H.264/AVC비디오의 고속 부호화를 위한 인트라모드 선택 알고리듬

정회원 김 동 형*, 종신회원 정 제 창**

Intra-mode Decision Algorithm for Fast Encoding of H.264/AVC Video

Donghyung Kim* ***Regular Member***, Jechang Jeong** ***Lifelong Member***

요 약

H.264 부호화 표준은 부호화 효율을 높이기 위해 1/4 화소 단위의 움직임 추정, 다중 참조 프레임, 인트라 예측, 루프 필터, 다양한 블록 크기의 지원 등과 같은 새로운 부호화 도구들을 사용한다. 이를 통해 이전의 비디오 부호화 표준들에 비해 율-왜곡(率-歪曲) 관점에서 높은 성능을 보이지만 그로 인해 부호기의 복잡도는 상당히 증가한다. 본 논문은 부호기 복잡도의 증가를 초래하는 주요 부호화 도구들 중 인트라 매크로블록 모드 선택의 복잡도 감소에 주안(主眼)점을 두며, 이를 위한 고속 부호화 알고리듬을 제시한다. 제안하는 알고리듬은 먼저 간단한 전처리과정을 통해서 Intra4x4 모드의 예측모드를 한정하고, 선택된 Intra4x4의 예측모드를 사용하여 다른 인트라 모드들의 예측모드를 결정한다. 실험결과 제안하는 알고리듬은 기존의 방법보다 높은 성능을 보이며 참조소프트웨어와 비교하여 약 82%의 속도향상을 가져오는 것으로 나타난다.

Key Words : fast encoding, H.264 Video, intra-mode selection, mode prediction, primary edge mode

ABSTRACT

For the improvement of coding efficiency, the H.264 standard uses new coding tools such as VBS, 1/4-pel accurate ME, multiple references, intra prediction, loop filter, etc. Using these coding tools, H.264 has achieved significant improvements from rate-distortion point of view compared to existing standards. However, the encoder complexity is greatly increased due to these coding tools. We focus on the complexity reduction method of intra-mode decision. Our algorithm first restricts selective prediction modes of intra4x4 using a simple preprocessing. The prediction modes of intra4x4 are used for restricting those of the other inter-modes. Simulation results show that the proposed method outperforms other conventional methods and save about 82% of total encoding time.

Ⅰ. 서 론

H.264/AVC 표준은 ITU-T의 VCEG(Video Coding Experts Group)과 ISO/IEC의 MPEG (Moving Picture Expert Group)의 공동 작업으로 제정되었다[1]. H.264/AVC 비디오 부호화 방식은 이전의 비디오 부호화 표준들(MPEG2, MPEG4 Part2, H.263)에는 없던 새로운 부호화 도구들이 추가되었으며, 이러한 도구들은 H.264 부호화 표준이 이전의 비디오 부호화 표준과 비교하여 보다 높은 압축효율을 갖게 하였

※ 본 연구는 서울시 산학연협력사업으로 구축된 서울 미래형콘텐츠컨버전스 클러스터 지원으로 수행되었습니다.
* 한양대학교 전자통신전파공학과 영상처리 및 신호처리 연구실 (kimdh@ece.hanyang.ac.kr)
** 한양대학교 전자전기컴퓨터 공학부 (jjeong@ece.hanyang.ac.kr)
논문번호 : KICS2007-04-172, 접수일자 : 2007년 04월 12일, 최종논문접수일자 : 2007년 6월 21일

다. 새롭게 추가된 부호화 도구들에는 4×4 단위의 정수변환(integer transform), CAVLC(Context-based Adaptive Variable Length Coding) 또는 CABAC (Context-based Binary Arithmetic Coding)을 통한 엔트로피부호화, 루프필터(loop-filter), 인트라예측(intra-prediction)부호화, 다중참조 영상의 사용, 다양한 블록 크기(VBS: variable block size), 적응적 프레임/필드 부호화(Adaptive Frame/Field Coding) 및 NAL(Network Abstraction Layer)등의 있다.

이렇게 새롭게 추가된 부호화 도구들은 H.264 부호화 표준이 이전의 여러 동영상 부호화 표준들에 비해 보다 높은 압축 효율을 가질 수 있게 하는데, 다섯 장의 참조 프레임의 사용과 32 화소의 움직임 추정 영역을 사용하여 메인 프로파일로 부호화 한 경우, H.264/AVC 부호화 표준은 MPEG-2와 비교하여 약 48~78%까지의 비트율 감소효과가 있다고 알려져 있다[3]. 하지만 비트율 감소 효과와 더불어 부호기의 복잡도는 상당히 높아진다. 특히 새로이 추가된 부호화 도구들 중 매크로블록모드 선택과정은 높은 복잡도를 가지는데 이는 선택가능한 모든 예측모드에서의 비용값을 계산하여야 하기 때문이다.

