

H.264 기반 선택적인 미세입자 스케일러블 코딩 방법

(A H.264 based Selective Fine Granular Scalable Coding Scheme)

박 광 훈[†] 유 원 혁^{**} 김 규 현^{***}
(Gwang Hoon Park) (Won Hyuck Yoo) (Kyuheon Kim)

요약 본 논문에서는 선택적으로 강화계층의 시간적 예측 정보를 사용하는 H.264 기반 선택적인 미세입자 스케일러블 비디오 코딩 방법을 제안한다. 제안된 방법의 기본계층은 최근 표준화되었으며 고압축률이 특징인 H.264 (MPEG-4 Part 10 AVC) 알고리즘으로 코딩한다. 강화계층은 기본적으로 국제표준인 비트플레인 기반 MPEG-4 (Part 2) 미세입자스케일러블 코딩 방법으로 코딩한다. 본문에서는 엔코더측에서 강화계층간 시간적 예측 방법에 의해 발생된 효과적인 영상 정보를 드리프트 현상이 최소화 하는 방향으로 선택적으로 적용하여 디코더측에 전송하는 방안을 제안하였다. 강화계층간 예측 방법만을 추가했을 때는 시간적 중복성을 줄여주는 효과를 볼 수 있지만 저비트율 대역에서 엔코더와 디코더간의 참조 저장 영상 불 일치로 인한 드리프트 현상이 심하게 발생한다. 제안된 알고리즘은 시간적 예측 결과가 우수하여 코딩효율을 현저히 높여줄 수 있는 경우에만 선택적으로 강화계층간의 시간적 예측 정보를 사용하였으며, 이로 인하여 저비트율 대역에서의 드리프트 현상을 현저하게 줄일수 있었으며, 전반적으로 코딩 효율을 높여주는 효과를 가져왔다. 여러 영상 시퀀스를 대상으로 실험한 결과, 제안된 코딩 방법은 현존하는 국제표준인 MPEG-4 기반 미세입자 스케일러블 코딩 방법보다 같은 비트율 대역에서 영상화질이 약 3~5 dB 높은 성능을 보여주고 있으며, H.264를 기반으로한 미세입자 스케일러블 코딩 방법보다도 약 1~3 dB 높은 성능을 보여주고 있음을 발견할 수 있었다.

키워드 : 동영상코딩, MPEG-4, H.264, 비디오코딩

Abstract This paper proposes the H.264-based selective fine granular scalable (FGS) coding scheme that selectively uses the temporal prediction data in the enhancement layer. The base layer of the proposed scheme is basically coded by the H.264 (MPEG-4 Part 10 AVC) visual coding scheme that is the state-of-art in codig efficiency. The enhancement layer is basically coded by the same bitplane-based algorithm of the MPEG-4 (Part 2) fine granular scalable coding scheme. In this paper, we introduce a new algorithm that uses the tempoal prediction mechanism inside the enhancement layer and the effective selection mechanism to decide whether the temporally-predicted data would be sent to the decoder or not. Whenever applying the temporal prediction inside the enhancement layer, the temporal redundancies may be effectively reduced, however the drift problem would be severly occurred especially at the low bitrate transmission, due to the mismatch bewteen the encoder's and decoder's reference frame images. Proposed algorithm selectively uses the temporal-prediction data inside the enhancement layer only in case those data could significantly reduce the temporal redundancies, to minimize the drift error and thus to improve the overall coding efficiency. Simulation results, based on several test image sequences, show that the proposed scheme has 1~3 dB higher coding efficiency than the H.264-based FGS coding scheme, even 3~5 dB higher coding efficiency than the MPEG-4 FGS international standard.

Key words : MPEG-4, H.264, Video Signal Processing, Moving Picture Coding

[†] 종신회원 : 경희대학교 전자정보학부 교수

ghpark@khu.ac.kr

^{**} 학생회원 : 경희대학교 컴퓨터공학과

wonhy@video.khu.ac.kr

^{***} 정회원 : 한국전자통신연구원 대화형미디어연구팀 팀장

kkim@etri.re.kr

논문접수 : 2004년 2월 4일

심사완료 : 2004년 3월 30일

1. 서론

현재의 정보기술분야는 빠르게 통합되어 가는 추세에 있고, 방송/통신망의 융합이 진행됨에 따라, 방송망, 유/무선 인터넷 등의 서로 다른 대역폭을 가지는 다양한

통신망과 디지털 TV, HDTV, PDA, 노트북, IMT-2000 등과 같은 다양한 화질과 연산능력을 가지는 단말기들이 혼재하고 있는 상황에서, 다양한 통신 환경에서 각각의 사용자에게 주어진 환경에 맞도록 동영상을 전송할 수 있는 스케일러블 비디오 코딩 방법이 필요하다.

MPEG-4(Part 2 Visual) 기반 미세입자 스케일러블 코딩(fine granular scalability: FGS) 방법은 ISO/IEC의 국제 표준안으로서 가변적인 네트워크 전송 선로의 상황에 주어진 대역폭(bandwidth)에서 최적의 화질 향상을 얻을 수 있도록, 국제적으로 최초로 설계된 망상태 적응적인(network-adaptive) 스케일러블 비디오 코딩 방법이지만[1-3], 코딩 효율면에 있어서 최적화된 단일계층(single layer) 코딩 방법에 비하여 성능이 현저히 떨어진다는 평가를 받고 있다[4-5]. 따라서, MPEG-4 기반 FGS와 같이 망상태 적응적이라는 장점을 그대로 유지하면서 단일계층 코딩 방법에 버금가는 코딩효율을 유지할 수 있는 방법론이 지속적으로 요구되고 있다.

본 논문에서는 강화 계층내에서 시간적 예측(temporal-prediction) 정보를 선택적으로 이용하여 코딩효율이 향상된 FGS 코딩 방법을 제안한다. 제안된 코딩 방법의 기본 계층(base layer) 코딩 방법으로 H.264(MPEG-4 Part 10 AVC)를 사용하며, 강화계층(enhancement layer) 코딩 방법으로서 국제 표준안인 MPEG-4 기반 FGS 방법의 비트플레인(bitplane) 기반의 강화계층 코딩 방법을 기본적으로 사용하고, 여기에 강화계층간의 시간적 예측 방법을 적용하여 시간적 예측 정보를 추출하고 이를 코딩효율을 극대화하는 방향으로 선택적으로 이용하는 방법을 제안하여 사용하였다.

본 논문의 구성은 2~3장에서 MPEG-4와 H.264기반의 FGS를 소개하였으며, 4장에서 제안된 선택적인 FGS 방법을 소개하였고, 실험결과를 5장에서 다루었으며 6장에서 결론을 맺었다.

