

통계적 분석을 통한 고속 휘도 화면내 4x4 예측 모드 결정 방법

*노대영 **김지은 ***정광수 ****오승준

광운대학교 방통융합 플랫폼 센터

*kurtnoh@media.kw.ac.kr, **jieun@media.kw.ac.kr

A fast luminance intra 4x4 prediction mode decision method by statistical analysis

*Noh, Dae-Young **Kim, Ji-Eun ***Chung, Kwang-Sue ****Oh, Seoung-Jun

BnC Convergence Platform Center, Kwangwoon University

요약

율-왜곡 최적화 기법을 통한 화면내 예측 모드 결정 방법은 부호화 효율이 높지만 복잡도가 크다. 본 논문에서는 H.264/AVC의 율-왜곡 값의 특성을 분석하여 율-왜곡 최적화 기법에 사용되는 예측 모드의 후보를 줄임으로써 보다 빠른 휘도 화면내 4x4 예측 모드 결정 방법을 제안한다. 제안된 방법은 균일한 차분 블록값을 가지는 예측 모드와 균일하지 않은 차분 블록값을 가지는 예측 모드의 통계적 분석을 통해 부호화 속도를 향상시킨다. H.264/AVC의 참조 소프트웨어 JM 14.2와 비교하여 0.04[dB]라는 무시할 수 있는 PSNR의 손실을 가지면서도 0.3[%]의 비트율 절감과 19.6[%]의 부호화 속도 향상을 가져왔다.

1. 서론

비디오 압축 부호화 기술인 H.264/AVC는 ISO/IEC MPEG과 ITU-T VCEG 공동으로 결성된 JVT(Joint Video Team)에 의해 표준화가 완료되었고 부호화 효율이 높기 때문에 많은 비디오 부호화 장치에 사용되고 있다[1]. H.264/AVC의 화면내 예측 방법은 모든 예측 모드에 대하여 부호화 과정을 수행한 후 율-왜곡 최적화(RDO)기법을 이용하여 얻어진 율-왜곡 값(RDCost)이 가장 작은 모드를 최종 예측 모드로 선택하기 때문에 복잡도가 크다. 부호화 효율을 높이기 위하여 고속 휘도 화면내 예측 모드 결정 방법들이 연구되어왔지만 PSNR과 비트율의 성능을 감소시키는 결과를 가져왔다[2]. 따라서 H.264/AVC의 부호화 손실을 최소화하면서 부호화 속도를 증가시키기 위해 고속 휘도 화면내 4x4 예측 모드 결정 방법을 제안한다.

2. 본론

본 논문에서는 율-왜곡 값의 통계적인 특성을 분석하고 그 특성에 따라 율-왜곡 최적화 기법을 적용할 후보 모드들을 선출하여 부호화 속도를 향상시키기 위한 방법을 제안한다.

가. 율-왜곡 값의 통계적 특성 분석

H.264/AVC의 휘도 4x4 화면내 예측 모드를 결정하기 위한 율-왜곡 최적화 기법은 Lagrangian Cost Function을 이용하여 식(1)과 같이 계산되어진다[3].

$$J(s, c, IMODE | QP, \lambda_{MODE}) = SSD(s, c, IMODE | QP) + \lambda_{MODE} R(s, c, IMODE | QP) \quad (1)$$

s 와 c 는 각각 현재 부호화하는 4x4블록 휘도 성분의 원영상과 복원된 영상의 값을 나타낸다. QP 는 양자화 계수를 나타내며 $IMODE$ 는 화면내 예측 모드를 가리킨다. $SSD(\bullet)$ 는 원영상 블록과 복원된 블록의 제곱 차이 합으로써 왜곡값을 나타내고 $R(\bullet)$ 는 선택된 $IMODE$ 에 따른 비트량을 나타낸다. λ_{MODE} 는 모드 결정을 위한 lagrange 계수(multiplier)를 나타낸다.

