

능률적 AZCB 예측 기반 H.264 적응 모드 결정 개선 알고리즘

論 文

56-11-28

Enhanced Adaptive Mode Decision of H.264 Based on Efficient AZCB Prediction

金 敷 洙* · 金 勇 求** · 崔 榮 鎬† · 崔 潤 植***

(Yang-Soo Kim · Yong-Goo Kim · Yung-Ho Choi · Yoon-Sik Choe)

Abstract - This paper proposes an enhanced adaptive mode decision scheme for fast H.264 encoders. By efficiently predicting AZCB (All Zero Coefficient Block), the proposed scheme can encode motion pictures in H.264 up to 2.86 and 1.68 times faster than JM9.3 and AMD [1], respectively. Besides, this scheme significantly reduces the encoding performance fluctuation of AMD across tested bit-rates and video sequences.

Key Words : H.264, AZCB, 부호화 모드, Encoder, Image Compression

1. 서 론

H.264 부호기는 다양한 부호화 모드를 탐색하여 영상 압축에 최적 모드를 선택함으로써 높은 영상 압축 효율을 얻는다. 그러나 이와 같은 다양한 모드의 탐색은 H.264 부호기의 복잡도를 크게 증가시킴으로써 고 품질 실시간 부호기의 구현에 어려움을 유발하고 있다.

이러한 부호 모드 선택의 복잡도 문제를 해결하기 위하여, 다양한 고속 모드 결정 알고리즘들이 제안되었다 [1-4]. 이들 중 적응적 모드 결정(Adaptive Mode Decision, AMD) 알고리즘은 코딩 성능 및 속도 개선 면에서 매우 뛰어난 성능을 보여주었다. 이 알고리즘은 움직임 추정/복원 및 비트율-왜곡 최적화(Rate-Distortion Optimization, RDO) 로 이루어진 부호화 루프에서 초기 종료를 허용함으로써 부호화 복잡도를 경감시켰다. 보다 자세하게 설명하자면, AMD 알고리즘은 가장 큰 크기(16x16)의 블록 파티션으로부터 가장 작은 크기(4x4)의 블록 파티션 순으로 움직임 추정(Motion Estimation, ME) 및 비트율-왜곡 최적화 연산을 수행하면서, 매 단계에서 모든 이산여현변환(Discrete Cosine Transform, DCT) 계수가 0이 되는지를(All-Zero Coefficient Block, AZCB) 탐색한다. 이는 큰 블록 파티션 모드에서 AZCB가 생성되는 경우 이보다 더 작은 블록 파티션을 가지는 모드들은 ME의 정확도 증가로 AZCB가 될 것이며, 따라서 이들 모드들에 대한 ME 및 RDO 연산은 생략해도 코딩 성능에 차이가 발생하지 않아, 결과적으로 완전 모드 결정(Full Mode Decision, FMD) 과 같은 부호 성능을

보다 빠른 부호화 속도로 얻을 수 있게 하기 위함이다.

그러나, AMD 알고리즘은, 위와 같은 성능상의 이점에도 불구하고, 그 동작 원리가 가지는 한계에 의하여 섬세하면서 동작이 많아 대부분의 매크로블록이 작은 블록 파티션 모드로 부호화되는 경우 그 부호화 속도 개선에 한계를 나타낸다. 이는 원본 비디오 특성에 따른 부호화 속도의 차이를 갖게 하고, 일반적으로 이와 같은 차이가 크게 나타나는 경우 실시간 부호기 설계에 심각한 문제를 유발한다. 이러한 문제는 연산 능력이 제한된 장치에서 비디오 부호화를 수행하는 경우 특정 장면에 대해서는 매우 낮은 프레임 율을 가지는 부호화 결과를 보이게 되고, 반대로 일정 프레임 율 이상의 압축을 제공하는 장치의 개발을 위해서는 가장 낮은 부호 속도를 고려하여 매우 높은 연산 능력을 갖춘 연산 장치를 채용해야 하는 것을 의미한다. 따라서 본 논문에서는 이와 같은 문제를 해결하기 위한 효과적이면서도 빠른 모드 결정 알고리즘을 제시하고자 한다.

