

주파수 특성의 제약 조건들을 이용한 H.264/AVC를 위한 고속 화면 내 모드 선택 방법

정회원 진순중*, 준회원 박상준, 정회원 정제창

Fast Intra Mode Selection Algorithm for H.264/AVC Using Constraints of Frequency Characteristics

Soonjong Jin* *Regular Member*, Sang-Jun Park *Associate Member*,
Jechang Jeong *Regular Member*

요 약

H.264/AVC 비디오 부호화 표준은 MPEG-2, H.263 또는 MPEG-4와 같은 기존의 비디오 부호화 표준과 비교하여 상당한 부호화 효율의 향상을 제공한다. 부호화 효율의 향상을 수행하기 위해서 H.264/AVC는 매크로블록 단위로 최적의 움직임 벡터, 참조 화면, 매크로블록 모드를 선택하는데 있어서 율-왜곡 최적화 (Rate-distortion optimization) 기법을 이용한다. 결과적으로 높은 부호화 효율을 제공함과 동시에 복잡도 역시 상당히 증가하였다. 본 논문에서는 화면 내 부호화 모드들에서 나타나는 주파수 특성의 제약 조건들을 이용하여 고속 화면 내 모드 선택 방법을 제안한다. 먼저 영상의 주파수 해석을 통해 H.264/AVC의 화면 내 모드 별 특성을 살펴본다. 다음 제안하는 주파수 에러 비용 (Frequency Error Cost, FEC)을 계산하여 최소 비용을 가지는 모드와 그 주변 모드 2개만을 후보 모드로 선택하고 후보 모드의 율-왜곡 비용만을 계산하여 최적의 화면 내 모드를 결정하여 복잡도를 감소시킨다. 실험 결과는 제안하는 방법이 H.264/AVC와 비교하여 비슷한 율-왜곡 성능을 보이면서 복잡도를 상당히 감소시킴을 보여준다.

Key Words : Intra Prediction, Fast Mode Selection, Frequency Analysis, H.264/AVC

ABSTRACT

H.264/AVC video coding standard enables a considerably higher improvement in coding efficiency compared with previous standards such as MPEG-2, H.263 and MPEG-4. To achieve this, for each macro-block in H.264/AVC, Rate-Distortion Optimization (RDO) technique is employed to select the best motion vector, reference frame, and macro-block mode. As a result, computational complexity is increased significantly whereas RDO achieve higher improvement. This paper presents fast intra mode selection algorithm based on constraints of frequency characteristics which are derived from intra coding modes of H.264/AVC. First of all, we observe the features of each intra mode through the frequency analysis of image. And then proposed Frequency Error Costs (FECs) are calculated to select the best mode which has minimum cost. Computational complexity is considerably reduced because rate-distortion costs only calculate the candidate modes which are set of best mode and its neighbouring two modes. Experimental results show that proposed algorithm reduces the complexity dramatically maintaining the rate-distortion performance compared with H.264/AVC reference software.

※ 본 연구는 서울시 산학연 협력 사업으로 구축된 서울 미래형 콘텐츠 컨버전스 클러스터 지원으로 수행되었습니다.

* 한양대학교 전자컴퓨터통신공학과 영상통신및신호처리 연구실 (sjin@hanyang.ac.kr)

논문번호 : KICS2007-11-523, 접수일자 : 2007년 11월 20일, 최종논문접수일자 : 2008년 4월 16일

I. 서 론

새로운 차세대 비디오 압축 표준이 ITU-T의 VCEG(Video Coding Expert Group)과 ISO/IEC의 MPEG(Moving Picture Expert Group)이 함께 구성한 JVT(Joint Video Team)에 의해 제정되었다^[1]. 이 표준은 ISO/IEC와 ITU-T에서 각각 사용되는 명칭인 MPEG-4 Part 10 Advanced Video Coding (AVC)와 ITU-T 권고안 H.264를 근거로 H.264/AVC라고 불린다. H.264/AVC 비디오 부호화 방식은 이전의 비디오 부호화 표준들(MPEG-2, H.263, MPEG-4)에는 없던 새로운 부호화 도구들이 추가 되었으며 이러한 도구들은 H.264 부호화 표준이 이전의 비디오 부호화 표준과 비교하여 보다 높은 압축 효율을 갖게 하였다. 이러한 높은 압축 효율로 인해 H.264/AVC는 저장 매체, 모바일 화상 통신, 인터넷, 휴대 이동 방송, HD-DVD 등 여러 산업의 영역에서 이용되고 있다^[2].

새롭게 추가된 부호화 도구들에는 화면 내 예측 부호화, 4x4 단위의 정수 변환, 향상된 엔트로피 부호화, 향상된 루프 필터, 다중 참조 영상의 사용, 다양한 블록 크기 등이 있다. 이러한 새롭게 추가된 부호화 도구들은 H.264 부호화 표준이 이전의 여러 동영상 부호화 표준들에 비해 보다 높은 압축 효율을 가질 수 있게 한다. 다섯 장의 참조 프레임의 사용과 32 화소의 움직임 추정 영역을 사용하여 메인 프로파일로 부호화 한 경우, H.264/AVC 부호화 표준은 MPEG-2와 비교하여 저 비트율에서 약 45~65%, 고 비트율에서 약 25~45% 까지 비트율 감소 효과가 있다고 알려져 있다^[3]. 하지만 비트율 감소 효과와 더불어 부호기의 복잡도는 상당히 높아진다. 특히 새로이 추가된 부호화 도구들 중 매크로블록 모드 선택 과정은 높은 복잡도를 가지는데 이는 선택 가능한 모든 예측 모드에서의 윌-왜곡 비용을 계산하여야 하기 때문이다.