2. MPEG-4 기반 FGS 코딩 방법

MPEG-4 기반 FGS 코딩 방법은 스케일러블 코딩을 수행할 때, 전송선로의 상태가 안정적이지 못한 경우에, 예를 들어 유/무선 인터넷 등과 같이 전송선로의 상태가 급변하여 각 사용자에게 할당할 수 있는 대역폭의 변동이 심한 경우에 적용하기 위하여 설계된 코딩 방법으로, 강화계층의 영상정보를 비트플레인 단위로 코딩하여 전송하는 특징을 갖고 있다. 강화계층에서 강화정보를 비트플레인별로 나누어 가장 중요한 비트(most significant bit: MSB)를 최우선적으로 전송하고, 그 다음 중요한 비트를 비트플레인별로 나누어 연속적으로 전송하는 방식이다. 수신단측에서 전송선로의 대역폭이

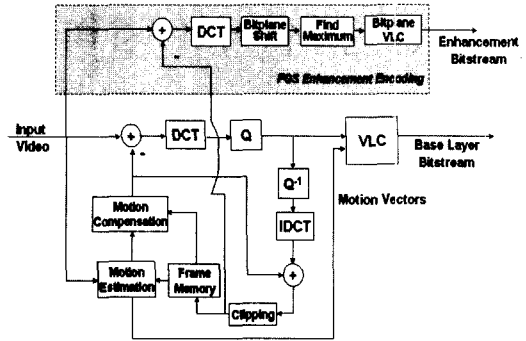


그림 1 MPEG-4 기반 FGS 코딩 방법의 엔코더 구성도

급변하여 영상 복원에 필요한 모든 비트를 수신하지 못한 경우에도, 그때까지 수신된 부분적인 영상 비트스트림만을 이용하여 최적의 화질 향상을 얻을 수 있도록 설계되었다.

MPEG-4 기반 FGS 코딩 방법의 엔코더단 구조는 그림 1에 도시된 바와 같으며, 기본계층은 단일계층에서 최적화된 MPEG-4 Part 2 visual 엔코딩 방법인 복합화된(hybrid) 움직임 추정(motion estimation: ME) 및 움직임 보상(motion compensation: MC)과 이산적 코사인 변환(discrete cosine transform: DCT) 방식을 이용하고 있다. 강화계층에서 발생하는 영상정보를 코딩하기 위하여 FGS 코딩방법은 원 영상(original image)과 기본계층에서 재생된 영상(reconstructed image)과의 오차영상(residual image)을 구하는 과정을 거친다. DCT를 수행하는 과정에서는 앞에서 결정된 오차영상 값을 블록 단위(8*8 단위) DCT를 이용하여 DCT 계수로 변환한다. 이때, 선택적으로 좋은 화질을 갖는 블록이 필요하다면 해당 블록에 위치한 영상정보를 최우선적으로 전송하는 과정을 선택적으로 수행할 수 있다. 이를 선택적 강화(selective enhancement: SE)라고 정의하며, 비트플레인 쉬프트('bitplane shift') 단계에서 수행한다. 최대값 찾기('find maximum') 단계에서는 DCT를 수행한 모든 값들의 절대치에 대한 최대값을 구하며, 그 값들은 해당 영상 프레임을 전송하기 위한 최대 비트플레인의 갯수를 구하는데 사용된다. 비트플레인 가변장 부호화(bitplane variable length coding: VLC) 단계에서는 비트플레인별로 블록 단위로 구해진 64개의 DCT 계수들(DCT 계수의 해당 비트플레인의 비트: 0 또는 1)이 지그재그 스캔(zigzag scan) 순서로 하나의 행열에 들어가게 되며, 각 행렬들은 가변장 길이 코드표(VLC table)에 따라 런-길이(run-length) 부호화된다.

그림 2는 MPEG-4 기반 FGS 코딩 방법의 디코더단의 구조도이다. 그림에 도시된 바와 같이, 기본계층과 강화계층으로 나뉘어져서 전송선로를 통하여 수신된 비

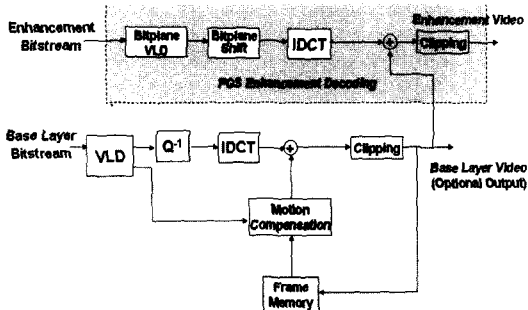


그림 2 MPEG-4 기반 FGS 코딩 방법의 디코더 구성도

트스트림의 디코딩을 그림 1에 도시된 엔코더와 역순으로 수행하게 된다. 기본계층에서는 MPEG-4 영상 디코딩 방법인 역방향 hybrid MC/DCT방식을 그대로 사용한다. 강화계층에서는 디코더단에 입력된 강화 비트스트림(enhancement bitstream)을 비트플레인별로 가변길이 디코딩('bitplane VLD' 단계)을 수행한 후, 선택적으로 좋은 화질을 갖는 블록의 위치 등이 전송되었다면, 비트플레인 쉬프트('bitplane shift' 단계)를 선택적으로 수행한다. 이후, 비트플레인별로 VLD가 수행되고 선택적으로 비트플레인 쉬프트가 수행되어 도출된 값에, 블록 단위의 역 이산적 코사인 변환('IDCT' 단계)을 수행하여 강화계층에서 전송된 영상을 복원한 후, 기본계층에서 복호된 영상과 합치고 해당 합산 값들을 0과 255 값 사이로 클립핑('Clipping' 단계)하여 최종적으로 향상된 화질의 영상을 복원하게 된다[1-3].

3. H.264(MPEG-4 AVC) 코딩 방법

차세대 동영상 압축 기술로 주목을 받고있는 H.264는 MPEG-4 Part 10 AVC(advanced video coding)라고도 불리는 코딩 방식으로, ISO/IEC 산하의 MPEG(Moving Picture Expert Group)과 ITU-T 산하의 VCEG(Video Coding Expert Group)이 하나의 동영상 압축 표준을 제정하기로 합의하여 표준화를 완료한 것이다. H.264는 압축효율의 개선에 초점이 맞추어져 있으며, 특히 H.263+과 비교하여, 50%이상의 압축효율 개선을 목표로 개발되었다[6-10].

그림 3은 H.264의 엔코더 구조를 보여주고 있다. 영상 압축 단위는 16*16 크기의 매크로블록(macroblock) 단위이며, 각 단위의 부호화는 인접 영상과의 상관 관계에 따라 결정되어 '인트라(intra)' 모드 혹은 '인터(inter)' 모드로 부호화하게 된다.