식(1)과 같이 율-왜곡 최적화 기법은 비트율값과 왜곡값으로 구성된다. 이들을 분석하기 위해 식(2)와 식(3)을 정의하고 그 결과를 정규화한 후 대표적인 값을 그림 1에 나타낸다.

$$B_m = (b_m + 1.0) / (b_0 + 1.0) - 1.0 \quad (2)$$

$$D_m = (d_m + 1.0) / (d_0 + 1.0) - 1.0 \quad (3)$$

b_m 와 d_m 은 각각 오름차순으로 정렬된 비트율값과 왜곡값을 나타내며 m 은 H.264/AVC 화면내 모드 수로써 0부터 8까지의 값을 가진다.

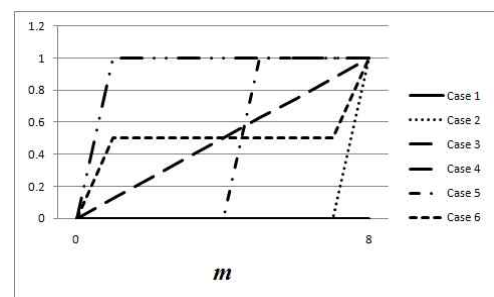


그림 1. 정규화된 비트율값과 왜곡값의 대표 경우

그림 1과 같이 예측 모드의 비트율값과 왜곡값은 동일한 부분(Case 1)

과 증가하는 부분(Case2-6)으로 나뉜다. 즉 예측 모드의 값이 균일한 부분과 균일하지 않은 부분으로 나누어 질 수 있다.

나. 균일한 값을 가지는 예측 모드

예측 모드들의 왜곡값이 동일하거나 유사한 경우, 울-왜곡 최적화 기법에서는 비트율을 줄이기 위하여 MPM(Most Probable Mode)을 최적 모드로 선택한다. 따라서 왜곡값이 유사한 모드들을 예측할 수 있다면 MPM만을 후보 모드로 하여 울-왜곡 최적화를 하면 부호화 속도를 향상시킬 수 있다[4].

부호화되는 블록의 왜곡값은 양자화 에러가 가장 큰 비중을 차지하며 식(4)와 같이 근사화할 수 있다[5].

$$D = (Q_{step})^2 / \epsilon \tag{4}$$

Q_{step} 은 양자화 구간의 폭이고 ϵ 은 12로 근사화가 가능하다.

예측 모드들의 왜곡값들간의 유사성을 표현하기 위한 측정자를 식(5)와 같이 정의한다.

$$\Delta_D = D_{max} - D_{min} \tag{5}$$

D_{max} 와 D_{min} 은 예측 모드들의 왜곡값 중에서 가장 큰 값과 가장 작은 값을 나타낸다. 식(4)와 식(5)를 이용하여 Δ_D 는 식(6)과 같이 나타낼 수 있다.

$$\begin{aligned} \Delta_D &= (Q_{step})^2 / \epsilon_{max} - (Q_{step})^2 / \epsilon_{min} \\ &= (\epsilon_{min} - \epsilon_{max}) (Q_{step})^2 / (\epsilon_{max} \times \epsilon_{min}) \\ &= (Q_{step})^2 / \epsilon_D \end{aligned} \tag{6}$$

예측모드들의 왜곡값이 유사한 경우, Δ_D 의 범위는 0부터 $(Q_{step})^2 / \epsilon_D$ 까지이고 울-왜곡 최적화 기법을 통해 선택되는 최적 모드는 MPM과 같다는 가설을 세운다. 가설 검증을 위해 울-왜곡 최적화 기법을 통해 선택된 최적 모드가 MPM인 경우와 MPM이 아닌 경우의 Δ_D 값의 확률 분포를 조사하였다. 그림 2는 위 가설의 결과를 나타낸 것이다.

