

MPEG-4 Video Stream의 Scalable Coding을 위한 최적화 방안에 관한 연구

곽무진, 한승균, 서덕영

경희대학교 전자정보학부

A Study on optimized method of Scalable Coding of MPEG-4 Video Stream

Moo-Jin Kwak, Seung-Gyun Han, Doug-Young Suh

Dept. of Electronic Eng., Kyunghee-Hee University

mjkwak@newmedia.kyunghee.ac.kr

sghan@newmedia.kyunghee.ac.kr

suh@nms.kyunghee.ac.kr

요약

본 논문에서는 동영상의 계층적 부호화의 효율을 높이기 위한 방안에 대해 연구하였다. 단일 계층부호화에 비해 다 계층부호화는 계산량이 많아진다. 따라서 계층적 부호화의 장점을 살리고 단점을 보완하는 방안을 제시하였다. 우선 인코더에서 고급계층의 복잡도를 줄이기 위하여 고급계층의 참조 형태를 P-VOP(Prediction-Video Object Plane)만으로 정한다. 고급계층의 참조 영역으로 사용되는 업샘플링된 VOP의 횟수를 줄여서 업샘플링에 따른 계산량을 줄인다. 그리고 고급계층의 비트율을 조절하여 Traffic shaping 효과도 얻을 수 있다. 이러한 방법들을 통해 단일 계층 부호화에 비해 다 계층부호화의 장점을 살리고 단점을 보완하는 코덱을 제안한다

1. 서론

최근 들어 컴퓨터와 통신 기술의 비약적 발달로 인해 정보 통신망 사용자들은 다양한 정보 서비스를 접할 수 있게 되었다. 게다가 동영상 압축 기술의 발달로 적은 대역폭을 사용하여 대용량의 비디오 서비스를 가능토록 했다. MPEG-4(Moving Picture Coding Experts Group)의 다양한 기능·중에 계층화된 부호화 방식은 입력 영상을 기본계층과 고급계층으로 나누어서 차별적 코딩이 가능하다[1][2]. 이러한 방식의 계층화된 부호화 방식은 단일 계층부호화 방식에 비해 장단점을 동시에 가지고 있다[3]. 우선 장점으로는 계층화된 서비스를 사용자에게 차별적으로 제공할 수 있다는 점이다. 하지만 다 계층부호화 방식은 단일 계층부호화에 비해 다 계층으로의 부호화에 따른 계산량이 많아진다. 따라서 본 논문에서는 계층부호화의 장점을 살리고 단점 중에 하나인 계산량을 줄이는데 주안점을 둔다.

본 논문에서는 공간적 계층화(Spatial Scalability)를 이용한 부호화의 최적화 방안에 대해 제안하고 그 성능을 평가 분석한다. 2장에서는 계층부호화의 종류, 특징과 구조에 대해 알아보고 3장에서는 제안하는 공간적 계층화의 형식에 대해 설명한다. 4장에서는 기존 방법과 제안하는 방법 간의 성능을 비교, 분석하며 마지막으로 5장에서 본 논문의 결론을 맺는다.

2. 계층부호화

2.1 계층 부호화 종류

MPEG-4에서는 SNR 계층부호화(SNR Scalability), 공간적 계층부호화(Spatial Scalability), 시간적 계층부호화(Temporal Scalability) 등이 있다. SNR 계층부호화는 기본계층과 고급계층의 양자화 계수를 달리하여 계층을 구분한다. 공간적 계층부호화는 공간적 해상도를 달리하여 계층을 나눈다. 인코딩에서는 계층화에 따른 업샘플링 방식이 추가로 들어간다[4]. 또 시간적 계층부호화는 시간적인 해상도를 달리하여, 고급계층이 기본계층에 비해서 초당 프레임율을 늘린다.

2.2 계층 부호화의 구조

기본계층 인코더를 통해서 인코딩된 비트열을 통해 재조립된 VOP를 고급계층의 참조 영역으로 하기 위하여 업샘플링을 한다. 두 개의 비트열로 나뉜 후 디코더에서는 기본계층 비트열만 가지고서 디코딩할 수도 있고, 기본계층 비트열과 고급계층 비트열을 같이 가지고 디코딩해서 시간적, 공간적 해상도가 뛰어난 영상을 만들어 낼 수 있다.

