

# 시간 적응 3차원 Subband 부호화 기법

正會員 김 용 관\*, 김 인 철\*\*, 이 상 옥\*

## Temporal Adaptive 3D Subband Image Sequence Coding Technique

Yong Kwan Kim\*, Rin Chul Kim\*\*, Sang Uk Lee\* *Regular Members*

### 요 약

본 논문에서는 입력 동영상의 시간적인 변화량을 감안한 적응 변환을 이용하여 부호화 효율을 높일 수 있는 시간 적응 3차원 SBC 기법을 제안하였다. 본 논문에서 제안한 적응 3차원 SBC에서는 근사식을 이용하여 시간 적응 변환에 필요한 기준을 설정하였다. 이러한 기준을 이용한 적응적인 시간 변환에 의해 인접한 시간 대역과 상관도가 높은 시간 기저 대역 프레임에 대하여는 H.261과 같은 MC-DCT 부호화 기법을 취하고, 시간 상위 대역 프레임들에 대하여 움직임 성분에 의한 다양한 에너지 분포를 고려하여, 공간상의 적응적인 변환 기법인 wavelet packet 변환을 수행하였다. 변환 대역 신호들에 대한 부호화 과정에서, 대역의 특성을 고려하여 적응적으로 재배열한 뒤, 균일 양자화 및 가변 길이 부호화하였고, 쿼드트리(quadtree) 기법을 이용하여 부가정보를 감축하였다. 컴퓨터 모의실험을 통해 제안하는 시간 적응 부호화 기법을 사용함으로써 H.261 또는 고정 변환 3차원 SBC에 비하여 주관적인 화질에서 뿐만 아니라 PSNR 측면에서 약 0.29~3.14 dB 까지 성능이 향상됨을 알 수 있었다.

### Abstract

In this paper, we propose a temporal adaptive transform 3D SBC coder with motion compensation, exploiting redundancy in the temporal domain. We propose a temporal adaptivity measure, by which the R-D optimal temporal transform can be chosen. The base temporal subband frame is coded using H.261-like MC-DCT coder, while the higher temporal subband frames are coded using the 2D adaptive wavelet packet bases, considering the various energy distribution which results from the temporal variation. In encoding the subbands, we employ adaptive scanning methods, uniform step-size quantization with VLC, and coded/not-coded flag reduction technique using the quadtree structure. From the simulation results, the proposed adaptive 3D subband coder shows about 0.29~3.14 dB gain over the H.261 and the fixed 3D subband coder techniques.

\*서울대학교 전기공학부 신호처리 연구실  
Signal Processing Lab, School of Electrical Engineering

\*\*한성대학교 정보전산학부  
School of Information and Computer Engineering

論文番號: 95246-0720

接受日字: 1995年 7月 20日

## I. 서 론

최근 동영상 부호화에서 주파수 대역 분할 부호화(SBC:subband coding) 방법들이 제안되었다. SBC에서는 영상을 여러개의 주파수 대역(subband)들로 분할한 후에, 각각의 특성에 따라 부호화하고 다시 합성한다[1, 2]. 이러한 SBC는 영상 전체에 대하여 변환하므로 저전송을 ब्ल록 변환 부호화에서 나타나는 ब्ल록효과가 발생하지 않는 장점이 있다[2].

3차원 SBC는 [2]에서 처음 제안된 것으로서, 동영상을 시공간상의 3차원 신호로 간주하여 시공간 영역에서 일정한 크기로 대역을 분할한 다음 각 대역 신호를 부호화한다. 즉, 시간 영역과 공간 영역이 분리 가능(separable)하다는 가정하에서, 먼저, 시간 영역에서 영상 신호를 대역 분할하고, 이를 다시 공간 영역에서 분할한다. 이러한 방법은 시공간 영역에서 대역을 분할함으로써 영상의 시공간상의 상관성을 모두 이용할 수 있다는 장점이 있다. 또한, 3차원 SBC는 시공간 주파수 영역에서 나타나는 인간의 시각 감도 특성을 쉽게 적용할 수 있다[2]. 그러나 이러한 3차원 SBC는 움직임이 많은 영상에서는 성능이 저하되는 단점을 나타낸다. 따라서 3차원 SBC에서 움직임 보상을 효과적으로 수행하기 위한 기법들이 제안되었다[3-6].

3차원 SBC에서 이동 보상 기법을 적용하는 방법에 따라서 여러가지 방법들이 제안되었다. [3]에서는 시간 영역에서의 움직임 보상을 수행하는 3차원 SBC 방법이 제안되었다. 그러나 이러한 방법들은 기본적으로 ब्ल록 단위로 시간적인 움직임 보상을 하므로 공간적으로 불연속적인 부분이 발생하여 SBC의 성능이 저하된다는 단점이 있다. 움직임 보상 ब्ल록들에 대한 SBC의 단점을 해결할 수 있는 방법으로서, 전체 영상에 대한 카메라 panning을 보상하는 방법을 이용한 3차원 SBC 부호화 방법 [4]이 제안되었다. 또 다른 방법으로서, [5]에서는 먼저 입력 프레임들에 대하여 시간 대역 분할한 후에, 에너지가 집중되어 있는 시간 기저 대역 신호는 움직임 보상과 2차원 DCT 변환 부호화하고, 움직임 보상에 따른 부호화 효율이 낮은 시간 상위 대역 신호는 직접 2차원 SBC를 적용하는 방법이 제안되었다. 그러나 위와 같은 3차원 SBC 기법들은 시간 영역에서 고정된 시간 주파수 대역 분할

기법을 사용하므로 다양한 동영상의 시간적인 변화를 충분히 고려하기 어렵다는 단점이 있다.

[6]에서는 실험적인 비교를 통하여 3차원 SBC 부호화 기법에서 시간적인 적응 변환 기법에 의해서 부호화 성능이 향상될 수 있음을 보였다. [7]에서는 동영상 부호화에서 이동 보간 방법과 관련하여 시간적으로 일정한 변화량을 갖는 시간 영역을 구분하는 적응적인 알고리즘을 제안하였다. 그러나, [7]에서는 적응적인 시간 영역 구분을 위하여 문턱값을 이용하므로 해석적이지 못하다는 단점이 있다.

본 논문에서는 입력 동영상의 시간적인 변화량을 감안한 시간 적응 변환 기법을 이용하여 부호화 효율을 향상시킬 수 있는 3차원 SBC 부호화 기법을 제안한다. 일반적으로 동영상에서는 시간적인 변화가 심한 부분과 그렇지 않은 부분이 나타난다. 그림 1에서와 같이 입력 동영상을 시간 영역에서 1차원 변환하고 시간 기저 대역에 대해서만 움직임 보상을 하는 경우에, 동영상의 시간적인 변화량이 작은 경우에는 긴 탭(tap)수를 갖는 시간 변환을 적용하는 것이 여러 프레임들에 걸친 상관성을 효과적으로 이용할 수 있으므로 더욱 유리하다. 반대로 시간적인 변화가 많은 경우에는 짧은 탭수를 갖는 변환을 사용함으로써 부호화 효율을 향상시킬 수 있다. 이러한 개념은 공간적인 가변 ब्ल록 크기 변환 개념 [8]을 동영상 신호의 시간적인 영역에 적용한 것으로 생각할 수 있다.

