

주관적 관심영역 중요도를 고려한 화면내 예측 간소화 방법

이호영[†], 권순각^{††}

요 약

최신 비디오 부호화 표준인 H.264에서는 화면내 예측으로 화소블록의 신호값을 예측하기 위하여 9개의 모드를 사용한다. 이러한 과정을 거쳐 화면내 부호화신호에서 우수한 압축율을 가져올 수 있으나 9가지 예측 모드 전체 사용은 인접화소 비교로 인해 발생하는 연산량 또는 탐색건수로 복잡도가 증가되는 비효율성을 내재하고 있다. 본 논문에서는 주관적 관심영역을 고려한 화면내 예측모드의 간소화 방법을 제안한다. 비디오 신호의 각 화면에는 관심이 주어지는 특정영역이 존재한다. 이러한 영역은 다른 영역에 비하여 더 좋은 주관적 화질 제공이 필요하다. 제안된 방법은 주관적 관심이 높은 영역에 비하여 주관적 관심이 낮은 영역에는 9개의 모드중에 필수모드만을 제공하여 예측모드 간소화를 높여준다. 예측특성만을 고려하여 화면 전체에 간소화를 주는 기존 방법에 비하여 제안된 방법은 추가적으로 11%~15%의 예측모드 간소화가 가능하다.

A Simplification Method of Intra Prediction Considering Importance of Subjective Interest Region

Ho-young Lee[†], Soon-kak Kwon^{††}

ABSTRACT

In H.264 as the newest video standard, 9 modes are used in order to predict the signal values of a block composed with several pixels by intra prediction. From these process, H.264 can bring high compression ratio in the encoded signal but the use of total 9 modes can give the inefficiency of the increase of the complexity induced by the amount of operation processing or the number of searching which is applied to compare adjacent pixels. This paper proposes a simplification method of prediction mode for the intra-picture coding by considering subjective interest region. There are certain region being interested within a picture of the video sequence. This region requires better subjective picture quality than the other regions. The proposed method increases the simplification of prediction mode by providing just essential modes of total 9 modes for less interest regions compared with the interest region. It is possible to get the additional 11% ~ 15% simplification of the prediction mode by the proposed method, compared with the conventional method which simplifies the prediction mode for all of the picture by using the prediction characteristics only.

Key words: Intra-picture coding(화면내 부호화), Interest region(관심영역)

1. 서 론

최근에 H.264[1-3] 부호화에서 제안된 화면내 예

측 방법은 현재 화면내의 신호의 중복성을 제거하기 위하여 먼저 부호화되고 복원된 이웃한 화소를 사용하여 부호화 효율을 증대시키고 있다[4].

※ 교신저자(Corresponding Author): 권순각, 주소: 부산시 부산진구 엄광로 995(614-714), 전화: 051)890-1727, FAX: 051)890-2629, E-mail: skkwon@deu.ac.kr
접수일: 2009년 1월 20일, 완료일: 2009년 4월 22일

[†] 정회원, 알림프로모션 팀장
(E-mail: wincomys@dreamwiz.com)

^{††} 종신회원, 동의대학교 컴퓨터소프트웨어공학과 부교수

화면내 예측을 통한 높은 압축 효율은 매크로블록 단위로 예측을 수행할 때 16×16 블록에 대해 4가지 예측모드, 4×4 블록에 대해서는 9가지 예측모드를 지원한다. 이 과정에서 모든 모드를 비교하여 오차 값을 계산하게 된다. 표준에서 제시된 화면내 예측은 화질은 우수하나 그 계산과정이 복잡하고 처리의 효율성이 떨어지기 쉽다. 화면내 예측부호화 복잡도를 줄이기 위한 방법으로 블록크기를 빠르게 결정하기 위해 16×16 블록과 4×4 블록의 경계선 신호크기 값과 SATD(Sum of Transformed Difference) 사이의 임계값을 사용하는 방법[5]과 모드 결정을 위해서 Sobel 연산자를 이용하는 방법[6], 매크로블록의 평탄 정도를 이용하여 블록크기를 결정하고 결정된 블록 크기 안에서 대표화소를 이용하는 방법[7], 4×4 블록에서 사용되는 예측모드를 16×16 블록 및 색차블록에 확대 적용하는 방법[8] 등이 있다. 또한, 예측모드 사용빈도를 측정하여 빈도가 낮은 예측모드를 사용하지 않고 생략하는 예측모드 간소화 방법[4]이 제안되었다.

