

FGS를 이용한 계층적 코딩기법에 관한 연구

김종욱, 이배호
전남대학교 컴퓨터공학과
e-mail: jukim69@moiza.chonnam.ac.kr

A Study on Scalable Video Coding Technique with Fine Granularity Scalability

Jong-Uk Kim, Bae-Ho Lee
Dept. of Computer Engineering, Chonnam National University

요약

본 논문에서는 MPEG-4 표준에서 채택된 FGS(Fine Granularity Scalability)를 이용한 계층적 비디오 부호화 방법을 적용하였다. MPEG-4 FGS는 인터넷에서 이용되는 다양한 특성의 단말기들과 대역폭 변동에서 유연성 있는 기술이다. 적용된 MPEG-4 FGS 기술은 DCT 기반 비트 평면 부호화를 이용한다. 실험에서는 일반적인 SNR 계층 비디오 코딩(Multi-layered SNR scalable video coding)과의 비교를 통해 FGS의 효율성을 평가한다. 비교는 두 부호화 기술의 PSNR 값의 분석을 가지고 이루어졌고 MPEG-4 FGS 기술이 이전의 다중 SNR 계층 비디오 부호화 방법보다 평균 1~2dB 정도의 높은 값으로 나타났다. 이를 통해 MPEG-4 FGS가 효율성면에서 다중 SNR 계층 비디오 부호화 방법보다 우수함을 알 수 있지만, 단일 계층부호화(Single-layered scalable video coding)에 비해서는 효율성이 낮게 나타났다.

1. 서론

인터넷 비디오 멀티캐스팅은 최근 많은 멀티미디어 응용을 위한 중요한 기술 영역이다. 이 영역의 응용은 올림피아나 월드컵 결승전 같은 주요한 이벤트를 수백만 인터넷 이용자들이 동시에 시청하는 것을 포함한다. 이런 서비스를 수행하기 위해, 멀티캐스트 인터넷 비디오는 고품질 계층비디오부호화 알고리즘의 개발을 요구한다. FGS는 스트리밍 응용을 위한 MPEG-4 표준에 의해 채택되어온 새로운 비디오 압축 구조이다. FGS는 일반적인 IP기반 네트워크들과 같이 글로벌 인터넷을 특정화하는 넓은 범위의 대역폭 변환(bandwidth variation)들을 지원할 수 있다. 또한 FGS는 인터넷상에서 흔하게 발생하는 패킷 손실(packet losses)을 복원 할 수 있는 기능을 제공한다.[2][6][7].

FGS가 MPEG-4에 처음으로 소개 되었을 때, 세 가지의 주요한 접근방법들이 제안 되었다[1][6]. 크게 zero-tree 연산에 기반한 wavelet 부호화 방식, DCT(Discrete Cosine Transform)계수들을 비트 평면 단위로 부호화 하는 방식, 그리고 오차 신호에 대해 mat

-ching pursuit을 이용한 부호화 방식 등이 제안되었다. 비트 평면 DCT 기반 부호화(bit plane DCT-based coding)와 웨이블릿 압축 방법의 다양한 변환들에 대한 많은 연구가 진행되어 왔다. MPEG-4 FGS 상위 계층 신호의 전체적인 분석에 기반한, 몇몇 연구는 비트 평면 DCT 부호화와 임베드드 제로트리 웨이블릿 기반 압축(embedded zero-tree wavelet based compression)은 비슷한 결과를 보였다. 그러나 부호화 효율과 구현 가능성 측면을 고려함으로써 DCT 기반의 비트 평면 부호화 방식을 가장 적합한 방식으로 채택하였다[1][5][6].