실시간 H.264/AVC 부호화의 실현을 위해, 부호기의 복잡도 감소에 대한 연구는 2003년 표준이 제정된 이후 다각적으로 연구되어 왔다. Yin 등은 인터프레임(inter-frame)내에서의 고속 매크로블록 모드 선택 방법과 부화소(sub-pixel) 단위에서의 고속 움직임추정 기법을 제안하였다[4]. Huang 등은 다중 참조 프레임을 사용하는 H.264/AVC의 부호화 과정의 복잡도를 줄이기 위해 이전 한 프레임에서의 움직임 추정 및 인트라 예측을 통해 얻어진 정보를 이용하여 잔여 참조 프레임의 사용여부를 결정하는 알고리듬을 제안하였다[5]. Park 등이 제안한 고속 부호화 알고리듬은 인트라모드의 고속 선택을 통하여 고속 부호화를 실현하는 방법으로 이는 휘도(Luminance)블록의 인트라예측에 사용되는 인트라 모드 중 4×4 블록단위로 예측을 수행하는 Intra4×4 모드의 예측모드를 이미 계산된 Intra16×16의 예측모드를 사용하여 추정함으로써 연산량을 감소시키며, 색차(Chrominance) 블록의 예측모드 또한 계산된 Intra16×16의 예측모드를 사용한다[6]. Pan 등이 제안한 고속 부호화 알고리듬 또한 고속 인트라 모드 선택을 통해 구현되며, 이는 각 인트라모드에서의 예측모드를 결정하기 위해서 소벨연산(Sobel Operating)을 이용하여 에지맵(Edge Map)을 산출하고 이를 토대로 각 예측모드에서의 선택가능한 예측모드를 한정함으로써 인트라모드선택 과정에서의 복잡도를 감소시킨다[7].

본 논문에서 제안하는 H.264/AVC 고속 부호화 알고리듬 또한 앞서 Park 그리고 Peng이 제안한 것과 같이 인트라모드의 고속 선택을 통해 고속 부호화를 실현하는 방법으로서 이는 적은 연산량만을 가지는 전처리과정(Preprocessing)을 통하여 Intra4×4 모드에서의 선택가능한 예측모드를 한정한다. 또한 선택된 Intra4x4의 예측모드를 이용하여 Intra16×16의 예측모드 및 색차블록에 대한 예측모드를 결정한다.

논문의 구성은 2절에서 참조 소프트웨어(JM 10.1[8])에 구현되어 있는 매크로블록 모드 선택 알고리듬에 대해서 기술하고, 3절에서 논문에서 제안하는 고속인트라모드의 선택방법의 기술과 함께 연산량을 산술적으로 비교한다. 4절에서는 8개의 서로 다른 크기의 표준테스트 영상에 대해서 제안하는 알고리듬을 적용함으로써 효용성을 보이고 마지막 절에서 결론을 맺는다.

Ⅱ. 참조소프트웨어에서의 인트라모드 선택 기법

H.264/AVC FREext(Fidelity Range Extension)이 사용되지 않은 경우 인트라프레임(Intra-frame)에 포함되어있는 모든 매크로블록들은 휘도(luminance) 블록의 부호화를 위한 두 개의 인트라모드(Intra-mode) 즉, Intra16×16 및 Intra4×4 모드와 색차(chrominance) 블록의 부호화를 위한 하나의 인트라모드 즉, ChromaIntra8x8(컬러포맷이 4:2:0의 경우)을 사용하여 부호화된다. 각 인트라모드는 서로 다른 예측모드의 수를 가지는데, 그림 1은 Intra4x4 모드에서 사용하는 9가지의 예측모드를 나타내며, 그림 2는 Intra16x16 모드에서 사용되는 4가지의 예측모드를 나타낸다. ChromaIntra8x8의 예측모드는 Intra16x16과 동일하며 통계적 발생빈도에 따라 그 순서만을 달리한다[1].

율-왜곡 최적화(RDO: Rate Distortion Optimization) 기법이 사용되는 경우 참조소프트웨어는 색차블록의 각 예측모드에 대해서 Intra16×16과 Intra4×4의 율-왜곡 비용값(rate-distortion cost)을 비교하여 최소의 비용값을 갖는 모드를 해당 매크로블록의 휘도 및 색차블록의 인트라모드로서 선택한다. 이 과정에서 Intra4×4는 각 4×4 블록마다 율-왜곡 비용값을 계산하며 매크로블록 내에 포함되어 있는 16개의 모든

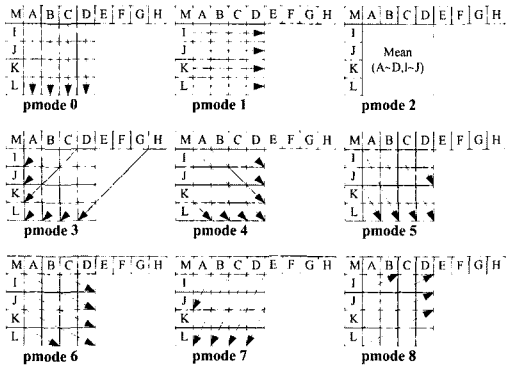


그림 1. Intra4x4의 9가지 예측모드
Fig. 1. Nine prediction modes of Intra4x4

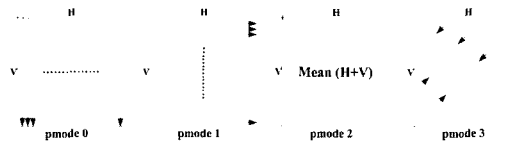


그림 2. Intra16x16 및 ChromaIntra8x8의 4가지 예측모드
Fig. 2. Four prediction modes o Intra16X16 and ChromaIntra8x8

4×4 블록에서의 율-왜곡 비용값의 합을 Intra16×16의 율-왜곡 비용값과 비교한다. 율-왜곡 비용함수는 식 (1)과 같이 계산될 수 있는데, 여기에서 왜곡(Distortion) 값은 각 매크로블록 모드에서의 SNR을 구함으로써 계산할 수 있고, 율(Rate) 값은 매크로블록의 부호화가 끝나는 시점에서의 압축된 비트열(Bitstream)의 비트율을 나타낸다.

