일반논문-10-15-3-07

H.264/AVC에서의 최고 확률 모드를 이용한 고속 화면 내 예측 모드 결정

김 대 연^{a)}, 김 정 필^{a)}, 이 영 럽^{a)†}

Fast Intra Prediction Mode Decision using Most Probable Mode for H.264/AVC

Dae-Yeon Kim^{a)}, Jeong-Pil Kim^{a)}, and Yung-Lyul Lee^{a)†}

요 약

가장 최근에 표준화가 완료된 동영상 압축 코덱인 H.264/AVC는 율-왜곡 최적화를 사용하여 압축률이 상당히 향상되었다. 율-왜곡 최적화는 다수의 후보 모드들 중에서 최적의 모드를 결정하기 위한 수단이며, 모든 모드에 대하여 라그랑지안 비용을 계산하여 그 비용이 가장 적은 모드를 최적의 모드로 결정한다. 하지만 이 방법의 사용으로 인하여 H.264/AVC 부호기의 계산 복잡도가 상당히 증가하였다. 본 논문에서는 H.264/AVC의 화면 내 예측 부호화의 율-왜곡 최적화의 사용으로 인한 계산 복잡도를 감소시키는 고속 화면 내 예측 모드 결정 방법을 제안한다. 제안된 방법은 Intra4×4와 Chroma Intra의 부호화 속도를 향상시키기 위하여 최고 확률 모드(MPM)가 가장 최적의 예측 모드인 경우를 미리 판단하여, 이 경우에는 율-왜곡 최적화 과정을 생략하고 최고 확률 모드를 사용하여 부호화한다. 또한 이 방법을 색차 화면 내 예측 모드에 유사하게 적용하였다. 실험 결과, 제안된 방법은 PSNR (Peak Signal-to-Noise Ratio) 손실이 거의 없이 평균 약 63%의 부호화 시간을 절약하였다.

Abstract

The most recent standard video codec, H.264/AVC achieves significant coding efficiency by using a rate-distortion optimization (RDO). The RDO is a measurement for selecting the best mode which minimizes the Lagrangian cost among several modes. As a result, the computational complexity is increased drastically in encoder. In this paper, a method for fast intra prediction mode decision is proposed to reduce the RDO complexity. To speed up Intra4×4 and Chroma Intra encoding, the proposed method decides the case that *MPM* (Most Probable Mode) is the best prediction mode. In this case, the RDO process is skipped, and only *MPM* is used for encoding the block in Intra4×4. And the proposed method is also applied to the chroma Intra prediction mode in a similar way to the Intra4×4. The experimental results show that the proposed method achieves an average encoding time saving of about 63% with negligible loss of PSNR (Peak Signal-to-Noise Ratio).

Keywords : H.264, MPEG-4 Part 10, Intra prediction, Fast mode decision

a) 세종대학교 컴퓨터공학과 DMS 연구실

DMS Lab. Dept. of Computer Engineering, Sejong University

† 교신저자 : 이영렬(ylllee@sejong.ac.kr)

※ 이 논문은 2009년도 세종대학교 교내연구비 지원에 의한 논문임.

· 접수일(2010년1월29일), 수정일(2010년4월20일), 게재확정일(2010년4월23일)

1. 서 론

H.264/AVC는 Moving Picture Experts Group (MPEG) 과 Video Coding Experts Group (VCEG)으로 구성된 Joint

Video Team (JVT)에서 MPEG-4 Part 10과 ITU-T Recommendation H.264로 공동 발표 되었다^{[1][3]}. H.264/AVC는 기존의 압축 표준들과 비교하여 방향성 기반의 화면 내 예측, 1/4 단위의 가변 블록 크기의 움직임 예측 및 보상, CAVLC (Context-based Adaptive Variable Length Coding)와 CABAC (Context-based Adaptive Binary Arithmetic Coding)^[4]을 이용한 엔트로피 부호화, 그리고 De-blocking filter^[5] 등의 기술을 이용하여 MPEG-2 보다 약 2배, MPEG-4 ASP (Advanced Simple Profile) 보다 약 1.5배의 압축률 향상을 보였다^[6].

위의 기술들 중 방향성 기반의 화면 내 예측은 Intra4×4, Intra8×8 (FRExt 프로파일에서만 사용됨), Intra16×16과 같이 3가지 블록 모드를 이용하여 예측이 수행된다. 또한 각 블록 모드는 블록 경계의 복호화된 인접 픽셀을 이용하여 방향성 예측을 수행하는 다수의 화면 내 예측 모드들을 포함한다. H.264/AVC에서는 다양한 블록 모드와 다수의 예측 모드 중 최적의 모드를 선택하기 위하여 율-왜곡 최적화 (Rate-Distortion Optimization)를 이용한다^[7]. 이 방법은 모든 후보 모드에 대하여 실제로 부호화와 복호화를 수행하여 율-왜곡 비용을 계산하여 그 값이 가장 적은 모드를 최적의 모드로 선택하는 방법이다. 따라서 높은 압축 효율의 향상을 가져왔지만 부호화 시간을 상당히 증가시키는 문제를 초래하였다. 이는 실시간 부호화를 요구하는 응용분야에서 큰 단점으로 작용한다. 이러한 단점을 해결하기 위하여 화면 내 예측 모드를 고속으로 선택하기 위한 방법들이 연구되어 왔다. 에지 검출 방법을 이용한 방법, 통계 기반의 방법 등의 다양한 알고리즘을 이용한 고속 화면 내 예측 모드 방법들이 연구되었으나, 몇몇 알고리즘들은 복잡도가 높아서 부호화 속도의 향상이 높지 않고 어떠한 알고리즘들은 부호화 속도를 상당히 높였으나 PSNR을 감소시키는 결과를 가져왔다^{[8]-[10]}.

