

일반논문-08-13-1-18

이종 해상도 스테레오 비디오 코딩 시스템

박시내^{a)}, 심동규^{a)†}

Heterogeneous Resolution Stereo Video Coding System

Seanae Park^{a)}, and Donggyu Sim^{a)†}

요 약

본 논문에서는 스테레오 시점 및 디스플레이 특성을 고려한 효율적인 스테레오 비디오 코딩 방법을 제안한다. 현재 많은 스테레오 비디오 디스플레이 장치는 기존의 단일 시점 비디오 출력과 스테레오 비디오 출력이 모두 가능하다. 이러한 장치에서 스테레오를 비디오 출력 시에는 스테레오 디스플레이 패널의 특성으로 인하여, 두 개 시점 비디오 각각을 수평축으로 1/2 다운샘플링된 비디오를 입력받는다. 본 논문에서는 기존의 단일 비디오와 스테레오 비디오 시스템을 동시에 고려하기 위하여, 좌측영상은 원래의 해상도로 코딩하고, 우측영상은 가로 방향으로 1/2 다운 샘플링 하여 코딩하는 방법을 제안한다. 우측 영상의 코딩 시 좌측 영상을 참조 하는 경우 참조 영상과 원 영상사이의 해상도가 일치하지 않아 시점 간 예측이 불가능하므로, 본 논문에서는 이를 위해 시점간 예측방법을 변경하여 부호화하는 방법을 제안한다. 실험결과 제안된 방법은 기존의 스테레오 비디오 부호화 방법에 비하여 0.5~0.8 dB 이상의 코딩 이득을 얻을 수 있었다.

Abstract

In this paper, we propose an effective stereo-view video coding method that considers stereo-view and displayer characteristics. Current many stereo video displayers are designed for not only stereo display but also conventional single view display. In these systems, the resolution of two input videos for a stereo mode is half of that of single view for compatibility with conventional single view video services. In this paper, we propose a stereo video codec to deal with both single view and stereo view services by encoding whole left image and down-sampled right image. However, direct disparity estimation is not possible between two views because the resolution of a left image is different from that of the corresponding right image. So, we propose a disparity estimation method to make use of full information of the left reference image without down-sampling. In experimental result, we achieved 0.5~0.8 dB coding gain, compared with several conventional algorithms.

Keyword : stereo-view, stereo video coding, different resolution, efficient coding

1. 서 론

영상기술의 발달로 기존의 아날로그 방식의 영상처리 방

법이 디지털 기반의 방식으로 변화하고 있으며, 그 결과 사용자는 보다 선명하고 고화질의 영상을 제공받을 수 있게 되었다. 이로 인하여 사용자는 영상매체를 이용한 서비스를 통해서 전보다 더 높은 현장감을 느낄 수 있게 되었으나, 사용자의 요구는 현재의 기술에 만족하지 않고 끊임없이 증가되고 있다. 현재 우리나라에서는 고화질의 영상을 방송하는 HDTV 서비스가 제공되고 있으며, 2010년까지 전

a) 광운대학교

Kwangwoon University

† 교신저자 : 심동규(dgsim@kw.ac.kr)

※ 본 연구는 "서울시 산학연 협력사업"을 통하여 이루어졌음

공중과 방송서비스를 기존의 아날로그 방송에서 HDTV 방송 서비스로 전환할 예정이다. 또한, 디지털 방식으로 영상을 전송하여 어디서나 영상서비스를 제공 받을 수 있는 Digital Multimedia Broadcast (DMB) 서비스가 상용화 되었다. 이처럼 디지털 기반의 비디오 서비스를 통해 사용자가 고화질 영상을 손쉽게 이용할 수 있게 되었지만, 단일 시점의 영상의 한계로 인하여 사용자가 느끼는 입체감과 현장감을 높이기 위해 한계가 있다. 사람이 입체감을 느끼게 되는 요소는 주로 두 눈을 통해 입력되는 영상의 양안시차 때문인데^[1], 현재 서비스되고 있는 디지털 비디오 방송은 단일 비디오만을 전송하기 때문에 사용자는 입체감을 느낄 수 없다. 이렇게 사용자의 요구가 디스플레이 관점에서 단순히 고화질의 2D 영상매체를 통한 서비스에서 벗어나 보다 현장감 있는 3차원 비디오 서비스로 변화됨에 따라, 이와 관련된 다양한 분야의 연구가 진행되고 있다. 이에 대한 연구의 일환으로 스테레오 비디오의 코딩기술 연구가 활발하게 진행되고 있으며^[2-9], 3D 압축 기술에 대하여 Joint Video coding Team (JVT)의 주도하에 Multi-view Video Coding (MVC) 코덱에 대한 표준화가 진행 중이다.

스테레오 시점 디스플레이 시스템은 인간의 시각시스템의 특징을 이용한 기술로, 앞서 설명한 양안시차를 고려하여 둘 혹은 그 이상의 시점에서 촬영된 영상을 동시에 디스플레이 함으로써 사용자가 입체감을 느낄 수 있도록 하는 장치이다. 최근에 시판되고 있는 많은 스테레오 비디오 디스플레이 패널은 기존의 2D 비디오와 스테레오 비디오를 모두 처리할 수 있도록 설계되어 있다. 단일 비디오가 입력될 경우 기존 패널과 같이 동작하고, 스테레오 모드에서 두 시점 비디오가 동시에 입력될 경우, 각 시점의 비디오가 사용자의 좌, 우 각 눈에 상이 맺히도록 하여 사용자가 입체감을 느끼도록 한다. 이 때, 스테레오 패널의 화소수의 제한으로 인하여 스테레오 비디오의 경우 단일 비디오 해상도의 절반에 해당하는 해상도를 가진 양안 비디오를 입력 받게 된다. 본 논문에서는 스테레오 비디오를 코딩하는 방법으로 기존 시점 비디오는 기존의 H.264/AVC 기반으로 부호화하고^[10], 확장 시점 비디오는 기존시점 공간 해상도의 절반에 해당하는 해상도를 가지도록 다운 샘플링한 후에 기존 시점 비디오를 참조하여 부호화한다. 이때 참조 영상과

원 영상간의 해상도 차이로 인하여 기존의 움직임 예측 방법^[10]으로는 시점 간 참조가 불가능한데, 본 논문에서는 새로운 시점 간 예측 방법을 제안하여 이중 해상도의 시점이 다른 두 개의 비디오를 이용하여 시점 간 중복성을 제거하는 방법을 제안한다.

본 논문은 먼저 2장에서 스테레오 시점 비디오의 부호화에 대해서 설명하고, 3장에서는 제안한 스테레오 비디오 부호화 방법에 대해서 설명한다. 4장에서는 기존의 방법 및 제안된 방법의 실험결과를 비교하고 마지막으로 5장에서 결론을 맺는다.

