

논문 2007-44SP-5-9

# H.264/AVC에서 부호화 효율 개선을 위한 매크로 블록 기반 적응 보간 필터 방법

(Macroblock-based Adaptive Interpolation Filter Method for Improving Coding Efficiency in H.264/AVC)

윤 근 수\*, 김 재 호\*\*

(Kun Su Yoon and Jae Ho Kim)

## 요 약

본 논문은 H.264/AVC에서 부호화 효율 개선을 위한 매크로 블록 기반의 적응 보간 필터 방법을 제안한다. 제안 방법은 다양한 방향의 움직임을 세밀하게 보상하는 9가지 분리 가능한 2차원 보간 필터들을 적용한다. 그리고 매크로 블록이 부호화되기 위한 비트율과 왜곡을 고려한 최적의 비용 함수를 정의하고 정의된 비용 함수를 최소화시키는 필터를 매크로 블록 당 적응적으로 선택한다. 실험 결과, 다양한 표준 QCIF(176x144)/CIF(352x288) 동영상 테스트 시퀀스들에 대해서 제안 방법이 기존 방법들에 비하여 항상 우수한 부호화 효율을 지니고 있으며 H.264/AVC 보다 평균 6.25%(참조 영상 프레임: 1개), 3.46%(참조 영상 프레임: 5개)의 비트율이 절감된다.

## Abstract

In this paper, we propose macroblock(MB)-based adaptive interpolation filter method for improving coding efficiency in H.264/AVC. In the proposed method, nine separable two-dimensional(2D) interpolation filters are applied for precisely compensating motions in various directions. The optimal cost function which considers the bit rate and distortion for coding the MB is defined. The filter is adaptively selected per MB for minimizing the defined cost function. In the experimental results, the proposed method shows more excellent in coding efficiency than the conventional methods for the various standard QCIF(176x144)/CIF(352x288) video test sequences. It leads to about 6.25%(1 reference frame) and 3.46%(5 reference frames) bit rate reduction on average compared to the H.264/AVC.

**Keywords :** Adaptive Interpolation Filter, Motion Compensated Prediction, Prediction Mode, H.264/AVC

## I. 서 론

H.264/AVC는 이전의 혼성(hybrid) 동영상 압축 부호화 표준과 같이 움직임 보상 예측(motion compensated prediction, MCP)을 기반으로 하고 있다<sup>[1]</sup>. H.264/AVC에서는 부호화 효율을 개선하기 위해 다양한 블록 크기, 1/4 화소 정밀도, 다중 참조 프레임 등의 움직임 보

상 예측 방법들을 적용하였다<sup>[2~4]</sup>. 그러나 전체 부호화 비트율에 대해서 예측 오차(prediction error)에 대한 비트율은 평균 70% 이상(표준 QCIF/CIF 동영상 테스트 시퀀스들, Y-PSNR: 36 dB)을 차지한다. 따라서 부호화 효율을 개선하기 위해서는 예측 오차에 대한 비트율을 효과적으로 감소시키는 움직임 보상 예측 방법이 개발되어야 한다.

H.264/AVC에서는 1/4 화소 움직임 벡터 해상도(motion vector resolution)를 가지는 움직임 보상을 수행한다. 소수 화소의 움직임을 추정하고 보상하기 위해서 참조 영상의 소수 화소 위치들을 보간한다. 이때 1/2 화소 위치들을 보간하기 위하여 6탭 Wiener 필

\* 정회원, 부산대학교 전자공학과  
(Department of Electronics Engineering, Pusan National University)  
\*\* 정회원, 부산대학교 컴퓨터 및 정보통신연구소  
(Research Institute of Computer Information and Communication, Pusan National University)  
접수날짜: 2007년6월14일, 수정완료일: 2007년8월17일

터를 이용한다. 이 필터는 움직임 보상 예측 효율을 저하시키는 엘리어싱(aliasing) 요소들을 감소시키기 위해 설계되었다<sup>[5]</sup>. 그러나 보간 과정에서 고정된 필터 계수들을 적용하기 때문에 엘리어싱, 양자화 오차, 움직임 추정 오차 등과 같은 동영상 신호의 non-stationary 통계적 특성을 고려하지 못한다. 이를 해결하기 위해 매크로 블록(macroblock, MB) 또는 프레임(frame)기반의 적응 보간 필터(adaptive interpolation filter) 방법들이 연구되었다<sup>[6-11]</sup>.

Chono는 매크로 블록 단위로 적응되고 분리 가능한(separable) 보간 필터들을 제안하였다<sup>[6, 7]</sup>. 이 필터들은 대칭(symmetric) 또는 비대칭(non-symmetric)이고 한정된 움직임 벡터 해상도 보다 움직임을 세밀하게 보상할 수 있다. 그리고 움직임 벡터가 부호화되기 위한 비트율과 예측 오차를 고려하여 비용 함수(cost function)로 정의하였고 정의된 비용 함수를 최소화시키는 필터를 매크로 블록 당 적응적으로 선택하였다. 또한 영상 복원을 위하여 선택된 필터의 정보를 전송하였다.

Chono 이후로 움직임 적은 표준 동영상 테스트 시퀀스에서 특히 부호화 효율을 개선하는 프레임 기반의 적응 보간 필터 방법들이 제안되었다<sup>[8-11]</sup>. Wedi는 프레임 단위로 적응되고 분리 가능한 공간-시간 보간 필터를 제안하였다<sup>[8, 9]</sup>. 이 필터는 예측 오차가 최소화되도록 필터 계수들을 계산하였다. 그리고 영상 복원을 위해 계산된 필터 계수들을 차분 부호화(differential coding)하여 전송하였다. 그러나 이 필터는 움직임 벡터 추정 오차에 민감한 약점이 있다고 알려졌다<sup>[10]</sup>. 이를 보완하기 위하여 Vatis는 프레임 단위로 적응되고 분리 불가능한(non-separable) 2차원 보간 필터들을 제안하였다<sup>[10, 11]</sup>. 이 필터들은 각각의 소수 화소 위치를 보간하고 예측 오차가 최소화되도록 필터 계수들을 계산하였다. 그리고 계산된 필터 계수들은 양자화, 예측, 엔트로피 부호화(entropy encoding)되어 전송하였다.

Wedi, Vatis는 보간 과정에서 프레임 단위로 동영상 신호의 non-stationary 통계적 특성을 고려하여 부호화 효율을 개선하였다. 그러나 매크로 블록과 같은 국부적인 동영상 통계를 고려하는데 효과적이지 않다. 본 논문에서는 Chono의 필터들보다 다양한 방향의 움직임을 세밀하게 보상하기 위해 9가지 분리 가능한 2차원 보간 필터들을 적용한다. 그리고 매크로 블록이 부호화되기 위한 비트율과 복원 오차(reconstruction error)를 고려하여 울-웨곡 측면에서 최적의 비용 함수를 정의하고 정의된 비용 함수를 최소화시키는 필터를 매크로 블록

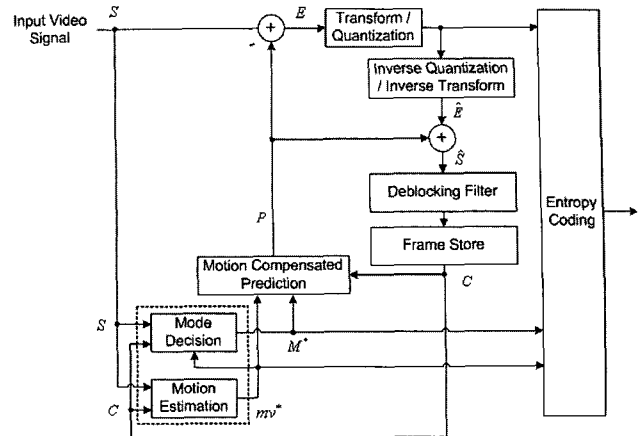


그림 1. H.264/AVC 부호화기의 블록 다이어그램

Fig. 1. Block diagram of H.264/AVC encoder.

