

스케일러블 비디오 부호화 Phase 1 기술

최해철 · 강정원 · 배성준 · 유정주(ETRI 전파방송연구단)

1. 서론

근래의 무선망 및 인터넷의 초고속화에 따른 정지영상 및 동영상 등의 정보를 포함한 다양한 멀티미디어 서비스가 통방융합 현상과 더불어 새로운 패러다임으로 활성화 되고 있으며, BcN(Broadband convergence Network) 및 통방융합망의 등장으로 압축 부호화 기술만을 개발하던 시기와 달리 최근의 디지털 멀티미디어의 생성, 전송 및 소비 환경은 광범위하게 급변하고 있다. 예를 들어 서로 다른 특성 전송 용량 및 에러 발생율을 갖는 이중망이 혼재해 있으며, DMB(Digital Multimedia Broadcasting) 서비스와 같은 통신과 방송의 융합화가 빠르게 진행되고 있다. 더욱이 비디오를 소비하는 단말 또한 다양한 화면 크기와 연산 능력을 가진 새로운 제품들을 주변에서 흔히 볼 수 있다. 지금까지 산업에서 주로 쓰인 비디오 압축 기술은 고정된 영상 포맷과 비트율(bit-rate)로 압축/전송하는 단일 계층 압

축 부호화(single layer video coding) 방식이었다. 즉 주변에서 흔히 볼 수 있는 mp4, avi, wmv 형식의 비디오 파일들은 고정된 영상 크기 및 화질을 가지고 있다. 앞서 언급한 환경에서 이러한 단일 계층 압축 부호화 방식은 멀티미디어의 생성에서부터 사용자가 직접 시청하는 단말의 소비까지 끊김 없는 품질의 보장에 대한 사용자의 요구를 효과적으로 충족시킬 수 없다.

이러한 환경에서 각 네트워크의 조건과 비디오를 시청하는 단말기의 성능에 맞추기 위해 변환부호화(transcoding)¹⁾나 비트량변환(transrating) 방식 등이 개발되어 왔다. 그러나 이 방식들은 제한된 범위에서의 영상 형식 및 비트율 적응이 가능하지만 큰 복잡도로 인해 실시간 적응이 요구되는 환경에서 널리 쓰이지 못했다. 따라서 언제 어디서나 어떠한 단말로든 효과적으로 동영상을 전송/소비하기 위해서는 앞서 기술한 전송 및 소비 환경에 따라 적응적으로 변환할 수 있는 스케일러블리티

1 韓國電子通信研究院: 본 연구는 정보통신부 및 정보통신연구진흥원의 IT 신성장동력핵심기술개발 사업의 일환으로 수행하였음.
[2005-S-103-02, 통방융합 환경에서의 유비쿼터스 콘텐츠 서비스(UCA) 기술 개발]

(scalability)라는 개념이 생겨 났으며 최근 이에 대해 활발한 연구가 진행 중이다.

동영상에 있어서 스케일러빌리티 기능을 위한 기존 부호화 표준은 MPEG-2^[2], H.263^[3] 그리고 MPEG-4 (part2) Visual^{[4][5]}까지 계층 부호화에 기반한 여러 차례의 시도가 있었다. 하지만 이들은 하나의 계층으로 부호화하는 기존 압축 부호화 기법에 비해 부호화 효율이 낮고, 여러 종류의 스케일러빌리티를 종합적으로 지원하지 못하는 취약점이 있었다. 전송/소비 환경 측면에서, 과거에는 전송망이 고정된 처리율(throughput)을 가지며 패킷의 전송 성공 혹은 실패로 국한한 정책으로 설계되었기에, 비트율을 세밀히 변화 시키면서 전송망의 상황에 맞춘 서비스의 필요성이 대두되지 않았으며, 단말 시장 또한 고정된 TV 혹은 PC(personal computer)에서 비디오를 시청하였기에 이 표준들이 산업적으로 성공을 거두지 못하였다.