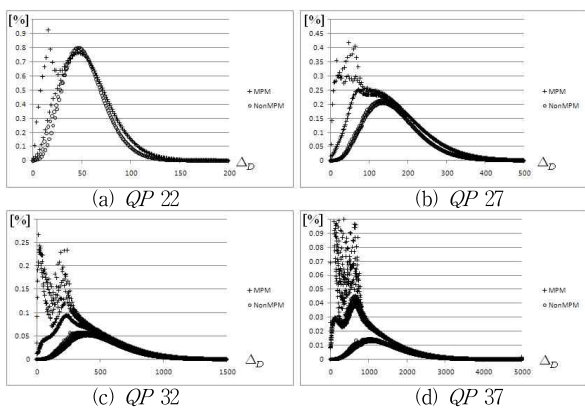


그림 2. Δ_D 의 확률 분포

Δ_D 값이 작을수록 MPM이 최적 모드가 될 확률이 크고 최적모드가 MPM인 경우와 MPM이 아닌 경우의 확률 밀도 함수가 교차하는 부분의 Δ_D 값은 QP가 커질수록 커진다는 것을 알 수 있다.

H.264/AVC의 부호화 속도 향상을 위해 식(5)에서 정의한 Δ 에 차분 블록값을 적용한 후(Δ_R) 그림 2의 확률 분포조사와 같은 방법으로 Δ_R 의 범위가 0부터 $(Q_{step})^2 / \epsilon_R$ 까지이면 울-왜곡 최적화 기법

을 통해 선택되는 최적 모드는 MPM과 같다는 가설을 설정할 수 있다. 이 가설의 검증을 위하여 Δ_R 값의 확률 분포를 조사하여 그림 3에 나타냈다. 그림 3에서 Δ_R 값이 약 $(Q_{step})^2 / 3$ 지점부터 MPM이 선택되지 않을 확률이 급격히 증가한다는 것을 알 수 있다. 따라서 실험적으로 Δ_R 이 $(Q_{step})^2 / 3$ 보다 작거나 같은 경우 최적 모드를 MPM으로 선정할 수 있다.

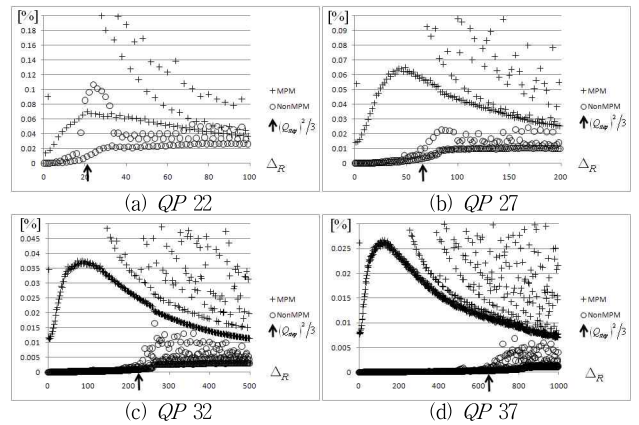


그림 3. Δ_R 의 확률 분포

다. 균일하지 않은 값을 가지는 예측 모드

영상이 균일하지 않은 경우 예측 모드들간의 왜곡값은 차이가 존재한다. 이 왜곡값과 차분 블록값간의 관계 여부를 조사하기 위한 측정자를 식(7)과 같이 정의한다.

$$\Psi_R^n = R'_n / R'_0 - 1 \tag{7}$$

R' 은 차분 블록값 R 을 오름차순으로 정렬한 값으로써 R'_0 가 가장 작고 R'_8 가 가장 크다. n 은 1부터 8까지의 값을 가진다.

통계적 조사를 하기 위해 식(8)과 같은 두 가지 가설을 설정한다.

$$\begin{aligned} H_0 &: Mode(R'_0) \neq Mode(\min RDCost) \\ H_1 &: Mode(R'_0) = Mode(\min RDCost) \end{aligned} \tag{8}$$

Mode는 입력값과 같은 값을 가지는 예측 모드를 나타내고 minRDCost는 울-왜곡 최적화 기법에 통해 얻어진 값들 중 가장 작은 값을 나타낸다. Ψ_R^n 과 설정된 가설과의 확률 분포를 조사하여 그림 4와 그림 5에 나타내었다.

