

H.264/AVC 기반 계위 부호화를 위한 효과적인 양자화 계수 결정법

김 승 환, 호 요 성

광주과학기술원 정보통신공학과

전화 : 062-970-2258 / 핸드폰 : 011-9052-5739

An Efficient Method to Select Quantization Parameters for Scalable Extension of H.264/AVC Video Codec

Seung-Hwan Kim, Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)

E-mail : {kshkim, hoyo}@gist.ac.kr

Abstract

For scalable extension of the H.264/AVC video coding standard, motion-compensated temporal filtering (MCTF) has been investigated. In MCTF, two-channel decomposition is achieved by the lifting representation including prediction and update steps. In this paper, we consider the importance of each subband frame obtained by MCTF and propose an efficient method to select quantization parameters for each subband frame. Experimental results demonstrate the proposed method increases the average PSNR value by up to 0.5dB.

I. 서론

인터넷을 통한 비디오 스트리밍과 같이 네트워크 채널 용량이 넓은 범위에서 시시각각 변하는 네트워크 비디오 응용에서 기존의 비디오 부호기는 시간에 따라 변하는 채널용량을 정확히 예측할 수 없다. 이러한 인터넷 스트리밍을 위해 제안된 방법이 계위 비디오 부호화(scalable video coding)이다. 계위 비디오 부호기들은 부호화를 수행하기 전에 미리 전송채널의 용량을 예측하여 최소 채널용량보다 작은 비트율로 기본계층(base layer)을 부호화하므로 기본계층 데이터의 전송은 항상 보장된다 [1]. 또한 원래 영상과 기본계층 영

상의 차이 값인 잔여 영상은 네트워크 채널용량이 허용되는 범위에서 향상계층(enhancement layer)에서 부호화되어 전송된다. 향상계층은 기본계층에 있는 비트열을 개선하기 위하여 사용되는 비트열이다.

향상계층의 종류로는 공간적, 시간적, 그리고 SNR (Signal-to-Noise Ratio; 신호 대 잡음비)계위가 있다. 공간적 계위는 크기가 작거나 해상도가 낮은 화면의 크기를 늘리거나 해상도를 높이기 위해 사용된다. 시간적 계위는 시간적 해상도를 높이기 위한 방법으로 향상계층을 추가하여 시간 해상도를 높여주는 것으로, 예를 들어, 초당 15화면 영상을 30화면의 영상으로 만들어 준다. SNR 계위는 화질을 좋게 하는 방법으로 기본계층에 가변길이 복호화 역스캔, 그리고 역양자화를 수행한 값에 향상계층의 값을 더한다 [2].

지난 제66차 MPEG 회의에서 계위 비디오 부호화 표준에 대한 연구가 본격적으로 시작되었고, 제69차 회의 이후 HHI에서 제안한 H.264/AVC 확장 계위 부호화 방법(scalable extension of the H.264/AVC)이 WD(Working Draft)로 결정되었다 [3].

본 논문에서는 H.264/AVC 확장 계위 부호화 표준 소프트웨어에서 부호화 효율을 높이기 위해 사용하는 움직임 보상된 시간적 필터링(motion-compensated temporal filtering, MCTF)의 성능 향상을 위한 방법을 제안한다. 제안된 방법에서는 MCTF 과정에서 생성되는 각 단계(stage)별 고주파 및 저주파 화면의 중요성을 고려하여 적절한 양자화 변수값을 결정하는 방법에 대해서 논의한다.

II. H.264/AVC 기반의 계위 부호화

2.1 리프팅(Lifting) 기법

최근에 제안된 계위 비디오 부호화 표준 소프트웨어(JSVM)에서는 리프팅(Lifting)기법을 이용하여 MCTF를 구현하였다 [3]. 그림 1에서는 JSVM에 구현된 리프팅 기법에 대해 묘사하였다.

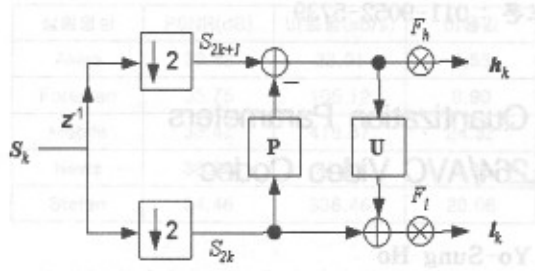


그림 1. Lifting 기법

리프팅 기법에서는 먼저 영상의 화면들을 짝수 화면(S_{2k})과 홀수 화면(S_{2k+1})으로 분리한다. 그리고 두 분리된 영상들을 예측 단계(prediction step)와 갱신 단계(update step)를 거쳐서 각각 고주파 화면(h_k)과 저주파 화면(l_k)으로 분리(decomposition)한다. 예측 단계는 참조 화면으로 원 영상을 사용한다는 것과 개방 루프(open loop) 방식을 사용한다는 점을 제외하면 기존의 블록기반 비디오 부호화 방식에서 움직임 예측 과정과 동일하다. 즉, 예측 단계를 통해 예측 움직임 벡터와 잔여 영상과 같은 고주파 화면이 얻어진다. 갱신 단계에서는 예측 단계에서 얻은 예측 움직임 벡터를 이용하여 짝수 화면에 고주파 화면을 더하여 새로운 저주파 화면을 생성한다 [4][5].

