, 0]}{R_{xx}[0, 0]} \text{ and } q_y = \frac{R_{yy}[0, 1]}{R_{yy}[0, 0]} \quad (2)$$

x-와 y- 방향에서 서브밴드에 대한 웨이블릿 계수들의 상관관계에 대하여 측정한다. 만약 계수들이 x- 방향으로 충분히 전체상관 된다면, $R_{xx}[1, 0]$ 와 q_x 는 제로가 된다. 만약 $q_x = 1$ 이라면, 가장 강한 상관을 획득한다. 동시에 y- 방향에서의 상관에 대하여 참을 유지한다.



(그림 1) 제안된 3DWCVC 비디오 코딩 알고리즘의 블록 다이어그램

(그림 1)은 제안된 3DWCVC(3-D Wavelet Color Video Coding) 비디오 코딩 알고리즘의 블록 다이어그램을 보여준다. (그림 2)의 예제에서는 2가 원소의 분해를 보여준다. 첫째, 최초의 영상은 4개의 서브밴드(LL, LH, HL, HH)로 분해된다. L은 저역 통과 필터로부터 산출된 서브밴드를 표시하고, H는 고역 통과 필터링으로부터 산출된 서브밴드를 표시한다. 이러한 서브밴드는 가장적인 1/4 트리 구조를 구성하는 4로 간주된다. LL-서브밴드에 대한 상관관계의 계산은 필요하지 않고, 반면에 세 개의 다른 서브밴드에 대한 상관비율 q_x 와 q_y 가 계산된다. 유일하게 값이 주어지지 않는 해상도보다 크다면 분해가 수행된다.



(그림 2) x-y-적응적인 3단계 웨이블릿 변환

성공적인 역 웨이블릿 변환에 대하여 분해 구조의 전송이 필요하다. 이것은 4-심볼-알파벳(분해하지 않은 것, x-방향으로 분해, y-방향으로 분해, 두 가지 방향에서 분해)를 사용하여 수행된다. 전송된 심볼의 숫자

는 모든 1/4 트리 단계에서 4번 거듭되어진다. 적응성 있는 산술적인 코더를 전송 비용의 최소화를 위해 사용한다.

3.1.2 웨이블릿 양자화

최적의 양자화 전략은 웨이블릿 변환된 이미지에서 정보를 제거할 목적으로, 주어진 목표 비트율의 강제하에, 재구성된 이미지에서 결과물은 투명도의 쓰레스홀드(threshold) 아래에 있는 경향이 있다. 이러한 문제의 해결책은 인간 시각적 인식에 관한 확실한 수학적 모델의 설계를 포함하는 물론 접근하는 범위이상이다. 그러나, 가장 단순한 일률적인 스칼라 양자화를 총체적인 단계 $size^2 \cdot q$ 로 하는 모델과 함께 시작하고, 주어진 부대역 $W_{i,k}$ 의 주파수 내용에 의존하는 “특성 요소(quality factor)” q 에 관한 지각있는 가중치에 대한 선택사항을 유지한다

그것들의 웨이블릿 계수들을 흡수하기 위하여 본질적으로 잡음과 관련되어 있는 쓸모없는 지역 즉, 큰 제로 상자 $[-r, r]$ 를 구현하였다. 단계 크기에 대한 제로 상자 크기에 대한 비율 $\eta = (2r/q)$ 는 인트라프레임 코딩을 위하여 $\eta_{intra} = 1.5$ 로 선택되어졌고, 인터프레임 코딩을 위하여는 $\eta_{inter} = 2.0$ 을 사용한다. 다양한 비트율로 테스트된 모든 비디오 소스들에 대하여 경험적으로 만족한 선택이 되었음이 발견되었다

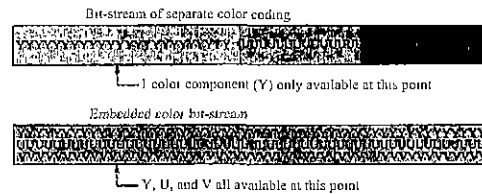
3.2 컬러 비디오 코딩

3.2.1 컬러가 내장된 비디오 코딩방법

지금까지는 휘도(luminance)라 불리는 유일한 하나의 컬러 플레인을 고려했다. 이 절에서는 3-D SPIHT에 대한 단순한 애플리케이션을 어떤 컬러 비디오 코딩으로 고려할 것이고, 반면에 여전히 전체 내장성을 유지하고 정확한 비율을 관리한다.