본 논문에서는 시간 적응 변환 기법을 용이하게 적용할 수 있는 움직임 보상 3차원 SBC 부호화기를 제안한다. 제안한 부호화기에서 R-D(rate-distortion) 측면에서 변환 프레임 수에 따른 부호화 성능을 측정할 수 있는 기준을 선정하고, 이 기준을 이용하여 최적의 변환 프레임 수를 결정할 수 있는 시간 적응 변환 알고리즘을 제안한다. 그리고, 공간 주파수 영역상의 특성을 효율적으로 이용할 수 있는 wavelet packet 변환을 수행한다.

본 논문은 서론에 이어 II장에서는 본 논문에서 구상한 적응적인 3차원 SBC 기법을 제시한다. III장에서는 3차원 대역 분할된 신호들에 대한 부호화 기법을 설명한다. 그리고, IV장에서는 모의실험 결과를 검토하고, V장에서는 본 논문의 결과를 제시한다.

## II. 적응 3차원 SBC 기법

본 논문에서 고려하는 3차원 SBC 기법에서는 para-unitary 필터 बैं크를 이용하여 입력 신호를 먼저 시간 방향으로  $M$ 개의 균일 대역폭을 갖는 시간 주파수 대역 프레임들로 나눈다. 이때, 동영상의 시간적인 변화량을 감안한 적응적인 변환 기법을 용이하게 적용하기 위하여, 시간 영역에서 오버랩 필터링(overlap filtering)을 하는 대역 분할 방법[5] 대신에, 시간 영역의  $M$ 개 프레임들에 대하여  $M$ 개의 대역으로 변환하는 DCT 변환을 이용한다. 이러한 DCT 변환은 para-unitary 필터 बैं크의 특수한 경우로서 필터의 차수와 필터 बैं크의 수가 동일한 경우에 해당한다[10].

이러한 블럭 변환을 이용하는 시간적인 적응 변환 기법을 이용하여 입력 영상을 시간 주파수 대역들로 분할하고, 시간적인 변환으로 에너지가 집중된 저대역 프레임들에 대해서는 기존의 H. 261과 같이 MC (motion compensation)-DCT 부호화 기법을 적용한다. 시간 기저 대역에서 움직임 보상에 의한 오차 신호(DFD: displaced frame difference) 전체에 대한 SBC 방법을 고려할 수 있다[1]. 그러나 모의 실험에서 이러한 방법을 적용한 결과, 움직임 보상 블럭들 간의 불연속적인 경계에 의하여 비효율적인 고주파 신호 성분이 발생하여, 평균 자승 오차(MSE: mean square error) 측면에서 부호화 효율이 저하되는 결과를 확인하였다. 따라서 본 논문에서는 H.261과 같은 움직임 보상 블럭 신호에 대한 DCT 부호화 방법을 시간 기저 대역에 대한 부호화 방법으로 취한다. 나머지 시간 상위 대역 프레임들은 에너지가 작고 상호 상관성

이 비교적 낮으므로 움직임 보상을 하지 않는다. 따라서, 신호 전체에 대하여 처리하여 MSE 측면의 효율 뿐만 아니라 시간적으로도 우수한 오버랩 필터링을 이용하는 SBC 기법을 이용한다. 이때, 각 시간 상위 대역의 공간 주파수 영역에서 나타나는 다양한 에너지 분포를 고려하기 위하여 적응적인 대역 분할 기법인 wavelet packet 알고리즘[11]을 적용한다. 3차원 변환 대역 신호들은 양자화한 후에 VLC(variable length coding)를 하고 버퍼를 통하여 고정 전송율로 전송한다. 양자화 스텝 크기(step size)는 버퍼의 상태에 따라서 조절하는 되먹임 버퍼 제어 기법을 도입한다.

### II.1 시간 적응 변환 기법

일반적으로 잘 알려진 MC-DCT 부호화 기법은 동영상을 효과적으로 부호화한다. 그러나, 움직임이 비교적 작은 동영상인 경우에는 3차원 SBC 기법의 성능이 MC-DCT 부호화 기법에 비해 우수하다. 즉, 3차원 SBC는 시간축상으로 다수의 프레임을 이용하여 부호화하는 경우, 프레임간 상관도를 효율적으로 이용할 수 있으므로 움직임이 비교적 작은 영상들에 대해 높은 성능을 보인다. 그러나, 움직임이 복잡해질수록 프레임들 간의 상관도가 낮아지기 때문에 적은 수의 프레임들을 이용하여 부호화하는 것이 효과적이다. 따라서, 입력 영상의 움직임 복잡도를 고려하여 3차원 SBC의 시간축상의 변환 프레임 수를 변화시킨다면 부호화 효율을 향상시킬 수 있을 것이다. 본 절에서는 이러한 목적을 위해, R-D 성능 측면에서 변환 프레임 수에 따른 부호화 성능을 측정할 수 있는 기준을 선정하고, 이를 이용하여 최적의 변환 프레임

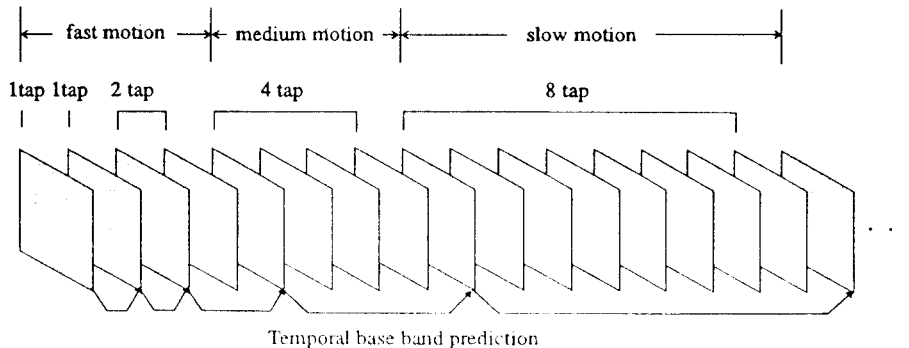


그림 1. 시간 적응 변환 기법의 예.

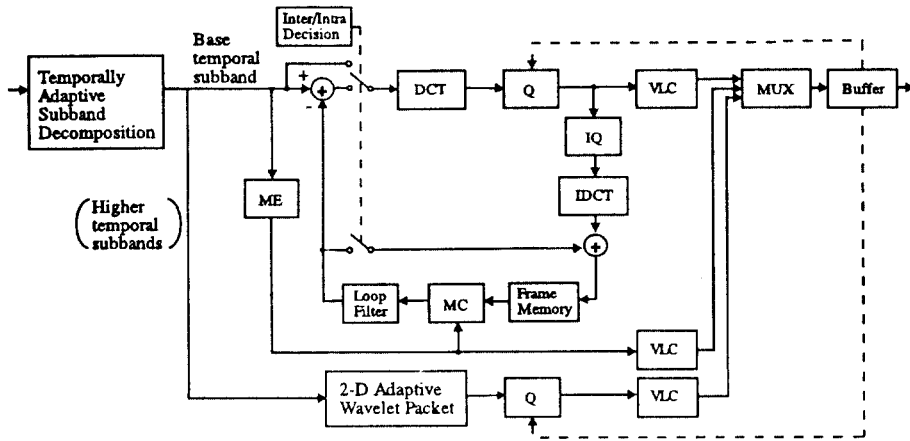


그림 2. 제안하는 적응 3차원 SBC 블록 선도.