율-왜곡 최적화(RDO: Rate Distortion Optimization) 기법이 사용되는 경우 참조소프트웨어는 색차블록의 각 예측모드에 대해서 Intra16×16과 Intra4×4의 율-왜곡 비용값(rate-distortion cost)을 비교하여 최소의 비용값을 갖는 모드를 해당 매크로블록의 휘도 및 색차블록의 인트라모드로서 선택한다. 이 과정에서 Intra4×4는 각 4×4 블록마다 율-왜곡 비용값을 계산하며 매크로블록 내에 포함되어 있는 16개의 모든 4×4 블록에서의 율-왜곡 비용값의 합을 Intra16×16의 율-왜곡 비용값과 비교한다. 율-왜곡 비용함수는 식 (1)과 같이 계산될 수 있는데, 여기에서 왜곡(Distortion) 값은 각 매크로블록 모드에서의 SNR을 구함으로써 계산할 수 있고, 율(Rate) 값은 매크로블록의 부호화가 끝나는 시점에서의 압축된 비트열(Bitstream)의 비트율을 나타낸다.

$$RDcost = Distortion + lambda \times Rate \qquad (1)$$

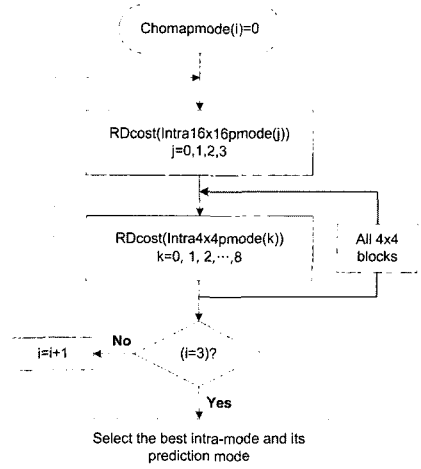


그림 3. H.264/AVC 참조소프트웨어에서의 Intra-mode 선택 과정
Fig. 3. The process of Intra-mode selection in the H.264/AVC reference software

앞서 언급한 바와 같이 각 인트라모드는 서로 다른 예측모드의 가지 수를 가지기 때문에, 하나의 인트라 매크로블록에 대한 최적의 인트라모드를 선택하기 위해서는 모든 예측모드들의 조합에 대한 율-왜곡 비용값을 계산하여야하며, 최종적으로 부호화에 사용된 해당 매크로블록의 예측모드는 정보는 부호화된 비트스트림에 포함되게 된다. 따라서 Intra16×16의 경우 매크로블록 당 하나의 예측모드 정보를 포함하게 되며, Intra4×4의 경우에는 하나의 매크로블록 당 16개의 예측모드 정보를 포함하게 된다.

그림 3은 율-왜곡 최적화기법이 사용되는 경우 참조소프트웨어서의 인트라모드 선택과정을 나타내는 블록 다이어그램이다. 그림 3에서 볼 수 있는 바와 같이 하나의 색차성분의 예측모드에 대해서 먼저 Intra16×16의 네 개의 예측모드들 중 최소의 비용값을 갖는 예측모드를 선택하고, 이후 Intra4×4의 비용값을 계산한다. 이러한 과정은 네 가지의 색차성분에 대해서 수행되며 최종적으로 최소의 비용값을 갖는 모드를 휘도 블록 및 색차 블록의 매크로블록 모드로서 선택하게 된다. 표 1은 그림3에 포함되어 있는 각 인트라 모드의 예측모드를 나타내고 있으며, 이때 Chromapmode(i), Intra16×16pmode(i), 그리고 Intra4×4pmode(i)는 각각 ChormaIntra8×8, Intra16×16 및 Intra4×4의 예측모드를 나타낸다.

참조소프트웨어에서의 인트라모드 선택과정에서 볼 수 있듯이 좌측 및 상측, 그리고 상우측의 블록이 모두 사용가능한 경우, 인트라프레임내에 포함되어

표 1. ChromaIntra8×8(Apmode), Intra16×16(Bpmode) 및 Intra4×4(Cpmode)의 예측모드
Table 1. Prediction modes of ChromaIntra8×8(Apmode), Intra16×16(Bpmode), and Intra4×4(Cpmode)

i	Apmode	Bpmode	Cpmode
0	DC	Vertical	Vertical
1	Horizontal	Horizontal	Horizontal
2	Vertical	DC	DC
3	Plane	Plane	Diag.-downleft
4	-	-	Diag.-downright
5	-	-	Vert.-right
6	-	-	Hor.-down
7	-	-	Vert.-left
8	-	-	Hor-up

있는 하나의 매크로블록에 대한 최적의 인트라모드를 결정하기 위해서는 ChromaIntra8×8 색차 신호의 4가지의 예측모드에 대해서 Intra16×16 및 Intra4×4 휘도 신호의 4가지 및 9가지의 모든 조합 가능한 예측모드에 대한 율-왜곡 비용함수 계산과정이 필요하다. 이는 하나의 매크로블록에 대해서 최적의 Intra-mode 선택을 위해서는 최대 4×(1×4+16×9)=592회의 율-왜곡 비용 함수 계산과정이 필요함을 의미한다.