컬러 비디오에 대한 3-D SPIHT의 단순한 애플리케이션은 전통적인 컬러 비디오 코더가 했던 것처럼 각 컬러 플레인을 개별적으로 코드화한다. 그러므로, 각 플레인의 생성된 비트 스트림은 연속적으로 연결된다. 그럼에도 불구하고, 이 단순한 방법은 컬러 컴포넌트 사이에서 비트의 할당을 요구할 것이고, 정확한 비율 제어에 실패하고, 또한 비디오 코덱의 완전한 내장성에 대한 요구사항을 만족시키는 데 실패할 것이고, 이후 디코더는 재구성하여 보여주기 위해 전체 비트-스트림이 도착할 때까지 기다릴 필요가 있다. 그 대신에, 하

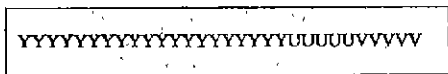
나는 모든 컬러 플레인을 코딩 단계에서 하나의 단위처럼 다룰 수 있고, 비트 스트림의 어떤 지점에서 멈출 수 있으며, 또한 주어진 비트율에서 최고 품질의 컬러 비디오를 재구성하기 위해서 하나의 혼합된 비트-스트림을 생성한다. 게다가, 컬러 플레인 사이에서 비트를 가장 바람직하게 자동적으로 할당하는 알고리즘을 위하여 완전 내장성과 정확한 비율 제어(rate control)를 보존할 것이다. 두가지 방법에 의해 생성된 비트스트림은 (그림 3)에서 보여준다. 즉, 첫 번째 방법은 전통적인 컬러 비트스트림을 보여주고, 두 번째 방법은 컬러가 내장된 비트스트림이 생성된 방법을 보여준다. 비트스트림의 어떤 지점에서 멈출 수 있고, 첫 번째 경우와는 대조적으로 그때의 비트율에서 컬러 비디오를 여전히 재구성할 수 있다



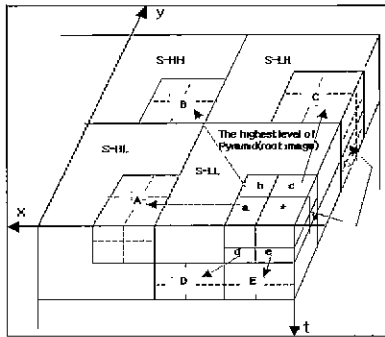
(그림 3) 비트스트림의 두가지 다른 방법 · 독립적인 컬러 코딩(Y,U,V의 정확한 순서)과 내장된 컬러 코딩(Y,U,V는 어떤 지점에서 선택됨)

YUV, YCbCr, 기타등등 처럼 3개의 컬러 공간은 휘도 성분 Y 플레인으로 간단하게 고려한다. 각각의 컬러 플레인은 독립적으로 웨이블릿 변환될 것이고, 그것의 독자적인 피라미드 구조를 가지고 있다. 현재는 모든 컬러 플레인을 함께 코드화하여 3-D SPIHT 알고리즘은 세개 플레인 모두에서 초기 단계에 적합한 좌표로 LIP(the list of insignificant pixels)와 LIS(the list of insignificant sets)를 초기화할 것이다. (그림 4)는 LIP와 LIS에 대한 최초의 내부 구조를 보여주고, 나타난 Y, U, V는 각 컬러 플레인내의 각 루트 픽셀에 대한 좌표를 나타낸다. 각 컬러 플레인이 그것의 독자적인 공간적 중심의 트리를 가지기 때문에, 그것은 컬러 플레인 사이에서 서로 배타적이면서 포괄적이다. 그것은 그들의 독자적인 좌표 크기에 대한 의미성을 따르는 플레인 사이의 비트를 자동적으로 할당한다. 각 컬러 플레인의 루트 픽셀이 초기화되는 것이 어떤 것이든지 순서에 대한 효과는 주로 낮은 비트율에서 코딩할 때를