본 논문에서는 H.264/AVC의 압축률 손실을 최소화하면서 부호화 속도를 증가시키는 고속 화면 내 예측 모드 결정 방법을 제안한다. 제안된 방법은 Intra4×4 모드에서 MPM (Most Probable Mode)이 가장 최적의 예측 모드인 경우를 사전에 판단하여 이 경우에는 율-왜곡 최적화 과정을 생략하고 MPM 모드를 이용하여 부호화한다. 또한 색차 화면

내 (Chroma Intra) 예측 모드에서는 DC 예측 모드가 최적인 경우를 사전에 판단하여 이 경우에는 DC 예측 모드로 부호화하는 방법을 제안한다.

본 논문은 다음과 같은 순서로 구성된다. 2장에서 H.264/AVC 화면 내 예측 코딩에 대하여 설명하고, 3장에서는 제안된 고속 화면 내 예측 모드 결정 방법에 대하여 설명한다. 4장에서는 제안된 방법을 H.264/AVC 참조 소프트웨어와 기존의 고속 화면 내 예측 모드 결정 방법과 성능을 비교하며, 마지막으로 5장에서 결론을 맺는다.

II. H.264/AVC의 화면 내 부호화

1. 화면 내 예측

화면 내 부호화는 공간상으로 인접한 화소 값들이 높은 유사성을 가지고 있는 특징을 이용한 압축방법이다. H.264/AVC에서의 화면 내 예측 부호화는 블록 크기에 따라 휘도 신호의 경우 3가지 블록 모드 (Intra16×16, Intra8×8, Intra4×4), 색차 신호의 경우 1가지 블록 모드 (Chroma Intra)의 부호화 방법이 존재한다. Intra16×16 모드와 Chroma Intra 모드의 경우는 각각 4가지 화면 내 예측 모드가 존재하며, Intra8×8 모드와 Intra4×4 모드의 경우는 각각 9가지 화면 내 예측 모드가 존재한다.

그림 1. Intra4×4 예측 (a) Intra4×4에 사용되는 화소 (b) Intra4×4 예측 방향
Fig. 1. Intra4×4 prediction (a) Pixels used in Intra4×4 (b) Prediction directions of Intra4×4

그림 1의 (a)는 H.264/AVC의 Intra4×4 모드에 사용되는 화소들의 예를 보여주는 것이다. 소문자 a 부터 p 까지의 16개의 화소들은 현재 부호화 될 4×4 블록의 화소들이고,

대문자 A 부터 M 까지의 13개 화소들은 이전에 부호화 된 블록의 화소들이다. Intra4×4 예측 부호화는 이미 부호화된 A 부터 M 의 화소 값을 이용하여 그림 1의 (b)와 같이 8가지의 방향성을 이용한 예측과 평균을 예측 값으로 사용하는 DC 예측을 사용하여 현재 부호화 될 화소들과의 차분 과정을 통해 차분 계수를 생성한다. 예를 들어, Intra4×4 예측 방법 중, 모드 3인 Diagonal down left 예측이 사용될 경우, 첫 번째 행의 a, b, c, d 화소들은 식 (1)과 같이 차분 계수 a', b', c', d' 가 생성된다. Intra8×8 모드와 Intra16×16 예측도 이와 유사한 방법으로 예측을 수행하여 차분 계수를 생성한다.

$$\begin{aligned} a' &= a - [(A + 2B + C) / 4] \\ b' &= b - [(B + 2C + D) / 4] \\ c' &= c - [(C + 2D + E) / 4] \\ d' &= d - [(D + 2E + F) / 4] \end{aligned} \quad (1)$$

2. H.264/AVC의 변환과 양자화

화면 내 예측이 수행된 후, 부호기는 차분 계수에 대하여 정수 이산 여현 변환 (DCT)과 양자화 (Quantization)를 수행한다^[11]. 4×4 차분 계수 블록에 대한 정방향 정수 이산 여현 변환은 다음 식 (2)와 같이 수행된다.

$$\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \mathbf{R} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \quad (2)$$

여기서 행렬 $\mathbf{R}$ 은 4×4 차분 계수 블록이고 행렬 $\mathbf{C}$ 는 4×4 변환 계수 블록이다.

H.264/AVC는 변환 계수 블록에 대하여 스칼라 양자화를 수행하여, 행렬 $\mathbf{C}$ 는 다음 식 (3)과 같이 양자화 된다^[11].

$$|Z(i, j)| = \text{sign}\{C(i, j)\} \times \left[\left(|C(i, j)| \cdot A(Q_M, i, j) + f 2^{L+Q_E} \right) \gg (L + Q_E) \right] \quad (3)$$

여기서 Z 는 행렬 $\mathbf{C}$ 의 양자화된 계수 블록이고, $\text{sign}(a)$ 는 a 값의 부호를 반환하는 함수이다. A 는 양자화 테이블이고, f 는 양자화 간격을 조절하는 값이다. $Q_M = Q \bmod 6$, $Q_E = Q/6$, $L=15$ 로 정의되며, 이 때 Q 는 양자화 파라미터를 나타낸다.

식 (3)에 의하여, 양자화 에러가 발생하기 때문에 원본 화소와 복원된 화소가 서로 동일하지 않다.

3. 율-왜곡 과정과 엔트로피 코딩

율-왜곡 최적화 과정은 모든 후보 모드에 대하여 실제로 부호화와 복호화를 수행하여 라그랑지안 함수를 이용하여 율-왜곡 비용을 계산하고 그 값이 가장 적은 모드를 최적의 모드로 선택한다. Intra4×4에서의 율-왜곡 비용은 다음 식 (4)를 통하여 계산된다.

$$\begin{aligned} J(s, c, \text{IMODE} | QP, \lambda_{\text{MODE}}) \\ = D(s, c, \text{IMODE} | QP) + \lambda_{\text{MODE}} \cdot R(s, c, \text{IMODE} | QP) \quad (4) \\ \text{IMODE} \in \{\text{all nine Intra4x4 modes}\} \end{aligned}$$

여기서 QP 는 양자화 파라미터이고, λ_{MODE} 는 라그랑지안 상수이고, IMODE 는 Intra4×4의 9가지 예측 모드 중 하나의 모드를 나타낸다. D 는 원본 화소 블록 s 와 복원된 화소 블록 c 와의 왜곡량을 나타내고, 이것은 SSD (Sum of Squared Difference)를 이용하여 계산한다. R 은 IMODE 를 이용하여 부호화하였을 때 발생하는 비트량을 나타낸다.