본 논문에서는 MPEG-4 FGS와 일반적인 SNR 계층 비디오 코딩(SNR scalability video coding)의 효율성을 비교한다. 특히, 다중 SNR 계층 비디오 코딩(multi-layered SNR scalable video coding)의 PSNR의 결과가 일정 범위의 비트율에서 FGS 결과들과 비교되어진다. 또한 단일 계층 부호화 방식과의 비교를 통해 효율성 면에 있어서 MPEG-4 FGS와의 차이를 보인다. 본 논문의 구성은 다음과 같다. 2장에서, 스케일러빌리티의 기능과 구조를 소개한다. MPEG-4 FGS의 기본형태, 인코더, 그리고 디코더의 구조는

3장에서 기술하였다. 4장에서는 MPEG-4의 핵심기술인 DCT 기반 비트 평면 부호화 기술을 나타낸다. 실험 결과는 5장에서 비교 분석하였다. 마지막으로 7장에서는 결론을 맺었다.

2. 스케일러빌리티 기능과 구조

시간적(temporal), 공간적(spatial) 그리고 SNR 스케일러빌리티(scalability)들은 MPEG-2, MPEG-4, H.263 같은 비디오 압축 표준들에서 정의 되었다[3]. 계층비디오는 기본계층(base layer)과 하나 혹은 여러 개의 상위계층들(enhancement layers)로 구성된다. 일반적으로 계층 비디오 스트림-scalable video stream)의 기본 계층부분은 디코딩을 위해 요구된 최소량의 데이터를 나타낸다. 상위계층 부분은 추가적인 정보를 표현한다. 따라서 상위계층은 디코드 될 때 영상 신호를 강화(enhancement)한다.

1) 공간 스케일러빌리티

공간 스케일러빌리티는 비트율은 같지만 다른 공간적 해상도(spatial resolution)에서 일련의 영상장면들을 두 개의 층으로 부호화하기 위한 기술이다. 공간해상도가 낮은 계층을 "기본 계층(BL)", 높은 계층을 "상위 계층(EL)"이라 부른다. 기본계층은 비 계층 인코딩(non-scalable encoding)처리와 같고 보다 낮은 공간 해상도에서 부호화된다. 상위계층은 P 나 B 프레임중 하나로서 인코드 된다. 기본계층에서 I 프레임과 시간적으로 일치하는 VOP(Video Object Plane)는 P 프레임으로서 인코드 된다. 기본계층에서 P 프레임과 시간적으로 일치하는 VOP는 B 프레임으로서 인코드 된다. 기본계층에서 디코드된 VOP는 예측을 위한 참조(reference)로서 이용된다. 기본계층에서 시간적으로 일치하는 VOP는 상위계층의 VOP가 인코드 되기 전에 부호화 되어야 한다. 고위계층 비트스트림(bitstream)은 이미지의 해상도를 증가시키기 위해 이용된다. 그림1의 (a)는 공간 스케일러빌리티의 구조를 보인다. 기본 계층의 영상을 업 샘플링해서 상위계층과 같은 크기의 영상을 만들어내고, 상위계층의 영상으로부터 뿐만 아니라 이 보간 된 영상으로부터도 예측을 적용함으로써 보다 효율이 높은 부호화를 실현한다.

2) 시간 스케일러빌리티

공간 해상도는 같지만 1초당 프레임 주파수가 다른 것들을 효율적으로 보낼 때 시간적 스케일러빌리티를 사용한다. 시간해상도가 낮은 계층을 "기본계층", 높은 계층을 "상위계층"이라 부른다. 기본계층

은 낮은 프레임 율에서 부호화되고 상위계층은 고위 계층 중의 I, P, B 프레임을 사용할 뿐 아니라 기본 계층의 영상을 써서 예측할 수 있다. 시간적 스케일러빌리티의 효율성은 높고 비 계층적 부호화(non-scalable coding)에 매우 근접한다. 그림1의 (b)는 시간적 스케일러빌리티의 구조를 보인다. P 프레임 예측은 기본계층에서만 이용된다. 상위계층 예측은 기본계층으로부터의 P 프레임이나 B 프레임 어느 것을 이용한다.

