

論文2000-37SP-5-5

동영상 부호화기 내부에서의 효과적인 DCT영역 전처리 필터링

(A DCT-Domain Pre-filtering Scheme in a Video Encoder)

金成得*, 羅鍾範*

(Sung Deuk Kim and Jong Beom Ra)

요 약

전처리 필터링은 카메라로부터 들어오는 잡음을 제거하여 부호화 효율을 높여주기 때문에, 전처리 과정의 효과적인 구현은 동영상 압축에서 중요한 연구분야 중의 하나였다. 본 논문에서는 근사화된 일반화 위너 필터링(approximated generalized Wiener filtering)과 이차원 DCT의 분해(factorization)를 바탕으로, 부호화기 내부에서 수행되는 효과적인 전처리 필터링 방법을 제안한다. 제안한 전처리 필터링은 원 영상 블록(original image block)과 움직임 보상된 차 영상 블록(motion-compensated error block)의 DCT계수들에 적절한 값들을 곱하는 것으로 수행된다. 전처리 필터링이 동영상 압축기에 깊이 파묻혀 있지만, 전처리 과정으로 인한 연산량의 증가는 전체 부호화 과정에 비해 크지 않으며, 전통적인 블록 부호화 기법의 틀은 그대로 유지한다. 간결한 구조와 연산에도 불구하고, 제안한 방법은 잡음이 있는 동영상에 대해서 좋은 필터링 및 부호화 성능을 나타낸다.

Abstract

Efficient implementation of pre-filtering has been an important issue in video sequence coding, because it can remove camera noise and improve coding efficiency dramatically. This paper introduces a novel pre-filtering scheme that is performed inside a video encoder. The proposed pre-filtering is based on the approximated generalized Wiener filtering and two-dimensional discrete cosine transform (DCT) factorization; and is achieved by scaling the DCT coefficients of original image blocks for intra block coding and those of motion-compensated error blocks for inter block coding, respectively. Even though the pre-filtering operation is embedded in a video encoder, its additional computational complexity is marginal compared to the encoding process, and the overall architecture of the conventional video encoder is maintained. In spite of its simplicity, the proposed pre-filtering scheme provides good filtering and coding performance for noisy video sequences.

I. 서 론

최근, MPEG-1, MPEG-2, H263과 같은 블록기반의 동영상 압축기들은 동영상 정보를 저장하거나 전송하

기위해 널리 사용되고 있다. 이러한 동영상 압축기들은 동영상에 내재하는 중복적인 정보를 제거해 줌으로써 좋은 압축 효율을 얻게 된다. DCT, 움직임 보상, 가변 부호화는 각각 공간적, 시간적, 통계적 중복성을 제거해 줌으로써 부호화 효율을 높여주는 효과적인 도구들로 잘 알려져 있다. 동영상은 공간적 시간적으로 연관성이 매우 높으므로, 이러한 도구들은 동영상의 압축효율을 크게 높이는 역할을 한다. 그러나 실제적인 동영상 부호화기에 있어서 카메라로부터 들어온 잡음 성분은 그

* 正會員, 韓國科學技術院 電子電算學科
(Department of Electrical and Electronic Engineering,
Korea Advanced Institute of Science and Technology)
接受日字:2000年4月25日, 수정완료일:2000年6月26日

불규칙성으로 인해 영상의 화질 뿐만 아니라 부호화 효율을 나쁘게 한다. 따라서 잡음이 있는 영상이나 동영상 부호화할 때는 영상의 화질과 부호화 효율을 개선하기 위해 전처리 필터링을 적용하는 것이 일반적이다^{[1],[2]}.

동영상에서의 잡음제거 기법들은 많이 연구되어왔다. 그 중에서도, 공간영역의 적응적 위너 필터링과 움직임 보상을 이용한 시-공간적 필터링이 동영상에서 좋은 잡음 제거성능을 보였다^{[3],[4]}. 그러나 이러한 필터링 방식들은 전처리 필터링을 가진 동영상 압축시스템 전체를 최적화하는 입장에서보다 잡음제거 필터링 그 자체적으로 연구되어왔다. 이것은 잡음제거과정을 동영상 압축과는 독립적인 과정으로 생각해 왔기 때문이다. 따라서, 전처리 과정을 동영상 압축에 접목하기 위해서 전처리 필터링과 동영상 압축기를 순차적으로 실행시키는 방식이 일반적이었다. 그러나 이러한 순차적인 구조에서는 추가적인 전처리 과정으로 인해 연산량이 증가한다.

본 논문은 잡음이 있는 동영상을 위해서 전처리 필터링을 포함하는 동영상 압축기를 최적화하는 것을 목표로 한다. 동영상 압축기는 DCT에 기반을 두고 있기 때문에, 동영상 처리를 DCT영역에서 최적화하려는 연구들이 있었다^{[5],[6]}. 이와 유사하게 이 논문에서는, DCT 영역 필터링을 기반으로, 블록기반의 동영상 압축기에서의 효과적인 전처리 필터링 방안을 제안한다. DCT영역에서의 빠른 전처리 필터링을 위한 효과적인 부호화기 구조를 찾기 위해, 근사화된 일반화 위너 필터링의 개념을 활용한다^{[7],[8]}. 제안한 전처리 필터링 연산은 DCT계수들에 적절한 값들을 곱하는 것으로 수행된다. 동영상 압축기는 하나의 블록을 인트라(intra) 혹은 인터(inter) 블록으로 나누어서 DCT영산을 수행하기 때문에 필터링도 또한 두 경우로 구분된다. 인트라블록에서는, 필터링이 원영상 블록에 대해 적용된다. 반면, 인터블록에서는, 블록기반으로 움직임 보상된 차 영상 블록이 필터링된다. 따라서, 전처리 과정이 삽입되었지만 일반적인 동영상 부호화기의 구조는 그대로 유지된다.

근사화된 일반화 위너필터는 평균값이 0인 영상 데이터에 기반을 둔다. 따라서, 평균값이 0이 아닌 영상 데이터를 위해서는, 먼저 평균값을 추정하여 원영상에서 그 값을 빼고 필터링을 수행한다. 마지막으로, 평균값을 필터링된 결과에 더한다. 효과적인 DCT영역에서

의 필터링을 위해서 우리는 곱하기 연산에 기반한 DCT영역 평균값 제거 방식을 이용한다. 여기서 평균값 추정 및 뺄셈, 필터링, 평균값 더하기의 모든 과정이 일원화된 곱하기 과정으로 수행된다. 필터링이 DCT계수에 어떤 값을 곱하는 것으로 수행되기 때문에, 우리는 스케일화 DCT (scaled DCT)를 가지고 이 필터링 과정을 보다 최적하는 방안을 또한 살펴본다^[9].