Intra4×4 모드와 Intra8×8 모드는 9가지의 예측 모드를 사용하기 때문에 예측 모드를 표현하기 위하여 많은 비트가 필요하다. 따라서 H.264/AVC에서는 예측 모드들 중 최적의 모드로 선택될 확률이 가장 높은 모드 (MPM)를 정의하였다. MPM 은 좌측 블록과 상단 블록의 예측 모드 중 모드 번호가 작은 모드로 결정된다.

화면 내 예측 모드를 엔트로피 코딩하는 과정에 있어서 율-왜곡 최적화에서 결정된 최적의 예측 모드와 MPM 모드

가 동일할 경우, 1 비트의 신호만을 복호기에 전송하여 압축 성능을 향상시켰다. 만약 동일하지 않을 경우, 1 비트 신호 외에 나머지 8가지 모드 중 어떤 모드인지를 나타내는 3 비트 신호를 추가적으로 복호기에 전송한다.

III. 제안하는 고속 화면 내 예측 모드 선택 방법

1. 제안하는 방법의 동기

율-왜곡 최적화에서 율-왜곡 비용이 서로 동일한 모드들이 존재할 경우에는 그 모드들 중 임의의 한 모드만을 율-왜곡 최적화에서 사용하여도 압축률의 감소가 없다. 따라서 율-왜곡 비용이 동일할 것으로 예상되는 모드들을 사전에 검출한다면 율-왜곡 최적화에서 사용되는 후보 모드의 개수를 줄일 수 있다. 하지만 율-왜곡 비용이 서로 정확히 동일한 경우는 빈번히 발생하지 않기 때문에 전체적인 부호화 속도의 향상이 거의 없다. 예측 모드들의 율-왜곡 비용이 서로 정확히 동일하지 않고 유사할 경우는 좀 더 빈번히 발생하기 때문에 압축 성능의 손실이 조금 발생하지만 전체적인 부호화 속도를 향상시킬 수 있다. 본 논문에서는 예측 모드들의 율-왜곡 비용이 서로 유사할 것으로 예상되는 모드들을 사전에 검출하여 율-왜곡 최적화에서 사용되는 후보 모드의 개수를 줄임으로 인해 부호화 속도를 향상시키는 방법을 제안한다.

어떠한 두 모드의 율-왜곡 비용이 동일하기 위해서는 식 (4)의 왜곡량 D 와 비트량 R 이 동일해야한다. 여기서 왜곡량이 동일하기 위해서는 두 모드의 복원된 블록이 서로 동일해야하며, 비트량이 동일하기 위해서는 양자화된 변환 계수 블록과 예측 모드를 표현하기 위한 비트량이 서로 같아야한다.

식 (5)는 Vertical 예측과 Horizontal 예측을 이용하여 임의의 블록을 부호화하였을 때의 각 예측의 복원된 블록을 표현한 식이다.

$$\left. \begin{aligned} R_{VER} &= IQ \cdot IDCT \{ DCT \cdot Q(S - P_{VER}) \} + P_{VER} \\ R_{HOR} &= IQ \cdot IDCT \{ DCT \cdot Q(S - P_{HOR}) \} + P_{HOR} \end{aligned} \right\} \quad (5)$$

여기서 R_{VER} 은 Vertical 예측을 사용하였을 때의 복원된 블록을 나타내고 R_{HOR} 은 Horizontal 예측을 사용하였을 때의 복원된 블록을 나타낸다. S 는 원본 화소 블록을 나타내고, P_{VER} 과 P_{HOR} 는 각각 Vertical, Horizontal 예측에 의한 예측 화소 블록을 나타낸다. $DCT \cdot Q(x)$ 는 식 (2)와 (3)을 이용하여 블록 x 의 양자화된 변환 계수 블록을 유도하는 함수이다. 또한 $IQ \cdot IDCT(x)$ 는 블록 x 에 대하여 역양자화 및 역변환된 계수 블록을 유도하는 함수이다.

$$\left. \begin{aligned} T_{VER} &= DCT \cdot Q(S - P_{VER}) \\ T_{HOR} &= DCT \cdot Q(S - P_{HOR}) \end{aligned} \right\} \quad (6)$$

$$\left. \begin{aligned} R_{VER} &= R_{HOR} \\ T_{VER} &= T_{HOR} \end{aligned} \right\} \text{if } P_{VER} = P_{HOR} \quad (7)$$

식 (6)은 두 예측 모드의 차분 신호에 대한 양자화된 변환 계수 블록 T_{VER} , T_{HOR} 을 나타낸 식이다. 식 (5)와 식 (6)을 통하여 만약 식 (7)과 같이 P_{VER} 과 P_{HOR} 이 동일한 값을 갖는 블록일 때, 복원된 블록 R_{VER} 과 R_{HOR} 이 서로 동일하고 양자화된 변환 계수 블록 T_{VER} , T_{HOR} 또한 동일하다는 것을 알 수 있다. 따라서 이 경우에는 두 모드의 예측 모드를 표현하기 위한 비트량만 서로 동일하다면 율-왜곡 비용이 동일하기 때문에 율-왜곡 최적화 과정에서 두 모드 중, 어느 하나의 모드만 사용하여도 압축 성능의 감소 없이 부호화 속도를 향상시킬 수 있다. 하지만 예측 화소 블록이 서로 정확히 동일한 경우는 확률적으로 많이 발생하지 않기 때문에 전체적으로 높은 부호화 속도 향상을 기대하기 어렵다.

