

3DTV 방송을 위한 프레임 호환 패킹기술

원치선 (동국대학교)

I. 서론

영화로부터 시작된 3D 입체 영상물에 대한 최근의 흥행 성공은 3DTV의 도입을 재촉하고 있다. 실제로 TV 제조업체들은 이미 3D 영상 포맷을 디스플레이 할 수 있는 3DTV를 출시하고 있다. 그러나 시간축의 각 시점별로 우안과 좌안에 해당되는 스테레오 영상의 존재 자체로부터 기본적으로 3DTV 영상 전송의 대역을 배가시키고 3DTV 신호의 전송을 위한 기존의 영상 코덱 및 전송 인프라의 재사용을 어렵게 하고 있다. 따라서 3D 영상 포맷에 대한 새로운 압축 표준이 필요하게 되었으며 이를 위해 MPEG (Moving Picture Expert Group)에서는 다차원 비디오 부호화 MVC(Multiview Video Coding)의 압축 표준화를 진행하고 있다. 이와 같은 새로운 영상 압축 및 방송 시스템의 도입 시에 가장 문제가 되는 것은 역시 기존의 2D 디지털 TV(DTV)와의 호환성 문제이다.

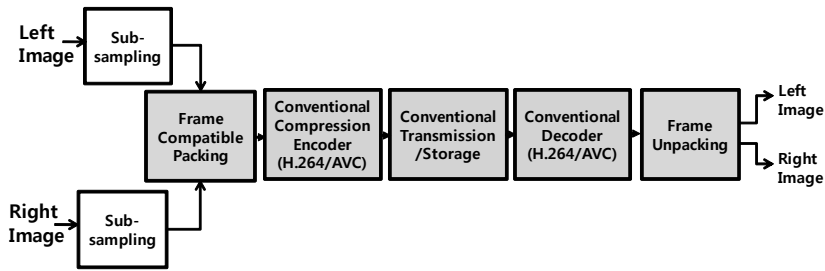
호환성을 염두에 둔 경우 3DTV (셋톱박스)는 소위 3D-Ready TV에서 3D-Capable TV로 진화할 것이다^[1]. 3D-Ready TV는 비압축 양안 영상을 입력 받을 수 있도록 되어 있으므로 HDMI1.4 규격과 같이 3D 영상의 디스플레이 방식에 따른 기본적인 스테레오 영상 포맷을 지원한다. 반면에 3D-Capable TV는 압축된 3D 비디오를 입력 받을 수 있도록 되어 있으므로 자체적으로 압축을 해제할 수 있는 디코더를 내장하고 있고 압축 비트 스트림의 출처에 따라 프레임 호환 포맷 (Frame-compatible format)을 지원하는 TV와 MVC와 같은 새로운 다차원 비디오 압축 포맷을 지원하는 Full HD 해상도의 3DTV로 나눌 수 있다. 프레임 호환 포맷을 지원하는 3D-Capable TV의 경우 양안 영상을 기존의 2D 비디오의 해상도에 맞게 축소하여 짜 맞추는 비디오에 대해 기존의 DTV(Digital TV) 비디오의 압축 방법에 의해 압축된 압축

비트 스트림을 받으므로 기존 DTV의 하드웨어와 전송 시스템을 상당부분 활용할 수 있는 장점이 있다. 반면에 양안 영상들의 해상도가 각각 절반으로 줄어들어 해상도의 손실을 감수해야 하는 단점이 있으며 이와 같은 해상도의 손실은 완전한 해상도(Full resolution)를 갖는 새로운 영상압축 방식을 채택하는 3DTV로 진화하면서 해소 될 것이다. 또 다른 시나리오로 고려될 수 있는 것으로 3DTV를 위해 MVC를 초기부터 직접적으로 도입하는 것인데 이 방법은 기존 DTV와의 호환성 문제를 어렵게 할 수 있다.