인트라 매크로블록은 시간적으로 인접하는 영상과의 시간적 상관 관계가 적거나 또는 의도적으로 움직임 보상(MC)을 하지 않고 부호화하는 경우로, 움직임 벡터를

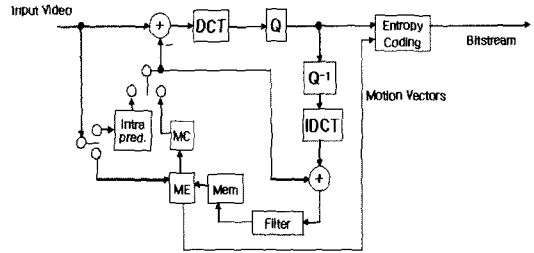


그림 3 H.264(MPEG-4 Part 10 AVC) 엔코더 구성도

전송하지 않고 16*16블록 또는 4*4블록 단위로 인트라 예측(intra prediction)을 수행한 후 예측 값을 입력 매크로블록에서 차분을 취하여 이 차분 값을 4*4 정수 변환(integer DCT), 양자화(quantization: Q), 엔트로피 부호화(entropy coding)의 일련의 부호화 과정을 거쳐 전송하게 된다. 반면에 인터모드의 경우는 인접하는 영상과의 시간적 상관 관계를 이용하기 위하여 각 매크로블록의 움직임 벡터를 입력된 현재 영상과 이미 코딩 완료되어 전송된 후 저장된 참조 프레임 영상(reference frame image)으로부터 산출하여 움직임 보상이 수행된 예측 영상을 생성하게 된다. 움직임 추정시 다양한 블록 크기(16*16, 16*8, 8*16, 8*8, 8*4, 4*8, 4*4)를 이용하여 1/4 픽셀의 정밀도를 기본으로 최대 1/8 픽셀의 정밀도로 ME 및 MC를 수행할 수 있다. 그리고 코딩효율을 극대화하기 위하여 시간적으로 인접한 여러 개의 영상을 참조할 수 있다(multiple reference frames)[6-7].

4. 제안된 H.264 기반 FGS 코딩방법

본 논문에서는 2장에 소개된 MPEG-4기반 FGS 코딩 방법의 기본계층 코딩방법을 3장에서 소개된 H.264 알고리즘으로 대체한 H.264 기반 FGS 코딩 방법을 우선적으로 구현하여 비교 대상으로 하였다.

본 논문에서 제안하는 H.264 기반 선택적인 FGS(selective FGS: SFGS) 코딩 방법은 앞서 구성된 H.264기반 FGS코딩 방법에서, 강화계층의 영상정보에 대한 코딩 효율을 높이기 위하여 시간적 예측(temporal prediction) 알고리즘을 도입하였으며, 이에 저역 비트율 대역에서 필연적으로 심각하게 발생하는 드리프트(drift) 현상을 최소화하는 방향으로 제안된 방법이다.

4.1 시간적 예측을 수행하는 FGS 코딩 방법

강화계층에서 시간적 예측을 수행하는 FGS 코딩 방법은 강화계층내 각각의 영상 프레임 정보의 시간적 중복성(temporal redundancy)을 줄임으로써 코딩 효율을 극대화하기 위해 구성된 방법이며 본 논문에서는 임의로 H.264 기반 예측 FGS(predictive FGS: Pr-FGS) 코딩 방법이라고 명명한다.

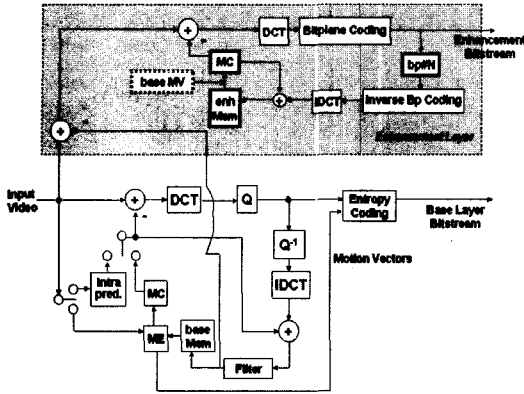


그림 4 H.264 기반 Pr-FGS 방식 엔코더 구성도

그림 4는 H.264 기반 Pr-FGS 코딩방법의 엔코더 구성도를 보여주고 있다. 기본계층 방법은 3장에서 설명된 H.264로 구성되어 있고, 강화계층 방법도 2장에서 설명된 FGS 코딩 방법에 시간적 예측 알고리즘이 추가된 구조로 이뤄져 있다. 강화계층에 추가된 시간적 예측 알고리즘은 다음과 같은 단계로 코딩을 수행한다.

Step 1: 첫번째 영상의 제한된 강화영상 정보 구축 단계

- 1-1: 처음 원 영상(original image)이 입력되어 해당 영상에 대한 기본 계층의 엔코딩이 완료되어 재구성한 영상과 입력된 원 영상간의 오차영상(residual image)이 강화 계층에 입력된다.
- 1-2: 입력된 오차영상 정보에 대하여 'DCT' 단계와 'Bitplane Coding' 단계를 수행하여 강화정보를 생성한다.
- 1-3: 'bp#N' 단계에서 엔코딩이 완료된 비트플레인 정보 중 중요한 비트순으로 최상위 N개의 비트플레인 정보를 추출한다.
- 1-4: 최상위 N개의 비트플레인에 포함된 영상정보를 'Inverse Bp Coding' 단계와 'IDCT' 단계를 수행하여 역 코딩(복호화)하여 해당 결과를, N개의 비트플레인 영상 정보만으로 구성된 강화계층의 부분 참조 영상 정보를 수록한 'enhMem' 버퍼에 저장한다.

Step 2: 후속 영상의 강화계층내의 영상정보 처리 단계

- 2-1: 후속 원 영상이 입력되어 해당 영상에 대한 기본 계층의 엔코딩이 완료되어 재구성한 영상과 후속 원 영상간의 오차영상이 강화계층에 입력된다.
- 2-2: 기본계층에서 입력된 원영상과 'baseMem' 버퍼에 저장된 이전 영상에 대한 재생영상간의 움직임 예측을 수행하여 도출된 움직임 벡터(그림에서 'base MV')를 이용하여 'enhMem'버퍼에 저장된

N개의 비트플레인 영상정보만으로 구성된 이전 프레임의 강화정보에 대한 움직임보상('MC' 단계)을 수행하여 현재 입력 영상 프레임에 대한 시간적 예측 정보를 구성한다.

- 2-3: 강화계층에 입력된 2-1과정의 오차영상과 2-2과정에서 구성된 시간적 예측 정보와의 차를 구하여 해당 정보에 대한 'DCT' 단계와 'Bitplane Coding' 단계를 수행하여 생성된 강화계층의 비트스트림을 디코더단으로 전송한다.

- 2-4: 비트스트림 전송 후 'bp#N' 단계에서 최상위 비트플레인 N개를 추출하고, 이 N개의 비트플레인을 'Inverse Bp Coding' 단계와 'IDCT' 단계를 수행하여 역코딩하여 도출된 결과에 2-2의 수행 결과로서 도출된 현재 프레임의 움직임 보상(MC)이 완료된 현재 영상 프레임에 대한 시간적 예측 정보('MC' 단계 완료 후)가 더해져서 반복적으로 'enhMem' 버퍼에 저장되며, 해당 영상 정보는 후속 영상이 엔코더에 입력되면 이전 영상의 부분적인 강화 정보로서 이용된다.

Step 3: Step 2를 입력영상 시퀀스가 종료할 때까지 반복한다.

본 논문에서 기존의 FGS코딩 방법에 추가하여 구성된 시간적 예측 알고리즘은 강화계층에 입력된 영상정보간의 시간적 중복성을 줄이는 것으로써 강화계층의 비트스트림 발생 양을 현저하게 줄이는데 목적을 두고 있다. 하지만 충분히 예상된 바와 같이 드리프트 현상이 심각하게 발생한다는 치명적인 문제점을 안고 있다. MPEG-4 FGS에서도 드리프트 현상이 발생하는 것을 막기위하여 강화계층내의 시간적 중복성을 제거하는 방법론을 강하게 배제했었지만 본 논문에서는 강화 계층내의 코딩효율을 향상시키기 위하여 다시 도입한 것이다.