수를 효율적으로 추정할 수 있는 시간 적응 변환 기법을 제시한다.

그림 2에 제시한 움직임 보상 3차원 SBC에서는 시간축상으로  $M$ 개의 프레임들에 대한 변환을 수행한 후, 각 시간 대역별로 그 특성에 따라 부호화한다. 이때, 시간 상위 대역 신호들은 공간 영역상의 상관도가 비교적 낮고, 시간 기저 대역 신호는 움직임 보상을 수행하므로 이때 산출되는 DFD 신호 역시 상관도가 낮다. 따라서, 시간축상으로 변환된 신호들은 memoryless Laplacian 확률 밀도 함수로 표현된다고 간주할 수 있다. 만약 각 대역 신호를 균일 양자기(uniform threshold quantizer)로 양자화한 후, 엔트로피 부호화한다면,  $i$ 번째 시간 대역의 왜곡  $d_i$ 는 다음의 식 (1)과 같이 표현할 수 있다[12].

$$d_i = \epsilon_i^2 2^{-2r_i} \sigma_i^2, \quad (1)$$

여기서,  $\epsilon_i^2$ 와  $r_i$ 는  $i$ 번째 시간 대역에 대한 양자기 성능 계수와 전송율을 각각 나타낸다. 그리고,  $\sigma_i^2$ 은  $i = 0$ 이면 시간 기저 대역에서 DFD 신호의 분산을 나타내고,  $i > 0$ 이면  $i$ 번째 시간 대역의 분산을 나타낸다. 이때, 각 시간 대역에 대하여 최적의  $r_i$ 를 얻는 방법은 각 대역에서 발생하는 왜곡을 일정값 이하로 동일하게 유지시키는 것이다[12]. 한편, 각 시간 대역 신호는 각각의 국부적인 시간 대역 특성에 따라, 공간상

에서 다양한 변환이 수행된 후 부호화된다. 만약 입력 신호에 대하여 수행되는 모든 변환들이 변환 후에 에너지가 보존되도록 정규화되어 있고, 변환에 의한 대역 신호들을 균일 양자기를 이용하여 양자화한다면, 모든 대역 신호들을 동일한 균일 양자기로 양자화함으로써 최적 비트 할당이 수행된다. 이러한 사실은 [12]에서 적용하는 최적 비트 할당식 유도 과정으로부터 쉽게 증명할 수 있으므로, 본 논문에서는 생략한다. 위의 최적 비트 할당 조건을 이용하여, 모든 시간 대역 신호들을 동일한 양자기로 양자화한다. 이때, 각 시간 대역 신호들은 동일한 Laplacian 확률 분포를 갖는다고 간주할 수 있으므로 각 시간 대역에 대한 양자기 성능 계수는  $\epsilon^2$ 으로 같다고 볼 수 있다. 따라서, 시간축상으로  $M$  프레임 변환을 수행하는 3차원 SBC의 MSE 측면에서의 왜곡값  $d$ 는 다음의 식 (2)와 같이 전송율  $r$ 의 함수로 나타낼 수 있다.

$$d = \epsilon^2 2^{-2r} \sigma_i^2, \quad (2)$$

여기서  $\sigma_i^2$ 은 시간 대역 신호 분산값들에 대한 기하평균을 나타내며 [12], 다음의 식 (3)과 같다.

$$\sigma_i^2 = \left[ \prod_{i=0}^{M-1} \sigma_i^2 \right]^{\frac{1}{M}}. \quad (3)$$

식 (2)와 식 (3)을 참조하면, 전송율이 일정하다면, 부

호화 왜곡은  $\sigma_i^2$ 의 함수로 주어지고,  $\sigma_i^2$ 은  $M$ 과 밀접한 관계에 있음을 알 수 있다. 따라서, 본 논문에서는 R-D 측면에서 변환 프레임 수에 따른 부호화 성능을 측정할 수 있는 기준으로서 식 (3)을 선정한다. 이때, 영상의 특성에 따라  $M$ 을 적절히 선택함으로써 최적의 부호화 성능을 나타내는 3차원 SBC를 구현할 수 있을 것이다.

이제, 식 (3)을 참조하여 영상의 특성에 따라 최적의 시간축상의 대역수  $M$ 을 결정하는 시간 적응 변환 기법을 설명한다. 시간 적응 변환 기법을 적용하는 경우, 영상내의 물체의 움직임이 없어 프레임간 상관도가 높을 때는  $M$ 이 커질수록 부호화 효율이 증가한다. 그러나 부호화 및 복호화 과정에서 각각  $M$ 개의 프레임 메모리(frame memory)가 필요하므로 최소한  $2M$ 개의 프레임 메모리와 그에 해당하는 지연시간  $((2M-1)/\text{frame-rate}$  초)이 요구되는 단점이 있다[4]. 또한,  $M$ 의 종류가 많아지면, 이에 따라 부호화 복잡도가 크게 증가하는 단점이 있으므로, 부호화 성능을 참조하여  $M$ 의 최대 크기 및 종류를 제한할 필요가 있다. 본 논문에서는, 부호화 지연시간을 1초 이내로 하면서, 부호화 복잡도를 감안하여  $M$ 의 종류를 {1,

2, 4, 8}로 제한하였다. 사실, 본 논문에서의 시간 적응 3차원 SBC를 수행하기 위해서는 기존의 H. 261에서와 같은 MC-DCT 부호화 기법에 비하여 많은 프레임 메모리, 부호화 지연 및 계산량이 필요하다. 그러나, 본 논문에서는 허용한 만한 수준의 시스템 복잡도 증가를 감수하면서, 부호화 성능 향상에 중점을 두고 있다.