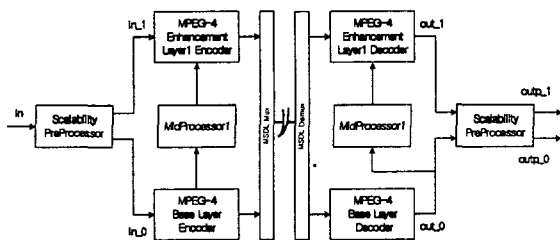


그림 1. 계층 부호화의 구조

2.3 공간적 계층부호화의 형식

공간적 계층부호화의 형식은 고급계층의 참조 형식에 따라 PPP(Prediction)와 PBB(Bidirectionally Prediction)로 나뉘어진다. P-VOP일 경우는 순방향 예측이, B-VOP일 경우는 순방향, 역방향 예측이 동시

에 사용되어진다.

2.3.1 고급계층의 형식 : PPP

고급계층의 형식이 P-VOP일 경우는 순방향 예측이 사용되어지는데 아래 표 1.에서처럼 여러 방향으로의 참조를 할 수 있다.

ref_select_code	forward prediction reference
00	Most recently decoded enhancement VOP belonging to the same layer.
01	Most recently VOP in display order belonging to the reference layer.
10	Next VOP in display order belonging to the reference layer.
11	Temporally coincident VOP in the reference layer(no motion vectors)

표 1. P-VOP의 예측 방식

ref_select_code가 3일 경우의 모식도는

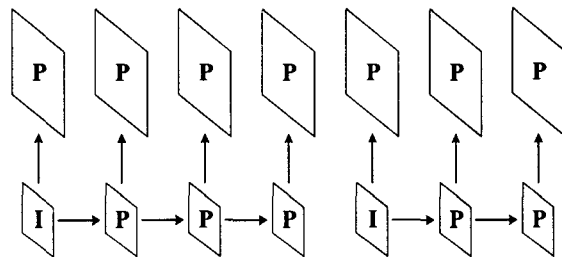


그림 2. 고급계층의 형식 : PPP

2.3.2 고급계층의 형식 : PBB

고급계층의 형식이 B-VOP일 경우는 순방향, 역방향 예측이 동시에 사용되어진다. 아래 표 2.에서 나오는 것처럼 여러 방향으로의 참조를 할 수 있다.

ref_select_code	forward temporal reference	backward temporal reference
00	Most recently decoded enhancement VOP belonging to the same layer.	Temporally coincident VOP in the reference layer(no motion vectors)
01	Most recently decoded enhancement VOP belonging to the same layer.	Most recently VOP in display order belonging to the reference layer.
10	Most recently decoded enhancement VOP belonging to the same layer.	Next VOP in display order belonging to the reference layer.
11	Most recently VOP in display order belonging to the reference layer.	Next VOP in display order belonging to the reference layer.

표 2. B-VOP의 예측 방식

ref_select_code가 0일 경우의 모식도는

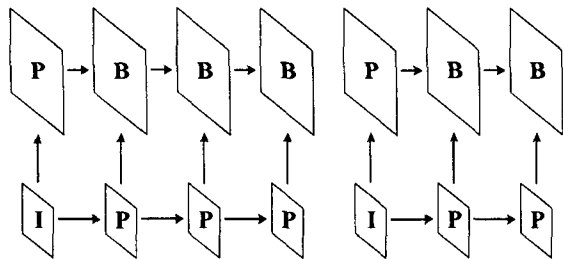


그림 3. 고급계층의 형식 : PBB

3. 제안하는 공간적 계층부호화의 형식

고급계층의 형식은 예측에 따른 계산량을 줄이기 위해서 P-VOP인 경우만을 제시한다.