상기 언급한 바와 같이 기존 TV와의 호환성을 염두에 둔 경우 3DTV를 도입하는 과정에서 단계적인 진화의 시나리오가 불가피하며 이 경우 각각의 단계에서 출시한 3DTV가 향후 호환성을 유지하며 계속해서 활용될 수 있도록 고려하는 것이 중요하다. 반대로 초기에 도입된 방식이 향후 Full 3DTV 혹은 무안경 3DTV까지 발전할 수 있는 확장성도 갖추어야 한다. 이와 같은 관점에서 프레임 호환 포맷 (Frame-compatible format)을 지원하는 3D-Capable TV에서 Full HD 3D-Capable TV로의 진화를 위해 스케일러블 (Scalable) 부호화 기술을 도입하는 SREFS (Scalable Resolution Enhancement of Frame-Compatible Stereoscopic 3D Video) 부호화 접근 방법이 MPEG에서 거론되기 시작하고 있다^[2]. 본 논문에서는 SREFS 부호화 접근 방법의 개념과 간단한 사례를 소개한다.

II. 프레임 호환 포맷

프레임 호환 포맷 (Frame-compatible format)을 지원하는 3DTV는 <그림 1>과 같이 우안 및 좌안 영상에 대해 부표본화



〈그림 1〉 프레임 호환 포맷을 지원하는 방송 시스템 블럭도

(sub-sampling)를 적용한 이 후 공간 (혹은 시간) 해상도가 절반으로 줄어든 좌우 영상을 붙여 하나의 프레임을 형성하고 이 후에 기존의 2D DTV에 적용된 압축 코덱 및 전송 인프라를 이용하는 구조이다. 이때 적용된 부표본화 패턴에 따라 수직 라인을 부표본화 하여 수평 해상도가 절반씩 줄어든 좌우영상을 붙이는 Side-by-side packing 방법과 수평라인에 대해 부표본화를 적용한 영상을 상하로 붙인 Top-bottom packing, 그리고 수평 및 수직 라인 별로 한 픽셀씩 어긋나게 부표본화 하여 전체적으로 마름모꼴의 표본화 패턴을 갖는 Quincunx (혹은 Checkerboard) packing 방법, 그리고 시간축 방향으로 좌우 영상 프레임을 번갈아 가면서 표본화하는 Frame Sequential packing 등이 가능하며 어떤 패킹방법을 사용했는지를 나타내기 위해 H.264/AVC 표준에서 Supplementary Enhancement Information (SEI) 메타데이터를 정의한바 있다^[3]. 즉, SEI 메시지를 번역할 수 있는 TV 수신기는 프레임 호환 포맷에 의해 패킹된 비디오로부터 좌우 양안영상을 분리해서 3D 영상을 적절히 디스플레이 할 수 있다.

프레임 호환 포맷 패킹 방법들이 기존의 영상 코덱과 전송 인프라를 그대로 활용할 있다는 장점이 있는 반면에 패킹된 영상에 새로운 왜곡이 나타날 수 있다. 즉, 패킹 방법 마다 차이가 좀 있지만 전반적으로 패킹 후 고주파 성분이 증가하여 압축 효율을 저해할 수 있으며 cross-talk과 color-bleeding의 부작용이 발생할 수 있다. 단순히 표본화의 관점에서 보면 Quincunx 부표본화 방법이 원 신호의 주파수 성분을 더 잘 보존하는 것으로 알려져 있다. 그러나 수평 및 수직 방향의 한 픽셀 오프셋을 갖는 Quincunx 부표본화를 좌우 영상에 적용한 후에 포개어 패킹하는 방법은 수신단에서 패킹을 해제하여 좌우 영상을 분리하는 과정을 복잡하게 할 수 있는 단점이 있으므로 Quincunx 부표본화 후에 Side-by-side나 Top-bottom으로 패킹하는 방법이 Quincunx 부표본화와 Side-by-side/Top-bottom 패킹의 장점을 모두 살리는 방법이 될 수 있다^[4]. 이와 같이 다양한 부표본화를 통한 프레임 호환 패킹 방법의 여러 변형이 가능한 가운데 일부에서는 포맷에 대한 표준을 구체화하는 작업에 착수하고 있다. 예를 들

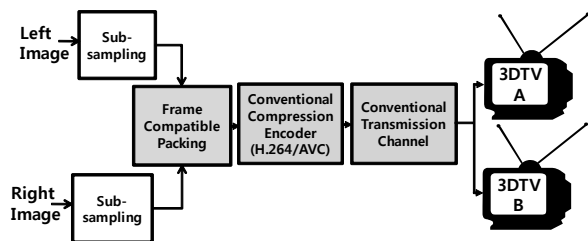
어, 최근에 Cable Television Laboratories, Inc.에서 프레임 호환 3D 서비스를 위한 포맷 규정을 발표하였다^[5].