근래의 환경은 앞서 기술한 바와 같이 다양한 단말과 전송망이 혼재해 있다. 따라서 비디오 부호화에서의 스케일러빌리티의 필요성이 점점 절실하게 되었으며, 과거 취약점을 해결하면서 실시간 비디오 전송에서의 다양한 스케일러빌리티를 지원하기 위해 현재 ISO/IEC Moving Picture Experts Group (MPEG)과 ITU-T Video Coding Experts Group (VCEG)의 Joint Video Team (JVT)은 Scalable Video Coding (SVC)^{[6][7]} 표준화를 MPEG-4 part 10 MPEG-4 AVC/H.264^[8]의 수정판(Amendment)으로 진행하고 있다. 여기에는 유럽, 미국, 한국, 일본, 중국 등이 적극적으로 참여하고 있다.

본 고에서는 스케일러블 비디오 부호화의

개요에 대해 설명하고, 이제 표준화가 거의 완료된 SVC phase 1의 표준 채택 기술들을 소개한다.

II. 스케일러블 비디오 부호화 개요

SVC 표준에서 스케일러빌리티(scalability)란 'MPEG-4 AVC/H.264 부호화와 비교 시 비슷한 비트율-왜곡(Rate-Distortion) 성능을 가짐과 동시에 공간, 시간, 화질 측면에서 비트스트림의 한 부분을 제거할 수 있는 기능(functionality)'으로 정의한다. SVC는 여러 개의 비디오 계층을 하나의 비트열로 부호화하며, 각 층은 각각의 비트율, 프레임율, 영상 크기 및 화질을 가지고 있다. 즉 비디오를 공간적/시간적/화질적 차원의 임의 값을 가지는 비트열로 부호화하며, 그 세 가지 차원의 조합에 따라 폭넓은 스케일러빌리티를 제공할 수 있다. 기술적인 관점에서, 하나의 스케일러블 비트스트림은 두 개 혹은 그 이상의 의존적인 계층으로 구성될 수 있다. 이 경우, 스케일러블 코덱(codec)은 하나의 기본 계층과 스케일러블 상위 계층들로 구성된다. 최상위 계층의 복호 시 자신 및 자신 보다 낮은 계층의 정보들이 함께 이용된다. 화질 스케일러빌리티를 예로 든다면, 일반적으로 기본 계층은 기본적인 비디오 화질을 제공하고, 연속된 상위 계층은 이전 계층들로 만들어진 비디오보다 높은 화질을 갖도록 부호화한다. 마찬가지로 시간 및 공간 해상도에서도 동일한 원리를 적용하여 스케일러빌리티를 지원한다. 이는 하나의 차원에 대한 스케일러빌리티로는 최적의 가변성과 적응성을 제공할 수 없는 응용분야가

있을 수 있으며, 이는 시간, 공간, 화질 측면 모두의 스케일러빌리티를 필요로 하기 때문이다.

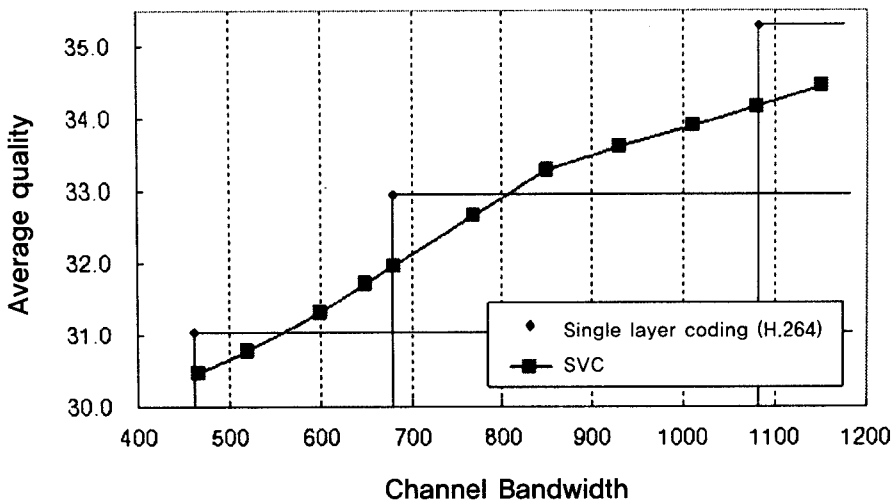
그림 1은 MPEG-4 AVC/H.264와 SVC 기술을 이용할 때의 비트율 대비 화질 곡선을 보여준다. 그림에서 단일 계층 압축 부호화 방식은 채널의 대역폭이 증가했음에도 불구하고 더 좋은 화질을 제공할 수 없지만, SVC는 대역폭에 따라 보다 좋은 화질의 비디오를 지원할 수 있다. SVC는 단일 계층 부호화 방식의 R-D 곡선과 일치하는 연속적인 곡선을 얻는 것이 이상적인 목표이나, 현재 SVC 표준 기술은 기존 MPEG-4 AVC/H.264와 비교 시 20% 내외의 비트율을 추가적으로 요구한다.