II. 기존의 스테레오 비디오 부호화 방법

지금까지 영상 압축 및 전송에 관한 연구는 단일 시점 안에서 압축의 성능을 향상시키거나 화질을 개선하는 방향으로 진행되어 왔다. 하지만 사용자의 요구가 2D 비디오의 화질 향상에서 입체감과 실제감을 느낄 수 있는 입체영상 비디오로 변화에 감에 따라, 3차원 비디오 처리에 대한 연구가 시작되고 관련 표준화가 활발하게 진행되고 있다^[11-13]. 가장 단순하게 3D를 느낄 수 있는 스테레오 디스플레이의 경우, 이미 많은 디스플레이 장치가 개발 및 보급 되었고 관련 연구도 많이 진행된 상태이다. 스테레오 디스플레이를 위해서는 양안시차^[1]를 고려하여 좌, 우안 각각을 위해 촬영된 두 영상이 필요하고, 이 두 영상을 전송하기 위해서는 스테레오 비디오를 위한 부호화 알고리즘이 필요하다.

그림 1은 기존 시점이 되는 왼쪽 영상과 확장 시점이 되는 오른쪽 영상 두 개를 각각의 인코더를 통해 비트스트림으로 부호화 한 후에 전송하고, 이를 다시 각각의 디코더를 사용하여 디코딩 한 후에 디스플레이 하는 구조의 블록 다이어그램이다. 좌, 우 입력 영상이 각각 $N \times M$ 의 해상도를 가지며, 이를 각각의 부호화 및 복호화 장치를 통해 처리한다. 기준이 되는 좌측 영상은 2D 디스플레이 패널에서는 복호화 이후 $N \times M$ 의 해상도의 영상이 그대로 출력되고, 스테레오 디스플레이 패널에서는 가로방향의 필터링과 2:1 다운 샘플링을 통해 가로 방향으로 1/2 크기의 해상도를 가지는 영상으로 변환된다. 그리고 확장 시점의 우측영상 역

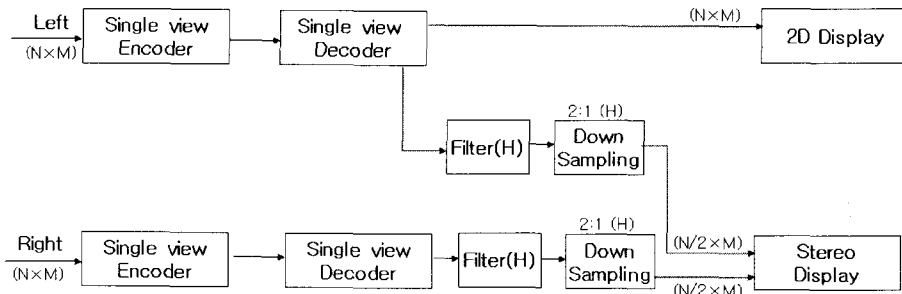


그림 1. 좌, 우측 영상을 각각 인코딩하는 스테레오 비디오 코딩 시스템
Fig. 1. Stereo video coding system that encodes each view image.

시 스테레오 디스플레이를 위하여 복호화 이후 가로 방향의 필터링과 2:1 다운 샘플링을 통해 가로방향으로 1/2 크기의 영상으로 변환되고, 변환된 좌측 영상과 함께 스테레오 디스플레이 장치를 통해 출력되어진다. 현재 대부분의 스테레오 디스플레이 장치가 이와 같이 좌, 우측 영상을 가로 방향으로 1/2 다운 샘플링하여 각각 절반의 영상만을 사용하는 구조로 이루어져 있다. 비디오 통신 시스템이 스테레오 비디오만을 지원 한다면 인코딩 이전 단계에서 영상의 크기를 줄여서 코딩의 효율을 높일 수 있으나, 시스템이 스테레오 서비스를 지원하더라도 기존의 2D 디스플레이 장치와의 호환성을 유지하여야 하기 때문에, 좌, 우측 영상 중 한 개의 영상은 반드시 원래의 영상 해상도로 코딩하여야 한다. 그림 1의 시스템은 기존의 2D 비디오 부호화 기술을 그대로 스테레오 시점에 적용시킨 것으로, 별도의 어떤 기술도 추가되지 않고 스테레오를 지원한다는 장점을 가진다. 하지만 시점 간의 참조를 전혀 하지 않고 각각 부호화하기 때문에, 시점 간에 중복되는 데이터가 많이 남게 되어

부호화 효율이 높지 않은 단점이 있다^{7,12}.

그림 2는 기준 시점이 되는 좌측 영상은 기존의 2D 디스플레이 장치와의 호환성 유지를 위해 원래 크기의 영상을 부호화 및 복호화 하고, 확장 시점이 되는 우측 영상은 인코딩 이전 단계에서 다운 샘플링하여 부호화 및 복호화하는 스테레오 비디오 시스템의 블록 다이어그램이다. 기준 시점의 영상은 2D 디스플레이 패널과의 역방향 호환성을 위하여 원래의 해상도인 $N \times M$ 의 크기로 부호화하고, 확장 시점 영상은 가로방향 필터링과 2:1 다운샘플링을 통해 $N/2 \times M$ 의 해상도를 가지는 영상으로 변환하고 이를 부호화한다. 이는 스테레오 디스플레이 패널이 해상도의 제한으로 인하여 출력 영상을 디스플레이 이전단계에서 가로방향으로 1/2 다운샘플링 하는 특성을 이용한 것으로, 부호화 이전에 영상을 다운 샘플링하여 코딩함으로써 코딩의 효율을 증가시킬 수 있다. 그러나 확장 시점 영상의 코딩 시 기준 시점 영상인 좌측 영상을 참조하지 않기 때문에, 그림 1에서의 시스템과 마찬가지로 각 시점 영상 사이의 중복성

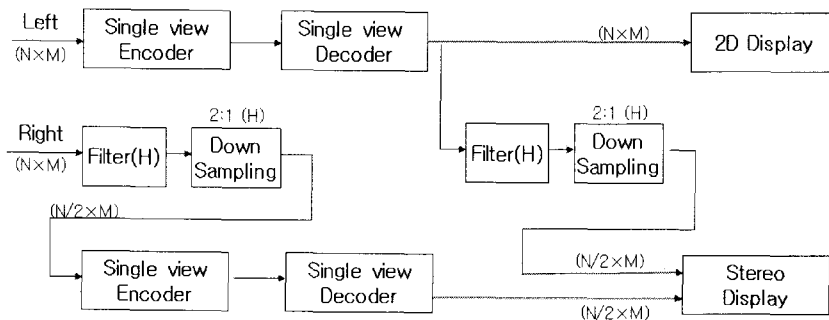


그림 2. 우측 영상을 다운샘플링 한 후에 각각 인코딩하는 스테레오 비디오 코딩 시스템
Fig. 2. Stereo video coding system that codes the original resolution for a left view video and the down sampled resolution for right view

