동적 적응을 위한 움직임 보상 계층형 동영상 부호화

정회원 이 재 용*, 박 희 라**, 고 성 제**

Motion-Compensated Layered Video Coding for Dynamic Adaptation

Jae-Yong Lee*, Hee-La Park**, and Sung-Jea Ko** *Regular Members*

요 약

본 논문에서는 이종 환경에서 효과적으로 사용할 수 있도록 다중 계층형 비트열을 생성할 수 있는 계층형 동영상 부호화 방식을 제안한다. 제안하는 부호화 기법에서는 시축 해상도의 확장성(scalability)을 지원하기 위해 시축 계층구조에서 새로운 움직임 추정 구조를 가지도록 하고, 웨이블릿 분할을 이용하여 공간 확장성까지도 지원하도록 한다. 움직임 추정과 보상을 사용하여 시축 중복성을 더욱 줄여 조건 갱신 방식보다 압축률을 향상시킬 뿐만 아니라 '분산 부대역 인트라 갱신' 기법을 사용하여 동적 적응이나 에러에 대해서도 효과적으로 동작할 수 있다. 또한 EZW (embedded zerotree wavelet) 기법을 사용하여 임베디드 비트열을 생성함으로써 추후에 데이터율 확장성을 쉽게 얻을 수 있도록 하였다. 따라서 제안하는 부호화 기법은 인터넷이나 ATM, 무선망과 같이 상호 운용과 확장성을 필요로 하는 네트워 상에서 매우 효과적으로 사용될 수 있을 것이다.

ABSTRACT

In this paper, we propose a layered video coding scheme which can generate multi-layered bitstream for heterogeneous environments. A new motion prediction structure with temporal hierarchy of frames is developed to afford temporal resolution scalability and the wavelet decomposition is adopted to offer spatial scalability. The proposed scheme can have a higher compression ratio than replenishment schemes by using motion estimation and compensation which can further reduce the temporal redundancy, and it effectively works with dynamic adaptation or errors using dispersive intra-subband update (DISU). Moreover, data rate scalability can be attained by employing embedded zerotree wavelet (EZW) technique which can produce embedded bitstream. Therefore, the proposed scheme is expected to be effectively used in heterogeneous environments such as the Internet, ATM, and mobile networks where interoperability and scalability are required.

I. 서 론

동영상 서비스와 같은 연속 매체를 다루는 멀티미디어 응용에서는 적절한 시간에 필요한 정보를 가시화하기 위하여 시간 제약 조건이 필요하다. 만일 통신망이 QoS를 보장하지 못하거나 운영체제에서 실시간 스케줄링을 지원하지 못한다면 지연이나 지터(jitter)를 응용계층에서 조절하여, 주어진 자원하에서 시간 제약 조건을 충족시킬 수 있도록 할 필요성이 있다[1]. 이것은 동영상의 화질과 처리 부하의 절충을 하는 적응기법에 의하여 가능하다. 적응기법에 의하여 재생 단말은 단지 전체 동영상 정보 중 일부분만을 사용하게 되는데 이것으로 의미있는 복원을 가능하도록 하기 위해서는 특별한 동영상 부호화 기법이 필요하다. 뿐만 아니라 앞으로

* 현대전자산업(주) 정보통신연구소 (jlee@ieee.org),
** 고려대학교 전자공학과 (sjko@dali.korea.ac.kr)
논문번호 : 99149-0413, 접수일자 : 1999년 4월 13일

디지털 동영상 서비스는 디지털 방송, 동영상 압축 저장, 비디오 온 디맨드 및 화상회의 시스템 등과 같이 요구조건이 서로 다른 여러 이질적인 응용 간에 상호 운용(interoperability)을 필요로 할 것이다. 따라서 이종 환경에서 하나의 비트열을 이용하여 이렇게 다양한 동영상 서비스가 요구하는 여러 조건을 만족시키기 위해서는 동시에 개개의 서비스에 알맞은 비트율을 통합적으로 제공할 수 있도록 하는 동영상 부호화 기술이 필요하다.

이러한 목적을 위한 한 가지 방법은, 동영상 정보를 스케일러블한(scalable) 형태로 기술하여 수신단에서 생성된 비트열의 일부만을 선택하여 주어진 자원을 이용하여 복호화할 수 있도록 하는 것이다. 확장성(scalability)을 지원하기 위한 한가지 방법은 계층형 부호화 기법을 사용하는 것이다. 계층형 부호화 방식에서는 동영상 정보가 다중 비트율을 제공할 수 있도록 여러 계층으로 분산되어진다. 가장 기본적인 정보를 전달하는 기저 계층(base layer)과 이에 추가되어 향상된 화질을 얻을 수 있도록 하는 복수개의 개선 계층(enhancement layer)들로 구성될 수 있다. 이 때에는 복원시에 많은 계층을 이용할수록 더 우수한 화질을 얻을 수 있다.

