

MPEG-4 FGS 비디오 전송을 위한 향상 계층 데이터 추출 기법

*박형미 **문주희 ***김현철 ****김규현

* **세종대학교 정보통신연구소 *** ****한국전자통신연구원 방송미디어연구그룹

*loogin@sju.ac.kr

Enhancement Data Extraction Algorithm for MPEG-4 FGS Video Transmission

*Park, Hyoung-Mee **Moon, Joo-Hee ***Kim, Hyun-Cheol ****Kim, Kyuheon

* **Information and Telecommunication Research Institute, Sejong University

*** ****Broadcastin Media Research Group, ETRI

요약

MPEG-4 FGS 비트 스트림을 서버측에서는 각 클라이언트의 채널환경에 맞게 적절한 비트율로 추출하여 각 클라이언트에 게 전송해야 한다. 부분적인 비트 스트림만으로도 영상 복원이 가능하지만, 같은 비트율에서 전송화질의 극대화를 위해서는 비트 스트림의 추출방법에 대한 연구가 필요하다. 이에 본 논문은 MPEG4 FGS 비디오 스트림을 네트워크 상에서 동적인 통신용량의 변화에 맞춰 적절한 비트율로 추출하는 방법에 대해 제안한다. 제안하는 방법은 같은 비트율에서 전송화질의 극대화를 위해 향상계층의 각 비트 평면(bit-plane)이 손실됐을 때 품질에 미치는 영향을 고려하여 비트율을 줄인다. 이렇게 함으로써 동일한 비트율이 주어진 경우 얻을 수 있는 종단간(end-to-end) 비디오 품질의 향상을 비교 분석하였다.

1. 서론

이전의 통신방식은 회선교환(Circuit Switching)기반 점대점 연결로써, 통화가 시작되기 전 사용자 간에 전송속도에 대한 사전협의를 마친 후 서비스가 개시되는 구조였으며, 전송대상은 주로 텍스트와 정지영상이었다. 통신부호화 시스템에서는 사전에 협의된 채널용량과 같거나 더 적은 비트율 안으로 사용하여 입력 비디오 신호를 부호화 하였다. 또한 복호기에서는 부호화된 모든 비디오 신호를 수신할 수 있으며, 수신된 모든 비트들을 복호할 수 있다고 가정하였다[1]. 그러나 1990년 중반이후 인터넷은 급속한 속도로 확산되어 현재는 점차 인터넷 망과 같은 패킷교환(Packet switching)에 근간을 둔 통신이 보편화되었고 인터넷이나 인트라넷과 같은 IP기반 망은 그 특성상 서로 이질적인 사용자들에게 서비스 품질(Quality of Service)이 보장되지 않는 채널을 통하여 음악이나 동영상과 같은 다양한 멀티미디어 서비스를 실시간으로 제공해야 한다[2][3]. 인터넷 비디오 스트리밍과 같이 네트워크 채널용량이 어느 주어진 시간에 대해서 넓은 범위에서 변하는 네트워크 환경에서 부호화기는 시간에 따라 변하는 채널용량을 정확히 예측할 수 없으며, 대역폭에 스트림의 비트율을 맞추어 줄 수 없게 되었다.

인터넷 망이 실시간 동영상 서비스의 QoS (Quality of Service)를 보장해 줄 수 없기 때문에 영상 부호화 관점에서 해결하기 위해서 MPEG-4 표준에서는 스트리밍 비디오를 위해 미세 계층적(Fine Granular Scalable: FGS) 부호화 기법[4][5][6] 이라는 새로운 프로파일을 제공하였다. MPEG-4 미세 계층적 부호화 기법은 스트림의 비트율을 망의 대역폭에 간단하고 유연하게 맞추는 것을 목표로 만들어졌다. 이렇게 함으로써 스트리밍 서비스를 이용하는 많은 수신측에 각각의 대역폭을 최대한 이용하여 전송화질을 극대화 할 수 있다. 미세 계층적 부호화 기법은 비트율을 수 비트 단위로 세분하여 부호화함

으로써 가용전송 비트율이 크게 변동되더라도, 재 부호화 없이 사용자가 어떤 비트율에서도 복호 가능한 서비스를 제공할 수 있다. 미세입자 스케일러블(FGS) 부호화 기법에서는 강화계층의 전송 중에 대역폭이 급변하여 강화계층의 비트스트림을 모두 수신하지 못한 경우에도, 그때 가지 수신된 부분적인 비트스트림 만을 이용하여 수신된 영상의 화질 향상이 가능하다.

