

고화질 확장성을 위한 FGS 코딩 구조의 개선

부 희 형[†] · 김 승 호^{††}

요 약

MPEG-4 파트 2에서 화질 확장성을 지원하는 FGS (fine granularity scalability) 코딩 기법은 가변적인 네트워크 대역폭에서 최적의 화질을 구현할 수 있는 스케일러블 비디오 코딩 기법이다. 본 논문에서는 기본 FGS 코딩 구조의 향상 계층에서, 영상의 잔여 신호를 다시 한 번 더 비트 평면 (bit-plane) 코딩을 해주는 FGS 코딩 구조를 제안하였다. 실험에서는 제안된 FGS 코딩 구조와 MPEG-4 VM (verification model) FGS 코딩 구조를 비교하여 제안된 FGS 코딩 구조의 화질 확장성을 평가하였다. 비교는 테스트된 3개의 비디오의 PSNR의 분석으로 이루어졌다. 결과는 제안된 구조가 비트율 제어 기법 VM5+를 이용했을 때, VM-FGS 코딩 구조보다 Y, U, V PSNR이 평균적으로 각각 0.4 dB, 9.4 dB, 9 dB 더 높게 나타났고, QP (quantization parameter) 값을 17로 고정시킨 경우는, 제안된 구조가 4.61 dB, 20.21 dB, 16.56 dB 더 높게 나타났다. 결과에 의해, 제안된 구조가 VM-FGS 코딩 구조보다 최저부터 최대의 화질을 얻을 수 있는 화질 확장성이 더 높은 구조임을 알 수 있었다.

키워드 : FGS 코딩, 고화질 확장성, 비트플레인 코딩

An improvement in FGS coding scheme for high quality scalability

Hee-Hyung Boo[†] · Sung-Ho Kim^{††}

ABSTRACT

FGS (fine granularity scalability) supporting scalability in MPEG-4 Part 2 is a scalable video coding scheme that provides bit-rate adaptation to varying network bandwidth thereby achieving of its optimal video quality. In this paper, we proposed FGS coding scheme which performs one more bit-plane coding for residue signal occurred in the enhancement-layer of the basic FGS coding scheme. The experiment evaluated in terms of video quality scalability of the proposed FGS coding scheme by comparing with FGS coding scheme of the MPEG-4 verification model (VM-FGS). The comparison was conducted by analysis of PSNR values of three tested video sequences. The results showed that when using rate control algorithm VM5+, the proposed FGS coding scheme obtained Y, U, V PSNR of 0.4 dB, 9.4 dB, 9 dB averagely higher and when using fixed QP value 17, obtained Y, U, V PSNR of 4.61 dB, 20.21 dB, 16.56 dB averagely higher than the existing VM-FGS. From results, we found that the proposed FGS coding scheme has higher video quality scalability to be able to achieve video quality from minimum to maximum than VM-FGS.

Keywords : FGS Coding, High Quality Scalability, Bit-plane Coding

1. 서 론

FGS 코딩 기법은 인터넷 비디오 스트리밍 기술에서 최적 전송을 위한 표준의 필요성이 높아짐에 따라 관심 받는 주제가 되었다. FGS 코딩 구조는 기본 계층과 향상 계층의 2개의 층으로 구분되는 MPEG-4 비트 스트림을 생성한다. FGS 인코더의 기본 계층은 각각의 프레임에 고정된 비트율 (bit-rate)로 인코딩하고, 향상 계층은 원 영상의 프레임과

기본 계층에서 움직임 보상된 프레임 사이의 차이 값을 인코딩한다. 향상 계층은 비트 평면 코딩 구조에 의해 부호화 되고, 전송 대역폭에 따라 디코더에 전송될 수 있는 비트 평면의 수는 달라질 수 있다. 즉, 하나의 압축된 비트 스트림으로부터 여러 비트율의 전송용 비트 스트림을 생성하는 것이 가능하고, 비트율의 미세 변경이 가능하므로 가용 전송 대역폭에 최적화된 데이터의 전송을 실현할 수 있다[1,9].

FGS 코딩 기법은 제안된 이후 테스트 과정에서 몇 가지 문제가 제기되었다[3]. 그 가운데 화질이 예상보다 좋지 않은 문제점도 지적되어, 화질 향상을 위한 여러 방법들이 제안되고 있었다[4,5,6]. 대표적인 방법으로는 Schaar et al (2000)이 제안한 향상 계층에서 움직임 보상 루프를 갖춘 FGS (MC-FGS: motion compensation FGS) 구조를 들 수 있다[4].