스케일러블 동영상 압축 부호화 기법은 크게 부대역 기반 부호화와[2]-[4] 예측 기반 부호화[5]-[10] 두 부류로 나눌 수 있다. 시공간 부대역 부호화 기법이 예측 기반 부호화 기법보다 일반적으로 좀 더 세밀한 확장성을 제공해 줄 수는 있지만 복잡도와 계산량이 크고, 그로 인한 지연으로 인해 현재 실용적으로 사용하기에는 어려움이 있다. 그러므로 전통적인 움직임 보상 기반 부호화 기법들을 확장하여 확장성을 얻으려고 하는 노력들이 활발히 이루어져 왔다. 그러나 다양한 움직임 추정 및 보상 방식이 그동안 효과적으로 동영상 압축 부호화에 사용되어 온 반면에, 기존의 (1계층) 동영상 부호화에서와는 달리 움직임 보상을 계층형 표현 방식에 기반을 둔 스케일러블 동영상 부호화에 사용하기에는 어려운 점이 있다[4],[9]. 복호시에 개선 계층을 사용하지 않는 경우 복호기에는 움직임 보상을 위한 기준 정보가 일부 계층에 대해서만 존재하게 되므로 이후에 새로운 계층을 추가하고자 할 때 움직임 보상에 의하여 완전한 화질을 복원할 수 없기 때문이다. 따라서 조건 갱신(CR: conditional replenishment) 방식이 계층형 동영상 부호화를 위하여 제안되었다[8]. 여기에서는 동영상 프레임을 블럭으로 분할하고 각각을 이전 프레임의 해당 블록과 비교하여 일정 기준치 이상으로 변화가 발생한 블럭에 대해서만 인트라 부호화하는 방법을 사용한다. 하지만 이 방식으로는 움직임 보상 예측을 이용한 방법과 같이 높은 압축 성능을 얻기 어려울 뿐만 아니라 부호화된 블럭들과 갱신되지 않은 블럭들 사이에서 불연속이 나타나게 된다.

본 논문에서는 위에서 언급한 여러 문제들을 해결하면서 확장성을 제공하기 위하여 프레임들을 시축 계층 구조로 배열하고 이 계층 구조 상에서 새로운 움직임 예측을 수행하는 계층형 동영상 부호화 방식을 제안한다. 또한 영상을 웨이블릿 피라미드로 분할하고 각 계층마다 최선의 움직임 벡터를 추출할 수 있도록 움직임 벡터를 하위 움직임 벡터를 이용하여 점진적으로 개선하는 기법을 이용하여 공간 해상도 확장성을 지원할 수 있도록 한다. 복원시 동적 적응을 가능케하고 또한 에러가 전파되는 것을 막기 위해서 주기적인 인트라 갱신을 수행한다. 이 때에는 급격한 비트율의 증가가 일어나지 않도록 각 계층으로 인트라 갱신을 분산시키는 방법을 제안한다. 제안하는 방법을 이용하여 이종 환경을 위해 확장 가능한 다중 계층 비트열을 생성할 수 있으며 공간 해상도와 프레임율을 다양한 사용자의 요구에 따라 독립적으로 조절할 수 있다. 게다가 움직임 추정과 보상을 이용하여 시축 중복성을 좀 더 줄일 수 있으므로 조건 갱신 방식을 이용한 경우보다 높은 압축률을 얻을 수 있다.

다음 II 장에서는 시축 계층화와 이에 기반한 움직임 추정 방식을 서술하고, III 장에서는 영상 피라미드 분할과 다중 해상도 움직임 벡터를 사용하여 공간 해상도 확장성을 지원하는 것과 에러의 전파를 막고 동적 적응을 하기 위한 주기적인 인트라 갱신을 제안한다. IV 장에서 실험결과를 보이고 그에 대한 고찰을 하였으며 마지막으로 V 장에서 전체 내용을 요약하고 결론을 맺는다.

II. 시축 계층화와 움직임 추정

동영상의 경우 공간 해상도, 시축 해상도, 데이터율 확장성 등 몇가지 확장성이 있다. 공간 해상도 확장성이란 영상을 동영상의 탐색에 사용하는 작은 엄지손톱만한 크기로부터 HDTV의 크기까지 복원할 수 있는 기능을 의미하고, 시축 해상도 확장성이란 프레임의 갱신율을 사용자의 요구에 의하여 조

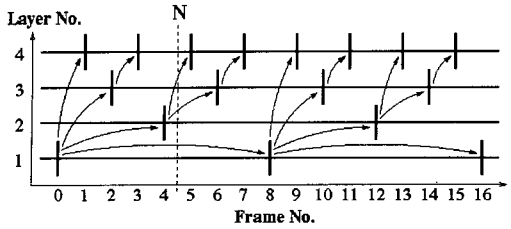


그림 1. 시축 계층 상에서의 움직임 추정.

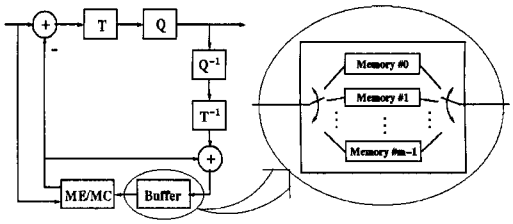


그림 2. 시축 계층화를 위한 부호화기 및 버퍼의 구조.

절할 수 있는 것을 의미한다. 신호 대 잡음비 확장성이라고도 불리우는 데이터율 확장성은 하나의 비트열로부터 사용자가 원하는 임의의 비트율로 동영상을 재생할 수 있도록 하는 것을 의미한다. 본 장에서는 먼저 시축 해상도의 확장성을 지원하는 시축 계층화를 제공하는 방법을 설명한다.

