

## H.264의 고속 인트라 예측 알고리즘

조량석, 원치선  
동국대학교 전자공학과

### Fast Algorithm for Intra Prediction in H.264

Lyang Seok Jo, Chee Sun Won  
Dept. of Electronics, Dongguk University  
E-mail : jolyang@dongguk.edu

#### 요약문

최근에 표준화된 H.264의 인트라 예측은 I-프레임을 압축하는데 사용된다. 최대의 코딩효율을 위해서 부호화기는 모든 인트라 예측 모드를 반복적으로 찾는다. 그러나 이것은 H.264의 부호화기 복잡도를 매우 크게 하는 단점이 있다. 모든 경우의 모드중에 영상에 따라서는 특히 그 발생 가능성이 높거나 낮은 경우의 모드가 발생한다. 본 논문에서는 예지맵을 이용하여 H.264의 인트라 예측 모드를 선별적으로 선택하는 방법을 제안한다.

#### I. 서론

H.264는 ISO/IEC의 MPEG과 ITU-T의 VCEG 두 그룹이 공동 연구 기관 JVT(Joint Video Team)을 창설하여 새롭게 제안한 동영상 압축에 관한 국제 표준이다. H.264 동영상 부호화[1]의 기본 개념은 기존의 동영상 압축 알고리즘과 비슷하지만, 우수한 화질을 얻기 위해 이전과는 다른 몇 가지 특징들이 있으며, 그것은 다음과 같다.

첫째, 범용 가변장 부호화(UVLC : universal variable

length coding)라는 하나의 가변 길이 부호화 테이블을 사용한다. 기존의 H.263과 같은 표준에서는 부호화하는 과정에 따라서 서로 다른 가변 길이 부호화 테이블을 제공하였으나, H.264는 부호화 과정에 상관없이 하나의 가변 길이 부호화 테이블만을 사용한다.

둘째로, 1/4 화소 움직임 정확도(1/4-pel motion accuracy)로 움직임 보상을 수행한다. 기존의 표준들이 정수 또는 반 화소 단위의 움직임 정확도로 움직임 보상을 하는데 비하여 H.264는 더 정확한 해상도로 움직임 보상을 하기 때문에 보다 높은 화질의 예측 영상을 얻을 수 있다. 그리고 선택적으로 1/8 화소 움직임 정확도까지 움직임 보상이 가능하다.

셋째로, 변환 블록 크기(variable block size)에 의해서 부호화를 수행한다. 이전의 표준에서는 주로 16x16 크기로 고정된 매크로 블록(MB : macroblock)을 사용하여 부호화를 한 반면에, H.264에서는 16x16 크기의 기존의 MB를 여러 개의 블록 크기로 세분화 한후, 각각에 대해서 부호화를 한다. 그리고 가장 높은 성능을 보여주는 블록 크기를 비용함수를 통하여 계산한 후 최적의 MB 크기로 결정한다.

넷째로, 기존의 DCT와 비슷하지만 차이점이 있는 4x4 정수 변환 부호화(4x4 integer transform based coding) 방법의 의해서 데이터를 변환한다. 이는 H.264의 블록

의 크기 중 가장 작은 것이 4x4 이기 때문이며, 또한 정수 변환을 통하여 변환 과정의 속도 향상을 얻는 것과 실수 연산의 시스템에 대한 의존성을 제거하는 것을 목표로 하고 있다.

다섯번째로, 다중 참조 프레임의 사용이다. 기존의 표준은 k 번째 프레임  $F_k$  을 부호화하기 위해서  $F_{k-1}$  만을 참조하였다. 하지만 H.264 에서는 최대 5 프레임 즉,  $F_{k-1}$ ,  $F_{k-2}$ ,  $F_{k-3}$ ,  $F_{k-4}$ ,  $F_{k-5}$  의 정보를 이용해서 움직임을 보상이 가능하다. 참조되는 영상의 정보가 기존의 방법에 비해서 많기 때문에 더 좋은 부호화 효율을 기대할 수 있다. 하지만 복호기에서 기억 용량을 많이 필요로 하는 단점이 있다.