본 논문에서는 속도 개선 성능에 큰 변동을 보이는 AMD 기법보다 안정적인 부호화 속도 변화를 가지는, 향상된 적응 모드 결정 기법을 제안한다. 본 논문에서 제안하는 알고리즘에서는 AZCB 평가 연산을 직접적으로 수행하는 대신, DCT 계수의 통계적 분산을 활용하여, 조정 가능한 신뢰 매개 변수(Configurable Confidence Parameter) 에 의해 계산된 최대 계수 레벨을 추정함으로써, AZCB를 예측하였다. 이러한 AZCB 예측 기법은 신뢰 매개 변수를 변경함에 따라 부호화의 비트율-왜곡 품질과 부호화 속도간의 점진적 절충(graceful trade-off)이 가능하다. 적절한 성능 절충을 통해, 본 논문의 알고리즘은 AMD 알고리즘의 비트율-왜곡 성능을 거의 그대로 유지하면서, 부호화 속도 변동 문제를 해결할 수 있었다. 게다가, 예측에 사용된 알고리즘의 구현을 극도로 단순화시킴으로써 AMD 보다 최대 1.68 배 정도 더 빠른 부호화 속도를 갖게 하였다.

* 正 會 員 : 延世大學 電氣電子工學部 博士課程

** 正 會 員 : 延世大學 電氣電子工學部 研究教授

† 교신저자, 正會員 : 建國大學 電氣工學科 助教授 · 工博

E-mail : yunghoch@konkuk.ac.kr

*** 正 會 員 : 延世大學 電氣電子工學部 教授 · 工博

接受日字 : 2007年 7月 31日

最終完了 : 2007年 8月 24日

2. AZCB의 예측

움직임 보상 잔여 신호의 모델링에 일반적으로 사용되는 라플라시안(Laplacian) 분포 가정으로부터 잔여신호 블록의 절대치 합(Sum of Absolute Differences, SAD)은 식(1)과 같이 나타낼 수 있다.

$$SAD / B_{SIZE} = 2 \times \int_0^{\infty} x \cdot \frac{\alpha}{2} \exp(-\alpha \cdot |x|) dx = \frac{1}{\alpha} = \frac{\delta_f}{\sqrt{2}} \quad (1)$$

여기서 Bsize는 잔여신호 블록의 화소수이며, δ_f 는 라플라시안 분포에 의해 가정된 잔여신호의 분산(variance)이다. 중심한계정리(Central Limit Theorem)에 의해, NxN 블록의 각 DCT 계수는 가우시안 분포[6]로 근사될 수 있으며, 이 때 가우시안 분포의 분산은 다음과 같이 표현될 수 있다.

$$\delta_f^2[\mu, \nu] = \delta_f^2 \times M[\mu, \nu], \quad \mu = 0 \sim N; \nu = 0 \sim N \quad (2)$$

여기서 $\delta_f[\mu, \nu]$ 는 $[\mu, \nu]$ 번째 DCT 계수의 분산을 나타내고, 분산 변환 가중치(Variance Translation Weight)인 $M[\mu, \nu]$ 는 DCT 커널에 의해 결정된다. 4X4 DCT의 경우, 분산 변환 가중치들은 아래의 행렬과 같이 주어진다.

$$M = \begin{bmatrix} 5.607424 & 2.125210 & 1.060864 & 0.678503 \\ 2.125210 & 0.805453 & 0.402067 & 0.257152 \\ 1.060864 & 0.402067 & 0.200704 & 0.128365 \\ 0.678503 & 0.257152 & 0.128365 & 0.082099 \end{bmatrix} \quad (3)$$

식 (2)와 (3)을 이용하여, 각 DCT 계수를 잔여신호 블록을 부호화할 양자화 스텝과 비교함으로써, ‘ $[\mu, \nu]$ 번째 DCT 계수는 0으로 양자화 된다.’라는 가설 검증(Hypothesis Test)을 수행할 수 있다. 이 때, (3)으로부터 가장 큰 분산값을 DC 계수에 해당하는 $\delta_f[0,0]$ 이며, 따라서 AZCB 가설은 식 $Q_{STEP} > c \cdot \delta_f[0,0]$ 또는 식 (4)를 통하여 검증될 수 있다.