예측 모드들의 율-왜곡 비용이 서로 유사할 경우는 좀 더 빈번히 발생하기 때문에 이 경우에 후보 모드의 수를 줄이면 압축 성능의 손실이 조금 발생할 수 있지만 전체적인

부호화 속도를 향상시킬 수 있다. 어떠한 예측 모드들의 울-왜곡 비용이 서로 유사한 값이 발생하는 경우를 판단하기 위한 식을 다음 식 (8)로부터 유도하였다. 식 (8)은 두 예측 모드의 차분 신호에 대한 양자화된 변환 계수 블록들이 서로 동일한 경우를 나타낸 식이다. 이 식이 성립하는 경우에는 두 모드의 예측 모드를 표현하기 위한 비트량만 서로 동일하다면 울-왜곡 비용 중, 울(Rate)에 대한 비용이 서로 동일하다.

$$DCT \cdot Q(S - P_{VER}) = DCT \cdot Q(S - P_{HOR}) \quad (8)$$

식 (8)에서 이산 여현 변환이 분배법칙이 성립된다는 특성을 이용하여 양변에 변환 및 양자화를 하면 식 (9)를 유도할 수 있다. 이 때, 양자화로 인한 양자화 에러가 존재하고 양자화는 분배법칙이 성립하지 않기 때문에 양변은 서로 정확히 동일하지는 않다. 하지만 H.264/AVC에서는 스칼라 양자화(Scalar Quantization)를 사용하고 있기 때문에 서로 유사한 특성이 존재한다. 식 (9)를 이용하여 식 (10)을 유도할 수 있으며, 다시 이산 여현 변환의 분배법칙을 이용하여 식 (11)을 유도할 수 있다.

$$DCT \cdot Q(S) - DCT \cdot Q(P_{VER}) \cong DCT \cdot Q(S) - DCT \cdot Q(P_{HOR}) \quad (9)$$

$$DCT \cdot Q(P_{HOR}) - DCT \cdot Q(P_{VER}) \cong 0 \quad (10)$$

$$DCT \cdot Q(P_{HOR} - P_{VER}) \cong 0 \quad (11)$$

식 (11)은 두 모드의 예측 화소 블록의 차분 블록을 변환 및 양자화를 할 경우에 영행렬(0)과 유사하다는 것을 의미한다. 다시 말해, 식 (11)이 성립하는 경우는 두 모드의 예측 화소 블록이 서로 유사한 경우이며, 특히 양자화 매개변수가 작을수록 상당히 유사한 블록을 갖는다. 결국 식 (11)이 성립할 경우, 식 (12)와 같이 두 예측 화소 블록이 서로

유사하기 때문에 복원된 블록 또한 서로 유사한 특성을 갖는다. 여기서 식 (11)은 식 (8)에서 유도되었기 때문에 식 (11)이 성립하다면 울-왜곡 비용 중, 울(Rate)에 대한 비용은 서로 동일(두 모드의 예측 모드를 표현하기 위한 비트량이 서로 동일할 경우)하고 복원된 블록은 서로 유사하다. 따라서 이 경우에는 두 예측 모드의 울-왜곡 비용이 서로 유사한 값을 갖기 때문에 울-왜곡 최적화 과정에서 두 예측 모드 중, 하나의 예측 모드만을 사용하면 압축 성능의 손실이 조금 발생할 수 있지만 전체적인 부호화 속도를 향상시킬 수 있다. 이와 같은 접근 방법을 Intra4×4에서의 9가지 예측 모드와 Chroma Intra에서의 4가지 예측 모드에 적용하여 모드 결정 과정의 속도를 향상시켰다.

2. 제안된 고속 Intra4×4 예측 모드 결정 방법

본 절에서는 Intra4×4의 9개의 예측 모드를 대상으로 앞 절의 방법을 적용하여 화면 내 부호화 속도를 향상시키는 방법을 제안한다. 앞서 설명한 바와 같이, 식 (11)이 성립할 경우, 두 예측 모드가 유사한 울-왜곡 비용을 갖는다. 하지만 Intra4×4는 9개의 예측 모드가 존재하기 때문에 식 (13)과 같이 하나의 예측 블록과 나머지 8개의 예측 블록과의 차분 계수 블록을 변환 및 양자화를 수행하여 그 결과가 전부 영행렬일 경우에 모든 예측 모드가 유사한 울-왜곡 비용을 갖게 된다.

$$\begin{aligned} 0 &= DCT \cdot Q(P_1 - P_2) = DCT \cdot Q(P_1 - P_3) \\ &= DCT \cdot Q(P_1 - P_4) = \dots = DCT \cdot Q(P_1 - P_9) \end{aligned} \quad (13)$$

하지만 2장에서 언급한 바와 같이 MPM은 예측 모드를 표현하는데 있어서 1 비트만 사용하고 다른 모드들은 4비트를 사용한다. 따라서 식 (13)이 성립할 경우, MPM의 비트량이 MPM이 아닌 모드와 비교하여 상대적으로 항

$$\left. \begin{aligned} P_{HOR} &\cong P_{VER}, \\ IQ \cdot IDCT \{ DCT \cdot Q(S - P_{VER}) \} &\cong IQ \cdot IDCT \{ DCT \cdot Q(S - P_{HOR}) \}, \\ \therefore R_{VER} &\cong R_{HOR} \end{aligned} \right\} \text{ if (11) is satisfied} \quad (12)$$

상 적기 때문에 MPM 이 최적의 모드로 선택될 확률이 높다. 이러한 특징을 이용하기 위하여, 식 (11)에서의 대표 모드를 MPM 으로 사용하여, 최종적으로 식 (14)를 유도하였다.

$$\mathbf{O} = DCT \cdot Q(P_{MPM} - P_i), \text{ for } i=0,1,\dots,7,8(i \neq MPM) \quad (14)$$

여기서 P_{MPM} 은 MPM 을 사용하였을 때의 예측 화소 블록이며, P_i 는 그림 1 (b)에 표시된 번호 상에서의 i 번째 예측 모드를 사용하였을 때의 예측 화소 블록이다. 이 때, i 번째 예측 모드가 MPM 일 경우에는 항상 영행렬이 나오기 때문에 계산할 필요가 없다.