근사화된 일반화 위너필터는 데이터의 공분산(covariance)에 의존한다. 따라서 평균값을 뺀 인트라 및 인터블록에 대한 적절한 공분산 추정치가 필요하다. 움직임 보상된 차 영상의 공분산 모델을 찾기 위한 몇몇 시도가 있었다^{[10],[11]}. 다소 실험적인 접근을 취한 [10]의 방법과 이론적인 접근의 [11]가 그 예이다. 이 논문에서는, 필터링을 위한 적절한 공분산 추정치를 얻기 위해서, [10]과 유사한 방법으로 MPEG-4 실험영상에 대해 실제적인 공분산 추정치를 구했다.

비록 제안한 전처리 필터링이 동영상 압축기에 깊이 연관되어서 단지 DCT계수에 어떤 값을 곱하는 것으로 수행되지만, 제안한 방식은 특히 인터블록에서 공간영역의 적응적 위너 필터링과 동영상 압축기를 순차적으로 이은 방법에 비해 좋은 필터링 및 부호화 성능을 나타낸다.

이 논문은 다음과 같이 구성된다. 2절에서는 근사화된 위너필터링을 간략히 살펴보고, 평균값이 0이 아닌 영상블록을 위해서 곱하기 연산에 기반한 전체적인 필터링 구조를 제안한다. 3절에서는 잡음이 있는 동영상을 위한 제안한 동영상 압축기에 대해 자세히 다루고 4절에서는 제안한 방법에 대한 실험결과를 보인다. 마지막으로 5절에서 결론을 맺는다.

II. 근사화된 일반화 위너 필터링

1. 개 관

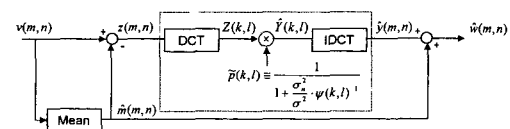


그림 1. 평균값이 0이 아닌 영상에 대한 근사화된 일반화 위너 필터링
Fig. 1. Approximated generalized Wiener filtering for non-zero mean images.

일반화 위너 필터링(generalized Wiener filtering)은 DCT와 같은 빠른 유니터리 변환(unitary transform)을 이용해서 근사적으로 위너 필터를 구현하는 하나의 효과적인 방법이다^[7]. 그림 1은 평균값이 0이 아닌 영상 데이터를 위한 근사화된 일반화 위너 필터링의 블록도를 보여준다. 먼저 v , w , \hat{w} 를 각각 관측된 잡음이 있는 영상 블록, 원하는 영상 블록, 필터링된 영상블록을 나타내는 행방향 열벡터들로 정의하자. v 는 대개 평균값이 0이 아닌 영상 데이터이므로, 관측된 데이터로부터 평균값들을 추정해서 뺀 후에 필터링을 적용한다. 마지막으로 추정된 평균치가 필터링된 데이터에 더해진다. 여기서 \hat{m} , z , y , \hat{y} 는 각각 평균값, 평균값이 빼진 관측 데이터, 평균값이 빼진 원하는 데이터, 평균값이 빼진 필터링된 데이터를 나타내는 행방향 열벡터들이다. 블러(blur)가 없을 경우, 일반화 위너 필터링은 평균값이 0인 영상 모델에 대해서 다음과 같이 기술된다.

$$\hat{y} = A^*T[ALA^*T]Az \equiv A^*T\tilde{L}Z, \quad (1)$$

여기서 $\tilde{L} = ALA^*T$, $L = [I + \sigma_n^2 R^{-1}]^{-1}$, $Z = Az$, $R = E[yy^T]$ 이고 σ_n^2 는 잡음분산(noise variance)을 의미한다. A 는 하나의 유니터리 변환이다. 우리의 경우, $A = (C_8 \otimes C_8)$ 이고 여기서 C_8 와 \otimes 는 각각 8×8 DCT 행렬과 크로네커 곱하기 연산(Kronecker product operator)을 의미한다. 많은 유니터리 변환에서 \tilde{L} 은 근사적으로 대각화가 되기 때문에, (1)은 다음과 같은 식으로 근사화 될 수 있다.

$$\hat{y} = A^*T \hat{Y}, \quad (2)$$

여기서 $\hat{Y} = \tilde{L}Z \approx [\text{diag}\tilde{L}]Z$ 이다. \otimes 블록에 (2)를 대응시키면,

$$\hat{Y}(k,l) = \tilde{p}(k,l)Z(k,l), \quad (3)$$

여기서

$$\tilde{p}(k,l) \equiv \frac{1}{1 + \frac{\sigma_n^2}{\sigma^2} \cdot \psi(k,l)^{-1}} \quad (4)$$

$\psi(k,l)$ 은 ARA^*T 의 대각선에 놓이는 정규화된 요소값 들이고 σ^2 는 y 의 분산을 나타낸다. σ^2 은 대개 z 의 분산에서 잡음분산을 뺀 값으로 추정된다. (3)에서 보듯이, 평균값이 0인 영상모델에 있어서 근사화된 일반화 위너 필터링은 실제적으로는 이차원 DCT계수들을 $\tilde{y}(k,l)$ 로 곱하는 것임을 알 수 있다. 일단 $\hat{y}(m,n)$ 이 얻어지면, 최종적인 필터링 결과 영상 $\hat{w}(m,n)$ 은 $\hat{m}(m,n)$ 과 $\hat{y}(m,n)$ 을 더해서 얻어낸다.

2. 평균값이 0이 아닌 영상 블록에 대해 제한한 필터링 구조

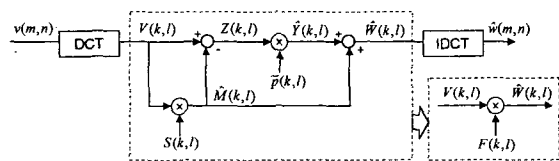


그림 2. 평균값이 0이 아닌 영상에 대한 근사화된 일반화 위너 필터링의 DCT영역 표현

Fig. 2. DCT-domain representation of the approximated generalized Wiener filtering for non-zero mean images.