III. 스케일러블 비디오 부호화 기술

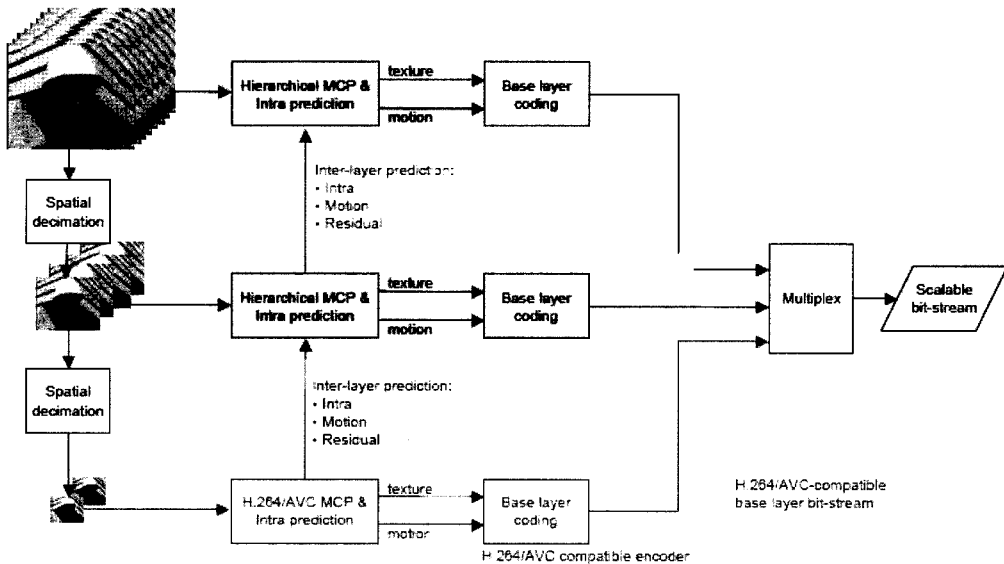
SVC의 시간적 스케일러빌리티는 Hierarchical B

picture¹⁰⁾ 구조로써 지원하고, 공간적인 스케일러빌리티는 과거 스케일러블리티를 제공하는 표준들에서 활용되었던 상향표본화(up-sampling)를 통한 다 계층 방식에 기반한다. 그리고 화질(quality 혹은 SNR) 스케일러블리티를 위해서는 CGS(Coarse Grain Scalability)의 경우 공간 스케일러블리티를 위한 기법과 동일한 방식에 양자화 계수(quantization parameter)만을 변경하며, MGS(Medium Grain Scalability)의 경우 양자화 에러에 대한 점진적 부호화 기법을 이용한다¹⁰⁾.

JSVM 부호기 구조는 제공하려는 스케일러블리티의 종류 및 단계에 따라 달라진다. 3개의 공간적 계층으로 구성하는 경우 SVC 부호기 구조는 그림 2와 같다. 공간적 스케일러블리티를 위해 입력 영상의 공간 해상도를 줄여 작은 크기의 영상을 하위 계층 각각에서 입력 받은 후, 각 계층에서 Hierarchical B picture를