드리프트 현상은 그림 4에 도시된 강화계층의 'bp#N' 단계에서 N개의 비트플레인 수만큼의 영상 정보가 추출된 후, 강화계층용 참조 영상 정보로서 'enhMem' 버퍼에 저장된 영상정보 모두가 전송선로상에서 무사히 전송되어 강화계층용 디코더에서 전부 디코딩이 가능한가 여부에 따라 발생된다. 만약 엔코더측에서 전송된 해당 정보(N 개의 비트플레인 정보) 모두가 디코더에서 복호화가 가능하다면 엔코더측과 디코더측의 'enhMem' 버퍼에 저장된 참조 영상간 불일치에 따른 드리프트 현상이 전혀 발생되지 않는다. 하지만 FGS 코딩 방법은 전송선로의 대역폭이 급변하여 영상 복원에 필요한 모든 비트스트림을 수신하지 못한 경우에도, 그때까지 수신된 부분적인 영상 비트스트림만을 이용하여 최적의 화질향상을 얻을 수 있도록 설계되어있기 때문에 시간적 예측 방법을 적용하게 되면 드리프트 현상이 발생하는 것을

피할 수 없으며 특히 저역 비트를 대역에서는 치명적인 화질 저하가 발생된다. 실험적으로 엔코더 측에서 비트 플레인 3개(N=3) 정도를 강화계층의 참조 영상 정보로 사용해야 전송 영상의 화질 개선이 가능한데, 디코더측에서 해당 비트플레인 3개에 대한 정보를 모두 수신할 수 있는 전송 환경이 구현 되어야만 엔코더측의 강화계층과 동일한 영상정보를 보유한 디코더측의 'enhMem' 버퍼를 구현할 수 있다.

그림 5는 Foreman QCIF 영상을 본 논문에서 구현한 H.264 기반 FGS 코딩 방법과 H.264 기반 Pr-FGS 코딩 방법(N=3)으로 디코딩을 수행한 결과를 비교한 그림이다. 그림 5의 가로축은 디코딩이 완료된 비트율(bitrate: bits per second)이며, 그림의 세로축은 해당 비트율에서 복원된 영상의 Y 콤포넌트에 대한 평균 PSNR(peak signal to noise ratio)이다(단위 dB). 기본 계층은 56 kbps로 부호화와 복호화를 수행하였으며, 화질향상을 구현하기 위한 강화 계층에서 전송된 비트스트림 중 수신되어 복호가 완료된 수신 비트율을 32 kbps에서 1024 kbps까지 단계적으로 증가시키며 실험하였다.

그림 5에 표기된 A 기점을 기준으로 왼쪽의 저역 비트율 대역에서는 Pr-FGS가 FGS에 비해 약 4 dB 까지 나 낮은 PSNR을 보여주고 있으며, A기점을 중심으로 오른쪽의 중/고역 비트율 대역에서는 Pr-FGS가 FGS에 대해 약 1.5 dB까지 높은 PSNR을 보여주고 있다. Pr-FGS가 A 기점의 왼쪽 구역에서 FGS에 비하여 낮은 PSNR을 갖는 이유는 Pr-FGS의 디코더가 엔코더측의 'bP#N' 단계에서 사용된 3개의 비트플레인 정보를 수신받지 못하여 참조영상 불일치에 따른 드리프트 현상이 심하게 나타는 것으로 추정할 수 있다. 반면에, A 기점을 중심으로 오른쪽의 중/고역 비트율 대역에서는 오히려 높은 PSNR을 보여주는 이유는, 비트플레인 3개 만큼의 데이터를 디코더에서 수신하여 시간적 예측정보의 사용을 효과적으로 수행하여 코딩 효율이 그 만큼

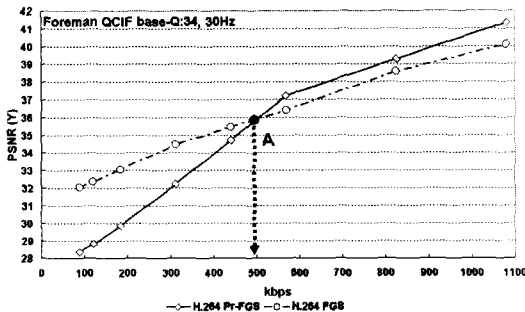


그림 5 3개의 비트플레인 정보로 강화계층 참조영상 버퍼를 구성한 H.264 기반 Pr-FGS 실험 결과의 예

높아졌다고 판단할 수 있다.

결과적으로 강화계층에서 시간적 예측 방법을 사용하는 것이 기존의FGS 코딩방법의 코딩 성능을 향상시킬 수 있는 방법이지만, 일정량의 전송 비트율 이상이 되어야만 효과를 볼 수 있다는 것을 알 수 있다.

4.2 H.264 기반 선택적인 FGS 코딩방법

본 논문에서는 강화계층에서 시간적 예측 알고리즘을 사용함으로써 중/고역 비트율 대역에서의 코딩성능 향상을 도모하고, 저역 비트율 대역에서 드리프트 현상이 발생하는 것을 최소화하여 코딩 성능 저하를 막는 방안을 제안하였으며 이를 H.264 기반 선택적인 FGS(selective FGS: SFGS) 코딩 방법이라 명명한다.

그림 6은 본 논문에서 제안한 H.264 기반 SFGS 코딩 방법의 엔코더 구성도를 도시하고 있다. 제안된 방식은 강화계층 내의 영상 정보로부터 시간적 예측 정보를 구하고, 해당 예측정보를 필요에 따라 선택적으로 사용하여 엔코딩을 수행하고 전송하는 것을 특징으로 한다.

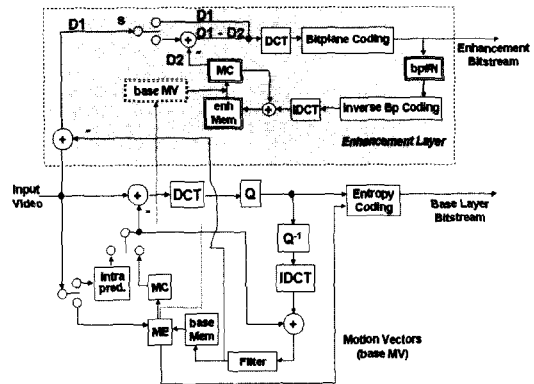


그림 6 H.264 기반 SFGS 코딩 방법의 엔코더 구성도

그림 6에서 'D1' 정보는 원 입력 영상과 기본계층에서 엔코딩을 수행한 후의 재생영상('baseMem' 버퍼에 저장된 영상정보와 동일)과의 매크로블록 단위의 오차영상이다. 'D2' 정보는 강화계층내에서 이전 프레임의 코딩 결과 중 N개의 비트플레인에 해당하는 영상정보('enhMem' 버퍼에 저장된 영상정보와 동일)를 이용하여 기본계층의 움직임 벡터값('baseMV')을 적용하여 움직임 보상(강화계층내의 'MC' 단계)을 수행한 매크로블록 단위의 결과 영상이다.