그림 1에 보여진 평균값이 0이 아닌 영상에 대한 근사화된 일반화 위너 필터링의 블록도는 DCT영역에서 $\hat{M}(k,l) = S(k,l) \cdot V(k,l)$ 의 제한조건하에 그림 2처럼 새롭게 그려질 수 있다. 이 제한조건은 평균값 데이터의 DCT계수값들에 어떤 가중치 행렬 $S(k,l)$ 을 곱함으로써 얻어진다는 것을 의미한다. 이 가정과 식 (3)으로부터, 필터링된 영상데이터는 DCT영역에서 다음과 같이 표현된다.

$$\begin{aligned} \hat{W}(k,l) &= \hat{Y} + \hat{M}(k,l) \\ &= (\tilde{p}(k,l) \cdot (1 - S(k,l)) + S(k,l) \cdot V(k,l)) \\ &= F(k,l) \cdot V(k,l) \end{aligned} \quad (5)$$

여기서

$$\begin{aligned} F(k,l) &= \tilde{p}(k,l) \cdot (1 - S(k,l)) + S(k,l) \\ &= \frac{1 + S(k,l) \cdot \frac{\sigma_n^2}{\sigma^2} \cdot \frac{1}{\psi(k,l)}}{1 + \frac{\sigma_n^2}{\sigma^2} \cdot \frac{1}{\psi(k,l)}} \end{aligned} \quad (6)$$

식 (5)으로부터 전체 필터링 과정이 $F(k, l)$ 와의 하나의 곱하기 연산으로 합쳐지는 것을 볼 수 있다. $F(k, l)$ 는 신호대 잡음비, 공분산 추정치, 평균값 추정 방법에 의존한다.

따라서, 남은 작업은 평균값 추정을 위한 적절한 $S(k, l)$ 를 찾는 것이다. 여기서 우리는 DCT영역에서 $\hat{M}(k, l) = S(k, l) \cdot V(k, l)$ 의 가정을 만족하는 두가지 종류의 $S(k, l)$ 를 선택한다. 가장 간단하고 효과적인 선택은 평균값 블록으로서 88 블록의 DC값을 사용하는 것이다. 즉,

$$S(k, l) = S_1(k, l) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7)$$

또 하나의 $S(k, l)$ 은 평균값 블록이 관측된 블록과 다음의 5×5 평탄화 커널과의 콘볼루션(convolution)을 통해 얻어진다고 가정함으로써 얻어진다.

$$avg(m, n) = \frac{1}{25} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (8)$$

콘볼루션 과정에서, 88 블록 바깥쪽에 있는 픽셀들에 대해서는 경계 픽셀들이 거울에 비친 듯 존재한다고 가정한다. 식 (8)의 콘볼루션 커널은 세퍼러블(separable)한 형태를 가지므로, 이러한 평탄화 과정은 다음과 같이 행렬형태로 표현될 수 있다.

$$\hat{m} = hvh^T, \quad (9)$$

여기서

$$h = \frac{1}{5} \cdot \begin{bmatrix} 2 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & 2 \end{bmatrix} \quad (10)$$

식 (9)를 DCT영역에서 다시 쓰면 다음을 얻는다.

$$\hat{N} = HVH^T, \quad (11)$$

여기서 $\hat{M} = C_8 \hat{m} C_8^T$, $V = C_8 v C_8^T$, $H = C_8 h C_8^T$. 다행히도, 식 (8)에 대한 평탄화 커널에 있어서, H는 대각선 행렬(diagonal matrix)가 되고, 식 (11)은 다음과 같이 보다 간략하게 기술된다.

$$\hat{M}(k, l) = H(k, k) \cdot H(l, l) \cdot V(k, l) = S(k, l) \cdot V(k, l), \quad (12)$$

여기서

$$S(k, l) = S_1(k, l) = \begin{bmatrix} 1.00000 & 0.85239 & 0.48284 & 0.07023 & -0.20000 & -0.23592 & -0.08284 & 0.11329 \\ 0.85239 & 0.72658 & 0.41157 & 0.05986 & -0.17048 & -0.20109 & -0.07061 & 0.09657 \\ 0.48284 & 0.41157 & 0.23314 & 0.03391 & -0.09657 & -0.11391 & -0.04000 & 0.05470 \\ 0.07023 & 0.05986 & 0.03391 & 0.00493 & -0.01405 & -0.01657 & -0.00582 & 0.00796 \\ -0.20000 & -0.17048 & -0.09657 & -0.01405 & 0.04000 & 0.04718 & 0.01657 & -0.02266 \\ -0.23592 & -0.20109 & -0.11391 & -0.01657 & 0.04718 & 0.05566 & 0.01954 & -0.02673 \\ -0.08284 & -0.07061 & -0.04000 & -0.00582 & 0.01657 & 0.01954 & 0.00686 & -0.00939 \\ 0.11329 & 0.09657 & 0.05470 & 0.00796 & -0.02266 & -0.02673 & -0.00939 & 0.01283 \end{bmatrix} \quad (13)$$

III. 잡음이 있는 동영상을 위한 제안한 동영상 압축기

1. 부호화기 내부에서의 전처리 필터링
 평균값이 0이 아닌 영상 데이터에 대해서도 근사화된 일반화 위너 필터링은 변환된 계수에 대한 곱하기 연산으로 수행될 수 있음을 보였다. 이 사실을 바탕으로, 이러한 전처리 필터링이 동영상 부호화기의 내부구조와 어떻게 잘 조화되는가를 보인다.

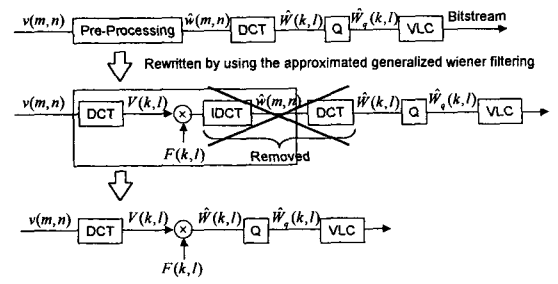


그림 3. 제안된 방식에서의 인트라 블록 부호화
 Fig. 3. Proposed architecture for encoding intra blocks.

그림 3은 인트라 블록들에 대한 전체적인 부호화 과정을 보여준다. 전처리 필터링 과정을 동영상 부호화기

에 결합시킬 때, 인트라 블록 처리에 있어서 워너 필터링에 사용되는 IDCT가 동영상 부호화기에 있는 DCT에 의해 상쇄된다. (이것은 DCT가 근사화된 일반화 워너 필터링을 위한 유니터리 변환으로 선택되었다는 가정에 기반을 둔다.) 이것은 근사화된 일반화 워너 필터링이 동영상 압축기에 결합될 때 단지 한번의 DCT연산만이 필요하다는 것을 의미한다.

그림 3의 개념은 입력잡음이 움직임 보상 데이터 $p(m, n)$ 에 남아있지 않다는 가정하에서 인터블록의 처리에도 유효하다. 따라서, 전통적인 동영상 압축기의 전체적인 구조는 변환계수들이 $F(k, l)$ 라는 가중치 행렬에 의해 곱해진다는 사실을 제외하고는 그대로 유지된다. 그림 4는 인터블록을 처리하기 위한 블록도를 보여준다. 공분산은 인트라 및 인터 블록들에 대해 다르게 얻어지므로 식 (6)에서 $F(k, l)$ 역시 해당 블록의 모드 선택에 의존한다.