$$SAD < \frac{B_{SIZE} \times Q_{STEP}}{c \cdot \sqrt{2} \times M[0,0]} \quad (4)$$

여기서, Q_{STEP} 은 현재 잔여신호 블록의 부호화에 사용될 양자화 스텝 크기(Quantization Step Size)를 의미하며, c는 AZCB 가설 테스트의 정확도 제어를 위한 신뢰 매개 변수이다. 예를 들어, c=1로 하면, 현재 잔여신호 블록이 AZCB라는 가설은 69.2%의 신뢰도를 가지게 되며, 이는 가우시안 분포로 모델링된 $[0,0]$ 번째 DC 계수가 $(-\delta_f[\mu, \nu], \delta_f[\mu, \nu])$ 범위 내에 69.2% 존재하게 되기 때문이다.

3. 알고리즘 구현

이 장에서는 본 논문에서 제안한 AZCB 예측(AZCB Prediction, AZCBP) 방법을 H.264 고속 모드 결정에 적용하기 위해 고려된 두 가지 사항 - 테이블 참조 기반 AZCB 고속 구현기법과 H.264의 최적 모드들의 전역 통계방법을 설명한다.

Step1: SKIP 모드에 대해 AZCBP 수행. 만약 조건 (4)를 만족시킨다면, SkipInter16x8, SkipInter8x16, SkipInter8x8, SkipIntra4x4를 1로 설정.

Step2: P16x16 모드에 대해 움직임 추정과 AZCBP 수행. 만약 조건 (4)를 만족시킨다면, SkipInterP8x8, SkipIntra4x4를 1로 설정.

Step3: 만약 SkipInterP16x8이 1이라면, step 7로 이동. 그렇지 않다면, P16x8 모드에 대해 움직임 추정과 AZCBP 수행. 만약 조건 (4)를 만족시킨다면, SkipIntra4x4를 1로 설정.

Step4: 만약 SkipInterP8x16가 1이라면, step 7로 이동. 그렇지 않다면 P8x16 모드에 대해 움직임 추정과 AZCBP 수행. 만약 조건 (4)를 만족시킨다면 SKIPIntra4x4를 1로 설정.

Step5: 만약 SkipInterP8x8이 1이라면, step 7로 이동.

Step6: 각 8x8 블록에 대해 아래의 sub-block 모드 결정을 4번 반복.

Step6.1: P8x8, P8x4, P4x8 모드에 대해 각각 움직임 추정과 AZCBP 수행. 만약 모드 중 어느 하나라도 조건 (4)를 만족시킨다면, SkipInter4x4를 1로 설정.

Step6.2: 만약 SkipInter4x4가 1이라면, step6.4로 이동.

Step6.3: P4x4 모드에 대해 움직임 추정 수행.

Step6.4: 만약 위의 과정을 수행하지 않은 8x8 블록이 남아 있다면, step6.1로 이동.

Step7: SkipFlag가 설정되지 않은 모든 모드들에 대해 비트율-왜곡치 연산 수행.

Step8: 최소의 비트율-왜곡치를 갖는 모드를 최적의 모드로 선택.

그림 1 제안된 모드 결정 알고리즘

Fig. 1 Implementation details of the proposed enhanced adaptive mode decision algorithm

우선, AZCB 테스트 수행에 사용되는 식(4)의 우변 값을 한계 SAD(Cut-Off SAD)라 하자. 이 한계 SAD는 신뢰 매개변수 c가 주어지는 경우 각 양자화 스텝 크기에 따라 미리 계산될 수 있다. 이 때, Bsize는 16x16으로 계산하여 양자화 스텝 크기를 인덱스로 하는 한계 SAD 테이블에 미리 계산된 값을 저장하고, 각 서브 블록 파티션의 SAD를 더해 16x16을 단위로 AZCB 가설 검증을 수행한다. 따라서 각 매크로블록 당 소요되는 AZCB 가설검증의 총 연산 요구량은 1번의 테이블 검색과 1번의 비교 연산에 불과하게 된다.