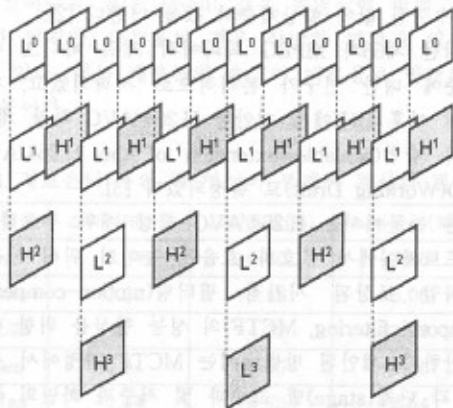


그림 2. MCTF 구조

2.2 움직임 보상된 시간적 필터링 (MCTF)

H.264/AVC 확장 계위 부호화 표준에서 MCTF 과정은 부호화 효율 향상에 가장 기여가 많은 부분이다. 앞에서 설명한 리프팅 기법이 두 개의 영상을 각각 고주파, 저주파 영상으로 분리하는 과정임을 확인하였다. MCTF는 이러한 리프팅 기법의 연속된 조합이라고 볼 수 있다. 그림 2에서와 같이 화면들의 그룹(group of picture, GOP)을 설정하고 설정된 GOP 단위 내에서 MCTF를 실행한다[6]. GOP내에서 연속하는 두 화면이 리프팅 기법을 통해 각각 저주파 및 고주파 화면으로 분리된다. 분리된 화면들 중에서 저주파 화면들에 대해서만 위의 과정을 반복한다. 그림 2에서는 GOP 크기가 12이며, 각각 3단계의 과정을 거쳐 MCTF가 적용된 예를 나타내고 있다.

III. 제안한 알고리즘

3.1 계층별 부호화

그림 3에서 보는 바와 같이, MCTF 과정을 통해 분리된 각 영상들은 각 계층별로 분리된다. 각 계층의 형성은 MCTF 과정에서 리프팅 기법의 적용된 횟수에 의해 결정된다. 예를 들면, 그림 3에서 L_3 와 H_3 는 리프팅 기법을 3번 적용하여 얻은 결과 화면들을 나타낸다. 또한 각 계층별 양자화 변수는 MCTF 과정에서 계층간의 독립적으로 결정된다. 현재 JSVM에서는 각 계층간의 양자화 변수설정에 대한 방법을 제시하지 않고 단지 모든 계층별로 동일한 양자화 변수를 사용한다. 따라서 본 논문에서는 각 계층별 화면들의 중요도를 고려하여 적절한 양자화 변수값 설정에 대한 방법을 제안한다.

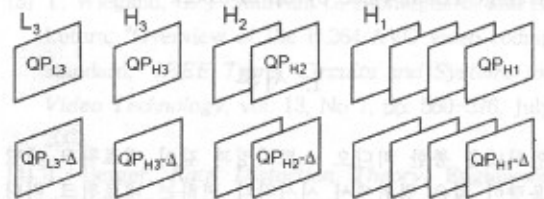


그림 3. 계층별 부호화

3.2 계층별 가중 부호화 방법

본 논문에서는 MCTF 결과로 얻어진 각 계층별 화면의 중요도를 고려하기 위해 분리된 각 계층들이 복원된 영상을 형성하는 과정에서 각 복원된 영상에 기여하는 정도에 따라 중요도를 고려한다.

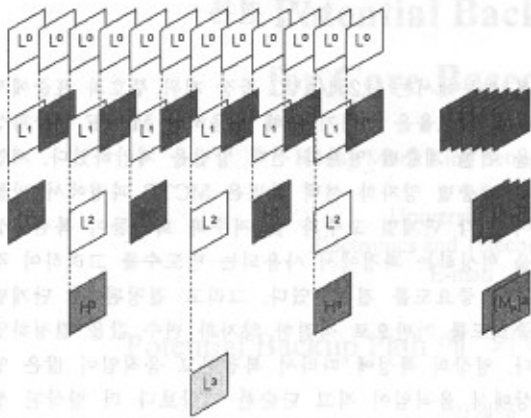


그림 4. MCTF 복원 과정

그림 4는 MCTF 과정을 통해 분리된 화면들이 움직임 예측 벡터(M_p)들을 이용하여 복원된 영상을 형성하는 과정을 보여주고 있다. 각 계층별 가중값을 계산 위해 각각 L^3 화면에 대한 가중값을 W_{L3} , H^3 화면에 대한 가중값을 W_{H3} , H^2 화면에 대한 가중값을 W_{H2} , H^1 화면에 대한 가중값을 W_{H1} 으로 설정하였다. 또한 각 계층간에 상관도를 $K(0 < k < 1)$ 라 가정하였다.

