

레터논문 (Letter Paper)

방송공학회논문지 제20권 제5호, 2015년 9월 (JBE Vol. 20, No. 5, September 2015)

<http://dx.doi.org/10.5909/JBE.2015.20.5.782>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

스크린 콘텐츠 비디오의 압축을 위한 인접 화소의 가중 합을 이용한 적응적 Residual DPCM 기법

강 제 원^{a)†}

Adaptive Residual DPCM using Weighted Linear Combination of Adjacent Residues in Screen Content Video Coding

Je-Won Kang^{a)†}

요 약

그래픽스 등 컴퓨터로 합성한 영상을 일컫는 스크린 콘텐츠 비디오의 경우는 색대비가 높은 그래픽 요소로 인하여 예측 후에도 잔차 신호의 공간적 연관성이 여전히 높게 나타나므로 현 샘플의 가장 인접한 화소를 이용하는 residual differential pulse-code modulation (RDPCM) 기법을 효율적으로 적용할 수 있다. 본 논문에서는 잔차 신호의 주위 인접 화소의 가중 합으로 보다 정확한 예측의 RDPCM을 수행하는 부호화 기법을 제안한다. 각 가중 계수는 현재 잔차 신호의 인접 영역에서 왜곡 값에 L_1 정규화를 포함한 비용함수를 최소화 하여 추정한다. 제안 기법은 스크린 콘텐츠 동영상 압축 참조 코덱 대비 약 3.1%의 부호화 성능 향상을 보인다.

Abstract

In this paper, we propose a novel residual differential pulse-code modulation (RDPCM) coding technique to improve coding efficiency of screen content videos. The proposed method uses a weighted combination of adjacent residues to provide an accurate estimate in RDPCM. The weights are trained in previously coded samples by using an L_1 optimization problem with the least absolute shrinkage and selection operation (LASSO). The proposed method achieves BD-rate saving about 3.1% in all-intra coding.

Keyword : HEVC, screen content video coding, RDPCM

a) 이화여자대학교 전자공학과(Ewha W. University)

† Corresponding Author : 강재원(Je-Won Kang)

E-mail: jewonk@ewha.ac.kr

Tel: +82-2-3277-2347

ORCID: <http://orcid.org/0000-0002-1637-9479>

※ 이 논문은 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2014R1A1A2056587).

· Manuscript received August 20, 2015; Revised September 16, 2015;

Accepted September 16, 2015.

1. 서 론

최근 컴퓨터 합성 기술 및 디지털 콘텐츠 제작 기술의 발달과 더불어 컴퓨터의 합성에 의하여 제작한 스크린 콘텐츠 비디오가 새로운 멀티미디어 서비스에 널리 사용되고 있다^[1]. 그러나 스크린 콘텐츠 비디오는 일반 비디오와 달리 색상 대조비가 높은 그래픽, 문자, 숫자 등을 다수 포함

하고 있어 HD/UHD 방송을 위하여 개발이 된 High Efficiency Video Coding (HEVC) 동영상 압축 표준^[2]과 같은 기존의 압축 기술로는 효율적인 압축이 용이하지 않다. 이에 따라 동영상 압축에 관한 국제 표준화 기구인 JCT-VC (Joint Collaborative Team on Video Coding)에서는 스크린 콘텐츠 비디오의 압축 성능 제고를 위하여 스크린 콘텐츠 비디오 압축 표준^[3](HEVC/SCC)을 제정 중으로 스크린 콘텐츠 비디오의 특성을 효율적으로 이용하는 부호화 기술을 개발 중이다.

Residual DPCM (RDPCM)은 비디오 신호의 예측 후 남은 잔차 신호를 DPCM하는 방식으로 공간상에서 가장 인접한 화소로부터 예측 부호화를 수행하는 무손실 압축 기법이다^[4]. 본 논문에서는 보다 정확한 예측을 수행하기 위한 방식으로 주위 잔차 성분의 가중 합의 선형 조합으로 DPCM을 수행하는 손실 압축에서의 새로운 방식의 RDPCM 기법을 제안한다.

II. RDPCM

HEVC 확장표준에서는 손실 압축의 경우 transform unit (TU)가 변환 생략 (transform skipping, TS) 모드로 부호화할 때 잔차 신호에 대해서 RDPCM을 수행하는 방식이 개발되었다. 변환 생략 모드는 잔차 신호를 변환하지 않고 직접 엔트로피 부호화하는 기법으로 색대비가 높은 그래픽 요소의 경계에서 발생하는 높은 주파수 영역 대의 잔차 성분이 많아 스크린콘텐츠 비디오 압축에 유용하게 사용할 수 있다. RDPCM은 변환 생략 과정에서 엔트로피 부호화를 위한 잔차 성분의 에너지 총량을 줄여 보다 뛰어난 압축 성능을 제공하도록 한다.