Ⅲ. 제안하는 알고리듬

제안하는 고속 인트라모드 선택 알고리듬은 각 인트라모드에서의 선택가능한 예측모드를 한정하여 율-왜곡 비용값의 계산과정을 줄임으로써 인트라프레임의 고속 부호화를 실현한다. 그림 4는 제안하는 알고리듬의 전체적인 구성도를 나타내는 것으로 2절에서 소개한 그림 3에 나타나 있는 참조소프트웨어에서의 방법과는 다르게 각 색차성분의 예측모드에 대해서 Intra4×4에 대한 율-왜곡 비용함수의 연산을 먼저 수행한 후 Intra16×16에 대한 율-왜곡 비용값을 계산한다. 이때 모든 인트라모드에서의 율-왜곡 비용계산은 한정된 예측모드에서만 수행된다.

3.1 Intra4×4 고속 예측모드 선택 알고리듬

제안하는 알고리듬은 Intra4×4 모드에서의 선택가능한 예측모드 후보를 결정하기 위해서 4×4 블록내의 에지방향을 추정하고, 이를 이용하여 Intra4×4 모드에서의 선택가능한 예측모드 후보를 결정한다.

먼저 블록내의 에지를 추정하기 위한 방법으로 그림 5에 보이는 바와 같이 4×4 블록의 좌측 및 상측에 위치하는 화소위치에서의 에지의 크기 및 방향을 Prewitt 미분기를 이용하여 계산한다. 이때 좌측 및 상측 화소위치만을 고려하는 이유는 연산

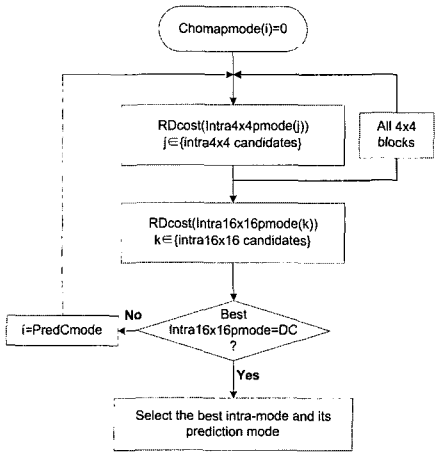


그림 4. 고속 Intra-mode 선택 알고리듬의 블록도
Fig. 4. Block diagram of the fast Intra-mode selection method

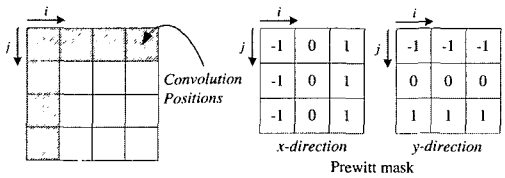


그림 5. 에지크기 및 각도 계산의 화소 위치 및 Prewitt 미분기
Fig. 5. Position to calculate edge magnitude and phase, and Prewitt mask

량을 최소화함과 동시에 4×4블록의 예측부호화 과정에 사용되는 주변화소값들의 영향을 함께 고려하기 위함이다.

Prewitt 미분기를 이용한 에지의 크기 및 방향은 해당 4×4 블록에 x 및 y 방향의 Prewitt 미분기를 콘볼루션(Convolution)함으로써 구할 수 있다. 예를 들어, (i,j)위치의 화소값 $p_{i,j}$에서의 콘볼루션 결과는 식 (2)와 같으며, 이때 에지의 크기($mag_{i,j}$)및 방향($ang_{i,j}$)은 각각 식(3) 및 식(4)와 같다.

$$\begin{aligned} dx_{i,j} &= (p_{i-1,j+1} - p_{i-1,j-1}) + (p_{i,j+1} - p_{i,j-1}) + (p_{i+1,j+1} - p_{i+1,j-1}) \\ dy_{i,j} &= (p_{i+1,j-1} - p_{i-1,j-1}) + (p_{i+1,j} - p_{i-1,j}) + (p_{i+1,j+1} - p_{i-1,j+1}) \end{aligned} \quad (2)$$

$$mag_{i,j} = |dx_{i,j}| + |dy_{i,j}| \quad (3)$$

$$ang_{i,j} = \arctan\left(\frac{dy_{i,j}}{dx_{i,j}}\right) \quad (4)$$

위의 에지의 크기 및 각도의 계산과정에서 에지의 각도는 0~2π(0~360°)내의 모든 값을 가질 수 있지만 제안하는 알고리듬에서는 단지 예측모드를 추정하기 위해 사용되기 때문에 인접한 예측방향을 갖는 예측모드의 각도로 고려한다. 가령 에지의 각

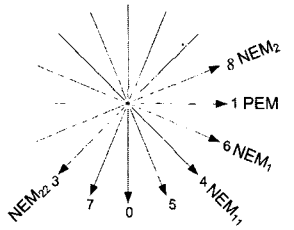


그림 6. PEM으로 pmode1이 선택된 경우 NEM1~22의 선택
Fig. 6. Selection of NEM1~22 when PEM is pmode1

도가 π/10(18°)의 경우 가장 근접한 예측모드의 예측방향인 π/8(22.5°)인 것으로 고려한다. 따라서 실제 적용에 있어서는 $\tan(ang_{i,j}$=0°, 22.5°, 45°, 67.5°, 90°, 112.5°, 135°, 157.5°)의 값을 미리 계산해 놓음으로써 식 (4)에 포함된 *arctan( )*함수의 계산과정은 생략할 수 있다[7].