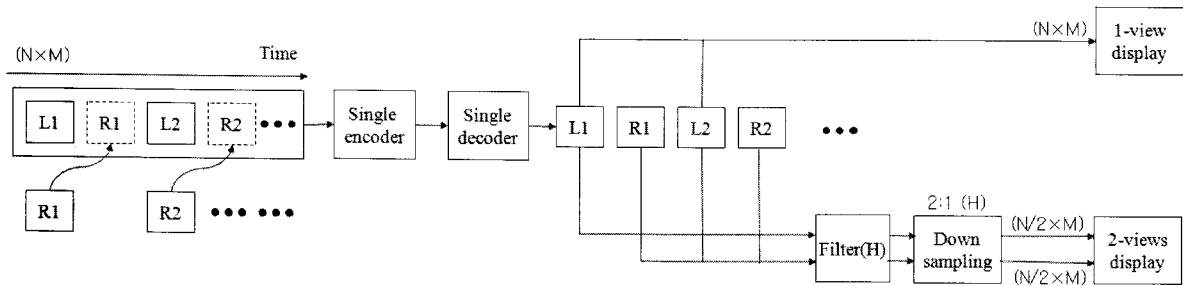


그림 3. 확장 시점 영상이 기준 시점 영상을 참조하는 H.264/AVC 기반 스테레오 비디오 코딩 시스템
 Fig. 3. Stereo video coding system that an extended-view is coded by referencing a base-view with the H.264/AVC codec

을 이용하지 못하게 된다^{[7][12]}.

그림 3은 하나의 인코더와 디코더를 사용하는 스테레오 비디오 시스템의 블록도를 나타낸다. 기준 시점인 좌측 영상과 확장 시점인 우측 영상을 하나의 영상으로 재구성한 후 부호화 및 복호화 하고, 이를 디스플레이 이전 단계에서 두 개의 시점으로 나누어 디스플레이 하는 구조의 블록다이어그램이다. 그림에서와 같이 부호화시 확장 시점 영상이 기준 시점 영상을 참조하기 위하여, 원래의 해상도인 $N \times M$ 의 크기를 가지는 영상을 부호화 한다. 그리고 디스플레이 이전 단계에서 복호화 된 좌, 우측 영상 각각을 가로 방향으로 필터링 한 고 1/2 다운 샘플링 하여 $N/2 \times M$ 의 해상도를 가지는 영상으로 변환하여 출력하게 된다. 이 시스템 역시 기존의 인코더 및 디코더를 그대로 사용하는 예로 두 개의 영상을 하나의 영상으로 재구성하였기 때문에, 확장 시점 영상인 좌측 영상이 기준 시점 영상인 우측 영상을 참조하는 것이 가능하다는 점이 이전의 시스템들과 차이점이다. 두 개의 시점이라 하더라도 같은 시간에 비슷한 위치에서 촬영된 두 개의 영상은 공간적인 중복성을 가지게 된다. 이를 이용하여 비디오 코딩을 하게 되면 두 개의 시점 사이의 중복성 제거를 통하여 코딩의 효율을 향상시킬 수 있게 된다. 하지만 두 개의 시점을 하나의 인코더 및 디코더를 사용하여 부호화 및 복호화하기 때문에, 두 시점의 해상도가 일치해야 할 것이다. 이때, 확장 시점 비디오에서 불필요한 데이터를 부호화하게 됨으로써, 부호화 효율이 저하되는 단점을 가지고 있다. 또한, 기준 시점 영상과 확장 시점 영상을 하나의 영상으로 재구성하여 하나의 영상으로 부호화 하고 복호

화 하였기 때문에, 프레임별로 기준 시점과 확장 시점에 대한 정보를 포함하여야 한다. 이는 기존의 2D 디스플레이 장치와의 호환성 문제를 발생시킨다. 또한, 확장 시점 영상이 기준 시점 영상을 참조함으로써, 병렬처리가 불가능하여 두 시점을 부호화 및 복호화하는데 지연이 발생할 수 있다^{[7][12]}.

그림 4는 이중 해상도를 가진 스테레오 비디오를 부호화하는 방법의 블록 다이어그램이다^[14]. 2D 디스플레이 패널에 대한 호환성을 위하여 좌측 영상은 기존의 영상사이즈인 $N \times M$ 의 크기를 유지한 채 단일 시점 인코더를 통해 인코딩 된 후에 디코더부로 전송된다. 반면 우측영상은 좌측 영상에 대하여 1/2 다운 샘플링되어 $N/2 \times M$ 크기로 입력된다. 본 시스템은 우측영상의 공간 중복성 제거를 위하여 현 시점인 우측 영상 뿐 아니라 좌측 영상도 참조하게 되는데, 이때 좌측 영상을 우측영상의 사이즈와 맞추기 위하여 가로 방향으로 1/2 다운샘플링 한 후에 참조하게 된다. 우측 영상의 인코딩 시에 좌측 영상의 재구성 영상을 가로 방향으로 1/2 다운 샘플링 한 후에 우측영상의 참조를 위한 DPB에 저장하고, 우측영상의 코딩 시에 이를 참조하여 코딩하는 방법으로 우측영상의 코딩이 이루어진다. 실제 좌, 우측 영상은 동시에 약간의 시점 변화만을 가진 채로 입력되기 때문에 공간적인 중복성이 뚜렷하게 나타난다. 그림 4의 스테레오 비디오 부호화 시스템은 스테레오 디스플레이 패널에서 쓰이지 않는 우측영상의 정보를 제거한 후에 좌측영상과의 시점간 중복성까지 고려하여 부호화함으로써 효율적으로 스테레오 시점의 영상을 부호화한다. 그러나 기준 시점의 비디오를 참조영상으로 사용하여 시