제외하고 무시해도 좋다. (그림 5)는 피라미드의 가장 높은 단계에서 부모-자식 관계를 표현하고, 루트의 차원은 4*4*2이다. S-LL, S-LH, S-HL, 그리고 S-HH는 수직과 수평방향에서 주파수의 서브밴드를 나타내는 'spatial-low-low', 'spatial-low-high', 'spatial-high-low', 'spatial-high-high'를 나타낸다. S-LL에서 'a', 'b', 'c', 'd', 'e', 'f'에 의해 지시된 8 픽셀의 그룹이 있고, 픽셀 'f'는 픽셀 'b' 아래에 감추어져있다. 루트노드에서 'a' 위치에 위치한 모든 픽셀은 이번 자식도 가지지 않는다. 2*2*2 노드가 지시하는 루트 픽셀로부터 생긴 각 화살표는 부모-자식의 연결을 보여준다. (그림 5)에서 픽셀 'f'에 대한 자식 노드 'F'는 노드 'B' 아래에 감추어진 것이고, 그것은 'b'의 자식 노드이다. 집합 분할 (set partitioning)은 3-D 영역에서 즉각적으로 수행된다.



(그림 4) LIP와 LIS에 대한 최초의 내부적인 구조, 여기서 U와 V 플레인인 Y 플레인의 1/4 크기임



(그림 5) 피라미드의 가장 높은 단계에서 부모-자식 관계

3.2.2 움직임 평가와 보상

(1) 움직임 평가 : 최초의 프레임은 중복되지 않은 16*16 매크로 블록으로 분할된다. H.263 권고 안에서처럼 각 매크로 블록은 제로, 하나, 또는 네 개의 움직임 벡터를 가진다. 정수 픽셀의 해상도를 가진 완전 검색의 블록 매칭 알고리즘은 휘도 성분으로 하나의 움직임 벡터를 평방 평균 에러인 MSE(Mean-squared error)를 기준으로 사용하여 매크로블록마다 결정하기 위해 사용된다. 현재의 구현에서는 최초의 프레임 대신에 제

구성된 프레임이 더 나은 성능을 제공한 이후로 참조로써 사용된다. 검색범위는 수평적이면서 수직적인 두가지 방향에서 ±15 픽셀이다. 잘 알려진 것처럼 블록 매칭의 성능은 서브픽셀의 해상도를 사용함에 의해 실질적으로 향상될 수 있다. H.263 권고 안에서 명확히 설명됨에 따라, 움직임 벡터의 최초의 경수는 계산적으로 값싼 쌍일차 보간법의 관점에서 두 가지 방향내의 반 픽셀 해상도로 강제될 수 있다. 그리고, 이전의 재구성된 프레임은 참조로써 사용된다. 각 매크로 블록은 4개의 8*8 블록으로 분해되고, 블록마다 하나의 움직임 벡터는 반 픽셀 해상도로 동일한 절차를 정확하게 사용함에 따라 검색되고 정제된다. 네 개의 블록 움직임 벡터는 매크로 블록의 움직임 벡터와 독립적으로 결정된다. 이후에, 매크로 블록 당 제로 움직임 벡터는 $MSE_{zero} < MSE_{one} + T_{null}$ 일 때 결정되고, 이 수식에서 MSE_{zero} 와 MSE_{one} 는 각각 제로 움직임 벡터와 하나의 움직임 벡터를 매크로 블록마다 사용함에 따라 매크로 블록의 결과로 구성된 MSD를 나타낸다. 그리고, T_{null} 은 Null 가장자리(margm)를 명확하게 나타낸다.