H.264/AVC의 매크로블록 모드 선택 과정에 있어서 복잡도를 감소시키기 위한 많은 연구들이 진행되어 왔다. 특히 기존에도 존재했던 주제인 움직임 예측/추정에 관한 내용뿐만 아니라 기존에 존재하지 않아서 H.264/AVC에서 새로운 분야가 된 고속 화면 내 모드 결정 알고리즘들에 대한 연구도 진행되었다. Park 등이 제안한 고속 화면 내 모드 선택 알고리즘은 휘도 (Luminance) 블록의 화면 내 모드 중 4x4 블록 단위로 예측을 수행하는 LumaIntra4x4 모드의 예측 모드를 이미 계산된 16x16블록 단위로 예측을

수행하는 LumaIntra16x16 모드를 사용하여 추정함으로써 연산량을 감소시키며, 색차 (Chrominance) 블록의 예측 모드인 ChromaIntra8x8 모드 또한 계산된 LumaIntra16x16의 예측 모드를 사용하여 추정하므로 연산량을 감소시킨다^[4]. Pan 등이 제안한 고속 화면 내 모드 선택 방법은 소벨 연산을 이용하여 에지 맵을 산출하고 이를 토대로 각 예측 모드에서의 선택 가능한 예측 모드를 한정함으로써 화면 내 모드 선택 과정에서의 복잡도를 감소시킨다^[5].

Park의 알고리즘은 LumaIntra16x16모드를 이용하는 간단한 전처리 과정으로 LumaIntra4x4모드 선택에 제한을 두지만 LumaIntra4x4 모드 선택에서 영상의 지역적 특성을 반영하지 못한다는 단점이 있다. 즉 LumaIntra4x4모드를 LumaIntra16x16의 최종 모드로만 예측하므로 세밀한 지역 특성을 고려할 수 없고 이로 인해 속도 향상 대비 윌 왜곡 성능이 떨어진다. 또한 pan의 알고리즘은 각각의 픽셀에서 에지를 구하고 히스토그램을 구하는 복잡한 전처리 과정으로 인해 속도 향상이 높지 않다는 단점이 있다. 따라서 영상의 지역 특성을 정확하게 분석하고 모드를 제한 할 수 있는 전처리 과정을 통해 윌-왜곡 성능을 유지하면서 속도 향상을 가져올 수 있는 고속 화면 내 예측 방법이 필요하다.

본 논문에서는 부호화할 4x4 영상의 간단한 주파수 분석을 통해 각 LumaIntra4x4 모드별로 나타나는 제약 조건을 수학적으로 유도한다. 제안된 FEC를 통해 최소 비용을 가지는 모드를 현재 4x4영상의 최적 모드로 정하고 (primary predicted mode, PPM) 이 PPM과 DC mode, most probable mode, PPM 양쪽에 위치한 모드들을 후보 모드 집합 (candidate mode set, CMS) 이라 정한다. 이 CMS에 포함된 모드만의 윌-왜곡 비용을 계산하고 이 중에서 최종 모드를 선택한다. 또한 앞에서 구한 지역 영상의 모드별 FEC 값을 이용하여 LumaIntra16x16 모드별 복잡도를 계산해 내고 LumaIntra16x16 모드의 후보 모드를 제한한다. 마지막으로 LumaIntra16x16의 최종 모드를 통해 ChromaIntra8x8모드를 예측하므로 복잡도를 감소시킨다. 이러한 과정을 통해서 각 4x4 영상의 지역적 특성을 반영하므로 정확한 4x4 크기의 화면 내 예측을 수행할 수 있고 각각의 특성을 16x16 화면 내 예측에 이용하므로 정확한 16x16 예측을 수행할 수 있다.

본 논문의 구성은 다음과 같다. II장에서는 H.264/AVC의 화면 내 예측 모드들에 대해서 간단히 설명한다. III장에서는 본 논문의 주 아이디어인 주파수 분석을 통한 화면 내 모드 별 AC/DC 계수

제약 조건을 이용한 FEC를 유도한다. IV 장에서 제안하는 화면 내 모드 선택 방법을 설명하고 V장의 실험 결과에서 알고리즘의 속도향상과 그에 따른 율-왜곡 성능을 보인다. 마지막으로 VI장에서 결론을 맺는다.

II. H.264/AVC의 화면 내 예측

일반적으로 H.264/AVC baseline 프로파일을 이용하여 부호화 할 경우 모든 매크로블록들은 휘도 블록의 부호화를 위한 두 개의 화면 내 모드 즉, LumaIntra16x16 및 LumaIntra4x4 모드와 색차 블록의 부호화를 위한 하나의 화면 내 모드 즉, ChromaIntra8x8을 사용하여 부호화된다. 각 화면 내 모드는 서로 다른 예측 모드의 수를 가지는데, 그림 1은 LumaIntra4x4 모드에서 사용하는 9가지의 예측 모드를 나타내며, 그림 2는 LumaIntra4x4 예측 모드의 공간적 방향을 나타낸다.