* 이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No.2011-0004184).

† 준 회 원 : 경북대학교 전자전기컴퓨터학부 박사과정

†† 종신회원 : 한국과학기술원 전산학과 공학박사

논문접수 : 2011년 2월 2일
수정일 : 1차 2011년 6월 10일
심사완료 : 2011년 6월 16일

본 논문에서는, MPEG-4 표준화 과정에서 처음 제안되었던 기본 FGS 코딩 구조를 이용한 방법으로서, 향상 계층의 DCT 변환 과정에서 발생하는 영상의 잔여 신호를 다시 한번 더 비트 평면 코딩을 해주는 FGS 코딩 구조를 제안하였다. 결과에 의해 제안된 구조가 VM5+를 이용했을 때, VM-FGS 코딩 구조보다 Y, U, V PSNR이 평균적으로 각각 0.4 dB, 9.4 dB, 9 dB 더 높게 나타났고, QP 값을 17로 고정시킨 경우는, 제안된 구조가 4.61 dB, 20.21 dB, 16.56 dB 더 높게 나타났다.

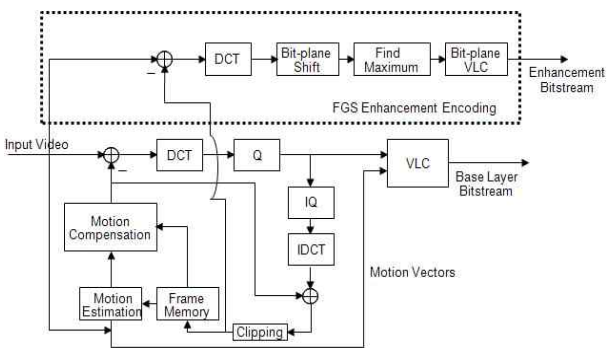
다음 장에서는 기본 FGS, VM-FGS, MC-FGS 코딩 구조와 FGS의 핵심 기법인 비트 평면 코딩 방법을 설명하였다. 3장에서는 본 논문에서 제안된 FGS 코딩 구조에 대하여 상세히 서술하였고, 4장에서는 3개의 비디오에 대하여 테스트 방법 및 결과를 보였다. 실험 결과는 테스트된 비디오들의 PSNR을 그래프로 표현하였다. 마지막으로 5장에서는 실험 결과로부터 결론을 맺는다.

2. FGS 코딩 구조와 비트 평면 코딩 기법

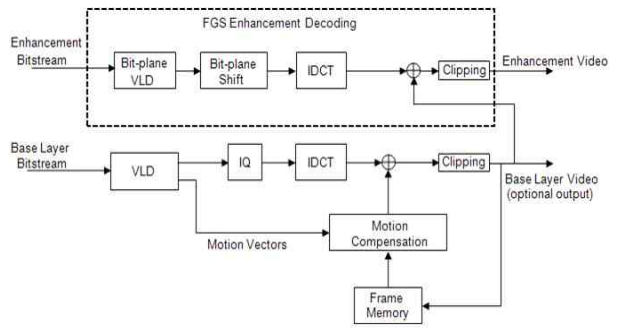
2.1 FGS 코딩 구조

기본 FGS 인코더와 디코더의 구조는 아래 (그림 1)과 (그림 2)와 같다. (그림 1)의 인코더 구조에서, FGS 향상 계층의 입력은 기본 계층의 입력 프레임과 기본 계층에서 움직임 보상된 프레임의 차 신호 (residual signal)가 된다. 이 신호는 8x8 블록 별로 DCT되고, 비트 평면 쉬프트, 최대값 구하기, 비트 평면 VLC 단계를 거쳐 최종 향상 계층의 비트 스트림을 생성한다. 여기서, 비트 평면 쉬프트는 특정 부분의 화질을 향상시키기 위해 특정 블록들의 비트 평면을 상위 MSB 쪽으로 상향 천이해주는 기능으로서, 채널 상황이 좋지 않은 경우에도 이 부분의 전송은 보장된다[1,9].

