스크린 컨텐츠 동영상 압축을 위한 적응적 색 공간 변환 기법

강 제 원
이화여자대학교 전자공학과

요 약

본 논문에서는 스크린 컨텐츠 동영상의 효율적인 압축을 위한 부호화 기법을 제안한다. 제안 부호화 기법은 RGB 색 요소 간의 신호의 통계적 중복성을 줄이기 위한 적응적 색 변환 기법에 기반한다. 기존의 색 변환 기법들은 스크린 컨텐츠 동영상이 포함하는 신호의 급격히 변화하는 톤계적 특성에 적용하기 어려운 반면 제안 방식에서는 부호화 블록의 시간적/공간적 주위 픽셀의 정보를 이용한 성분 분석을 통해 색 공간 변환 행렬식을 유도하여 부호화에 사용한다. 주위 픽셀 정보는 부호화 블록의 예측 방식에 따라 결정하며 움직임 벡터를 이용한 시간/변위 예측 부호 방식에서는 참조 블록의 정보를 이용하고 공간적 예측 방식에서는 블록의 인접 픽셀를 이용한다. 부호기 측에서는 비트율-왜곡 최적화를 통하여 예측 후 잔여신호의 부호화를 원래의 RGB 공간에서 수행할지 또는 색 변환한 공간에서 수행할지를 결정하고 관련 정보를 표시하여 블록 단위로 전송함으로써 부호화 효율을 증대한다. 실험에 의하면 제안 기법은 기존 압축 방식 대비 탁월한 부호화 성능을 제공함을 보인다.

1. 서론

스크린 컨텐츠 동영상은 컴퓨터의 합성에 의하여 제작한 동영상을 지칭하며 증강현실과 같은 새로운 멀티미디어 어플리케이션에 함께 사용할 수 있어 그 이용이 급속도로 증가하고 있다. 이와 같은 동영상은 원 색상의 정확한 표현을 위하여 RGB 각 색상 요소를 동일한 샘플링 비율을 이용한 RGB 4:4:4 동영상 형식으로 컴퓨터의 그래픽 장치로부터 취득한다. 일반 동영상의 경우 취득한 RGB 색상 요소들은 대게 YCbCr 색 공간으로 변환하여 압축을 수행하지만 스크린 컨텐츠 동영상의 경우 색 변환 과정에서 발생하는 왜곡이 컨텐츠의 화질에 심각한 열화를 야기할 수 있기 때문에 원래의 RGB 색 공간에서 주로 압축을 수행한다. 그러나 RGB 4:4:4 동영상 형식은 색상 간 표현 정보의 중복성이 크기 때문에 압축에 적합하지 않으며 동일한 샘플링 비율로 정보의 크기가 매우 커지는 단점이 있으므로 이를 해결하기 위한 효율적인 압축 기법의 개발이 요구되고 있는 실정이다.

본 논문에서는 스크린 컨텐츠 동영상을 압축하기 위하여 RGB 4:4:4 의 색 공간에서 중복되는 정보를 효율적으로 줄이기 위한 적응적 색 공간 변환 기법을 제안한다. 제안 기법에서는 부호화 블록의 시간적/공간적 주위 픽셀 정보를 이용하여 주성분 분석을 수행하고 이로부터 색상 간 연관 정보를 줄이기 위한 변환 행렬의 기저함수를 유도한다. 유도한 변환 행렬은 부호기의 비트율-왜곡 최적화 과정을 통하여 부호화 하는 블록의 색 공간을 결정하게 된다. 제안 기법은 최근의 동영상 압축 표준인 High Efficiency Video Coding (HEVC) [1]의 4:4:4 동영상 포맷을 위한 확장인 HEVC/RExt.[2]와 비교하여 탁월한 압축 성능을 보인다.

2. 기존 기법

최근의 동영상 압축 표준인 HEVC 는 일반 동영상에 대하여 기존 압축 표준 대비 우수한 부호화 성능을 보이지만 그 주요 부호화 기법들은 스크린 컨텐츠의 압축에 적합하지 않으므로 이에 관한 압축 표준 개발이 Moving Picture Expert Group (MPEG)에서 현재 진행 중이다 [3]. 현재 논의되고 있는 주요 기술로는 화면 내 블록 예측기법과 색 요소 간 예측 기법이 있다. 화면 내 블록 예측 기법[4]에서는 움직임 벡터를 사용하여 동일 화면 내에서 앞서 부호화한 블록의 정보를 예측 부호화하며 특히 컨텐츠의 반복되는 패턴이나 글자의 부호화에 효율적이다. 색 요소 간 예측 기법에서는 비율 변수를 이용하여 앞서 부호화한 색상 요소로부터 다른 요소를 예측하고 그 잔여신호와 비율 변수를 복호기로 전송한다 [5].