1. 시축 계층화

현존하는 계층형 부호화 방식은 MPEG이나 H.26x와 같은 기존의 표준 동영상 부호화 기법을 채용하여 단지 두 개의 계층만을 지원하는 형태가 대부분이다[5],[11]. 시간적인 계층구조(temporal hierarchy)에서 다수의 계층을 지원하기 위해서 시간축의 부대역 분할(subband decomposition)이 사용될 수도 있지만[2],[3], 이는 시스템의 복잡도를 증가시키며 부가적인 지연을 야기하게 된다. 따라서 동영상 부호화를 위해 움직임 추정과 보상에 기반하여 시축 계층화를 하는 구조가 제안되었는데[12], 이 구조에서는 부표본화된 프레임들이 어떠한 계층들의 조합에도 각 프레임 사이에 일정한 시간 간격을 유지하도록 그림 1과 같이 배열된다. $N_{TL}=M+1$ 개의 시축 계층이 있다 할 때 $n\geq 0$인 프레임 n이 전달되는 계층번호를 $L_M(n)$으로 나타낸다면,

$$L_M(n)=M-r(n \bmod 2^M+2^M)+1 \tag{1}$$

이며,

$$r(m)=\min\{k>0:\ \lfloor m/2^k \rfloor 2^k \neq m\}-1 \tag{2}$$

이다. 여기에서 $r(m)$은 m의 이진 표현에서 우측으로부터 시작하여 (제일 우측 비트의 위치를 0이라 하고) 0이 아닌 첫 번째 비트의 위치를 나타낸 다. 그림 1에 4개의 계층($M=3$)으로 이루어진 시축 계층화가 도시되어있다. 위의 시축 계층화 기법에서 움직임 벡터는 기존의 예측 기반 부호화에서와 같이 바로 이전 프레임에서 추정될 수가 없다. 이는 바로 이전 프레임이 디코더에서 이용하지 않는 상위의 계층을 통하여 부호화 될 수 있기 때문이다. 따라서 프레임 n의 움직임 추정 기준 프레임은 그림 1에서 보여지는 바와 같이 계층 $L_M(n)$이나 그보다 하위의 계층에 존재하는 프레임들 중 시간적으로 프레임 n에 가장 가까운 프레임으로 결정된다. 전체 계층의 개수가 $M+1$일 때 프레임 n의 기준 프레임의 일련번호는

$$R_M(n)=n-2^{M-w(n)} \tag{3}$$

와 같이 주어지며, 여기에서

$$w(n)=\min\{k\geq 0: n \bmod 2^{M-k}=0\} \tag{4}$$

이다.

2. 부호화기의 구성과 시축 계층의 특성

1) 기준 영상을 위한 다중 메모리

시간적인 계층 구조에서 프레임들이 다수의 계층으로 분산되어져 있기 때문에 예측 부호화를 위해서는 다수의 움직임 추정 기준 프레임들의 저장을 위한 저장 공간을 필요로 하게된다. 예를 들어, 그림 1을 보면 점선으로 표시된 N이라는 순간에는 두 개의 기준 프레임이 저장되어 있어야 함을 알 수 있다. 즉 프레임 4는 프레임 5와 6을 위해, 프레임 0은 프레임 8의 예측 부호화를 위해 저장되어 있어야 한다. 따라서 이후에 있을 움직임 추정 과정에 사용하기 위하여 이들 여러 개의 기준 프레임들을 저장할 수 있는 특별한 버퍼를 필요로 한다. 제안하는 계층형 부호화기가 그림 2에 도시되어 있다.

이러한 구조에서 이산여현변환(DCT)이나, 이산웨이블릿변환(DWT) 등과 같은 여러 변환 기법들이 T의 위치에서 사용될 수 있다. 기준 프레임을 위한 버퍼는 여러 개의 프레임 메모리와, 그림 2에서와

같이 움직임 추정시에 적절한 해당 기준 프레임 메모리에 읽고 쓰기 위한 스위칭부로 이루어진다. 움직임 추정을 위해 필요한 프레임 메모리의 개수는 $\lceil (M+1)/2 \rfloor$ 이다. 여기서 $\lceil \cdot \rfloor$ 은 반올림을 의미한다. 움직임 보상을 위해 기준 프레임을 읽어 올 때나, 복원된 영상이 이후 프레임들의 부호화를 위해 프레임 메모리에 씌여질 때 여러 개의 프레임 메모리 중 해당 프레임 메모리가 적절히 선택되어야 한다. 기준 프레임이 읽어지게 될 프레임 메모리의 번호는 기준 프레임이 속한 계층의 번호에 따라 결정된다. 프레임 n의 부호화를 위한 기준 프레임이 저장되어 있는 프레임 메모리 번호는

$$B(n) = \lfloor L_M(R_M(n))/2 \rfloor \quad (5)$$

으로 주어진다. 이 때 $\lfloor m \rfloor$ 은 m보다 크지 않은 최대 정수를 의미한다. 여기서 주목할 것은 기준 프레임과 현재 프레임의 거리가 매 프레임에 따라 달라지게 되어 기준 프레임으로부터의 움직임 예측값이 비교적 클 수도 있으므로 움직임 벡터 탐색 영역이 현재 프레임과 기준 프레임의 거리에 비례하여 조절되어야 한다는 것이다.

