

H.264/AVC에서 화면간 예측 모드의 적응적 움직임 벡터 해상도 부호화 방법

*이주욱 **문주희

세종대학교 정보통신학과 정보통신연구소

*ljoblue@sju.ac.kr **jhmoon@sejong.ac.kr

An Adaptive Motion Vector Resolution Coding Scheme of Inter Prediction Mode in H.264/AVC

*Lee, Juock **Moon, Joo-hee

Dept. of Information & Communications Engineering,

Information & Telecommunication Research Institute, Sejong University

요약

기존의 비디오 부호화 표준에서는 참조영상을 보간하여 해상도를 증가시킨 후, 고정된 움직임 벡터 해상도로 영상 전체를 부호화 한다. 참조 영상의 해상도를 증가시킨 만큼 움직임 보상에 의하여 예측에러가 줄어들지만, 움직임 벡터 해상도가 증가한 만큼 움직임 벡터의 부호화 비트량이 증가한다. 고정된 해상도의 움직임 벡터로 부호화하는 경우, 영상의 지역적인 움직임 특성이 다른 경우 부호화 효율이 떨어질 수 있다. 따라서 본 논문에서는 기존의 비디오 부호화 표준들이 영상의 지역적인 특성을 고려하지 않고 고정된 해상도의 움직임 벡터를 사용하여 부호화하는 문제점을 극복하기 위하여 슬라이스 단위로 1/4 화소 해상도 또는 1/8 화소 해상도 또는 움직임 벡터 단위로 적응적으로 화소 해상도를 결정하는 것 중에서 최적의 슬라이스 움직임 벡터 해상도를 결정하여 부호화하는 방법을 제안한다. 제안한 방법을 사용하여 부호화하면 움직임 벡터의 부호화 비트의 낭비를 막고, 예측 에러도 줄어들어 부호화 효율을 높일 수 있다. 제안하는 방법을 사용하여 부호화 하는 경우 H.264/AVC와 비교하여 평균 1.97%의 BD-RATE를 감소한다.

1. 서론

기존의 비디오 부호화 표준들은 부호화 효율을 위해 움직임 보상 예측과 예측 에러를 부호화한다. 움직임 보상 예측을 수행하는 경우 성능을 높이기 위해 참조 영상의 해상도를 증가시켜 움직임 예측을 한다. 이러한 경우 예측 에러가 줄어서 텍스처 비트는 감소하지만 차분 움직임 벡터의 비트가 증가하는 문제점이 발생할 수 있다. 또한 해상도가 커질수록 메모리 사용량이 증가하고 움직임 예측을 수행하는 영역이 넓어지는 것과 같은 효과가 있으므로 움직임 예측 시간과 부호화 시간이 증가한다. 효율적인 부호화 성능과 복잡도 측면의 이유로 MPEG-4 part2[1] 및 H.264/AVC[2-3]에서는 가로와 세로를 각각 4배씩 보간을 수행하여 1/4 화소 해상도로 움직임 예측과 움직임 보상을 수행한다.

과거의 연구에서도 참조 영상의 해상도 및 보간 필터를 위한 많은 연구가 진행되었다. T. Wedi[4-6]는 H.26L에 1/2, 1/4, 1/8 화소 해상도로 움직임 예측을 수행하고 보간 필터를 비교했고, J. Ostermann[7]은 H.264에서 1/8 화소 해상도의 움직임 벡터로 움직임 예측 및 움직임 보상을 하고 참조 영상의 가로와 세로를 각각 8배로 해상도를 높이는 8탭 보간 필터를 ITU-T SG16 Q6 VCEG에 기고하여 KTA[8] 기술로 채택되도록 하였다. 그러나 J. Ostermann의 기술은 영상에 따라

성능의 차이가 컸으며, 특히 텍스처가 복잡한 영상에서는 성능이 좋으나 움직임이 큰 영상 또는 텍스처가 단순한 영상에서는 차분 움직임 벡터의 비트가 많이 발생하므로 부호화 효율이 감소하였다. 또한 큰 영상에서는 부호화 효율이 감소하여 차세대 비디오 부호화 표준[9]에 적절하지 않다.