예측모드 사용빈도에 의한 예측모드 간소화 방법에 의해 영상특징별로 예측모드의 생략은 부호화 시점에서 연산시간을 줄이는 효과가 있어 화면내 예측의 고속화 또는 간소화에 기여하고 있음을 알 수 있었다. 이러한 결과를 근거로 본 논문에서는 주관적 관심영역에 의한 슬라이스별 예측모드의 차등적용으로 영상별 예측모드 간소화 방법을 제안한다. 즉, 예측모드의 사용빈도는 인접화소를 비교하는 탐색 건수를 얼마나 줄이느냐가 관건인데 기존의 예측모드 사용빈도를 고려한 간소화 방안을 전체 화면에 적용하지 않고 화면내 관심영역으로 분류되는 영역 또는 슬라이스에는 기존에 제안된 예측모드 간소화 방안에서 도출한 예측모드를 적용하고 그 외 비관심영역 또는 슬라이스에는 기본 예측모드 0, 1, 2만을 사용한다.

본 논문의 구성은 2장에서 관련 연구로서 예측모드 사용빈도에 의한 예측모드 간소화를 위한 기존 방안에 대해 알아보고, 3장에서는 주관적 관심영역의 화질개선을 위한 기존 방안에 대해 알아본다. 4장에서는 주관적 관심영역의 중요도를 고려한 예측모드 간소화 방안을 제안하며 5장에서 모의실험의 결과를 분석하며 6장에서 결론에 대해 기술한다.

2. 화면내 예측 간소화를 위한 기존 방법

H.264의 화면내 예측에서는 각 매크로블록에 적용될 화소 값을 결정하기위해 JM[9] 부호화기에 의한 9개의 예측모드를 그림 1과 같이 모두 적용한다.

그림 1과 같이 9개의 예측모드 적용에 따른 복잡도를 감소시키기 위하여 예측모드의 사용빈도를 고려한 간소화 방법이 제안되었다[4]. 예측모드 사용빈도에 의한 간소화 방법은 7% 이하의 사용빈도를 보이는 예측모드를 생략하는 것으로서, Bus 영상의 경우는 예측모드 0, 1, 2만 사용하고, Flower 영상의 경우는 예측모드 0, 1, 2, 8만을 사용한다. 중앙의 움직임이 많은 Foreman 영상은 예측모드 3, 5, 6, 7 등 4개의 예측모드를 생략하고, Waterfall 영상과 같이 중앙의 움직임 보다 주변 배경의 움직임이 많은 영상은 예측모드 3, 5, 7 등 3개의 예측모드를 생략한다.

예측모드 사용빈도 7% 이하의 예측모드를 간소화시키면 화질의 변화가 미미한 상태에서 기존 9개 예측모드에서 3개 또는 6개의 예측모드만의 사용으로 연산량 또는 탐색건수에서 33%~67%의 간소화가 가능한 것으로 나타났다.