2) 오류 전파의 지역성

전송시에 발생할 수 있는 채널 에러에 의하여 수신 정보가 손상되었거나 혹은 적응기법에 의하여 수신단에서 특정 계층을 사용하지 않는 경우 일반적인 예측 기반 부호화기에서는 이후에 에러의 전파가 큰 문제가 된다. 하지만 위에서 설명한 계층구조를 사용할 경우에는 프레임들 간의 예측이 하나의 계층구조로 엮어져 그 전파에 한계가 주어진다. 즉, 기저 계층에 위치하는 하나의 프레임으로부터 2^M개의 프레임을 단위로 하여 하나의 움직임 예측 구조가 이루어지며 이 예측 구조 내에서 발생하는 모든 정보의 손실은 그 구조 내에서만 전파가 될 뿐 새로운 예측 구조에는 영향을 미치지 못한다. 예를 들어 3번째 계층에서 발생한 에러의 경우 그에 따른 손실은 그 다음 프레임까지만 전파될 뿐 그 이후에는 정상적인 복원에 아무 문제가 없다. 다만 기저 계층에서의 손실만이 계속해서 이후 프레임에 영향을 미치며 이어져 나갈 수 있지만 이는 다음 장에서 설명할 주기적인 갱신 방법에 의해 해결될 수 있다. 가능하면 이러한 시축 계층화는 우선순위에 기반한 전송방식 혹은 처리방식과 함께 사용되어야 하는데 이는 정체시에 발생하는 피할 수 없는 오류가 복호화시 사용하는 모든 계층 중 상위에서 일어나게 하여 에러 전파를 최대한 줄이도록 해야 하기 때문이다.

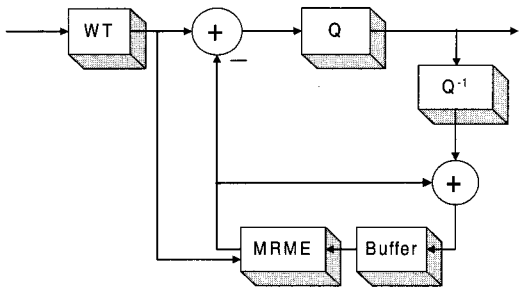


그림 3. 웨이블릿 변환을 이용한 공간 계층형 부호화기.

Ⅲ. 공간 계층화와 동적 적응을 위한 대역내 갱신의 분산

본 장에서는 웨이블릿 변환과 다중 해상도 움직임 추정을 이용하여 공간 확장성을 제공함과 동시에 복원시에 동적인 스케일링을 가능케 하기 위하여 주기적인 인트라 갱신을 분산시키는 방법을 제안한다.

1. 공간계층화

다중 해상도의 영상을 생성할 수 있는 웨이블릿 영상 피라미드 분할 기법을 사용하여 II 장에서 기술한 시축 계층형 동영상 부호화 방법을 공간 확장성을 지원하도록 확장할 수 있다. 피라미드 구조로부터 얻어진 다중 계층은 원신호를 서로 다른 공간 해상도에 대하여 특성화한다.

시공간 계층형 부호화기는 그림 2의 시축 계층형 부호화기를 수정하여 그림 3과 같이 구성될 수 있다. 하나의 동영상 프레임은 우선 부호화기의 첫 부분에서 웨이블릿 피라미드 구조의 다른 해상도를 가지는 여러 개의 계층으로 분할된다. 웨이블릿 분할이 모든 부호화 과정보다 선행되기 때문에 움직임 추정은 웨이블릿 영역에서 수행되어야 한다. 웨이블릿 피라미드에서 다른 계층의 움직임들은 서로 다르지만 결국에는 같은 움직임을 단지 다른 척도와 주파수 대역에서 표현하고 있는 것이기에 상관성이 매우 크다. 이에 기반하여 다중 해상도 움직임 추정(MRME: multiresolution motion estimation)[6]이 제안되었으며 여기에서 높은 해상도에서의 움직

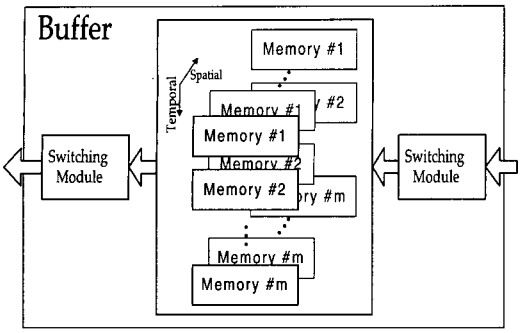


그림 4. 공간 계층화를 위한 버퍼의 구조

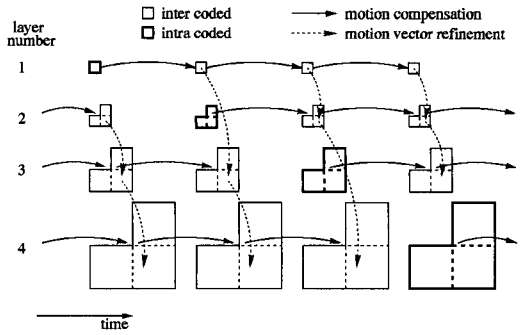


그림 5. 시축 기저 계층에서의 DISU.

임 벡터는 하위 해상도에서 얻어진 움직임 정보를 이용하여 세밀화 함으로써 얻어질 수 있다.

웨이블릿 분할과 MRME를 이용한 부호화기 구조에서 그림 3의 프레임 버퍼는 그림 4와 같이 다중 면을 가지는 구조로 구성된다. 대응하는 수신단의 복호화기에서도 동일한 구조의 기준 프레임 버퍼를 가지고 있을 것이며 복호화 수행시에 사용자의 요구나 사용할 수 있는 주어진 자원에 근거하여 프레임 메모리 전체 혹은 그 중 일부가 기준 프레임의 정보로서 선택될 것이다. 이러한 방식으로 움직임 예측 기반 부호화 방식이 계층형 동영상 부호화기에 활용될 수가 있다.