〈그림 1〉 비디오 부호화 방식에 따른 채널 대역폭별 화질 성능 그래프



〈그림 2〉 공간적 3계정의 스케일러빌리티를 제공하는 스케일러블 부호기 예

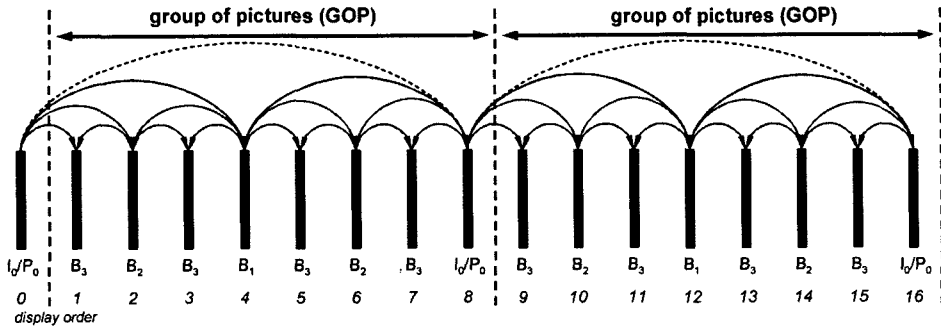
적용하여 시간적 스케일러빌리티를 지원하며, 움직임 정보와 잔여 텍스처 정보는 MPEG-4 AVC/H.264 알고리즘을 기초로 부호화한다. 부가적으로 화질 스케일러빌리티를 제공하기 위해 각 공간 계층에서 MGS 기법이 적용 될 수 있다. 각 공간 계층에서의 부호화 시, 하위 계층에서 이미 부호화된 화면내 부호화 정보, 움직임 벡터, 잔여 신호 등은 현재 계층의 예측 신호로 활용될 수 있다. 공간적 첫 번째 계층은 반드시 MPEG-4 AVC/H.264와 완벽한 호환성을 가져야 하는 제약이 있다.

1. 시간 스케일러빌리티 관련 기술

MPEG-4 AVC/H.264 에서 모든 영상은 참조 영상으로 사용할 수 있고, 다음에 오는 영상의 움직임 보상 예측에 이용될 수 있다.

Memory management control operation (MMCO) commands는 복호화된 영상을 저장할 수 있는 decoded picture buffer (DPB)를 적절하게 제어하며, DPB에 저장된 다른 영상의 움직임 보상 예측에 이용되는 참조 영상들은 reference picture list re-ordering (RPLP) commands를 사용하여 임의적으로 선택이 가능하다. 이러한 MPEG-4 AVC/H.264의 특징들은 이전 비디오 부호화 방식들이 지원하지 못했던 임의적인 부호화 및 참조 영상의 선택을 가능하게 한다. 이를 기반으로 B 영상의 참조를 허용하는 General B picture를 정의함으로써 시간적 스케일러빌리티를 지원하는 Hierarchical B picture가 제안되었다.

그림 3은 4단계를 갖는 전형적인 hierarchical B picture 구조를 보여준다. 비디오의 첫 영상은 IDR (Instantaneous Decoder Refresh) 영상으



〈그림 3〉 GOP 크기가 8이며, 4개의 시간적 레벨을 갖는 hierarchical B picture 구조

로서 화면내 부호화 된다. 이전에 부호화된 모든 영상들이 현재 부호화되는 영상보다 화면 표시 순서(display order) 상 앞서 위치할 때, 이것을 key picture(그림 3의 8번째, 16번째 프레임)라고 부른다. 시간적으로 현재 key picture와 이전 key picture 사이에 위치하고 있는 모든 non-key picture들을 하나의 group of pictures (GOP)라고 한다. Key picture는 임의 접근을 위해 화면내 부호화되거나, 이전 key-picture를 참조 영상으로 움직임 보상 예측을 통해 P 영상으로 화면간(inter) 부호화 된다. GOP의 영상들은 그림 3과 같이 계층적으로 참조되어 다양한 프레임율을 가질 수 있다. 예를 들어 B3 영상들을 모두 보내지 않는 경우 원래 프레임율의 반으로 복호화 가능하다.