마지막으로, 공간영역에서의 인트라 예측이다. 인트라 예측을 공간영역에서 다양한 예측 모드를 조합하여 최소 오차를 갖는 모드를 예측 모드로 사용함으로써 부호화 효율을 높이고 있다. 하지만 최소 오차인 모드를 찾기 위해 블록크기와 그에 따르는 예측모드들의 모든 조합에 대해 반복적인 탐색으로 부호화 복잡도를 높이는 문제가 있다.

본 논문에서는 인트라 예측시 부호화기의 복잡도 문제를 개선하기 위해 예지법을 이용하여 인트라 예측 모드를 선별적으로 선택하는 방법을 제안한다.

## II. H.264 의 인트라 예측

매크로블록이 인트라 모드로 부호화될 때, 예측블록은 이전 부호화되어 재구성된 블록으로부터 예측된다. 이후 예측된 블록과 현재블록의 차분을 부호화 하게 된다.

<표 1>~<표 3>에 나타난 대로, 밝기 성분은 4x4 부블록 또는 16x16 매크로블록 단위로 예측을 수행하며, 4x4 밝기 블록에서는 9 가지, 16x16 밝기 블록에서는 4 가지의 예측모드가 있다. 색차성분은 8x8 블록에 대해 4 가지의 예측모드가 있다.

<표 1> Intra 4x4 luma prediction mode

Intra4x4 PredMode	Name of Intra4x4PredMode
0	Intra_4x4_Verical (prediction mode)

1	Intra_4x4_Horizontal (prediction mode)
2	Intra_4x4_DC (prediction mode)
3	Intra_4x4_Diagonal_Down_Left (prediction mode)
4	Intra_4x4_Diagonal_Down_Right (prediction mode)
5	Intra_4x4_Verical_Right (prediction mode)
6	Intra_4x4_Horizontal_Down (prediction mode)
7	Intra_4x4_Verical_Left (prediction mode)
8	Intra_4x4_Horizontal_Up (prediction mode)

<표 2> Intra 16x16 luma prediction mode

Intra16x16 PredMode	Name of Intra16x16PredMode
0	Intra_16x16_Verical (prediction mode)
1	Intra_16x16_Horizontal (prediction mode)
2	Intra_16x16_DC (prediction mode)
3	Intra_16x16_Plane (prediction mode)

<표 3> Intra 8x8 chroma prediction mode

intra_chroma pred_mode	Name of intra_chroma_pred_mode
0	Intra_Chroma_DC (prediction mode)
1	Intra_Chroma_Horizontal (prediction mode)
2	Intra_Chroma_Verical (prediction mode)
3	Intra_Chroma_Plane (prediction mode)

하나의 매크로블록에서, 4x4 인트라 예측에서 수행되는 조합의 경우의 수는 (모드수 9 × 부블록수 16) = 144 가지가 되며, 16x16 인트라 예측에서 수행되는 조합의 경우의 수는 (모드수 4) = 4 가지, 8x8 인트라 예측에서는 (모드수 4) = 4 가지가 된다. 그러므로, 모든 경우의 수는 (144 + 4) × 4 = 592 가지가 된다. 그러나 실제로 주어진 영상에 따라서는 발생가능성이 특히 높거나 낮은 모드가 존재하며, 이런 경우 발생 가능성이 높은 경우의 모드들에 대해서만 최적 모드를 탐색하여 수행시간을 감축 시킬 수 있다.

### III. 고속 인트라 예측 모드 선택

본 장에서는 앞서 설명한 인트라 예측의 계산량을 줄일수 있는 방법을 제안한다.