2. 주기적인 분산 부대역내 갱신

움직임 추정에 기반한 스케일러블 동영상 부호화기/복호화기의 설계 목표는 최적의 예측을 찾는 것이다. 통신망 대역폭이나 단말 시스템의 능력과 같은 사용 가능한 자원에 따라 복호시에 동적 적응을 하기 위해서는 복호기는 공간 계층을 추가하거나 버리게 된다. 동영상 재생시 화질을 증가시키기 위해 개선 계층을 추가한 경우 움직임 보상에 사용할 수 있는 기준 정보가 최적이지 못하다. 즉, 추가된 계층에는 움직임 보상에 기준으로 사용할 이전 프레임에 대한 정보가 없으며 복원 화질은 기대한 것만큼 높지 못하다. 주어진 조건하에서 잠재적인 최선의 화질을 회복하고 따라서 동적 적응을 가능케 하는 방법으로 MPEG의 I-프레임과 같은 프레임내(intraframe) 정보를 주기적으로 삽입하는 것을 생각해 볼 수 있다. 하지만 이 경우에 비트열에 상당히 큰 피크가 발생한다는 단점이 있다. 이 문제를 해결하기 위하여 본 논문에서는 분산 부대역내 갱신(DISU: dispersive intra-subband update)을 제안한다. 여기에서는 하나의 전체 프레임 대신 그림 5와 같이 하나의 공간 계층을 이루는 웨이블릿 피라미드에서 한 단계의 부대역들만이 시축 기저 계층에서 인트라 부호화 되어 갱신되고 이후 차례로 각 단계의 부대역들이 인트라 갱신된다. DISU를 이용하여 복호기가 새롭게 공간 개선 계층을 추가할 경우 복원 영상의 화질은 움직임 예측 기준 정보가 최적이 될 때까지 점차적으로 증가하게 된다.

이 경우 인트라 코딩되는 부대역에서는 움직임 추정을 수행하지 않으므로 MRME시에 상위 계층의 움직임 벡터 추출에 필요한 움직임 벡터의 추가 정보가 만들어지지 못한다. 따라서 이 때에는 그림 5의 수직 방향 화살표가 나타내듯이 부호화시에 그보다 하위 계층에서 찾아진 움직임 벡터를 추가 정보 없이 바로 이용하여 움직임 벡터의 세밀화를 이룰 수 있도록 한다. DISU에서는 각 계층별 갱신 주기에 따라 다음과 같은 두가지 방법을 생각해 볼 수 있다.

- 단순 부대역 갱신 (simple update)

각 계층이 순차적으로 인트라 갱신이 되도록 한 방식이다. 한번씩 모두 갱신이 이루어지면 다시 기저 계층부터 반복적으로 갱신이 이루어진다. 전체 공간 계층수를 N_{SL}이라 할 때 모든 계층에서 갱신 주기는

$$P_s = 2^M \times N_{SL} \tag{7}$$

이 된다. 예를 들어 시축 계층이 4개, 공간 계층이 4개인 경우 갱신 주기는 $2^3 \times 4 = 32$ 프레임이다. 이것은 최악의 경우 화질이 복구되는데 걸리는 시간이 대략 1초 정도에 해당한다는 것을 의미한다. 복호기에서 상위 계층을 추가할 확률 혹은 에러가 날 확률이 모든 프레임에서 동일하다고 가정하면 인트

라 갱신에 의하여 문제가 해결될 때까지 소요되는 시간은 1프레임에서 31프레임까지 동일하게 분포하므로 평균적으로 16프레임에 해당하는 시간이 된다.

• 계층형 부대역 갱신 (hierarchical update)

상위 계층일수록 복호시에 버려질 확률이 커지므로 이들 계층의 인트라 갱신 빈도를 높이는 것을 고려해 볼 수 있다. 식 (1)로 주어지는 시축 계층화와 동일한 방식의 계층구조를 가지도록 하면 i번째 공간 계층의 갱신 주기는

$$P_h=\begin{cases} 2^M\times 2^{N_{SL}-1} & (i=1) \\ 2^M\times 2^{N_{SL}-i+1} & (i\geq 2) \end{cases} \quad (8)$$

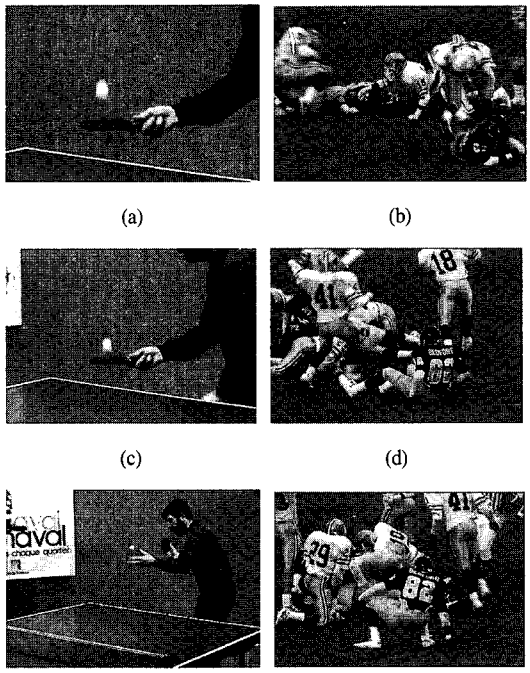
로 주어진다. 앞의 예에서와 마찬가지로 $N_{TL}=4$, $N_{SL}=4$라면 인트라 갱신의 평균 주기는,

첫 번째 계층: $2^3\times 2^3=64$
두 번째 계층: $2^3\times 2^3=64$
세 번째 계층: $2^3\times 2^2=32$
네 번째 계층: $2^3\times 2=16$