제안된 방법은 강화계층내에서 시간적으로 예측한 'D2' 정보를 드리프트 현상의 발생을 최소화하는 범위내에서 선택적으로 이용하는 방법론으로서, 강화계층내의 시간적 예측 정보인 'D2' 정보를 엔코딩에 이용하지 않으면 'D1' 정보가 엔코딩에 사용되며, 'D2' 정보를 사용

하면 'D1' 정보 대신 'D1-D2' 정보가 강화계층의 엔코딩의 입력 정보로 사용된다.

강화계층내 시간적 예측 정보를 엔코딩에 적용 가능한지 여부를 판단하는 방법은 그림 7에 도시한 바와 같이 매크로블록 단위로 해당 예측 정보를 사용했을 때의 변화를 평가하여 긍정적인 변화가 크게 발생된 경우에만 시간적 예측 정보를 사용하도록 한다.

매크로블록 단위의 시간적 예측 정보의 사용여부를 결정하기 위해서, 'D1'과 'D1-D2'의 SAD(sum of absolute difference)를 구한다. 매크로블록단위의 'D1'과 'D1-D2'는 각각 영상간의 차분 값(difference)으로 구성되어 있으며, 각각의 차분값에 절대값을 취하고 모두 더하여 SAD를 구한다. 이렇게 구해진 SAD는 각각 SAD_{D1}: SAD(D1)과 SAD_{DD}: SAD(D1-D2)로 명한다. 예측 정보를 강화계층의 엔코딩에서 사용하는 경우는 그림 7(a)에 도시된 바와 같이, SAD(D1-D2)의 값

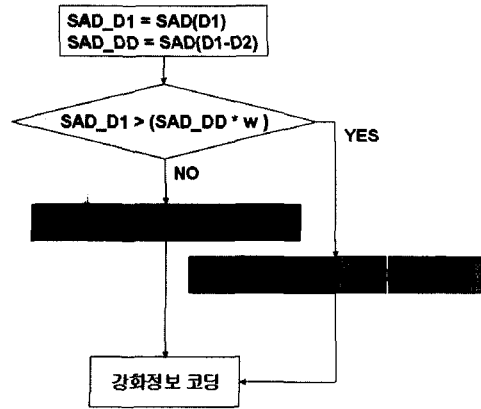


그림 8 예측정보 사용여부 판단 방법

이 SAD(D1)의 값보다 큰 쪽으로 작을 때이다. 이 경우는 강화계층내에서 이전 영상정보를 이용하여 현 영상 정보에 대한 정확한 시간적 예측이 이루어진 것을 의미하며, 가중치 w 값을 믿음요소(belief factor)로서 이용한다. 가중치 w값은 1보다 큰값을 가지며, 본 논문의 실험에서는 w = 1.8을 사용하였다.

매크로블록 단위의 시간적 예측정보를 사용하지 않는 경우에는 그림 7(b)와 같이 D1정보를 강화계층의 엔코딩에 적용하여 전송 비트스트림을 생성한다.

그림 8에 시간적 예측 정보의 사용 여부를 결정하는 방법을 구체적으로 도시하였다. SAD_{D1}값이 작으면 'D1'을 선택하고 그렇지 않으면 'D1-D2'를 선택하여 강화계층의 엔코딩을 위한 매크로블록 단위의 입력 정보로 결정한다.

강화 계층 비트스트림에는 매크로블록 단위로 예측 정보의 사용여부를 디코더측에 통보해주는 정보로서 선택 플래그 비트('S': selection flag bit)를 매크로블록 단위로 한 비트씩 할당하였다. 그림 9는 S 비트를 매크로블록 단위로 추가한 강화 비트스트림을 도식적으로 보여주고 있다.

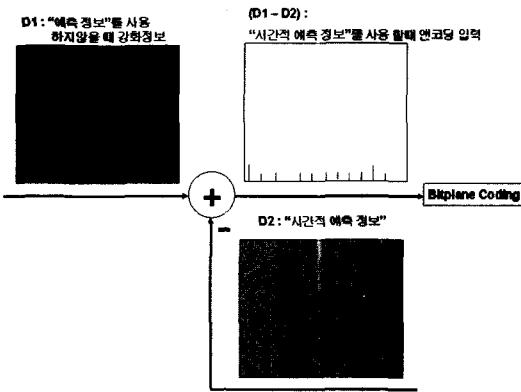
제한된 H.264 기반 SFGS 코딩 방법에서 강화계층내의 부호화방식은 단계별로 다음과 같다.

Step 1: 첫번째 영상의 제한된 강화영상 정보 구축 단계

1-1: 처음 원영상이 입력되면 기본 계층 부호화가 수행된 후, 입력된 원영상과 기본계층에서 엔코딩 완료 후 재생된 영상과의 오차 영상에 대한 정보 'D1'이 강화 계층에 매크로블록 단위로 입력된다.

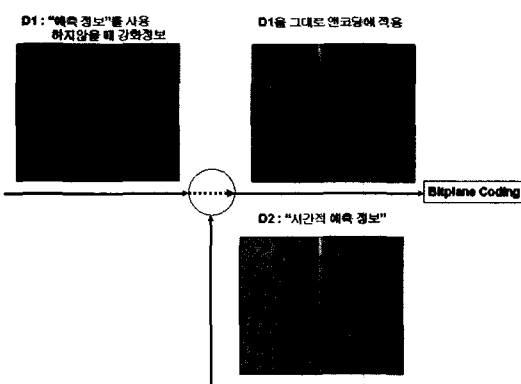
1-2: 해당 'D1' 정보는 'DCT' 단계에서 이상 여현 변환과 'Bitplane Coding' 단계에서 비트플레인 단위의 부호화를 수행하여 비트플레인 단위로 구성된 강화계층 비트스트림을 생성하고 디코더 측에 전송

Case #1: SAD(D1) > SAD(D1-D2) * w



(a) 시간적 예측 정보를 엔코딩에 적용하는 경우

Case #2: SAD(D1) ≤ SAD(D1-D2) * w



(b) 시간적 예측정보를 엔코딩에 적용하지 않는 경우

그림 7 시간적 예측 정보의 적용 상황

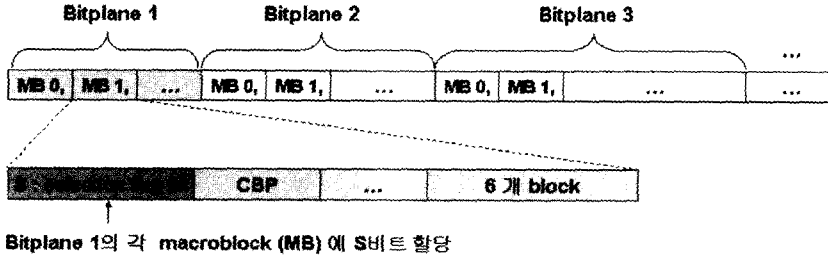


그림 9 플래그 비트 추가한 강화계층 비트스트림 구조

한다.