일반적으로 식 (3)의  $\sigma_i^2$ 은 프레임간 상관도 혹은 프레임간 움직임 복잡도에 따라 그 크기가 달라진다. 만약 영상의 움직임이 없는 경우에는  $M$ 을 크게할수록 성능이 개선되고, 움직임 복잡도가 높은 경우에는  $M$ 이 작을수록 성능이 우수하다. 만약 장면 변화(scene change) 등으로 인해 움직임이 급격히 변화하는 부분에서는 움직임이 비슷한 프레임들만을 따로 구분하여 시간 변환을 수행한다면, 그 성능이 개선될 수 있을 것이다. 이러한 점들을 참조하면, 부호화 왜곡은  $M$ 의 크기에 따라 국부적인 비교구간내에서 하나의 최소값을 갖는 컨베ックス(convex)한 형태의 함수로 나타난다고 볼 수 있다. 이때,  $M$ 을 증가시키면서 부호화 왜곡 측정자  $\sigma_i^2$ 을 순차적으로 비교함으로써 최적의  $M$ 을 추정할 수 있으므로 계산 시간을 크게 개선시킬

```

[단계 1] 반복 횟수 초기화  $k \leftarrow 1$ ;  $M(k) \leftarrow 1$ ;
        현재 프레임  $f_c$ 에 대한 움직임 보상 오차(DFD)를 구한 뒤에,
        식 (3)의  $\sigma_i^2(k)$ 을 계산한다;  $k \leftarrow k + 1$ ;
[단계 2]  $k$ 번째 변환 대수를  $M(k) = 2^{k-1}$ 로 설정한다.
[단계 3] 현재 프레임  $f_c$ 부터 프레임  $f_{c+M(k)-1}$ 까지  $M(k)$ 덱 변환하고,
        시간 거지 대역 프레임에 대하여 움직임 보상 오차(DFD)를 구한다.
[단계 4] 식 (3)을 이용하여  $\sigma_i^2(k)$ 을 계산한다
[단계 5] If  $\sigma_i^2(k) > \sigma_i^2(k-1)$ ,
        Then {  $M(k-1)$ -덱 변환을 최적의 변환으로 선택한다; Exit; }
        Else If  $M(k) = 8$ ,
        Then {  $M(k)$ -덱 변환을 최적의 변환으로 선택한다; Exit; }
[단계 6] 연속하는  $f_{c+M(k)}$ 부터  $f_{c+2M(k)-1}$ 까지  $M(k)$  프레임들에 대하여,
        [단계 2]와 [단계 3]의 방법을 이용하여  $\hat{\sigma}_i^2(k)$ 를 구한 뒤에,
 $\sigma_i^2(k) \leftarrow \frac{\sigma_i^2(k)+\hat{\sigma}_i^2(k)}{2}$ ;
 $k \leftarrow k + 1$ ; [단계 2]로 간다.
    
```

그림 3. 순차적인 최적 시간 변환 추정 알고리즘.

수 있다. 본 논문에서는 위의 가정을 이용하여 그림 3과 같은 알고리즘을 제시하고, 이를 통하여 입력 영상에 대한 최적의  $M$ 을 선정한다.

그림 3의 알고리즘에서는  $M=1$ 일 때부터 시작하여 순차적으로 최적의  $M$ 을 찾는다. 먼저  $k=1$ 에서 초기화를 한 이후에,  $k$ 번째 비교에서,  $M(k)=2^{k-1}$  프레임들에 대한 시간 대역 변환을 하고 시간 기저 대역에 대한 이동 보상 오차(DFD)를 구한 뒤에, 식 (3)을 이용하여  $\sigma_i^2(k)$ 를 계산한다. 그리고,  $\sigma_i^2(k)$ 를  $\sigma_i^2(k-1)$ 과 비교한다. 만약,  $\sigma_i^2(k)$ 가  $\sigma_i^2(k-1)$ 에 비하여 크면  $M(k-1)$ 을 최적의  $M$ 으로 선택한다. 그렇지 않다면, 현재  $M(k)$ 가 최대값인 8인지를 조사하고,  $M(k)$ 가 8이 아니면, 다음의  $(k+1)$ 번째 비교 과정에서 부호화 성능을 공정히 비교하기 위해서  $M(k+1)=2^k$  프레임 영역에 대한  $\sigma_i^2(k)$ 를 준비한다. 즉, 현재의  $M(k)=2^{k-1}$  프레임들과 연속하는  $M(k)=2^{k-1}$  프레임들에 대하여 식 (3)의  $\sigma_i^2$ 을 각각 구하고 그들의 평균을 취함으로써  $M(k+1)$  프레임 영역에 대하여 적절한  $\sigma_i^2(k)$ 을 구할 수 있다. 그러나, DFD의 순환적인 구조로 인해서, 처음  $M(k)$  프레임들을 부호화하여야 두번째  $M(k)$  프레임 영역에 대한  $\sigma_i^2(k)$ 을 계산할 수 있다. 일반적으로 서로 이웃하는 시간 기저 대역의 DFD 분산값은 비슷하다고 가정할 수 있으므로, 본 논문에서는 계산의 복잡도를 감소시키기 위하여 처음의  $M(k)$  프레임 영역에서의 DFD 분산을 이용하여 연속하는  $M(k)$  프레임들에 대한  $\hat{\sigma}_i^2(k)$ 를 계산하고, 그 값을 이용하여  $M(k+1)$  프레임 영역에 대하여 적절한  $\sigma_i^2(k)$ 를 구한다. 한편,  $M=1$ 일 때에는 시간축상의 변환을 수행하지 않기 때문에 동영상 부호화에 널리 사용되고 있는 MC-DCT 기법으로 귀착된다. 이때,  $M=1$ 일 때에는  $M$ 이 클 때에 비해서 움직임 벡터에 의한 부가정보량이 많기 때문에 실제로 양자화에 할당되는 정보량이 적어진다. 이러한 점을 참조하여  $M=1$ 일 때에는 식 (3)에서 계산된 왜곡값에 실험적으로 구한 계수 1.25를 곱하여 다른 변환과 비교하였다.

제안하는 알고리즘을 실행하기 위한 계산량은 다음과 같다. 한 프레임당 화소수가  $N$ 이고, H.261에서 한 프레임당 이동 보상에 따른 계산량을  $C_{mc}$ 로 표시하는 경우, 제안 알고리즘에 의해서  $M=1, 2, 4, 8$ 이 각각 선택되는 경우, 이동 보상에 따른 계산량은 한 프레임당 각각  $2C_{mc}, 1.5C_{mc}, 1C_{mc}, 0.5C_{mc}$ 이 된다. 만

일, 각 프레임에 대응하는  $M=1, 2, 4, 8$ 이 서로 균일한 빈도수를 나타낸다면, 한 프레임당 이동 보상에 따른 계산량은 평균적으로  $1.25C_{mc}$ 가 된다. 또한, 제안 알고리즘에서 시간 변환에 따른 곱셈수 및 변환 프레임들의 분산값 계산을 위한 곱셈수는 각각  $O(M)$ 에 비례한다.

### II.2 적응적인 wavelet packet 변환 기법

앞에서 설명한 바와 같이, 시간 기저 대역에서는 인접한 시간 기저 대역 신호들 간의 상관도가 높기 때문에 움직임 보상 및 DCT 부호화하는 것이 효율적이다. 그러나, 시간 상위 대역에서는 인접한 신호들 간의 상관도가 낮기 때문에 움직임 보상을 수행하지 않는 것이 적합하다. 따라서, 시간 상위 대역 신호들은 시각적으로 보다 우수한 성능을 보이고, 주파수상의 특성에 따라 적응적으로 부호화할 수 있는 wavelet packet 변환 [11]을 수행한다.

본 논문에서 고찰하는 wavelet packet 변환에서는 1차원상의 2대역 분할 과정을 수평과 수직 방향으로 가분할(separable)하게 적용함으로써 공간 영역의 2차원 신호를 균등한 크기의 4 대역들로 분할하는 과정을 기본으로 한다. 이러한 기본적인 구조를 이용하는 쿼드트리(quadtree) 구조[8]를 사용하여 2차원 공간 영역에 대하여 다양한 대역 분할 형태를 구현할 수 있다. 그림 4(a)에서는 쿼드트리 구조의 한 예를 나타내고 있으며, 그림 4(b)는 그림 4(a)에 대응하는 대역 분할 형태를 나타내고 있다.