따라서 식 (14)가 성립할 경우에는 Intra4×4에서 하나의 4×4 블록의 예측 모드를 결정하기 위한 유효-왜곡 최적화 과정을 생략하고 대표 모드인 MPM 을 현재 블록의 최적의 예측 모드로 사용하였다. 그림 2는 제안된 고속 Intra4×4 예측

모드 결정 방법을 나타낸 흐름도이다. 그림 2를 보면, Intra4×4에서 하나의 4×4 블록의 예측 모드를 결정하기 위한 유효-왜곡 최적화 과정에 앞서 식 (14)의 성립 여부를 판단한다. 식 (14)가 성립할 경우에는 유효-왜곡 최적화를 생략하고 MPM 으로 현재 블록을 부호화한다. 식 (14)가 성립하지 않을 경우에는 기존의 유효-왜곡 최적화를 이용하여 최적의 모드를 결정하여 현재 블록을 해당 예측 모드를 이용하여 부호화한다. 이 때, 정수 이산 역변환과 양자화를 예측 모드 개수만큼 수행해야하기 때문에 계산량이 많다. 이 문제를 해결하기 위하여 식 (14)를 계산하는 과정 중, 0 이 아닌 계수가 발생했을 경우에는 항상 식이 성립하지 않기 때문에 계산 과정을 중단하여 계산량을 줄였다.

표 1은 4장의 실험 조건 하에서 휘도 성분에 대하여 식 (15)와 같이 모든 예측 모드의 예측 화소 블록이 모두 동일한 경우와 제안된 방법이 사용하는 식 (14)가 성립하는 경우의 확률을 나타낸 표이다. 표에서 보이는 바와 같이 식

그림 2. 제안된 고속 Intra4×4 예측 모드 결정 방법의 흐름도

Fig. 2. The flowchart of the proposed fast Intra4×4 mode decision

표 1. 식 (14)와 식 (15)가 성립하는 경우의 확률

Table 1. The probabilities of the case that equation (14) or (15) is satisfied.

Resolution	Sequences	식 (15)가 성립하는 경우	식 (14)가 성립하는 경우
CIF	Coast guard	3.2%	9.7%
	Foreman	9.1%	29.2%
	News	17.8%	38.2%
	Soccer	7.3%	21.2%
4CIF	City	4.3%	14.8%
	Harbour	2.3%	7.5%
	Ice	29.3%	68.1%
	Crew	9.1%	33.6%
720p	City	4.7%	15.4%
	Bigships	7.6%	26.5%
	Night	7.6%	24.4%
	Shuttle start	33.4%	67.7%
Average		11.3%	29.7%

(15)가 성립하는 경우의 확률이 약 11.3%로 매우 낮기 때문에 이것을 이용하더라도 높은 부호화 속도의 향상을 기대하기 어렵다. 하지만 식 (14)가 성립하는 경우의 확률은 약 29.7%로 좀 더 빈번히 나타나기 때문에 이 방법을 이용하면 부호화 속도의 향상을 기대할 수 있다.

$$\mathbf{O} = P_{MPM} - P_i, \text{ for } i = 0, 1, \dots, 7, 8 (i \neq MPM) \quad (15)$$

3. 제안된 고속 Chroma Intra 예측 모드 결정 방법

하나의 MB (Macro-Block)에는 8×8 크기의 색차 블록 Cb, Cr이 존재하지만 H.264/AVC에서는 이 두 블록에 대하여 동일한 예측 모드를 사용하도록 정의 되어있다. 본 장의 1절에 설명한 접근 방법을 적용하여 4가지의 Chroma Intra 예측 모드 (DC, Horizontal, Vertical, Plane)를 대상으로 화면 내 부호화 속도를 향상시키는 방법을 제안한다. 고속 Chroma Intra 예측 모드 결정을 위하여 각 Cb, Cr 블록에 대하여 다음 식 (16)의 성립 여부를 판단하

였다. 만약 Cb, Cr 블록 중 한 블록이라도 식 (16)이 성립한다면 윗-왜곡 최적화를 생략하고 대표 모드인 DC 예측 모드를 최적의 예측 모드로 사용하였다. Chroma Intra 예측 모드들은 예측 모드를 표현하기 위한 비트량이 모두 동일하기 때문에 식 (16)이 성립할 경우, 모든 예측 모드들의 윗-왜곡 비용이 거의 유사하다. 이 경우 Intra4×4에서는 상대적으로 윗-왜곡 비용이 작은 MPM을 대표 모드로 선택하였으나 Chroma Intra 예측에서는 그러한 모드가 존재하지 않기 때문에 확률적으로 발생 빈도수가 가장 높은 예측 모드를 대표 모드로 선택하였다. 표 2는 4장의 실험 조건 하에서 Chroma Intra 예측 모드의 통계를 나타낸 표이다. 표에서 알 수 있듯이 DC 모드가 평균적으로 약 90.5%로 가장 발생 확률이 높았다. 따라서 DC 모드를 대표 모드로 선택하였다.

표 2. 윗-왜곡 최적화에 의해 결정된 chroma Intra 예측 모드의 통계

Table 2. The probabilities of the Chroma Intra prediction mode decided by the RDO.