최근에 진행되는 차세대 부호화 표준에서는 전술한 문제점을 해결하기 위하여 기존 비디오 부호화 표준들이 고정된 움직임 벡터 해상도를 사용하여 영상의 텍스처가 복잡한 영역 또는 움직임이 큰 영역의 지역적 특성을 반영하지 못하는 단점을 극복하기 위해 움직임 벡터 단위로 움직임 벡터의 해상도를 변경하는 기술이 발표되었다. 하지만 이는 움직임 벡터마다 해상도를 알려줘야 하기 때문에 부호화 효율이 감소하는 문제점이 있다. 이러한 문제점을 극복하기 위하여 본 논문에서는 슬라이스 단위로 1/4 화소 고정 해상도로 움직임 벡터를 부호화하는 것 또는 1/8 화소 고정 해상도로 움직임 벡터를 부호화하는 것 또는 움직임 벡터 단위로 적응적으로 해상도를 결정하여 움직임 벡터를 부호화하는 것 중 최적의 해상도를 결정하여 부호화하는 방법을 제안한다. 제안된 방법을 사용하여 부호화하면 움직임 벡터의 부호화 비트의 낭비를 막고, 예측 에러도 줄어들어 부호화 효율을 높일 수 있다.

2장에서는 고정된 해상도에 따른 성능에 대하여 설명 및 분석 하

고, 3장에서는 제안하는 방법에 대하여 설명하고, 4장에서는 실험결과를 분석하고, 마지막으로 5장에서 결론으로 끝을 맺는다.

2. 고정된 해상도에 따른 성능

표 1. 실험환경

SW	JM11KTA2.7
GOP 구조	IPPP
엔트로피 부호화 방법	CABAC
참조 영상의 수	4장
움직임 추정 영역	±128
움직임 추정 방법	EPZS[10]
QP	22, 27, 32, 37 (P slice +1)
울-왜곡 최적화 방법	고 복잡도 모드
부호화 영상의 수	100장
매크로블록 크기	16x16

표 2. 움직임 벡터 해상도에 따른 성능

Sequence		KTA(1/4) vs KTA(1/8)	
		BD-PSNR [dB]	BD-RATE [%]
Class D	RaceHorses	-0.41	8.06
	BlowingBubbles	0.00	0.07
	BasketballPass	-0.35	7.60
Class C	RaceHorses	-0.43	11.28
	PartyScene	0.05	-1.45
	BQMall	-0.35	8.19
Class B	BasketballDrill	-0.24	6.51
	BasketballDrive	-0.22	7.31
	Cactus	-0.19	6.72
	ParkScene	-0.29	8.07
	Kimono	-0.34	9.43
Total Average		-0.25	6.53

기존의 비디오 부호화 표준인 MPEG-4 part2 또는 H.264/AVC에서는 1/4 화소 고정 해상도로 움직임 벡터를 영상전체를 부호화하고, KTA 기술에서는 1/8 화소 고정 해상도로 움직임 벡터를 영상전체를 부호화한다. 이러한 과거의 표준들이 움직임 벡터 해상도에 따라 어떤 성능 차이가 있는지 확인하기 위하여 영상전체를 1/4 화소고정 해상도로 움직임 벡터를 부호화하는 KTA와 영상전체를 1/8 화소 고정 해상도로 움직임 벡터를 부호화하는 KTA와 비교하였다. 세부 실험 조건은 표 1과 같은 조건으로 실험 하였다. 표 2는 실험결과이다. 객관적 성능 측정법은 BD-RATE (Bjontegaard-Delta Rate)와 BD-PSNR (Bjontegaard-Delta PSNR)으로 측정하였다[11-12]. 표 2의 실험결과를 분석하여 보면, 대다수의 영상에서 1/8 화소 고정 해상도 부호화한 경우가 1/4 화소 고정 해상도로 부호화한 경우보다 성능이 떨어진다. 그러나 일부 영상에서는 1/8 화소 고정 해상도로 부호화 하는 것이 성능이 상당히 좋은 것을 확인 할 수 있다.