3. 주관적 관심 영역의 화질 개선을 위한 기존 방법

해상도와 화질의 무시할 수 있는 범위내에서 양자

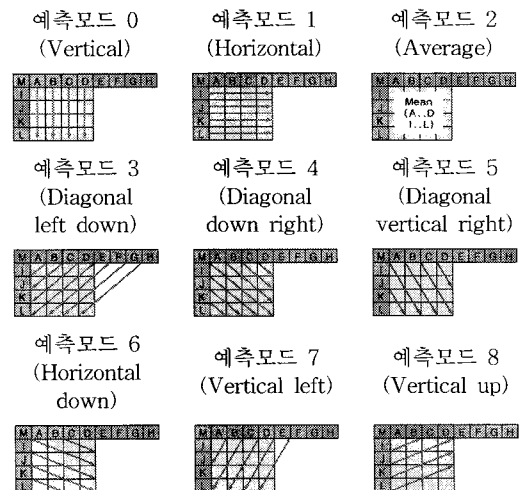


그림 1. 화면내 예측모드

화계단의 크기가 결정되는데 최근 다양한 멀티미디어 기기들의 해상도는 기존 모니터 해상도와 많은 차이를 보이고 있으며 해상도별 양자화 계단크기도 적용적으로 선택되어야한다. 이 과정에서 주관적 화질 개선의 필요성이 제기되었다. 이로부터 화면내 관심영역을 고려한 양자화 계단크기를 달리 적용하는 방안이 제안되었다[10]. 움직임이 많은 영역은 양자화 계단크기를 상대적으로 작게 부여하며 움직임이 작은 영역은 양자화 계단크기를 상대적으로 크게 부여한다. 이를 위하여 영상을 그림 2와 같이 5개의 관심영역으로 분류하였다. 관심영역별로 화면내의 양자화 계단크기를 차등적으로 설정하면 화면전체의 평균값인 객관적 PSNR 값은 다소 감소하지만, 화면내의 관심을 갖는 부분에 초점을 갖는 주관적 화질은 오히려 더 좋아짐을 알 수 있다. 즉, 객관적 화질 측정만으로는 인간의 시각적 화질 수준에 완전히 만족을 줄 수 없는 것으로 판단되었으며 주관적 화질평가 결과를 객관적 화질측정 결과에 추가적으로 적용하여 비트율의 변동이 없는 가운데 화질개선의 효과가 나타남을 알 수 있었다.

4. 주관적 관심영역 중요도를 고려한 화면내 예측 간소화방법

해상도와 예측모드 간소화 방법에 의해 영상특징

별로 예측모드의 생략은 부호화 시점에서 연산시간을 줄이는 효과가 있어 화면내 예측의 고속화 또는 간소화에 기여하고 있음을 알 수 있었다. 또한, 화면내 관심영역의 슬라이스별로 양자화변수 값을 차등 적용함으로써 주관적 화질개선이 가능함을 알 수 있었다.

이러한 특징을 고려하여 주관적 관심영역에는 화면내 예측시 상대적으로 많은 수의 예측모드를 적용하고, 반대로 주관적으로 관심이 별로 없는 영역에는 화면내 예측시 상대적으로 예측모드 수를 줄이는 간소화 방법을 제안한다. 제안방법은 기존방법[10]에서 제시된 것처럼 영상을 관심영역에 따라 분류하고, 관심영역에는 기존방법[4]에서 제시된 것처럼 예측모드의 사용빈도가 7%미만이 되는 예측모드는 생략시키고, 그렇지 않는 영역에는 기본이 되는 예측모드 0, 1, 2만 적용시킨다.

그림 3과 같이 주관적 관심영역 분류에 의해 중앙 집중방식은 중앙에, 외곽집중방식은 외곽에 예측모드를 많이 적용하고, 그렇지 않는 영역에는 기본으로 사용되는 예측모드를 적용한다. 주관적 관심영역과 그렇지 않은 영역을 한 화면에서 슬라이스로 나눈다. 한 화면을 구성하는 슬라이스가 18개인 경우에 각각 1/3씩 6개 슬라이스로 나누어 중앙집중방식은 중앙 1/3영역에 예측모드 빈도 7% 이하를 제외한 예측모