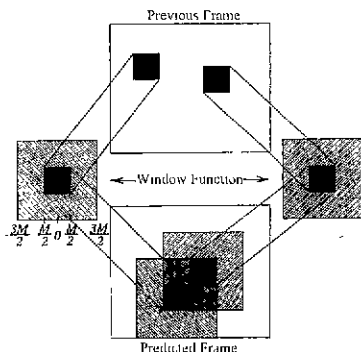
또한, $MSE_{join} + T_{split} < MSE_{one}$ 일 때, 매크로 블록마다 네 개의 움직임 벡터가 사용되고, 여기서 MSE_{join} 은 매크로 블록마다 네 개의 움직임 벡터가 사용될 때를 나타내고, 또한 T_{split} 은 미리 정해진 분열된 가장자리(split margin)를 나타낸다. 반면에, 매크로 블록마다 하나의 움직임 벡터가 사용된다. 물론, 각 매크로 블록에 대하여 사용된 움직임 벡터의 수는 측면 정보로써 디코더에 전송되어지는 것이 필요하다.

널 가장자리(null margin) T_{null} 이 증가함에 따라, 제로 움직임 벡터를 가진 매크로 블록의 수는 증가한다. 이러한 결과는 움직임 벡터의 전송에서 다 써버린 대역폭의 감소를 초래하고, 움직임 보상된 에러 프레임 에너지의 증가로 끝난다. 분열한 가장자리 T_{split} 의 값이 축소됨에 따라서 매크로 블록 당 하나의 움직임 벡터 위에 네 개의 움직임 벡터는 호감이 있고, 움직임 벡터 정보의 증가를 초래하고, 움직임 예인 에러 에너지의 감소를 가져온다. 몇 개의 테스트 영상 순서를 사용함에 따라, 그것은 실험적으로 $T_{split} \approx 6 \cdot T_{null}$ 이 균일한 움직임 보상된 에러 프레임용 의미있는 MSE의 축소를 산출하는 완전한 OBMC 알고리즘과 잘 어울린다는 사실이 발견되어졌다.

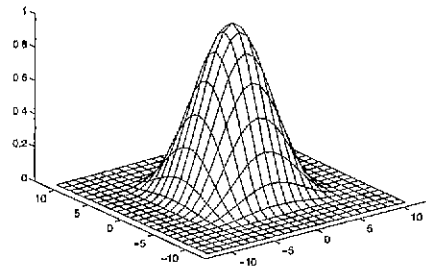
(2) 움직임 보상 : 블록킹 효과에 특별한 규정을 받지 않는 균일한 움직임 보상은 알고리즘의 성공에 결

정적이다 DCT-기반의 하이브리드 비디오 코딩 알고리즘에서, DCT의 효율성은 움직임 블록 경계선이 DCT 블록 경계선으로 잘 연합된다는 사실 때문에 블록 기반의 움직임 평가와 움직임 보상에 의해 인공적인 블록킹 효과에 의해서 의미적으로 강등되지 않는다. 그러나, 웨이블릿 변환처럼 진역적인 변환의 경우에 인공적인 블록킹 효과는 불연속의 고주파 컴포넌트가 생성될 것이고, 따라서 웨이블릿 변환의 효율성은 에너지 컴팩션에 기초를 두고서 의미적으로 강등된다. 개선 알고리즘 방법에 따라 완전하게 중복된 블록 움직임 보상은 움직임 보상된 여러 프레임에 대한 블록킹 효과를 경감시키는 것이 사용된다. OBMC는 균일한 움직임 보상된 여러 프레임을 제공할 뿐만 아니라 또한 움직임 보상된 예견 에러를 감소시킨다. 그러므로, 간단한 OBMC 알고리즘은 H.263 권고안의 부분으로 포함된다.