III. 3D-Capable TV의 진화 시나리오

II 장에서 소개한 프레임 호환 포맷을 사용한 3DTV 전송 방법은 기존 DTV의 코덱과 전송 인프라를 그대로 활용할 수 있다는 장점이 있는 반면 수신된 양안 영상들의 해상도에 손실이 발생할 수 있다는 단점이 있다. 따라서 완전 HD 해상도의 3DTV 코덱과 전송 방식이 도입되기 이전까지 기존 2DTV의 프레임 호환 포맷 (Frame-compatible format)을 지원하는 3DTV가 다양한 형태로 진화할 것으로 예상된다.

1. 1세대 3D-Capable TV

프레임 호환 패킹으로 기존의 2D DTV의 해상도와 동일한 형태로 변환된 스테레오 영상은 기존 DTV의 압축코덱으로 압축되어 전송되며 1세대 3D-Capable TV에 입력된다. 1세대 3D-Capable TV는 기본적으로 다양한 프레임 호환 패킹의 포맷 방식(즉, Side-by-side 혹은 Top-bottom 등의 서로 다른 형태)을 구분하기 위해 전송되는 SEI 메시지를 번역할 수 있으며 번역된 프레임 패킹 방식에 따라 압축을 해제한 영상으로부터 좌우 양안영상을 추출할 수 있다 (〈그림 2〉의 3DTV-A). 이때 추출된 양안 영상과 사용된 프레임 패킹의 기법을 참조하여 줄어든 해상도를 원래 3D 양안 영상의 해상도로 복원하려는 기능을 탑재하거나 패킹에 의한 압축 왜곡



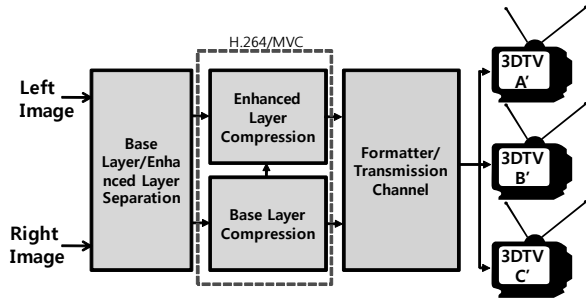
〈그림 2〉 1단계 3D-Capable TV

을 완화시키는 기능을 갖는 TV로 발전할 수 있다 (그림 2)의 3DTV-B). 예를 들어, 사용된 프레임 패킹 기법에 따라 삭제된 픽셀들의 값들을 이웃하는 삭제되지 않고 보존된 픽셀의 밝기 값으로 추정하여 복원하는 인터플레이션(interpolation) 기법을 시도하는 기능을 탑재할 수 있다.