두 번째로, 본 논문에서 제안하는 모드 결정 방법의 효율성을 극대화하기 위해 최적 모드의 전역 통계 특성을 이용하였다. 이는, FMD를 수행하는 경우 최적 모드로 선택될 확률이 높은 매크로블록 모드를 먼저 테스트함으로써, H.264 최적 모드 선택 프로세스가 보다 빠르게 종료될 수

있도록 하여 부호화 속도를 높이려는 것이다. 참고문헌 [2-3]에서 수집된 최적 모드의 전역 통계 특성에 따라, 큰 블록 파티션으로 분할된 매크로블록 모드들이 먼저 움직임 추정 및 윌-왜곡치 연산의 대상이 되도록 하였는데, 그림 1에 자세한 알고리즘을 설명하였다. 그림 1의 제안된 모드 결정 방법에서, 인트라 4x4 모드는 오직 8x8 또는 이보다 큰 블록 파티션을 가지는 매크로블록 모드에서 AZCB가 감지되지 않았을 경우에만 수행되도록 하였는데, 이는 인트라 4x4 모드가 전역적인 통계상 가장 적게 선택되는 모드일 뿐 아니라, 오직 복잡한 경계선이나 새롭게 나타난 장면을 부호화할 경우에만 선택되는 모드이므로 일반적으로 큰 블록 사이즈로 움직임 추정이 잘 되지 않는다고 볼 수 있기 때문이다. 마지막으로, 제안된 그림 1의 모드 결정 방법에서 특정 모드가 AZCB로 판명되었을 경우 한 단계 더 낮은 레벨의 모드에 대해서도 비트율-왜곡치 연산을 수행하도록 하였는데, 이는 가설 검증 오차로 인한 부호화 품질 저하 문제를 최소화하는 동시에 RDO 측면에서 보다 적합한 모드가 선택될 수 있는 여지를 제공하기 위함이다.

4. 시뮬레이션 및 결과

제안된 알고리즘의 성능을 평가하기 위해, 다양한 조건 하에서 시뮬레이션을 수행하였다. 시뮬레이션에 사용된 공통 부호화 파라미터는 [7] 베이스라인 프로파일, CIF 사이즈 이미지의 ±32 크기 움직임 추정 탐색 범위(Search Range) 또는 QCIF 크기 이미지의 ±16 크기의 탐색범위, 고속 움직임 추정 [8], 하나의 움직임 추정 기준 영상, 하다마드 변환 사용 및 비트율-왜곡치 최적화 사용 등이다. 테스트 영상으로는 Foreman(F)(QCIF, 10, 150)¹⁾, Susie(S)(QCIF, 15, 150), Mobile(M)(CIF, 30, 300), Coastguard(C)(CIF, 15, 150)가 사용되었으며, 양자화 인덱스 28, 32, 36, 40에서 부호화/비교 되었다. 지면을 절약하기 위해, 본 논문에서는 제안된 알고리즘과 AMD 알고리즘의 성능만을 직접 비교하였다. 하지만 [1]에서 AMD 기법이 다양한 기존의 모드 결정 기법들과 비교되었기 때문에, 본 논문의 AMD와의 성능 비교는 [1]을 통하여 기존의 다른 모드 결정 기법들과의 객관적 상대 비교가 가능하다.

표 1에 본 논문에서 제안하는 알고리즘의 성능 평가 결과를 요약하였다. 성능 평가에 사용된 지표로는 평균 비트율 차이(DBR: Delta Bit-Rate), 평균 PSNR 차이(DPSNR: Delta PSNR) 및 평균 부호화 속도 향상비(SR: Speed-up Ratio) 등이 사용되었는데, DBR 및 SR의 경우에는 JM9.3 레퍼런스 코덱에 대한 AMD 및 제안 알고리즘의 성능 차이를 JM9.3을 기준으로 하는 퍼센트(%) 단위로 나타내었고, DPSNR의 경우에는 JM9.3에 대한 평균 PSNR 차이를 데시벨(dB) 단위로 나타낸 것이다. 즉, 표 1에서 Mobile 시퀀스를 QP=28로 부호화 한 경우 AMD 기법의 부호화 속도 향상비 8.92는 JM9.3에 동일한 공통 부호화 파라미터를 적용하여 부호화 했을 때 보다 8.92% 더 빠르게 부호화 된다는 의미이고, Mobile 시퀀스를 QP=40으로 부호화 하는 경우 제안 알고리즘의 평균 PSNR 차이 -0.03은 제안 알고리즘이