고려하는 7개의 화소위치에서의 에지크기 및 각도의 계산과정이 끝나면, 각 에지각도에서의 에지크기의 합을 구하고 최대 에지크기의 합을 가지는 에지각도와 가장 인접한 예측방향을 갖는 예측모드를 PEM(Primary Edge Mode)로 정의한다. 또한 PEM과 예측방향 측면에서 이웃한 두 개의 예측모드를 NEM(Neighboring Edge Mode)라 정의하며 이때 시계방향으로 이웃한 예측모드를 NEM_1, 그리고 반시계방향으로 이웃한 예측모드를 NEM_2라 정의한다. 마지막으로 NEM_1에 시계방향으로 이웃한 예측모드를 NEM_{11}, NEM_2에 반시계방향으로 이웃한 예측모드를 NEM_{22}라 정의한다.

그림 6은 0°의 예측방향을 갖는 예측모드 pmode1(0°)이 PEM으로 선택된 경우 NEM_1, NEM_2, NEM_{11}, NEM_{22} 선택의 예를 나타낸다.

이상의 과정이 완료되면 제안하는 알고리듬은 Intra4×4에서 사용 가능한 예측모드를 세 가지 조건의 만족 여부에 따라 {DC, PEM, NEM_1, NEM_2, NEM_{11}, NEM_{22}}의 부분집합 중 하나로 선택한다.

먼저 DC 및 PEM 예측모드에서의 율-왜곡 비용함수를 계산하며, 만일 DC에서의 비용값이 PEM에서의 비용값보다 충분히 작을 경우 즉, 식 (5)가 만족되는 경우에는 해당 4×4 블록이 에지성분이 거의 없는 평탄한 영역인 것으로 판단하여 더 이상의 예측모드에 대한 비용값 계산과정은 생략된다.

$$RDcost(DC) \times K < RDcost(PEM) \qquad (5)$$
$$K = (1 + QP/(51 \times 2)).$$

식 (5)에서 K는 양자화 계수에 따라 다르게 선택되며 0~51의 양자화계수(QP)를 갖는 H.264/AVC 부호화 표준의 특성을 고려하면 K는 1~1.5의 값을 가짐을 알 수 있다. 영상이 낮은 비트율로 부호화되는 경우 즉, 큰 양자화 계수가 사용되는 경우 K값은 1.5에 가까운 값을 가지며, 높은 비트율로 부호화되는 경우 즉, 작은 양자화 계수가 사용되는 경우 K의 값은 1에 가까워진다. 이러한 K값의 식은 실험적으로 결정되었으며, QP값과 1/54의 기울기로 비례할 때 가장 성능을 보이는 것으로 나타났다.

식 (5)가 만족되지 않는 경우는 해당블록내 에지가 있는 경우로 판단되어 PEM 주변의 예측모드에서도 더 작은 율-왜곡 비용값을 예상할 수 있으며, 제안하는 알고리듬은 PEM에 이웃한 두 개의 예측모드인 NEM_1, NEM_2에서의 율-왜곡 비용값을 추가적으로 계산한다. 이후 PEM, NEM_1, NEM_2의 율-왜곡 비용값을 비교하여 PEM의 율-왜곡 비용값이 최소를 갖는 경우 즉, 식 (6)이 성립하는 경우에는 더 이상의 예측모드를 고려하지 않는다.

$$RDcost(PEM) = min(RDcost(PEM, \neq M_1, \neq M_2)) \qquad (6)$$

마지막으로 식 (6)이 만족하지 않는 경우 즉, NEM_1 또는 NEM_2에서 최소의 율-왜곡 비용값이 발생한 경우 하나의 예측모드를 추가적으로 고려하는데, 식 (7)과 같이 NEM_1의 율-왜곡 비용값이 NEM_2의 율-왜곡 비용값보다 같거나 작을 경우에는 NEM_{11}을 추가적으로 고려하며, 반대의 경우에는 NEM_{22}를 고려한다.

$$RDcost(\neq M_1) \leq RDcost(\neq M_2) \qquad (7)$$

따라서, 제안하는 알고리듬은 각 조건의 성립여부에 따라 Intra4×4의 예측모드를 {DC, PEM}, {DC, PEM, NEM_1, NEM_2}, 그리고 {DC, PEM, NEM_1, NEM_2, NEM_{11} or NEM_{22}} 중 하나로 제한한다. 그림 7는 이상에서 기술한 고속 Intra4×4 고

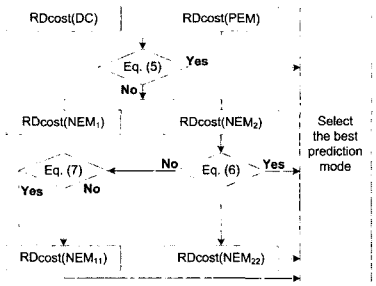


그림 7. Intra4×4 예측모드 후보를 선택을 위한 블록 다이어그램
Fig. 7. Block diagram for selection of Intra4×4 prediction mode candidates

속 예측모드 선택을 위해 예측모드를 한정하는 과정을 나타내는 블록다이어그램이다.