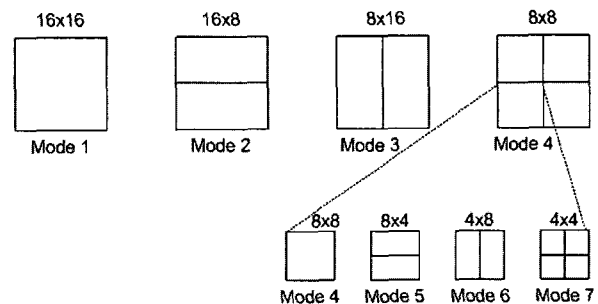


그림 2. H.264/AVC에서 다양한 블록 크기의 모드

Fig. 2. Modes of various block sizes in H.264/AVC.

우수당 적응적으로 선택한다. 실험 결과, 제안 방법은 기존 매크로 블록 기반의 방법과 프레임 기반의 방법들 보다 한 부호화 효율을 지니고 있다.

본 논문의 구성은 다음과 같다. II장에서는 기존의 움직임 보상 예측에 대하여 언급하고, III장에서는 제안 방법을 자세히 설명한다. IV장에서는 제안 방법에 대한 실험 결과를 보여주고, V장에서는 결론을 맺는다.

## II. 기존의 움직임 보상 예측

### 1. H.264/AVC에서의 움직임 보상 예측

그림 1은 H.264/AVC 부호화기의 블록 다이어그램을 나타낸다. 현재 영상  $S$ 는 참조 영상  $C$ 로부터 움직임 보상 예측되고 이 결과 예측 영상  $P$ 가 생성된다. 그리고 현재 영상과 예측 영상이 차분되어 얻어지는 예측 오차  $E$ 와 움직임 정보  $M^*$ ,  $mv^*$ (최적의 예측 모드와 움직임 벡터)는 엔트로피 부호화하고 전송되어진다.

움직임 보상 예측은 현재 매크로 블록을 그림 2와 같이 다양한 블록 크기의 모드들로 분할한다<sup>[2]</sup>. 그리고 해당 블록에 대하여 참조 영상에서 가장 유사한 정합 블

록을 찾는 움직임 추정(motion estimation)을 통해 움직임 벡터를 계산한다. 또한 모드 결정(mode decision)을 통해 최적의 예측 모드와 움직임 벡터를 결정한다.

최적의 예측 모드 및 움직임 벡터를 결정하기 위하여 비트율과 왜곡 정도를 동시에 고려한 율-왜곡 최적화(rate-distortion optimization) 기법<sup>[12]</sup>을 적용한다. 시간적 중복성을 감소시키고 움직임 추정 효율을 개선하기 위하여 식 (1)의 예측 모드(prediction mode)  $M$ 을 정의한다. 현재 매크로 블록이 주어질 경우, 식 (1)의 비용 함수  $J_{MOTION}(m|B, M)$ 을 이용하여 각각의 예측 모드에서 블록  $B$ 에 대한 움직임 벡터를 계산한다.

$$mv^*(B, M) = \arg \min_m A J_{MOTION}(m|B, M) \quad (1)$$

$$= \arg \min_m \left\{ D_{DFD}(m|B, M) + \lambda_{MOTION} \cdot R_{MOTION}(m|B, M) \right\}$$

$$where M \in \left\{ \begin{array}{l} SKIP, 16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, \\ 8 \times 4, 4 \times 8, 4 \times 4, INTRA 4 \times 4, \\ INTRA 16 \times 16 \end{array} \right\}$$

여기서  $mv^*(B, M)$ ,  $\lambda_{MOTION}$ ,  $D_{DFD}(m|B, M)$ ,  $R_{MOTION}(m|B, M)$ 은 예측 모드에 대한 최적의 움직임 벡터, 움직임 추정을 위한 라그랑지안(Lagrangian) 계수, 현재 블록과 참조 블록간의 예측 오차, 움직임 벡터의 비트율을 각각 나타낸다.

예측 모드에 대한 최적의 움직임 벡터를 계산한 이후 식 (2)의 비용 함수  $J_{MODE}(M|MB)$ 을 이용하여 최적의 예측 모드를 매크로 블록 단위( $MB$ )로 결정한다.

$$M^*(MB) = \arg \min_M A J_{MODE}(M|MB) \quad (2)$$

$$= \arg \min_M \left\{ D_{REC}(M|MB) + \lambda_{MODE} \cdot R_{REC}(M|MB) \right\}$$

여기서  $M^*(MB)$ ,  $\lambda_{MODE}$ ,  $D_{REC}(M|MB)$ ,  $R_{REC}(M|MB)$ 는 최적의 예측 모드, 모드 결정 위한 라그랑지안 계수, 현재 매크로 블록과 복원 매크로 블록 간의 복원 오차, 모드  $M$ 과 연관된 비트율을 각각 나타낸다.  $R_{REC}(M|MB)$ 은 예측 오차, 움직임 벡터, 예측 모드, 참조 영상 프레임 등에 대한 비트율을 포함한다. 최적 예측 모드와 움직임 벡터  $mv^*(B, M^*)$ 는 현재 매크로 블록의 예측을 위해 이용되고 엔트로피 부호화하여 전송된다.

식 (1)의  $mv^*$ 을 계산하는데 있어서, H.264/AVC에서는 움직임 벡터 해상도를 1/4 화소 정밀도까지 수행한다<sup>[5]</sup>. 소수 화소 정밀도의 움직임을 추정하고 보상하기

위해서 그림 3에서 나타낸 참조 영상의 소수 화소 위치들을 보간한다. 보간 과정은 다음과 같은 2 단계로 수행된다.

1) 정수 화소로부터 1차원 6탭 Wiener 필터를 수평과 수직 방향으로 필터링하여 1/2 화소 위치들을 보간한다. 이 필터는 예측 효율을 저하시키는 엘리어싱 요소들을 감소시키기 위해 설계되었다<sup>[5]</sup>.

2) 2탭 쌍일차(bilinear) 필터를 이용하여 1/4 화소 위치들을 보간한다.