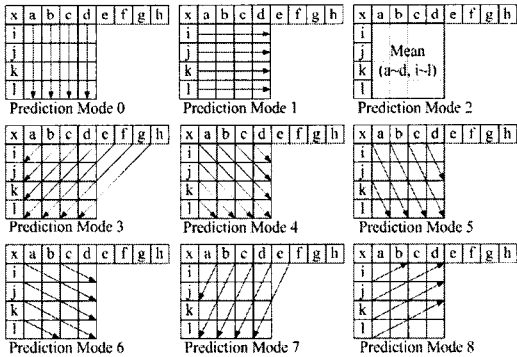


그림 1. 9가지 LumaIntra 4x4 예측 모드 방향

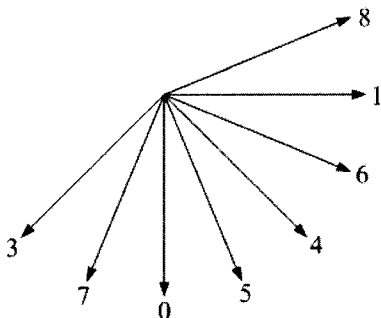


그림 2. LumaIntra4x4 예측 모드들의 공간적 방향

LumaIntra4x4 에서 예측 가능한 모드들은 다음과 같다.

- ◆ 모드 0: Vertical prediction
- ◆ 모드 1: Horizontal prediction
- ◆ 모드 2: DC prediction
- ◆ 모드 3: Diagonal Down-left prediction
- ◆ 모드 4: Diagonal Down-right prediction
- ◆ 모드 5: Vertical-right prediction
- ◆ 모드 6: Horizontal-down prediction
- ◆ 모드 7: Vertical-left prediction
- ◆ 모드 8: Horizontal-up prediction

즉 LumaIntra4x4 모드는 모드 2를 제외하고 그림 2에 나타낸 8가지 방향에 맞게 4x4 단위로 화면 내 예측을 수행한다. 위에 언급한 4x4 단위로 예측하는 LumaIntra4x4 예측과 비슷하게 매크로블록 전체가 16x16 단위로 예측이 될 수도 있다. 그림 3은 LumaIntra16x16 모드에서 사용되는 4가지의 예측 모드를 나타낸다.

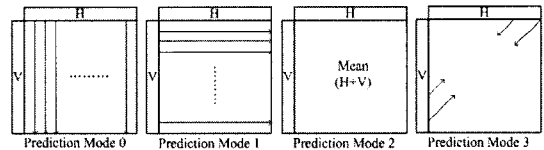


그림 3. 4가지 LumaIntra16x16 예측 모드 방향

LumaIntra16x16에서 예측 가능한 모드들은 다음과 같다.

- ◆ 모드 0: Vertical prediction
- ◆ 모드 1: Horizontal prediction
- ◆ 모드 2: DC prediction
- ◆ 모드 3: Plane prediction

마지막으로 ChromaIntra8x8의 예측 모드는 LumaIntra16x16과 동일하나 색차 성분의 통계적 발생빈도에 따라 그 순서만을 달리한다^[1]. ChromaIntra8x8에서의 모드 순서는 다음과 같다.

- ◆ 모드 0: DC prediction
- ◆ 모드 1: Horizontal prediction
- ◆ 모드 2: Vertical prediction
- ◆ 모드 3: Plane prediction

율-왜곡 최적화 기법이 사용되는 경우 H.264/AVC는 색차 블록의 각 예측 모드에 대해서 LumaIntra16x16

과 LumaIntra4×4의 울-왜곡 비용을 비교하여 최소의 비용을 갖는 모드를 해당 매크로블록의 휘도 및 색차 블록의 화면 내 모드로 선택한다. 이 과정에서 LumaIntra4×4는 각 4×4 블록마다 울-왜곡 비용을 계산하며 매크로블록 내에 포함되어 있는 16개의 모든 4×4 블록에서의 울-왜곡 비용의 합을 LumaIntra16×16의 울-왜곡 비용과 비교한다. 울-왜곡 비용은 식 (1)과 같이 계산될 수 있는데, 여기에서 왜곡 (distortion) 값은 각 매크로블록 모드에서의 SNR을 구함으로써 계산할 수 있고 울 (rate) 값은 매크로블록의 부호화가 끝나는 시점에서의 부호화된 비트 스트림의 비트율을 나타낸다.

$$RDCost = Distortion + \lambda \times Rate \quad (1)$$

일반적으로 화면 내 프레임 안에 포함되어 있는 하나의 매크로블록에 대한 최적 모드를 결정하기 위해서는 ChromaIntra8×8의 4가지의 예측 모드에 대해서 LumaIntra16×16 및 LumaIntra4×4의 4가지 및 9가지의 모든 조합 가능한 예측에 대한 울-왜곡 비용 계산 과정이 필요하다. 이는 하나의 매크로블록에 대해서 최적 모드 선택을 위해서는 최대 $4 \times (1 \times 4 + 16 \times 9) = 592$ 회의 울-왜곡 비용 계산 과정이 필요함을 의미한다.

III. 주파수 분석 유도 과정

본 논문에서 $t(j,i)$ 와 $\hat{t}(j,i)$ 는 각각 원본 영상과 H.264 화면 내 예측에 의해서 예측된 영상을 나타낸다. $e(j,i)$ 는 에러 영상 즉 $t(j,i) - \hat{t}(j,i)$ 를 나타내며 $a_{k,l}(j,i)$ 는 주파수 분석을 위해 사용된 변환 커널을 나타낸다.