Resolution	Sequences	DC	Vertical	Horizontal	Plane
CIF	Coast guard	97.0%	1.7%	0.8%	0.5%
	Foreman	90.2%	4.0%	3.3%	2.5%
	News	75.4%	12.9%	8.1%	3.6%
	Soccer	91.8%	3.8%	3.3%	1.1%
4CIF	City	96.6%	1.0%	1.6%	0.8%
	Harbour	94.3%	1.3%	3.2%	1.2%
	Ice	95.9%	0.9%	2.5%	0.7%
	Crew	78.7%	4.2%	11.4%	5.7%
720p	City	95.1%	2.0%	2.0%	0.9%
	Bigships	90.1%	7.1%	1.8%	1.0%
	Night	82.4%	7.2%	7.2%	3.2%
	Shuttle start	98.2%	0.7%	0.6%	0.5%
Average		90.5%	3.9%	3.8%	1.8%

또한, Chroma Intra 모드 결정에서도 마찬가지로 식 (16)을 계산하는 과정 중, 0 이 아닌 계수가 발생했을 경우에는 항상 식이 성립하지 않기 때문에 계산 과정을 중단하여 계산량을 줄였다.

$$\begin{aligned} \mathbf{O} &= DCT \cdot Q(CP_{DC} - CP_1) = DCT \cdot Q(CP_{DC} - CP_2) \\ &= DCT \cdot Q(CP_{DC} - CP_3) \end{aligned} \quad (16)$$

여기서 CP_{DC} 는 DC 예측을 사용하였을 때의 예측 화소 블록이며, CP_1 , CP_2 , CP_3 은 각각 Horizontal 예측, Vertical 예측, Plane 예측을 사용하였을 때의 예측 화소 블록이다.

IV. 실험 결과

제안된 방법의 성능을 평가하기 위하여 H.264/AVC의 참조 소프트웨어인 JM (Joint Model) 11.0에 제안된 방법을 구현하였다^[12]. 그리고 제안된 방법의 성능을 검증하기 위하여 Pan의 방법^[8]과 La의 방법^[13]과 성능을 비교하였다. Pan의 고속 화면 내 예측 모드 결정 방법은 원본 화소 블록에 대하여 에지 검출 알고리즘을 적용하여 에지 방향과 유사한 예측 모드들을 울-왜곡 최적화에서 후보 예측 모드로 사용하는 방법이며 La의 방법은 *MPM*이 아닌 모드가 가질 수 있는 최소 울-왜곡 비용과 *MPM*의 실제 울-왜곡 비용을 비교하여 최소 비용보다 작을 경우 울-왜곡 최적화를 생략하고 *MPM*만을 이용하여 부호화하는 방법이다. 실험은 펜티엄 IV-2.67 GHz,

4Gbyte RAM 사양의 PC에서 수행되었다. 다양한 해상도의 영상에 대하여 각각 100 프레임을 화면 내 슬라이스로 부호화하였다. 자세한 실험 조건은 표 3에 나타내었다.

표 4는 실험 결과를 나타낸 표이다. 표 4에서 ΔP 는 BD-PSNR을 의미하며 ΔB 는 BD-RATE를 나타내고 각각 VCEG-M33^[14]에 정의되어 있으며, 울-왜곡 곡선 상에서의 평균 PSNR 증가량과 평균 비트율 증가량을 나타낸다. 따라서 BD-PSNR은 양수일 때, BD-RATE는 음수일 때 화질이 향상 또는 압축률이 향상되었다고 평가할 수 있다. 부호화 시간 절약을 ΔT 는 다음 식 (17)과 같이 계산하였다.

$$\Delta T(\%) = \frac{T_{JM} - T_{method}}{T_{JM}} \times 100 \quad (17)$$

표 4에서 알 수 있듯이, Pan의 방법은 평균적으로 2.83%의 비트율이 증가하거나 0.18 dB의 PSNR이 감소하였으며 La의 방법은 평균적으로 0.18%의 비트율이 증가하거나 0.01 dB의 PSNR이 감소하였다. 제안된 방법은 0.91%의 비트율 증가 하거나 혹은 0.06 dB의 PSNR이 감소되는 것으로 나타났다. 부호화 시간 절약을 Pan의 방법이 평균적으로 53.6%, La의 방법은 평균적으로 19.0%, 제안된 방법은 약 62.6%의 시간을 절약하였다. 따라서 제안된 방법은 다른 방법들과 비교하여 부호화 시간 절약을 가장 높은 것으로 나타났다. 압축 효율면에서는 La의 방법보다 제안된 방법이 좋지 못하지만 La의 방법은 부호화 시간이 거의 절약되지 않는 단점이 있다.

그림 3은 3개의 영상에 대한 실험 결과를 울-왜곡 곡선으로 나타낸 그림이다. 울-왜곡 곡선 상에서 제안된 방법은 JM과 거의 유사한 성능을 보였고 Pan의 방법과 비교하여 최대 0.3 dB의 PSNR이 향상되었다. La의 방법은 JM의 결과와 매우 유사하기 때문에 표시하지 않았다.

울-왜곡 최적화에서 선택된 최적의 모드와 제안된 방법(PM) 혹은 Pan의 방법의 사용으로 인해 선택된 모드가 서로 일치할 확률을 알아보기 위하여 표 3의 실험 조건하에서 그 확률을 측정하여 표 5에 나타내었다. 표 5를 통하여 제안된 방법은 Pan의 방법보다 최적의 예측 모드를 결정하는

표 3. 실험 조건
Table 3. Experimental conditions

Sequences	Resolution	Experiment conditions
Coast guard	CIF (352x288)	Baseline 프로파일, 화면 내 부호화, QP={22, 27, 32, 37}, CAVLC, 울-왜곡 최적화 사용, Intra16×16, Intra4×4 사용
Foreman		
News		
Soccer		
City	4CIF (704x576)	
Harbour		
Ice		
Crew		
City	720p (1280x720)	
Bigships		
Night		
Shuttle start		

표 4. 실험 결과 (* PM: 제안된 방법)

Table 4. Experimental results (* PM: Proposed method)