### 2. 적응 보간 필터를 이용한 움직임 보상 예측

H.264/AVC에서는 1/2 화소 위치들을 보간하기 위하여 위 단계 1과 같이 6탭 Wiener 필터를 이용한다. 그러나 이 필터는 고정된 필터 계수들을 적용하기 때문에 보간 과정에서 엘리어싱, 양자화 오차, 움직임 추정 오

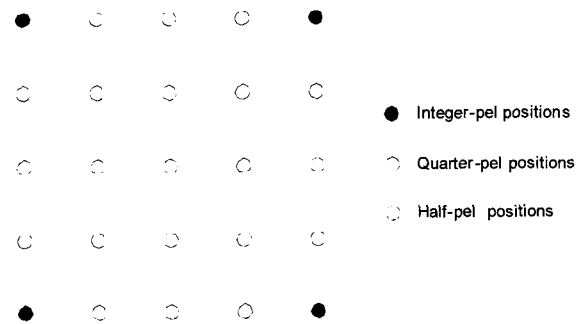


그림 3. 1/4 화소 휘도 보간을 위한 정수와 소수 화소 위치

Fig. 3. Integer-pel and fractional-pel positions for the quarter-pel luma interpolation.

표 1. H.264/AVC와 기존의 적응 보간 필터 방법들을 위한 보간 과정

Table 1. Interpolation process for H.264/AVC and the conventional adaptive interpolation filter methods.

Resolution Method	1/2-pel	1/4-pel	1/8-pel
H.264/AVC	Fixed 6-tap Wiener filter	Fixed 2-tap bilinear filter	-
Wedi	Frame-based adaptive spatial-temporal filter	Fixed 6-tap Wiener filter	Fixed 2-tap bilinear filter
Vatis	Frame-based adaptive 2D non-separable filters		-
Chono	MB-based adaptive symmetric and non-symmetric filters	Fixed 2-tap bilinear filter	-

표 2. Chono가 제안한 대칭과 비대칭 필터

Table 2. Chono proposed a symmetric and non-symmetric filters.

Filter number ( $i$ )	Filter Coefficients
1	[1, -5, 20, 20, -5, 1] / 32
2	[1, -5, 27, 12, -4, 1] / 32
3	[1, -4, 12, 27, -5, 1] / 32

차 등과 같은 동영상 신호의 non-stationary 통계적 특성을 고려하지 못한다. 이를 해결하기 위해 표 1과 같은 매크로 블록 기반 또는 프레임 기반의 적응 보간 필터 방법들이 연구되었다<sup>[8~11]</sup>.

#### (1) 매크로 블록 기반의 적응 보간 필터 방법

Chono는 1/2 화소 위치들을 보간하기 위해 매크로 블록 단위로 적응되고 분리 가능한 보간 필터들을 제안하였다<sup>[6, 7]</sup>. 그러나 1/4 화소 위치들은 고정된 2탭 쌍일차 필터가 이용된다. 표 2는 미리 계산된 Chono의 필터들을 나타낸 것이다. 필터 번호 1은 대칭 필터이고 H.264/AVC의 필터와 동일하다. 하지만 필터 번호 2, 3은 비대칭 필터이고 한정된 움직임 벡터 해상도 보다

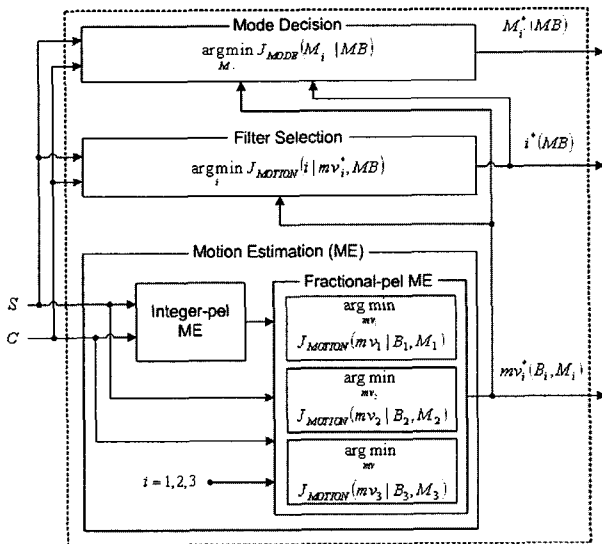


그림 4. Chono의 움직임 보상 예측. 필터 번호  $i$ 는 1/2 화소 위치를 보간하기 위한 각각의 소수 화소 움직임 추정에 이용된다. 1/4 화소 위치들은 쌍일차 필터에 의해 보간된다. 필터 선택은 모드 결정 이전에 수행됨을 주목해야 한다.

Fig. 4. Chono's MCP. The filter number  $i$  is used in each fractional motion estimation for interpolating at half-pel positions. Quarter-pel positions are interpolated by the bilinear filter. Note that the filter selection is performed before the mode decision.

움직임을 세밀하게 보상할 수 있다고 알려져 있다<sup>[7]</sup>.

그림 4는 Chono의 움직임 보상 예측을 나타낸 것이다. 필터 선택은 움직임 추정에서 획득되어진 필터들에 대한 최적의 움직임 벡터를 이용한다. 따라서 모드 결정 이전에 수행됨을 주목해야 한다. 움직임 보상 예측은 다음과 단계로 수행된다.

#### 1) 움직임 추정

현재 매크로 블록과 참조 영상 프레임들을 이용하여 정수 화소 위치에서 최적의 움직임 벡터를 계산한다. 그리고 표 2의 3가지 필터들을 이용하여 1/2 화소 움직임 벡터를 계산한다. 또한 고정된 2탭 쌍일차 필터를 이용하여 1/4 화소 움직임 벡터를 계산한다. 이를 위해서 식 (3)의 비용 함수  $J_{MOTION}(mv_i | B_i, M_i)$ 가 이용된다. 식 (3)은 H.264/AVC의 움직임 추정을 위한 식 (1)을 3번 수행한 것이다. 이는 Chono가 3가지 필터들을 적용했기 때문이다.

$$\begin{aligned}
 mv_i^*(B_i, M_i) &= \arg \min_{mv_i} J_{MOTION}(mv_i | B_i, M_i) \\
 &= \arg \min_{mv_i} \left\{ D_{DFD}(mv_i | B_i, M_i) \right. \\
 &\quad \left. + \lambda_{MOTION} \cdot R_{MOTION}(mv_i | B_i, M_i) \right\} \quad (3) \\
 &\quad \text{for } i=1, 2, 3
 \end{aligned}$$

여기서  $mv_i^*(B_i, M_i)$ 는 필터 번호에 대한 최적의 움직임 벡터를 나타낸다.

#### 2) 필터 선택

그림 5는 필터 선택을 위한 비용 함수의 계산 과정을 묘사한 것이다. 필터는 식 (4)의 비용 함수  $J_{MOTION}(i | mv_i^*, MB)$ 을 이용하여 매크로 블록 단위로 선택된다.

$$\begin{aligned}
 i^*(MB) &= \arg \min_i J_{MOTION}(i | mv_i^*, MB) \\
 &= \arg \min_i \left\{ D_{DFD}(i | mv_i^*, MB) \right. \\
 &\quad \left. + \lambda_{MOTION} \cdot R_{MOTION}(i | mv_i^*, MB) \right\} \quad (4) \\
 &\quad \text{for } i=1, 2, 3
 \end{aligned}$$

여기서  $i^*(MB)$ ,  $D_{DFD}(i | mv_i^*, MB)$ ,  $R_{MOTION}(i | mv_i^*, MB)$ 는 선택된 필터 번호, 위 단계 1에서 계산되어진 움직임 벡터를 이용한 예측 오차, 필터 번호에 대한 최적 움직임 벡터의 비트율을 각각 나타낸다.