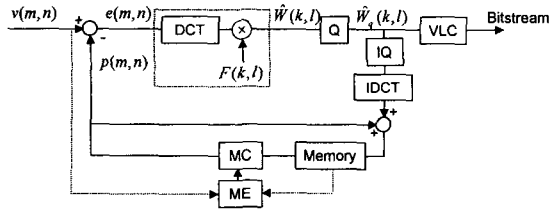


그림 4. 제안된 방식에서의 인터블록 부호화

Fig. 4. Proposed encoder architecture for encoding inter blocks.

2. 전처리 필터링과 스케일화 DCT와의 결합

스케일화 DCT는 잘 알려진 고속 DCT알고리즘으로 DCT를 핵심 연산부, 곱하기 부, 위치 재배열부로 분해한다^[9]. 근사화된 일반화 워너 필터링이 변환 계수들에 대한 곱하기 연산으로 수행되기 때문에 필터링을 위한 곱하기 연산과 DCT자체를 위한 곱하기 연산을 묶어서 보다 최적화할 수 있다. 또한 위치 재배열부는 지그재그 스캔순서(zigzag scanning order)를 적절히 조작함으로써 흡수될 수 있다.

스케일화 DCT에 따르면, 8×8 DCT 행렬 C_8 는 다음과 같이 분해된다.

$$C_8 = P_8 D_8 R_{8,1} M_8 R_{8,2}, \quad (14)$$

여기서, $R_{8,2} = \tilde{B}_1 B_2 B_3$,

$$P_8 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$M_8 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma(4) & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \gamma(4) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma(6) & \gamma(2) \\ 1 & 0 & 0 & 0 & 0 & 0 & -\gamma(2) & \gamma(6) \end{bmatrix}$$

$$R_{8,1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 1 \end{bmatrix}$$

$$\tilde{B}_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$B_3 = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 \end{bmatrix}$$

$$D_8 = \frac{1}{4} \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\gamma^{-1}(4) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma^{-1}(6) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma^{-1}(2) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\gamma^{-1}(5) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\gamma^{-1}(1) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma^{-1}(3) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma^{-1}(7) \end{bmatrix}$$

(15)

$\gamma(k) = \cos(2\pi k/32)$. 여기서 $R_{8,1}M_8R_{8,2}$, D_8 , P_8 는 각각 핵심연산부, 곱하기 부, 위치 재배열부를 의미한다. C_8 는 $R_{8,1}M_8R_{8,2}$ 의 결과에 곱하기 연산을 가하고 그 결과의 위치를 재조정하는 것으로 쉽게 구현된다. 식 (14)는 이차원 DCT로 확장될 수 있다. 즉,

$$\begin{aligned} C_8 \otimes C_8 &= (P_8 D_8 R_{8,1} M_8 R_{8,2}) \otimes (P_8 D_8 R_{8,1} M_8 R_{8,2}) \\ &= ((P_8 D_8) \otimes (P_8 D_8)) ((R_{8,1} M_8 R_{8,2}) \otimes (R_{8,1} M_8 R_{8,2})) \quad (16) \\ &= (P_8 \otimes P_8) (D_8 \otimes D_8) (R_{8,1} \otimes R_{8,1}) (M_8 \otimes M_8) (R_{8,2} \otimes R_{8,2}) \\ &= RD(R_1 MR_2) \end{aligned}$$

따라서, $C_8 \otimes C_8$ 는 $R_1 MR_2$ 연산후의 이차원 곱하기 연산과 위치 재조정연산으로 구현된다.

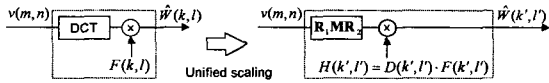


그림 5. 스케일화 DCT와의 최적화
Fig. 5. Joint optimization with the scaled DCT.

따라서, 그림 5에 보여진 것처럼, 스케일화 DCT에 내재하는 곱하기 성분 $D(k',l')$ 와 전처리 필터링을 위한 곱하기 성분 $F(k',l')$ 은 하나의 통일된 곱하기 성분 $H(k',l')$ 으로 표시될 수 있다. 여기서 (k',l') 라는 표시가 (k,l) 를 대신해서 사용되었다. 이것은 스케일화 DCT에서 $DR_1 MR_2$ 의 연산이 끝난 후에도 여전히 위치 재조정부의 연산은 남아 있기 때문이다. 하지만 위치 재조정부는 단지 변환계수의 절대적인 위치를 바꾸는 것이므로, 이러한 위치 재조정 연산은 VLC의 지그

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

(a)

0	14	5	27	1	6	28	15
10	39	23	52	19	32	54	45
3	25	12	41	8	17	43	30
21	50	37	59	34	47	61	56
2	16	7	29	4	13	42	26
9	31	18	44	11	24	53	40
35	57	48	62	36	49	63	58
20	46	33	55	22	38	60	51

(b)

그림 6. (a) 일반적인 지그재그 스캔순서, (b) 스케일화 DCT를 위한 변형된 지그재그 스캔순서

Fig. 6. (a) The normal zigzag scanning order, and (b) the combined scanning order for the permutation of the scaled DCT and the zigzag scanning of the VLC.

재그 스캔 연산과 합쳐 질 수 있다. VLC는 스케일화 DCT의 위치 재조정과 유사한 지그재그 스캔 과정을 필요로 한다. 따라서, 만약 위치 재조정과정이 그림 6(a)에서 보이는 원래의 지그재그 스캔 순서에 반영된다면, 변형된 지그재그 스캔 순서는 그림 6(b)와 같이 유도된다.

요약해 보면, 필터링을 위한 $H(k',l')$ 은 다음과 같이 정리될 수 있다.

$$\begin{aligned} H(k',l') &= D(k',l') \cdot F(k',l') \\ &= \frac{D(k',l') + D(k',l') \cdot S(k',l') \cdot \frac{\sigma_n^2}{\sigma^2} \cdot \frac{1}{\Psi(k',l')}}{1 + \frac{\sigma_n^2}{\sigma^2} \cdot \frac{1}{\Psi(k',l')}} \quad (17) \end{aligned}$$

여기서

$$\begin{aligned} \sigma^2 &= \max(\sigma_z^2 - \sigma_n^2, 0), \\ \sigma_z^2 &= \sum_{k=0}^7 \sum_{l=0}^7 Z^2(k,l) \\ &= \sum_{k=0}^7 \sum_{l=0}^7 (1 - S(k,l'))^2 \cdot D^2(k,l') \cdot V_c^2(k',l'). \quad (18) \end{aligned}$$

$V_c(k',l')$ 는 스케일화 DCT에서 $R_1 MR_2$ 의 연산이 끝난 직후의 변환계수들을 의미한다. 여기서 $D(k',l')$, $S(k',l')$, $\Psi(k',l')$ 및 이것들과 연관된 연산들은 고정된 값을 가지기 때문에 사전에 미리 계산해서 작은

메모리에 저장해 두면 되므로 계산량의 증가는 크지 않다.