3.1 주파수 분석

먼저 저주파수부터 고주파수까지 정렬된 하다마드 변환 커널을 식 2에 나타내었다.

$$a_{k,l}(j,i) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \quad (2)$$

식 (2)에 나타난 정렬된 하다마드 변환 커널은 뒤어

난 주파수 분석 성능을 보여주며 덧셈과 뺄셈만으로 구성되어 복잡도가 상당히 적다는 장점이 있다. 만약 완전 예측 조건 즉 $e(j,i) = t(j,i) - \hat{t}(j,i) = 0$ 인 상황을 가정한다면 $t(j,i) = \hat{t}(j,i)$ 이고 따라서 $T(k,l) = \hat{T}(k,l)$ 과 같다. 이것은 완전 예측 조건일 경우 예측된 영상의 주파수 특성은 원본 영상의 특성과 거의 동일하다는 것을 의미한다. 따라서 (2)를 이용해서 원본 영상의 주파수 계수들을 구하고 그 특성을 적절히 잘 이용하면 정확한 예측을 수행할 수 있다.

3.2 AC 계수 제약 조건

LumaIntra4x4의 모드 0에 의해 예측된 \hat{t}_{M0} 은 (3)과 같고 \hat{t}_{M0} 에 대한 \hat{T}_{M0} 는 식 (4)와 같다.

$$\hat{t}_{M0} = \begin{bmatrix} a & b & c & d \\ a & b & c & d \\ a & b & c & d \\ a & b & c & d \end{bmatrix} \quad (3)$$

$$\hat{T}_{M0} = 4 \begin{bmatrix} a+b+c+d & a+b-c-d & a-b-c+d & a-b+c-d \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4)$$

즉 $\sum_{j=1}^3 \sum_{i=0}^3 |\hat{T}_{M0}(j,i)| = 0$ 이다. 완전 예측 조건이라 가정하면 $T(k,l) = \hat{T}(k,l)$ 과 같으므로 모드 0에 대한 AC 계수 제약 조건은 식 (5)와 같이 구할 수 있다.

$$\Delta AC_{M0} = \frac{1}{12} \sum_{j=1}^3 \sum_{i=0}^3 |T(j,i)| \quad (5)$$

이와 마찬가지로 \hat{T}_{M1} 과 \hat{T}_{M2} 를 구하면 식 (6)과 (7)의 같은 완전 예측 조건을 얻는다.

$$\sum_{j=0}^3 \sum_{i=1}^3 |\hat{T}_{M1}(j,i)| = 0 \quad (6)$$

$$\sum_{j=0}^3 \sum_{i=0}^3 |\hat{T}_{M2}(j,i)| = 0, \quad \neq i \text{ and } j = 0 \quad (7)$$

따라서 LumaIntra4x4의 모드 1과 모드 2에 대한 AC 계수 제약 조건들은 식 8과 9와 같이 나타낼 수 있다.

$$\Delta AC_{M1} = \frac{1}{12} \sum_{j=0}^3 \sum_{i=1}^3 |T(j,i)| \quad (8)$$

$$\Delta AC_{M2} = \frac{1}{15} \sum_{j=0}^3 \sum_{i=0}^3 |T(j,i)|, \quad \epsilon i \text{ and } j = 0 \quad (9)$$

모드 3과 모드 4는 수직 선(Vertical line)에 대해서 예측 방향이 대칭이며 또한 서로 직교한다. 이는 특정 AC 계수들의 위치에서 기 함수 대칭 특성을 보인다. 즉 \hat{T}_{M3} 이 특정 위치 (j, i)와 (k, l)위치에서 식 10과 같은 특성을 가질 때 \hat{T}_{M4} 는 식 11과 같은 특성을 가진다.

$$\hat{T}_{M3}((j,i)|m) = \hat{T}_{M3}((k,l)|m) \quad (10)$$

$$\hat{T}_{M4}((j,i)|m) = -\hat{T}_{M4}((k,l)|m) \quad (11)$$

식 10과 11에서 각각 특정 위치 (j, i)와 (k, l)은 index m에 따라서 정해지는데 이 위치는 표 1에 나타내었으며 그림 4와 같다.

표 1. Index m에 따른 특정 위치 (j,i)와 (k,l)의 위치

m	(j,i)	(k,l)
0	(0,1)	(1,0)
1	(0,2)	(1,3)
2	(2,0)	(3,1)
3	(1,2)	(2,1)
4	(0,3)	(3,0)
5	(2,3)	(3,2)

T(0,0)	T(0,1)	T(0,2)	T(0,3)
T(1,0)	T(1,1)	T(1,2)	T(1,3)
T(2,0)	T(2,1)	T(2,2)	T(2,3)
T(3,0)	T(3,1)	T(3,2)	T(3,3)

그림 4. 모드 3과 4에 해당하는 특정 위치

따라서 모드 3과 모드 4에 대한 AC 계수 제약 조건들은 식 12와 13과 같이 나타낼 수 있다.

$$\Delta AC_{M3} = \frac{1}{6} \sum_{m=0}^5 |T((j,i)|m) - T((k,l)|m)| \quad (12)$$

$$\Delta AC_{M4} = \frac{1}{6} \sum_{m=0}^5 |T((j,i)|m) + T((k,l)|m)| \quad (13)$$