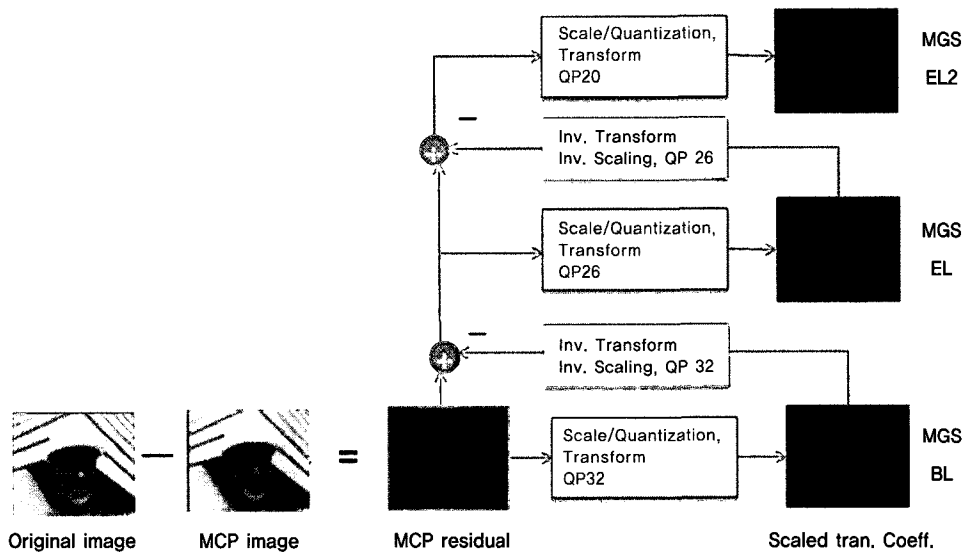
2. 공간적 스케일러빌리티 기술

공간적 스케일러빌리티는 여러 영상 크기의 계층을 쌓는 다 계층 부호화 기법에 기반한다.

이 때 각 해상도의 계층은 독립적으로도 부호화 할 수 있지만, 하위 계층의 부호 결과물인 복호된 화소값, 움직임 벡터, 및 잔여 신호 정보를 상위 계층에서 이용 함으로써 부호화 효율을 높일 수 있다. CGS 스케일러빌리티도 공간적 스케일러빌리티와 동일한 방식으로 지원한다. 단지 다른 점은 각 계층의 입력 영상 크기가 동일하는 것이다. 본 절에서는 각 공간적 또는 CGS 계층의 계층간(inter layer) 상관 관계를 이용하기 위해서 제안된 3가지 계층간 예측 기술을 설명한다.

가. Inter-layer intra texture prediction

하위 계층의 복호된 화소 정보를 이용하는 계층간 화면내 텍스처 예측(Inter-layer intra texture prediction)을 Intra_BL 모드라 한다. 현재 표준에서는 복호기의 과도한 연산을 유발하는 움직임 보상을 여러 계층에서 수행하는 것을 금지하기 위해, 하위 계층의 대응 위치



〈그림 4〉 화질 스케일러빌리티를 위한 부호화기 구조

블록이 화면내 부호화된 매크로블록인 경우에만 Intra_BL 모드가 적용될 수 있다.

공간 계층간 영상의 크기가 다른 경우 Intra_BL 모드에서 하위 계층으로부터 예측 신호를 얻기 위해서, 하위 계층의 동일 대응 지점에 위치한 블록의 복원된 화소 값에 대해 상향 표본화(up-sampling)를 수행한다. 이 때 밝기 값에 대해서는 4-tap cubic spline 필터, 컬러 값에 대해서는 bi-linear 필터가 적용된다. 그 상향 표본화된 신호를 현재 블록의 참조 신호로 활용하여 그 차이 값만을 부호화 한다. 이 Intra_BL 모드는 SVC에서 부호화 효율을 높이는 주된 기술 중 하나이다.

나. Inter-layer motion prediction

하위 공간 계층의 대응 블록이 화면간 부호

화 되었다면 상위 계층의 블록은 하위 계층으로부터 움직임 정보 및 모드 정보를 참조할 수 있다. 따라서 상위 계층에서는 움직임 정보 및 매크로블록의 분할(partitioning) 정보, 참조 영상 정보 등을 매크로블록 헤더에 부호화하지 않을 수 있다. 영상 시퀀스(sequence)를 낮은 비트율로 부호화할 때, 이러한 syntax 정보들의 비중이 상당히 높기 때문에 이들에 할당된 비트를 대폭 줄일 수 있는 이 기법은 높은 성능을 가질 수 있다.