IV. 실험결과

실험을 위해 H.263 동영상 압축기를 사용했다. 첫번째 프레임만 인트라 프레임으로 부호화되고 나머지는 인터 프레임으로 부호화된다. H.263의 추가 모드들은 모두 사용되지 않는다. 잡음이 있는 동영상은 MPEG-4 실험영상에 AWGN잡음을 섞어서 만들었다. 인트라 및 인터 프레임에 대해서 고정된 양자화 변수를 사용했고 300프레임을 10Hz의 프레임률로 부호화했다. QCIF형태의 *Hall monitor*, *Mother and daughter*, *Foreman*, *Coast guard* 영상을 실험영상으로 사용했다. 실험결과를 비교하기 위해 공간영역 적응적 위너 필터링을 사용했다^[3].

$$\hat{u}(n_1, n_2) = \hat{m}(n_1, n_2) + \frac{\hat{\sigma}^2(n_1, n_2)}{\hat{\sigma}^2(n_1, n_2) + \sigma_n^2} (v(n_1, n_2) - \hat{m}(n_1, n_2)), \quad (19)$$

여기서

$$\begin{aligned} \hat{m}(n_1, n_2) &= \frac{1}{(2M+1)^2} \sum_{k_1=n_1-M}^{n_1+M} \sum_{k_2=n_2-M}^{n_2+M} v(k_1, k_2), \\ \hat{\sigma}^2(n_1, n_2) &= \max(\hat{\sigma}_z^2(n_1, n_2) - \sigma_n^2, 0), \\ \hat{\sigma}_z^2(n_1, n_2) &= \frac{1}{(2M+1)^2} \sum_{k_1=n_1-M}^{n_1+M} \sum_{k_2=n_2-M}^{n_2+M} (v(k_1, k_2) - \hat{m}(n_1, n_2))^2 \end{aligned} \quad (20)$$

$\hat{u}(n_1, n_2)$ 은 관측된 잡음이 있는 영상 $v(n_1, n_2)$ 의 필터링된 결과를 나타낸다. 공간영역 적응적 위너 필터링의 결과를 얻기 위해서, 전통적인 동영상 부호화기 앞 단에서 잡음이 있는 영상의 각 프레임을 필터링하고 이렇게 필터링된 영상을 부호화한다. 여기서 M 으로는 2가 사용되고, 잡음분산은 안다고 가정한다. 정량적으로 잡음의 양과 복원된 영상의 화질을 측정하기 위해 $\text{SNR} = 10 \log_{10}(\sigma^2/\sigma_n^2)$ 과 $\text{PSNR} = 10 \log_{10}(255^2/\sigma_n^2)$ 을 사용한다. 여기서 σ^2 는 원영상의 분산이고 σ_n^2 는 원영상과 획득된 영상사이의 차영상에 대한 분산이다.

제한된 부호화기를 위해서, 먼저 평균값을 빼 인트라 및 인터 블록에 대한 실제적인 공분산 추정치를 구한

다. 평균값이 빼진 8×8 크기의 블록데이터를 S 로 표현하면, 공분산 행렬의 추정치는 [10]에서 다음과 같이 주어진다.

$$\hat{C} = \frac{1}{2} \frac{N}{\text{trace}(S^T S)} (S^T S + S S^T) \quad (21)$$

이것은 wide sense stationary의 경우 단위분산 값으로 정규화된다. 여기서 블록의 크기 N 은 8이다. 인트라 및 인터 블록에서 평균값을 빼기 위해서 식 (7)과 (13)에 해당되는 평균값 추정방법을 사용하고, 각 방법을 방법1과 방법2로 부른다. 공분산의 추정을 위해서 QCIF형태의 *Akiyo*, *Container ship*, *News*, *Silent voice*라는 4개의 다른 MPEG-4실험 영상을 사용한다. 여기서 300 프레임의 각 동영상은 10Hz의 프레임률로 부호화되고 양자화 변수는 2를 사용했다. 정규화된 공분산 추정치는 프레임단위로 평균해서 얻는다. 그리고 평균값을 빼 인트라 및 인터블록에 대한 정규화된 공분산 추정치를 4가지 영상의 각 프레임에 대해서 최소값, 최대값, 및 전체 평균값으로 그림 7과 그림 8에 각각 나타내었다. 우리는 공분산 추정치의 전체 평균값을 이번 실험에서 사용하였다.

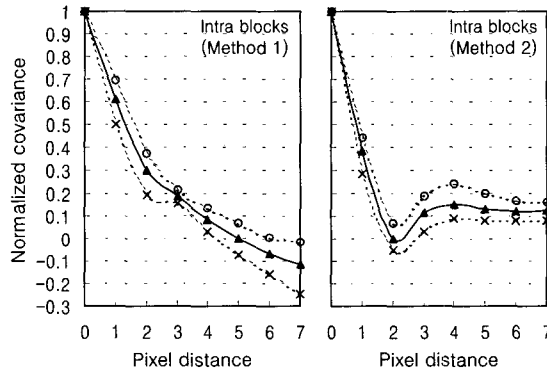


그림 7. 인트라 블록에 대한 정규화된 공분산 추정치. 프레임단위로 평균해서 얻은 공분산추정치의 최소값, 최대값, 및 전체 평균값이 그림에 나타나 있음

Fig. 7. Normalized covariance estimates after frame-based averaging for intra blocks. Depicted are minimal, maximal, and averaged block covariance estimates for all the intra frames in the four sequences.