앞서 설명한 바와 같이 MPEG-4 미세입자 스케일러블 부호화기법에서는 부분적인 비트 스트림만으로도 영상 복원이 가능하지만, 같은 비트율에서 전송화질의 극대화를 위해서는 가용 대역폭에 맞춰 비트 스트림을 추출하는 방법에 대한 연구가 필요하다. 이에 본 논문은 네트워크 전송을 위해서 미세입자 스케일러블 향상계층 데이터를 선택적으로 추출하는 방법을 제안한다. 본 논문의 2장에서는 스케일러블 비디오 통신 시스템에서 추출기와 부호화기/복호기 간의 구조를 설명하고, MPEG-4 FGS 데이터의 기본구조와 부호화/복호화 방식에 대해서 설명을 한다. 3장에서는 MPEG-4 FGS 향상계층 데이터를 선택적으로 추출하는 방식에 대하여 설명을 하고, 4장에서는 제안한 향상계층 데이터를 방식과 MPEG-4 FGS 레퍼런스 소프트웨어에서 사용되는 추출 방식의 성능을 비교 분석하였다.

2. FGS의 기본 구조

멀티미디어 콘텐츠를 계층적 부호화/복호화 하는 통신 시스템은 부호화기, 추출기, 복호화기로 나타낼 수 있으며, <그림 1>은 이들 간의 관계를 나타낸다. 이들 중 추출기(Extractor)는 전송 단에서 부호화기에서 생성한 비트스트림을 사용자의 통신용량에 맞춰 적절한 비트율로 향상계층의 데이터를 추출하는 역할을 한다.

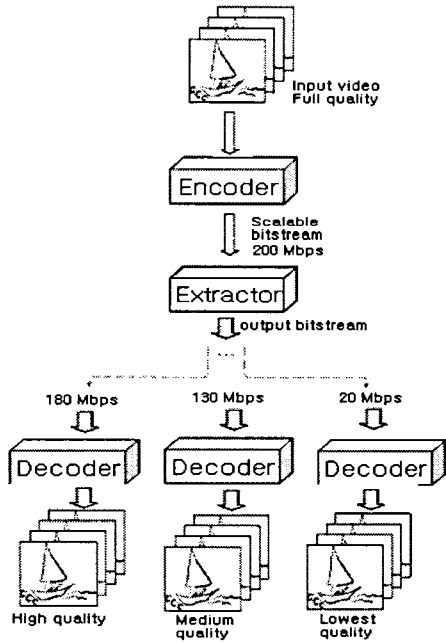


그림 1. 계층적 통신부호화 시스템의 구성도

기본적인 FGS 코딩방법의 구조를 살펴보면 <그림 2>와 같다. 그림에서 보는 바와 같이 기본계층과 함께 강화계층으로서 FGS 계층을 두고 있다. 기본계층은 기존의 MPEG-4 비디오 코딩기법을 그대로 사용하고 있다. 강화계층인 FGS 계층 간에는 계층 간의 움직임 추정 및 보상 등 코딩효율을 증가시키는 방법론을 전혀 사용하지 않고 단지 기본 계층 간으로부터 코딩 효율을 증가시키는 방법만을 고려하고 있는 것이 특징이다.