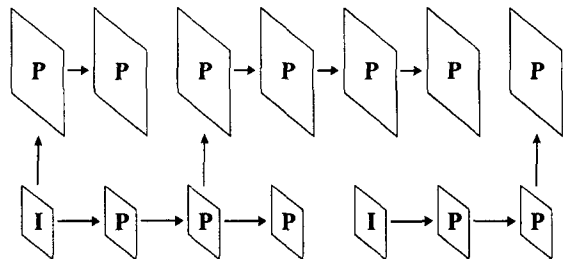


그림 4. 제안한 고급계층의 형식 : PPP

우선 고급계층의 첫 번째 P-VOP는 기본계층의 I-VOP(Intra)를 참조한다. 그리고 인트라 주기 (Intra_period) 값에 따라 업샘플링 횟수를 제한한다. 만약 인트라 주기가 10일 경우는 6번째 VOP에서 업샘플링을 수행해서 고급계층의 참조 VOP로 사용되어 지고, 나머지는 이전 VOP를 참조 영역으로 한다. 이러한 참조 VOP를 고급계층의 이전 VOP로 함으로써 업샘플링으로 인한 계산량을 줄일 수 있다.

4. 실험방법 및 실험결과

본 실험에서 Momusys MPEG-4 코덱을 사용했다. MPEG-4에서는 영상의 움직임과 복잡도의 정도에 따라 Class A~C로 구분한다. 본 실험에서는 Class A의 Akiyo, Class B의 News, Class C의 Stefan 영상을 사용하였다. 실험의 조건으로는 총 프레임은 30 frame, 프레임율은 10Hz, 그리고 인트라 주기는 10이다.

4.1 Akiyo의 비트율에 따른 비교 분석

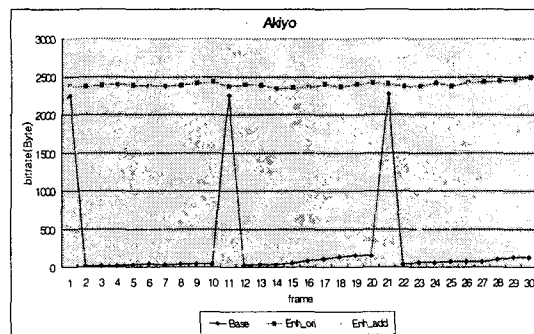


그림 5. Akiyo 비트율

원래 고급계층은 기본계층에서의 업샘플링한 VOP만을 참조하므로 일정한 비트율을 나타낸다. 하지만 제안하는 고급계층은 업샘플링에 따른 횟수를 인트라 주기 동안 한번으로 제한했다. 따라서 프레임 수가 6일 때 업샘플링된 VOP를 참조하므로 원래 고급계층과 같은 비트율을 나타내고, 고급계층의 이전 VOP를 참조하는 나머지 프레임들은 비트율이 현저히 낮다. 게다가 기본계층의 I-VOP의 비트율이 최대치를 나타냄으로 고급계층에서는 업샘플링된 프레임의 비트율이 최대치를 나타냄으로 업샘플링 프레임 인트라 주기 중간에 위치하게끔 해 Traffic shaping의 효과를 보았다.