1-3: 후속 정보 처리를 위하여, 'bp#N' 단계에서 비트플레인 단위의 강화계층 정보를 상위 비트플레인부터 차례로 N개 추출하여 'Inverse Bp Coding' 단계에서 역 비트플레인 코딩과 'IDCT' 단계에서 역 이산 여현 변환을 수행한다. 비트플레인 N개를 역 코딩하여 구성된 강화계층의 부분적인 참조 영상 정보를 'enhMem' 버퍼에 저장한다.

Step 2: 후속 영상의 강화계층내의 영상정보 처리 단계

2-1: 후속 원 영상이 입력되어 기본 계층 부호화가 수행되고, 입력된 원영상과 기본계층에서 엔코딩 완료 후 재생된 영상과의 오차 영상에 대한 정보 'D1'이 강화 계층에 매크로블록 단위로 입력된다.

2-2: 기본계층에서 입력된 원영상과 'baseMem' 버퍼에 저장된 이전 영상에 대한 재생영상간의 움직임 예측을 수행하여 도출된 움직임 벡터(그림에서 'base MV')를 이용하여 'enhMem' 버퍼에 저장된 N개의 비트플레인 영상정보만으로 구성된 이전 프레임의 강화정보에 대한 움직임보상('MC' 단계)을 수행하여 'D1'을 예측한 매크로블록 단위의 예측정보 'D2'를 구성한다.

Step 3: 시간적 예측 정보의 사용여부 판별 단계

3-1: 매크로블록단위의 시간적 예측 정보인 'D2'를 엔코딩에 적용할 것인지 판단한다.

3-2: 'D2' 정보를 적용하기로 결정되었다면, 매크로블록 단위의 'S' 비트에 '1'의 값을 부여하고 'D1-D2' 정보를 엔코더 입력 정보로 결정한다.

3-3: 'D2' 정보를 적용하지 않기로 결정되었다면, 'S' 비트에 '0'의 값을 부여하고 'D1' 정보를 엔코더 입력 정보로 결정한다.

Step 4: Step 3에서 결정된 매크로블록 단위의 입력 정보를 'DCT' 단계에서 이산 여현 변환과 'Bitplane Coding' 단계에서 비트플레인 부호화를 수행하여 비트플레인 단위로 구성된 강화계층 비트스트림을 생성하며 해당 비트스트림을 디코더 측에 전송한다.

Step 5: 후속된 영상정보 처리를 위하여, 'bp#N' 단계에서 비트플레인 단위의 강화정보를 상위 비트플레인부터 차례로 N개 추출하여 'Inverse Bp Coding' 단계에서 역 비트플레인 코딩과 'IDCT' 단계에서 역 이산 여현 변환을 차례대로 수행한다. Step 3에서 결정된 입력 정보가 'D1' 정보라면, 비트플레인 N개를 역코딩하여 얻은 현재 강화정보를 그대로 'enhMem' 버퍼에 저장한다. 그렇지 않고 'D1-D2' 정보가 결정되었다면, 비트플레인 N개를 역코딩하여 얻은 데이터에 'D2' 정보를 더한 후 'enhMem' 버퍼에 저장한다. 버퍼에 저장된 영상 정보는 후속 영상이 엔코더에 입력되면 이전 강화 정보로서 이용 된다.

Step 6: Step 2~Step 5를 입력 영상 시퀀스가 종료할 때까지 반복한다

그림 10은 본 논문에서 제안된 H.264 기반 SFGS 코딩 방법의 디코더단의 구성도이다. 그림에 도시된 바와 같이, 기본계층과 강화계층으로 나뉘어져서 전송선로를 통하여 수신된 비트스트림의 디코딩을 그림 6에 도시된 엔코더와 역순으로 수행하게 된다.

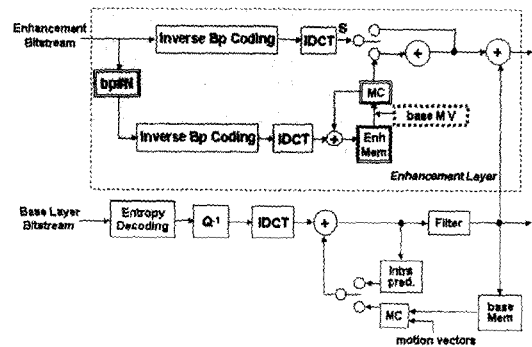


그림 10 H.264 기반 SFGS 코딩방법의 디코더 구성도

지금까지 본 논문에서 제안한 H.264 기반 SFGS에 관하여 설명하였다. 제안된 방식은 H.264를 기본계층을 위한 코딩 방법으로 정하고, 강화계층 코딩 방법으로 기

존의 FGS 코딩 방법에 강화계층내 시간적 예측 알고리즘과 선택적인 예측 정보 판별 알고리즘이 추가된 코딩 방법이다. 이는 강화 계층내 시간적 예측 알고리즘이 적용되었을 때의 코딩성능 향상을 이룰수 있는 장점을 활용하고 그에 따른 드리프트 현상이 심각하게 발생하는 단점을 극복하기 위하여 확실하고 정확한 시간적 예측 정보만을 엔코딩에 사용하는 선택적인 예측 정보 판별 알고리즘으로 극복함으로써 전반적인 코딩 성능의 향상을 가져올 수 있었다.

5. 실험 결과

본 논문에서 제안된 H.264 기반 SFGS 코딩방법을 구현하고, 제안된 방법의 성능 비교를 위하여 기존의 국제 표준안인 MPEG-4 기반 FGS 코딩 방법과, 본 논문에서 보다 발전적으로 구성한 H.264 기반 FGS 코딩 방법, 그리고 과도기적인 제안 방법인 H.264 기반 Pr-FGS 코딩 방법 등을 구현하여 각각의 코딩 성능을 심도있게 비교하였다.

실험에서 비교된 각각의 알고리즘은 MPEG-4 FGS의 경우 Momusys ver. 0.9 S/W를 기반으로 구현되었으며, H.264 기반 FGS, Pr-FGS, 그리고 SFGS는 JM 6.1e S/W를 기반으로 구현되었다. 또한 실험조건으로는 표 1에 명시하였다.