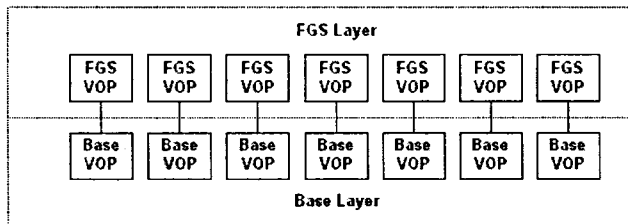


그림 2. 기본적인 FGS 코딩 방법 구조

<그림 3>는 FGS 향상계층을 위한 부호화기 구조를 보이고 있다. FGS 기본계층은 기존의 블록기반 부호화 방법과 동일하게 부호화되며, FGS 향상계층은 입력으로 입력영상과 기본계층에서는 재생된 신호의 차 신호(residual signal)가 입력으로 들어가게 되고 이 신호는 8x8 블록별로 DCT(Discrete Cosine Transform) 된다. 현재의 화면 전체에 대한 DCT(Discrete Cosine Transform) 영역 차 신호가 얻어지면 절대치가 가장 큰 계수를 구하여 비트평면의 계수를 정하게 된다. 그 후 비트평면 부호화 과정을 통해 FGS 향상계층의 비트스트림이 만들어지게 된다. <그림 3>에서 비트평면 쉬프트는 특정부분의 화질을 향상시키기 위해 그 부분에 해당하는 블록들의 비트 평면을 MSB (most significant bit) 쪽으로 상향 천이 해주는 기능을 한다. 상위 MSB 평면의 경우 비트스트림 전송 시 먼저 전송되므로 채널상황이 좋은 낮은 상황에서도 이 부분의 전송은 보장되게 된다. <그림 3>에서 비트평면의 각 블록에 대해서 엔트로피 부호화를 수행하는데, MPEG-4 FGS에서는 엔트로피 부호화 방식으로 가변 길이 부호화

(VLC, Variable Length Coding) 방식을 사용한다. FGS에서는 상위 네 개의 비트평면에 대해 각각 하나씩의 VLC (Variable Length Coding) 테이블을 이용한다. 네 번째 이하의 비트평면에 대해서는 네 번째 비트평면의 VLC (Variable Length Coding) 테이블을 그대로 적용하게 된다. 클리핑 (Clipping) 블록은 재생된 픽셀 값의 범위가 0~255범위를 넘어가게 될 경우 값을 0~255범위 안으로 만들어준다.