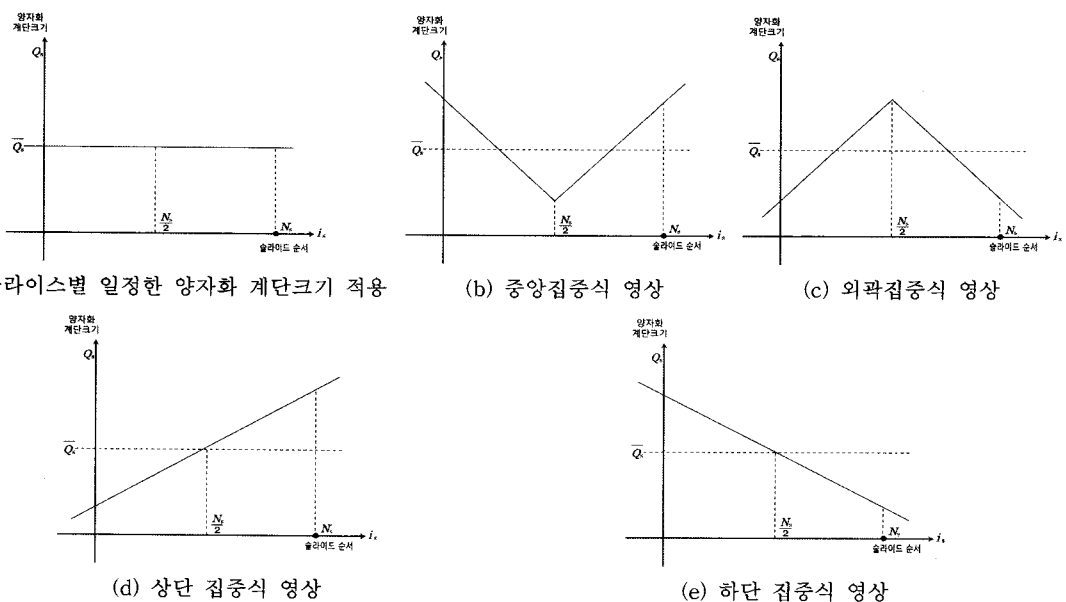


그림 2. 관심 영상형태별 양자화 계단크기의 차등 적용

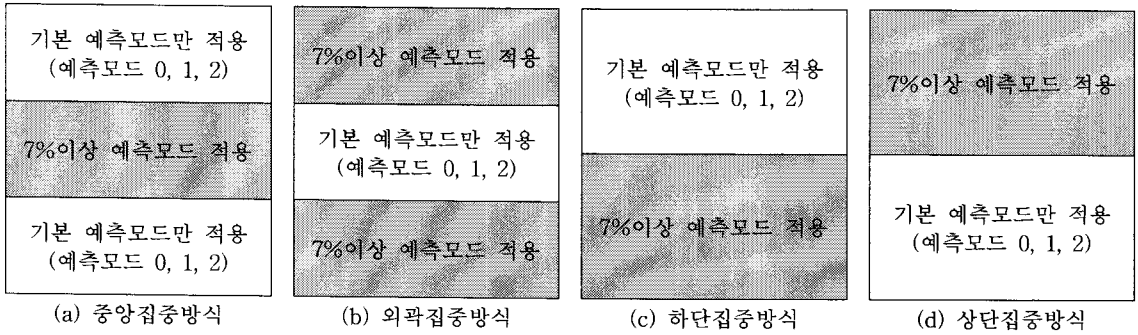


그림 3. 주관적 관심영역에 따른 화면내 예측모드 간소화

를 모두 적용하며, 외곽부분에는 기본적으로 가장 많이 사용되는 0, 1, 2 예측모드만 적용한다. 외곽집중방식은 18개 슬라이스중 중앙부분 1/3인 6개 슬라이스를 제외한 외곽영역에 기존방법과 같이 예측빈도 7% 이하를 제외한 모든 예측모드를 적용하며 나머지 중앙부분의 슬라이스에는 기본 예측모드인 0, 1, 2만 적용한다.

마찬가지로, 하단집중방식은 하단에, 상단집중방식은 상단에 예측모드 사용빈도 7% 이상의 모드를 적용하며 그 외 부분에는 기본모드만 적용한다.