으로 구해진다. 이로부터 인트라 갱신에 필요한 평균 시간은 다음과 같이 구해진다. 먼저 모든 프레임에서 오류가 날 확률이 동일하다 가정하면 어떤 오류가 특정 계층에 있을 확률은 특정 프레임이 그 계층에 있을 확률과 같으므로, 그 오류가 첫 번째 계층에 있을 확률은 1/8, 두 번째에 있을 확률은 1/8, 세 번째는 1/4, 마지막으로 네 번째에 있을 확률은 1/2이다. 따라서 임의의 오류 발생 후에 그로 인해 화질이 저하되는 문제가 소멸될 때까지 소요되는 평균 시간은 네 개의 모든 계층을 평균하여 $(1/8\times 64+1/8\times 64+1/4\times 32+1/2\times 16)/2 = 16$ 프레임으로 단순 부대역 갱신의 경우와 동일하다. 하지만 앞서 말한 바와 같이 각 계층이 사용될 확률은 동일하지 않으므로 (상위 계층으로 갈수록 사용될 확률이 낮아짐) 사용 조건에 따라 실제로 갱신에 필요한 시간은 달라지게 된다. 예를 들어 전체 계층을 모두 받으면서 한 개층 정도를 수시로 버리거나 추가하는 경우에는 계층형 부대역 갱신이 유리한 반면 하위 계층 일부만을 복원시에 사용하는 경우에는 계층형 부대역 갱신의 경우 부대역 갱신 주기가 길어지므로 불리하다.

Ⅳ. 실험 결과 및 고찰

제안한 동영상 부호화 기법의 성능을 평가하기 위하여 제안 방식을 352x240 (SIF) 크기의 두 실험 동영상, 'Table Tennis'와 'Football'에 적용하였다. 그림 6에서 이들 동영상의 몇 프레임들을 보여주고 있다.



(a) (b) (c) (d) (e) (f)

그림 6. 실험 동영상 - 'Table Tennis': (a) 10번째, (c) 40번째, (e) 70번째 프레임, 'Football': (b) 10번째, (d) 50번째, (f) 90번째 프레임.

실험에서는 4개의 시간 계층과 4개의 공간 계층이 생성되었으며 Daubechies 4-tap FIR 필터[13]를 사용하여 웨이블릿 변환을 영상 피라미드 분할에 이용하였다. 영상 분할은 3번 수행하여 4단계의 웨이블릿 피라미드를 생성하였다.

1. 시축 계층화의 영향

시축 계층화가 부호화에 미치는 영향을 알아보기 위하여 두 실험 동영상을 그림 2의 부호화기를 이용하여 부호화 하였다. 실험에는 웨이블릿 변환을 사용하였고 양자화기로는 EZW[14] 대신 압축 효율을 좀 더 개선시킨 IEZW (improved EZW)를[15] 사용하여 임베디드(embedded) 비트열을 생성하였으며, 그 결과를 QM 부호화기[16]로 통계적 부호화 하여 최종 비트열을 생성하였다. 그림 7과 그림 8은 시축 계층화에 기반한 움직임 추정을 한 경우와 계층화 없이 기존 방식대로 부호화 경우 프레임별 PSNR 그래프를 보이고 있다. 제안한 계층화 방식의 경우

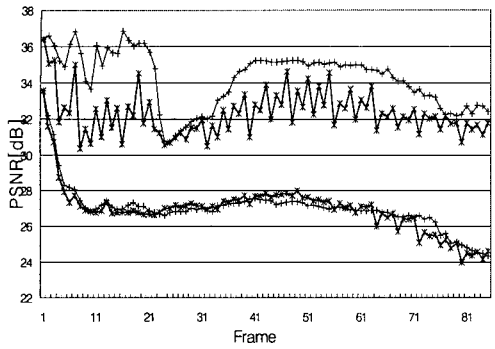


그림 7. 'Table Tennis'의 시축 계층화 유무에 따른 화질 비교 (+: 계층화 없음, *: 계층화 수행, 데이터율: 위에서부터 각각 1.0, 0.2bpp).

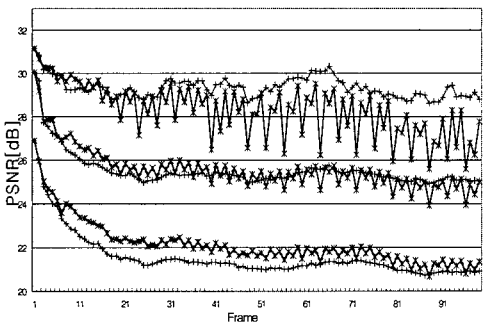


그림 8. 'Football'의 시축 계층화 유무에 따른 화질 비교 (+: 계층화 없음, *: 계층화 수행, 데이터율: 위에서부터 각각 1.0, 0.6, 0.2bpp).

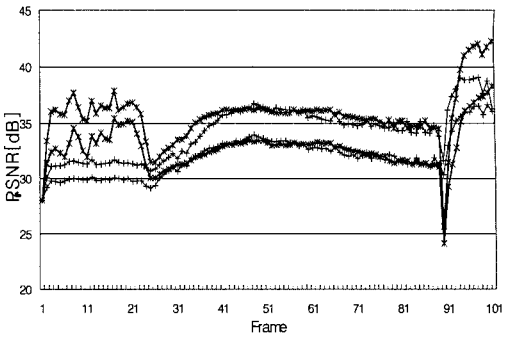


그림 9. 움직임 추정 방법에 따른 'Table Tennis'의 복원 화질 비교 (+: MRME 사용, *: ME 사용, 위에서부터 각각 1.2bpp와 0.8bpp임).