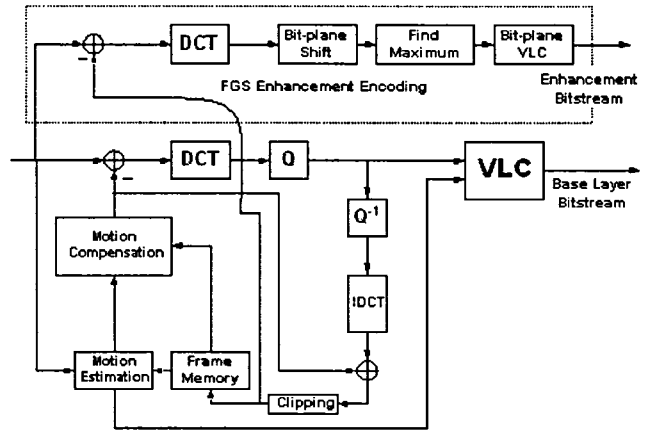


그림 3. MPEG-4 FGS 부호화기 구조

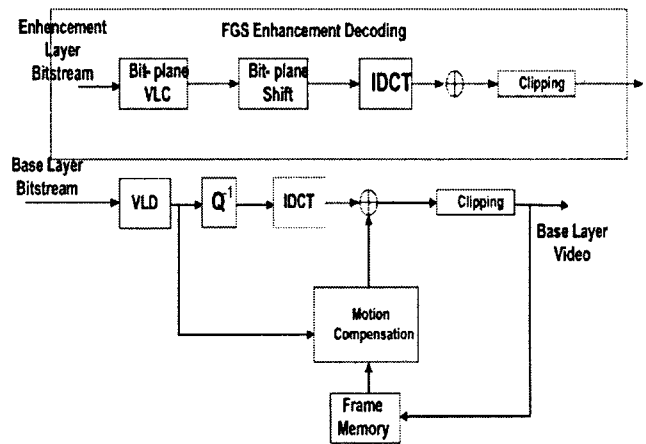


그림 4. MPEG-4 FGS 복호화기 구조

FGS 코딩 기법의 복호화기 구조는 <그림 4>에 도시된 것과 같이, 기본계층에서 MPEG-4 비디오 디코딩 기법을 그대로 사용하고 있다. FGS 향상계층에서는 입력된 향상계층 비트스트림을 비트플레인 별로 가변길이 디코딩 (VLD, Variable Length Decoding)을 수행한 후, 선택적으로 좋은 화질을 갖는 블록의 위치 등이 전송되었다면, 비트 플레인 쉬프트를 선택적으로 수행한다. 비트 플레인 별로 VLD(Variable Length Decoding)가 수행되고 선택적으로 쉬프트가 수행되어 도출된 값에 블록 단위의 IDCT (Inverse Discrete Cosine Transform)를 수행하여 강화계층에서 전송된 영상을 복원하고 이를 기본 계층에서 복호된 영상과 합친 후에 해당 값들을 0에서 255값 사이로 클리핑 (Clipping) 하여 최종적으로 향상된 화질의 영상을 복원하게 된다.

3. 향상계층 데이터 추출 방법

제안하는 추출방법은 m ($1 \leq m \leq M$)인 임의의 정수, M : 전체 프

레이의 개수) 개의 VOP (Video Object Plane) 그룹 내에서 부화된 각 비트평면이 손실되었을 때 발생하는 왜곡을 측정하여 비트율을 줄이는 방법이다. 왜곡을 최소화하도록 하기 위해서는 각 비트평면 별 손실 왜곡을 계산한 후 추출 시 우선순위를 결정하고, 우선순위가 높은 순서대로 채널 상황에 맞춰 향상계층 데이터를 추출한다. 자세한 설명은 다음과 같다.

1 단계. 그룹 당 추출해야 할 목표 비트 양을 구한다.

여기서의 그룹이란 m개의 VOP (Video Object Palen)를 뜻하며, m은 $1 \leq m \leq M$ (M: 전체 VOP의 개수)을 나타낸다. 그룹 당 목표 비트 양과 추출해야 할 목표 비트 양은 <식 1>을 이용하여 계산한다.

$$\{G\}_T = \frac{b_r \cdot N \cdot \{G_{remain}\}}{fr} - \{G_{current}\} - \{G\}_{EXT} \quad \text{<식 1>}$$

<식 1>의 각 변수에 대한 설명은 다음과 같다.

- $\{G_{current}\}$: 현재 그룹의 비트 양
- b_r : 목표 비트 율 (target bit-rate)
- fr : 프레임 율 (frame-rate)
- N : 그룹 당 프레임 개수
- $\{G_{remain}\}$: 이전 그룹에서 남은 비트 양
- $\{G\}_T$: 현재 그룹의 목표 비트 양
- $\{G_{EXT}\}$: 현재 그룹에서의 추출해야 할 목표 비트 양

2 단계. 그룹내의 VOP별 비트양과 각 비트 평면의 비트양을 구한다.

각 VOP는 FGS_VOP_START_CODE를 사용하여 구분할 수 있으며, 각 비트평면은 시작은 FGS_BP_START_CODE로 구분할 수 있다. 입력된 향상계층 데이터 비트 스트림 읽어 들이면서 각 START_CODE를 이용하여 VOP별 비트양과 비트평면 별 비트 양, 각 VOP별 비트평면의 개수를 센다. 이때 각 VOP와 비트 평면들의 스트림 내에서의 시작위치를 저장해놓는다. 2단계에서 구한 각각의 변수는 이하 다음과 같이 표기한다.