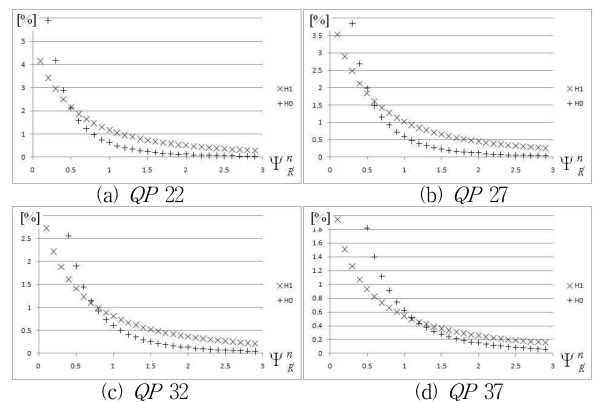


그림 4. Ψ_R^1 의 확률 분포

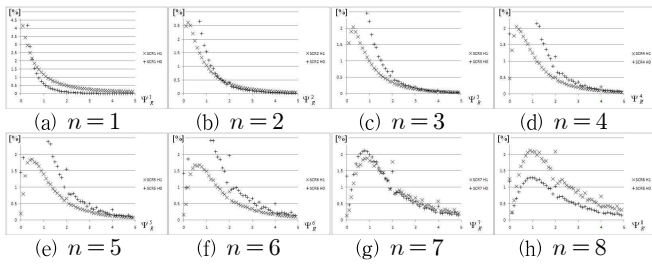


그림 5. Ψ_R^n 의 확률 분포

그림 4와 그림 5에서 Ψ_R^n 값이 커질수록 H_0 보다 H_1 의 확률이 커지는 것을 알 수 있다. 따라서 임의의 Ψ_R^n 값보다 작거나 같은 예측 모드들로 구성된 후보군을 윌-왜곡 최적화 기법에 사용하면 부호화 효율의 감소를 최소화 하면서 부호화 속도를 향상시킬 수 있다.

적절한 Ψ_R^n 값을 구하기 위하여 라플라스 분포 함수를 모델로 하는 LRT(Likelihood Ratio Test)진단법을 사용할 수 있다. 라플라스 분포 함수는 식(9)와 같고 LRT는 식(10)과 같다.

$$f_x(x) = \frac{1}{2\lambda} e^{-\frac{|x-\mu|}{\lambda}} \quad (9)$$

$$\ln A(\mathbf{R}) \geq \ln \eta \quad (10)$$

$$\frac{H_1}{H_0}$$

식(9)와 식(10)을 이용하여 식(11)을 유도할 수 있고 충분 통계량을 구할 수 있다.

$$\sum_{k=1}^K \mathbf{R}_k \geq \frac{H_1}{H_0} \frac{\lambda_1 \lambda_0}{\lambda_1 - \lambda_0} \left(\ln \eta - K \ln \left(\frac{\lambda_0}{\lambda_1} \right) - K \left(\frac{\mu_1}{\lambda_1} - \frac{\mu_0}{\lambda_0} \right) \right) \quad (11)$$

식(11)을 통해 구해진 값을 GOF(Goodness Of Fit)를 통해 모델링하면 식(12)와 같이 근사화가 가능하다.

$$\alpha Q_{step} + \beta \quad (12)$$

그림 6은 식(11)을 통해 구해진 값의 분포와 식(12)로 계산된 값의 분포를 보여준다.

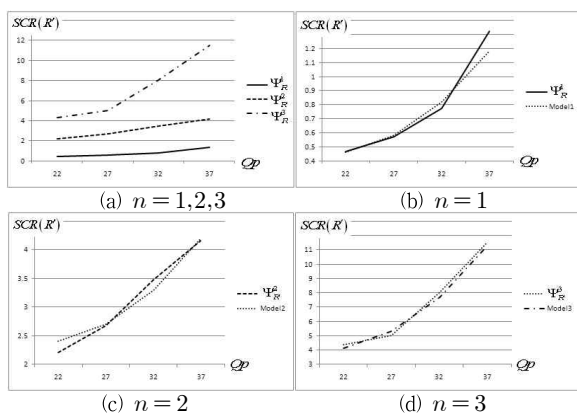


그림 6. LRT를 통해 구해진 임계값 및 근사화된 임계값

라. 효과적인 휘도 화면내 4x4 예측 모드 선택 과정

그림 7은 본 논문에서 제안하는 방법을 도식화한 흐름도이다.