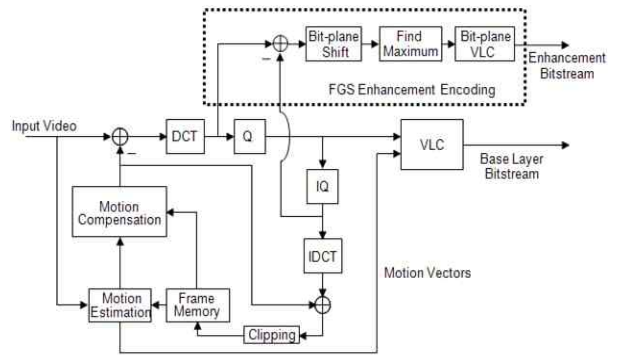
MPEG-4 VM에서는 아래 (그림 3)의 구조에 맞추어 코딩이 되어 있었다. 이 방법은 원 영상의 DCT 계수와 재구성된 DCT 계수 사이에서 잔여 값을 취하는 것은 원점을 재구성된 점으로 옮기는 것과 동일하다는 이론이 적용되었다. 이 구조는 코딩이 간단하고 처리 시간이 줄어든다는 장점이 있었지만 이후에 향상 계층에서 움직임 보상 문제가 논의되었다.



(그림 1) FGS 인코더의 기본 구조

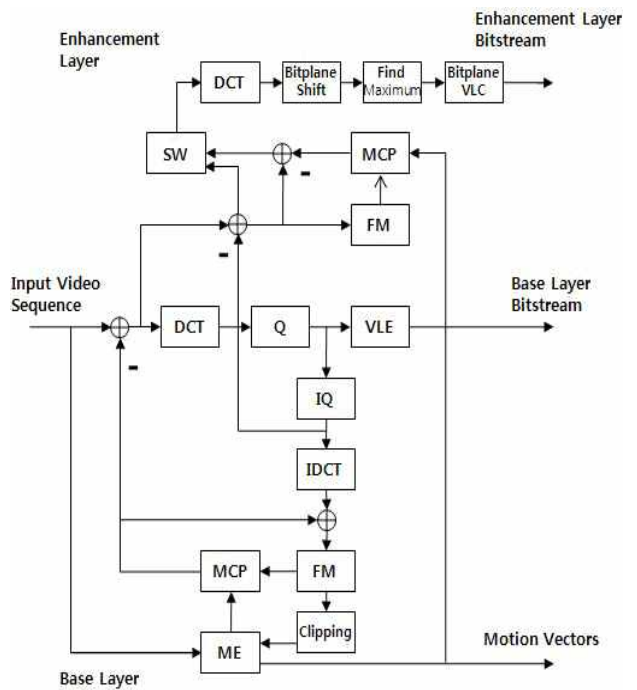


(그림 2) FGS 디코더의 기본 구조



(그림 3) 변형된 FGS 인코더 구조

아래 (그림 4)는 MC-FGS 구조를 나타낸다. 이 구조는 화질이 저하된 상태의 움직임 벡터와 움직임 보상을 이용하는 과정에서 발생하는, VM-FGS의 에러 전파 (drifting error)에 대한 해결 방안으로 제안된 구조이다.



(그림 4) MC-FGS 인코더 구조

<표 1> VM-FGS와 기본 FGS의 비교를 위한 Foreman 비디오의 PSNR (비트 평면 코딩: 3 level)

BL 비트율 [Kbits/s]	YUV	VM-FGS [dB]	기본 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>에서는 기본 FGS 코딩 구조가 VM-FGS 코딩 구조보다 더 높은 화질을 나타내는지 알아보기 위해 CIF (352 x 288) Foreman 비디오의 실험 결과를 나타내었다. PSNR에 의해 기본 FGS 코딩 구조는 예상대로 VM-FGS 코딩 구조보다 더 높은 화질을 보였다.

2.2 비트 평면 코딩

다음은 비트 평면 부호화 절차를 예를 들어 설명하였다 [1,2,9].

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

(그림 5) 양자화된 DCT 블록

1) 위 (그림 5)에서 2차원 DCT 배열을 zig-zag 순서로 1차원 배열한다.

값: 10 0 6 0 0 4 0 2 3 0 0 0 2 0 0 1 0...0
부호 : 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0...0

2) 가장 큰 절대치를 갖는 계수를 구하고, 이 계수를 2진수로 표현하기 위한 최소의 비트수 N을 구한다. N이 비트 평면의 개수가 된다.

최대값: 10, 이진수: 1010, N: 4

3) 네 개의 비트 평면을 구성한다.