4.2 News의 비트율에 따른 비교 분석

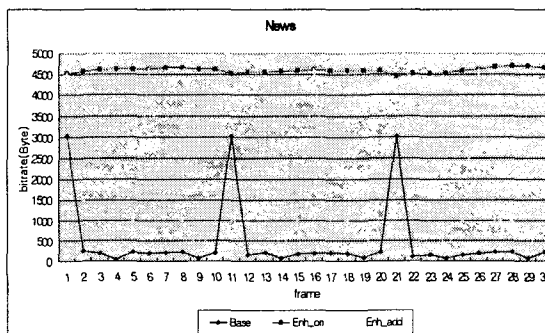


그림 6. News 비트율

Akiyo의 비트율과 같은 결과를 나타낸다. 단지 Akiyo 보다는 움직임이 많아서 기본계층의 비트율과 고급계층의 비트율이 차이가 난다

4.3 Stefan의 비트율에 따른 비교 분석

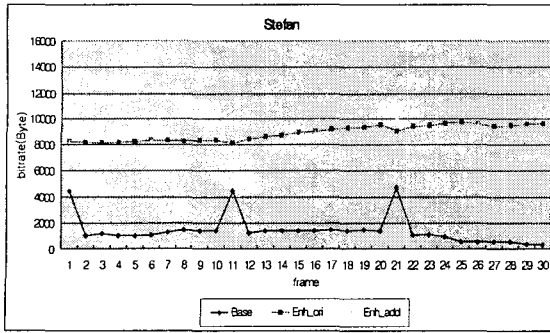


그림 7. Stefan 비트율

움직임이 많은 stefan의 경우는 제안하는 형식의 효과가 미비하다. 업샘플링 횟수 제한에 따른 시간 단축 이득은 있지만 움직임이 많은 경우는 고급계층에서 참조 VOP를 이전 VOP보다는 기본계층의 업샘플링된 VOP를 참조하는 게 더 낫다는 결과를 보여주고 있다.

4.4 업샘플링 횟수 제한 이득과 PSNR(Peak Signal to Noise Ratio)

업샘플링 횟수 제한에 따른 이득은 전체 인코딩 시간 중에서 총 업샘플링 시간을 1로 놓고, 제한된 횟수의 업샘플링 시간과 이전 VOP를 가져오는 시간을 더한 후 시간의 비를 타나낸 것이다.

	업샘플링 제한 이득	원래 고급계층 PSNR	제안한 고급계층 PSNR
Akiyo	50.83%	35.53	35.51
News	51.44%	33.44	33.46
Stefan	50.70%	30.85	31.28

표 3. 업샘플링 제한 이득과 PSNR 비교

표 3.과 같이 영상의 구분에 따른 차이 없이 50~5% 의 시간 이득을 본다. 그리고 원래 업샘플링만으로 인코딩, 디코딩된 원래 고급계층의 PSNR과 제한된 업샘플링과 이전 VOP 참조가 된 제안하는 고급계층의 PSNR은 큰 차이가 없다. 다만 4.3에서 살펴본 바와 같이 움직임이 많은 영상일 경우는 PSNR 차이보다는 비트율이 현저히 증가함을 알 수 있다.

5. 결론

MPEG-4의 계층부호화 중 공간적 계층부호화를 사용해서 고급계층에서 참조 VOP를 기본계층의 업샘플링

된 VOP로 할 경우, 업샘플링에 대한 계산량이 많아진다. 따라서 인코딩 시간이 길어진다. 제안하는 고급계층 방식은 업샘플링 횟수를 줄이고 대신에 참조 VOP를 이전 VOP로 하므로써 인코딩 시간을 줄이고, 비트율도 현저히 줄일 수 있는 장점을 보였다. 또한 인트라 주기의 중간에 업샘플링을 적용함으로써 기본계층과 고급계층의 프레임당 최대 비트율의 위치를 달리하여 Traffic shaping 효과를 보였다. 하지만 움직임이 많은 stefan 영상일 경우는 업샘플링 횟수 제한에 따른 이득은 볼 수 있지만, 고급계층의 비트율이 높아지는 단점을 보인다. 따라서 상대적으로 움직임이 적은 경우는 업샘플링 횟수에 제한을 하고, 움직임이 많은 경우는 업샘플링 횟수를 증가시켜 비트율을 낮추는 방법을 택해야 한다.

참고 문헌

- [1] "MPEG-4 Video Verification Model Version17.0", ISO/IEC JTC1/SC29/WG11 N3515
- [2] N. Wells, P. Tudor, "Standardization of scalable coding schemes," Proc. IEEE ISCAS 94, Tutorials pp121-130, May 1994.
- [3] B. Girod, K. Stuhlmiller, M. Link and U. Horn, "Packet Loss Resilient Internet Video Streaming," Part of the IS&T/SPIE Conference on Visual Communications and Image Processing '99 Jan. 1999
- [4] Qingwen Hu and Sethuraman Panchanathan, "Image/Video Spatial Scalability in Compressed Domain," IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, VOL. 45, NO. 1, FEBRUARY 1998