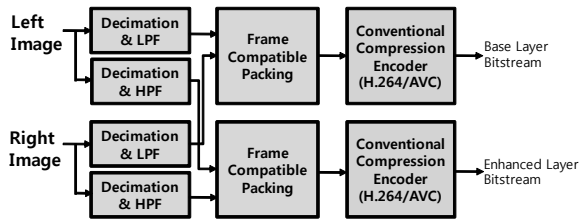
2. 2세대 3D-Capable TV

〈그림 1〉의 1세대 3D-Capable TV의 경우 기존의 DTV의 송신단(방송국) 부호화 시스템에 대한 변형이 거의 필요하지 않다. 반면에 2단계 3D-Capable TV의 경우 〈그림 3〉과 같이 송신단의 부호화 시스템에 대한 업그레이드가 필요하다. 2단계 3D-Capable TV는 기본적으로 〈그림 2〉의 3DTV-A와 3DTV-B와 호환성 보장을 전제로 한다. 따라서 기존의 압축 코덱과 프레임 호환 패킹에 기반한 압축 비트 스트림을 기본 층(Base layer)의 비트 스트림으로 제공할 수 있어야 한다. 이 기본 층의 비트 스트림은 3DTV-A와 3DTV-B의 수신단에서 복호화되어 종전의 화질로 3D 영상을 디스플레이 할 수 있다. 그러나 Full HD 해상도 수준의 추가적인 화질의 향상이 필요한 경우를 위해 송신단에서 기본 층의 화질과 양안 소스의 Full HD 사이의 화질 차이를 보상할 수 있는 정보를 포함한 여분의 정보 (side information)를 갖는 향상 층(Enhanced layer)의 비트 스트림을 추가적으로 제공할 수 있으며 이 향상 층의 비트스트림을 검출하여 복호화 할 수 있는 3DTV-C는 3DTV-A와 3DTV-B에 비해 향상된 화질의 3D 영상을 디스플레이 할 수 있다.

〈그림 3〉의 2단계 3D-Capable TV는 최근에 MPEG에서 거론되고 있는 SREFS (Scalable Resolution Enhancement of Frame-Compatible Stereoscopic 3D Video) [2] 부호화 접근 방법의 하나로 간주될 수 있다. 이 외에 또 다른 SREFS의 구조는 〈그림 4〉와 같이 MPEG의 MVC 표준을 활용하는 것이다. 〈그림 4〉에서 양안 영상은 우선 기본 층(Base Layer)과 향상 층(Enhanced Layer)의 프레임으로 분리된다. 예를 들어, 입력 소스 양안 영상의 홀수 번째 라인과 짝수 번째 라인



〈그림 4〉 2단계 3D-Capable TV (MPEG 고려 모델)

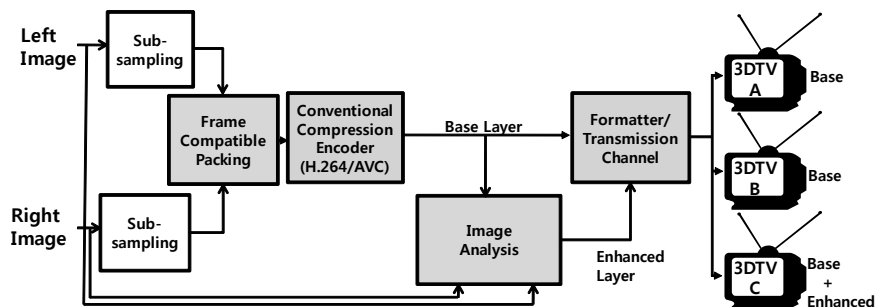


〈그림 5〉 Scalable Resolution Stereoscopic Video (SRSV) compression system

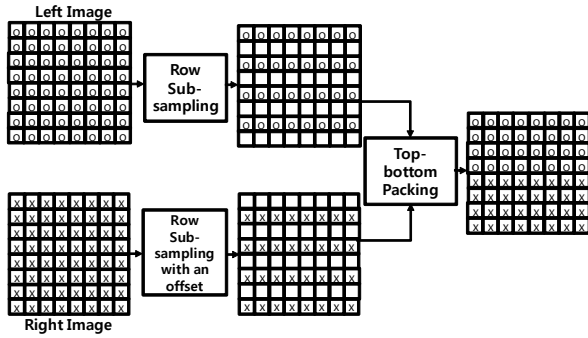
을 분리해서 생성한 두 쌍의 양안 영상에 대해 각각 H.264/MVC와 같은 압축 방법으로 기본 층과 향상 층을 압축하거나 〈그림 5〉와 같이 입력 소스의 양안 영상에 필터를 적용하여 저주파 성분과 고주파 성분으로 분리된 두 쌍의 양안 영상을 프레임 호환 패킹을 수행한 후 저주파 성분의 프레임은 기본 층으로 고주파 성분의 프레임은 향상 층의 압축을 적용하는 방법 등이 논의되고 있다 [6,7]. 현재 어떤 방식으로 3DTV의 진화 시나리오를 결정할 것인가에 대해 MPEG에서 논의되고 있으며 기본적으로 압축화질 뿐만이 아니라 호환성과 발전성의 측면에서 진화의 시나리오를 결정해야 할 것이다.