3. 제안하는 방법

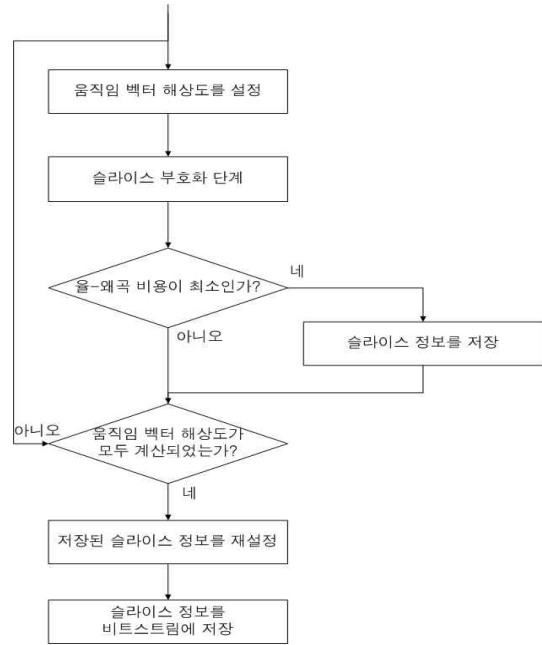


그림 1. 슬라이스 단위로 최적의 움직임 벡터 해상도를 결정하기 위한 순서도

기존의 비디오 부호화 표준과 같이 1/4 화소 고정 해상도 또는 1/8 화소 고정 해상도로 움직임 벡터를 영상전체를 부호화 한 결과 영상에 따라 성능차이가 나타났다. 기존의 비디오 부호화 표준들은 고정된 해상도로 움직임 벡터를 부호화하여 영상의 지역적 특성을 반영하지 않기 때문이다. 이러한 문제점을 해결하기 위해 본 논문에서는 슬라이스 단위로 움직임 벡터의 해상도를 1/4 화소 고정 해상도, 1/8 화소 고정 해상도, 움직임 벡터 단위로 해상도를 적용적으로 결정하여 움직임 벡터를 부호화하는 것 중 최적을 선택하여 부호화하는 방법을 제안한다.

제안하는 방법은 슬라이스 단위에서 1/4 화소 고정 해상도, 1/8 화소 고정 해상도와 움직임 벡터 단위로 적용적으로 해상도를 1/4 화소 또는 1/8 화소로 최적인 해상도를 결정하는 것 중 최적을 결정하므로 그림 1 과 같은 순서로 슬라이스의 최적의 움직임 벡터 해상도를 결정하여 부호화 한다. 첫 번째 단계로 1/4 화소 고정 해상도로 움직임 추정을 수행하고 움직임 보상을 하여 부호화 한 경우 슬라이스의 울-왜곡 비용(RDcost1)[13]을 계산한다. 두 번째 단계로 1/8 화소 고정 해상도로 움직임 추정을 수행하고 움직임 보상을 하여 부호화 한 경우 슬라이스의 울-왜곡 비용(RDcost2)을 계산한다. 세 번째 단계로 움직임 벡터 단위로 1/4 화소 해상도 또는 1/8 화소 해상도 중 최적의 해상도를 결정하여 부호화한 경우 슬라이스의 울-왜곡 비용(RDcost3)을 계산한다. 세 개의 울-왜곡 비용 중에서 작은 값의 움직임 벡터 해상도를 최적의 움직임 벡터 해상도로 결정하고, 최적의 해상도에 따른 슬라이스의 부호화 정보를 비트스트림에 출력한다. 각각의 움직임 벡터 해상도에 따라 움직임 벡터 차분값을 부호화하는 방법은 H.264/AVC와 동일하게 부호화 한다. 움직임 벡터 차분값을 계산하기 위하여 움직임 벡터 예측값을 계산하는 경우, 인접한 주변 블록의 움직임 벡터 해상도가 현재 블록의 해상도와 같지 않은 경우, 인접한 주변 블록의 움직임 벡터의 값을 곱셈, 나눗셈 및 반올림을 사용하여 변환한 다음 예측값으로 사용한다. 예를 들면, 현재 블록의 움직임 벡터가 1/8

화소 해상도이고, 인접한 주변 블록의 해상도가 1/4 화소 해상도인 경우, 주변 블록의 움직임 벡터를 곱하기 2를 하여 변환한 다음 미디언 필터링(Median Filtering) 하여 현재 블록의 움직임 벡터 예측값으로 사용한다. 또한, 슬라이스 단위의 움직임 벡터 해상도를 나타내는 정보는 슬라이스 헤더에 1 비트 또는 2 비트로 1/4 화소 고정 해상도, 1/8 화소 고정 해상도, 또는 움직임 벡터 단위로 해상도를 적응적으로 결정한 경우를 부호화한다. 또한, 움직임 벡터 단위로 해상도를 적응적으로 결정한 경우에는 움직임 벡터 단위로 해상도를 나타내는 정보를 1 비트로 부호화한다. 움직임 벡터의 해상도를 나타내는 정보는 Skip 매크로블록을 제외한 모든 인터 예측 블록에서 부호화한다.