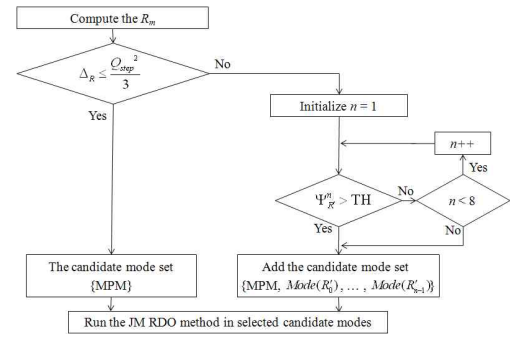


그림 7. 휘도 화면내 4x4 예측 모드 결정 방법

각 예측 모드별 차분 블록값을 구하고 이를 이용해 Δ_R 값을 구한다. 이를 오름차순으로 정렬한다. Δ_R 이 $(Q_{step})^2/3$ 이하이면 최종 모드를 MPM모드로 결정한다. 그렇지 않은 경우 Ψ_R^n 이 식(12)를 통해 구해진 임계값보다 작은 경우에 해당하는 예측모드들을 후보군으로 선정하고 후보군의 예측모드들과 MPM모드중에서 윌-왜곡 기법을 통해 최종 모드를 선정한다.

3. 실험 결과

제안된 방법의 성능을 평가하기 위한 실험 환경은 표 1과 같다.

표 1. 실험 환경

Parameter	Settings
Reference S/W	JM 14.2
Profile	Baseline
QP	22, 27, 32, 37
GOP Structure	III...
Sequences	CIF - Football, Foreman, Mobile, Silent, Coastguard 4CIF - City, Crew, Soccer, Harbour 720p - Bigships, City, Crew, Shuttlestart, Harbour
RDO mode	High Complexity Mode
VLC	CAVLC

제안된 방법의 성능을 검증하기 위하여 H.264/AVC 부호화기의 성능을 기준으로 하여 기존 고속화 방법[2]과 비교하였다. 검증 방법은 휘도 성분 PSNR, 비트율, 부호화 시간의 차이를 식(13)을 이용하여 비교하였고 그 실험 결과를 표 2에 정리하였다.

$$\Delta P = PSNRY_{proposed} - PSNRY_{JM14.2} [dB] \quad (13)$$

$$\Delta B = \frac{Bitrate_{proposed} - Bitrate_{JM14.2}}{Bitrate_{JM14.2}} \times 100 [\%]$$

$$\Delta T = \frac{Time_{proposed} - Time_{JM14.2}}{Time_{JM14.2}} \times 100 [\%]$$

표 2. JM 14.2기준 실험 결과

실험 영상		기존 방법 [2]			제안한 방법		
		ΔP [dB]	ΔB [%]	ΔT [%]	ΔP [dB]	ΔB [%]	ΔT [%]
CIF	Football	-0.05	0.31	19.36	-0.05	-0.30	17.67
	Foreman	-0.04	0.40	15.60	-0.04	-0.27	19.10
	Mobile	-0.05	0.32	17.46	-0.06	-0.09	19.16
	Silent	-0.05	0.36	15.28	-0.05	-0.29	17.19
	Coastguard	-0.05	-0.08	23.77	-0.05	-0.28	20.44
4CIF	City	-0.05	0.50	20.10	-0.04	-0.26	19.03
	Crew	-0.03	0.53	13.92	-0.03	-0.38	17.71
	Soccer	-0.05	0.12	18.63	-0.04	-0.35	17.86
	Harbour	-0.06	0.71	26.97	-0.04	-0.22	24.71
720p	Bigships	-0.04	0.40	18.81	-0.04	-0.29	18.77
	City	-0.05	0.62	20.40	-0.04	-0.29	18.86
	Crew	-0.03	0.47	16.54	-0.03	-0.38	19.45
	Shuttlestart	-0.03	0.38	13.80	-0.03	-0.50	20.07
	Harbour	-0.05	0.92	26.45	-0.04	-0.20	23.81
Average		-0.04	0.43	19.08	-0.04	-0.29	19.56

제안한 방법이 기존 방법에 비해 비트율은 0.72% 감소되었으며, PSNR의 변화 없이 부호화시간의 속도가 약 0.5% 향상되었음을 알 수 있다. 그림 8과 그림 9는 Silent 영상과 Soccer 영상에서 JM 14.2와 기존 방법, 그리고 제안한 방법의 RD곡선을 보여준다. 제안한 방법의 RD곡선이 기존 방법에 비해 JM 14.2 방법의 RD곡선에 훨씬 가깝게 위치한다.