그림 9와 그림 10은 각각 20dB와 10dB의 SNR을 가

진 잡음이 있는 인트라 프레임에 대한 부호화 성능을 보여준다. 이 그림으로부터, 인트라 프레임에 대한 부호화 효율은 다소 영상의 특성과 잡음의 양에 의존하는 것을 알 수 있다. 그림이 평탄할수록, 공간영역 위너 필터링이 보다 좋은 성능을 주는 반면, *Coast guard*와 같이 세밀한 부분이 많은 복잡한 영상에 대해서는 공간영역 위너 필터링이 다소 나쁜 결과를 낸다. 일반적인 관점에서 세 방법은 PSNR관점에서 비슷한 결과를 내지만, 필터링을 하지 않은 경우에 비해서는 확실히 좋은 결과를 낸다고 말할 수 있다. 그림 10에서 필터링을 수행하지 않는 경우의 부호화 효율은 비트율이 낮을수록 좋아진다. 이것은 저 전송률에서의 넓은 양자화 간격이 필터링 효과를 주어서 잡음이 부분적으로 제거되기 때문이다. 그림 11은 인트라 프레임으로 부호화된 *Hall monitor* 영상을 보여준다. 동일한 양자화 변수에 대해 공간영역 위너 필터링의 경우가 보다 많은 비트를 한 인트라 프레임에서의 영상은 블록단위의 필터링으로 인해 88 블록경계에서 블록화 현상이 나타난다. 블록화 현상은 방법1이 방법2보다 다소 눈에 띈다. 잡음의 양이 적어질수록, 블록화 현상은 눈에 덜 띄게 된다. 양자를 발생시킨다. 그림을 보면, 제안된 방법으로 처리화 간격이 넓어질수록, 블록화 현상을 일으키는 주요원인은 블록단위의 필터링에서 블록단위의 양자화로 바

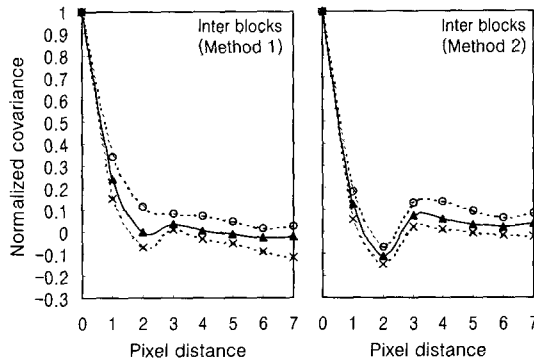


그림 8. 인터 블록에 대한 정규화된 공분산 추정치. 프레임단위로 평균해서 얻은 공분산에측치의 최소값, 최대값, 및 전체 평균값이 그림에 나타나 있음

Fig. 8. Normalized covariance estimates after frame-based averaging for inter blocks. Depicted are minimal, maximal, and averaged block covariance estimates for all the inter frames in the four sequences.

뀌게 된다. 블록화 현상은 MPEG-4에서 권고되었던 후처리 필터링과 같은 방법을 통해 효과적으로 감쇄시킬 수 있다^{[12], [13]}.

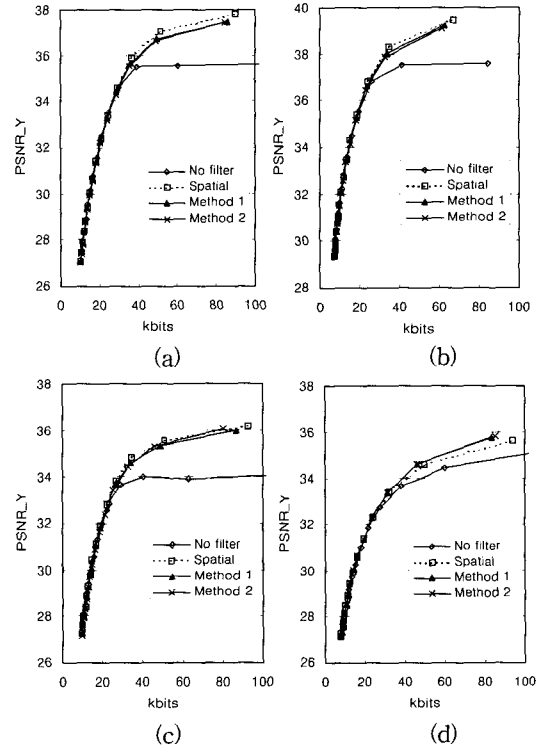
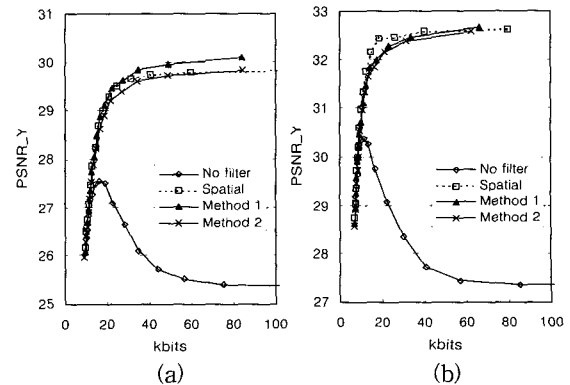


그림 9. 인트라 프레임에 대한 부호화 성능. 잡음이 있는 입력영상의 각 프레임은 20dB의 SNR을 가진다. (a) *Hall monitor*, (b) *Mother and daughter*, (c) *Foreman*, (d) *Coast guard*

Fig. 9. Encoding performance of intra frame. Each frame has an SNR of 20dB. (a) *Hall monitor*, (b) *Mother and daughter*, (c) *Foreman*, (d) *Coast guard* sequences.



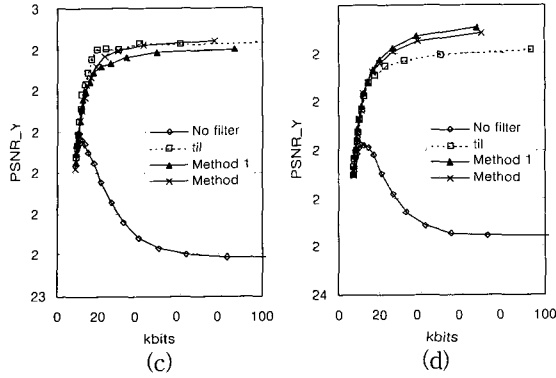


그림 10. 인트라 프레임에 대한 부호화 성능. 잡음이 있는 입력영상의 각 프레임은 10dB의 SNR을 가진다. (a) *Hall monitor*, (b) *Mother and daughter*, (c) *Foreman*, (d) *Coast guard*

Fig. 10. Encoding performance of intra frame. Each frame has an SNR of 10dB. (a) *Hall monitor*, (b) *Mother and daughter*, (c) *Foreman*, (d) *Coast guard* sequences.