#### 3) 모드 결정

선택된 필터 번호  $i^*$ 에 대한 최적의 예측 모드 결정

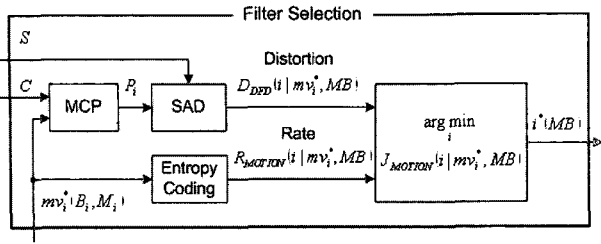


그림 5. 필터 선택을 위한 Chono의 비용 함수의 계산.  
SAD는 sum of absolute difference의 약자이다  
Fig. 5. Computation of Chono's cost function for filter selection. SAD is the sum of absolute difference.

은 식 (5)의 비용 함수  $J_{MODE}(M_i | MB)$ 가 이용된다. 식 (5)은 H.264/AVC에서 모드 결정을 위한 식 (2)와 같고 선택된 필터 번호에 대해서 나타낸 것이다.

$$M_i^*(MB) = \arg \min_{M_i} \left\{ J_{MODE}(M_i | MB) + \lambda_{MODE} \cdot R_{REC}(M_i | MB) \right\} \quad (5)$$

여기서  $M_i^*(MB)$ 는 선택된 필터 번호에 대한 최적의 예측 모드를 나타낸다. 따라서 이 단계를 통해 선택된 필터 번호에 대한 최적의 예측 모드와 움직임 벡터  $mv_i^*(B_i, M_i)$ 가 결정되어 진다.

4) 예측

현재 매크로 블록은  $mv_i^*(B_i, M_i)$ 을 이용하여 예측 된다.

움직임 보상 예측 이후 선택된 필터 번호와 엔트로피 부호화된  $mv_i^*(B_i, M_i)$ 는 영상 복원을 위해 전송한다.

(2) 프레임 기반의 적응 보간 필터 방법

Chono 이후 프레임 기반의 적응 보간 필터 방법들이 제안되었다<sup>[8~11]</sup>. 이는 움직임이 적은 표준 동영상 테스트 시퀀스들에서 Chono 보다 부호화 효율이 개선되는 측면이 있다. Wedi는 보간 과정에서 엘리머싱이 예측 오차에 미치는 영향을 수학적으로 분석하였다<sup>[8]</sup>. 그리고 프레임 단위로 적용되고 분리 가능한 공간-시간 보간 필터를 제안하였다<sup>[8,9]</sup>. 이 필터는 예측 오차를 최소화 하는 필터 계수들을 계산하고 표 1과 같이 1/2 화소 위치들을 보간할 때 이용된다. 그러나 1/4, 1/8 화소 위치는 고정된 6탭 Wiener 필터와 2탭 쌍일차 필터를 각각 이용한다. 그리고 계산된 필터 계수들은 프레임 단위로

차분 부호화하고 전송된다. 그러나 Wedi의 필터는 움직임 벡터 추정 오차에 민감하다는 약점을 가지고 있었다. 그러므로 부호화 효율 개선은 1/8 화소 움직임 벡터 해상도와 결합될 때 성취할 수 있다고 알려졌다<sup>[10]</sup>.

Vatis는 Wedi의 약점을 보완하기 위하여 프레임 단위로 적용되고 분리 불가능한 2차원 보간 필터들을 제안하였다<sup>[10,11]</sup>. 이 필터들은 표 1과 같이 각각의 소수 화소 위치를 보간하고 전체 54개의 필터 계수들을 가진다. 그리고 예측 오차를 최소화 하는 필터 계수들을 계산한다. 또한 54개의 필터 계수들을 양자화 및 공간-시간 예측하고 이 결과를 엔트로피 부호화하여 전송한다.

III. 제안 방법

프레임 기반의 기존 방법들은 보간 과정에서 프레임 단위로 동영상 신호의 non-stationary 통계적 특성을

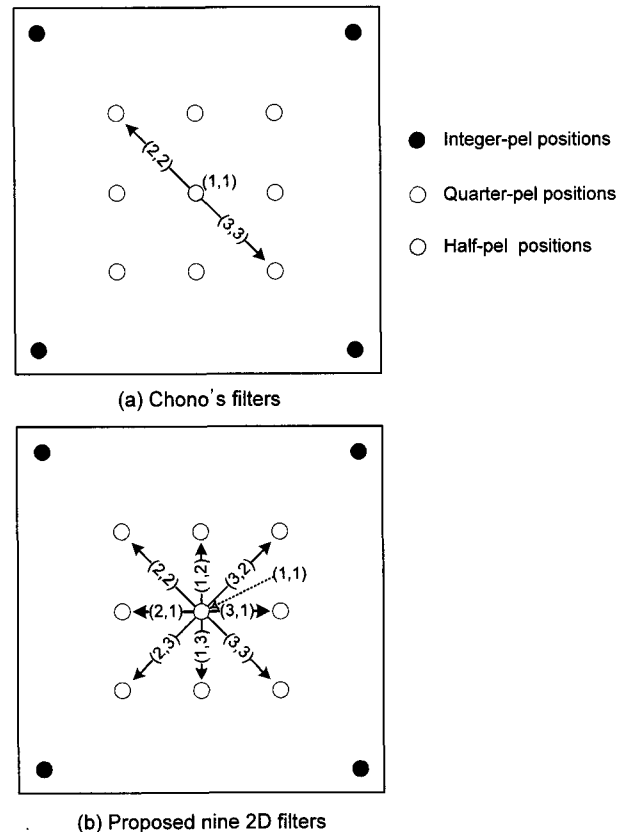


그림 6. Chono 필터들과 제안된 9가지 2차원 필터들의 움직임 보상. m과 n은 수평과 수직 방향의 필터 번호를 각각 나타낸다

Fig. 6. Motion compensation with Chono's filters and the proposed nine 2D filters. Note that m and n are the filter number of horizontal and vertical direction, respectively.

고려하여 부호화 효율을 개선하였다. 그러나 매크로 블록과 같은 국부적인 동영상 신호의 통계를 고려하는데 효과적이지 않다. 따라서 본 논문에서는 Chono가 제안한 방법에서 다음의 내용을 개선하는데 목적을 둔다.

- 다양한 방향의 움직임을 세밀하게 보상하기 위해 9 가지 분리 가능한 2차원 보간 필터들을 적용한다.
- 최적의 필터 선택을 위해 매크로 블록이 부호화되기 위한 비트율과 왜곡을 고려한 비용 함수를 정의한다.