4. 실험결과

표 5는 제안하는 방법을 사용한 성능이다. 제안하는 방법을 사용한 성능을 평가하기 위하여 2장에서 실험하였던 표 1과 동일한 환경으로 KTA의 1/4 고정 해상도와 비교하였다. 표 6의 실험결과를 통해 모든 영상에서 KTA의 1/4 화소 고정 해상도 보다 성능이 향상되었으며 평균 1.97%의 성능향상을 얻을 수 있었다.

표 3. 제안하는 방법의 성능

Sequence		KTA(1/4) vs 제안하는 방법	
		BD-PSNR [dB]	BD-RATE [%]
Class D	RaceHorses	0.02	-0.35
	BlowingBubbles	0.19	-4.60
	BasketballPass	0.04	-0.83
Class C	RaceHorses	0.01	-0.34
	PartyScene	0.22	-5.78
	BQMall	0.06	-1.32
	BasketballDrill	0.06	-1.59
Class B	BasketballDrive	0.03	-0.98
	Cactus	0.04	-1.19
	ParkScene	0.17	-4.59
	Kimono	0.00	-0.08
Total Average		0.08	-1.97

5. 결론

기존의 비디오 부호화 표준의 화면간 부호화 방법은 참조영상의 해상도를 1/4 화소 또는 1/8 화소로 고정된 해상도를 사용하여 움직임 추정 및 보상을 수행하였다. 그러나 고정된 해상도로 영상전체를 부호화하는 것은 항상 최적의 성능을 발휘하지 못하는 것을 확인 할 수 있었다. 따라서 제안하는 방법인 슬라이스 단위로 1/4 화소 고정 해상도, 1/8 화소 고정 해상도 또는 움직임 벡터 단위로 해상도를 적응적으로 결정하는 것 중에서 영상의 지역적인 특성에 따라 최적의 해상도를 결정하여 부호화 하는 방법을 제안하였다. 그리고 제안하는 방법을 사용하여 부호화하는 경우, 최대 5.78%, 최소 0.08%, 평균 1.97%의 BD-RATE 감소가 있었다.

참고문헌

- [1] Int. Standards Org./Int. Electrotech. Comm.(ISO/IEC) JTC 1, ISO/IEC 14496-2 (MPEG-4 visual version 1), Coding of Audio-Visual Objects Part 2: Video 1999.
- [2] ITU-T Recommendation H.264 and ISO/IEC 14496-10 AVC, "Advanced video coding for generic audiovisual services," version 3: 2005.
- [3] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560~576, 2003. 7.
- [4] T. Wedi, "Motion and aliasing compensated prediction for H.26L", ITU-T SG16 Q15-I-35, 1999, 10.
- [5] T. Wedi, "Hybrid video coding based on high-resolution displacement vectors", Visual Communications and Image Processing 2001, SPIE Vol. 4310, pp.186-193, 2001.
- [6] T.Wedi, "Adaptive interpolation filters and high-resolution displacements for video coding", IEEE Transactions and circuits and systems for video technology, Vol16, No.4., 2006. 4.
- [7] J. Ostermann and M.Narroschke, "Motion compensated prediction with 1/8-pel displacement vector resolution", ITU-T SG16 Q6 VCEG-AD09, 2006. 10.
- [8] K. Suehring, KTA Software Coordination, [Online]. Available: <http://iphone.hhi.de/suehring/tml/download/KTA/>
- [9] J. Ohm, G. Shullivan, V. Baroncini, "Draft call for proposals on high-performance Video Coding", ISO/IEC JTC1 SC29 WG11 N10926, 2009. 10.
- [10] A. M. Tourapis, "Enhanced predictive zonal search for single and multiple frame motion estimation", Visual Communications and Image Processing 2002(VICIP-2002, pp. 1069~1079, 2002. 1.
- [11] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves", ITU-T SG16 Q6 VCEG-M33, 2001. 4.
- [12] Stéphane Pateux, "Calculation An excel add-in for computing Bjontegaard metric and its evolution," ITU-T Q.6/SG16 Document VCEG-AE07, 2007. 1.
- [13] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression", IEEE Signal Proc. Mag., pp. 74~90, 1998.