다. Inter-layer residual prediction

공간 계층 간 움직임 정보에 차이가 크면, 잔여 정보를 예측하는 것은 부호화 효율에 도움이 될 수도 있고 그렇지 않을 수도 있으나, 현재 계층 블록의 움직임 정보가 하위 계층의

대응 위치 블록의 움직임 벡터와 동일하거나 거의 유사하면 이전 계층의 잔여신호가 현재 계층의 잔여 신호를 위한 예측으로 사용할 때 부호화 효율이 증대될 수 있다. 따라서 SVC는 하위 계층의 복원된 잔여 계수를 현재 계층의 잔여 계수를 부호화 할 때 예측 신호로 활용하는 것을 적응적으로 사용할 수 있다.

3. 화질 스케일러빌리티 관련 기술

SVC는 화질적 스케일러빌리티를 위해서 그림 4와 같이 양자화에 의한 오차 신호를 이전 계층 보다 적은 양자화 계수 값으로 보정하여 부호화 하는 개념이다. 즉 첫번째 계층은 큰 양자화 간격으로써 듬성듬성하게 표현한 후, 계층이 올라갈수록 양자화 간격의 폭을 세밀하게 조정하여 보다 좋은 화질의 계층을 쌓는 개념이다. 기술적으로 살펴보면, 앞서 설명한 Hierarchical B picture 에 의한 시간적 분할 (decomposition) 이후의 신호들을 최고 양자화 계수값(Quantization Parameter)으로써 최소 화질을 갖는 기본 계층을 부호화 한 후, 이 기본 계층을 다시 복호하여 원 신호와의 차이값을 구한다. 이 차이값, 즉 양자화 에러를 다시 DCT 변환하여 기본 계층보다 적은 양자화 계수(Quantization parameter)로 양자화 한다. 이때 QP는 이전 계층보다 6 적은 값을 선택하는데, 이는 MPEG-4 AVC/H.264에서 QP 값이 6이 줄면 양자화 간격 크기가 반 정도로 감소 (비트량은 약 두 배 증가)되기 때문이다.

IV. 결론

현재 ISO/IEC MPEG과 ITU-T VCEG의 JVT에서는 멀티미디어를 다양한 망 및 단말 환경에서 서비스할 때 QoS를 보장하기 위한 목적으로 스케일러빌리티를 제공하기 위한 Scalable Video Coding에 대한 표준화를 진행하고 있으며 2007년 7월 현재 SVC phase 1에 대해 거의 완료하였다. 따라서 본 고에서는 스케일러빌리티의 필요성과 SVC의 개요에 대해 소개하고, 스케일러빌리티를 지원하기 위한 대표적인 기술들에 대해 설명하였다.

결론적으로 SVC는 이종망 환경에서 발생하는 사용자의 선호도, 소비 단말기와 네트워크의 조건 등 다양한 전달 및 소비 환경에 효과적으로 대처할 수 있으며, 스케일러빌리티를 갖는 계층들의 조합을 하나의 비트열로 부호화하여 적응적인 전달/소비를 제공함으로써 여러 소비 환경 각각에 QoS를 보장하는 콘텐츠를 서비스할 수 있다. 더욱이 동일한 생산 비용이 여러 개의 전달/소비 환경에 공유될 수 있기 때문에 비용과 효율의 생산성 측면에서도 SVC는 매우 매력적이다. 추후 스케일러블 비디오는 최근의 멀티미디어 서비스 환경에서 보다 향상된 비디오 서비스와 새로운 시장의 기회를 제공할 수 있는 중요한 기술 중 하나일 것이다.