3.2 Intra16×16 고속 예측모드 선택 알고리듬

Intra16×16모드는 네 가지 예측모드를 가진다. DC를 제외한 각 예측모드에서의 예측방향은 그림 8에 나타나있는 것처럼 가로, 세로 및 좌우대각선 방향을 나타낸다.

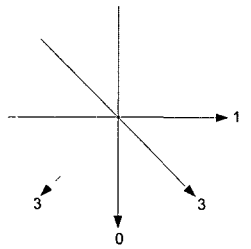


그림 8. Intra16×16의 예측모드의 예측방향
Fig. 8. Prediction directions of prediction modes in Intra16×16

Intra16×16의 고속 예측모드 선택을 위해서 제안하는 알고리듬은 이전의 Intra4×4의 예측모드 선택 과정에서 이미 구한 16개의 4×4블록들에 대한 예측모드 정보를 사용한다. 먼저 Intra4×4의 16개 4×4 블록들 중 pmode0(90°)로 선택된 블록의 수를 num(V)로 정의하고 pmode1(0°)으로 선택된 블록의 수를 num(H), pmode2로 선택된 블록의 수를 num(DC), 마지막으로 pmode3(45°) 또는 pmode4(135°)로 선택된 블록의 수를 num(P)으로 정의한다.

만일 식 (8)과 같이 DC로 선택된 블록의 개수가 80% 이상의 경우에는 해당 매크로블록을 평탄영역인 것으로 판단하여 예측모드로서 DC만을 고려한다.

$$\frac{\nu m(DC)}{\nu m(V)+\nu m(H)+\nu m(DC)+\nu m(P)} > 0.8 \qquad (8)$$

식 (8)이 만족하지 않는 경우에는 해당 매크로블록 내에 에지가 있는 것으로 판단하여 Intra16×16의 예측모드로서 DC 이외에 num(V), num(H), num(P) 중 최대값을 갖는 예측모드를 고려한다. 따라서 제안하는 알고리듬은 Intra16×16의 예측모드의 수를 1개 또는 2개로 한정한다.

3.3 ChromaIntra8x8 고속 예측모드 선택 알고리듬

앞서 기술한 Intra4×4 및 Intra16×16은 계산과정은 먼저 색차신호인 ChromaIntra8x8의 예측모드가 DC의 경우에 수행된 결과이며, 위의 과정을 나머지 ChromaIntra8×8의 예측모드에서도 반복해서 수행하여야 한다. 제안하는 알고리듬은 ChromaIntra8×8의 고속 예측모드 선택을 위해서 ChromaIntra8×8의 예측모드가 DC일 때의 Intra16×16의 예측모드를 사용하여 이후의 ChromaIntra8×8 예측모드에서의 계산 여부를 결정한다. 만일 Intra16×16에서 선택된 예측모드가 DC의 경우에는 더 이상의 ChromaIntra8×8에서의 예측모드는 고려하지 않으며, 만일 Intra16×16의 예측모드가 DC 이외의 값을 가지는 경우 ChromaIntra8×8의 예측모드는 Intra16×16에서 선택된 예측모드만을 추가적으로 고려한다. 따라서 제안하는 알고리듬은 ChromaIntra8×8의 예측모드의 수를 1 또는 2개로 한정한다.

3.4 연산량 분석

제안하는 알고리듬에서 고려하는 예측모드의 수는 가변적이며, 낮은 비트율 즉, 양자화 계수가 큰 경우일수록 더 적은 연산량을 가진다. 이는 Intra4×4의 경우 식 (5)에서 볼 수 있는 바와 같이 양자화 계수가 큰 경우 선택가능한 예측모드의 수가 더욱 제한되며, Intra16×16의 예측모드 또한 제한된 예측모드만을 사용하는 Intra4×4의 예측모드정보를 사용하여 추정함으로써 예측모드의 수가 제한되기 때문이다.

표 2. 고려한 예측모드 개수의 비교
Table 2. Comparison of the number of considered prediction modes

Intra-mode	JVT Model	Proposed
Intra4x4	9	2, 4, 5
Intra16x16	4	1, 2
ChromaIntra8x8	4	1, 2

표 2는 H.264/AVC 참조 소프트웨어에서의 방법과 제안하는 알고리듬에서의 고려하는 예측모드의 수를 비교하여 나타내고 있다. 따라서 H.264/AVC 참조 소프트웨어에서 하나의 매크로블록당 수행되는 율-왜곡 비용함수의 계산회수가 최대 592 (4·(9×16+4×1)) 회인데 반하여 제안하는 방법은 최소 33(1·(2×16+1×1))회의 율-왜곡 계산회수를 가지며, 최대 164(2·(5×16+2×1))회의 계산회수를 가진다.

Ⅳ. 실험 및 결과 분석

제안하는 알고리듬의 타당성을 보이기 위해서, ITU-T에서 권고하는 H.264/AVC 참조소프트웨어

(JM10.1[8])에 적용하여 실험하였으며, 실험환경은 다음과 같다.