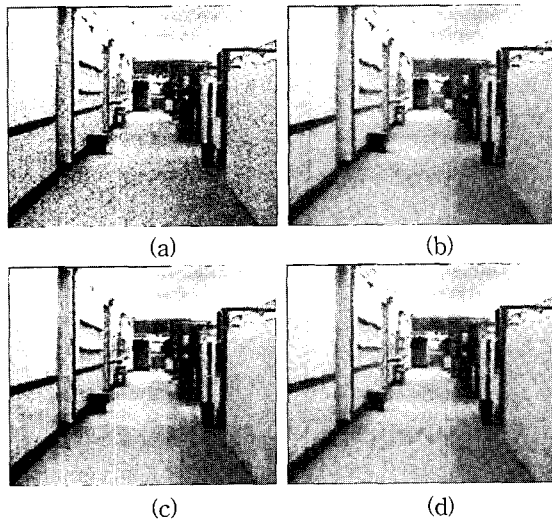


그림 11. 인트라 프레임으로 부호화된 *Hall monitor* 영상. 잡음이 있는 입력영상은 10dB의 SNR을 가진다. 양자화 변수로는 2를 사용. (a) 필터링이 없는 방법 (25.4dB, 182kbits), (b) 공간영역 적응적 위너 필터링 (29.8dB, 107kbits), (c) 방법 1 (30.1dB, 84kbits), (d) 방법 2 (29.8dB, 84kbits)

Fig. 11. Encoded images of the intra frame in *Hall monitor* sequence which is corrupted with a noise of 10dB SNR. A fixed quantization parameter of 2 is used. (a) No filtering (25.4dB, 182kbits), (b) spatial-domain adaptive Wiener filtering (29.8dB, 107kbits), (c) method 1 (30.1dB, 84kbits), and (d) method 2 (29.8dB, 84kbits).

그림 12와 그림 13은 각각 20dB와 10dB의 SNR을 가진 잡음이 있는 인터 프레임에 대한 부호화 성능을 보여준다. 인트라 프레임에서의 결과와 비교하면, 제안된 방법이 동영상 부호화기 앞 단에서 수행되는 공간영역 적응적 위너 필터링에 비해 보다 나은 결과를 보이는 것을 알 수 있다. 이것은 제안된 방법이 인트라 프레임에서보다 인터 프레임에서 보다 좋은 결과를 나타내는 것을 의미한다. 이것은 아마 인터 프레임에서는 움직임 보상을 통해 공간적인 연관성이 이용되기 때문일지도 모른다. 잡음의 양이 많아질수록 두 방법의 PSNR차이는 보다 넓어진다. 영상의 특징에 꽤 의존하던 인트라 프레임에서의 결과와는 달리, 방법 1이 일반적으로 가장 좋은 결과를 낸다. 그러나, 인트라 및 인터

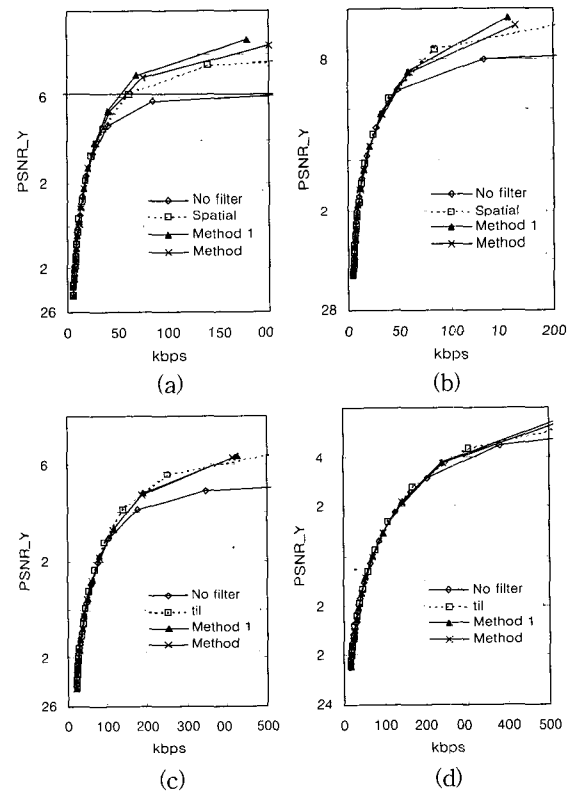


그림 12. 인터 프레임에 대한 부호화 성능. 잡음이 있는 입력영상의 각 프레임은 20dB의 SNR을 가진다. (a) *Hall monitor*, (b) *Mother and daughter*, (c) *Foreman*, (d) *Coast guard*

Fig. 12. Encoding performance of inter frames. Each frame has an SNR of 20dB. (a) *Hall monitor*, (b) *Mother and daughter*, (c) *Foreman*, (d) *Coast guard* sequences.

프레임에 대한 필터링을 전체적으로 고려할 때, 방법 1 과 방법 2의 차이는 그렇게 크지 않다. 이 결과로부터, $S(k, l)$ 의 선택이 전체 필터링 성능에는 크지 않은 영향을 준다고 예측할 수 있다. 그림 13에서는 그림 10과 같이, 필터링이 수행되지 않을 경우, 비트율이 올라갈수록 잡음이 있는 동영상에 대한 PSNR은 낮아지는 것을 알 수 있다.

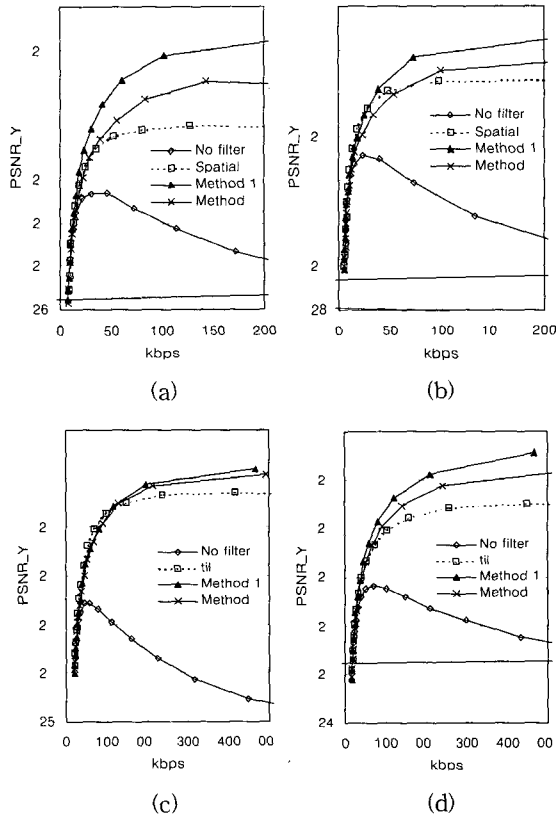


그림 13. 인터 프레임에 대한 부호화 성능. 잡음이 있는 입력영상의 각 프레임은 10dB의 SNR을 가진다. (a) Hall monitor, (b) Mother and daughter, (c) Foreman, (d) Coast guard
 Fig. 13. Encoding performance of inter frames. Each frame has an SNR of 10dB. (a) Hall monitor, (b) Mother and daughter, (c) Foreman (d) Coast guard sequences.