완전한 OBMC 알고리즘에 대한 연산은 (그림 6)에서 소개되었다. 프레임은 8*8 픽셀의 중복되지 않는 블록으로 분열되고 블록 당 하나의 움직임 벡터가 할당된다. 그것은, 매크로 블록마다 하나의 움직임 벡터가 움직임 평가에서 결정되는 것이고, 움직임 벡터는 네 개의 구성 요소인 블록들의 각각에 대하여 중복된다. 각 중복된 블록은 이전의 재구성된 프레임으로부터 현재의 블록으로 할당된 9개의 움직임 벡터와 그것의 8개의 이웃하는 블록을 사용함에 의해 현재의 프레임을 전송하는 것에 의해 결정된 많은 9개 블록들에 대한 가중된 합계로써 구성된다. 몇몇의 테스트 영상의 순서에서 성능 평가는 4-픽셀이 겹쳐질과 함께 향상된 코사인 윈도우는 가중치를 위한 좋은 선택을 보여준다(그림 7)



(그림 6) 본 논문에서 사용된 완전한 중복 블록 움직임 보상 (M=8은 블록 사이즈)



(그림 7) 향상된 코사인 윈도우

4. 성능평가

다른 비디오 코딩 알고리즘에 대한 성능 비교는 상당히 어렵다. 하나의 원인은 MPEG-4 테스트 순서가 최초의 ITU-T 601 포맷으로 분배되고, 자신의 포맷 변환을 사용한다. 또 다른 원인은 다량의 다른 비율의 제어 알고리즘이다. H.263, ZTE, 그리고 3DWCVC 간의 정당한 성능 비교를 보증하기 위하여, 3DWCVC에서 사용된 테스트 순서는 [6]에서와 같고, 또한 3DWCVC와 H.263 두가지에서 어떤 비율의 제어도 적용되지 않는다. 대신에, 모든 프레임은 균일한 스칼라 양자화기로 양자화된다.

3DWCVC에서 움직임 평가는 유일한 휘도 성분으로 수행된다. 휘도(chrominance)성분에 대한 움직임 보상에 대하여, 대응하는 휘도 움직임 벡터는 두 가지에 의해 나누어지는데, 블록 크기 4:4인 OBMC와 두 개의 픽셀이 겹쳐지는 중대된 코사인 윈도우 함수가 적용된다. 그러므로, 4-규모 그리고 3-규모의 이진법의 웨이블릿 분해는 각각 휘도성분과 채도(색차)성분 위에서 수행된다. 휘도 성분과 채도 성분은 균일한 스칼라 양자화기로 양자화된다. 모든 실험에서 4:4 평방을 구성하는 요소가 사용되고, 세 개의 의미있는 계수보다 더 적게 가진 클러스터는 제거된다. 객관적인 성능은 PSNR에 의해 측정되고, 다음처럼 정의된다.

$$PSNR[dB] = 20 \log_{10} \frac{255}{RMSE} \quad (3)$$

식 (3)에서 RMSE는 최초의 프레임과 재구성된 프레임간의 MSE(평방평균 에러)의 루트(근)이다. 비트율과 PSNR 성능에 대한 모든 결과는 디코딩된 비트 스트림으로부터 계산된다.

성능 비교는 QCIF 해상도내에서 수행되고, 첫 번째

인트라 프레임 코딩 결과는 주이졌다. 따라서, 전체 순서의 코딩 결과는(첫 번째 인트라프레임은 인터프레임에 의해 뒤따르는 것을 포함하여)소개되었다.

4.1. 인트라프레임 코딩

인트라프레임 코딩 비교는 14kbits(0.55bpp)와 28kbits(1.10bpp)에서 8개의 모든 테스트 순서들의 첫 번째 프레임에서 완성된다. 첫 번째로, H.263이 실행되고, 즉양자화기 스텝 크기는 H.263에 의해 획득된 비트율과 정확하게 어울리도록 새로운 알고리즘내에서 조정되었다. H.263, ZTE, 그리고 새로운 알고리즘이 적용된 3DWCVC의 결과가 <표 1>과 <표 2>에서 요약되었다. <표 1>에서는 14kbits에서 휘도 성분에 대한 3DWCVC가 H.263 보다 평균적으로 1.79dB 만큼 성능이 뛰어나고, 또한 0.64부터 1.39dB까지 ZTE를 둘러싸는 성능을 초과한다. 28kbits에서는 3DWCVC과 H.263 간의 차이가 증가함을 볼 수 있다. 즉, 3DWCVC는 1.31dB부