HEVC/SCC의 RDPCM은 수평 및 수직 방향의 화면 내 예측 후 동일 방향으로 RDPCM을 수행하는 implicit RDPCM 방식^[6]과 화면 간 예측 후 수평 (H-RDPCM) 혹은 수직 (V-RDPCM) 방향의 RDPCM을 선택적으로 수행하고 RDPCM 예측 방향 정보를 비트스트림을 통해 복호 측으로 전송하는 explicit RDPCM 방식^[7]을 이용한다. 이 때 앞서 부호화한 샘플 중 RDPCM의 예측 방향에 따라 가장 인접

한 왼쪽 열 혹은 상위 행의 잔차 성분을 이용하여 예측을 수행한다. $N \times N$ 크기 블록의 잔차 신호 $r_{i,j}$ 에 대하여 수직 방향의 RDPCM 적용 후 남은 2차 잔차 신호 $\widetilde{r}_{i,j}$ 는 수식으로 다음과 같이 표현할 수 있다.

$$\widetilde{r}_{i,j} = \begin{cases} r_{i,j} & , i = 0 \\ r_{i,j} - Q(r_{i-1,j}) & , 0 < i < N \dots \end{cases} \quad (1)$$

이 때 $Q(r)$ 은 양자화 잡음을 포함하는 복원한 잔차 신호이다. 식 (1)에서 보듯이 수직 방향 RDPCM의 경우 부호기측에서는 2차 잔차신호 $\widetilde{r}_{i,j}$ 를 엔트로피 부호화 후 복호기측으로 전송하고 다음 행의 잔차 신호 예측을 위하여 복원하여 예비한다. 해당 예측 과정은 블록의 모든 행에 대해서 순차적으로 진행한다. 반대로 복호 측에서는 식 (2)에서 보듯이 복원한 2차 잔차 신호를 순차적으로 더하여 현재 i 번째 행의 잔차 신호를 다음과 같이 복원한다.

$$Q(r_{i,j}) = \sum_{k=0}^i Q(\widetilde{r}_{k,j}), \quad 0 < i < N \dots \quad (2)$$

III. 제안 기법

기존 기법에서는 수평 혹은 수직 방향의 RDPCM 예측을 수행한 반면 제안 기법에서는 인접하는 잔차 성분의 가중 합으로부터 보다 정확한 예측을 수행하는 RDPCM 예측 방식을 제안한다. 부호기 측에서는 비트율-왜곡 최적화를 통하여 제안 RDPCM 기법을 수행할지 여부를 결정하여 비트스트림으로 전송하는 explicit RDPCM 방식을 따른다.

제안 기법의 예측은 일반적으로 식 (3)과 함께 표현할 수 있다.

$$\widetilde{r}_{i,j} = \begin{cases} r_{i,j} & , i = 0 \\ r_{i,j} - \sum_{(k1,k2) \in K} \alpha_{i-k1,j-k2} Q(r_{i-k1,j-k2}) & , o.w \dots \end{cases} \quad (3)$$

이 때 $k1$ 와 $k2$ 는 현재 부호화를 수행하는 잔차 신호 r_{ij} 의 예측 잔차 신호의 위치 인덱스이고, K 는 앞서 부호화를

완료한 인접 영역, 그리고 $\alpha_{i-k_1, j-k_2}$ 는 각 예측 성분의 가중 계수를 의미한다. 예측 신호를 선택하기 위한 영역 K 는 부호화 성능, 복잡도, 병렬 처리 등 여러 구현의 용이성의 trade-off를 고려하여 결정할 수 있는데 본 논문에서는 예측 신호의 위치를 좌측 ($r_{i, j-1}$)과 상측 ($r_{i-1, j}$), 좌상측 ($r_{i-1, j-1}$), 우상측 ($r_{i-1, j+1}$) 샘플의 집합을 K 로 선택하고 그에 따른 계수를 구한다.

식 (3)의 예측 기법은 기존 기법 대비 보다 많은 샘플을 이용하여 예측의 정확도를 증가할 수 있지만 해당 계수를 전송하는 것은 부호화 효율 측면에서 좋지 않다. 따라서 가중 계수를 인접 영역에서 추정한다. 먼저 현재 신호 $r_{i, j}$ 의 가중치 학습을 위한 샘플 영역 Ω_{ij} 은 그림 1과 같이 정의한다. 실험에서는 $M=L=2$ 로 정하였다. 추정에 사용하는 $r_{\Omega_{ij}}^l \in \Omega_{ij}, l=1, \dots, |\Omega_{ij}|$ 의 잔차 신호 샘플에 대해서 왜곡 값 $D_{\Omega_{ij}}$ 는 다음 식과 같이 정의한다.