그림 11에는, 각각 구현된 코딩 방법의 성능 비교를 Akiyo QCIF 30Hz 영상 시퀀스(sequence)에 적용하여 비교한 결과를 보여주고 있다. 그림에서, 현존하는 국제 표준안인 MPEG-4 기반 FGS코딩 방법의 성능이 가장 낮고, 본 논문에서 성능 비교를 위하여 구현된 H.264 기반 FGS 코딩 방법의 성능이 MPEG-4 기반 FGS 코딩 방법의 성능에 비하여 약 1~3 dB 높고, 과도기적으로 구현된 H.264기반 Pr-FGS의 코딩 성능이 저역 비트율 대역에서는 H.264 기반 FGS 코딩 방법의 성능에 비하여 약 1~2 dB 낮지만, 256kbps 이상의 중/고역 비트율 대역에서는 무려 3~6 dB의 성능 향상을 가져오고 있음이 관찰되었다. 저역 비트율 대역의 코딩성능 향

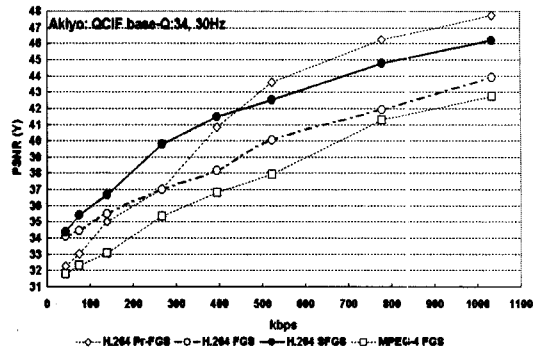


그림 11 Akiyo QCIF 30 Hz 영상 시퀀스에서 MPEG-4 기반 FGS, H.264 기반 FGS, H.264 기반 Pr-FGS, H.264 기반 SFGS의 코딩 성능

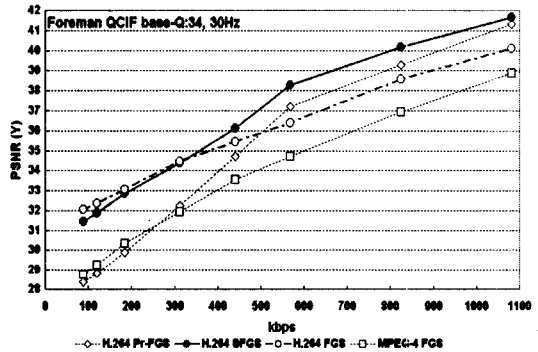


그림 12 Foreman QCIF 30 Hz 영상 시퀀스에서 MPEG-4 기반 FGS, H.264 기반 FGS, H.264 기반 Pr-FGS, H.264 기반 SFGS의 코딩 성능

상을 위하여 본논문에서 도입된 H.264 기반 SFGS 코딩 방법의 성능이 전체 비트율 대역에서 H.264 기반 FGS 방법보다 약 1~3 dB 향상되었고, MPEG-4 기반 FGS기법과 비교해서는 약 3~5 dB의 코딩 효율 향상을 얻어낼 수 있음을 확인할 수 있었다.

그림 12에는, 각각 구현된 코딩 방법의 성능 비교를

표 1 실험 조건

| Video Codec | Parameter Setting | Meaning |
|--------------------------------|--|---|
| MPEG-4 FGS | - RateControl = 3 - SearchRange = 16 - Value of M = 1 | - Uses TM5 rate-control algorithm for base layer - Motion vector displacement range - Frame Format (IPPPPPPPPPPP...) |
| H.264 based FGS, Pr- SFGS, FGS | - IntraPeriod = 0 - FrameSkip = 0 - NumberReferenceFrames = 1 - NumberBFrames = 0 - SymbolMode = 1 - RDOptimization = 1 | - Uses only one I frame for best coding efficiency - No frame Skip - Total number of reference frame range - Frame Format (IPPPPPPPPP...) - Uses CABAC - Uses RDOptimization |

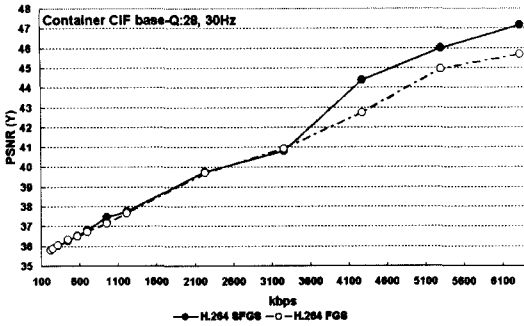


그림 13 Container CIF 30 Hz 영상 시퀀스에서 H.264 기반 FGS와 H.264 기반 SFGS의 코딩 성능

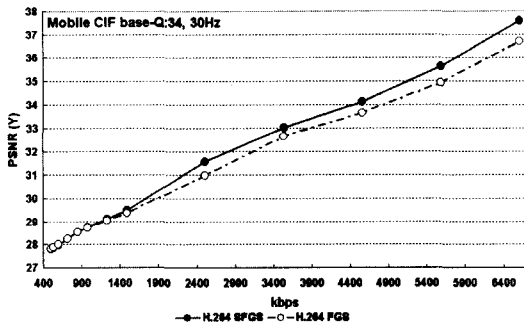


그림 14 Mobile CIF 30 Hz 영상 시퀀스에서 H.264 기반 FGS와 H.264 기반 SFGS의 코딩 성능 비교

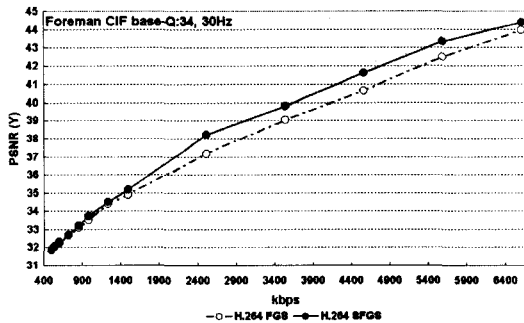


그림 15 Foreman CIF 30 Hz 영상 시퀀스에서 H.264 기반 FGS와 H.264 기반 SFGS의 코딩 성능 비교

Foreman QCIF 30 Hz 영상 시퀀스로 비교한 결과를 보여주고 있다. 그림 11의 결과와 유사하게 그림 12에서

도 MPEG-4기반 FGS코딩 방법의 성능이 가장 낮고, 본 논문에서 성능 비교를 위하여 구현된 H.264 기반 FGS 코딩 방법의 성능이 MPEG-4 기반 FGS코딩 방법의 성능에 비하여 약 2~4 dB 높고, H.264기반 Pr-FGS의 코딩성능이 저역 비트율 대역에서는 드리프트 현상이 발생됨에 따라 H.264기반 FGS코딩 성능에 비하여 무려 4 dB까지 코딩 성능의 하락이 뚜렷하게 관찰되었으며 약 500 kbps이상의 중/고역비트율 대역에서는 1.5 dB까지의 성능 개선 효과도 동시에 관찰되었다. 반면, H.264 기반 SFGS의 성능은 전체 비트율 대역폭에서 H.264 기반 FGS 기법보다 약 2 dB까지 코딩성능이 향상됨이 관찰 되었고, MPEG-4 기반 FGS 코딩 방법과 비교해서는 약 3~5 dB의 코딩 효율 향상을 얻어 낼 수 있었다.

그림 13~15는 각각 CIF급 30Hz의 Container, Mobile, Foreman 영상 시퀀스에서 H.264 기반 FGS와 H.264 기반 SFGS 코딩 방법의 성능비교를 각각의 수신 비트율 대역마다 수행한 결과를 도시하고 있다. 각각의 그림에서, MPEG-4기반 FGS 코딩 성능은 앞서 그림 11~12에 도시한 바와 같이 H.264 기반의 FGS 코딩 방법에 비하여 현저하게 떨어지기 때문에 비교하지 않았다. 그림 13~15에서도 역시 본 논문에서 제안된 H.264 기반 SFGS의 코딩 성능이 전체 비트율 대역에서 H.264 기반 FGS의 성능보다 약 2 dB까지 코딩 효율이 향상됨을 쉽게 관찰할 수 있다.