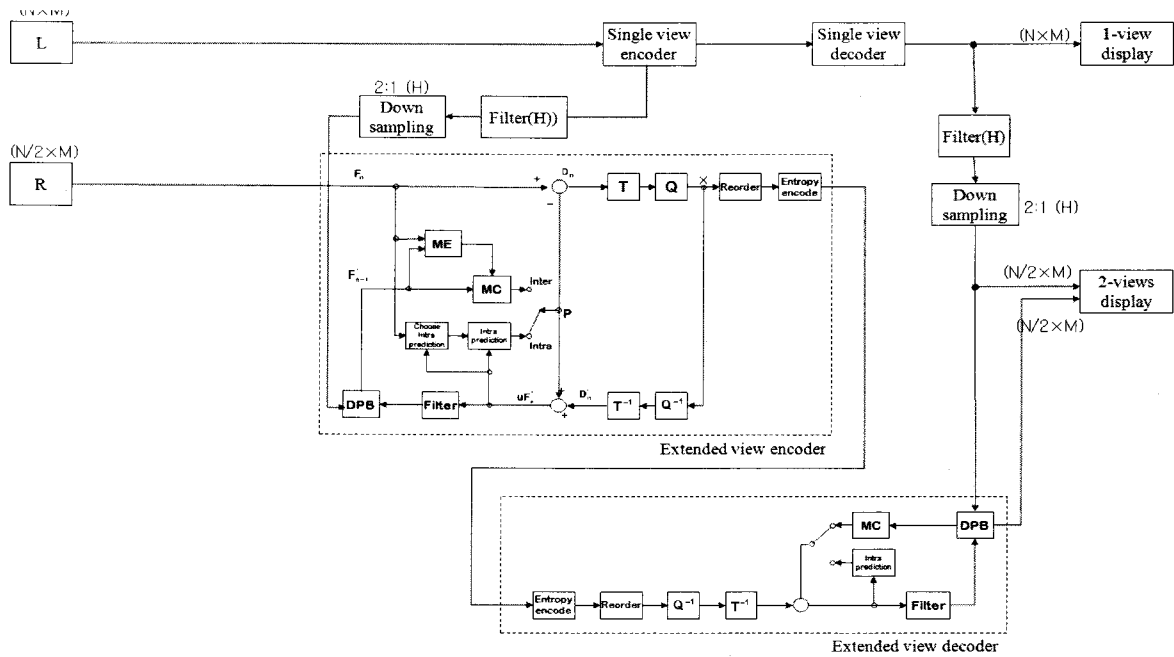


그림 4. 이종 해상도 스테레오 비디오 부호화 시스템
 Fig. 4. Heterogeneous resolution stereo video coding system

점 중복성을 제거함에 있어서, 필터링 및 다운샘플링을 하게 됨으로 인하여 정보의 손실이 발생하는 단점을 가지게 된다.

III. 제안한 이종 해상도 스테레오 비디오 부호화 방법

지금까지 단일 비디오 부호화를 위한 MPEG-2, MPEG-4, 또는 H.264/AVC 코덱을 변형하지 않는 스테레오 디스플레이 시스템과 최근에 제안된 이종 해상도 스테레오 비디오 부호화 방법에 대하여 살펴보았다. 단일 비디오 부호화 방법을 그대로 사용하는 방법들은 공통적으로 기존의 코덱을 변형하지 않고 그대로 사용할 수 있다는 장점을 가진 대신, 스테레오 디스플레이 시스템의 특성을 충분히 고려하지 않았다고 할 수 있다. 최근에 제안되었던, 이종 해상도 비디오 부호화 방법이 동일 해상도 비디오 부호화 방법의 문제점을 어느 정도 해결하였으나, 확장 시점 비디오 부호화 효율이 충분하지는 않은 상태이다^[14].

그림 5는 본 논문에서 제안하는 스테레오 디스플레이 비디오 코덱 시스템의 블록도를 나타낸다. 전체적인 시스템의 구성은 기존의 이종 해상도 스테레오 비디오 부호화 방법과 유사하지만, 우측영상의 인코딩 시에 좌측 영상을 참조하는 방법에 차이가 있다. 기존의 방법에서는 우측영상의 부호화 시에 현 시점의 이전 영상 뿐 아니라 좌측 영상도 참조하기 위하여 좌측 영상을 가로방향으로 1/2 다운샘플링 한 후에 이를 참조하였으나, 제안한 방법에서는 좌측 영상을 다운 샘플링하지 않고 원 해상도 그대로의 영상을 참조하여 부호화한다. 기존의 이종해상도 스테레오 비디오 코딩 방법은 기준 영상을 필터링한 후 다운샘플링 함으로써 영상의 고주파 성분이 제거되는 단점을 가지고 있다. 이렇게 고주파 성분이 제거된 참조 영상을 이용하여, 저해상도의 영상을 예측시 예측 오차를 크게 하는 문제점을 가지고 있다. 이에 반해, 제안한 방법은 기준 영상의 고주파 성분을 제거하지 않아, 저해상도 영상의 압축시 압축 효율이 증대될 수 있다.

움직임 예측 및 시점 간 예측을 수행하기 위해서는 각 픽셀 값들 사이의 오차의 합 (SAD)을 구하고, 이를 최소화

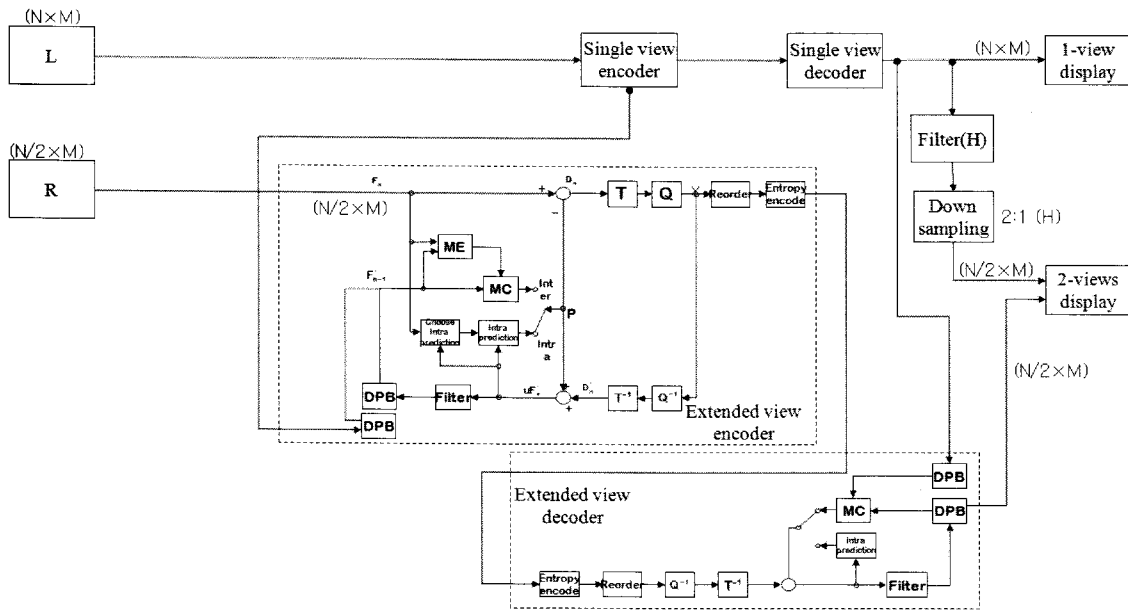


그림 5. 제안한 이종 해상도 스테레오 비디오 부호화 시스템
 Fig. 5. Proposed system for heterogeneous resolution stereo video coding