10 0 6 0 0 4 0 2 3 0 0 0 2 0 0 1 0...0
MSB : 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0...0
MSB-1: 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0...0
MSB-2: 1 0 1 0 0 0 0 1 1 0 0 0 1 0 0 0 0...0
MSB-3: 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0...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) 2차원 심볼과 부호를 부호화한다.

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

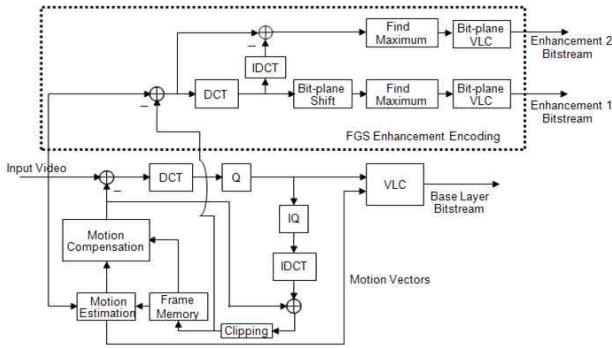
FGS에서는 각 비트 평면마다 PDF (probability distribution function) 특성을 고려하여 독립적인 VLC 테이블을 구성하고 있다[2].

3. 제안된 FGS 코딩 구조

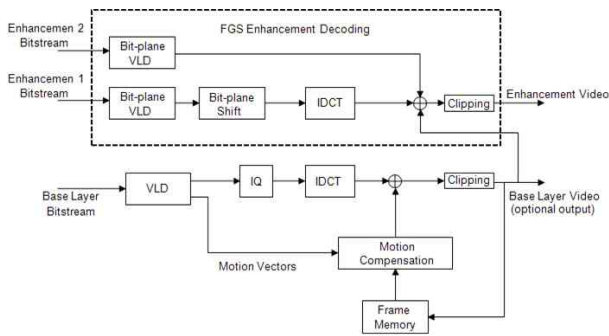
본 논문에서는 기본 FGS 코딩 구조의 항상 계층에서 영상의 잔여 신호를 다시 한 번 더 비트 평면 코딩을 해주는 FGS 코딩 구조를 제안하였다. 구조는 아래 (그림 6)과 (그림 7)과 같다.

(그림 6)의 제안된 인코더 구조는, 항상 계층 1에서 발생하는 영상의 잔여 신호를 항상 계층 2에 포함시키기 위한 방법으로 제안되었다. 이 방법은 항상 계층 2에서, 항상 계층 1의 DCT와 IDCT의 차이를 미리 계산하고 비트 평면 코딩을 하여 항상 계층 1에서 생성된 비트 스트림과 함께 보내주는 방법이다. 항상 계층 2의 입력 신호는 항상 계층 1의 입력신호와 그 입력 신호의 DCT, IDCT 과정을 수행한 신호와의 차이 신호가 된다. 이 차이 신호를 입력으로 하여, 최대값 구하기, 비트 평면 VLC 단계를 수행하면 최종 항상 계층 2의 비트 스트림을 생성하게 된다.

(그림 7)의 제안된 디코더 구조는, 제안된 FGS 인코더 구조의 역 과정으로서, 기본 계층에서 생성되는 신호와 항상 계층 1에서 생성되는 신호, 항상 계층 2에서 비트 평면 VLD (variable length decoder) 과정을 거쳐 생성되는 신호, 총 3개의 신호를 더하여 국제표준에서 권고하는 값의 범위로 클리핑 (Clipping) 단계를 수행하여 최종 향상된 비디오 신호를 생성한다.



(그림 6) 제안된 FGS 인코더 구조



(그림 7) 제안된 FGS 디코더 구조

즉, 제안된 FGS 인코더 구조는 기본 FGS 코딩 구조에서, 항상 계층의 DCT 변환 과정에 발생하는 영상의 잔여 신호를 다시 한 번 더 비트 평면 코딩을 수행하는 구조이다. 제안된 구조에 의해 생성된 비트 스트림은 스트리밍 전송 시 비트 평면의 잘림 현상이 없을 경우, 완전한 복원이 가능할 수 있다. 제안된 구조의 항상 계층 2에서, 영상의 잔여 신호에 대해 필요한 비트 평면의 개수는 보통 1~2개가 필요하였다. 실험 및 결과는 다음 장에서 설명하였다.