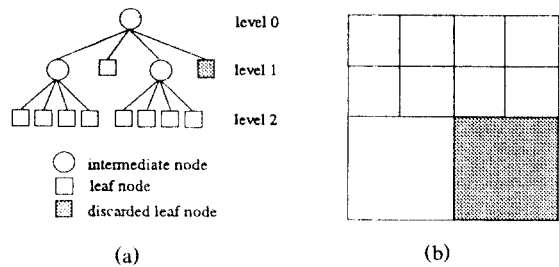


그림 4. 3단계 쿼드트리(quadtree) 구조와 그에 대응하는 대역 분할 형태, (a) 쿼드트리 구조, (b)(a)에 대응하는 대역 분할 형태.

일반적으로, 가능한 모든 wavelet packet 변환들 중에서 적합한 wavelet packet 변환을 찾는 것은 매우 복잡하다. 따라서, 본 논문에서의 wavelet packet 변환은 시간 상위 대역의 휘도 신호에 대하여 기본적으로 쿼드트리 레벨(level) 1에 해당하는 4 대역들로 분할한 상태에서, 각 대역 신호들을 추가적으로 나눌 것인지를 결정하는 방법을 이용하였다. 이때, 최대 쿼드트리 레벨에 해당하는 대역들로 모두 분할된 경우, 균일 16 대역 분할 형태가 되도록 정하였다. 색 신호는 균일 4 대역들로 분할한다. 본 논문에서의 적응적인 공간 대역 분할을 위한 기준으로서는 다음과 같은 부호화 이득식 (4)를 이용한다[12].

$$G_{sbc} = \frac{(\frac{1}{4} \sum_{i=0}^3 \sigma_i^2)}{(\prod_{i=0}^3 \sigma_i^2)^4} \quad (4)$$

식 (4)의 부호화 이득값은 추가적으로 4개의 대역들로 분할함으로써 생기는 부호화 이득을 나타낸다. 본 논문에서는 식 (4)의 부호화 이득값이 실험적으로 정한 일정한 문턱값(5.0)을 넘는 경우, 해당 노드(node)의 대역을 추가적으로 분할한다. 이러한 방법을 이용하면, 특정 방향 공간 주파수 성분이 많은 경우에는 대역 분할에 따르는 부호화 이득식 (4)가 높아져서 추가적인 대역 분할이 이루어지므로 시간 상위 대역에서의 공간적인 에지 성분들을 효과적으로 표현할 수 있다.

적용 wavelet packet 변환 과정의 결과, 분할된 대역들에서 대역의 분산이 실험적으로 정한 일정한 문턱값(10.0) 이상인 경우만 부호화하고 그렇지 않은 대역은 양자화하는 경우 거의 영에 가까워지므로 전송하지 않는다.

### III. 부호화 기법

앞장에서 제안한 시간 적응 변환 알고리즘을 이용하여 움직임 보상을 고려한 변환 부호화 이득이 최대가 되는 M-탭(tap) 변환을 선정한다. M-탭 시간 변환에 의한 기저 대역 프레임은 기존의 H.261과 같은 MC-DCT 부호화 기법을 이용하고, 나머지 시간 상위 대역 프레임들은 공간상의 SBC 기법을 적용한다. 본

절에서는 이와같이 변환된 신호들을 부호화하는 방법을 설명한다.

#### III.1 계수 양자화

시간 기저 대역 신호는 8×8 DCT 변환 후, 그 계수를 지그재그(zigzag) 형태로 재배열하고, 균일 양자화, 그리고 2차원 Huffman 부호화한다. 이에 반해, 시간 상위 대역 신호들에 대해서는 제 II.2절에서 설명한 바와 같이 각 시간 대역에서의 공간적인 주파수 특성을 고려하여 적응적으로 wavelet packet 변환을 수행한다. 그런 다음, 변환 대역 신호를 인접한 위치에 있는 화소들로 구성되는 8×8 혹은 16×16 크기의 블록 단위로 양자화하고, 가변 길이 부호화 후 전송한다. 본 논문에서는 가변 길이 부호화 효율을 보다 향상시키기 위해 wavelet packet 변환 후에, 쿼드트리 레벨 1과 레벨 2에 해당하는 대역들에 대해서 각각 16×16 및 8×8 블록 단위로 부호화한다. 변환에 의한 대역 신호들은 주파수 영역상의 특성을 고려하여 그림 5와 같이 적응적으로 재배열한다. 2차원 변환 대역상에서 수평과 수직 주파수가 동일한 대역은 블록내의 신호가 비교적 균일한 특성을 보이므로 일반적인 재배열 방법인 그림 5(a)와 같은 지그재그 재배열 방법을 사용한다. 그리고, 수직 주파수가 상대적으로 높은 대역에서는 수평 방향 에지가 존재하므로 그림 5(b)와 같이 수평 방향으로 재배열하고, 그 반대의 경우에는 그림 5(c)와 같이 수직 방향으로 재배열한다. 이와 같은 재배열 방법을 사용함으로써 양자화된 신호에서 연속적으로 영이 발생할 확률을 높여, 부호화 성능을 향상시킬 수 있다[1, 13]. 재배열된 대역 신호는 2차원 가변 길이 부호화의 효율을 높이

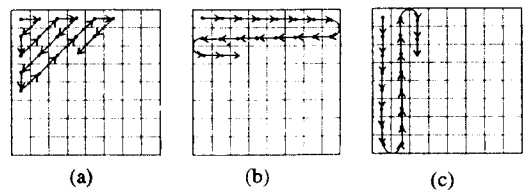
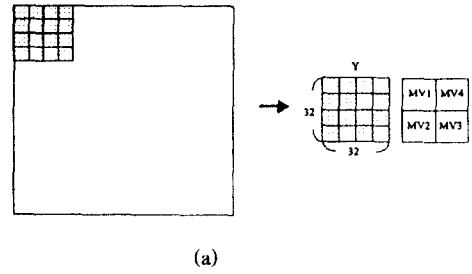


그림 5. 8×8 블록 신호들에 대한 재배열 방법의 예, (a) low-low 또는 high-high 대역인 경우, (b) low-high 대역인 경우, (c) high-low 대역인 경우.

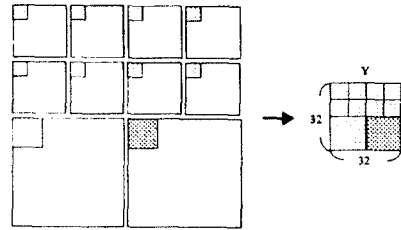
기 위해 양자기에 스텝 크기 만큼 dead-zone을 설정한 균일 양자기 [11]로 양자화한다. 양자화된 신호는 연속된 영의 수와 영이 아닌 값의 크기를 동시에 부호화하는 2차원 가변 길이 부호화되어 전송된다. 그리고, 블럭내 재배열된 나머지 신호가 모두 영인 경우에는 EOB(end of block) 부호를 전송한다. 이때, 가변 길이 부호책은 실험 영상이 아닌 6종류의 MPEG-I 동영상들을 이용하여 Huffman 알고리즘 [13]을 이용하여 구성하였다.