- CODEC: JVT Model (JM10.1)
- Operating System: Window XP Professional
- CPU: Intel Pentimum-IV
- Memory: 1,024 MB
- Quantization Parameter: 20~30
- Rate-Distortion Optimization: used

다양한 환경에서 제안한 알고리듬의 성능평가를 위해서 화면내의 다양한 방향성 성분을 가지는 네 개의 QCIF(176×144) 크기의 테스트 영상(akiyo, carphone, claire, coastguard)과 네 개의 CIF(352×288) 크기의 테스트 영상(news, silent, stefan, table)을 사용하여 실험하였다. 각 테스트 영상의 최초 10 프레임을 인트라프레임으로 부호화하였으며, 이때 제안한 알고리듬의 성능비교를 위해서 Park이 제안한 방법[6] 및 Pan이 제안한 방법[7], 그리고 참조소프트웨어에서의 방법을 함께 비교하였다.

표3은 RDO를 사용한 경우와 사용하지 않는 경우에 대해서 28의 QP값을 사용하여 부호화 하였을 때 제안하는 방법을 비롯하여 참조소프트웨어에서의 방법, Park의 방법, 그리고 Pan의 방법에서의 객관적 화질을 나타내고 있다. RDO를 사용하지 않는 참조소프트웨어에서의 방법은 RDO를 사용한 고속 부호화 알고리듬들과 비교하여 높은 속도향상을 가져오지만 두 배 이상의 객관적 화질 열화가 발생함을 알 수 있다. 이 결과에서 볼 수 있듯이 제안하는 방법은 RDO의 사용여부와 관계없이 다른 방법과 비교하여 유사하거나 높은 객관적 화질을 가짐을 알 수 있다.

또한 표 4는 참조소프트웨어를 이용한 비트스트림의 비트율을 1로 하였을 때각 방법에서의 상대적인 비트율을 나타내는 것으로 제안하는 방법은 비트율면에서 높은 성능을 가지는 Pan의 방법과 거의 유사한 비트율을 가짐을 알 수 있다.

마지막의 표 5는 참조소프트웨어에서의 부호화 시간을 1로 하였을 때의 상대적인 부호화시간을 나타내는 것으로 이때 제안하는 방법에서의 부호화시간은 전처리과정을 포함한 전체 부호화시간을 나타낸다. 이 결과에서 제안하는 알고리듬은 RDO의 사용여부와 관계없이 다른 방법들과 비교하여 높은 속도향상을 가져옴을 알 수 있다.

V. 결 론

H.264/AVC 비디오 부호화 표준은 현존하는 비디오 부호화 표준들 중 가장 높은 부호화 효율을 나타내지만 부호기는 매우 높은 복잡도를 가진다. 본 논문에서는 이러한 복잡도를 줄이기 위한 하나의 방법으로서 고속 인트라모드 선택 방법을 제안하였다. 제안하는 고속 인트라모드 선택 방법은 먼저 Intra4x4의 예측모드를 한정된 예측모드 후보들 중에서 선택하고, 이 결과를 Intra16x16 및 ChromaIntra8x8의 예측모드 결정에 사용한다. 이는 각 인트라모드가 가지는 예측모드의 상관성을 이용한 것으로 제안하는 알고리듬을 적용하였을 때 미약한 화질 열화 및 비트율 증가를 보이면서도 RDO 기법이 사용되는 경우 평균 82%까지의 속도향상을 보이며, RDO가 사용되지 않는 경우에도 30%이상의 속도향상을 보인다. 따라서 만일 제안하는 알고리듬을 RDO를 사용하지 않은 경우에 대해서 적용한다면 RDO를 사용하지 않음으로서 단축되는 시간에 추가적인 속도향상 또한 가져올 수 있게 된다.

이러한 고속 부호화 방법은 방송편집 등에 사용되는 인트라프레임들의 고속부호화에 더욱 효율적으로 사용될 수 있을 것이라 기대한다.

표 3. 양자화 파라미터 28의 값에 대하여 RDO를 사용여부에 따른 객관적 화질(dB)의 비교
Table 3. Comparison of PSNR(dB) for QP=28 according to the use of RDO

size	sequences	RDOptimization = ON				RDOptimization = OFF			
		JVT	Park's	Pan's	Proposed	JVT	Park's	Pan's	Proposed
QCIF	akiyo	38.93	38.92	38.87	38.96	38.83	38.83	38.82	38.83
	carphone	38.13	38.06	38.08	38.07	37.96	37.97	37.95	37.97
	claire	40.82	40.70	40.78	40.76	40.70	40.62	40.66	40.66
	coastguard	35.50	35.44	35.45	35.44	35.35	35.30	35.33	35.33
CIF	news	38.95	38.91	38.92	38.92	38.80	38.74	38.79	38.78
	silent	36.36	36.32	36.33	36.32	36.24	36.19	36.22	36.22
	stefan	37.14	37.04	37.04	37.03	36.77	36.76	36.77	36.77
	table	35.20	35.10	35.09	35.07	35.01	34.97	34.98	34.97