모드 5는 모드 7과, 모드 6은 모드 8과 각각 예측 방향이 수직 선과 수평 선 (Horizontal line)에 각각 대칭이며 따라서 특정 AC 계수들 위치에서 기 함수 대칭 특성을 각각 가진다. 각 모드의 주파수 특성을 14-17에 나타내었다.

$$\hat{T}_{M5}((j,i)|n) = -\hat{T}_{M5}((k,l)|n) \quad (14)$$

$$\hat{T}_{M7}((j,i)|n) = \hat{T}_{M7}((k,l)|n) \quad (15)$$

$$\hat{T}_{M6}((j,i)|o) = -\hat{T}_{M6}((k,l)|o) \quad (16)$$

$$\hat{T}_{M8}((j,i)|o) = \hat{T}_{M8}((k,l)|o) \quad (17)$$

식 14~17에 있는 index n과 o를 위한 특정 위치는 표 2와 같다. 그림 5의 (a)는 모드 5와 7을 위한 주파수 계수에서 특정 위치를, (b)는 모드 6과 8을 위한 특정 위치를 나타낸다.

표 2. Index m에 따른 특정 위치 (j,i)와 (k,l)의 위치

모드 5와 7			모드 6과 8		
n	(j,i)	(k,l)	o	(j,i)	(k,l)
0	(0,2)	(1,1)	0	(0,1)	(3,0)
1	(1,0)	(0,3)	1	(1,1)	(2,0)
2	(2,1)	(3,2)	2	(0,2)	(3,3)
3	(2,0)	(3,3)	3	(1,2)	(2,3)

T(0,0)	T(0,1)	T(0,2)	T(0,3)	T(0,0)	T(0,1)	T(0,2)	T(0,3)
T(1,0)	T(1,1)	T(1,2)	T(1,3)	T(1,0)	T(1,1)	T(1,2)	T(1,3)
T(2,0)	T(2,1)	T(2,2)	T(2,3)	T(2,0)	T(2,1)	T(2,2)	T(2,3)
T(3,0)	T(3,1)	T(3,2)	T(3,3)	T(3,0)	T(3,1)	T(3,2)	T(3,3)

(a) 모드 5와 7을 위한 위치 (b) 모드 6과 8을 위한 위치

그림 5. 예측 방향이 선 대칭을 이루는 각 모드 5와 7, 6과 8을 위한 특정 위치

위의 각 모드에 대한 AC 계수 제약 조건을 식 18-21에 나타내었다.

$$\Delta AC_{M5} = \frac{1}{4} \sum_{n=0}^3 |T((j,i)|n) + T((k,l)|n)| \quad (18)$$

$$\Delta AC_{M7} = \frac{1}{4} \sum_{n=0}^3 |T((j,i)|n) - T((k,l)|n)| \quad (19)$$

$$\Delta AC_{M6} = \frac{1}{4} \sum_{o=0}^3 |T((j,i)|o) + T((k,l)|o)| \quad (20)$$

$$\Delta AC_{M8} = \frac{1}{4} \sum_{o=0}^3 |T((j,i)|o) - T((k,l)|o)| \quad (21)$$

3.3 DC 계수 제약 조건

앞에서 구한 AC 계수의 특성뿐만 아니라 DC 계수에 있어서도 제약 조건을 구할 수 있다. 식 (2)를 이용하여 현재 4x4영상의 DC 계수 $T(0,0)$ 는 식 (22)와 같이 구할 수 있다.

$$T(0,0) = \sum_{j=0}^3 \sum_{i=0}^3 t(j,i) \quad (22)$$

LumaIntra4x4의 모드 0의 예측에 의한 DC 계수는 식 23과 같다.

$$\hat{T}_{M0}(0,0) = 4(a+b+c+d) \quad (23)$$

이 두 DC 계수들 $\hat{T}_{M0}(0,0)$ 와 $T(0,0)$ 의 차이는 현재 영상과 모드 0으로 예측된 영상과의 DC 에러를 예측할 수 있는 좋은 척도가 된다. 따라서 DC 계수 제약 조건에서는 예측을 수행하기 전에 그림 1에 나타난 예측에 사용할 값 a-1, x 값을 이용하여 각 모드별 DC 값을 구하고 이를 이용하여 각각의 모드별 DC 계수 제약 조건을 구할 수 있다. 각 모드별 DC 값은 다음과 같다.

$$\hat{T}_{M1}(0,0) = 4(i+j+k+l) \quad (24)$$

$$\hat{T}_{M2}(0,0) = 2(a+b+c+d+i+j+k+l) \quad (25)$$

$$\hat{T}_{M3}(0,0) = \frac{a+4b+8c+12d+14e+12f+8g+5h}{4} \quad (26)$$

$$\hat{T}_{M4}(0,0) = \frac{d+4c+8b+12a+14x+12i+8j+4k+l}{4} \quad (27)$$

$$\hat{T}_{M5}(0,0) = \frac{16a+15b+10c+3d+5i+3j+k+11x}{4} \quad (28)$$

$$\hat{T}_{M6}(0,0) = \frac{5a+3b+c+12i+17j+14k+3l+9x}{4} \quad (29)$$

$$\hat{T}_{M7}(0,0) = \frac{6a+13b+13c+12d+12e+7f+g}{4} \quad (30)$$

$$\hat{T}_{M8}(0,0) = \frac{3i+10j+15k+36l}{4} \quad (31)$$

각 모드 별 DC 계수 제약 조건들은 식 (32)와 같이 정의할 수 있다.