영상을 에지 검출기[3]를 통해 에지의 크기  $E(x,y)$ 와 방향  $D(x,y)$ 을 다음과 같이 구한다.

$$E(x,y) = \sqrt{G_x^2 + G_y^2} \quad (1)$$

$$D(x,y) = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (2)$$

여기서,  $G_x$  는 픽셀위치  $(x,y)$ 에서의 수평방향 그래디언트,  $G_y$  는  $(x,y)$ 에서의 수직방향 그래디언트이다.

예측 블록에 따라 식(2)의 값을 모드의 방향으로 양자화하고, 크기에 대한 에지 방향 히스토그램  $H(m)$ 을 구한다. 즉, 그림 1의 8 가지 방향중에  $m$  번째 방향에 해당되는 픽셀의 해당 에지크기의 누적값이 히스토그램의 빈  $H(m)$ 이 된다.

좀더 자세히 9 개의 히스토그램 빈의 값은 아래의 식으로부터 얻을 수 있다.

$$H(0) = \sum_{\substack{(x,y): \\ -\frac{9}{16}\pi < D(x,y) < -\frac{7}{16}\pi}} E(x,y)$$

$$H(1) = \sum_{\substack{(x,y): \\ -\frac{\pi}{16} < D(x,y) < \frac{\pi}{16}}} E(x,y)$$

$$H(3) = \sum_{\substack{(x,y): \\ \frac{13}{16}\pi < D(x,y) < \frac{11}{16}\pi}} E(x,y)$$

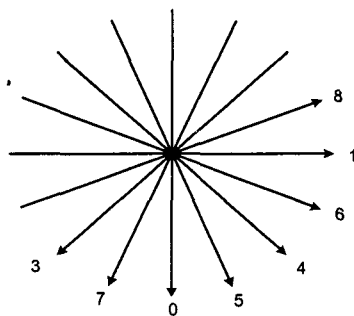
$$H(4) = \sum_{\substack{(x,y): \\ -\frac{5}{16}\pi < D(x,y) < -\frac{3}{16}\pi}} E(x,y)$$

$$H(5) = \sum_{\substack{(x,y): \\ -\frac{7}{16}\pi < D(x,y) < -\frac{5}{16}\pi}} E(x,y)$$

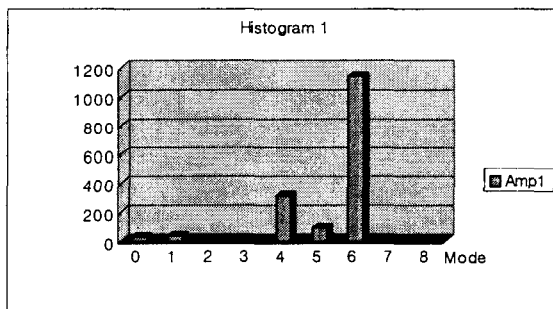
$$H(6) = \sum_{\substack{(x,y): \\ -\frac{3}{16}\pi < D(x,y) < -\frac{\pi}{16}}} E(x,y)$$

$$H(7) = \sum_{\substack{(x,y): \\ -\frac{11}{16}\pi < D(x,y) < -\frac{9}{16}\pi}} E(x,y)$$

$$H(8) = \sum_{\substack{(x,y): \\ \frac{\pi}{16} < D(x,y) < \frac{3}{16}\pi}} E(x,y)$$



<그림 1> 인트라 4x4 예측 모드 방향



<그림 2> 에지 방향 히스토그램

그림 2 는 4x4 밝기 예측에 대한 예로서, 히스토그램의 빈은 예측모드, 세로는 해당 방향성을 가진 에지 크기의 누적합이다.

여기서 예측 모드의 후보를 히스토그램의 크기가 가장 큰 빈을 해당 블록에서 가장 우세한 모드 후보로 선택하고, 그 이웃하는 방향의 모드와 DC 모드를 후보로 선택한다. 위의 그림에서 후보 모드는 모드 6, 모드 1, 모드 4 와 DC 모드가 된다.

