

화질 확장성을 높인 FGS 코딩 기법에 관한 연구

부희형, 백낙훈
경북대학교 전자전기컴퓨터학부
e-mail : aqr724@gmail.com

A study of FGS coding with increased video quality scalability

Hee-Hyung Boo, Nak-Hoon Beak
School of EECS, Kyungpook National University

요 약

MPEG-4 Part 2에서 표준으로 채택된 FGS (fine granularity scalability) 코딩 기법은 전송 선로의 상태가 급변하는 경우에도 주어진 대역폭에서 최적의 화질 향상을 얻을 수 있도록 설계된 망 상태에 적응적인 스케일러블 비디오 코딩 기법이다 [1][2][9]. 본 논문에서는 기존의 FGS 향상 계층에서 영상의 잔여 신호를 다시 한 번 bit-plane 코딩을 해줌으로써 화질 확장성을 더 높인 Advanced FGS 코딩 구조를 제안하였다. 본 논문의 실험에서는 기존의 MPEG-4 VM (verification model)에서 사용된 FGS 코딩과의 비교를 통해 Advanced FGS 구조의 화질 확장성을 평가하였다. 비교는 두 부호화 기술의 PSNR 값의 분석으로 이루어졌고, 결과를 통해 Advanced FGS 구조가 고품질을 가지며, 화질 확장성이 더 높은 구조임을 알 수 있었다.

1. 서론

FGS 코딩 기법은 인터넷 스트리밍에 대한 비디오 코딩 표준의 필요성이 높아짐에 따라 나오게 되었다. FGS 인코딩 과정에서 생성되는 MPEG-4 비트스트림은 기본 계층과 향상 계층 2 개의 층으로 구분된다. 기본 계층은 각각의 프레임을 고정된 bit-rate 로 인코딩하고, 향상 계층은 원 영상의 프레임과 기본 계층에서 움직임 보상된 프레임 사이의 차이 값을 인코딩한다. 향상 계층은 bit-plane 코딩 구조에 의해 부호화됨으로써 전송 대역폭에 따라 디코더에 전송될 수 있는 bit-plane 수가 달라질 수 있다. 즉, 하나의 압축된 비트스트림으로부터 여러 비트율의 전송용 비트스트림을 생성하는 것이 가능하고, 비트율의 미세 변경이 가능하므로 가용 전송 대역폭에 최적화된 데이터의 전송을 실현할 수 있다 [1][9].

FGS 코딩 기법은 이후 테스트 과정에서 여러 가지 문제가 제기되었다 [3]. 그 가운데 예상보다 화질이 좋지 않은 문제점도 지적되었으며, 화질 향상을 위해 최근까지 여러 가지 방법들이 제안되고 있었다 [5][6]. 대표적인 방법으로는 Schaar 가 제안한 향상 계층에서 움직임 보상 루프를 갖춘 FGS (MC-FGS: motion compensation FGS) 구조를 들 수 있다 [4][그림 4].

본 논문에서는 표준화 초기에 제안되었던 Basic FGS 코딩 구조에서 향상 계층의 DCT 변환 과정 중에 발생하는 영상의 잔여 신호를 다시 한 번 bit-plane 코딩을 해주는 Advanced FGS 구조를 제안하였다. 다음 절에서는 Basic FGS, VM FGS, 그리고 MC-FGS 코딩 구조를 제시하였고, FGS 의 핵심 기법인 bit-plane 코딩 방법을 설명하였다.