Resolution	Sequence	PM		Pan		La		QP	Total encoding time(sec)				Time saving $\Delta T(\%)$		
		ΔP	ΔR	ΔP	ΔR	ΔP	ΔR		JM	PM	Pan	La	PM	Pan	La
CIF	Coast guard	-0.04	0.49	-0.12	1.72	-0.01	0.10	22	61.7	25.6	25.8	62.9	53.0	58.2	3.9
								27	53.8	16.8	23.0	53.5	68.1	57.3	6.7
								32	47.0	13.9	20.6	44.3	70.4	56.1	12.3
								37	42.2	11.6	18.6	35.4	72.9	55.9	22.4
	Foreman	-0.04	0.69	-0.14	2.30	-0.02	0.34	22	51.8	26.4	21.4	50.5	47.7	58.7	9.1
								27	46.1	17.8	19.2	41.8	58.7	58.4	16.2
								32	41.9	14.0	17.7	34.4	65.6	57.8	24.7
								37	39.0	10.7	17.0	29.2	72.1	56.5	32.3
	News	-0.07	0.89	-0.26	3.22	-0.01	0.12	22	50.3	32.0	21.0	47.9	35.7	58.4	12.3
								27	46.4	23.3	19.5	42.1	46.6	57.9	16.2
								32	42.6	18.1	18.1	36.3	55.0	57.4	22.1
								37	39.9	13.2	17.3	31.9	64.8	56.6	27.3
	Soccer	-0.07	1.11	-0.13	2.24	-0.01	0.12	22	57.8	29.2	24.6	59.6	49.6	57.3	4.5
								27	50.2	18.4	21.9	49.5	59.0	56.4	9.4
								32	44.0	13.2	19.7	39.7	68.5	55.3	17.0
								37	39.2	9.6	18.1	30.6	75.3	53.9	30.0
	CIF Average	-0.05	0.79	-0.16	2.37	-0.01	0.17						60.2	57.0	16.6
4CIF	City	-0.04	0.50	-0.24	3.20	-0.01	0.06	22	226.5	92.7	102.0	242.3	58.3	55.0	2.0
								27	200.5	59.9	93.2	210.2	68.3	53.5	5.1
								32	179.2	51.2	85.0	180.5	71.5	52.6	9.3
								37	163.4	43.8	80.2	154.2	73.6	50.9	15.9
	Harbour	-0.05	0.57	-0.24	3.03	-0.01	0.07	22	231.4	101.8	101.3	243.6	54.9	56.2	4.1
								27	207.0	66.8	93.2	215.1	65.4	55.0	6.0
								32	186.1	56.6	86.3	189.9	69.3	53.6	8.1
								37	170.1	49.5	81.2	168.5	71.0	52.3	11.9
	Ice	-0.07	1.23	-0.20	3.56	-0.02	0.31	22	163.9	55.8	92.2	133.2	65.9	43.8	27.5
								27	152.2	37.8	86.6	106.8	74.1	43.1	37.9
								32	146.8	32.5	83.8	96.9	77.5	42.9	41.6
								37	143.1	29.6	83.2	90.6	79.4	41.9	44.5
	Crew	-0.06	1.21	-0.17	3.65	-0.02	0.33	22	197.1	135.2	88.4	199.5	30.9	55.2	9.2
								27	173.5	85.4	80.3	162.3	46.2	53.7	17.3
								32	159.2	62.6	75.1	134.9	58.2	52.9	25.4
								37	150.9	44.0	71.2	112.1	69.1	52.8	33.8
	4CIF Average	-0.05	0.88	-0.21	3.36	-0.01	0.20						64.6	51.0	18.7
720p	City	-0.04	0.58	-0.22	3.08	0.00	0.05	22	527.4	256.0	234.7	535.4	50.4	55.5	4.4
								27	465.7	141.1	210.1	459.7	67.5	54.9	7.6
								32	415.3	116.4	193.3	392.0	71.8	53.5	12.4
								37	375.9	98.0	180.6	330.1	74.0	52.0	19.3
	Bigships	-0.05	0.93	-0.12	2.33	-0.01	0.15	22	488.8	268.8	218.0	490.2	43.6	55.4	5.8
								27	426.1	157.1	194.8	401.9	59.1	54.3	13.2
								32	385.1	109.6	180.7	321.3	69.7	53.1	23.4
								37	355.4	83.2	169.0	255.8	76.2	52.4	34.8
	Night	-0.11	1.51	-0.22	3.19	-0.01	0.16	22	490.8	344.6	221.5	494.2	27.0	54.9	6.7
								27	437.3	199.2	201.4	423.0	42.5	54.0	11.0
								32	398.2	142.0	187.4	362.3	56.8	52.9	16.7
								37	370.1	107.2	177.6	309.6	66.8	52.0	24.0
	Shuttle start	-0.05	1.16	-0.10	2.44	-0.01	0.32	22	378.0	111.1	188.1	295.2	68.5	50.2	28.8
								27	353.3	80.5	177.1	238.8	73.9	49.9	38.7
								32	339.4	66.7	171.3	201.2	78.0	49.5	46.5
								37	330.8	58.9	167.0	177.6	80.3	49.5	51.7
	720p Average	-0.06	1.04	-0.17	2.76	-0.01	0.17						62.9	52.7	21.6
Total average		-0.06	0.91	-0.18	2.83	-0.01	0.18						62.6	53.6	19.0

확률이 QP 22에서 평균적으로 약 28.7%, QP 32에서 약 28.3%가 높은 것을 알 수 있다.