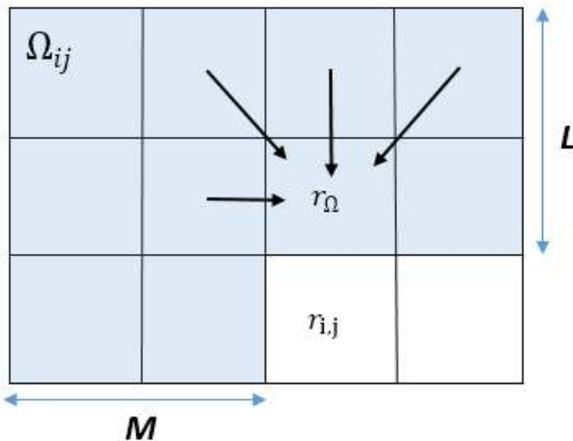


그림 1. 제안 모델의 계수를 추정하기 위한 학습 영역 Ω
 Fig.2. Previously coded samples used for estimating the weights

$$D_{\Omega_{ij}} = \sum_{l=1}^{|\Omega|} e^{-d(l)} |r_{\Omega_{ij}}^l|^2 \dots \quad (4)$$

이 때 $d(l)$ 은 현재 샘플의 위치와 l 번째 샘플의 위치 차이로 2차원 좌표계의 L_2 distance이고, 샘플 영역에서의 2차 잔차 신호 $r_{\Omega_{ij}}^l$ 는 다음 수식과 같다.

$$\widetilde{r_{\Omega_{ij}}^l} = Q(r_{\Omega_{ij}}^l) - \sum_{(k_1, k_2) \in K} \alpha_{i-k_1, j-k_2} Q(r_{\Omega_{i-k_1, j-k_2}}^l) \dots \quad (5)$$

따라서, 모델 계수 $\alpha_{k_1, k_2} (= \alpha_{i-k_1, j-k_2})$ 는 식 (4)를 최소화 하도록 다음과 같이 구할 수 있다.

$$\alpha_{k_1, k_2}^* = \operatorname{argmin}_{\alpha_{k_1, k_2} \in A} D_{\Omega_{ij}} \dots \quad (6)$$

식 (6)을 풀기 위한 최적화 방식이 몇 가지 있다. 그 중 최소제곱법 (least square optimization)은 간단한 계산량으로 closed-form의 해를 제공하기 때문에 효율적으로 사용할 수 있다. 그러나 최소 자승법은 이상치(outlier)의 근사에 취약하며 샘플 데이터 사이에 강한 공선 관계(multicollinearity)가 발생할 때 변수의 해석력이 저하되는 문제가 있다. 스크린 콘텐츠의 경우 그래픽 요소의 경계에서 발생하는 잔차 신호는 공간 상에서 인접 화소 사이에서 변화량이 매우 크고 경계 안에서는 작아지는 계단 함수 형태의 신호 특성 상 최소 제곱법은 적합하지 않다고 할 수 있다. 따라서 제안 기법에서는 계수에 L_1 penalty를 주어 계수의 추정 값이 0이거나 0에 근접하도록 하는 희소 모형이 선택 되도록 한다. 따라서, 식 (4)의 왜곡 함수는 다음과 같이 수정하여 사용한다.

$$D_{\Omega_{ij}} = \sum_{l=1}^{|\Omega|} e^{-d(l)} |Q(r_{\Omega_{ij}}^l) - \sum_{(k_1, k_2) \in K} \alpha_{k_1, k_2} Q(r_{\Omega_{i-k_1, j-k_2}}^l)|^2 + \lambda \|\alpha\|_1 \dots \quad (7)$$

이 때 $\|\alpha\|_1$ 는 계수 벡터 α 의 L_1 norm이고 λ 는 constant multiplier이다. 실험에서는 $\lambda=5$ 를 선택하였고 식 (7)의 구현으로 Least absolute shrinkage and selection operator (LASSO)^[8]를 사용하였다.

제안 기법은 부호기 측에서 RDPCM 적용 여부를 결정하여 비트스트림에 전송하는 explicit RDPCM 방식을 이용하므로 최적 모드 m 은 비트율-왜곡 최적화로부터 다음과 같이 결정한다.

$$m^* = \operatorname{argmin}_{m \in M} D + \gamma R \dots \quad (8)$$

이 때 부호화 모드 셋 M 은 제안 기법의 적용 여부를 결정하는 집합이며, D 는 해당 모드의 왜곡, R 은 비트율, 그리고 γ 은 모드 결정에 사용하는 Lagrangian 상수이다.