색 변환 기법은 원 영상을 표현하는 색 공간을 새로운 공간으로 변환하는 방식으로 적은 계산량 대비 높은 성능으로 얼굴인식과 같은 다양한 영상 어플리케이션에 활용되어왔다 [6]. 압축에 관해서는 YCgCo 색 변환 행렬 [7]이 4:4:4 동영상 형식을 부호화하는데 효율적으로 알려져 있으며 H.264/AVC 의 확장 표준 및 현 개발 중인 스크린 컨텐츠 압축 표준에도 최근 채택이 되었다. 그러나 기존의 색 변환 기법은 고정 값을 갖는 행렬 요소를 사용하여 스크린 컨텐츠 동영상이 포함하는 공간 상에서 통계적으로 급격하게 변하는 정보에 범용적으로 대응하기에 어려움이 있다. 색 변환 행렬의 적응적인 이용을 위하여 Marpe *et al.* 은 블록 기반의 비트율-왜곡 최적화를 통하여 서로 다른 변환 행렬을 적용하는 기법을 제안하였다 [8]. 그러나 행렬식은 여전히 고정 요소를 갖기 때문에 선택한 행렬식이 최적이라고 하기 힘들다. Kawamura *et. al* 은 Karhunan-Loeve (K-L) 변환을 이용하여 행렬식을 유도하였다 [9]. 그러나 K-L 변환의 경우 신호가 가우시안 통계 분포를 따르지 않을 경우 최적의 변환 부호 성능을 보이지 않으며 심지어 부호화 성능에 심각한 열화를 가져올 수 있기 때문에 부호화 성능이 뛰어나지 않았다.

3. 제안 기법

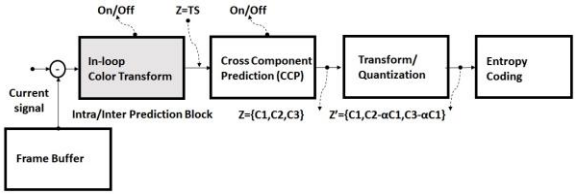

그림 1: HEVC/REext. 프레임웍에 적용한 제안 알고리듬

제안하는 압축 기법에서는 예측 후 잔여 신호를 색 공간 변환 기법을 사용하여 색상 요소 간의 통계적 중복성을 감소시키고 이를 기존의 부호화 방식에 따른 일련의 이산 코사인 변환, 양자화, 엔트로피 부호화를 통하여 부호화한다. 그림 1 은 HEVC/REext.의 부호기 프레임웍에 제안 기법을 통합하여 보인 블록다이어그램으로 예측 후 잔여 신호 $s_r \in \mathrm{R}^3$ 를 변환하고 새로운 색 공간에서의 3 차원 벡터 계수 $z \in \mathrm{R}^3$ 를 사용하여 부호화함을 보인다.

$$z = T_{3\times3} s_r$$

이 때 $T_{3\times3}$ 은 3x3 의 행렬식이다.

제안 기법에서는 위의 행렬식을 주성분 분석을 이용하여 유도한다. 주성분 분석에 의하면 변환 행렬식의 기저함수가 서로 연관되지 않도록 하는 선형 행렬식을 구할 수 있다. 주성분 분석을 유도하기 위하여 사용하는 주위 샘플 픽셀은 부호화 블록에 적용한 예측 기법에 기반하여 결정한다. 즉 공간적 예측을 수행하는 경우에는 공간 상의 주위 픽셀을 사용하고 시간적 예측 또는 화면 내 블록예측을 수행하는 경우에는 시간 상의 주위 픽셀을 사용한다. 이를 그림 2 에 보였다. 원본 신호를 s 로 예측 신호를 p 로 표현하였다.

예측 신호로부터 얻은 L 개의 샘플로부터 주성분을 구하기 위하여 평균 제곱 오차 기반의 식 (1)의 비용함수를 정의한다. 이 때 T 는 변환행렬이고 t_i 는 행렬식의 기저 벡터로 서로 직교한다.

$$J = \sum_{i=1}^{L} w(i) \,|\, p(i) - \sum_{j=1}^{3} t_j^{\,T} p(i) t_j \,|^2, \cdots (1)$$

이 때 $w(i)$는 i번째 샘플의 선택함수로 1 이면 해당 샘플을 주성분 분석에 포함 그렇지 않으면 포함하지 않도록 한다. 식 (1)의 비용함수는 다시 다음과 같이 식 (2)로 표현한다.

$$J = \mathrm{trace}(\mathrm{C_x}) - \sum_{j=1}^{3} t_j^{\,T}\, \mathrm{C_x}\, t_j \;, \cdots (2)$$