 참고문헌

- [1] A. Vetro, C. Christopoulos, and H. Sun. "Video transcoding architectures and techniques: An overview", IEEE Signal Processing Mag., vol. 20, pp. 18-29, Mar. 2003.
- [2] ITU-T and ISO/IEC JTC1, "Generic Coding of Moving Pictures and Associated Audio Information Part 2: Video," ITU-T Recommendation H.262 ISO/IEC 13818-2 (MPEG-2), 1994.
- [3] ITU-T, "Video coding for low bit rate communication," ITU-T Recommendation H.263, Version 1: Nov. 1995, Version 2: Jan. 1998, Version 3: Nov. 2000.
- [4] W. Li, "Overview of Fine Granularity Scalability in MPEG-4 Video Standard," IEEE Trans. on Circuit System and Video Technology. vol. 11, no. 3, pp. 301-317, Mar. 2001.
- [5] ISO/IEC 14496-2, "Information technology-Coding of audio-visual objects - part 2: Visual", International Standard, second edition, December 2001.
- [6] ITU-T document, "Joint Draft 10 of SVC Amendment," Joint Video Team JVT-W201, JVT 23rd meeting, San Jose, USA, April 2007.
- [7] ITU-T document, "Proposed modifications for Joint Scalable Video Model," Joint Video Team JVT-W202, JVT 23rd meeting, San Jose, USA, April 2007.
- [8] ITU-T and ISO/IEC JTC 1, "Advanced video coding for generic audiovisual services," ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG4-AVC), Version 1: May 2003, Version 2: Jan. 2004, Version 3: Sep. 2004, Version 4: July 2005.
- [9] H. Schwarz, D. Marpe, and T. Wiegand, "Analysis of hierarchical B pictures and MCTF," Proceedings of ICME' 06, Toronto, Canada, July 2006.
- [10] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable extension of the H.264/MPEG-4 MPEG-4 AVC/H.264 video coding standard,"

Joint Video Team JVT-U145, JVT 21st meeting, Hangzhou, China, October 2006.

 저자소개


 최해철

1997년 2월 경북대학교 전자공학과(학사)
 1999년 2월 KAIST전기 및 전자공학과(석사)
 2004년 8월 KAIST전기 및 전자공학과(박사)
 2004년 9월-현재 ETRI 전파방송연구단 방송미디어 연구그룹 선임연구원
 2007년 3월-현재 과학기술연합대학원대학교 이동통신 및 디지털방송공학 겸임 조교수
 주관심 분야 : 영상통신, 영상분할, 비디오 부호화, Scalable Video Coding, MPEG Video, ITU-T VCEG


 강정원

1993년 2월 한국항공대학교 항공전자공학과 (학사)
 1995년 2월 한국항공대학교 항공전자공학과 (석사)
 2003년 8월 Georgia Institute of Technology ECE (박사)
 2003년 10월-현재 ETRI 전파방송연구단 방송미디어 연구그룹 선임연구원
 2007년 3월-현재 과학기술연합대학원대학교 이동통신 및 디지털방송공학 겸임 조교수
 주관심 분야 : 비디오신호처리, 비디오 부호화, Scalable Video Coding, 비디오 적응 변환

저자소개



배성준

1997년 2월 고려대학교 전자공학과(학사)
 1999년 2월 KAIST전기 및 전자공학과(석사)
 2004년 8월 KAIST전기 및 전자공학과(박사)
 2004년 8월-2005년10월 하나로텔레콤 기술전략팀
 2005년 10월-현재 ETRI 전파방송연구단 방송미디어
 어연구그룹 선임연구원

주관심 분야 : 비디오 부호화, Scalable Video
 Coding, MPEG Video, ITU-T
 VCEG, 비디오스트리밍, 스케일러블스
 트리밍



유정주

1982년 2월 광운대학교 전자통신공학과(학사)
 1984년 2월 광운대학교 전자통신공학과(석사)
 2001년 7월 영국 LANCASTER대학교 컴퓨터공학과
 (박사)
 1984년 3월-현재 ETRI 전파방송연구단 방송미디어
 연구그룹 방통융합미디어연구팀장

주관심 분야 : 영상통신, 비디오 부호화, Scalable
 Video Coding, MPEG Video,
 ITU-T VCEG, 비디오스트리밍, 스케
 일러블스트리밍