터 3.25dB까지의 성능을 보여 H.263의 성능을 초과한다. 동일한 비트율에서 3DWCVC는 평균적으로 1.38dB에 대해서 ZTE 보다 우수하다. 14kbits에서 채도 성분에 대하여, 3DWCVC는 U와 V 성분 각각에 대하여 평균적으로 1.27과 1.09dB에 대해서 H.263 보다 우수하다. 28kbits에서, 3DWCVC는 U와 V 성분 각각에 대하여 평균적으로 1.62dB과 1.44dB로 H.263 보다 성능이 뛰어나다. 평균화된 채도 성분에 대하여, 3DWCVC는 14와 28kbits 각각에서 0.46~2.10dB과 1.23~2.23dB에 대해서 ZTE의 성능을 초과한다.

4.2 전체 순서 코딩

인터프레임 코딩 비교에 대하여, H.263에 대한 무제한의 움직임 벡터 모드와 향상된 예견 모드 두가지가 사용된다. <표 3>은 10fps 프레임율로 샘플화되고, 24 kbps 비트율로 인코딩된 코딩결과를 보여준다 <표 4>에서는 15fps로 샘플화되면서 112kbps로 인코딩된 코

<표 1> 14kbits에서 인트라프레임 코딩에 대한 성능비교 결과

순서	File 크기 [Bts]	Luminance, PSNR[dB]			Chrominance, PSNR[dB]				
		H.263	ZTE	3DWCVC	U		Avg	V	
					H.263	ZTE		U	V
Akiyo	14000	33.06	34.62	35.35	35.23	37.38	36.19	37.12	39.46
Coast Guard	13888	30.12		31.18	41.98	44.94		42.84	45.12
Container Ship	14088	28.81		30.69	36.53	35.15		37.36	36.59
Foreman	13976	30.11	30.86	32.26	38.22	38.31	38.69	39.27	39.04
Hall Monitor	14104	29.68		31.84	36.04	39.99		37.40	40.29
Mother & Daughter	13624	33.78		35.54	39.27	39.67		40.86	41.58
News	14080	28.60	29.38	30.03	33.24	34.40	33.47	34.91	35.79
Silent Voice	13688	30.34		32.03	34.76	36.53		35.75	37.26

<표 2> 28kbits에서 인트라프레임 코딩에 대한 성능비교 결과

순서	File 크기 [Bits]	Luminance, PSNR[dB]			Chrominance, PSNR[dB]				
		H.263	ZTE	3DWCVC	U		Avg	V	
					H.263	ZTE		U	V
Akiyo	28218	38.42	40.18	41.67	40.09	41.53	40.81	42.82	43.25
Coast Guard	29280	34.23		35.54	44.13	45.68		44.88	46.69
Container Ship	28088	31.33		36.49	39.36	38.71		40.46	39.85
Foreman	26968	35.05	35.27	37.02	40.19	41.23	40.76	41.46	42.51
Hall Monitor	29040	35.62		38.32	38.66	40.87		40.75	42.25
Mother & Daughter	28088	38.77		40.46	42.27	42.70		44.07	44.90
News	27320	33.38	34.49	35.38	36.87	37.90	36.81	38.34	39.26
Silent Voice	26536	34.27		35.72	37.44	38.90		39.20	40.32

〈표 3〉 10fps, 24kbps에서 코딩의 성능비교 결과

순서	Bit Rate [kbps]	Luminance, PSNR[dB]			Chrominance, PSNR[dB]				
		H.263	ZTE	3DWCVC	U		Avg	V	
					H.263	ZTE	3DWCVC		
Akiyo	22.29	37.46	36.64	37.98	41.77	42.53	44.02	42.31	43.05
Container Ship	23.30	32.84		33.20	39.05	38.20		38.97	37.82
Hall Monitor	21.41	34.46	34.11	34.74	38.35	40.41	39.63	38.77	41.15
Mother & Daughter	23.81	35.53		35.75	40.85	41.26		41.12	41.75