Chono는 그림 6(a)와 같이 가로 방향과 세로 방향으로 동일한 필터를 적용하여 1/2 화소 위치를 보간한다. 이는 한정된 움직임 벡터 해상도 보다 오직 2가지 방향의 움직임을 세밀하게 보상할 수 있다. 그러나 제안된 필터들은 다양한 방향의 움직임을 세밀하게 보상할 수 있다. 그림 6(b)에서 볼 수 있듯이 제안된 9가지 필터들은 표 2의 필터 번호를 가로와 세로 방향의 조합으로 획득된다. 이 필터들은 움직임 벡터 해상도 보다 주위 8 방향의 움직임을 세밀하게 보상할 수 있다. 따라서 보간 과정에서 Chono 보다 움직임 추정 오차를 효과적으로 감소시킨다. 본 논문에서는 제안된 필터들을 각각 필터 모드(filter mode)  $f$ 라 명시하고 표 3은 제안된 9가지 2차원 필터들과 이에 대응되는 필터 모드들을 나

표 3. 제안된 9가지 2차원 필터들과 필터 모드  
Table 3. Proposed nine 2D filters and these filter modes.

Filter mode ( $f$ )	Combination of the filter number (horizontal, vertical)
1	(1,1)
2	(2,2)
3	(3,3)
4	(2,3)
5	(3,2)
6	(2,1)
7	(3,1)
8	(1,2)
9	(1,3)

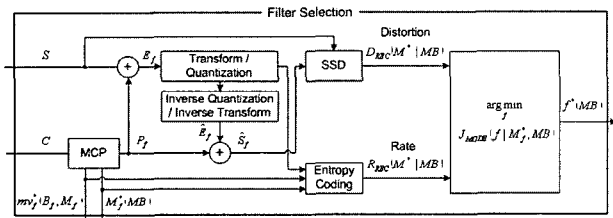


그림 7. 필터 선택을 위한 제안된 비용 함수의 계산. SSD는 sum of squared difference의 약자이다  
Fig. 7. Computation of the proposed cost function for filter selection. SSD is the sum of squared difference.

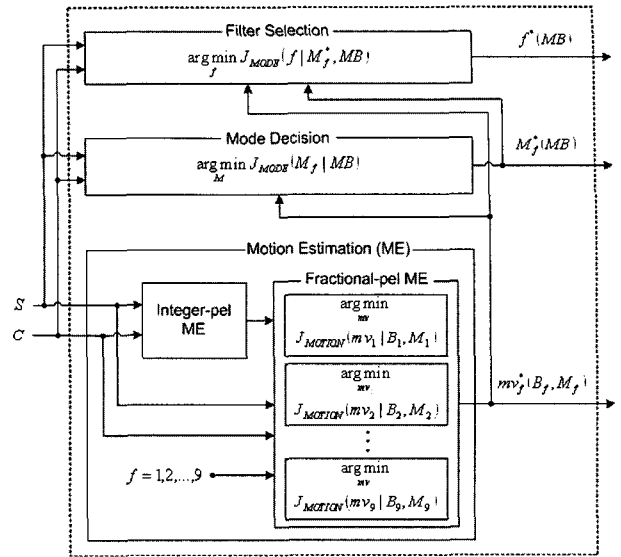


그림 8. 제안된 움직임 보상. 제안된 9가지 필터 모드  $f$ 는 1/2화소 위치들을 보간하기 위한 각각의 소수화소 움직임 추정에 이용된다. 1/4화소 위치들은 쌍일차에 의해 보간된다. 최적 필터는 모드 결정 이후에 선택됨을 주목해야 한다  
Fig. 8. Proposed MCP. The proposed nine filter mode  $f$  is used in each fractional-pel motion estimation for interpolating at half-pel positions. Quarter-pel positions are interpolated by bilinear filter. Note that the optimal filter is selected after the mode decision.

타낸다.

필터 선택을 위한 Chono의 비용 함수는 움직임 벡터가 부호화되기 위한 비트율과 예측 오차를 고려하였다. 이는 매크로 블록이 부호화되기 비트율과 왜곡 측면에서 최적이지 않다. 따라서 본 논문에서는 매크로 블록이 부호화되기 위한 실제 비트율과 복원 오차를 고려한 최적의 비용 함수를 제안한다. 그림 7은 제안된 비용 함수의 계산 과정을 묘사한 것이다. 제안된 비용 함수는 식 (6)의  $J_{MODE}(f | M_f^*, MB)$ 와 같이 정의할 수 있다. 식 (6)은 식 (2)를 9번 수행한 것과 같다.

$$f^*(MB) = \arg \min_f J_{MODE}^A(f | M_f^*, MB) = \arg \min_f \left\{ D_{DFID}(f | M_f^*, MB) + \lambda_{MODE} \cdot R_{REC}(f | M_f^*, MB) \right\} \text{ for } f=1,2,\dots,9 \quad (6)$$

여기서  $f^*(MB)$ ,  $D_{REC}(f | M_f^*, MB)$ ,  $R_{REC}(f | M_f^*, MB)$ 는 선택된 필터 모드, 현재 블록과 필터 모드에 대한 최적의 예측 모드의 복원 블록간의 복원 오차, 필터 모드에 대한 최적의 예측 모드의 비트율을 각각 나타낸다.

$R_{REC}(f|M_f^*, MB)$ 은 예측 오차, 움직임 벡터, 예측 모드, 참조 영상 프레임 등의 비트율을 포함한다.  $D_{REC}(f|M_f^*, MB)$ 는 필터 모드에 대한 예측 오차의 양자화 오차를 포함한다.

그림 8은 제안된 방법의 움직임 보상 예측을 나타낸 것이다. 필터 선택은 9가지 필터들에 대한 최적의 예측 모드를 결정한 이후 수행된다는 것을 주목해야 한다. 이는 매크로 블록이 부호화되기 위한 비트율과 왜곡 측면에서 최적의 필터를 선택하기 위한 것이며 Chono 방법과의 중요한 차이점이다. 제안된 움직임 보상 예측은 다음과 같은 단계로 수행된다.

1) 움직임 추정

Chono의 단계 1과 동일하고 9가지 필터 모드들에 대해서 적용한다. 식 (7)의 비용 함수  $J_{MOTION}(mv_f|B_f, M_f)$ 을 이용하여 필터 모드에 대한 최적의 움직임 벡터를 계산한다.

$$mv_f^*(B_f, M_f) = \arg \min_{mv_f} J_{MOTION}(mv_f|B_f, M_f) = \arg \min_{mv_f} \left\{ D_{DFD}(mv_f|B_f, M_f) + \lambda_{MOTION} \cdot R_{MOTION}(mv_f|B_f, M_f) \right\} \quad (3)$$

for  $f=1, 2, \dots, 9$

여기서  $mv_f^*(B_f, M_f)$ 는 필터 모드  $f$ 에 대한 최적의 움직임 벡터를 나타낸다.