모든 영상은 부호화 과정에서 움직임 벡터를 도출할 수 있으며 급격한 화면의 변화(scene change) 또한 발견해낼 수가 있다. 이러한 간소화 방법은 움직임 벡터 값을 이용하지만 최초 I-화면을 분석한 자료를 바탕으로 다음 I-화면부터 간소화 방안을 적용하여 화면의 변화, 즉 장면전환이 있기 전까지 지속적인 적용이 가능하다.

5. 모의실험 및 결과

본 논문에서 제안하는 방법에 대한 실험을 위해 H.264의 JM10.2[9]를 사용하였으며 실험 영상으로는 해상도가 수평 352화소, 수직 228화소를 갖는 Foreman, Waterfall 영상을 채택하였다. H.264 baseline profile을 사용하여 B-화면은 포함되지 않으며, 15개 화면이 1 개의 GOP로 구성되며 각 영상은 총 60개의 화면을 사용하였다. 한 화면당 슬라이스의 크기는 서로 같으며, 세로로 18개를 가진다.

제안된 방법은 9개의 예측모드를 모두 사용하는 방법, 예측모드 사용빈도에 따른 예측 간소화 방법 [4]과 부호화성능을 비교한다. Foreman 영상은 중앙

에 관심이 많은 영상으로서 중앙집중방식 간소화를 적용하였고, Waterfall 영상은 외곽에 관심이 많은 영상으로서 외곽집중방식 간소화를 적용하였다.

표 1과 표 2는 Foreman 영상에 대하여 I-화면간 화면의 간격이 2인 경우와 15인 경우에 대하여 2.5Mbps 비트율로 부호화하고 다시 복원한 영상에 대한 PSNR과 발생된 비트수를 각각 나타낸다. 기존 방법에서는 5개의 예측모드(모드 0, 1, 2, 4, 8)가 사용되고, 제안방법에서는 중앙부분에 기존방법과 같은 5개의 예측모드가 적용되고, 외곽부분에 기본 예측모드(모드 0, 1, 2)가 사용된다.

Dif PSNR은 예측모드를 전체 적용하는 방법에 의한 PSNR값과 기존 및 제안방법에 의한 PSNR값간의 차이를 나타낸다.

Foreman 영상은 표 1에서와 같이 I-화면거리 2에서는 PSNR이 다소 감소되는 현상을 보이고 있지만

표 1. Foreman 영상의 예측모드 간소화 결과(중앙집중방식, I-화면간 화면간격 : 2, 부호화율 : 2.5Mbps)

구 분	모드 (9개)	기존 방법 모드 (5개)	제안방법 중앙: 모드(5개) 외곽: 모드(3개)
PSNR(전체)	40.87	40.74	40.72
Dif PSNR(전체)	0.00	0.13	0.15
PSNR(I화면)	41.36	41.23	41.22
Dif PSNR(I화면)	0.00	0.13	0.14
발생비트수(전체)	5027192	5033928	5034720
발생비트수(I화면)	4145552	4203064	4214616
블록당 평균 예측모드 개수	9	5	3.67
예측모드 간소화 비율	-	44%	59%

표 2. Foreman 영상의 예측모드 간소화 결과(중앙집중방식, I-화면간 화면간격 : 15, 부호화율 : 2.5Mbps)

구 분	모드 (9개)	기존 방법 모드 (5개)	제안방법 중앙: 모드(5개) 외곽: 모드(3개)
PSNR(전체)	43.84	43.83	43.83
Dif PSNR(전체)	0.00	0.01	0.01
PSNR(I화면)	42.92	42.89	42.89
Dif PSNR(I화면)	0.00	0.03	0.03
발생비트수(전체)	5002312	5005928	5003016
발생비트수(I화면)	621096	637376	639000
블록당 평균 예측모드 개수	9	5	3.67
예측모드 간소화 비율	-	44%	59%

화면거리별 평균값은 0.1dB 내외로서 미미한 차이를 보였다. 이러한 경우 주관적 관심도가 중앙쪽으로 집중되어 있기 때문에 중앙집중방식을 적용한다면 수처리에 의한 화질평가가 비슷하므로 주관적으로 느끼는 화질에는 문제가 없음을 알 수 있다.