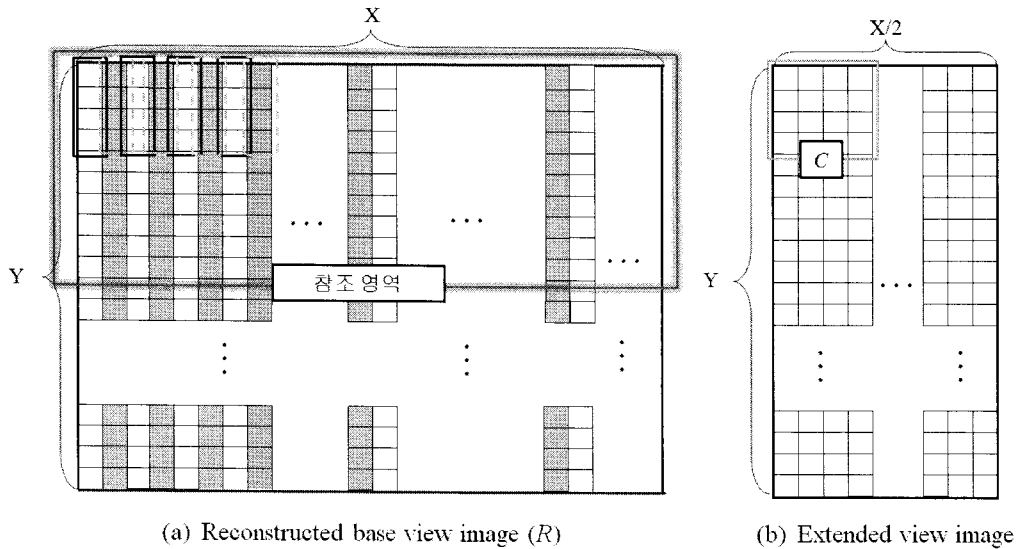


그림 6. 저해상도 확장 영상과 기준 영상으로부터 복원된 참조 영상
 Fig. 6. Low resolution extended-view image and reconstructed base-view image for reference image

하는 예측 블록을 찾게 되는데, 참조 영상과 원 영상간의 해상도가 다른 경우 지금의 방법으로는 올바른 예측이 수행될 수 없다. 이에 본 논문에서는 그림 6과 같은 새로운 시점 간 예측 방법을 제안한다. 그림 6에서 왼쪽 그림은 참

조를 위해 복원된 좌측 영상이고, 오른쪽 그림은 부호화를 위하여 다운샘플링 된 우측 영상이다. 우측 영상에서 현재 시점 간 예측을 하고자 하는 블록이 C이고, Search range로 표시된 영역이 복원된 좌측 영상에서의 참조영역이다. 이

때, 시점 간 예측은 먼저 흰색으로 표시된 픽셀들을 통해 수행되고, 그 이후에 회색으로 표시된 픽셀들을 통해 수행된다. 즉, 그림에서 실선으로 표시된 부분과 C 블록에 대한 오차를 계산하고, 그 후에 점선으로 표시된 부분과 C 블록에 대한 오차를 계산해 나아가는 방향으로 움직임 예측을 진행한다. 이는 참조 영역 전체에 대하여 슬라이딩 방식으로 진행되며, 참조 영역 내에서 가장 작은 SAD를 가지는 영역에 대한 예측 벡터를 구하고, 블록 C에 대한 시점 간 예측을 종료한다. 그림에서 나타낸 방법은 본 논문에서 제안한 방법을 정수 단위의 시점 간 예측으로 간단하게 표현한 것이며, 이를 기존의 H.264/AVC 6tap 필터를 적용하여 1/4 픽셀 시점 간 예측을 할 수 있다.

그림 6은 4x4 블록의 시점 간 예측을 예를 들어 나타낸 것으로, 이를 일반화 된 수식으로 표현하면 다음과 같다.

$$SAD = \sum_{x=0}^M \sum_{y=0}^N |(R(2x + mv_x, y + mv_y) - C(x, y))| \quad (1)$$

식 (1)에서 M과 N은 블록 모드에 따른 예측하고자 하는 블록의 크기를 나타내고, R은 참조영상, C는 현재 코딩하고자 하는 영상의 블록을 나타낸다. 위 식을 통해 계산된 가장 작은 SAD를 가지는 mvx와 mvy의 값이 움직임 벡터 MVx와 MVy값이 되고, 구해진 MVx 와 MVy값으로 아래 수식에 의해 예측 오차를 계산할 수 있다.

$$Residual(x, y) = (R(2x + MV_x, y + MV_y) - C(x, y)) \quad (2)$$

계산된 예측 오차는 기존의 단일 뷰 부호화 방법과 마찬가지로 이산코사인변환 후 양자화, 엔트로피 코딩 과정을 통해 부호화 된다.

제안한 방법은 기존의 방법에 비하여 필터링과 다운샘플링 과정이 없어 계산량이 적은 장점을 가지고 있다. 또한, 제안한 방법의 경우 원영상을 다운샘플링하지 않기 때문에, 저해상도 영상 압축시 참조영상으로 사용될 때, 1/4화소 보간 효과를 위하여 단지 1/2화소 보간만을 사용하면 되는 장점을 가지고 있다.

IV. 실험 결과

본 논문에서는 제안한 스테레오 비디오 부호화 시스템을 기존의 스테레오 부호화 시스템과 실험을 통해 비교하였다. 본 실험은 Pentium(R) D CPU 3.0 GHz, 2.0G RAM을 사용하였고, MVC 레퍼런스 소프트웨어 JMVM 3.0[15]에 제안한 방법을 구현하였다. 또한, 실험영상으로 “Ballroom”, “Exit”, “Akko & Kayo”, “Rena”, “Flamenco2”, “Race1”, “Uli”, “Break dancers” 총 8개의 영상을 사용하였다. 이 영상들은 MVC 표준화를 위하여 사용되는 시퀀스로 5~15개의 시점으로 이루어져 있으며, 각각 다른 영상크기와 프레임임을 가진다. 본 실험에서는 각 시퀀스에서 0번과 1번의 영상만을 이용하여 실험하였으며, 그 밖의 실험 환경은 MVC 표준화를 위한 실험에 준하였다^{[16][17]}.

표 1과 그림 7은 본 논문에서 제안한 방법과 기존의 스테레오 시점 코딩 시스템의 성능을 나타낸 것이다. 본 실험에서는 제안한 방법과 3개의 기존의 방법의 성능을 비교 평가하였다. 다음은 각 방법에 대한 설명이다.

| | |
|--------|---|
| 기존의 방법 | <ul style="list-style-type: none"> - Full & NVP : 그림 1의 시스템으로 좌, 우측 영상을 각각 인코딩한 결과 - Full & VP : 그림 3의 시스템으로 우측 영상을 좌측 영상을 참조하여 인코딩한 결과 - Half & VP : 그림 4의 시스템으로 우측 영상을 1/2 다운샘플링하고 좌측 영상을 다운샘플링 한 후에 참조하여 인코딩한 결과^[14] |
| 제안한 방법 | <ul style="list-style-type: none"> - Half & VP & change ME : 그림 5의 본 논문에서 제안한 스테레오 비디오 부호화 시스템으로 우측영상을 1/2 다운샘플링하고 좌측 영상을 다운샘플링 하지 않고 참조하여 인코딩한 결과 |

표 1은 MVC 표준화에서 사용한 8개의 시퀀스에 대하여 기존의 방법과 제안한 방법의 RD 성능을 보여준다. 본 논문에서 제안한 방법이 기존의 방법에 대하여 높은 코딩 효율을 보이는 것을 알 수 있다. 제안한 방법은 기존의 방법에 비하여 시점 간 중복성을 제거하면서, 참조 영상의 정보를 최대한 이용하여 부호화 효율을 얻을 수 있었다. 기존의 이종해상도 스테레오 비디오 코딩 방법은 다운샘플링 된 우측영상을 부호화하기 위하여 좌측영상을 참조할 시에, 좌