16x16 밝기 예측과 8x8 색차 예측의 경우는 4 가지 모드중, 히스토그램의 크기가 가장 큰 모드와 DC 모드를 후보로 선택한다.

H.264 의 인트라 예측 모드의 매크로블록당 모든 조합의 경우의 수는  $(9 \times 16 + 4) \times 4 = 592$  가지이며, 제안

한 방법의 조합의 경우의 수는  $(4 \times 16 + 2) \times 2 = 132$  가지가 된다.

#### IV. 실험결과

실험은 H.264 의 참조 소프트웨어 JM6.1[2]을 가지고 펜티엄 3-700Mhz 프로세서에서 320MB 메모리를 가진 시스템과 OS 는 윈도우 2000 하에서 CIF(352x288)과 QCIF(176x144) 영상에 대하여 100 프레임에 대하여 수행하였다. 결과치는 100 프레임에 대한 프레임당 평균치이며 부호화 시간 변화율에서는 제안된 방법의 예지 방향 히스토그램을 계산하는 시간을 포함하였으며 한장의 프레임을 부호화하는 시간을 측정하였다. 부호화기 파라미터는 다음과 같다.

Hadamard Transform ON  
 GOP IPP  
 Search Range  $\pm 32$   
 Reference Frames 1  
 Entropy Coding CABAC  
 QP 28

<표 4> Foreman QCIF

	I 프레임만	전체 시퀀스
비트량 변화율	2.28%	2.10%
프레임당 부호화 시간 변화율	-54.86%	-21.51%
SnrY (dB)	-0.0358	-0.0111

<표 5> Mobile CIF

	I 프레임만	전체 시퀀스
비트량 변화율	1.35%	1.10%
프레임당 부호화 시간 변화율	-54.96%	-27.61%
SnrY (dB)	-0.1186	-0.05

위의 결과로부터 QCIF 영상에서 I 프레임에 대한 비트량은 약 2.3%정도 증가한데 비해, 연산시간은 50% 이상 줄어들었으며, CIF 영상에서는 1.35%정도의 비트

량증가에 비해, 연산시간은 50%이상 줄어 들었음을 확인할수 있다. 특히 영상 크기가 커질수록 제안된 방법은 더욱 좋은 성능을 보인다.

한편 비트량의 변화가 중요시 되는 응용에서는, 가장 빈도가 낮은 모드를 후보에서 제외하는 방법-표 6-을 사용했을 때 CIF 영상에서 0.4%의 비트량 증가에 대해 연산속도는 20%정도 감소함을 확인하였다.

<표 6> 최소 빈도 모드 제외 방법

	QCIF Foreman I 프레임만	CIF Mobile I 프레임만
비트량 변화율	0.72%	0.43%
프레임당 부호화 시간 변화율	-20.17%	-17.94%
SnrY (dB)	0.01593	-0.0126

#### V. 결론

본 논문에서는 인트라 예측을 빠르게 수행할수 있는 모드 선택 방법을 예지 맵을 이용하여 제안하였다. 주요한 예지의 방향성을 참고하여 예측 모드 후보수를 줄임으로서 인트라 예측의 연산량을 크게 줄일수 있다. 또한 결과에서 알수 있듯이 비트량이나 화질은 거의 유지되고 있음을 알수 있다.

#### 참고문헌

- [1] ISO/IEC JTC1/SC 29/WG11/N5555 "FDIS 14496-10 JVT," Pattaya, March 2003
- [2] JM6.1, Reference Software of JVT
- [3] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing," Prentice Hall, 2001
- [4] Anil K. Jain, "Fundamentals of Digital Image Processing," Prentice Hall, 1989
- [5] Bojun Meng, Oscar C. Au, "Fast Intra-prediction Mode Selection for 4x4 Block in H.264," ICASSP, April, 2003