전체적으로 발생하는 비트수가 비슷하고 화질의 감소가 미미한 가운데 사용되는 예측모드 개수가 중앙부분에 해당되는 화면의 1/3영역에서는 기존방법 [4]의 간소화와 동일하지만 외곽부분에 해당되는 화면의 2/3영역에서 3/9의 추가적인 간소화가 가능하다. 즉, 주관적 관심영역에 의한 중앙집중방식을 Foreman 영상에 적용했을 때 화면내 예측을 위해서 사용되는 예측모드 개수는 중앙부분 44% 간소화, 외곽부분의 66% 간소화로 기존 간소화방법에 비해 약 15%의 추가적인 간소화가 가능하여 영상 전체적으로 약59% 간소화가 가능하다.

Waterfall 영상의 경우는 외곽집중방식으로 부호화 하였으며, 결과를 표 3과 표 4에 나타낸다. 기존방법에서는 6개의 예측모드(모드 0, 1, 2, 4, 6, 8)가 사용되고, 제안방법에서는 외곽부분에 기존방법과 같은 6개의 예측모드가 적용되고, 중앙부분에 기본 예측모드(모드 0, 1, 2)가 사용된다.

Waterfall 영상은 표 3, 표4에서와 같이 화면거리 2일 경우 PSNR값이 0.2dB 내외로 다소 감소되는 현상을 보이고 있지만 전체 화면거리를 감안하면 미미한 차이를 보였다. 화면거리 15에서는 기존 방법에 비해 0.1dB로 극히 미미한 차이를 보였다. 전체적으로 발생하는 비트수가 같고 화질의 감소가 미미한

표 3. Waterfall 영상의 예측모드 간소화 결과(외곽집중방식, I-화면간 화면간격 : 2, 부호화율 : 2.5Mbps)

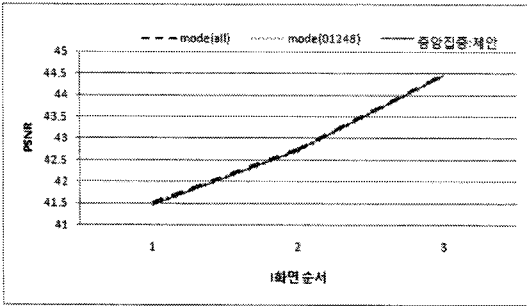
구 분	모드 (9개)	기존 방법 모드 (6개)	제안방법 외곽: 모드(6개) 중앙: 모드(3개)
PSNR(전체)	36.42	36.32	36.18
Dif PSNR(전체)	0.00	0.10	0.24
PSNR(I화면)	35.78	35.70	35.68
Dif PSNR(I화면)	0.00	0.08	0.10
발생비트수(전체)	5024576	5031016	5028152
발생비트수(I화면)	4543616	4582256	4555672
블록당 평균 예측모드 개수	9	6	5
예측모드 간소화 비율	-	33%	44%

표 4. Waterfall 영상의 예측모드 간소화 결과(외곽집중방식, I-화면간 화면간격 : 15, 부호화율 : 2.5Mbps)

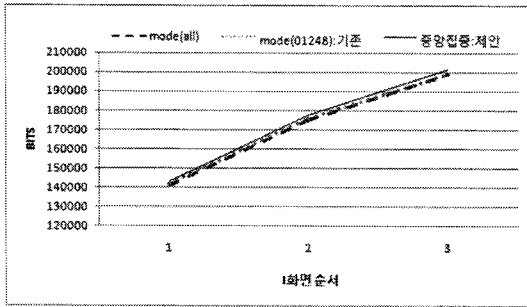
구 분	모드 (9개)	기존 방법 모드 (6개)	제안방법 외곽: 모드(6개) 중앙: 모드(3개)
PSNR(전체)	42.43	42.41	42.40
Dif PSNR(전체)	0.00	0.02	0.03
PSNR(I화면)	42.11	41.01	41.00
Dif PSNR(I화면)	0.00	0.10	0.11
발생비트수(전체)	5012392	5012376	5012008
발생비트수(I화면)	1108488	1117120	1138416
블록당 평균 예측모드 개수	9	6	5
예측모드 간소화 비율	-	33%	44%