4. 실험 및 결과

본 논문의 실험은 제안된 FGS 코딩 기법이 VM-FGS 코딩 기법보다 화질 확장성이 더 높은지 평가하기 위한 것이다. 실험 장비는 Intel® Core™ 2 Duo CPU E8400 @ 3.00GHz, 3.25GB RAM 을 사용하였다. 3개의 테스트 비디오는 아래 (그림 8)에 나타내었고, 테스트 비디오와 실험 조건은 <표 2>와 <표 3>에 나타내었다[7].



(그림 8) 실험에 이용된 Foreman, Soccer, Container의 비디오

<표 2> 테스트 비디오

비디오	특성
Foreman	CIF@30fps
Soccer	CIF@30fps
Container	CIF@30fps

<표 3> 실험 조건

항목	값	항목	값
Rate control	VM5+, QP 17고정	M	1
Deblocking Filter	사용	비트 평면 코딩	6 level
BL 비트율(kbps)	256, 512, 768, 1024	Intra period	8

* M : 1 (B-frame 사용 안 함)

테스트된 비디오의 화질 평가는 객관적 화질 평가인 PSNR (peak signal to noise ratio)을 이용하였다. 영상의 압축 부호화에서 복원 영상의 화질을 평가하기 위해 일반적으로 이용되고 있는 왜곡 척도는 아래 식 (1)의 평균제곱오차 (MSE: mean square error)가 있다[8].

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

위 식(1)의 $x(i,j)$ 와 $\hat{x}(i,j)$ 는 각각 크기가 $M \times N$ 인 원 영상과 복원 영상의 화소 값을 나타낸다. PSNR의 식은 아래 식 (2)와 같다.

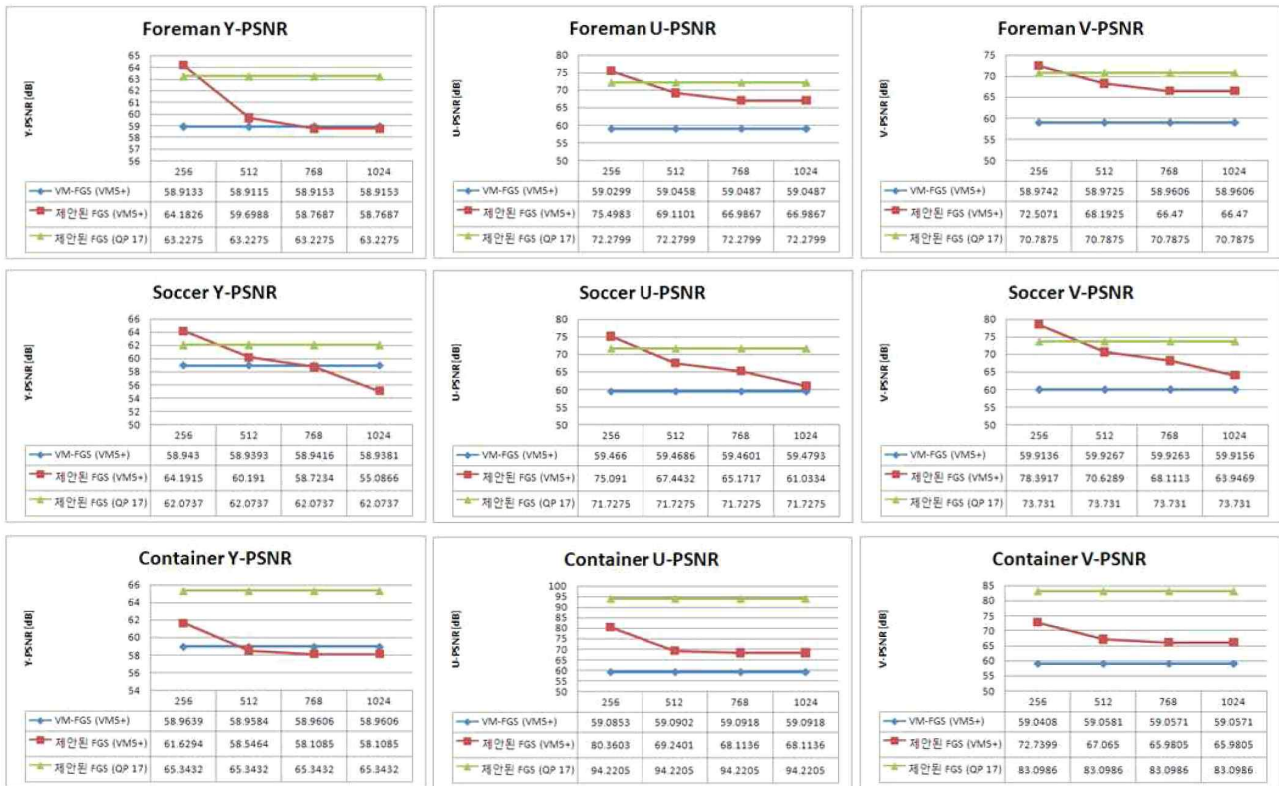
$$PSNR [dB] = 10 \log_{10} \frac{255^2}{MSE} \quad (2)$$