(a) Foreman 영상의 율-왜곡 곡선

(b) City 영상의 율-왜곡 곡선

(c) Night 영상의 율-왜곡 곡선

그림 3. 율-왜곡 곡선들
Fig. 3. Rate-Distortion curves

표 5. 최적 예측 모드의 적중 확률

Table 5. Hit ratio of the best prediction mode

Resolution	Sequences	QP 22		QP 32	
		PM	Pan	PM	Pan
CIF	Coast guard	71.3%	63.4%	93.5%	66.6%
	Foreman	72.1%	44.0%	72.7%	55.1%
	News	89.0%	49.2%	85.9%	55.1%
	Soccer	67.1%	47.2%	79.1%	53.1%
4CIF	City	76.6%	58.7%	90.2%	65.9%
	Harbour	83.6%	69.2%	96.2%	75.3%
	Ice	91.4%	24.5%	88.2%	31.8%
	Crew	80.9%	38.9%	73.5%	51.1%
720p	City	71.2%	52.4%	84.8%	61.3%
	Bigships	69.5%	41.3%	76.3%	57.6%
	Night	79.0%	48.3%	83.0%	57.9%
	Shuttle start	88.2%	57.9%	71.7%	25.0%
Average		78.3%	49.6%	82.9%	54.6%

V. 결 론

본 논문에서는 부호화 속도를 향상시키기 위한 새로운 고속 화면 내 예측 모드 결정 방법을 제안하였다. 제안된 방법은 Intra4×4와 Chroma Intra 예측 모드 결정 방법에 적용되었으며, 정수 이산 여현 변환과 양자화를 이용하여 모든 화면 내 예측 모드들 중 최고 확률 모드(휘도 성분의 *MPM*, 색차 성분의 DC 모드)가 가장 최적의 예측 모드인 경우를 미리 판단하여, 이 경우에는 율-왜곡 최적화 과정을 생략하고 최고 확률 모드만을 사용하는 방법이다. 실험 결과, 제안된 방법은 PSNR의 손실이 거의 없이 약 63%의 부호화 시간을 절약하는 것으로 나타났다. 따라서 제안된 방법은 실시간 부호화를 요구하는 분야에서 유용하게 사용될 수 있을 것으로 기대된다.

참 고 문 헌

- [1] ITU-T Recommendation H.264 and ISO/IEC 14496-10, "Advanced video coding for generic audiovisual services", May 2003.
- [2] Thomas Wiegand, Gary J. Sullivan, Gisle Bjntegaard, and Ajay Luthra, "Overview of the H.264/AVC video coding standard", IEEE Trans. Circuits Syst. Video Technol., pp. 560-576, July 2003.
- [3] Gary J. Sullivan and Thomas Wiegand, "Video compression - from concepts to the H.264/AVC standard", Proc. IEEE, pp. 18-31,

- January 2005.
- [4] Detlev Marpe, Heiko Schwarz, Thomas Wiegand "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard", IEEE Trans. Circuits Syst. Video Technol., pp. 620-636, July 2003.
 - [5] Peter List, Anthony Joch, Jani Lainema, Gisle Bjøntegaard, Marta Karczewicz "Adaptive Deblocking Filter", IEEE Trans. Circuits Syst. Video Technol., pp. 614-619, July 2003.
 - [6] Kamaci, N., Altunbasak, Y., "Performance Comparison of the Emerging H.264 Video Coding Standard with the existing standards", ICME, Vol.1, pp. 345-8, July 2003.
 - [7] Gary J. Sullivan and Thomas Wiegand, "Rate-Distortion Optimization for Video Compression", IEEE Signal Processing Magazine, pp 74-90, November 1998
 - [8] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu, and S. Wu, "Fast mode decision algorithm for intra-prediction in H.264/AVC video coding," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 7, pp. 813 - 822, July 2005.
 - [9] F. Pan, X. Lin, S. Rahardja, K. P. Lim, and Z. G. Li, "A Directional Field Based Fast Intra-Mode Decision Algorithm for H.264 Video Coding," IEEE Inter. Conf. on Multimedia and Expo, vol.2, pp. 1147-1150, June 2004.
 - [10] R. Garg, M. Jindal, M. chauhan, "Statistics Based Fast Intra-Mode Detection", Proc. of SPIE vol.5960, pp 2085-2091, 2005.
 - [11] Henrique S. Malvar, Antti Hallapuro, Marta Karczewicz, and Louis Kerofsky, "Low-Complexity Transform and Quantization in H.264/AVC", IEEE Trans. Circuits Syst. Video Technol., pp. 598-603, July 2003.
 - [12] http://iphome.hhi.de/suehring/tml/download/old_jm/jm11.zip
 - [13] Fast 4x4 Intra-prediction based on the most probable mode in H.264/AVC-IEICE Electronics express, Vol.5, No.19, p782-788
 - [14] Gisle Bjøntegaard, "Calculation of Average PSNR Differences between RD curves," ITU-T SC16/Q6, 13th VCEG Meeting, Austin, Texas, USA, Apr. 2001, Doc.VCEG-M33.

저 자 소 개

김 대 연

- 2005년 2월 : 세종대학교 인터넷학과 (학사)
- 2005년 3월 ~ 현재 : 세종대학교 컴퓨터공학과 (석박통합과정)
- 주관심분야 : Video compression, Image compression, Image processing

김 정 필

- 2010년 2월 : 세종대학교 컴퓨터공학과 (학사)
- 2010년 3월 ~ 현재 : 세종대학교 컴퓨터공학과 (석사과정)
- 주관심분야 : 영상압축, H.264/AVC

이 영 렬

- 1985년 2월 : 서강대학교 전자공학과 (학사)
- 1987년 2월 : 서강대학교 전자공학과 (석사)
- 1999년 2월 : 한국 과학기술원 전기·전자공학과 (박사)
- 1987년 1월 ~ 1994년 2월 : 삼성전자 R&D 센터 Digital Media Lab.
- 1999년 3월 ~ 2001년 8월 : 삼성전자 R&D 센터 Digital Media Lab. 수석연구원
- 2001년 9월 ~ 현재 : 세종대학교 전자정보대학 컴퓨터공학부 컴퓨터공학과 부교수
- 주관심분야 : Multi-view video coding, MPEG video compression, Multimedia(MPEG) systems, H.264/AVC, Image compression, Image processing, Multimedia data(A/V/System) over IP, Embedded system for A/V transport, Transcoding, Watermarking