L^3 화면은 복원된 영상을 형성하는 과정에서 제1단계에서는 L^3 화면, 제2단계에서는 H^3 와 역 갱신과정을 통해 두개의 L^2 화면, 제3단계에서는 H^2 와 역 갱신과정을 통해 3개의 L^1 화면을 형성하며, 마지막으로 제4단계 과정에서는 H^1 와 역 갱신과정을 통해 6개의 L^0 화면을 형성한다. 결과적으로 L^3 화면의 경우에 자신을 포함한 총 12개 화면의 복원에 기여한 것을 알 수 있다. 따라서 L^3 화면의 중요도는 식 (1)과 같다.

$$W_{L3} = 1 + 2k + 3k^2 + 6k^3 \quad (1)$$

이와같은 방식으로 H^3 , H^2 , H^1 화면에 대한 중요도는 식(2), 식(3), 식(4)와 같이 나타낼 수 있다.

$$W_{H3} = 1 + 2k + 3k^2 \quad (2)$$

$$W_{H2} = 1 + 2k \quad (3)$$

$$W_{H1} = 1 \quad (4)$$

위에서 제안한 가중값을 얻는 과정에서 각 계층별 상관도를 나타내는 k 값은 각 화면별 가중값을 얻기 위해 사용한 이론적인 값이다. 위에서 제시한 4가지 수식을 기반으로 저주파 화면에 대한 일반적인 가중값과 고주파 화면들의 각 단계별 가중값을 계산할 수 있다. 주어진 각 단계별 가중값 방법을 JSVM에 구현하기 위해서 본 논문에서는 각 단계별 양자화 변수값에 있어서 2 만큼의 균일한 차이를 두어 실험하였다.

IV. 실험 결과

본 논문에서 제안된 알고리즘의 성능을 검사하기 위해, JSVM 1.0을 사용하였으며, 3개의 서로 다른 특성을 가지는 비디오 영상들("FOREMAN", "AKIYO", "MOBILE")을 실험에 사용하였다. 실험에 사용된 모든 영상들은 352×288 화소의 해상도의 CIF 형식이며, 30fps로 부호화하였다. 실험에 필요한 중요 환경 변수 설정은 표 1에 정리하였다.

표 1. 실험 환경 변수 설정

GOP 크기	8
탐색 모드	Fast Search
시간적 계위 부호계층	1
FGS 계층	0, 0.5, 1, 1.5, 2, 3
양자화 계수(MCTF)	42

표 2와 표 3은 각 영상별로 FGS 계위 부호화율 0.5 단위로 증가시키면서 실험한 결과를 나타낸다. 표 2와 표 3에서 알 수 있듯이, 제안된 방법의 경우 저주파 화면에 많은 비트를 할당하면 FGS 계층이 증가함에 따라 더 나은 성능을 보인다. "MOBILE" 영상의 경우에 최대 0.8dB까지 PSNR 값이 향상된다.

표 2. FGS 계층에 따른 PSNR 값 비교("FOREMAN")

FGS 계층	JSVM		제안한 방법	
	비트율	PSNR	비트율	PSNR
0	94	30.82	104	31.01
0.5	140	31.69	148	31.92
1	205	33.20	210	33.51
1.5	324	34.20	323	34.62
2	452	35.42	469	36.29

표 3. FGS 계층에 따른 PSNR 값 비교("MOBILE")

FGS 계층	JSVM		제안한 방법	
	비트율	PSNR	비트율	PSNR
0	200	26.27	209	26.37
0.5	351	27.66	357	27.79
1	535	29.32	536	29.52
1.5	845	30.49	821	30.65
2	1292	32.35	1295	32.85

표 4에서는 "AKIYO" 영상에 대한 제안한 알고리즘의 실험결과를 나타내고 있다. FGS 계층은 3까지 사용하였다. 제안한 방법의 경우에 "MOBILE"과 같이 복잡하고 움직임이 많은 영상에서는 비교적 효율적이지만, "AKIYO"와 같이 움직임이 적고 단순한 영상의 경우에는 비교적 비효율적인 성능을 보였다.