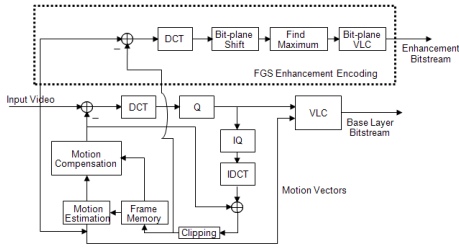
2. FGS 코딩 구조와 bit-plane 코딩 기법

(1) FGS 코딩 구조

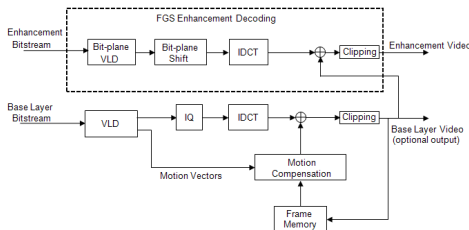
Basic FGS 인코더와 디코더의 구조는 (그림 1)과 (그림 2)와 같다. (그림 1)의 인코더 구조에서, FGS 향상 계층의 입력으로는 입력 프레임과 기본 계층에서 움직임 보상된 프레임의 차 신호 (residual signal)가 들어가게 되고, 이 신호는 8x8 블록 별로 DCT 된다. 이후 과정은 bit-plane 코딩이 이루어진다. Bit-plane 쉬프트는 특정 부분의 화질을 향상시키기 위해 특정 블록들의 bit-plane 을 상위 MSB 쪽으로 상향 천이 해주는 기능을 한다. 채널 상황이 좋지 않은 경우에도 이 부분의 전송은 보장된다.

MPEG-4 Part 2 의 VM 에서는 (그림 3)의 구조에 맞추어 FGS 코딩이 구현되어 있었다. (그림 3)의 FGS 코딩 방법은 원 영상의 DCT 와 재구성된 DCT 계수 사이에서 잔여 값을 취하는 것은 원점을 재구성된 점으로 옮기는 것과 동일하다는 이론이 적용되었다. 이 방법은 코드가 복잡하지 않고 짧으며, 처리 시간이 줄어든다는 장점이 있었다 [1]. 하지만, 이후에 향상 계층에서 움직임 보상 문제가 제기되었고, 대안으로 MC-FGS 기법이 제안되었다.

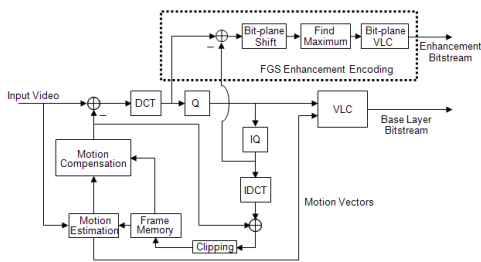
(그림 4)의 MC-FGS 구조는 향상 계층에서 화질이 저하된 상태의 기본 계층에서 나온 움직임 벡터와 움직임 보상을 이용한다는 데에서 drifting 에러에 문제점을 두어 나오게 되었다.



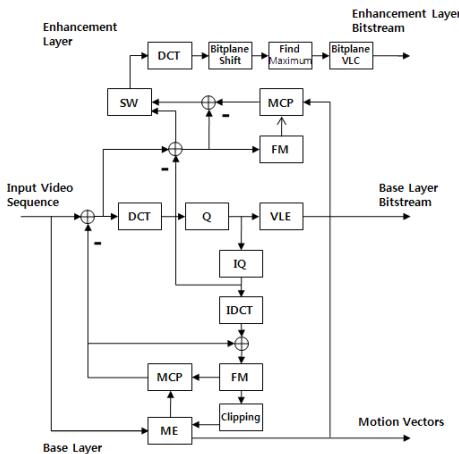
(그림 1) FGS 인코더의 기본 구조 [1]



(그림 2) FGS 디코더의 기본 구조 [1]



(그림 3) 변형된 FGS 인코더 구조 [1]



(그림 4) MC-FGS 인코더 구조 [4]

<표 1> Foreman의 PSNR에 의한 VM-FGS와 Basic FGS의 비교 (bit-plane coded: 3 level)

BL bit-rate [kbits/s]	YUV	VM FGS [dB]	Basic FGS [dB]
256	Y	32.4173	39.2463
	U	37.8538	42.8558
	V	39.3965	45.3244
512	Y	33.7618	44.8869
	U	38.7707	46.9766
	V	40.6092	48.6297
768	Y	34.7984	48.2503
	U	39.7428	49.5681
	V	41.6067	50.556
1024	Y	35.668	48.2456
	U	40.6093	49.5389
	V	42.4715	50.5369

위 <표 1>은 Basic FGS 코딩 구조를 이용하여 압축된 영상의 디스플레이되는 화질을 알아보기 위해 Foreman CIF 시퀀스로 우리가 실험한 결과를 나타내었다. PSNR 값에 의해 Basic FGS 코딩 구조는 예상만큼 고화질의 영상 압축 구조임을 알 수 있었고, VM-FGS와는 다른 결과를 나타냄을 볼 수 있었다.