식 (2)에서 255는 첨두치 간격 (peak-to-peak)을 나타낸다. 원 영상은 일반적으로 8비트로 양자화되기 때문에 첨두치는 255가 된다. PSNR은 최대 신호 전력 (peak signal power)과 잡음 전력 (noise power) 비율의 log 표현 방법으로서, PSNR 값이 클수록 원 영상에 근접함을 나타낸다.

본 논문의 실험에서는 기본 계층에 대하여 비트율을 256, 512, 768, 1024 Kbps로 다르게 조절하면서 Y, U, V 각각의 성분에 대하여 PSNR을 구하였다.

아래 (그림 9)는 제안된 FGS 코딩 구조와 VM-FGS 코딩 구조에서 테스트된 3개의 비디오의 Y, U, V에 대한 PSNR을 그래프로 표현하였다.

결과에 의해, 제안된 구조가 VM5+를 이용했을 때, VM-FGS 코딩 구조보다 Y, U, V PSNR이 평균적으로 각각 0.4 dB, 9.4 dB, 9 dB 더 높게 나타났고, QP 값을 17로 고정시킨 경우는, 제안된 구조가 4.61 dB, 20.21 dB, 16.56 dB 더 높게 나타났다. 그러나 VM5+를 이용한 경우,



(그림 9) Foreman, Soccer, Container 비디오의 Y, U, V 성분에 대한 PSNR

Y-PSNR이 기본 계층의 비트율을 높임에 따라 VM-FGS 코딩 구조보다 화질이 떨어지고 있다. 이것은 QP 값에 영향을 받기 때문이고, QP 값을 제안된 구조에 적합하게 조절한다면, Y, U, V 성분에서 모두 VM-FGS보다 고르게 높여 나타날 것이다.

5. 결론

본 논문에서는 기본 FGS 코딩 구조에서, 항상 계층의 DCT 변환 과정에서 발생하는 영상의 잔여 신호를 다시 한번 더 비트 평면 코딩을 해주는 FGS 코딩 구조를 제안하였다. 실험 결과, 제안된 구조가 VM5+를 이용했을 때, VM-FGS 코딩 구조보다 Y, U, V PSNR이 평균적으로 각각 0.4 dB, 9.4 dB, 9 dB 더 높게 나타났고, QP 값을 17로 고정시킨 경우는, 제안된 구조가 4.61 dB, 20.21 dB, 16.56 dB 더 높게 나타났다.

결과적으로, 제안된 구조가 VM-FGS 코딩 구조보다 최저부터 최고의 화질을 얻을 수 있는 화질 확장성이 더 높은 구조임을 알 수 있었다. 또한, 스트리밍 전송 시 비트 평면의 잘림 현상이 없을 경우, 완전하게 복원할 수 있는 가능성을 보여 주었다. 향후 연구는 제안된 FGS 코딩 구조에 적합한 QP값을 결정하는 비트율 제어 기법과 비트 평면의 알맞은 조절이 요구된다.

참고 문헌

- [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, 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, 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, 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] <http://trace.eas.asu.edu/yuv/index.html>
- [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 7권, 제 4호, pp.59-66, Dec., 2003.



부희형

e-mail : hhboo@knu.ac.kr

2004년 목포대학교 컴퓨터공학과(학사)
2006년 전남대학교 컴퓨터정보통신공학과
(공학석사)

2008년~현 재 경북대학교 전자전기
컴퓨터학부 박사과정

관심분야: 동영상 압축, 그래픽스, 컴퓨터 시뮬레이션 등



김승호

e-mail : shkim@knu.ac.kr

1981년 경북대학교 전자공학과(학사)

1983년 한국과학기술원 전산학과
(공학석사)

1994년 한국과학기술원 전산학과
(공학박사)

관심분야: 알고리즘, 멀티미디어, 다시점 동영상, 감시 시스템,
동기식 이더넷 등