- $b[VOP_j]$: 그룹 내에서 j번째 VOP의 비트양
- $b[BP_{jk}]$: 그룹 내에서 j번째 VOP의 비트평면의 비트양
- PT_{VOP_j} : 그룹 내에서 j번째 VOP의 스트림 내에서의 시작위치
- $PT_{BP_{jk}}$: 그룹 내에서 j번째 VOP의 k번째 비트평면의 스트림 내에서의 시작위치

3 단계. 손실 시 왜곡을 구해야 하는 비트평면 레벨을 구한다.

손실 시 왜곡을 구해야 하는 비트평면 레벨을 n이라 하면, 그룹 내 모든 VOP의 비트평면들은 데이터는 최하위 레벨부터 n+1레벨까지 모두 추출되고, n 레벨의 비트 평면들은 각 비트평면 손실 시 왜곡을 구하여 선택적으로 추출한다.

최하위 비트평면 (LSB, Least Significant Bit) 부터 그룹 내의 모든 k 번째 비트평면을 추출했을 때 추출된 비트양이 그룹 내의 목표추출 양보다 많아지는 비트평면 레벨 n을 구해낸다.

$$\{G_{EXT_k}\} = \{G_{EXT_{(k-1)}}\} - \sum_j^N \{BP_{jk}\} \quad \text{<식 2>}$$

<식 2>는 최하위 평면부터 k번째 평면까지 추출했을 때 추출된 비트양을 나타낸다. 즉, $\{G_{EXT_k}\}$ 이 0보다 작아지는 k가 n이 되며, 이는 n번째 비트평면 추출 시 그룹 내의 목표추출양보다 많이 추출되었음을 나타내며, 그룹 내의 n번째 비트평면 중 목표 비트 량에 맞춰 추출 시 손실이 적은 비트평면을 골라서 추출해야 함을 나타낸다.

4 단계. 그룹 내 각 VOP의 n번째 레벨의 비트평면들에 대해서, 손실 시 왜곡을 구한다.

추출하는 방법은 서버의 환경에 따라 정확한 손실 양을 측정하는 추출하는 방식과 단계2에서 구한 비트평면 별 비트 양을 이용하여 추출하는 방식 두 가지가 있다. 첫 번째 방식은 SNR측면에서는 이득이 되지만, 정확한 소실량을 계산하는데 시간이 소요된다. 사용자가 많은 경우 이는 전체적인 시스템에 부담이 될 수 있다. 이러한 상황에서는 계산량과 시간 측면에서 부담이 적은 추출방식이 필요하다. 이를 위해서 두 번째 방식은 비트평면 별 비트양 정보만을 이용하여 추출한다. 방법4-①은 정확한 손실 시 왜곡량을 측정하여 데이터를 추출하는 방식이고 방법 4-②는 비트평면별 비트양 정보만을 이용하여 추출하는 방식이다.

4- ①. 각 비트 평면의 NZBC와 비트양 정보를 사용하여 손실 시 왜곡을 구한다.

그룹 내 모든 n번째 비트 평면에 대해서 손실 시 왜곡을 구하기 위해서 각 비트평면에 대해서만 IDCT를 수행하여 비트평면의 이진 계수를 구한다. 이때 스트림 내에서 각 VOP와 비트평면은 (2단계)에서 구해놓은 VOP와 비트평면의 시작위치를 사용하여 찾는다. 구한 비트 평면의 이진계수를 <식 3-1>를 사용하여 손실 시 왜곡을 구한다.

$$J(\lambda) = D_j + \lambda \cdot b[BP_{jm}] \quad \text{<식 3-1>}$$

$$D_j = (NZBC_n \times 2^n) \quad (0 < j < N)$$

<식 3-1>의 각 변수에 대한 설명은 다음과 같다.

- λ : Lagrange Multiplier
- $b[BP_{jm}]$: 그룹내의 j번째 VOP의 n번째 비트 평면의 비트 양
- N : 그룹 내의 VOP개수
- $NZBC_n$: n번째 비트 평면의 0이 아닌 이진 계수의 개수

4-②. 각 비트양 정보를 사용하여 손실 시 왜곡을 구한다.