〈표 4〉 15fps, 115kbps에서 코딩의 성능비교 결과

순서	Bit Rate [kbps]	Luminance, PSNR[dB]			Chrominance, PSNR[dB]				
		H.263	ZTE	3DWCVC	U		Avg	V	
					H.263	ZTE	3DWCVC		
Coast Guard	111.63	31.50	31.01	31.54	40.74	42.57	41.90	40.66	42.37
Foreman	118.62	34.47		34.49	39.25	39.73		39.16	39.30
News	109.95	37.68	37.59	37.62	40.89	41.50	42.55	40.53	41.14
Silent Voice	119.38	38.33		38.57	41.10	41.70		41.71	42.43

당결과를 보여주고 있다. 코딩결과는 처음으로 H.263이 실행되었고, 그 밖에 3DWCVC에서 실행되었으며, 양자화기 간격 크기는 H.263의 비트율과 어울리도록 조정되었다. 10fps와 24kbps에서, 3DWCVC는 H.263과 ZTE 보다 각각 0.35dB과 0.99dB 만큼 더 우수하다. ZTE와 비교해보면, 3DWCVC는 7.5fps와 48kbps에서 또는 15fps와 112kbps에서 각각 0.24~0.62dB 그리고 0.03~0.53dB 만큼 더 우수하다. 5fps와 10kbps에서 테스트 순서에 대한 체도 성분에 대하여, 3DWCVC는 U와 V 컴포넌트에 대하여 평균적으로 각각 0.95와 0.50 dB 만큼 H.263 보다 성능이 뛰어나다. 3DWCVC는 평균 0.32dB 만큼 H.263 보다 우수하다. (그림 8)과 (그림 9)는 객관적인 실험영상을 재구성한 결과화면을 보여주고 있다. 저속 비트율의 비디오 영상을 재구성한 결과화면은 3DWCVC 방법이 H.263 보다 영상 품질이 향상됨을 보여주고 있다.



(그림 9) Foreman의 재구성된 프레임 50 at 115kbps and 15 fps (좌) H.263 (우) 3DWCVC



(그림 8) Container Ship의 재구성된 프레임 150 at 24kbps and 10fps (좌) H.263 (우) 3DWCVC

5. 결론 및 향후 연구방향

본 논문에서는 매우 낮은 비트율의 비디오 코딩을 위하여 컬러 비디오의 의미있는 링크로 연결된 컴포넌트를 분석하여 가상적인 혼성의 웨이블릿 기법의 코딩 알고리즘을 표현했다. 복원된 영상의 블록킹 현상과 복원된 영상의 선명하지 않은 흐림(blur) 현상을 제거하기 위하여 제안된 중복 블록 움직임 보상기법은 결합되어 정제된 움직임 평가가 움직임 보상에 에러 프레임과 긴밀하게 결합되어 의미적으로 축소된 프레임 에너지를 산출한다. 이런 지역에서는 웨이블릿 변환이 혁신적인 데이터 구조와 표현전략과 관련되어 성공적으로 적용될 것이다. 본 논문에서 새롭게 제안한 3DWCVC에서는 의미있는 링크로 연결된 컴포넌트를 분석하여 웨이블릿 변환된 에러 프레임을 조직하고 표현한다. 또한, 의미있는 링크로 연결된 컴포넌트들이나 각 서브