2) 모드 결정

식 (8)의 비용 함수  $J_{MODE}(M_f|MB)$ 을 이용하여 필터 모드에 대한 최적의 예측 모드 결정한다. 식 (8)은 식 (2)을 9번 수행한 것이다. 이는 제안된 9가지 필터 모드들이 적용되기 때문이다.

$$M_f^*(MB) = \arg \min_{M_f} J_{MODE}(M_f|MB) = \arg \min_{M_f} \left\{ D_{REC}(M_f|MB) + \lambda_{MODE} \cdot R_{REC}(M_f|MB) \right\} \quad (8)$$

for  $f=1, 2, \dots, 9$

여기서  $M_f^*(MB)$ 는 필터 모드에 대한 최적의 예측 모드를 나타낸다.

3) 필터 선택

식 (6)의 비용 함수  $J_{MODE}(f|M_f^*, MB)$ 을 이용하여

매크로 블록 단위로 최적의 필터 모드  $f^*$ 을 선택한다. 따라서 선택된 필터 모드에 대한 최적 예측 모드의 움직임 벡터  $mv_{f^*}^*(B_{f^*}, M_{f^*}^*)$ 가 결정되어 진다.

4) 예측

현재 매크로 블록은  $mv_{f^*}^*(B_{f^*}, M_{f^*}^*)$ 을 이용하여 예측된다.

움직임 보상 예측 이후 선택된 필터 모드와 엔트로피 부호화된  $mv_{f^*}^*(B_{f^*}, M_{f^*}^*)$ 는 영상 복원을 위하여 전송한다.

IV. 실험 결과 및 고찰

본 논문에서는 제안 방법의 우수성을 확인하기 위하여 H.264/AVC 참조 소프트웨어와 KTA(Key Technical Area) 소프트웨어가 합병된 jm11.0kta1.4<sup>[16]</sup>을 이용하여 기존 방법들과 제안 방법(PM) 간의 부호화 효율을 비교, 평가하였다. 실험은 300프레임으로 구성된

표 4. 실험 조건  
Table 4. Experimental conditions.

Frame rate	30Hz
QP (quantization parameter)	20, 24, 28, 32
Hadamard transform	used
Search range	±16
Total number of references	1, 5
Sequence type	I, P, P, ...
Motion vector resolution	1/4-pel
RD-optimized mode decision	used
Entropy coding method	CAVLC

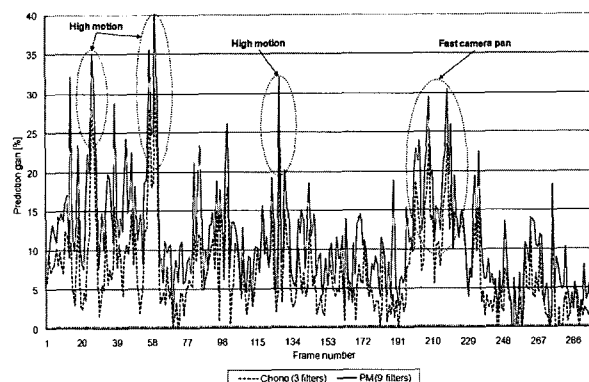


그림 9. Foreman 시퀀스에 대하여 Chono의 3가지 필터들과 제안된 9가지 필터들의 예측 이득 비교 (참조 영상 프레임: 1개)

Fig. 9. Comparisons of the prediction gain between Chono's three filters and the proposed nine filters for test sequence Foreman (1 reference frame).

QCIF 시퀀스들 (Foreman, Container, News, Silent)과 CIF 시퀀스들 (Mobile & Calendar, Tempete, Paris)에 대하여 수행되었다. 실험 조건은 표 4와 같다.

제안된 필터들에 대한 성능을 확인하기 위하여 그림 9는 Foreman 시퀀스에서 Chono와 제안된 필터들에 대한 예측 이득(prediction gain, PG)을 프레임 단위로 비교한 것이다. 예측 이득은 식 (9)과 같이 정의한다.

$$PG(\%) = \left[ 1 - \frac{\text{Prediction error (PM or Chono)}}{\text{Prediction error (H.264/AVC)}} \right] \times 100 \quad (9)$$

그림 9에서 표시한 높은 움직임(high motion), 빠른 카메라 팬(fast camera pan) 특성을 가지는 프레임 영

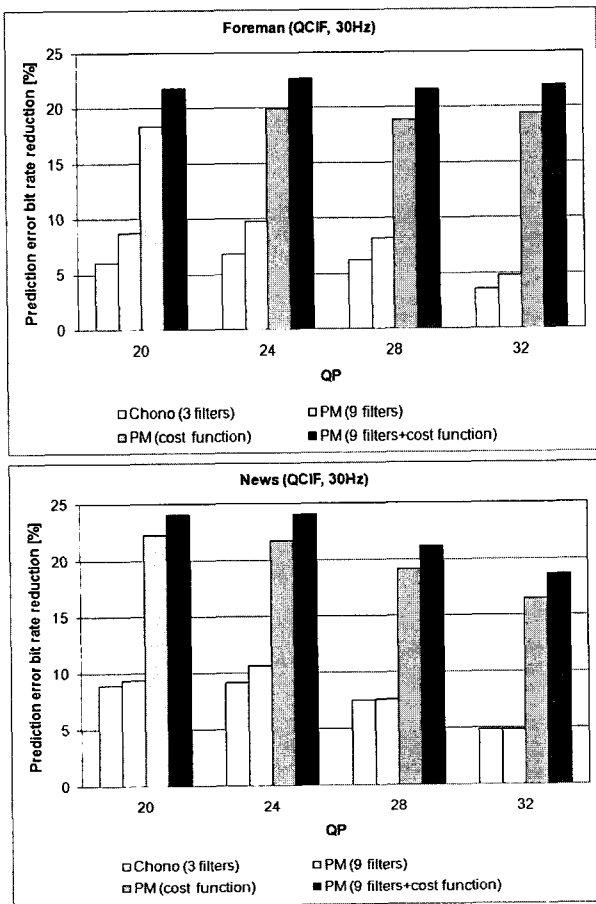


그림 10. Foreman, News 시퀀스에 대하여 Chono와 제안된 방법의 예측 오차에 대한 비트율 절감 비교 (참조 영상 프레임: 1개). 각각의 QP에서 PSNR은 고정하였다

Fig. 10. Comparisons the prediction error bit rate reduction between the proposed method and Chono's method for test sequence Foreman and News (1 reference frame). The PSNR is fixed in each QP.

역에서는 움직임 추정 오차가 크다. 이 영역에서 H.264/AVC 보다 Chono와 본 논문에서 제안된 필터들을 적용한 경우에 예측 이득이 향상된다. 이는 Chono와 제안된 필터들이 한정된 움직임 벡터 해상도 보다 움직임을 세밀하게 보상하기 때문이다. 그리고 제안된 필터들은 Chono 보다 다양한 방향의 움직임이 세밀하게 보상할 수 있기 때문에 모든 프레임에서 예측 이득이 향상된다.

그림 10은 Foreman과 News 시퀀스에서 Chono와 제안된 방법의 예측 오차에 대한 비트율 절감을 비교한 것이다. 그림에서 명시된 PM(9 filters), PM(cost function), PM(9filters+cost function)은 Chono의 필터들 대신 제안된 필터들, Chono의 방법에서 필터 선택을 위해 제안된 cost function, 제안된 필터들과 cost function을 각각 적용한 결과이다.