비트 양과 손실 시 왜곡간의 상관성을 이용하여 추출하는 방식으로 비트평면의 비트양이 많다는 것은 <식 3-1 >의 NZBC (Non-Zero Binary Coefficient)가 크다는 것을 의미한다. 그러므로 그룹 내의 n번째 모든 비트평면의 비트양을 손실 시 왜곡양으로 대체할 수 있으며, 이때의 비트평면 손실 시 왜곡은 <식 3-2>과 같다.

$$J(\lambda) = \lambda \cdot b[BP_{jm}], \quad (0 < j < N) \quad \text{<식 3-2>}$$

<식 3-1>의 각 변수에 대한 설명은 다음과 같다.

- λ : Lagrange Multiplier
- $b[BP_{jm}]$: 그룹 내의 j번째 VOP의 n번째 비트평면의 비트 양

6 단계) 목표 비트양이 되도록 4단계에서 구한 손실 시 왜곡이 작은 평면부터 데이터를 추출한다.

그룹 내의 비트평면 레벨이 n보다 큰 비트평면을 추출하고 n번째 비트평면에 대해서는 손실 시 왜곡의 합이 가장 작도록 목표 추출 비트량에 맞춰 추출한다.

4. 실험 결과 및 분석

제안된 향상계층 데이터 추출기법의 성능을 확인하기 위하여 MPEG-4 FGS 레퍼런스 소프트웨어를 이용하여 제안된 추출 방식과 비교하였다. MPEG-4 FGS 레퍼런스 소프트웨어에서 사용하는 추출 방식은 VOP당 추출해야 하는 평균 목표 추출 비트 양을 구한 후, 각 VOP마다 같은 양의 비트를 추출하는 방식이다. 추출된 비트 스트림을 MPEG-4 FGS 레퍼런스 소프트웨어를 사용하여 복호화 하였고, 복원된 영상 신호의 PSNR을 측정하였다. 테스트 시퀀스로는 MPEG4 표준에서 이용되는 Foreman과 Hall 동영상 시퀀스를 사용하였으며, 30fps에서 비트율을 달리 하였을 때 복원된 영상의 PSNR을 측정하였다. <그림 5, 6>에서 MS_extractor는 MPEG-4 FGS 레퍼런스 소프트웨어에서 사용되는 향상계층 데이터 추출방식이고, BIT_RATE는 제안한 방식 중에서 향상계층의 각 비트평면의 비트양 정보만을 사용하여 추출하는 방식이다. 손실 시 왜곡정보 이용은 비트평면의 NZBC를 사용하여 추출하는 방식이다.

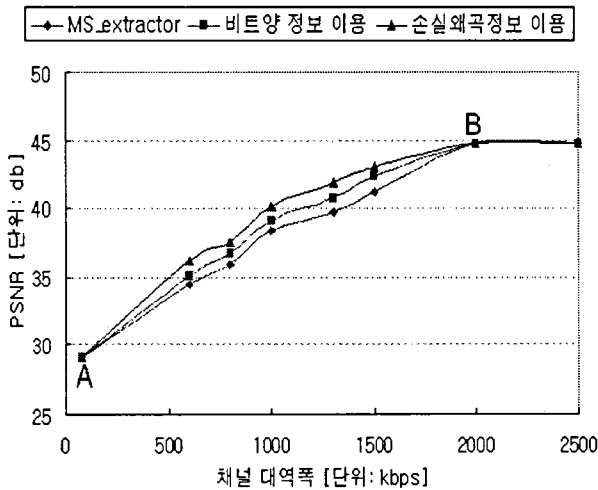


그림 5. Foreman, 300장, QCIF, 30fps

<그림 5>에서 점 A는 기본계층 데이터만 복호화 했을 때의 위치점이다. 즉, 대역폭이 80kbps일 때는 기본계층 비디오 데이터만을 수신하여 복호화 한다. 80kbps이상의 대역폭에서는 향상계층 비디오 데이터를 채널상태에 맞게 수신하여 기본계층 비디오 데이터 보다 더 나은 화질의 영상을 얻을 수 있다. 점 B는 기본계층과 향상계층의 모든 비디오 데이터를 수신했을 경우를 나타낸다.