맨드 내에 속하는 클러스터에 대한 정보는 네 개의 심볼을 가지고 적응적인 산술적 코딩에 의해 인코딩된다. 각 서브밴드 내의 의미있는 등급은 단지 두 개의 심볼 알파벳을 가진 것이 아니라면 적응적인 산술적 코딩에 의해 비트-플레인 순서로 인코딩된다. 적응적인 산술적 코딩에 의해 요구되는 조건적인 확률을 정의하기 위해 사용된 내용은 두 가지 경우에서 동일하다. 경험에 의한 뚜렷한 증거에 관하여, 혁신적인 데이터 구조와 표현 전략을 가진 웨이블릿 변환이 정지영상 코딩뿐만 아니라 비디오 코딩에 대한 매우 귀중한 가치있는 것을 표현한다는 것을 말할 것이다. 몇 개의 표준적인 테스트 순서로 광범위한 컴퓨터 실험은 웨이블릿이 아닌 낮은-비트율의 비디오 코딩 표준인 H.263과 높은 성능의 웨이블릿인 낮은 비트율의 비디오 코더 ZTE의 두 가지보다 빈함없이 성능이 뛰어남을 보여주었다. 그러므로 3DWCVC는 잘 알려진 최고의 낮은 비트율의 3-D 컬러 비디오 코더이다.

향후 연구방향은 좀더 향상된 계산적인 비용의 축소를 고려함과 동시에 신속하고 정확한 영상의 복원 알고리즘을 고려하고자 한다 또한 적응적인 웨이블릿 변환과 함께 더 나은 움직임 모델의 결합에 대하여 연구하고자 한다

참 고 문 헌

[1] C.I.Podilchuk, N. S. Javani, and N. Farvardin, "Three-dimensional subband coding of video," *IEEE Trans. Image Processing*, Vol.2, pp.125-139, Feb. 1995

[2] Y. Q. Zhang and S. Zafar, "Motion-compensated wavelet transform coding for color video compression," *IEEE Trans. Circuits Syst. Video Technol.*, Vol.2, pp.285-296, Sept. 1992.

[3] D. G. Sampson, E. A. B. da Silva, and M. Ghanbari, "Low bit-rate video coding using wavelet vector quantization," *Proc. Inst. Elect. Eng.*, Vol.142, pp.141-148, June 1995.

[4] R. W. Young and N. G. Kingsbury, "Frequency-domain motion estimation using a complex lapped transform," *IEEE Trans. Image Processing*, Vol.2, pp.2-17, Jan. 1993

[5] Y. Q. Zhang and S. Zafar, "Motion-Compensated wavelet transform coding for color video compression," *IEEE Trans. Circuits Syst. Video Technol.*,

Vol.2, pp.285-296, Sept. 1992

[6] S. A. Martucci, I. Sodagar, T. Chiang, and Y. Q. Zhang, "A zerotree wavelet video coder," *IEEE Trans. Circuits Syst. Video Technol.*, Vol.7, pp.109-118, Feb. 1997.

[7] S. Servetto, K. Ramchandran, and M. T. Orchard, "Wavelet based image coding via morphological prediction of significance," in *Proc. IEEE Int. Conf. Image Processing*, Oct. 1995, pp.530-533.

[8] B. B. Chai, J. Vass, and X. Zhuang, "Significance-linked connected component analysis for wavelet image coding," *IEEE Trans. Image Processing*, Vol.8, June 1999.

[9] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Process.*, Vol.1, pp.205-220, Apr. 1992.

[10] J. R. Ohm, "Three-dimensional subband coding with motion compensation," *IEEE Trans. Image Process.*, Vol.3, pp.559-571, Sept. 1994.



김혜경

e-mail hkkim@multi.soongsil.ac.kr
 1993년 서울산업대학교 전자계산학과(공학사)
 1997년 중앙대학교 컴퓨터학과(공학석사)
 1997년~현재 숭실대학교 컴퓨터학과 박사과정 재학중

관심분야 : 멀티미디어 시스템, 웨이블릿 영상코딩, 디지털 영상통신, 멀티미디어 응용



오해석

e-mail oh@computing.soongsil.ac.kr
 1975년 서울대학교 응용수학과(공학사)
 1981년 서울대학교 계산통계학과(이학석사)
 1989년 서울대학교 계산통계학과(이학박사)

1983~현재 숭실대학교 컴퓨터학과 교수
 1996년~1999년 숭실대학교 부총장직역임
 관심분야 : 멀티미디어 통신, 웨이블릿 영상 코딩, 멀티미디어 응용