### III.2 부가정보 부호화

시간 상위 대역 신호는 적응적인 wavelet packet 변환 후에, 양자화되어 전송된다. 이때, wavelet packet 변환 정보는 시간 대역에 따라 결정되므로, 매 시간 대역 프레임당 1회 전송된다. 한편, 각 시간 주파수 대역에서, 모든 공간 대역 신호들을 효율적으로 전송하기 위하여 블럭당 양자화 신호 전송 여부(coded/not-coded) 정보를 부호화한다. 본 논문에서는 부호화 효율을 높이기 위해 coded/not-coded 정보를 쿼드트리 구조를 이용하여 전송한다. 먼저, 매 시간 대역 프레임당 결정되는 wavelet packet 변환에 의해 분할된 모든 대역들에서 공간적으로 동일한 위치의 블럭들을 조합하여 그림 6(b)와 같이 매크로블럭(macroblock)을 형성한다. 이때, 블럭의 크기는 wavelet packet 변환 레벨 1에 해당하는 대역에서는  $16 \times 16$  이고, 레벨 2에 해당하는 대역에서는  $8 \times 8$  이므로 이들을 조합하면  $32 \times 32$  크기의 매크로블럭이 결정된다. 그런 다음, 매크로블럭에서의 모든 블럭들은 각 대역 신호의 특성에 맞도록 재배열하여 양자화하고, 양자화 신호가 모두 영이면 not-coded 정보를, 그렇지 않다면 coded 정보를 전송한다. 만약, 동일한 레벨의 모든 블럭들의 양자화 신호가 영이라면 그들의 부모 노드(parent node)를 not-coded로 표시하는 bottom-up 방식으로 coded/not-coded 정보를 부호화한다. 일반적으로, 시간 상위 대역들에서 매크로블럭의 모든 블럭들이 not-coded인 경우가 많으므로, 이런 경우, 레벨 0에서 not-coded로 선언함으로써 coded/not-coded 정보를 효율적으로 감축시킬 수 있다. 그림 7에서는 매크로블럭에 대한 coded/not-coded 쿼드트리 표현 방식을 보여주고 있다. 이때, II.2절에서 설명한 미리 버려진 대역들에 대해서는 미리 그 정보를 알 수 있으



(a)



(b)

그림 6. Y 프레임 신호에서 첫번째 매크로블럭 구성 방법의 예, (a)시간 기저 대역의 경우, (b)시간 상위 대역의 경우.

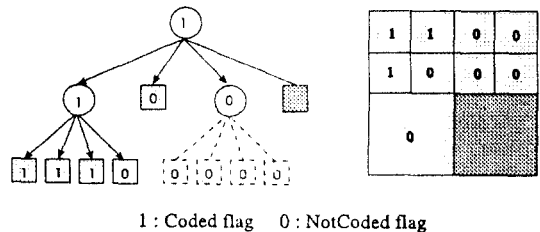


그림 7. 시간 상위 대역에서 한 매크로블럭에 대한 coded/not-coded 쿼드트리 표시 방법의 예.

므로, 이 대역들에 대한 coded/not-coded 정보를 따로 고려할 필요가 없다.

한편, 시간 기저 대역에서 매크로블럭은 그림 6(a)와 같이 시간 상위 대역에서의 크기와 동일하게 공간 영역에서 인접한 신호를 조합하여  $32 \times 32$  크기로 설정한다. 시간 기저 대역에 대한 움직임 보상은  $16 \times 16$  블럭에 대해 수행하므로 하나의 매크로블럭에는 4개



의 움직임 벡터가 대응된다. 움직임 벡터들은 공간적인 연속성을 고려하여 그림 8에서와 같은 순서로 1차원 DPCM 하고, H.261과 동일하게 가변 길이 부호화 하여 전송한다. 또한, 시간 상위 대역에서와 마찬가지로 양자화 신호의 전송 여부에 따라 coded/not-coded 정보를 표현함과 동시에 intra/inter/MC 등의 부호화 모드(mode)를 취한다. 시간 기저 대역 매크로블럭 모드의 표현은 H.261의 방법을 준용하였다.

### III.3 버퍼 제어

버퍼 제어는 버퍼의 충만도에 따라 양자기의 스텝 크기를 조절함으로써 이루어진다. 본 논문에서의 버퍼의 크기는 초당 전송율의  $\frac{1}{10}$  크기로 정하였다.

본 논문에서 고려하는 부호화기는 시간 대역에 따라 DCT 또는 wavelet packet 변환 부호화하기 때문에 각 대역들에 최적으로 전송율을 할당하는 것이 매우 중요하다. II.1절에서 간략히 살펴본 바와 같이 변환들이 정규화되어 있고, 균일 양자기를 사용하여 가변 길이 부호화하는 경우, 모든 시공간 대역 신호들을 동일한 스텝 크기로 양자화함으로써 최적으로 각 대역들에 전송율을 할당할 수 있다. 따라서, 본 논문에서는 공간 영역에서 동일한 위치에 있는 모든 시간 대역의 매크로블럭들을 모아서 GOB(group of block)을 설정하고, GOB 단위로 RM8[14]에서와 같이 버퍼의 상태에 비례하여 2~64 영역에서 양자기의 스텝 크기를 조절함으로써 버퍼를 제어한다.

## IV. 모의실험 결과

본 절에서는 본 논문에서 제안하는 적응 3차원 SBC 기법의 성능을 비교, 평가한다. 제안하는 적응 3차원 SBC 기법은 비교적 저전송율인 512 Kbps에서

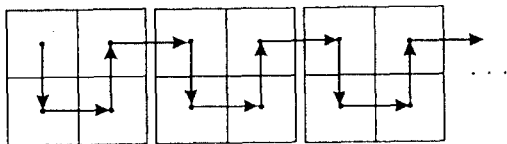
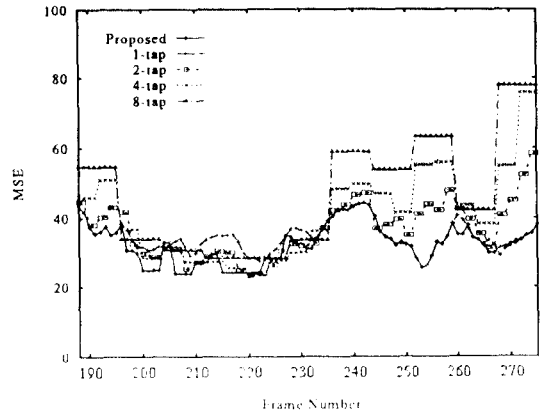
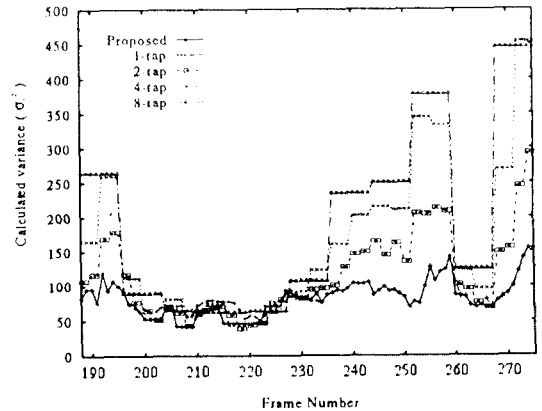


그림 8. 시간 기저 대역에서 연속하는 매크로블럭들에서 이동 벡터 전송 순서.