가운데 사용되는 예측모드 개수가 외곽부분에 해당되는 화면의 2/3영역에서는 기존방법[4]의 간소화와 동일하지만 화면의 1/3영역에서 3/9의 추가적인 간소화가 가능하다. 즉, 기존방법에서의 예측모드 사용빈도에 의한 33% 간소화가 전체 화면의 2/3 영역에서 가능하고 추가적으로 1/3 영역에서 다시 66%가 간소화되는 결과가 나타난다. 이를 종합하면 주관적 관심영역에 의한 외곽집중방식을 Waterfall 영상에 적용했을 때 중앙 1/3영역에서 66% 간소화, 외곽부분 2/3에서 33% 절감으로 평균 44% 간소화가 가능하였다.

Foreman 영상에 대한 중앙집중식 예측모드의 적용에 따른 PSNR 및 비트수의 변화는 그림 4에 나타

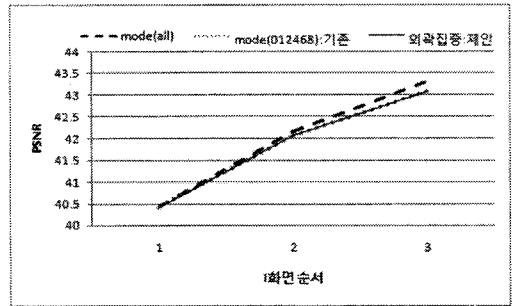


(a) PSNR

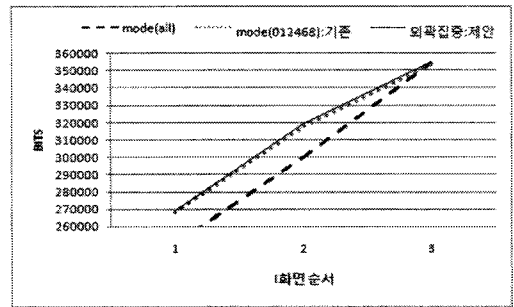


(b) 발생되는 비트수

그림 4. Foreman 영상의 중앙집중식 예측모드 적용시 PSNR값과 비트수(I-화면간 화면간격 : 15)



(a) PSNR



(b) 발생되는 비트수

그림 5. Waterfall 영상의 외곽집중식 예측모드 적용시 PSNR값과 비트수(I-화면간 화면간격 : 15)

낸다.

그림 4의 결과에서 나타난 바와 같이 비트수의 변화는 증감이 되지만 PSNR의 값은 기존 방법 그리고 예측모드 간소화방식에 중앙집중방식을 적용한 제안된 방법의 차이가 미미하다.

그림 5는 Waterfall 영상의 외곽집중식 예측모드를 적용하였을 때의 PSNR값과 비트수의 변화를 나타낸다.

그림 5에 나타난 바와 같이 Waterfall 영상은 외곽 집중방식으로 적용한 결과 Foreman 영상과 마찬가지로 제안된 방법이 예측모드 사용빈도 7% 이하의 예측모드를 간소화하는 기존 방법과 PSNR값에서 미미한 차이를 보인다.