3) SNR 스케일러빌리티

화질이 다른 두개의 영상 시퀀스(sequence)들을 동시에 효율적으로 보내기 위해서는 SNR 스케일러빌리티를 사용하는 것이 바람직하다. 일반적으로 SNR이 크면 압축률이 낮고 고화질이 된다. 그림1의 (c)는 SNR 스케일러빌리티 구조이다. 화질이 낮은 계층을 "기본계층", 화질이 높은 계층을 "상위계층"이라 부른다.

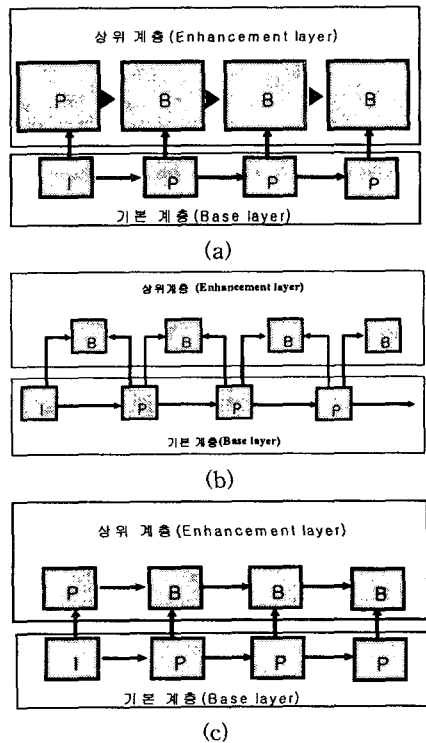


그림 1. 스케일러빌리티 구조 (a) 공간 스케일러빌리티 (b) 시간 스케일러빌리티 (c) SNR 스케일러빌리티

상위계층은 기본계층으로 부호화된 데이터를 재현하여 본래의 데이터로부터 재현된 것을 뺀 오차분만

을 기본계층의 양자화 스텝(quantization step) 보다 작은 양자화 스텝으로 양자화 해서 부호화 한다. 상위계층에서는 예측이 존재하지 않고 항상 이 오차분만이 부호화된다. SNR 스키일러빌리티를 이용하면 화질이 다른 두 종류의 그림을 간단히 부호화·복호화 할 수 있다.

3. FGS(Fine Granularity Scalability) 구조

FGS 기본구조는 기본계층(base layer)과 고위계층(enhancement layer)으로 구성된다. 기본계층은 낮은 비트율에서 비계층 부호화된다. 고위계층은 원래 영상(original picture)과 DCT 계수들의 비트평면 부호화를 이용하는 재구성된 영상(reconstructed picture)사이에서의 차이를 이용해 부호화 된다. 그림 4는 FGS의 기본구조를 나타내고 있다. 그림에서 보여 지는 것처럼 FGS는 각 VOP(Video Object Plane)에 대해 다양한 화질을 제공할 수 있다. 상위계층은 기본계층 VOPs의 화질을 강화하는 VOPs를 구성한다[4]. 그림 5와 6은 FGS 인코더와 디코더를 보이고 있다. FGS 상위 계층의 비트스트림은 인코딩이 완료된 후에 각 영상에 알맞은 임의의 비트수로 나누어진다. 디코더는 기본계층과 절단된 비트스트림(bitstream)들로부터 상위계층 영상을 재구성할 수 있어야 한다. 상위계층 영상 화질은 각 영상에 대해 디코딩된 비트 수에 비례한다.

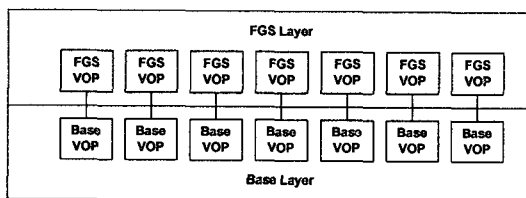


그림 4. FGS 기본구조

4. DCT 계수의 비트 평면 부호화

일반 DCT 코딩에서 양자화된 DCT 계수들은 run-level 부호화 기술을 이용한다. 영이 아닌 DCT 계수들 전까지의 연속된 영들을 "run"이라 부르고, 영이 아닌 DCT 계수의 절대 값을 "level"이라 부른다. 흔히 부르는 "2-D" VLC (Variable Length Code) 테이블을 이용하는 경우, (run, level)기호가 부호화 되고, 분리된 "EOB"기호는 DCT 블록의 끝을 의미한다. "3-D" VLC 테이블이 이용되는 경우, (run, level, eob)기호가 부호화되고, 여기서 "eob"는 DCT