표 1. 기존의 방법 및 제안된 방법에 대한 PSNR과 bitrate
 Table 1. PSNR and bitrate of the proposed and conventional algorithms

| | Tool | PSNR | Bitrate | PSNR | Bitrate | PSNR | Bitrate |
|---------------|-----------------------|-------|---------|-------|---------|-------|---------|
| Ballroom | Full & NVP | 35.98 | 679.24 | 35.03 | 547.51 | 33.47 | 386.08 |
| | Full & VP | 36.04 | 631.35 | 35.07 | 503.18 | 33.41 | 347.96 |
| | Half & VP | 35.51 | 544.12 | 34.49 | 435.12 | 32.77 | 301.97 |
| | Half & VP & change ME | 35.90 | 539.12 | 35.01 | 432.12 | 32.97 | 295.47 |
| Rena | Full & NVP | 44.15 | 669.50 | 41.33 | 358.86 | 38.38 | 196.55 |
| | Full & VP | 44.12 | 620.32 | 41.22 | 322.13 | 38.22 | 174.23 |
| | Half & P | 43.32 | 543.07 | 40.36 | 282.96 | 37.38 | 150.83 |
| | Half & VP & change ME | 43.63 | 533.07 | 40.70 | 277.96 | 37.73 | 150.83 |
| Flamenco2 | Full & NVP | 38.79 | 675.30 | 37.66 | 532.91 | 35.36 | 330.10 |
| | Full & VP | 38.81 | 667.34 | 37.67 | 525.64 | 35.37 | 323.43 |
| | Half & VP | 37.99 | 577.10 | 36.81 | 451.61 | 34.49 | 275.11 |
| | Half & VP & change ME | 38.40 | 572.10 | 36.89 | 446.61 | 34.63 | 271.11 |
| Exit | Full & NVP | 38.94 | 415.48 | 38.37 | 316.05 | 37.39 | 221.33 |
| | Full & VP | 38.93 | 406.58 | 38.34 | 304.62 | 37.31 | 211.01 |
| | Half & VP | 38.60 | 339.56 | 37.96 | 255.19 | 36.85 | 177.96 |
| | Half & VP & change ME | 38.67 | 339.06 | 38.16 | 254.19 | 37.16 | 177.96 |
| Race | Full & NVP | 39.72 | 1209.08 | 38.60 | 946.28 | 37.51 | 740.74 |
| | Full & VP | 39.74 | 1150.96 | 38.63 | 890.68 | 37.55 | 694.57 |
| | Half & VP | 39.45 | 920.33 | 38.30 | 711.76 | 37.19 | 555.00 |
| | Half & VP & change ME | 39.47 | 916.83 | 38.32 | 707.76 | 37.21 | 552.50 |
| Break dancers | Full & NVP | 39.74 | 1005.22 | 38.55 | 475.07 | 36.98 | 261.96 |
| | Full & VP | 39.74 | 991.13 | 38.53 | 462.40 | 36.88 | 250.67 |
| | Half & VP | 39.38 | 828.87 | 38.09 | 385.56 | 36.39 | 208.72 |
| | Half & VP & change ME | 39.38 | 825.37 | 38.09 | 382.06 | 36.39 | 206.72 |
| Akko & Kayo | Full & NVP | 41.06 | 992.00 | 38.23 | 561.80 | 34.04 | 280.74 |
| | Full & VP | 41.09 | 908.50 | 38.21 | 502.77 | 33.94 | 247.54 |
| | Half & VP | 40.39 | 796.05 | 37.43 | 440.28 | 33.12 | 212.26 |
| | Half & VP & change ME | 40.61 | 791.05 | 37.59 | 437.28 | 33.37 | 211.26 |
| Uli | Full & NVP | 36.12 | 2599.27 | 35.09 | 2071.90 | 31.76 | 1019.15 |
| | Full & VP | 36.12 | 2597.86 | 35.09 | 2070.75 | 31.74 | 1017.56 |
| | Half & VP | 35.46 | 2195.97 | 34.36 | 1737.24 | 30.94 | 836.19 |
| | Half & VP & change ME | 35.54 | 2192.97 | 34.51 | 1733.74 | 31.00 | 834.69 |

측 영상 역시 다운샘플링을 하여 두 영상의 해상도를 같게 한 후에 참조를 수행하였다. 기존의 2D 비디오 장치와의 호환성을 위하여 좌측 영상은 원래 해상도의 영상으로 복호화 한 것을 우측 영상과의 해상도를 맞추기 위하여 다운샘플링 한 것인데, 이는 다운샘플링을 통해 정보를 버리는

것으로 우측 영상을 다운샘플링 하여 정보의 크기가 줄었음에도 불구하고 코딩의 효율이 떨어지는 원인이 된다. 제안한 방법은 이를 보완하여 시점간 참조 시 좌측 영상의 정보를 최대한 활용하여 시점간 예측을 보다 정확하게 수행함으로써 코딩의 효율을 높일 수 있었다. 실험결과 평균