(a)



(b)

그림 9. 다양한 시간 변환들에 의한 3차원 SBC 성능 척도 (512 Kbps, "Foreman"), (a)부호화 MSE 결과, (b) 근사식에 의한 예상 부호화 왜곡 추정자  $\sigma_d^2$ .

시간적으로 고정된 3차원 SBC 기법 및 저전송율 동영상 부호화 표준안 중의 하나인 H.261[9]과 그 성능을 비교하였다.

본 논문의 3차원 SBC에서 공간 대역 분할/합성 필터로는 비교적 간단하면서도 성능이 우수한 것으로 알려진 Daubechies의 8-탭 wavelet 필터[15]를 사용하였으며, 순환 확장 방식으로 대역 분할/합성을 행하였다.

모의실험에서는 "Foreman" 그리고 "Mother & Daughter" 동영상상을 사용하였고, 동영상들은 휘도와

색 신호의 비율이 4:2:0인 352×288 크기의 CIF 영상으로 초당 25 프레임으로 표현된다. “Foreman” 영상은 순간적인 변화가 심하게 나타나며 프레임 번호가 증가함에 따라 뒤쪽 부분이 급격하게 움직이는 영상으로 이루어져 있다. “Mother & Daughter” 영상은 비교적 단순한 배경과 함께 움직임이 적게 나타나는 영상이다.

먼저, 그림 9에서는 II. 1절에서 제안한 시간 적응 알고리즘의 성능을 평가하였다. 그림 9에서는 “Foreman” 영상을 512 Kbps로 부호화할 때,  $M=1, 2, 4, 8$  인 고정 시간 대역 3차원 SBC와 제안한 적응 3차원 SBC 기법의 성능 척도를 제시한 것이다. 그림 9(a)에서는 각 시간 변환 방법을 이용하여 실제로 부호화하였을 때의 MSE를 도시하였고, 그림 9(b)에서는 식 (3)에서 제시된  $\sigma_i^2$  를 나타내었다. 그림 9(a)를 참조하면, 제안한 기법은 프레임간 움직임 복잡도에 따라 최적의  $M$ 을 적절히 추정하여 최적의 부호화 성능을 얻고 있음을 알 수 있다. 즉, 제안한 기법은 “Foreman” 영상의 전반부와 같이 움직임이 거의 없는 부분에서는  $M=8$  혹은 4인 3차원 SBC의 MSE와 비슷하고, “Foreman” 영상의 후반부와 같이 움직임 복잡도가 높은 경우에는  $M=1$  혹은 2인 3차원 SBC의 MSE와 유사함을 알 수 있다. 또한 그림 9의 (a)와 (b)를 비교하면,  $\sigma_i^2$  은 실제의 부호화 왜곡(MSE)을 충실히 표현하고 있음을 알 수 있다. 따라서, 제 II. 1절에 제시한 시간 적응 알고리즘을 이용한 3차원 SBC가 영상의 움직임 복잡도를 적절히 고려하여 최적의 부호화 성능을 나타냄을 알 수 있었다.

두번째로, 그림 10에서는 “Foreman” 영상을 512Kbps

로 부호화하였을 때, 본 논문에서 제안하는 적응 3차원 변환 기법의 성능을 비교 및 평가하였다. 여기서, 제안한 기법은 일반적으로 잘 알려져 있고 복잡한 움직임에 대하여 부호화 효율이 비교적 높은 H. 261과, 움직임이 작은 영상에서 우수한 성능을 보이는  $M=8$  일때의 3차원 SBC 기법과 함께 비교하였다. 그림 10을 참조하면, 제안한 기법은 프레임간 움직임 복잡도에 따라 최적의  $M$ 을 추정하여 부호화하기 때문에, 다른 기법들에 비해 부호화 성능이 우수한 것으로 나타났다. 표 1에서는 다양한 부호화 기법들을 이용하여 실험 동영상들을 512 Kbps로 부호화하였을 때, 평균적인 부호화 결과들을 나타내었다. 표 1을 참조하면, 제안한 기법은 “Foreman” 동영상을 512 Kbps 전송율로 부호화하였을 경우, H.261에 비해 평균 0.29 dB,  $M=8$ 인 3차원 SBC에 비해 약 1.32 dB 정도 향상된 성능을 보였다. 한편, 그림 11에서는 “Mother & Daughter” 영상을 512 Kbps로 부호화할 때의 성능을 제시하였다. 표 1을 참조하면, 제안한 기법의 PSNR 성능이 H.261 부호화 기법에 비하여 3.14 dB 우수함을 알 수 있었다. 즉, “Mother & Daughter” 영상은 움직임이 작은 영상이므로 제안한 기법은  $M=8$ 인 3차원 SBC의 성능을 따라감을 알 수 있다. 그러나, “Mother & Daughter” 영상의 뒷부분 등은 움직임 복잡도가 달라지는데, 이러한 경우 제안한 기법은 다른 기법에 비해 우수한 성능을 보임을 알 수 있다. 이러한 점들을 참조하면, 제안한 기법은 영상의 특성에 관계없이 견실한 성능을 보임을 알 수 있다.

마지막으로, 그림 12에서는 제안하는 부호화 기법에 의한 복원 영상들과 함께 H.261 부호화에 의한 복

표 1. 512 Kbps에서 다양한 시간 변환 방법들에 의한 부호화 결과 비교.

동영상	평균 결과	제안 방법	1-탭	2-탭	4-탭	8-탭	H.261
Foreman	Bits/frame	20382	20358	20539	20654	20714	20476
	PSNR(Y)	33.10	32.79	32.59	32.23	31.78	32.81
	PSNR(U)	40.32	39.64	39.89	39.78	39.36	39.03
	PSNR(V)	37.59	36.92	37.18	37.18	37.00	36.46
Mother & Daughter	Bits/frame	20065	20147	20176	20188	20057	20250
	PSNR(Y)	41.17	38.47	40.10	40.82	41.10	38.03
	PSNR(U)	44.36	41.75	43.06	43.90	44.32	41.01
	PSNR(V)	43.88	41.89	42.92	43.54	43.86	41.12

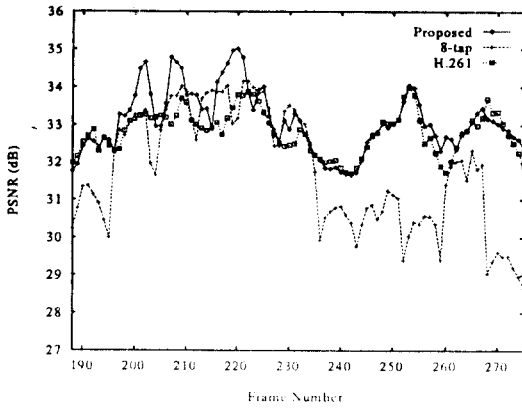


그림 10. "Foreman"에 대한 512 Kbps 부호화 PSNR 결과 비교.