(2) Bit-plane 코딩

Bit-plane 부호화 방법은 다음과 같다 [2][9].

$$\begin{bmatrix} 10 & 0 & -4 & 0 & 0 & -1 & 0 & 0 \\ 6 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(그림 5) 양자화된 DCT 블록

1) 2차원 DCT 배열을 zig-zag 순서로 1차원 배열한다.
 Value: 100600402300020010...0
 Sign: 000001001000000010...0

2) 가장 큰 절대치를 갖는 계수를 구하고, 이 계수를 2진수로 표현하기 위해 필요한 최소의 비트수 N을 구한다. N이 bit-plane의 개수가 된다.
 최대값: 10, 이진수: 1010, N: 4

3) 네 개의 bit-plane을 구성한다.
 100600402300020010...0
 MSB : 1000000000000000...0
 MSB-1: 0010010000000000...0
 MSB-2: 10100001100010000...0
 MSB-3: 000000001000000010...0

4) 2차원 심볼 (RUN, EOP)을 구성한다.
 MSB : (0,1)
 MSB-1: (2,0) (2,1)
 MSB-2: (0,0) (1,0) (4,0) (0,0) (3,1)
 MSB-3: (8,0) (6,1)
 - RUN: 1이 나오기 전의 연속된 0의 개수
 - EOP(end of plane): 현재 비트평면에서, 이후에 1의

값이 존재하는지의 여부

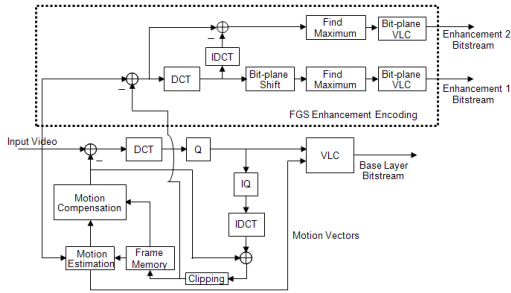
5) 2 차원 심볼과 sign 을 부호화한다.

- VLC(0,1)
- VLC(2,0), VLC(2,1)
- VLC(0,0), VLC(1,0), VLC(4,0), VLC(0,0), VLC(3,1)
- VLC(8,0), VLC(6,1), 0,0,1,0,1,0,1

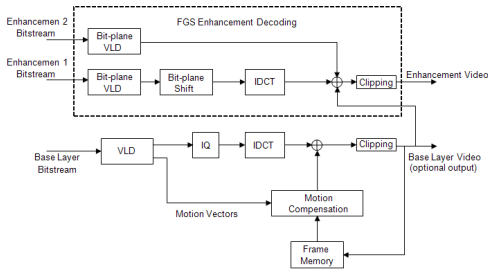
FGS 에서는 각 bit-plane 마다 PDF (probability distribution function) 특성을 고려한 독립적인 VLC 테이블을 구성하고 있다 [2].

3. 제안된 Advanced FGS 구조

본 논문에서는 Basic FGS 코딩 구조의 향상 계층에서 영상의 잔여 신호를 다시 한 번 bit-plane 코딩을 해주는 Advanced FGS 코딩 기법을 제안하였다. 구조는 (그림 6)과 (그림 7)과 같다.



(그림 6) Advanced FGS 인코더 구조



(그림 7) Advanced FGS 디코더 구조

제안된 Advanced FGS 인코더 구조는 향상 계층에서 DCT 변환 과정 중에 발생하는 영상의 잔여 신호를 다시 한 번 bit-plane 코딩을 해주는 구조를 갖는다. 제안된 구조에 의해 생성된 비트스트림은 스트리밍 전송 시 비트스트림에서 잘림 현상이 없을 경우, 완전한 복원이 가능하다. 향상 계층 2 에서 영상의 잔여 신호에 대해 필요한 bit-plane 의 개수는 보통 1~2 개가 필요하였다. 실험 및 결과는 다음 절에서 설명되었다.