JM9.3에 비해 PSNR이 0.03dB 나쁘다는 의미이다.

표 1의 DBR과 DPSNR 결과는 본 논문에서 제안된 알고리즘(또는 AMD 알고리즘)의 비트율-왜곡 성능이 JM9.3의 FMD와 거의 유사하다는 것을 보여주며, 이것은 비트율-왜곡 측면에서는 거의 아무런 품질 저하가 발생하지 않는다는 것을 의미한다. 하지만, 부호화 속도 면에서는 제안된 알고리즘에 의한 전반적인 속도향상비가 대략 56.19% 정도 되며, 이는 AMD 알고리즘의 41.84% 보다 뛰어난 것으로 나타나고 있다. 여기서, SR은 FMD에 의한 상대적 속도 향상 이득값이기 때문에, 이는 제안된 방법이 FMD와 AMD 보다 각각 2.28배 그리고 1.33배 더 빠르다는 것을 나타낸다. 뿐만 아니라, 본 논문에서 제안하는 방법의 SR 이득값은 모든 실험 영상 및 양자화 파라미터에 대해 매우 높은 속도 향상 성능을 보여 주고 있으며, 이는 적절한 신뢰 매개 변수(식 (4) 참조)를 선택함으로써 연산 능력이 상이한 다양한 장치에서 보다 안정적인 윌-왜곡 성능을 제공할 수 있음을 나타낸다. 반면, 표 1의 결과로부터, AMD 알고리즘의 경우 장면이나 움직임이 복잡한 Coastguard 혹은 Mobile 시퀀스의 경우 다른 비디오 시퀀스에 비해 대략 2배 정도 속도 개선 성능이 떨어지는 것을 볼 수 있으며, 낮은 양자화 파라미터를 사용하는 경우에 있어서도, 마찬가지로, 높은 양자화 파라미터를 사용하여 쉽게 AZCB가 검출되는 경우에 비해 2배 이상 속도 개선 능력이 떨어지는 것을 볼 수 있다. 비트율 별로(양자화 파라미터 변화에 따라) 테스트된 AMD 알고리즘의 SR 변동 폭((최대SR - 최소SR)/평균SR)은 각 시퀀스에 따라 M-131%, C-122%, F-41%, S-45%에 달하며, 반면 제안된 알고리즘의 경우에는 M-20%, C-20%, F-5%, S-3%에 불과하다. 또한 테스트 영상별로 테스트된 AMD 알고리즘의 SR 변동 폭은 QP=28, 32, 36, 40에 대해 AMD 알고리즘이 각각 120%, 71%, 30%, 21%인 반면에, 제안된 알고리즘은 46%, 25%, 27%, 32%였다. 부호기 연산 장치의 설계에 있어 어떤 알고리즘의 채용은, 그 알고리즘이 제공하는 평균 부호화 속도와 동작 상황에 따른 부호화 속도의 변화 폭이 함께 고려된 최소 성능 기준에 따라 결정되는데, 이는 저비용, 고품질의 고속 부호화기를 설계에 있어, SR의 변동 폭이 매우 제한적인, 제안 알고리즘이 매우 적합한 기법임을 강조하는 것이다.