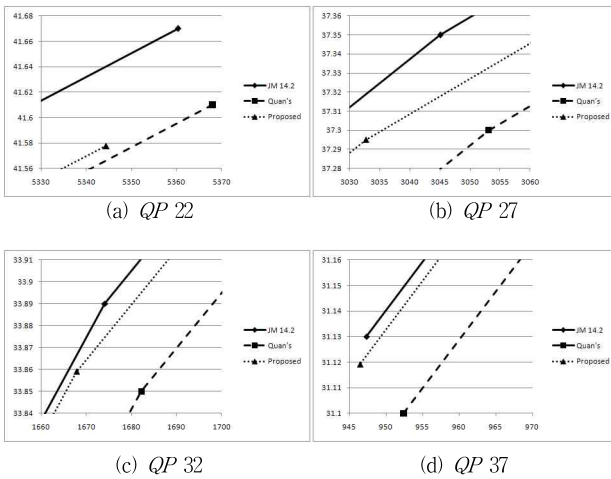


그림 8. Silent 영상의 RD곡선

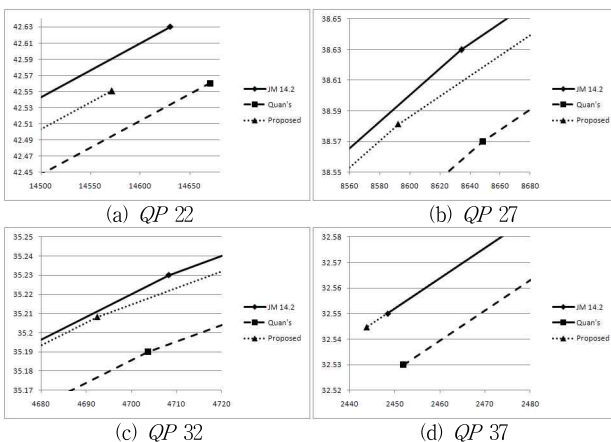


그림 9. Soccer 영상의 RD곡선

4. 결론

본 논문에서는 균일한 차분 블록값을 가지는 예측 모드와 비균일한 차분 블록값을 가지는 예측 모드의 통계적 특성을 분석하여 휘도 화면내 4x4 예측 모드의 후보를 줄임으로써 H.264/AVC의 부호화 속도를 향상시킬 수 있는 방법을 제안하였다. 제안한 방법은 JM 14.2 방법에 비해 화질의 감소가 거의 없이 평균 0.3%의 비트율 감소와 약 19.6%의 부호화 속도 향상을 가진다.

감사의 글

본 연구는 지식경제부 한국산업기술진흥원의 전략기술인력양성사업으로 수행된 결과임.

참고문헌

- [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] Do Quan, and Yo-Sung Ho, "Categorization for Fast Intra Prediction Mode Decision in H.264/AVC," IEEE Trans. Consumer Electronics, Vol. 56, No. 2, pp. 1049-1056, May 2010.
- [3] JVT-J049, "Joint model reference encoding methods and decoding concealment methods," December, 2003. -2.1.2.2.4
- [4] Dae-Yeon Kim, Jeong-Pil Kim, and Yung-Lyul Lee, "Fast Intra Prediction Mode Decision using Most Probable Mode for H.264/AVC" Journal of Broadcast Engineering, Vol. 15, No. 3, pp. 380-390, May 2010.
- [5] Ribas-Corbera, J. and Shawmin Lei, "Rate Control in DCT Video Coding for Low-Delay Communications" IEEE Trans. on Circuits and Systems for Video Technology, Vol. 9, No. 1, pp. 172-185, February 1999.