6. 결 론

H.264 동영상 부호화에서는 화면내 부호화 효율을 증대시키기 위하여 화면내 예측기법을 사용하고 있다. 화면내 예측을 위해서는 수평방향 4화소와 수직방향 4화소로 구성된 블록단위로 9개의 예측모드가 적용되고, 이중에서 가장 예측성능이 좋은 하나의

모드가 선택되어 부호화에 사용된다. 모든 블록에 대하여 9개의 예측모드를 사용하면 화면내 부호화를 위한 복잡도가 증가하게 된다. 따라서, 본 논문에서는 화면내 예측을 위해서 사용되는 예측모드 간소화를 통하여 화면내 부호화의 복잡도를 줄이기 위한 방법을 제안한다. 이를 위하여 화면을 관심영역과 그렇지 않는 영역으로 구분하고, 관심영역에는 좋은 화질을 제공하기 위하여 기존방법처럼 9개의 예측모드 중에서 출현빈도가 7%이상되는 5~6개의 예측모드를 적용한다. 그러나, 비교적 관심이 없는 영역에는 9개의 예측모드 중에서 중요한 3개의 예측모드만을 적용한다. 실험결과를 통하여 제안된 방법은 기존 방법과 비슷한 화질을 유지하면서, 예측모드 사용빈도 측면에서 11%~15%의 추가적인 간소화가 가능함을 확인하였다.

참 고 문 헌

[1] S.-k. Kwon, A. Tamhankar, and K. R. Rao, "Overview of H.264/MPEG-4 Part 10", *Journal of Visual Communications and Image*

Representation, Vol. 17, No. 2, pp. 186-216, 2006.

[2] ITU-T Rec. H.264 / ISO/IEC 11496-10, "Advanced Video Coding," *Final Committee Draft, Document JVT-F100*, 2002.

[3] T. Widegand, G. Sullivan, G. Bjonttegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard", *IEEE Trans. Circuit Syst. Video Technology*, Vol. 13, No. 7, pp. 560-576, 2003.

[4] 이호영, 권순각, 김태석, "화면내 부호화를 위한 예측특징 분석", *Journal of the Korean Data Analysis Society*, Vol. 11, No. 2, 2009.

[5] K. Zhang, C. Yuan, and Y. Zhong, "A Fast Block Type Decision Method for H.264/AVC IntraPrediction", *The 9th International Conference on Advanced Communication Technology*, pp. 673-676, 2007.

[6] F. Pans and H. J. Song, "Fast Mode Decision Algorithm for Intraprediction in H.264/AVC Video Coding", *IEEE Trans. Circuit Syst. Video Technology*, Vol. 15, No. 7, pp. 813-822, 2005.

[7] 이탁기, 김성민, 신광부, 정기동, "H.264에서 화소변화량을 이용한 빠른 인트라 예측", *멀티미디어학회논문지*, Vol. 10, No. 7, pp. 956-965, 2008.

[8] 김동형, 정제창, "H.264/AVC비디오의 고속 부호화를 위한 인트라모드 선택 알고리즘", *한국통신학회논문지*, Vol. 32, No. 6, pp. 956-995, 2007.

[9] JM Reference Software Version 10.2, <http://iphome.hhi.de/suehring/tml/download/>

[10] 이호영, 권순각, 이증화, "H.264 동영상 부호화에서 관심영역의 주관적 화질 개선 방법", *멀티미디어학회논문지*, Vol. 12, No. 7, 2009.



이 호 영

1983년 2월 중앙대학교 전자계산학과 졸업
 1994년 2월 한양대학교 언론홍보학과 석사
 2009년 2월 동의대학교 컴퓨터소프트웨어공학과 박사
 1996년 3월~현재 알림프로모션팀장

관심분야 : 영상편집, 영상신호처리



권 순 각

1990년 2월 경북대학교 전자공학과 졸업
 1992년 2월 KAIST 전기및전자공학과 석사
 1998년 2월 KAIST 전기및전자공학과 박사
 1997년 3월~1998년 8월 한국전자통신연구원 연구원

1998년 9월~2001년 2월 기술신용보증기금 기술평가센터 팀장
 2003년 9월~2004년 8월 Univ. of Texas at Arlington 교환 교수
 2001년 3월~현재 동의대학교 컴퓨터소프트웨어공학과 교수

관심분야 : 멀티미디어신호처리, 영상통신