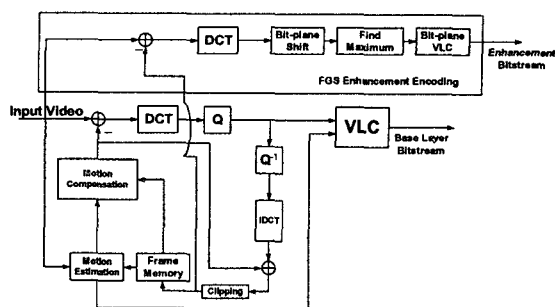


그림 5. FGS 인코더

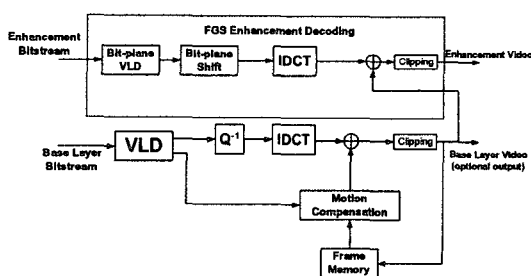


그림 6. FGS 디코더

블록의 끝을 의미한다. DCT 부호화 방법과 비트 평면 부호화 방법의 주요한 차이는 각각의 양자화된 DCT 계수를 어떤 십진 정수 값 대신 이진수 값으로 고려하는 것이다. 각 8*8 DCT 블록에 대해, 64개 절대 값들은 지그재그(zigzag)방법을 통해 하나의 배열로서 정렬된다. 블록의 비트 평면은 64비트들의 배열로서 정의 되고 그것은 중요도가 똑같은 위치에서 DCT 계수들 각각에 대한 절대 값으로부터 하나를 갖게 된다. 각 블록의 비트 평면에 대해, (RUN, EOP)기호들이 형성되고 가변길이는 출력 비트스트림(bitstream)을 넣기 위해 부호화 된다. 가장 중요한 비트평면(MSB plane)으로부터 시작함으로써, 2-D 기호들은 두개의 요소들을 형성 한다.

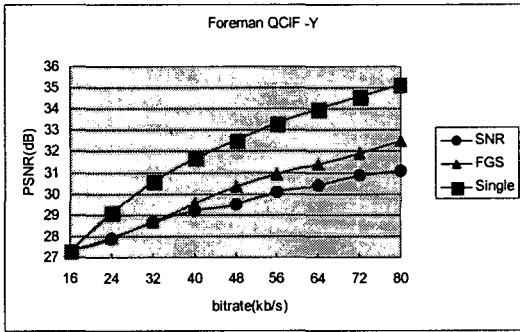
- 1) 1(RUN)의 앞에 있는 연속적인 0의 수
- 2) End-Of-Plane(EOP).

비트 평면이 모든 0을 포함한다면, 특별한 기호 ALL_ZERO는 그것을 표현하기 위해 만들어 진다. 이렇게 만들어진 기호들은 VLC(Variable Length Code) 테이블들을 이용함으로써 부호화 된다.

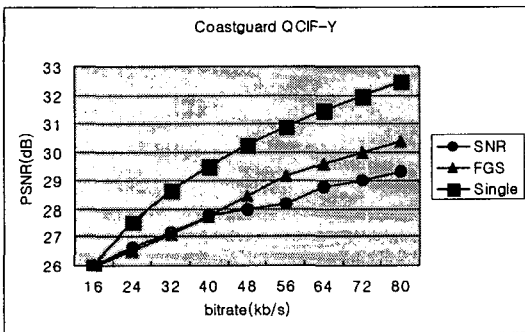
5. 실험 결과

본 논문에서는 새로운 인터넷 스트리밍(Internet streaming) 기술인 MPEG-4 FGS 부호화 기술을 적용하였다. 두개의 MPEG-4 테스트 영상(sequence)

들 중에서 QCIF 해상도(176*144)의 Coastguard 와 Forman이 이용되었다. 16, 24, 32, 40, 48, 56, 64, 72, 80kbps의 비트 율에서 PSNR 결과 값들이 계산되었고 이 결과 값들은 그림 7에서 다층 SNR 계층 부호화, 단일 계층부호 기술들과 함께 비교 되었다.