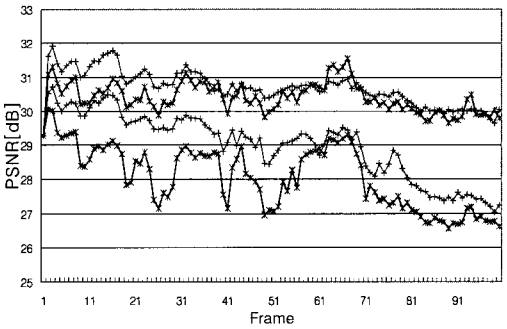


그림 10. 움직임 추정 방법에 따른 'Football'의 복원 화질 비교 (+: MRME 사용, *: ME 사용, 위에서부터 각각 1.2bpp와 0.8bpp임).

PSNR을 보면 그림 8의 1.0bpp 그래프에서 두드러지듯이 기존 계층화를 하지 않은 부호화 방식에 비하여 일정한 간격으로 오르내리는 것을 볼 수 있는데 이는 예측 프레임과 기준 프레임의 거리가 일정하게 변화하는데 기인한다. 실험을 통하여 압축 효율에서 약간의 양보를 통하여 다중 계층 비트열을 생성할 수 있음을 알 수 있었다.

2. MRME의 사용 및 공간 계층화

공간 확장성을 지원하기 위해서 그림 3의 부호화기를 구성하고 양자화는 앞의 실험에서 사용한 IEZW를 이용하였다. 부호화기를 그림 2와 같이 구성하여 원영상을 기반으로 움직임 추정한 경우와 새롭게 MRME를 사용한 방식을 두 실험 동영상에 적용하여 비교하였다. 그림 9와 그림 10에 보여진 결과에서 알 수 있듯이 낮은 비트율에서 MRME는 상대적으로 우수하게 동작하였고, 특히 'Football'과 같이 움직임이 큰 영상에서 그 특성이 두드러졌다. 하지만 'Table Tennis'의 앞부분이나 끝부분과 같이 움직임이 많지 않거나 영상의 일부분에서만 움직임이 발견되는 구간에서는 웨이블릿 영역에서 MRME를 사용하는 것 보다 영상 영역에서 움직임 추정을 하는 편이 우수하였다. 하지만 공간 계층화를 하고 동적으로 공간 해상도를 변경할 수 있도록 하기 위해서는 각 부대역별 움직임 벡터를 얻을 수 있도록 하는 MRME를 사용하는 편이 훨씬 유리하다.

공간계층화를 사용하면 복호시에 하나의 계층을 더 사용함으로써 공간 해상도는 두 배가 된다. 복호시에 공간 계층 중 일부 계층만이 사용되더라도 사용하지 않는 계층에 해당하는 웨이블릿 부대역의 계수값들을 0으로 치환하면 동영상을 전체 공간 해상도로 복원할 수가 있다. 이런 방법으로 데이터율 확장성까지도 얻을 수가 있으며 다음에서 이를 이용한 복호화시의 동적 적응 예를 보일 것이다.

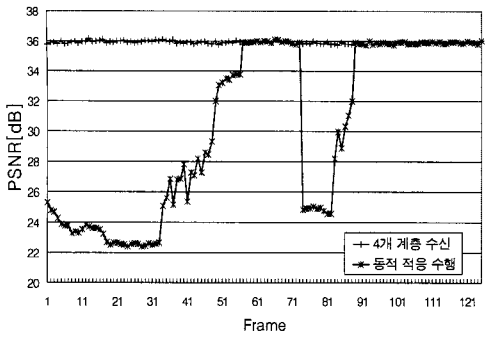


그림 11. 동적 적응의 예 (프레임 0-32: 2 계층, 33-72: 4 계층, 73-81: 3 계층, 82-124: 4 계층을 수신)

3. DISU를 사용한 예

그림 11은 공간계층을 이용하여 수신단에서 동적으로 데이터율을 변경하는 적응의 예를 보여주고 있다. 동영상을 전체 해상도로 복원함으로써 데이터율 적응의 효과를 얻을 수 있고 이 결과를 나타낸 것이다. 주기적인 갱신을 위해서는 단순 부대역 갱신을 사용하였다. 처음 두 개의 공간 계층으로부터 시작하여 복호기는 프레임 33에서 두 개의 계층을 추가하였다. 세 번째, 네 번째 계층이 프레임 48과 56에서 각각 인트라 갱신되면서 복원 화질은 최고 수준까지 도달하였다. 복호기는 프레임 73에서 계층 하나를 버려 화질을 떨어뜨렸으나 프레임 82에서 그 계층을 다시 추가함으로써 프레임 88에서 인트라 갱신이 이루어져 화질이 다시 최고 수준에 복귀하였다. 계층을 추가하고 화질 회복에 걸리는 시간은 추가한 계층에서 얼마나 빨리 인트라 갱신이 이루어지느냐에 달려있다. 위 예에서 최악의 경우에는 31프레임 기간에 해당하는 약 1초의 시간이 요구된다. 이는 적응이 수행되는 시간에 비하여 상대적으로 작은 값에 해당한다. 즉, DISU를 이용하여 비트율에서 피크가 발생하는 것을 억제하면서도 동적 적응에 대응할 수 있음을 알 수 있다.