$$\Delta DC_{Mi} = |T(0,0) - \hat{T}_{Mi}(0,0)|, \quad 0 \leq i \leq 8 \quad (32)$$

앞에서 구한 주파수 특성들을 이용하여 구한 DC와 AC 계수 제약 조건들을 가지고 LumaIntra4x4 모드를 위한 주파수 에러 비용 (Frequency Error Cost, FEC)는 식 33과 같이 정의할 수 있다.

$$FEC_{Mi}^{4 \times 4} = \Delta DC_{Mi} + \Delta AC_{Mi}, \quad 0 \leq i \leq 8 \quad (33)$$

또한 각각의 4x4 영상의 지역 특성을 반영하는 $FEC_{Mi}^{4 \times 4}$ 를 이용하여 LumaIntra16x16 예측 모드의 특성을 나타낼 수 있다. LumaIntra16x16의 모드 0, 1, 3에 대해서 각 모드의 $FEC_{Mi}^{16 \times 16}$ 는 식 34와 같다.

$$FEC_{Mi}^{16 \times 16} = \sum_{v=0}^3 \sum_{u=0}^3 FEC_{Mi}^{4 \times 4}(v,u), \quad i = 0, 1, \text{ and } 3 \quad (34)$$

식 34에서 u, v는 4x4 블록 번호를 나타내며 LumaIntra4x4의 0, 1, 3번 모드의 FEC 값들을 이용하여 LumaIntra16x16의 0, 1, 3번 모드의 FEC 값을 구한다. 각각의 4x4 블록 위치 u, v에 대해서 $FEC_{M2}^{4 \times 4}(u,v)$ 는 언제나 값이 $FEC_{M0}^{4 \times 4}$ 나 $FEC_{M1}^{4 \times 4}$ 보다 크거나 같다. 따라서 LumaIntra16x16의 DC 모드에 대한 FEC는 구하지 않고 모드 0, 1, 3의 FEC만을 구한다.

IV. 제안하는 고속 모드 예측 방법

먼저 LumaIntra4x4 예측을 위해 PPM을 식 35와 같이 정한다.

$$PPM = \arg \min_{0 \leq i \leq 8} FEC_{Mi}^{4 \times 4} \quad (35)$$

식 35에서 정한 PPM과 DC mode, most probable mode, 그리고 PPM과 이웃하는 그림 2의 주변 모드 2개를 LumaIntra4x4를 위한 CMS으로 놓고 이 CMS에 해당하는 모드들에 대해서만 울-왜곡 비용을 구하므로 LumaIntra4x4 모드의 복잡도를 감소시킨다. 만약 PPM 이 DC 모드일 때는 발생 확률에 의거해서 모드 0과 1을 주변 모드로 놓고 수행한다. LumaIntra16x16 예측을 위해 식 34에 의해 구한 $FEC_{Mi}^{16 \times 16}$ 값을 이용하여 식 36을 만족하는 모드 i 를 찾는다.

$$FEC_{Mi}^{16 \times 16} < FEC_{Mj}^{16 \times 16}, \quad i = 0, 1, 3, \quad j = 0, 1, 3 \neq i \quad (36)$$

식 36을 만족하는 i 가 존재하면 LumaIntra16x16 후보 모드로 모드 i 와 모드 2만을 테스트 하고 i 가 존재하지 않으면 모드 0, 1, 2, 3을 테스트 한다. 마지막으로 LumaIntra16x16 최종 모드를 이용하여 ChromaIntra8x8 모드 예측을 한다. ChromaIntra8x8 모드는 LumaIntra16x16 모드와 밀접한 관련이 있다. 또한 색차 영상의 발생 확률에 따라 살펴보면 통계적으로 DC 모드가 다른 모드에 비해서 상대적으로 많이 발생하는 것을 알 수 있다^[4]. 따라서 ChromaIntra8x8 의 후보 모드로 LumaIntra16x16 최종 모드와 DC모드만을 테스트 하여 복잡도를 감소시킨다.

V. 실험 결과

제안하는 (Prop.) 고속 화면 내 모드 선택 알고리즘의 성능을 살펴보기 위해 H.264/AVC 참조 모델 JM 12.4^[6]에 적용하여 실험을 하였으며 성능 비교를 위해서 참조 소프트웨어의 방법과 함께 Park의 방법^[4]과 Pan 방법^[5]을 함께 비교하였다. 실험 환경은 다음과 같다.

CPU - Intel(R) Core™2 CPU 6600@2.40GHz
 Memory - 2.00GB
 Operating system - Windows XP Professional
 CODEC - JM 12.4
 Encoding structure - All Intra
 Quantization parameter - 20/24/28/32/36/40/44
 Rate-Distortion Optimization - used

먼저 각각의 알고리즘의 모드 선택 횟수에 대해서 표 3 에 나타내었다.