이 때 $\mathrm{C_x}$ 는 선택함수에 의한 샘플 x의 공분산 행렬이다. 식 (2)의 비용함수를 최소화하는 기저벡터는 $\mathrm{C_x}$ 를 행렬 분해하여 얻는 고유 벡터 e_i 의 정규화로 다음과 같이 구한다.

$$U^T \mathrm{C_x} U = I \;, \cdots (3)$$

여기서 $U = (\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3)$ 이고, 구하고자 하는 변환 행렬식은 고유 값의 기하평균인 λ 로 정규화하여 다음과 같다.

$$T = \lambda U^T \cdots (4)$$

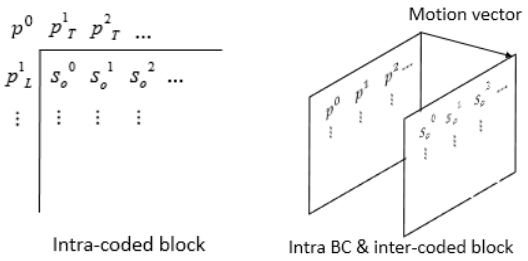

그림 2: 주성분분석을 위한 예측 방식에 따른 주위 픽셀의 선택. (a) 공간 예측 기법 및 (b) 공간/시간적 움직임 예측 기법.

그림 3 은 RGB 색상 공간에서 취득한 알파벳 H 의 32x32 블록의 잔여 신호를 제안 기법을 이용하여 변환하였을 때 얻은 계수를 보인다. 그림에서 픽셀 값이 밝아질수록 더욱 큰 값을 의미하며 RGB 공간에서 연관되어있는 정보는 제안 기법에 의하여 연관성이 줄었음을 알 수 있다.

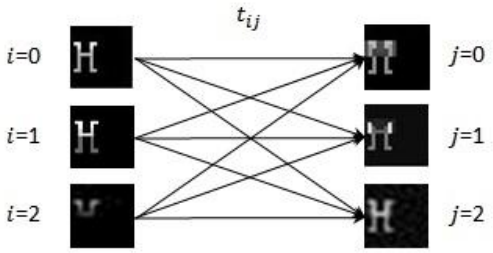

그림 3: 왼쪽 열에 표현한 RGB 색 공간에서의 픽셀 값과 오른쪽 열에 표현한 색 공간의 계수 값 비교.

제안 기법에서는 변환 행렬식을 HEVC 의 coding unit (CU) 마다 적응적으로 사용할 수 있도록 제어 플래그를 함께 부호화하여 전송한다. 즉, 부호기 측에서 비트율-왜곡 최적화를 통하여 최적의 변환 행렬식을 앞서 부호화한 시간적/공간적 주위 픽셀 정보를 이용하여 유도하고 그 결과를 플래그에 표시한다. 복호기 측에서는 수신한 플래그가 활성화 되어있는 경우 부호기에서 사용한 동일한 변환 행렬식을 유도하고 잔여신호의 색상 공간을 기존의 RGB 색 공간으로 역변환하여 복호화를 수행하며 그렇지 않은 경우는 역 변환을 생략하고 복호화를 수행한다. 이를 위하여 식 (5)와 같이 비트율-왜곡 최적화를 통하여 부호화 성능을 향상시키는 최적의 색 변환 행렬 T^* 을 사용한다.

$$T^* = \arg\min_{T\in\{T_D, \mathrm{I}\}} \{ |\, \mathrm{s} - T^{-1} \mathrm{z_q} \,| + \kappa(\mathrm{R_C} + \mathrm{R_H}) \}, \cdots (5)$$

이 때 s 는 원본 신호, z 는 양자화한 계수, T_D 는 유도한 색 변환 행렬식, I 는 단위행렬, κ 는 Lagrangian 상수, R_C 는 잔여신호의 부호화에 필요한 비트 수, R_H 는 플래그의 부호화에 필요한 비트 수이다.

3. 실험 결과

제안 기법의 효율적인 부호화 성능을 보이기 위하여 제안 알고리듬을 HEVC/RExt.의 참조 소프트웨어 모델인 HM13.0/RExt6.0 에 구현하여 실험을 하였다. 실험에서는 화면 내 예측 부호화 기법과 색 요소 간 예측기법을 활성화하였다. 테스트 동영상으로는 표 1 에서 보듯이 현재 스크린 컨텐츠 동영상 표준에서 사용하고 있는 다양한 해상도의 RGB 4:4:4 샘플 동영상을 사용하였다. 성능 평가의 구성은 부호화 성능 측정에서 보편적으로 사용하는 세가지 방식으로 각각 모든 동영상을 화면 내 부호화하는 방식, 일정 프레임 마다 화면 내 부호화 프레임을 삽입하여 임의 접근을 용이하게 하는 방식, 그리고 첫 프레임만을 화면 내 부호화하는 방식으로 구성하였다. 이를 각각 ‘ AI,’ ‘ RA, 그리고 ‘ LD’로 표 1 에 표현하였다.