그림 14는 고정된 양자화 변수에 대해서 인터 프레임으로 부호화된 Hall monitor 영상의 210번째 프레임에서의 영상을 보여준다. 제안된 방법이 공간 영역 적응적 위너 필터링에 비해 보다 적은 비트로 보다 편안

한 영상을 보이는 것을 알 수 있다. 인트라 프레임과는 달리 블록화 현상도 거의 눈에 띄지 않는다. 움직임 보상된 차영상 블록은 인트라 블록에 비해 보다 잡음과 유사하므로, DCT계수의 블록단위의 처리는 인트라 블록 보다 인트라 블록에서 심하게 블록화 현상을 유발한다.

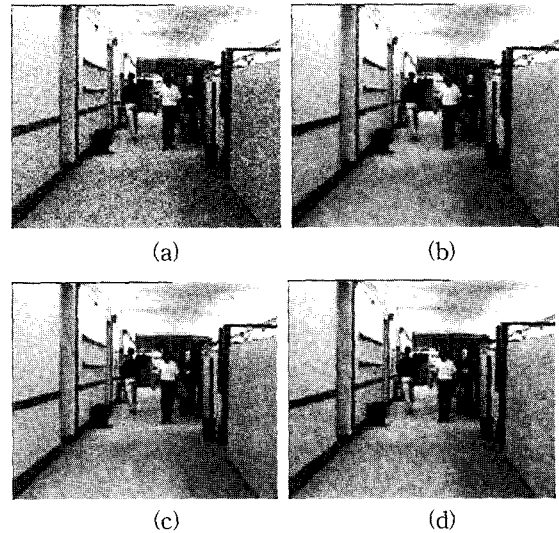


그림 14. 인터 프레임으로 부호화된 Hall monitor 영상의 210번째 프레임에서의 영상. 잡음이 있는 입력영상의 각 프레임은 10dB의 SNR을 가진다. 양자화 변수로는 2를 사용. (a) 필터링이 없는 방법 (25.4dB, 1631kbps), (b) 공간 영역 적응적 위너 필터링 (30.0dB, 826kbps), (c) 방법 1 (32.4dB, 248kbps), (d) 방법 2 (31.1dB, 345kbps)

Fig. 14. Encoded images of the 210th frame in Hall monitor sequence which is corrupted with a noise of 10dB SNR. A fixed quantization parameter of 2 is used. (a) No filtering (25.4dB, 1631kbps), (b) spatial-domain adaptive Wiener filtering (30.0dB, 826kbps), (c) method 1 (32.4dB, 248kbps), and (d) method 2 (31.1dB, 345kbps).

V. 결론

전통적인 동영상 부호화기와 밀접하게 연관된 전처리 필터링 과정을 제안했다. 전처리 필터링과 동영상 압축기를 단순히 나열하는 전통적인 방법과는 달리, 제안한 방법은 동영상 압축기내부에서 원영상 블록과 움직임 보상된 차영상 블록의 변환계수에 곱하기 연산을

수행하는 것으로 이루어진다. 이러한 필터링을 위해서, 근사화된 일반화 위너 필터링과 DCT영역 평균값 추정 방법을 사용하였고 전체 필터링 과정이 하나의 곱하기 연산으로 표현되었다. 또한 평균값이 빠진 인트라 및 인터 블록의 공분산 추정치를 조사하여 인트라 및 인터 블록의 필터링을 위해 사용되었다. DCT는 블록기반의 동영상 압축기에 필수적인 부분이고 전체적인 필터링이 복잡한 콘볼루션 연산대신에 변환계수에 대한 곱하기 연산으로 수행되기 때문에 전처리 필터링으로 인해서 증가되는 계산량은 많지 않고, 전통적인 동영상 부호화기의 구조가 그대로 유지된다.

제안된 방식은 특히 인터블록에서 공간영역의 적응적 위너 필터링과 동영상 압축기를 순차적으로 이은 방법에 비해 좋은 필터링과 부호화 성능을 나타낸다. 구조적인 간결함과 좋은 필터링 성능을 고려할 때 제안된 방법은 잡음환경에서 낮은 가격으로 동영상 압축기를 구현하고자 할 때 유용하게 사용될 것이다.

참 고 문 헌

- [1] O. K. Al-Shaykh and R. M. Mersereau, "Lossy compression of noisy images," *IEEE Trans. Image Processing*, vol. 7, no. 12, pp. 1641-1652, Dec. 1998.
- [2] N. Vasconcelos and R. Dufaux, "Pre and post-filtering for low bit-rate video coding," *Proc. Int. Conf. Image Processing*, vol. 1, pp. 291-294, Santa Barbara, California, Oct. 1997.
- [3] J. S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice-Hall, 1990.
- [4] K. J. Boo and N. K. Bose, "A motion-compensated spatio-temporal filter for image sequences with signal-dependent noise," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 3, pp. 287-298, June 1998.
- [5] R. H. Jansson, "Efficient DCT domain implementation of picture masking and compositing," *Proc. Int. Conf. Image Processing*, vol. 2, pp. 366-369, Santa Barbara, California, Oct. 1997.
- [6] N. Merhav and V. Bhaskaran, "Fast algorithms for DCT-domain image down-sampling and for inverse motion compensation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 3, pp. 468-476, June 1997.
- [7] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, 1989.
- [8] W. K. Pratt, "Generalized Wiener filter computation techniques," *IEEE Trans. Computers*, vol. 21, no. 7, pp. 636-641, July 1972.
- [9] E. Feig and S. Winograd, "Fast algorithms for the discrete cosine transform," *IEEE Trans. Signal Processing*, vol. 40, no. 9, pp. 2174-2193, Sept. 1992.
- [10] W. Niehsen and M. Brunig, "Covariance analysis of motion-compensated frame differences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 4, pp. 536-539, June 1999.
- [11] C.-F. Chen and K. K. Pang, "The optimal transform of motion-compensated frame difference images in a hybrid coder," *IEEE Trans. Circuits Syst. II: Analog Digital Signal Processing*, vol. 40, no. 6, pp. 393-397, June 1993.
- [12] "MPEG-4 video verification model V.8.0," ISO/IEC JTC 1/SC 29/WG 11/N1796, July 1997.
- [13] S. D. Kim, J. Yi, H. M. Kim, and J. B. Ra, "A deblocking filter with two separate modes in block-based video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 156 - 160, Feb. 1999.

저 자 소 개

金 成 得(正會員)

1994년 2월 경북대학교 전자공학과 졸업(공학사). 1996년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학석사). 2000년 2월 한국과학기술원 전자전산학과 졸업(공학박사). 2000년 3월~현재 (주) LG정보통신. 주관심분야는 디지털 신호처리, 영상처리, 영상시스템

羅 鍾 範(正會員)

1975년 2월 서울대학교 전자공학과 졸업(공학사). 1977년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학석사). 1983년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학박사). 1987년 7월~현재 한국과학기술원 전자전산학과 교수. 주관심분야는 디지털 영상 처리, 비디오 신호처리, 3차원 시각화