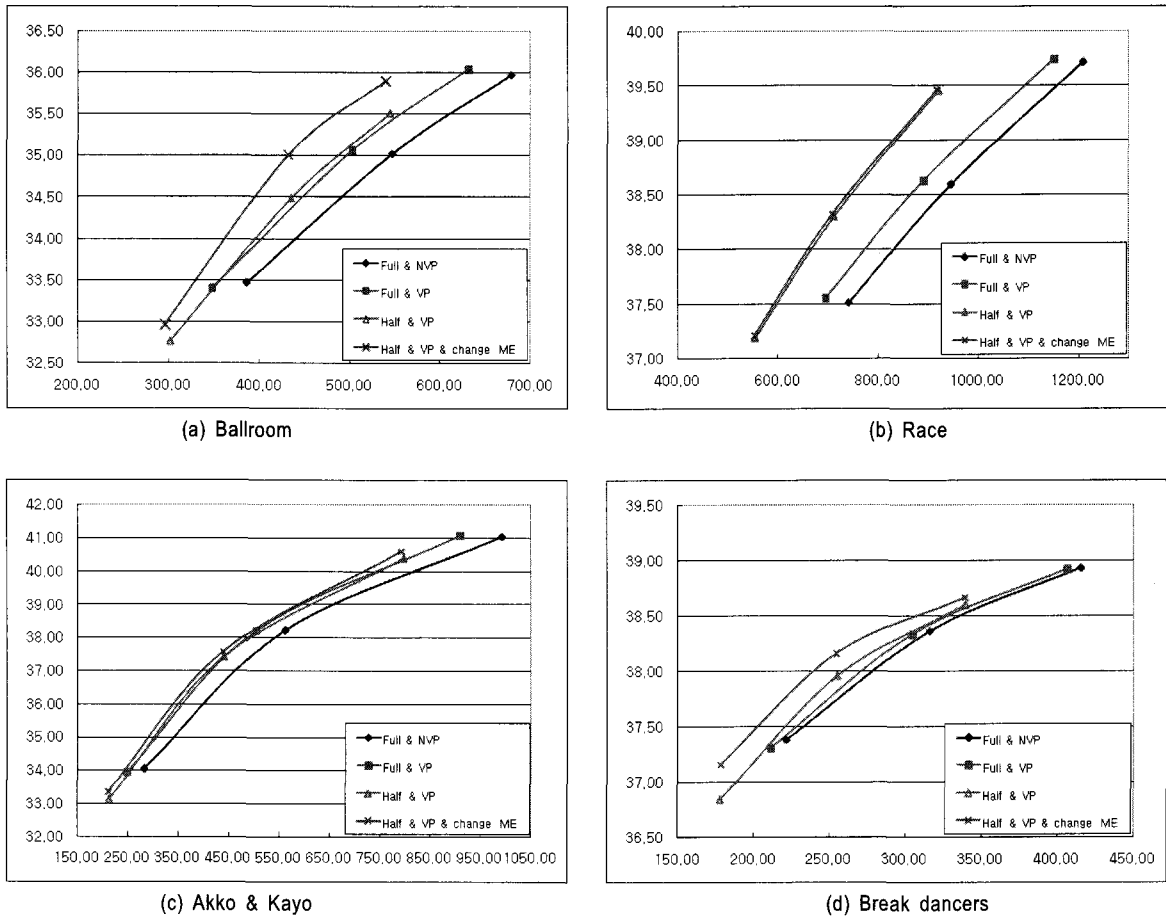


그림 7. 기존의 방법 및 제안된 방법에 대한 RD 성능비교
 Fig. 7. Rate distortion performance comparison of conventional and proposed algorithms.

적으로 0.5~0.8 dB의 코딩 이득을 얻었으며, 특정 영상에서 독립 부호화 방법에 비하여 1.2 dB 이상의 코딩이득을 얻을 수 있었다. 또한, 기존의 이중 해상도 스테레오 부호화 방법에 비하여서도 0.1~0.5 dB 정도의 부호화 이득을 얻을 수 있었다. 스테레오 비디오에 따라 제안한 방법이 상대적으로 높은 코딩 효율을 보이는 경우가 있다. 확장 영상의 경우 기준 영상에 비하여 다운 샘플링을 함에도 불구하고, 고주파 성분이 제거되지 않는 경우, 다운 샘플링과 기준 영상 참조에도 불구하고 부호화 효율이 크지 않은 경우도 있었다.

그림 7은 표 1의 실험결과 중 특색 있는 몇 개의 실험 결과를 RD 그래프로 나타낸 것이다. "Ballroom"과 "Break

dancers" 시퀀스에 대하여 제안한 알고리즘이 비교적 좋은 성능을 보이고 있다. 본 시퀀스의 경우 에지의 특징 등이 강하고, 뷰 간의 상관성이 높은 것을 특징으로 하여 제안한 알고리즘이 예측방법이 더 많이 선택이 되었다.

그림 8은 기존의 방법과 제안한 방법에 의하여 복원된 영상의 예를 보여주고 있다. 본 영상은 "Ballroom" 시퀀스의 7번째와 125번째 원영상을 다운샘플링 한 것과 이것을 기존의 방법과 제안한 방법으로 복원한 영상을 보여준다. 여기서 부호화는 약 300kbps로 하였으며, 제안한 방법이 기존의 방법에 비하여 각각 0.5dB와 0.3dB 향상되었다. 실제 화질 평가에서 고주파 에지 성분 지역 등에서 제안한 방법이 보다 자연스러움을 볼 수 있었다.

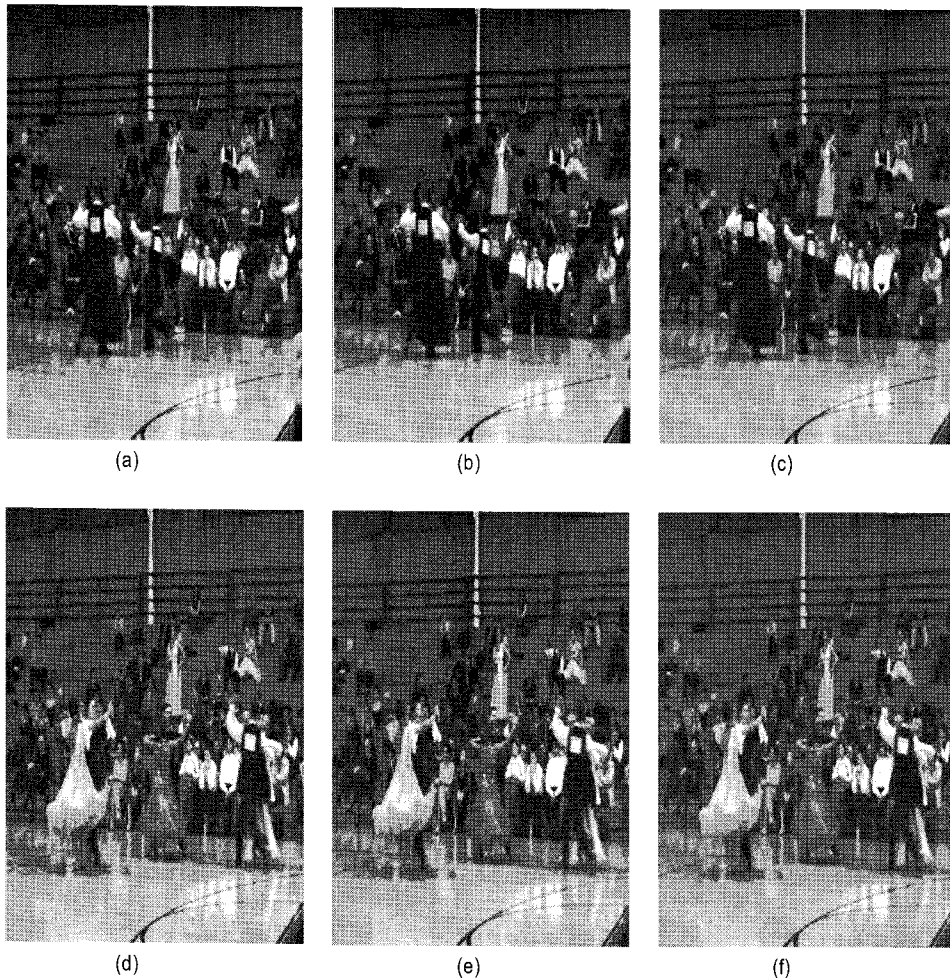


그림 8. 다운샘플된 원영상과 기존의 방법과 제안한 방법에 의하여 복호화된 영상 (a) Ballroom 시퀀스의 한 원본영상을 다운 샘플링 한 영상 (7번째 영상) (b) Half & VP 방법에 따른 복원 영상 (bitrate: 300kbps) (c) Half & VP & change ME 방법에 따른 복원 영상 (bitrate : 300kbps), (d) 125번째 원본 영상을 다운샘플링 한 영상 (e) Half & VP 방법에 따른 복원 영상 (bitrate : 300kbps) (f) Half & VP & change ME 방법에 따른 복원 영상 (bitrate : 300kbps)
 Fig 8. Down-sampled original image and reconstructed images from the conventional and proposed algorithms (a) Downsampled original image from "Ballroom" sequence (7th image) (b) Reconstructed image by the "Half & VP" method (bitrate: 300kbps) (c) Reconstructed image by the proposed method (bitrate: 300kbps), (d) Downsampled original image from "Ballroom" sequence (125th image), (e) Reconstructed image by the "Half & VP" method (bitrate: 300kbps) (f) Reconstructed image by the proposed method (bitrate : 300kbps)