표 1 제안된 알고리즘과 AMD 알고리즘의 성능 비교

Table 1 Performance of the proposed algorithm and AMD

SEQ	QP	28 (c=3)		32 (c=4)		36 (c=5)		40 (c=6)		Avg.	
		AMD	Prop.	AMD	Prop.	AMD	Prop.	AMD	Prop.	AMD	Prop.
M	SR	8.92	39.73	26.32	48.96	42.51	47.73	51.12	46.66	32.22	45.78
	DPSNR	0.04	-0.04	0.01	-0.06	-0.03	-0.06	-0.05	-0.03	-0.01	-0.05
	DBR	0.80	-0.28	0.49	-0.39	-0.54	-0.48	-0.33	0.19	0.11	-0.24
C	SR	15.04	48.18	23.72	52.03	40.07	54.33	56.27	58.59	33.78	53.23
	DPSNR	0.01	-0.03	0.03	-0.05	-0.01	-0.04	-0.03	-0.02	0.00	-0.04
	DBR	0.59	-0.06	0.88	-0.18	0.34	0.03	-0.70	-0.05	0.28	-0.07
F	SR	39.60	59.98	47.62	61.88	54.09	61.99	60.12	62.94	50.36	61.70
	DPSNR	-0.01	-0.05	0.01	-0.07	-0.01	-0.06	-0.01	-0.07	-0.01	-0.06
	DBR	0.29	0.34	0.14	0.82	0.16	0.00	0.11	0.04	0.18	0.30
S	SR	40.01	64.22	50.01	63.22	50.91	63.55	63.07	65.07	51	64.02
	DPSNR	-0.01	-0.04	0.03	-0.05	0	-0.05	-0.03	-0.02	0.00	-0.04
	DBR	0.13	0.08	-0.3	-0.56	0.09	0.13	-0.38	-0.06	0.05	-0.10
Avg	SR	25.89	53.03	36.92	56.52	46.90	56.90	57.65	58.31		
	DPSNR	0.01	-0.04	0.02	-0.06	-0.01	-0.05	-0.04	-0.04		
	DBR	0.45	0.02	0.30	-0.08	0.01	-0.08	-0.33	0.03		

1) 테스트 영상 뒤의 괄호 2개는 다음의 정보를 나타낸다.

(sequence 이름)(크기, 부호화 프레임-율, 부호화 된 프레임 수)

감사의 글

본 연구는 서울시 산학연 협력사업(10654)의 지원에 의하여 이루어진 연구로서, 관계부처에 감사드립니다.

참 고 문 헌

- [1] Y.-H. Kim, J.-W. Yoo, S.-W. Lee, J. Shin, J. Park and H.-K. Jung, "Adaptive mode decision for H.264 encoder", *Electronics letters*, Vol.40, No19, pp.1172-1173, Sept. 2004.
- [2] Lidong Xu and Xinggang Lin, "Fast mode decision for inter frames in H.264/AVC", in *Proc. IEEE ISIT*, Oct. 2005, vol.1, pp.433-436.
- [3] Dongming Zhang, Yanfei Shen, Shouxun Lin, and Yongdong Zhang, "Fast inter frame encoding based on modes pre-decision in H.264", in *Proc. IEEE ICME*, July 2005, pp.530-533.
- [4] Zhu Hong, Wu Cheng-ke, Wang Yang-li and Fang Young, "Fast mode decision for H.264/AVC based on macroblock correlation", *AINA'05*, Mar. 2005, Vol.1 pp.775-780.
- [5] Joint Video Team(JVT) of ISO/IEC MPEG&ITU-T VCEG, "Joint Video Team(JVT) of ISO/IEC MPEG&ITU-T VCEG", JVT-F100d2.doc, 6th JVT meeting, Awaji Island, JP, 5-13 Dec. 2002.
- [6] I. M. Pao and M. T. Sun, "Modeling DCT coefficients for fast video encoding", *IEEE Trans. Circuits Syst. Video Technol.*, vol.9, pp.608-616, June 1999.
- [7] G. Sullivan and G. Bjontegaard, "Recommended simulation common conditions for H.26L coding efficiency experiments on low-resolution progressive-scan source material", *ITU-T VCEG, Doc. VCEG-N81*, Sept. 2001.
- [8] Z. Chen, P. Zhou and Y. He, "Fast integer pel and fractional pel motion estimation for JVT", *JVT-F017*, 6th JVT meeting, Awaji Island, JP, Dec. 2002.