제안된 필터들은 H.264/AVC 보다 예측 오차에 대한 비트율이 약 평균 8% 절감된다. 특히 움직임이 많은 Foreman 시퀀스일 경우 움직임 적은 News 시퀀스 보다 개선됨을 알 수 있다. 또한 양자화 파라메타가 작을 수록 예측 효율이 증가한다.

제안된 비용 함수는 H.264/AVC 보다 예측 오차에 대한 비트율이 약 평균 19% 절감된다. 그리고 움직임이 많은 Foreman 시퀀스 뿐만 아니라 움직임 적은 News 시퀀스에서도 예측 효율이 현저하게 개선된다. 이는 제안된 비용 함수가 예측 오차에 대한 양자화 오차를 고려하기 때문에 움직임과 관계없이 개선된다.

제안된 필터와 비용 함수를 결합할 경우, H.264/AVC 보다 예측 오차에 대한 비트율이 약 평균 22% 절감된다.

그림 11은 제안 방법과 기존 방법들에 대한 윌-왜곡 곡선을 보여준다. 제안 방법이 기존 방법들 보다 항상 우수한 부호화 효율을 지니고 있음을 알 수 있다.

표 5, 6은 다양한 표준 동영상 시퀀스들에 대한 제안 방법과 기존 방법들과 전체 비트율 절감에 대한 비교를 보여준다. 표 5, 6의 값은 1개와 5개의 참조 영상 프레임 적용한 결과이다. 제안 방법은 기존 방법들과 비트율 절감을 비교해 볼 때 항상 우수하며 H.264/AVC 보다 평균 6.25%(참조 영상 프레임: 1개), 평균 3.46%(참조 영상 프레임: 5개)가 개선됨을 알 수 있다.

5개의 참조 영상 프레임을 적용할 경우, 기존 방법들은 H.264/AVC 보다 비트율이 절감되지 않는 부분이 있는 것을 알 수 있다. 이는 H.264/AVC에서 엘리어링과 움직임 추정 오차를 보상하기 위해 다중 참조 프레임 기법을 적용하기 때문이다. 따라서 5개의 참조 영상 프레임을 적용할 경우, 기존 방법들은 1개의 참조 영상



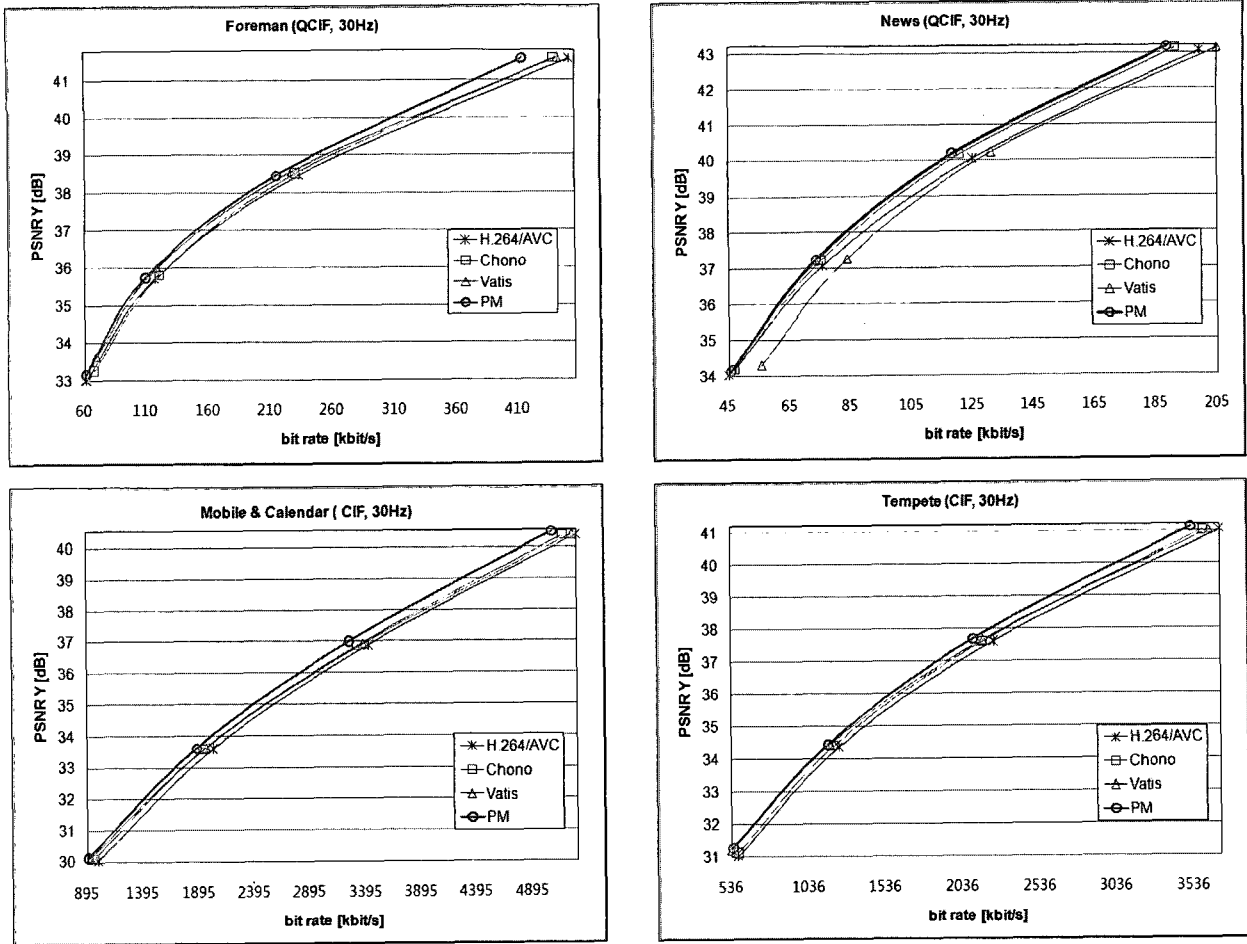


그림 11. 제안 방법과 기존 방법들에 대한 율-왜곡 곡선 (참조 영상 프레임: 1개)

Fig. 11. Rate-distortion curves for the proposed method and the conventional methods. (1 reference frame)

표 5. 전체 비트 율 절감 비교(참조 영상 프레임: 1개)  
Table 5. Comparisons of total bit rate reduction. (1 reference frame)

Sequence	Resolution	Chono <sup>[9]</sup>	Vatis <sup>[13]</sup>	PM
Foreman	QCIF	0.44%	4.35%	7.74%
Container	QCIF	2.40%	3.03%	10.93%
News	QCIF	3.60%	-5.97%	5.56%
Silent	QCIF	-2.18%	-7.80%	1.24%
Mobile&Calender	CIF	3.72%	3.77%	7.59%
Tempete	CIF	3.10%	5.56%	7.09%
Paris	CIF	0.78%	1.22%	3.59%

프레임을 적용한 결과 보다 부호화 효율이 감소된다. 그러나 제안 방법에서는 엘리머싱과 움직임 추정 오차 뿐만 아니라 예측 오차에 대한 양자화 오차를 고려하기 때문에 다중 참조 프레임을 적용할 경우에서도 부호화 효율이 개선된다.