IV. 사례소개

이 장에서는 프레임 호환 패킹 시에 부표본화 된 영상을 원래의 크기로 확대하기 위해 주어진 인터플레이션 기법들 중에 어떤 기법이 가장 적절한지 인코더에서 결정하고 결정



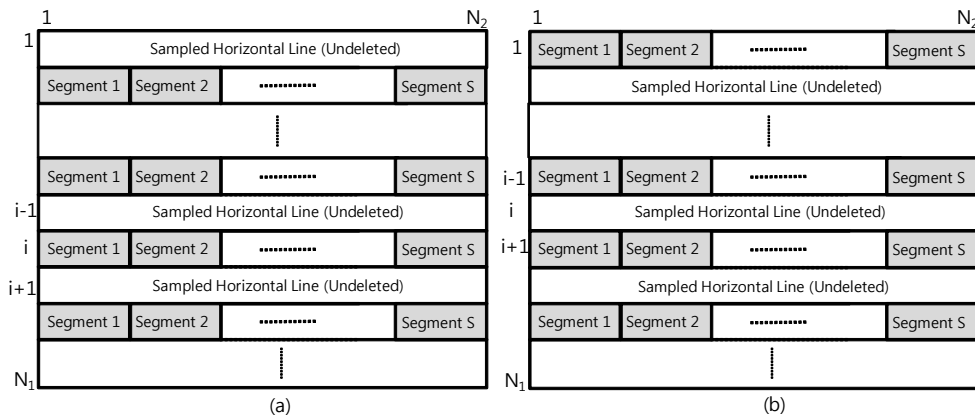
〈그림 3〉 2단계 3D-Capable TV



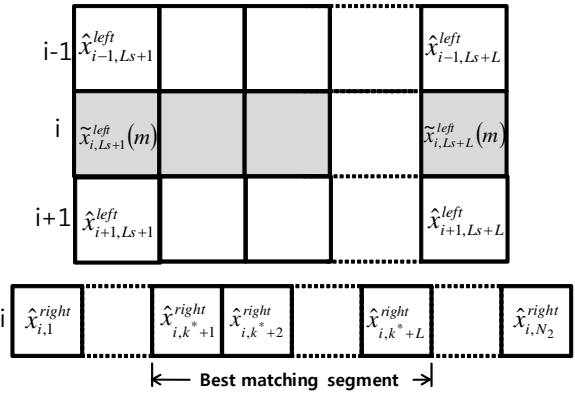
〈그림 6〉 Top-Bottom 패킹

된 인터플레이션 기법에 대한 정보를 여분의 정보(side information)로 수신 단에 전송함으로써 수신 단에서 이 여분의 정보를 번역할 수 있는 3DTV는 압축 해제 후 해상도 복원을 위해 각 부위별로 적절한 인터플레이션 기법을 적용할 수 있으므로 향상된 화질의 양안 영상을 복원할 수 있는 방법을 소개한다 [8].