표 1: 제안 기법의 HEVC/RExt 대비 부호화 성능 향상. 동일 PSNR 에서의 비트율 감소를 BD-rate 으로 표현.

해상도	테스트 동영상	AI	RA	LD
1080p	MissionControl	−17.2%	−18.5%	−17.7%
	Desktop	−11.2%	−6.3%	−5.9%
	Console	−9.8%	−6.2%	−4.4%
	Social Network	−17.3%	−19.1%	−19.2%
720p	Map	−15.6%	−15.3%	−15.4%
	Programming	−16.7%	−17.5%	−21.8%
	SlideShow	−23.8%	−23.8%	−22.0%
	Viking	−24.7%	−22.8%	−20.3%
평균		**−17.4%**	**−16.2%**	**−15.8%**

제안 기법의 HEVC/RExt. 대비 부호화 성능 향상을 동일 PSNR 에서의 비트율 감소 (BD-rate 감소)를 이용하여 측정하여 표 1 에 보였다. 표에서 보듯이 제안 기법은 기존 기법 대비 평균 약 16%의 비트율 감소를 통한 부호화 성능 향상을 보였다. 또한 제안 기법은 그림 4 에서 보듯이 주관적 화질 측면에서도 동일 비트율에서 기존 부호화 방식 대비 우수한 화질 개선을 보였다.

(a)

(b)

그림 4: (a) 기존 부호화 방식 및 (b) 제안 기법에 의한 주관적 화질 비교.

4. 결론

본 논문에서는 예측 후 잔여 신호의 색 간 정보의 중복성을 제거하기 위하여 적응적 색 공간에서 부호화를 수행할 수 있도록 하는 기법을 제안하였다. 제안 기법에 따르면 비트율-왜곡 최적화를 통하여 잔여신호는 RGB 원 색상 공간에서 부호화하거나 예측 방식에 따라 결정한 시간적 또는 공간적 주위 픽셀의 정보를 주성분 분석을 통하여 얻은 적응적 색 변환 행렬식을 통하여 부호화를 수행할 수 있다. 제안 알고리듬에 의하면 기존 부호화 방식 대비 약 16%의 동일 PSNR 대비 비트율 절약을 하도록 부호화 성능을 향상하였다.

5. 참고문헌

[1] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, “Overview of the High Efficiency Video Coding Standard,” *IEEE Trans. Circuits Syst. Video Tech.*, vol. 22, no. 12, pp. 1649- 1668, Dec. 2012.

[2] G. Sullivan, J. M. Boyce, Y. Chen, J. R. Ohm, C. A. Segall, and A. Vetro, “Standardized extensions of high efficiency video coding (HEVC),” *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, p. 1001-1016, Dec. 2013.

[3] ISO/IEC/JTC1/SC29/WG11 and ITU-T SG16 Q.6, “Joint call for proposals for coding of screen content,” Jan. 2014.

[4] C. Pang and J. Sole and L. Guo and M. Karczewicz and R. Joshi, “JCTVC-N0256: Non-RCE3: Intra Motion Compensation with 2-D MVs,” Aug. 2013.

[5] W. Pu, W.-S. Kim, J. Chen, K. Raspaka, L. Guo, J. Sole, and M. Karczewicz, “JCTVC-N0266: Non RCE1: Inter color component residual prediction,” in *ISO/IEC/JTC1/SC29/WG11 and ITU-T SG16 Q.6*, Aug. 2013.

[6] R. Hsu, M. Abdel-Mottaleb, and A. Jain, “Face detection in color images,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 5, p. 696-706, Aug. 2002.

[7] H. Malvar and G. Sullivan, “JVT-I014: YCoCg-R: a color space with RGB reversibility and low dynamic range,” in ISO/IEC/JTC1/SC29/WG11 and ITU-T SG16 Q.6, Sept. 2003.

[8] D. Marpe, H. Kirchhoffer, V. George, P. Kauff, and T. Wiegand, “An adaptive color transform approach and its application in 4:4:4 video coding,” in *Proc. EUSIPCO*, 2006, pp. 2005- 2008.

[9] K. Kawamura, T. Yoshino, and S. Naito, “JCTVC-L0371: In-loop color-space transformation of residual signals for range extensions,” in ISO/IEC/JTC1/SC29/WG11 and ITU-T SG16 Q.6, Jan. 2013.