III. 실험 결과

본 장에서는 HEVC/SCC 참조 코덱인 SCM 4.1 코덱 (screen content coding test model 4.1)^[5]을 이용하여 제안 기법의 부호화 성능을 평가하였다. 제안 기법은 화면 내 압축에 사용하기 때문에 실험은 "All intra" 환경에서 수행하였다. 테스트 시퀀스는 HEVC/ SCC에서 활용하는 비디오를 사용하였다. HEVC/SCC 표준에서는 picture parameter set을 이용하여 변환 생략 모드가 32×32까지 선택이 가능하다^[9]. 이에 따라 실험에서는 블록 내 샘플을 학습하는 제안 기법의 성능 평가를 위하여 참조코덱의 변환 생략 모드를 32×32까지 선택 가능하도록 구성을 변경하였다.

표 1. SCM4.1 대비 제안 알고리즘의 부호화 성능
 Table 1. Coding gain of the Proposed method VS SCM 4.1

Test sequence	Resolution	BD-rate saving (%)
PCB Layout	1920x1080	-5.7%
Waveform	1920x1080	-4.4%
Console	1920x1080	-2.0%
Desktop	1920x1080	-0.8%
Realtime Data	1920x1080	-3.1%
Flying graphics	1920x1080	-1.8%
PPT and XLS	1280x720	-3.6%
Video Conf.	1280x720	-3.2%
Average		-3.1%
Encoding Time		115%

표 1는 제안 기법의 부호화 성능을 보이며 스크린 콘텐츠 비디오 압축을 위한 state-of-the-art codec에서 평균적으로 3.1% BD-rate 감소를 보인다. SCM4.1은 손실 압축의 경우 implicit RDPCM방식을 사용하므로 표 1의 부호화 성능은 SCM4.1의 RDPCM 성능 대비 증가한 성능 향상을 보이는 것이다. 각 시퀀스 별 특징을 분석하면 PCB Layout의 경우

회로의 다양한 방향의 선들이 수직 혹은 수평의 단일 방향으로 예측이 힘들지만 제안 방식에 의하면 임의의 방향을 각 계수의 조합으로 선택하여 보다 효율적인 부호화 성능을 제공할 수 있었다. 반면 Desktop 비디오는 제안 기법의 부호화 성능 향상이 비교적 뛰어나지 않다. 그 이유는 Desktop 비디오의 경우 활자 이외에도 화면 내 평탄한 영역이 많이 나타남에 따라 RDPCM을 활용하기 위한 변환 생략 모드가 자주 선택이 되지 않기 때문이다. 이외에 숫자나 활자 등이 자주 나타나는 전자 문서로 구성된 시퀀스의 경우 3.1~ 3.6%의 비교적 우수한 부호화 효율을 보인다.

IV. 결론

본 논문에서는 효율적인 스크린 콘텐츠 비디오의 압축을 위하여 잔차 신호의 인접 성분으로부터 가중 계수를 유도하고 보다 정확한 예측을 수행하는 향상된 RDPCM 기법을 제안하여 참조 코덱 SCM4.1 대비 약 3.1%의 부호화 성능 향상을 보였다.

참 고 문 헌 (References)

- [1] ISO/IEC/JTC1/SC29/WG11 and ITU-T SG16 Q.6, "Joint call for proposals for coding of screen content," Jan. 2014.
- [2] 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.
- [3] R. Joshi, S. Liu, G. Sullivan, G. Tech, J. Xu, Y. Ye, "HEVC Screen Content Coding Draft Text 4, Document JCTVC-U1005," in ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, June. 2015.
- [4] Y. L. Lee, K.H. Han, and G. Sullivan, "Improved lossless intra coding for H.264/AVC," IEEE Trans. Image Processing, 2006.
- [5] SCM 4.1 software, available at: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.6+SCM-4.1
- [6] R. Joshi, J. Sole, and M. Karczewicz, "Extension of residual DPCM to lossy coding JCTVC-N0052", July 2013.
- [7] M. Nacarri, A. Gabriellini, M. Mrak, S. Blasi and E. Izquierdo, Inter prediction Residual DPCM, JCTVC-M0442, Apr. 2013.
- [8] R. Tibshirani, "Regression Shrinkage and Selection via Lasso" Journal of the Royal Statistics Society Series B, vol. 58, 1996.
- [9] X.Peng, J.Xu, L.Guo, J. Sole, and M. Karczewicz, "Non-RCE2: Transform skip on large Tus," JCTVC-N0288, July, 2013.