<그림 6>에서 대역폭이 5kbps일 때는 기본계층 비디오 데이터만을 수신하여 복호화 (점 A)한다. 향상계층 비디오 데이터는 1975 kbps이며, 기본계층과 향상계층 데이터를 모두 수신하기 위해서는 1800 kbps이상의 대역 폭이 되어야 한다.

실험 결과로부터 제안한 방법은 같은 비트율에서 기존의 MPEG-4 FGS 레퍼런스에서 제공하는 추출방식보다 1dB에서 2dB 정도의 성능

향상을 얻을 수 있음을 확인 할 수 있었다.

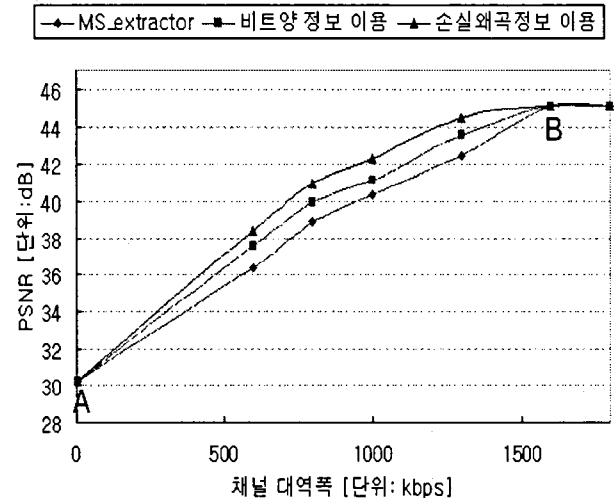


그림 6. Hall, 300장, QCIF, 30fps

5. 결론

인터넷 스트리밍 환경 뿐 아니라 앞으로 유비쿼터스(Ubiquitous) 및 무선통신 환경에서 다양한 단말과 네트워크 성질에 맞는 서비스를 위해서는 MPEG-4 FGS와 같은 스케일러블 코딩 기술과 부호화된 비트스트림을 각 채널 대역폭에 맞게 적절한 비트율로 추출하여 전송하는 기술이 요구된다. 본 논문에서는 MPEG-4 FGS 향상계층의 각 비트평면 데이터 추출 시 왜곡의 정도에 기반한 추출 알고리즘을 제안하였다. 제안된 알고리즘의 기본 핵심은 같은 비트율에서 전송화질의 극대화를 지원하면서 서버의 상황에 따른 추출방식을 달리하여 실시간 처리를 지원하도록 하는 것이다. 제안된 추출기법은 기존의 추출방식에 비하여 같은 수행시간을 유지하면서 더 나은 성능을 보여준다.

참고문헌

- [1] W. Li, "Overview of Fine Granularity Scalable in MPEG-4 Video Standard," IEEE Trans. Circuit and System for Video Technology, vol. 11, no. 3, pp. 301~317, March 2001.
- [2] Y.Zhang, "Streaming video over internet-issues and new deleopment," Proc, IEEE Wokshop on Signal Processing System, Taipei, Tawan, Oct 2000.
- [3] T. .Hou, D. Wu, W. Zhu, H. Lee, T. Chiang, Y. Zhang, "An end-to-end architecture for MPEG-4 video streaming over the Internet," Proc. Int. Conf. Image Processing Vol. 1, Kobe, Japan, pp. 254~257, 1999
- [4] W. Li. J. Ohm, M. V. Scahaar, H. Jiang., and S. Li. MPEG-4 Video verifaicon model ver. 18.0, ISO/IEC/JTC1/SC29/N3908, 2001.
- [5] H.Radha, Y. Chen, K. Parthasarathy, and R. "Scalable internet video using MPEG-4", Signal Processing: Image Communication, vol. 15, pp, 95~126, 1999.
- [6] Streaming video profile for FGS RDAM, ISO/IEC/JTC1/SC29/WG11/M5673, 2000