제안된 방법은 Chono 보다 부호화 효율을 개선하지

표 6. 전체 비트 율 절감 비교(참조 영상 프레임: 5개)  
Table 6. Comparisons of total bit rate reduction. (5 reference frames)

Sequence	Resolution	Chono <sup>[9]</sup>	Vatis <sup>[13]</sup>	PM
Foreman	QCIF	-4.77%	1.33%	6.47%
Container	QCIF	-1.29%	-7.95%	5.38%
News	QCIF	-0.77%	-6.66%	3.62%
Silent	QCIF	-4.06%	-8.40%	0.48%
Mobile&Calender	CIF	-0.62%	-12.73%	3.71%
Tempete	CIF	-1.84%	4.46%	3.66%
Paris	CIF	-2.04%	-5.02%	0.89%

만 부호화기의 계산량이 증가된다. 소수 화소 움직임 추정과 모드 결정 과정이 증가한다. 특히 후자의 경우, 전자 보다 계산량의 부담을 더욱 초래한다. 그러나 모드 결정의 계산량을 감소하기 위해 고속 모드 결정 방법들이 연구되고 있다<sup>[14~15]</sup>. 이를 적용한다면 계산량 부담을 감소시킬 수 있다.

## V. 결 론

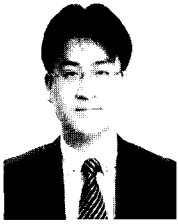
본 논문에서는 H.264/AVC에서 부호화 효율을 개선하기 위한 매크로 블록 기반의 적응 보간 필터 방법을 제안하였다. 제안된 방법에서는 한정된 움직임 벡터 해상도 보다 주위 8 방향의 움직임을 세밀하게 보상할 수 있는 9가지 분리 가능한 2차원 필터들을 생성하였다. 그리고 울-왜곡 측면에서 최적인 필터를 선택하기 위한 새로운 비용 함수를 정의하였다.

실험 결과, 제안된 필터들은 움직임이 많은 시퀀스에서 예측 효율을 더욱 개선하였다. 그리고 제안된 비용 함수는 움직임이 많은 시퀀스 뿐만 아니라 움직임이 적은 시퀀스에서도 예측 효율을 개선하였다. 전체 비트율에 대한 절감을 비교해 볼 때, 제안 방법은 다양한 표준 QCIF/CIF 동영상 테스트 시퀀스들에 대하여 항상 우수하다. 그리고 H.264/AVC 보다 평균 6.25%(참조 영상 프레임: 1개), 평균 3.46%(참조 영상 프레임: 5개)가 개선된다. 또한 다중 참조 프레임을 적용할 경우에서도 기존 방법들 보다 부호화 효율이 개선되는 장점을 지니고 있다.

## 참 고 문 헌

- [1] Int. Telecommun. Union-Telecommun. (ITU-T) and Int. Standards Org./Int. Electrotech. Comm. (ISO/IEC) JTC 1, Rec. H.264 and ISO/IEC 14496-10 (MPEG-4) AVC, Advanced Video Coding for Generic Audiovisual Services 2003.
- [2] M. Wien, "Variable Block-Size transforms for H.264/AVC", IEEE Trans. Circuits Syst. Video Technol, vol. 13, no. 7, pp. 604-613, Jul. 2003.
- [3] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol, vol. 13, pp. 560-576, July 2003.
- [4] T. Wiegand, X. Zhang, and B. Girod, "Long-Term Memory Motion-Compensated Prediction," IEEE Trans. Circuits Syst. Video Technol, vol. 9, no. 1, pp.70-84, Feb. 1999.
- [5] O. Werner, "Drift analysis and drift reduction for multiresolution hybrid video coding," Signal Process.: Image Commun., vol. 8, no. 5, Jul. 1996.
- [6] K. Chono and Y. Miyamoto, "MB basis Adaptive Motion Interpolation," JVT-D040, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG Meeting, 2002.
- [7] K. Chono and Y. Miyamoto, "Modified Adaptive Interpolation Filter," JVT-D078, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG Meeting, 2002.
- [8] T. Wedi and H. G. Musmann, "Motion and Aliasing-Compensated Prediction for Hybrid Video Coding", IEEE Trans. Circuits Syst. Video Technol, vol. 13, no. 7, July 2003.
- [9] T. Wedi, "Adaptive interpolation filters and high-resolution displacements for video coding," IEEE Trans. Circuits Syst. Video Technol, vol. 15, no. 4, pp.484-492, April 2006
- [10] Y. Vatis, B. Edler, I. Wassermann, D.T.Nguyen, J. Ostermann, "Coding of Coefficients of two-dimensional Adaptive Wiener Interpolation Filter", Visual Communications and Image Processing, Beijing, China, July 2005.
- [11] Y. Vatis, B. Edler, D. T. Nguyen, J. Ostermann, "Motion and Aliasing-Compensated Prediction Using a Two-dimensional Non-separable Adaptive Wiener Interpolation Filter", in Proc. IEEE Int. Conf. Image Processing, Genova, Italy, September 2005.
- [12] G. J. Sullivan and T. Wiegand, Rate-Distortion Optimization for Video Compression IEEE Signal Processing Magazine, Vol. 15, Num. 6, pp. 74-90, Nov. 1998
- [13] P. List, A. Joch, J. Lainema, G. Bjontegaard, and M. Karczewicz, "Adaptive deblocking filter," IEEE Trans. Circuits Syst. Video Technol, vol. 13, pp. 614-619, July 2003.
- [14] Libo Yang, Keman Yu, Jiang Li, and Shipeng Li, "An Effective Variable Block-Size Early Termination Algorithm for H.264 Video Coding", IEEE Trans. Circuits Syst. Video Technol, vol. 15, no. 6, pp.784-788, Jun. 2005
- [15] Y. K. Tu, J. F. Yang, M. T. Sun, "Fast Inter Mode Selection", IEEE Trans. Circuits Syst. Video Technol, vol. 16, no. 5, pp.600-611, Jun. 2006
- [16] KTA reference model 1.2, available at <http://www.tnt.uni-hannover.de/~vatis/kta/jm11.0kta1.2.zip>.

저 자 소 개



윤 근 수(정회원)  
 2000년 경남대학교 전자공학과  
 학사 졸업.  
 2001년 경남대학교 정보통신  
 공학과 석사 졸업.  
 2004년 부산대학교 전자공학과  
 박사 수료.

2004년~현재 부산대학교 전자공학과 박사 과정.  
 <주관심분야 : Image processing, Video coding,  
 Multimedia system, SoC>



김 재 호(정회원)  
 1980년 부산대학교 전기기계  
 공학과 학사 졸업.  
 1982년 한국과학기술원 산업전자  
 공학과 석사 졸업.  
 1990년 한국과학기술원 전기및  
 전자공학과 박사 졸업.

1990년~1991년 삼성전자 정보통신연구소  
 책임연구원.

1992년~현재 부산대학교 전자공학과 교수.  
 <주관심분야 : Video coding, High resolution  
 image processing, Animation automation>