4. 실험 및 결과

실험은 본 논문에서 제안된 Advanced FGS 코딩 기법이 기존의 방법보다 화질 확장성이 높은지 평가하기 위한 것이다. 실험 장비는 Intel® Core™ 2 Duo CPU E8400 @ 3.00GHz, 3.25GB RAM 을 사용하였다. 테스트 시퀀스와 실험 조건은 아래 <표 2>와 <표 3>에 나타내었다. 실험 후 영상의 화질을 평가하기 위하여 객관적 화질 평가인 PSNR (peak signal to noise ratio)을 적용하였다 [8]. 영상의 압축 부호화에서 복원 영상의 화질을 평가하기 위해 일반적으로 사용되고 있는 왜곡측도는 아래 식(1)의 평균제곱오차 (MSE: mean square error)가 있다.

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (x(i,j) - x'(i,j))^2 \quad (1)$$

위 식(1)에서 x(i, j)와 x'(i, j)는 각각 크기가 MxN 인 원 영상과 복원 영상의 화소 값을 나타낸다.

$$PSNR[dB] = 10 \log_{10} \frac{255^2}{MSE} \quad (2)$$

위 식(2)에서 255 는 첨두치 간격 (peak-to-peak)을 나타낸다. 원 영상은 일반적으로 8 비트로 양자화되기 때문에 첨두치는 255 가 된다 [7][8]. PSNR 은 peak signal power 와 noise power 비율의 log 표현 방법으로 PSNR 값이 클수록 원 영상에 근접함을 나타낸다. 본 논문의 실험에서는 기본 계층에 대한 비트 레이트를 256, 512, 768, 1024 로 다르게 조절하면서 YUV 각각의 성분에 대하여 PSNR 값을 구하였다.



(그림 8) 실험에 사용된 Foreman, Soccer, Container 의 프레임

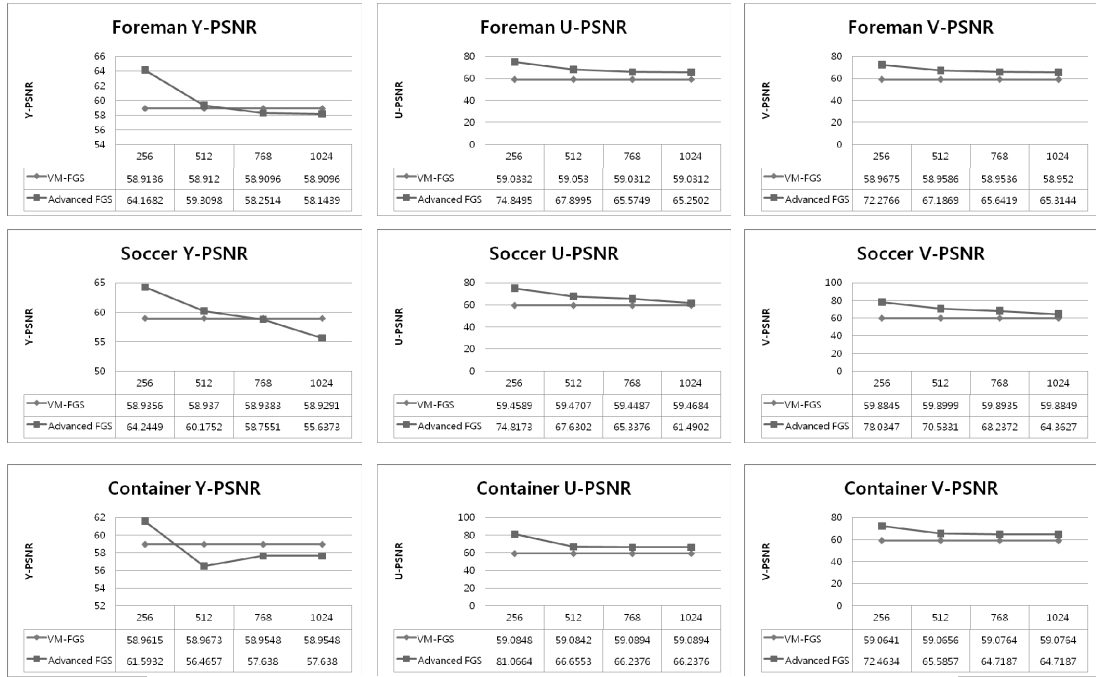
<표 2> 테스트 시퀀스

Sequences	Characteristics
Foreman	CIF@30fps
Soccer	CIF@30fps
Container	CIF@30fps

<표 3> 실험 조건

항목	값	항목	값
Rate control	VM5+	M	1
Deblocking Filter	use	bp-coded	6 level
BL bit-rate (Kbps)	256, 512, 768, 1024		