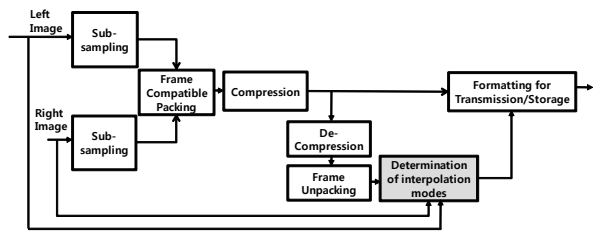
우선 〈그림 6〉과 같이 주어진 양안 영상에 대해 수평라인에 부표본화를 적용하는 Top-bottom 프레임 패킹을 적용하면 수신단 3DTV에서는 부표본화로 제거된 수평라인의 픽셀들을 인터플레이션을 통해 복원하며 이를 위해 〈그림 7〉과 같이 부표본화로 제거된 각 수평라인을 일정 단위(segment)로 나누어 각 세그먼트별로 미리 정한 일정 수의 인터플레이션 방법들 중에 가장 적절한 방법을 선택한다. 이때 각 세그먼트별로 미리 정해진 인터플레이션 함수들은 해당 세그먼트와 접하고 있는 상하의 수평라인 데이터와 반대쪽 View의 같은 위치의 수평라인 중에 수평 방향의 이동 (Parallax)을 탐색하여 연속성이 가장 장 보존되는 세그먼트의 데이터를 활용한다 (〈그림 8〉 참조). 최종적으로 각 세그먼트별로 선택된 인터플레이션 기법에 대한 정보는 여분의 정보(혹은 이 정보를 향상 층의 비트 스트림으로도 볼 수 있음)로 취급하여 수신 3DTV로 전송되며 디스플레이 전에 영상 확대를 위한 인터플레이션에 활용된다 (수신 단의 전체적인 구조는 〈그림 9〉 참



〈그림 7〉 세그먼트 구조 : (a) 좌안 영상, (b) 우안 영상

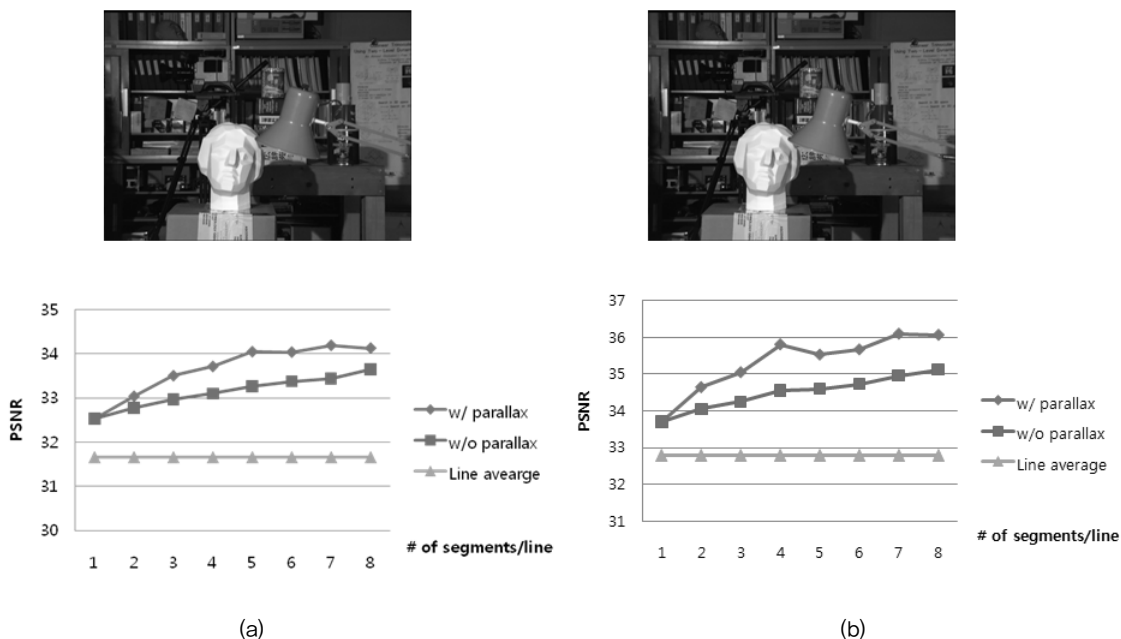


〈그림 8〉 좌안 영상의 i 번째 수평라인의 s 번째 세그먼트에 대한 인터플레이션 방법 m 을 결정하기 위해 참조하는 3개의 데이터 세트: (1) 좌안 영상 $i-1$ 번째 라인의 해당 위치 세그먼트, (2) 좌안 영상 $i+1$ 번째 라인의 해당 위치 세그먼트, (3) 우안 영상의 i 번째 수평라인의 수평 이동 예측된 (k^*) 세그먼트 데이터



〈그림 9〉 인터플레이션 모드 전송을 여분의 정보(혹은 향상 층)로 하는 경우의 예

조). 참고로 각 수평라인을 세그먼트로 나눌 때 짧은 길이로 나눌수록 〈그림 10〉과 같이 영상 확대 이후의 화질이 향상되나 그 만큼 수신 3DTV에 전달해야 할 여분의 정보가 많아진다. 수신 단에 제공해야 할 인터플레이션 기법에 대한 여분의 정보량이 많지 않은 경우 (압축을 적용하여) 기존 압축 포맷의 메타 데이터로서 보조 필드 (ancillary field)에 실어 전송할 수도 있다. 이 경우는 〈그림 3〉의 3DTV-C에 해당된다.