(a)



(b)

그림 7. PSNR 결과 비교 (a) Forman QCIF-Y
(b) Coastguard QCIF-Y

적용된 부호화 기술은 단일 계층 부호화에 비해서는 효율성이 낮지만 기존에 사용되어온 다층 SNR 계층 부호화보다는 평균적으로 1~2dB정도의 향상을 보이고 있다.

6. 결론

본 논문에서는 다양한 인터넷 환경에서 우수한 유연성(flexibility)을 보일 수 있는 MPEG-4FGS 부호화 기술을 나타냈다. 여기서는 계층 부호화 기술들 중 DCT기반 비트 평면 부호화 기술을 이용하였다. 그리고 이 기술의 평가를 위해 이전에 사용된 다층 SNR 계층 부호화 기술과의 비교 분석하였다. PSNR 결과 값이 보다 높은 비트 율(bit rate)로 가면서 평균적으로 1~2dB정도 더 높게 나타남을 볼

수 있었다. 그러나 단일 계층 부호화 기술에 비해서는 상당히 큰 부호화 효율성(coding efficiency)의 열세를 볼 수 있었다. 결국, 이전의 계층부호화 기술에 비해서는 효율성의 향상을 볼 수 있었지만, 단일 계층 부호화와의 큰 차이는 향상되어야 할 과제로 남는다. 차후의 연구는 [8]에서 제안된 방법과 같이 현재의 FGS 기술에 시간(temporal)과 SNR 스케일 러빌리티를 적용함으로써 화질(quality)을 향상시키고 또한 효율성면에서 단일 계층 부호화 수준에 근접할 수 있는 향상된 FGS 기술에 연구하고자 한다.

참고문헌

- [1] Weiping Li, "Overview of Fine Granularity in MPEG-4 Video Standard", IEEE Trans. on Circuits and Systems for Video Tech., Vol.11, No. 3, pp. 301- 317, March 2001.
- [2] W. Li, Jens-Rainer Ohm, Mihaela Van der Schaar, Hong Jiang, Shipeng Li, "MPEG-4 Video Verification Model version 18.0", ISO/IEC JTC1 /SC29/WG11 N3908, January 2001.
- [3] Coding of Audio-Visual Objects, Part-2 Visual, Amendment 3: Studio profile, ISO/IEC 14496-2: 1999 /FPDAM3, July 2000.
- [4] Coding of Audio-Visual Objects, Part-2 Visual, Amendment 4: Streaming Video Profile, ISO/IEC 14496-2:1999/FDAM, October, 2000.
- [5] W. Li, "Fine granularity scalability in MPEG-4 for streaming video", ISCAS 2000, Geneva, Switzerland, May 28-31, 2000.
- [6] ISO/IEC/JTC1/SC29/WG 11, "VM of ISO/IEC 14496-2 MPEG-4 Video FGS (v4.0)", Doc. N3317, Noordwijkerhout, Netherlands, March 2000.
- [7] Q. Wang, F. Wu, S. Li, Y. Zhong, and Y.-Q. Zhang, "Fine-granularity spatially scalable video coding", IEEE Int Conf on Acoustics, Speech and Signal Processing(ICASSP), pp. 1801-1804, Salt Lake City, May 2001.
- [8] F. Wu, Shipeng Li and Y.-Q. Zhang, "A frame-work for efficient progressive fine granular scalability video coding," IEEE Trans. on Circuits and Systems for Video Tech., Vol.11, No.3, pp. 332-344, March 2001.