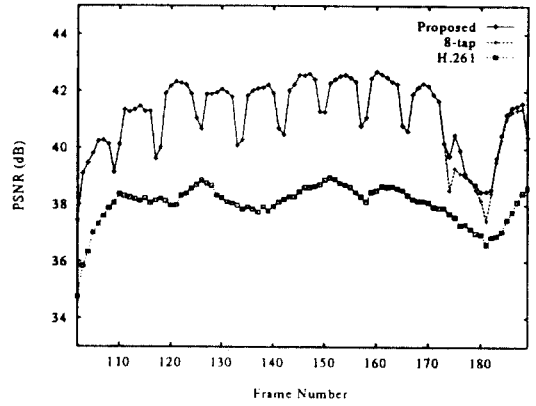


그림 11. "Mother & Daughter"에 대한 512 Kbps 부호화 PSNR 결과 비교.



(a)



(b)



(c)



(d)

그림 12. H.261 부호화 복원 영상과 제안 부호화 복원 영상 비교(512 Kbps), (a) "Foreman", H.261, b) "Foreman", 제안방법, (c) "Mother & Daughter", H.261 (d) "Mother & Daughter", 제안방법.

원 영상들을 비교하였다. 제안하는 부호화 기법에 의한 복원 영상들이 PSNR 측면에서 뿐만 아니라 시각적으로도 기존의 H.261 부호화 결과들에 비하여 우수함을 확인할 수 있었다.

### V. 결 론

본 논문에서는 동영상 부호화에서 시간적으로 적응적인 변환을 이용하는 움직임 보상 3차원 SBC 기법을 제시하고 모의실험을 통하여 그 성능을 비교 및 평가하였다. 본 논문에서 제안하는 움직임 보상 3차원 SBC 기법에서는 최적 비트 할당이 이루어진 것을 가정하였을 때의 시간 변환에 따른 부호화 왜곡 근사식으로부터, 미리 준비된 여러 변환들에 의한 예상 왜곡값들을 구하고, 적응적인 최적 변환 알고리즘을 이용하여 최적의 시간 변환을 추정하였다. 시간적인 기저 대역 프레임에 대하여는 기존의 H.261과 같은 부호화 방법을 취하였고, 나머지 시간 상위 대역 프레임들은 2차원 공간 영역에서 움직임에 의한 다양한 공간 주파수 성분들을 고려한 적응적인 wavelet packet 변환을 수행한 후 부호화하였다. 모의실험으로부터, 시간 적용 변환을 위한 근사화된 왜곡값이 실제의 MSE 왜곡값을 적절히 반영함으로써, 제안하는 시간 적용 변환 3차원 SBC 방법이 매우 효과적임을 확인할 수 있었다. 따라서, 본 논문에서 제시하는 적용 3차원 SBC 기법은 비교적 저전송율에서 full-frame 전송 속도로 동영상을 부호화하는 경우에 적합함을 알 수 있었다.

### 참 고 문 헌

- H. Gharavi, "Subband coding algorithms for video applications: videophone to HDTV-conferencing," *IEEE Trans. Circuit and Systems for Video Technol.* vol. 1, no. 2, pp. 174-183, June 1991.
- G. Karlsson and M. Vetterli, "Three dimensional subband coding of video," *Proc. IEEE Internat. Conf. on Acoust., Speech, and Signal Proc.*, pp. 1100-1103, 1988.
- J. R. Ohm, "Three-dimensional subband coding with motion compensation," *IEEE Trans. Image Processing*, vol. 3, no. 5, pp. 559-571, Sep. 1994.
- D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. Image Processing*, vol. 3, no. 5, pp. 572-588, Sep. 1994.
- K. N. Ngan and W. L. Chooi, "Very low bit rate video coding using 3D subband approach," *IEEE Trans. Circuit and Systems for Video Technol.*, vol. 4, no. 3, pp. 306-316, June 1994.
- Y. K. Kim and S. U. Lee, "3-D subband video coding technique using adaptive wavelet packet bases," *Proc. IEEE Internat. Symposium on Circuits and systems*, vol. 3, pp. 265-268, May 1994.
- J. Lee and B. W. Dickinson, "Temporally adaptive motion interpolation exploiting temporal masking in visual perception," *IEEE Trans. Image Processing*, vol. 3, no. 5, pp. 513-526, Sep. 1994.
- C. T. Chen, "Adaptive transform coding via quadtree-based variable blocksize DCT," *Proc. IEEE Internat. Conf. on Acoust., Speech, and Signal Proc.*, pp. 1854-1857, 1989.
- ITU-T Recommendation H. 261, "Video codec for audiovisual services at  $p \times 64$  kbit/s," Geneva, 1990.
- P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, New Jersey, 1993.
- R. R. Coifman, Y. Meyer and M. V. Wickerhauser, "Signal processing and compression with wavelet packets," *Numerical Algorithms Research Group, Yale University*, 1990.
- N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice-Hall, Englewood Cliffs, New Jersey, 1984.
- D. A. Huffman, "A method for the construction of minimum redundancy codes," *Proc. IRE*, vol. 40, pp. 1098-1101, Sep. 1952.
- CCITT, "Description of reference model 8 (RM8)," SGXV, *Specialist group on coding for visual telephony*, Doc. 525, June 1989.
- I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. on Pure and Applied Math.*, vol. XLI, pp. 909-996, 1988.



김 용 관(Yong-Kwan Kim) 정회원

1965년 3월 9일생

1988년 2월:서울대학교 제어계  
측공학과(학사)

1990년 2월:서울대학교 제어계  
측공학과(석사)

1990년 3월:현재 서울대학교 전  
기공학부 박사과정

※주요 관심 분야: Wavelet 변환 및 영상 압축기법등.



김 인 철(Rin-Chul Kim) 정회원

1985년 2월:서울대학교 제어계  
측공학과(학사)

1987년 2월:서울대학교 제어계  
측공학과(석사)

1992년 2월:서울대학교 제어계  
측공학과(박사)

1992년~1993년:서울대학교 자동  
화시스템 공동연  
구소 특별연구원

1992년~1994년:대우전자 선임연구원

1994년~현재:한성대학교 조교수

※주요 관심 분야:영상신호처리, 영상압축, VLSI 신  
호처리등.

이 상 옥(Sang-Uk Lee) 정회원

1973년:서울대학교 전기공학과(학사)

1976년:Iowa State University(석사)

1980년:University of Southern California(박사)

1980년:General Electric Company, Lynchburg, VA.

1981년~1983년:M/A-COM Research Center, Rockville,  
MD.(Technical Staff)

1983년~현재:서울대학교 전기공학부 교수.

※주요 관심 분야:영상 및 음성 신호처리, 컴퓨터 비  
전등.