본 논문에서 제안하는 H.264 기반 SFGS의 디코딩에 소요되는 계산 복잡도(Computational Complexity)를 H.264 기반 FGS와 비교측정을 수행하였다. 측정에 사용된 컴퓨터는 Pentium 4(CPU 2.4 Ghz, 512Mb)를 사용하였으며 Akiyo QCIF 영상 시퀀스의 디코딩 소요시간을 표 2에 제시하였다. 표에서 제시한 바와 같이, 제안된 방법의 디코딩 복잡도가 약 20% 정도 상승하는 것을 알 수 있었으며, 이는 강화계층에서 MC를 수행하는데 소요되는 시간의 영향이 큰 것으로 판단된다. 하지만, 제안된 방법의 코드 최적화를 수행하지는 않았지만, 평균 디코딩 시간이 31.3 msec/Frame으로서 초당 30 Hz의 영상 시퀀스를 디코딩하는 경우에도 충분한 것으로 판단된다.

6. 결론

기존의 스케일러블 비디오 코딩방법은 수신단측에서

표 2 1 프레임당 디코딩 평균 소요시간

(단위: msec)

| | 1회 | 2회 | 3회 | 4회 | 5회 | 6회 | 7회 | 8회 | 9회 | 10회 | 평균 |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| H.264 FGS | 25.15 | 25.05 | 25.05 | 25.05 | 25.05 | 25.05 | 25.05 | 25.05 | 25.05 | 25.1 | 25.07 |
| H.264 SFGS | 31.67 | 31.25 | 31.25 | 31.2 | 31.3 | 31.25 | 31.25 | 31.3 | 31.3 | 31.25 | 31.3 |

향상된 화질을 얻기 위해서는 엔코더측에서 전송된 강화 계층 비트스트림을 모두 수신하여야 하며, 수신 도중에 전송선로의 급격한 변화로 완전한 비트스트림을 수신하지 못하면 그때까지 수신된 정보의 복호화가 불가능하여 이를 화질 향상에 사용할 수 없게 된다. MPEG-4 기반 FGS 코딩 방법은 전송 대역폭이 급변하여 강화 계층에서 전송된 비트스트림을 모두 수신하지 못한 경우에도, 그때까지 수신된 부분적인 비트스트림만을 이용하여 수신영상의 화질향상이 가능하며, 최초의 망 상태 적응적인 스케일러블 비디오 코딩 방법이라 할 수 있다. 이로인하여 한개의 콘텐츠를 다양하게 사용할 수 있게 되어(one-Source, multi-use) 유비쿼터스 환경에 적합한 기술이라 평할 수 있으나 코딩효율면에서는 다소 미흡하다는 평가를 받고 있다.

본 논문에서 제안된 H.264 기반 SFGS 코딩 방법은 강화 계층내의 영상간 시간적 중복성을 제거하는 예측 방법을 사용하여 코딩효율을 증대시키며, 그로 인하여 심각하게 발생하는 드리프트 현상을 최소화 하는 방향으로 시간적 예측이 잘 수행된 영상정보만을 선택적으로 사용하도록 설계되었다. 즉, 시간적 예측이 정밀하여 확실하게 코딩효율을 높일 수 있는 경우에만 매크로블록단위로 예측정보를 사용하였으며, 사용 여부를 결정하는 S비트를 매크로블록당 한 비트씩 추가함으로써 수신단측에서도 동일한 작업을 수행할 수 있도록 하였다. 다양한 영상 시퀀스를 대상으로 실험한 결과, 본 논문에서 제안된 H.264 기반 SFGS 코딩 방법은 현존하는 국제 표준안인 MPEG-4 기반 FGS 코딩 방법보다 동일한 비트율 대역에서 영상화질이 약 3~5 dB 정도 높은 성능을 보여주고 있으며, 특히 H.264를 기반으로한 FGS 코딩 방법보다도 약 1~3 dB 높은 성능을 보여주고 있음을 확인할 수 있었다.

참 고 문 헌

[1] W. Li, J. Ohm, M. V. Schaar, H. Jiang, and S. Li, MPEG-4 Video verification model ver. 18.0, *ISO/IEC/ JTCl/ SC29/ WG11/ N3908*, January, 2001.

[2] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11, pp. 301-317, March, 2001.

[3] MPEG-4 Fine granular scalability verification model version 4.0, *ISO/ IEC/ JTCl/ SC29/ WG11/ N3317*, 2000.

[4] G. H. Park, Y. J. Lee, W. -S. Cheong, K. Kim, J. Kim, Y. K. Lim, "Water ring scan method for H.26L based FGS," *ITU-T SGI6 Q.6 JVT-B094*, Geneva, January 2002.

[5] G.H. Park, W. S. Cheong, K. Kim, Y.J. Lee, Y.K.

Lim, J. Kim, "Water ring scan method for MPEG-4 and H.26L based FGS methodologies," *ISO/ IEC/ JTCl/ SC29 WG11/ MPEG/ M 8023*, Jeju, Korea, March, 2002.

[6] 홍민철 전병우, "H.26L 동영상 부호화 표준방식의 배경 및 동향", 방송공학회, 2003.

[7] T. Wiegand, "Overview of the H.264/ AVC Video Coding Standard"

[8] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in Wireless Environments"

[9] T. Wiegand, G. Sullivan, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), Document JVT-G050.

[10] 김해광, 이상윤, "JVT 동영상 국제표준 프로파일/레벨 동향", 방송공학회, 2003.



박 광 훈

1985년 연세대학교 전자공학과(공학사)
 1987년 연세대학교 전자공학과(공학석사). 1991년 Case Western Reserve Univ., Dept. of EEAP(M.S.). 1995년 Case Western Reserve Univ., Dept. of EEAP(Ph.D.). 1995년~1997년 현대전자 멀티미디어 연구소 책임연구원. 1997년~2001년 연세대학교 문리대학 전산학과 부교수. 2001년~현재 경희대학교 컴퓨터공학과 부교수. 2004년 대한민국 특허기술상 대상(세종대왕상) 수상. MPEG-4(Part2) 및 H.264 국제표준안의 필수 특허 발명자. 국제등록특허 14개, 국내등록특허 37개 등. 관심분야는 멀티미디어, 비디오 신호처리, 패턴인식, 영상처리, 계산지능 등



유 원 혁

2002년 경희대학교 컴퓨터공학과(공학사). 2004년 경희대학교 컴퓨터공학과(공학석사). 관심분야는 동영상 압축기술, 스케일러블 코딩, 영상처리 등



김 규 현

1989년 한양대학교 전자공학과(공학사)
 1992년 영국 University of Newcastle upon Tyne 전기전자공학과(공학석사)
 1996년 영국 University of Newcastle upon Tyne 전기전자공학과(공학박사)
 1996년~1997년 영국 University of Sheffield, Reserach fellow. 1997년~현재 한국 전자통신 연구원 디지털 방송 연구단 방송 미디어 연구 그룹 대화형 미디어 연구팀장. 관심분야는 영상처리, 멀티미디어통신, 대화형 방송 시스템 등