표 4. 참조소프트웨어에서의 비트율을 1로 정규화하였을 때의 상대적 비트율의 비교
Table 4. Comparison of relative bitrates when bitrate from reference software is normalized

size	sequences	RDOptimization = ON				RDOptimization = OFF			
		JVT	Park's	Pan's	Proposed	JVT	Park's	Pan's	Proposed
QCIF	akiyo	1.000	1.074	1.014	1.029	1.000	1.124	1.011	1.016
	carphone	1.000	1.080	1.008	1.025	1.000	1.179	1.003	1.011
	claire	1.000	1.103	1.025	1.044	1.000	1.049	1.008	1.015
	coastguard	1.000	1.019	1.006	1.016	1.000	1.104	1.003	1.008
CIF	news	1.000	1.056	1.020	1.035	1.000	1.127	1.015	1.022
	silent	1.000	1.051	1.015	1.029	1.000	1.104	1.010	1.017
	stefan	1.000	1.040	1.015	1.024	1.000	1.131	1.011	1.014
	table	1.000	1.012	1.005	1.009	1.000	1.028	1.000	1.001

표 5. 양자화 파라미터 28의 값에 대하여 RDO를 사용여부에 따른 전체 부호화 시간 (sec)의 비교
Table 5. Comparison of total encoding time(sec) for QP=28 according to the use of RDO

size	sequences	RDOptimization = ON				RDOptimization = OFF			
		JVT	Park's	Pan's	Proposed	JVT	Park's	Pan's	Proposed
QCIF	akiyo	1.000	0.224	0.390	0.201	1.000	0.714	0.732	0.668
	carphone	1.000	0.213	0.369	0.180	1.000	0.710	0.704	0.655
	claire	1.000	0.189	0.336	0.156	1.000	0.734	0.759	0.714
	coastguard	1.000	0.185	0.340	0.167	1.000	0.736	0.736	0.687
CIF	news	1.000	0.219	0.329	0.181	1.000	0.730	0.729	0.679
	silent	1.000	0.217	0.336	0.184	1.000	0.723	0.727	0.680
	stefan	1.000	0.213	0.332	0.189	1.000	0.731	0.735	0.690
	table	1.000	0.185	0.324	0.158	1.000	0.740	0.745	0.698

참 고 문 헌

[1] JVT G050r1, "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14496-10 AVC)," May 2003.

[2] Thomas Wiegand, Gary J. Sullivan, "Overview of the H.264/AVC Video Coding Standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, pp. 560-576, July 2003.

[3] Thomas Wiegand, Heiko Schwarz, Anthony Joch, Faouzi Kossentini, "Rate-Constrained Coder Control and Comparison of Video Coding Standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, pp. 688-703, July 2003.

[4] Peng Yin, Hye-Yeon Cheong Tourapis, Alexis Michael Tourapis, Jill Boyce, "Fast Mode Decision and Motion Estimation for JVT/H.264," *Image Processing*, 2003 International Conference on, vol. 3 pp. 853-856, Sept. 2003.

[5] Yu-Wen Huang, Bing-Yu Hsieh, Tu-Chih Whang, Shao-Yi Chien, Shyh-Yih Ma, Chun-Fu Shen, Liang-Gee Chen, "Analysis and Reduction of Reference Frames for Motion Estimation in MPEG-4 AVC/JVT/H.264," *Acoustics, Speech, and Signal Processing*, IEEE International Conference on, vol. 3 pp. 145-148, April 2003.

[6] Jun-Sung Park and Hyo Jung Song, "Selective Intra Prediction Mode Decision for H.264/AVC Encoders," *Trans. on Computing and Technology*, vol. 13 pp. 51-55, May 2006.

[7] Feng Pan, Xiao Lin, Susanto Rahardja, Keng Pang Lin, Z.G. Li, Dajun Wu, and Si Wu, "Fast Mode Decision Algorithm for Intraprediction in H.264/AVC Video Coding," *Trans. on Circuit and Systems for Video Technology*, vol. 15, no. 7 pp. 813-822, July 2005.

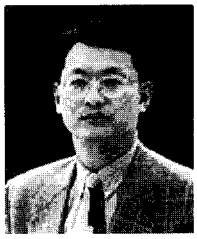
[9] JVT Model downloaded form http://bs.hhi.de/

suehring/tml/download/jm101.zip, accessed 2006.
〔10〕 문정미, 김재호, 문용호, "H.264/AVC 부호화기에 대한 효과적인 모드 결정 알고리즘" *한국 통신학회 논문지*, vol. 30, no. 3C, pp. 133-140, 2005.

김 동 형 (Donghyung Kim) 정회원

1999년 2월 충북대학교 전자공학과 졸업
2001년 8월 충북대학교 전자공학과 석사
2007년 2월 한양대학교 전자공학 박사
2007년 3월~현재 한양대학교 Post Doc.
<관심분야> 영상처리 및 영상압축

정 제 창 (Jechang Jeong) 종신회원



1980년 2월 서울대학교 전자 공학과 졸업
1982년 2월 KAIST 전기전자 공학과 석사
1990년 미국 미시간대학 전기공학과 공학박사
1980~1986 KBS 기술연구소 연구원 (디지털 TV 및 뉴미디어 연구)
1990~1991 미국 미시간대학 전기공학과 연구교수 (영상 및 신호처리 연구)
1991~1995 삼성전자 멀티미디어 연구소 (MPEG, HDTV, 멀티미디어 연구)
1995~현재 한양대학교 전자전기컴퓨터공학부 교수 (영상통신 및 신호처리 연구실)
1998년 11월 27일 과학기술자상 수상
1998년 12월 31일 정보통신부장관상 표창
<관심분야> 영상처리 및 영상압축