V. 결 론

본 논문에서는 움직임 추정과 보상을 시축 계층화와 더불어 사용함으로써 시축 확장성을 제공하고 웨이블릿 변환을 기반으로 하여 공간 확장성까지 지원할 수 있는 동영상 부호화 기법을 제안하였다. 제안하는 기법에서는 복원시 동적 적응을 가능케하고 또한 에러가 전파되는 것을 막기 위해서 주기적인 인트라 갱신을 수행하고, 이 때 급격한 비트율의 증가가 일어나지 않도록 각 계층으로 인트라 갱신을 분산시켰다. 실험을 통하여 기존의 움직임 보상 기반 부호화에 비견할만한 우수한 압축 성능을 보이면서 다중 계층을 제공하여 스케일러블한 시축 해상도 및 공간 해상도를 지원할 수 있고, 수신단의 동적 적응에서도 화질 유지가 가능하다는 것을 보였다. 따라서 제안하는 동영상 부호화 기법은 인터넷이나 ATM, 무선망과 같이 상호 운용과 확장성을 필요로 하는 이종 환경에서 멀티미디어 응용들을 위하여 매우 효과적으로 사용될 수 있을 것이다.

참 고 문 헌

[1] S. Jha, P. Wright, and M. Fry, "Playout management of interactive video - an adaptive approach," in *Proc. International Workshop on Quality of Service '97*, vol. 1, pp. 145-156, May 1997.

[2] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. Image Processing*, vol. 3, no. 5, pp. 572-588, Sept. 1994.

[3] J. Tham, S. Ranganath, and A. Kassim, "Highly scalable wavelet-based video codec for very low bit-rate environment," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 1, pp. 12-27, Jan. 1998.

[4] K. Uz, M. Vetterli, and D. LeGall, "Interpolative multiresolution coding of advanced television with compatible subchannels," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 1, no. 1, pp. 86-99, Mar. 1991.

[5] M. Ghanbari, "Two-layer coding of video signals for VBR networks," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 5, pp. 771-781, Jun. 1989.

[6] S. Zafar, Y. Zhang, and B. Jabbari, "Multiscale video representation using multiresolution motion compensation and wavelet decomposition," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 1, pp. 24-35, Jan. 1993.

[7] K. Illgner and F. Muller, "Spatially scalable video compression employing resolution pyramids," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 9, pp. 1688-1703, Dec. 1997.

[8] S. McCanne, M. Vetterli, and V. Jacobson, "Low-complexity video coding for receiver-driven layered multicast," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 6, pp. 983-1001, Aug. 1997.

[9] K. Rose and S. Regunathan, "Towards optimal scalability in predictive video coding," in *Proc. IEEE Int. Conf. Image Processing*, vol. III, pp. 929-933, Oct. 1998.

[10] K. Shen and E. Delp, "Wavelet based rate scalable video compression," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 9, no. 1, pp. 109-122, Feb. 1999.

[11] B.-R. Lee, K.-K. Park, and J.-J. Hwang, "H.263-based SNR scalable video codec," *IEEE Trans. Consumer Electronics*, vol. 43, no. 3, Oct. 1997.

[12] J.-Y. Lee, T.-H. Kim, and S.-J. Ko, "Motion prediction based on temporal layering for layered video coding," in *Proc. Int. Technical Conf. Circuit and Systems, Computers, and Communications*, vol. 1, pp. 245-248, Jul. 1998.

[13] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, no. 2, pp. 205-218, Apr. 1992.

[14] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3445-3462, Dec. 1993.

[15] E.-S. Kang, T. Tanaka, and S.-J. Ko, "Improved embedded zerotree wavelet coder," *Electronics Letters*, vol 35, no. 9, pp. 705-706, Apr. 1999.

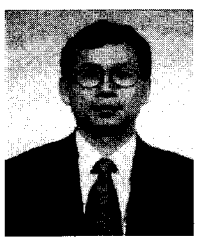
[16] W. Pennebaker and J. Mitchell, *JPEG still image data compression standard*, Van Nostrand Reinhold, New York, 1992.

이 재 용 (Jae-Yong Lee)

1993년 2월 : 고려대학교 전자공학과 졸업 (공학사)
1995년 2월 : 고려대학교 전자공학과 (공학석사)
1999년 8월 : 고려대학교 전자공학과 (공학박사)
1999년 8월~현재 : 현대전자 정보통신연구소 선임연구원
<관심분야> 영상처리, 비디오 코딩 및 전송, 멀티미디어 통신, 네트웍

박 희 라 (Hee-La Park)
1998년 2월 : 인하대학교 전자공학과 졸업 (공학사)
1998년 9월~현재 : 고려대 전자공학과 석사과정
<관심분야>
비디오 코딩 및 전송

고 성 제 (Sung-Jea Ko) 정회원



1980년 2월 : 고려대학교 전자공학과 졸업 (공학사)
1986년 5월 : State Univ. of New York at Buffalo, 전기 및 컴퓨터공학과 (공학석사)
1988년 8월 : State Univ. of New York at Buffalo, 전기 및 컴퓨터공학과 (공학박사)
1981년 8월~1983년 12월 : 대한전선 중앙 연구소 연구원
1988년 8월~1992년 5월 : The Univ. of Michigan Dearbon, 전기 및 컴퓨터 공학과 조교수
1992년 3월~현재 : 고려대학교 전자공학과 교수
1996년 11월 : IEEE APCCAS best paper award
1997년 12월 : 대한전자공학회 해동논문상 수상
IEEE Senior member, IEE member
<관심분야> 신호 및 영상 처리, 영상 압축 및 통신, 멀티미디어 통신