표 4. FGS 계층에 따른 PSNR 값 비교("AKIYO")

FGS 계층	JSVM		제안한 방법	
	비트율	PSNR	비트율	PSNR
0	34	33.79	35	33.85
1	98	36.86	98	36.91
2	226	40.26	225	40.29
3	473	43.36	466	43.37

그림 5와 그림 6에서는 "FOREMAN"과 "MOBILE" 영상에 대하여 FGS 계층의 증가에 따른 PSNR 값의 변화를 알기 쉽게 나타내고 있다. FGS 계층이 1미만인 경우에 제안한 방법과 기존의 JSVM과 비교하여 많은 차이가 없지만, FGS 계층이 1 이상인 경우에 제안한 방법이 매우 효율적임을 확인할 수 있다.

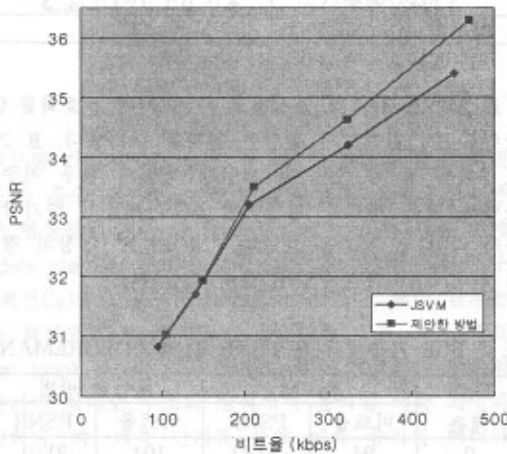


그림 5. PSNR 값 비교("FOREMAN")

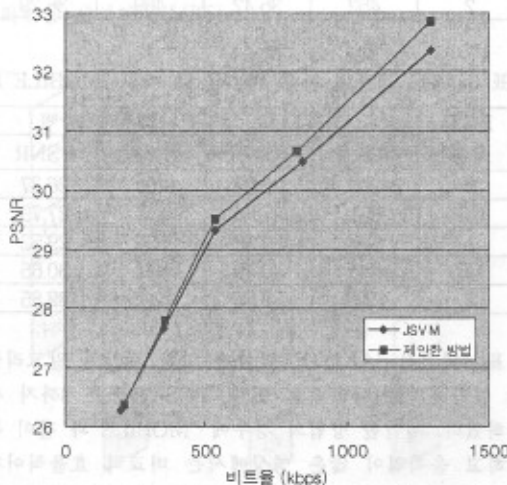


그림 6. PSNR 값 비교("MOBILE")

V. 결론

본 논문에서는 H.264/AVC 확장 계위 부호화 표준에서 부호화 효율을 높이기 위해 사용하는 MCTF 성능향상을 위한 계층별 양자화 선택 방법을 제안하였다. 제안된 계층별 양자화 선택 방법은 MCTF 과정에서 만들어진 각 단계별 고주파 및 저주파 화면들이 복원영상을 형성하는 과정에서 사용되는 빈도수를 고려하여 각각의 중요도를 결정하였다. 그리고 결정된 각 단계별 중요도를 기반으로 적절한 양자화 변수 값을 결정하였다. 영상의 특성에 따라서 복잡하고 움직임이 많은 영상에서 움직임이 적고 단순한 영상보다 더 향상된 성능을 보임을 실험을 통해 확인하였다. 또한, FGS 계위 부호화 실험환경에서, 제안된 방법의 경우 항상계층이 증가함에 따라 그 성능이 향상됨을 확인하였다.

감사의 글

본 연구는 광주과학기술원 실감방송연구센터(RBRC)를 통한 정보통신부 대학IT연구센터(ITRC), 교육인적자원부 두뇌한국21(BK21) 정보기술사업, 그리고 디지털콘텐츠협동연구센터(DCRC)의 지원에 의한 것입니다.

참고문헌

- [1] W. Li, "Overview of Fine Granularity Scalability in MPEG-4 Video Standard," *IEEE Trans. Circuit and System for Video Technology*, vol. 11, no. 3, pp. 301-317, March 2001.
- [2] D. Taubuman, "Successive Refinement of Video: Fundamental Issues, Past Efforts and New Directions" *Proc. of SPIE (VCIP 2003)*, vol. 5120, pp. 649-663, July 2003.
- [3] H. Schwarz, D. Marpe, and T. Wiegand, "Scalable Extension of H.264/AVC," ISO/IEC JTC1/WG11 Doc. M10569/SO3, Mar. 2004.
- [4] M. Flierl and B. Girod, "Video Coding with Motion-Compensated Wavelet Transforms," *Proc. of PCS*, pp.59-62, April 2003.
- [5] T. Wiegand, H Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-Constrained Coder Control and Comparison of Video Coding Standards," *IEEE Trans. on Circuits and Systems for Video Tech*, vol. 13, July 2003.
- [6] ISO/IEC JTC1, "Requirements and Application for Scalable Video Coding," ISO/IEC JTC1/WG11 Doc. N6025, Oct. 2003.