표 3. JM 12.4를 포함한 각각의 알고리즘의 모드 선택 횟수 비교

Intra Mode	JM	Park	Pan	Prop.
LumaIntra16x16	4	4	2	2 or 4
LumaIntra4x4	9	4-7	4	3 - 5
ChromaIntra8x8	4	1 or 2	2 or 3	1 or 2

즉 JM 12.4는 592번의 울-왜곡 비용 계산을 수행하며 Park의 방법은 최대 $2*(4+16*7)=232$ 번, 최소 $(4+16*4)=68$ 번의 울-왜곡 비용 계산을, Pan의 방법은 최대 $3*(2+16*4)=198$ 번, 최소 $2*(2+16*4)=132$ 번의 계산을, 제안하는 방법은 최대 $2*(4+16*5)=168$, 최소 $(2+16*3)=50$ 번의 울-왜곡 비용 계산을 수행하는 것을 알 수 있다.

다양한 특성의 영상에서 실험하기 위해 CIF 크기의 테스트 영상 Akiyo (300 frames), Bus (150 frames), Coastguard (300 frames), Container (300 frames), Football (90 frames), Foreman (300 frames) 을 이용하여 실험하였고 휘도 영상과 색차 영상의 평균제곱에러 (mean square error, MSE)를 모두 측정하기 위해 식 37과 같은 방법으로 aveMSE를 계산하였다.

$$aveMSE = \frac{4 \times MSE_U + MSE_V + MSE_V}{6} \quad (37)$$

이 aveMSE를 이용하여 객관적인 화질을 평가하기 위해 PSNR(peak to peak signal to noise ratio)을 계산하였으며 식 38과 같다.

$$PSNR = 10 \log_{10} \left(\frac{255^2}{aveMSE} \right) \quad (38)$$

H.264/AVC JM 12.4와 비교하여 각각의 알고리즘의 상대적인 PSNR 손실을 나타내기 위해 실험 결과를 식 39와 같이 변화량으로 나타내었으며 비트율과 시간을 식 (40)을 이용하여 상대적인 변화량을 측정하였다.

$$\Delta PSNR = PSNR_{JM} - PSNR \quad (39)$$

$$\Delta X = \frac{X - X_{JM}}{X_{JM}} \times 100 \quad (40)$$

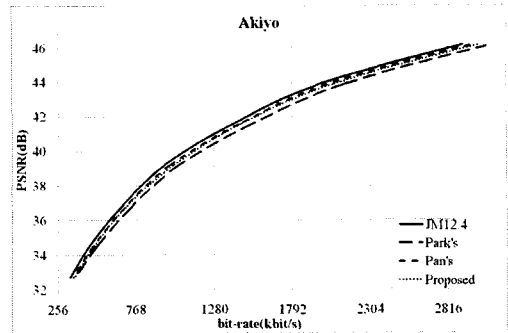
각각의 QP 20, 24, 28, 32, 36, 40, 44에 대한 평균적인 실험 결과를 JM 12.4와 비교하여 표 4에 나타내었다.

표 4에서 전체적으로 볼 수 있듯이 제안하는 알고리즘은 Park의 알고리즘에 비해 비슷한 속도 향상을 보이면서도 PSNR과 비트율 측면에서 우수한 결과를 보여준다. 또 Pan의 알고리즘에 비해 더 큰 속도 향상을 보이면서 PSNR과 비트율 측면에서는 비슷하거나 더 좋은 결과를 보인다. 특히 Foreman 영상의 경우 Park의 알고리즘은 67%의 속도 감소와 함께 6.128%의 비트율의 증가와 0.068dB의 PSNR 손실을 보인다. Pan의 알고리즘은 약 50%의 속도 감소와 함께 1.512%의 비트율 증가와 0.051dB의 PSNR 손실을 보인다. 이는 반해 제안하는 알고리즘은 66%의 속도 감소와 함께 1.260%의 비트율 증가와 0.026dB의 PSNR 손실을 보여 울-왜곡 측면에서 우수함을 알 수 있다.

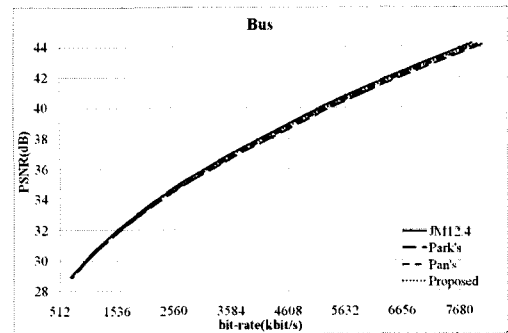
표 4. 평균 실험 결과

영상	방법	△시간 (%)	△PSNR (dB)	△비트율 (%)
Akiyo	Park	-64.610	-0.021	7.955
	Pan	-48.409	-0.034	2.460
	Prop.	-68.205	-0.034	4.079
Bus	Park	-68.168	-0.060	2.787
	Pan	-50.567	-0.062	1.723
	Prop.	-65.900	-0.048	0.899
Coastguard	Park	-67.739	-0.043	2.252
	Pan	-49.808	-0.047	1.024
	Prop.	-64.340	-0.029	0.701
Container	Park	-67.020	-0.036	3.588
	Pan	-49.661	-0.046	1.321
	Prop.	-64.440	-0.026	1.694
Football	Park	-67.028	-0.033	4.953
	Pan	-47.350	-0.045	1.615
	Prop.	-64.734	-0.028	2.116
Foreman	Park	-66.951	-0.068	6.128
	Pan	-49.423	-0.051	1.512
	Prop.	-65.629	-0.026	1.260