(그림 10) 인터플레이션 이후 화질 변화를 나타낸 그래프로 제거된 라인 당 세그먼트 수에 따른 화질(PSNR)을 나타냄. w/ parallax는 인터플레이션 대상 view의 상대쪽 같은 위치의 세그먼트 데이터를 수평 parallax 예측으로 결정한 방법을 나타냄. 반면 w/o parallax는 parallax 예측 없이 상대 view의 같은 위치에 있는 세그먼트 데이터를 활용함. ([8] 참조) : (a) 좌안 영상, (b) 우안 영상

V. 결론

방송의 관점에서 3DTV는 현존하는 DTV와 당분간 호환성을 유지하면서 발전하다가 궁극적으로 양안 영상에 대해 모두 Full HDTV를 지원하는 3DTV로 진화할 것이다. 이 과정에서 초기에 3DTV의 압축 및 전송 방식을 어떻게 설정할 것인가는 현존하는 DTV 및 향후 진화 과정의 3DTV와의 호환성과 궁극적으로 무안경 Full HD 3DTV로의 확장성을 모두 고려하는 신중한 선택이 되어야 할 것이다. 따라서 프레임 호환 패킹에 기반을 둔 새로운 스케일러블 코딩이나 프레임 호환 패킹 기법과 MVC의 적절한 결합의 형태가 될 것으로 예상된다.

감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2010-(C1090-1011-0003)).

참고문헌

- [1] A. Vetro, W.-H. Lou and M. Flynn, "TV Architecture Supporting Multiple 3D Services," IEEE ICCE, 2010.
- [2] ISO/IEC JTC1/SC29/WG11, "Use cases and

requirements for the enhancement of MPEG frame-compatible stereoscopic 3D video," N11681, October, 2010.

- [3] G. J. Sullivan, A. M. Tourapis, T. Yamakage, C. S. Lim, eds., "Draft AVC amendment text to specify Constrained Baseline profile, Stereo High profile, and frame packing SEI message," Joint Video Team, Doc. JVTAE204, London, United Kingdom, July, 2009.
- [4] A. Vetro, T. Wiegand, and G. J. Sullivan, "Overview of the Stereo and Multiview Video Coding Extensions of the H.264/AVC Standard," to appear in Proceedings of the IEEE, 2011.
- [5] Cable Labs, OpenCable Specifications Content Encoding Profiles 3.0 Specification OC-SP-CEP3.0-I01-100827, Sep., 1, 2010.
- [6] K. Minoo, et al, "On scalable resolution enhancement of frame-compatible stereoscopic 3D video," ISO/IEC JTC1/SC29/WG11 m18486, Guangzhou, China, Oct., 2010.
- [7] K. Minoo, et al, "Scalable resolution enhancement of frame-compatible stereoscopic 3D video," ISO/IEC JTC1/SC29/WG11 m19008, Daegu, Korea, Jan., 2011.
- [8] C. S. Won, "Adaptive Interpolation for 3D Stereoscopic Video in Frame-Compatible Top-Bottom Packing," IEEE ICCE, 2011.



원 치 선

1982년 2월 고려대학교 전자공학과 학사.
1986년 2월 University of Massachusetts, Amherst 석사.
1990년 2월 University of Massachusetts, Amherst 박사.
1989년 11월~1992년 8월 금성사 중앙연구소 선임연구원.
1992년 9월~현재 동국대학교 전자전기공학부 교수.
<관심분야> 영상처리, 영상분할, 영상검색, 워터마킹