V. 결론

본 논문에서는 스테레오 디스플레이 장치 및 스테레오 영상의 특성을 이용한, 역 방향 호환 스테레오 비디오 코딩 방법을 제안하였다. 제안한 알고리즘은 기존의 방법과 달리 확장 시점 비디오의 해상도를 기준 시점 비디오 해상도의 절반으로 하여 불필요한 신호의 부호화를 최대한 줄이

고, 확장 비디오 부호화시 기준 시점 비디오를 참조하는 방법으로 시점 간 중복성을 효과적으로 제거하였다. 확장 비디오 부호화시에 영상의 해상도가 다른 기준 비디오를 참조함으로써 발생하는 문제를 해결하기 위한 방법으로, 기준 비디오를 다운 샘플링하여 영상의 해상도를 맞추는 기존의 방법 대신, 원 영상의 해상도를 유지한 채로 시점간 예측 방법을 변경하여 참조하였다. 그 결과 제안된 방법이 기존

의 이종 해상도 스테레오 비디오 부호화 방법에 비하여 높은 코딩 효율을 보였다. 제안한 방법은 MVC 레퍼런스 소프트웨어 JMVM 3.0에 구현되었으며, MVC 표준화에서 테스트 영상으로 사용하는 8개의 영상에 대하여 실험되었다. 그 결과, 좌, 우측의 영상을 각각 부호화하는 방법에 대해서는 0.5~0.8dB 이상의 코딩이득을 얻었으며, 참조 영상 코딩 시에 기준 영상을 다운 샘플링하여 참조하는 방법에 대해서는 0.1~0.5 dB의 코딩 이득을 얻을 수 있음을 확인하였다.

본 논문에서 제안한 이종 해상도 스테레오 비디오 코딩 방법은 다시점 비디오 코딩에 확장 가능할 것이고, 이를 통해 효율적인 다시점 비디오의 부호화가 가능할 것이다.

참 고 문 헌

- [1] L. McMillan and G. Bishop, "Head-tracked stereoscopic display using image warping," SPIE Proc. 2409, pp. 21-30, Feb. 1995.
- [2] G. Perkins, "Data compression of stereo pairs," IEEE Trans. on Communications, vol. 40, no. 4, pp. 684-696, April 1992.
- [3] ISO/IEC JTC1/SC29/WG11, "Description of exploration experiments in 3DAV," N6194, Dec. 2003.
- [4] Li-Fu Ding, Shao-Yi Chien, Yu-Wen Huang, Yu-Lin Chang, and Liang-Gee Chen, "Stereo video coding system with hybrid coding based on joint prediction scheme," IEEE International Symposium on Volume Issue, vol. 6, pp. 6082-6085, May 2005.
- [5] F. Isgro, E. Trucco, P. Kauff, and O. Schreer, "Three-dimensional image processing in the future of immersive media," IEEE Trans. on Circuits and Systems for Video Technology, vol. 14, no. 3, pp. 388-303, Mar. 2003.
- [6] H.-Z. Jia, W. Gao, and Y. Lu, "Stereoscopic video coding based on global displacement compensated prediction," in International Conference on Information and Communications Security, vol. 1, pp. 61-65. Dec. 2003.
- [7] S. Pastoor, "3D-television: a survey of recent research results on subjective requirements," Signal Processing: Image Communication, vol. 4, no. 1, pp. 21 - 32, 1991.
- [8] H. Aydinoglu and M.H. Hayes III, "Stereo image coding," ISCAS, vol. 1, pp. 247-250, May 1995.
- [9] O. Faugeras, "Three-dimensional computer vision: A geometrical viewpoint," MIT Press, 1993.
- [10] Gary J. Sullivan, Pankaj Topiwala, and Ajay Luthra, "The H.264/AVC advanced video coding standard : overview and introduction to the fidelity range extension," applications of digital image processing XXVII, vol. 5558, pp. 454-474, Nov. 2004
- [11] R.-S. Wang and Y. Wang, "Multiview video sequence analysis, compression, and virtual viewpoint synthesis," IEEE Trans. on Circuits and Systems for Video Technology, vol. 10, no. 3, pp. 397-410, Apr. 2000.
- [12] ISO/IEC JTC1/SC29/WG11, "Report for EE3 in MPEG 3DAV," M9186, Dec. 2002.
- [13] ISO/IEC JTC1/SC29/WG11, "Requirements on multi-view video coding," N6501, Jul. 2004.
- [14] Joint Video Team (JVT) ISO/IEC MPEG & ITU-T VCEG, "Inter-view prediction with downsampled reference pictures," JVT-W079, Apr. 2007.
- [15] JVT MVC Reference Software version JMVM 3.0.
- [16] Joint Video Team (JVT) ISO/IEC MPEG & ITU-T VCEG, "Joint multi-view video model (JMVM) 4," JVT-W207-M, Apr. 2007.
- [17] Joint Video Team (JVT) ISO/IEC MPEG & ITU-T VCEG, "Joint draft 3 Multi-view video coding," JVT-V209-M, Jan. 2007.

저 자 소 개



박 시 내

- 2004년 2월 : 광운대학교 컴퓨터공학과 공학사
- 2006년 2월 : 광운대학교 컴퓨터공학과 공학석사
- 2006년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- 주관심분야 : MVC, FTV, H.264/AVC, 영상압축

저 자 소 개



심 동 규

- 1993년 2월 : 서강대학교 전자공학과 공학사
- 1995년 2월 : 서강대학교 전자공학과 공학석사
- 1999년 2월 : 서강대학교 전자공학과 공학박사
- 2000년 9월 : 현대전자, 기반기술 연구소, 선임연구원
- 2002년 3월 : 바로비전, 뉴미디어 연구소, 선임연구원
- 2005년 2월 : Univ. Washington, Senior research engineer
- 2005년 3월 : ~ 현재 : 광운대학교 전자정보대학, 컴퓨터공학과 조교수
- 주관심분야 : 영상처리, 컴퓨터비전, 멀티미디어