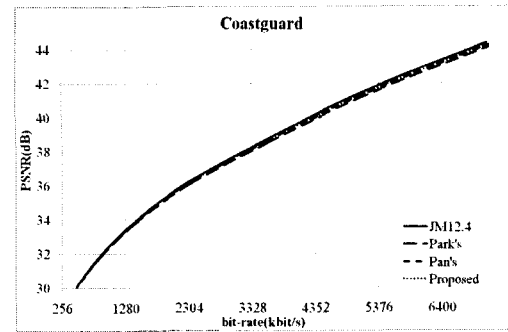
그림 6에 각 실험 영상에 대한 울-왜곡 성능을 나타내었다. 거의 모든 영상에서 제안하는 알고리즘이 다른 두 알고리즘에 비해 좋거나 비슷한 RD 성능을 나타내고 있음을 쉽게 확인 할 수 있다.



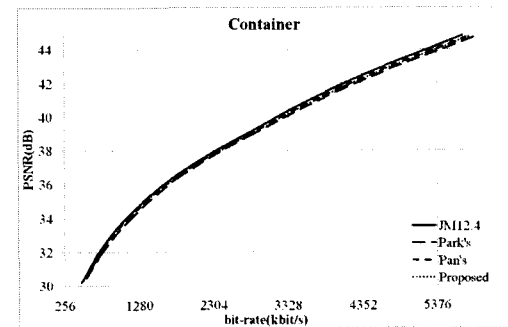
(a) Akiyo



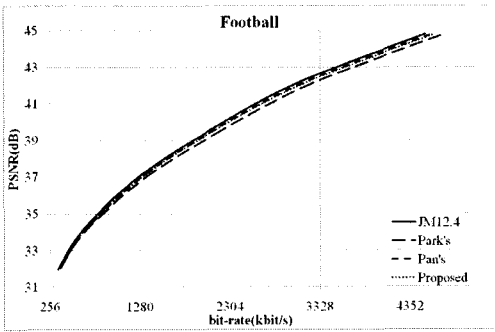
(b) Bus



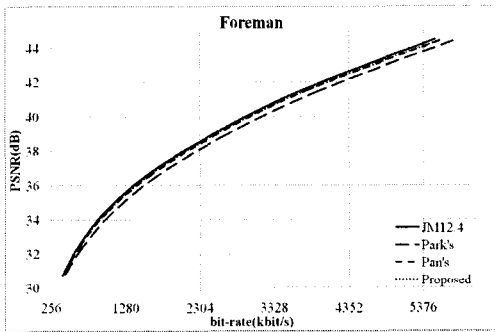
(c) Coastguard



(d) Container



(e) Football



(f) Foreman

그림 6. 각 영상 별 유효왜곡 성능 비교

VI. 결 론

본 논문에서는 H.264/AVC의 화면 내 예측의 모드별 특성을 주파수 측면에서 살펴보고 각 모드별 제약 조건을 유도하므로 고속 화면 내 모드 예측 방법을 제안하였다. 특히 4x4 예측 모드의 AC 계수 제약 조건과 DC 계수 제약 조건을 이용하므로 영상의 지역적 특성을 정확히 판단하여 4x4단위의 정확한 모드 예측을 가능하게 하였고 이 제약 조건들을 16x16 예측 모드에 이용하므로 속도 향상을 가져왔다. 실험 결과는 제안하는 알고리즘이 상당한 복잡도 감소를 보이면서 다른 알고리즘에 비해 유효왜곡 측면에서 더 좋은 성능을 나타냄을 증명한다.

참 고 문 헌

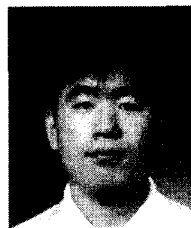
- [1] JVT G050r1, "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14496-10 AVC)," May 2003.
- [2] T. Wiegand, G. J. Sullivan, G. Bjontegaard,

and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Transactions on Circuits and Systems for Video Technology, Vol.13, pp.560 - 576, 2003.

- [3] Thomas Wiegand, Heiko Schwarz, Anthony Joch, Faouzi Kossentini, "Rate-Constrained Coder Control and Comparison of Video Coding Standard," IEEE Transactions on Circuits and Systems for Video Technology, Vol.13, pp.688-703, July 2003.
- [4] 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.
- [5] 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," IEEE Transactions on Circuits and Systems for Video Technology, Vol.15, No.7 pp.813-822, July 2005.
- [6] JVT Model downloaded from <http://bs.hhi.de/suehring/tml/download/jm1204.zip>, accessed 2007.
- [7] Avsnr4 tool downloaded from <http://ftp3.itu.ch/av-arch/video-site/h26L/avsnr4.zip>, accessed 2008.

진 순 종 (Soonjong Jin)

정희원



2004년 2월 한양대학교 전자컴퓨터공학부 졸업
2006년 3월 한양대학교 전자통신컴퓨터공학과 석사
2006년 3월~현재 한양대학교 전자통신전파공학과 박사과정
<관심분야> Image Compression,

Image Processing, H.264, Image Enhancement, Transcoding, MVC, SVC.

박 상 준 (Sang-Jun Park)

준희원

"한국통신학회 논문지 제 32권 제3C호"

정 제 참 (Jechang Jeong)

정희원

"한국통신학회 논문지 제 33권 제1C호"