3 개의 영상에서 YUV 각각의 성분에 대한 PSNR 값의 결과는 다음과 같다.



(그림 9) Foreman, Soccer, Container 시퀀스의 YUV 각 성분에 대한 PSNR

위 결과를 통해 bit-plane 코딩 6 레벨을 사용했을 경우, 3 개의 영상에서 Y-PSNR 은 평균적으로 약 0.4dB 높게 나타났고, U-PSNR 은 약 9.4dB 그리고 V-PSNR 은 약 9dB 높게 나타났다. 그러나 Y-PSNR 이 기본 계층에서 저 비트율로 압축될 경우에는 높은 수치를 보이지만, 기본 계층의 비트 레이트를 높임에 따라, VM-FGS 보다 화질이 떨어지고 있다. 이 부분은 제안된 구조에 적합한 레이트 컨트롤의 사용과 그에 따른 bit-plane 의 분할을 알맞게 조절해준다면, YUV 성분 모두 VM-FGS 보다 고르게 높이 나타날 것이다.

5. 결론

5 절의 실험 결과에 의해, Advanced FGS 코딩 구조는 최소~최대 비트율 범위에서 최저부터 최고의 화질을 가질 수 있는 화질 확장성이 높은 구조임을 알 수 있었다. 또한, 스트리밍 전송 시 비트스트림에서 잘림 현상이 없을 경우, 완전하게 복원할 수 있는 가능성을 보여 주었다. 향후 연구는 Advanced FGS 코딩 구조에 적합한 레이트 컨트롤의 설계와 bit-plane 의 알맞은 조절이 요구된다.

참고문헌

[1] W.Li, "Overview of fine granularity scalability in MPEG-4 video standard" IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 3, pp. 301-317, March 2001.
 [2] ISO/IEC JTC1/SC29/WG11, N3908, "MPEG-4 Video Verification Model version 18.0," Pisa, January 2001.

[3] ISO/IEC JTC1/SC29/WG11, N6710, "MPEG-4 Visual: List of Problems Reported v.14," Palma de Mallorca, October 2004.
 [4] M. van der Schaar and H. Radha, "Motion-Compensation based Fine-Granular Scalability(MC-FGS)", ISO/IEC JTC1/SC29/WG11, MPEG 2000/M6475, October 2000.
 [5] Su-Ren Chen, Chen-Po Chang and Chia-Wen Lin, "MPEG-4 FGS coding performance improvement using adaptive inter-layer prediction," Proc. ICASSP '04, IEEE International Conf. on Acoustics, Speech, and Signal Processing, ICASSP, 2004, vol. 3, pp. iii - 256-8, May 2004.
 [6] Jian Zhou, Huairong Shao Chia Shen and Ming-Ting Sun, "FGS enhancement layer truncation with minimized intra-frame quality variation," Proc. 2003 International Conf. on Multimedia and Expo, ICME, 2003, vol. 2, pp. II - 361-4, July 2003.
 [7] "Infomation technology - Coding of audio-visual objects - Part-2: Visual," International Standard, ISO/IEC JTC1/SC29/WG11, ISO/IEC 14496-2:2001(E), Dec. 2001.
 [8] D. K. Fibush, "Practical application of objective picture quality measurements," International Broadcasting Convention, no. 447, pp. 504-513, Sept. 1997.
 [9] 서광덕, 권순각, "MPEG-4 Fine Granular Scalability 기술을 이용한 차세대 비디오 스트리밍," 멀티미디어학회지, 제 7 권, 제 4 호, pp.59-66, Dec. 2003.