일반논문-09-14-6-04

실감형 다시점 스케일러블 비디오 코딩 방법의 설계 및 구현

박 민 우[a)], 박 광 훈[a)‡]

Design and Implementation of a Realistic Multi-View Scalable Video Coding Scheme

Min Woo Park[a)], and Gwang Hoon Park[a)‡]

요 약

본 논문에서는 3D 컨텐츠 서비스에 대한 사용자의 욕구를 만족시킴과 동시에 미래 컴퓨팅 환경에 적합한 새로운 동영상 코딩 방법으로서 실감형 다시점 스케일러블 비디오 코딩 방법을 제안하였다. 미래의 비디오 코딩 방법은 스테레오 스코픽 또는 다시점 비디오를 통하여 삼차원 실감형 입체영상을 사용자로 하여금 느끼게 하는 실감형 서비스를 지원함과 동시에 다양한 통신환경 및 다양한 종류의 단말을 통합적으로 지원하기 위한 'One-source Multi-use'를 달성할 수 있어야 한다. 지금까지 2차원 디스플레이만을 지원하는 동영상 코딩 방법과는 다르게 본 논문에서 제안하는 실감형 다시점 스케일러블 비디오 코딩 방법은 그러한 실감 서비스를 지원할 수 있는 방법이다. 제안된 코딩 방법은 다시점 비디오 코딩 방법과 스케일러블 비디오 코딩 방법의 기능을 통합하는 방향으로 설계되고 구현된 후 성능 평가를 통해 실제 3D 서비스에서의 응용 가능성을 살펴보았다. 성능 평가를 통해 본 논문에서 제안하는 코딩 구조가 코딩 효율을 효율적으로 유지하면서 시점간의 랜덤 액세스 성능을 크게 높여주는 것을 확인할 수 있었다.

Abstract

This paper proposes a realistic multi-view scalable video coding scheme designed for user's interest in 3D content services and the usage in the future computing environment. Future video coding schemes should support realistic services that make users feel the 3-D presence through stereoscopic or multi-view videos, as well as to accomplish the so-called one-source multi-use services in order to comprehensively support diverse transmission environments and terminals. Unlike the most of video coding methods which only support two-dimensional display, the proposed coding scheme in this paper is the method which can support such realistic services. This paper designs and also implements the proposed coding scheme through integrating Multi-view Video Coding scheme and Scalable Video Coding scheme, then shows its possibility of realization of 3D services by the simulation. The simulation results show the proposed structure remarkably improves the performance of random access with almost the same coding efficiency.

Keyword : Video compression, Multi-view Video Coding (MVC), Scalable Video Coding (SVC), realistic broadcasting

---

a) 경희대학교 전자정보대학 컴퓨터공학과 미디어랩
Media Lab., Dept. of Computer Engineering, College of Electronics and Information, Kyung Hee University
‡ 교신저자 : 박광훈 (ghpark@khu.ac.kr)
※ 본 논문은 교육과학기술부 및 한국연구재단의 국가지정연구실 지원 사업 (R0A-2005-000-10061-0)과 지식경제부 및 정보통신연구진흥원의 대학 IT 연구센터 지원 사업 (IITA-2009-C1090-0903-0011)의 연구 결과로 수행되었음.
· 접수일(2009년7월24일),수정일(2009년10월13일),게재확정일(2009년10월28일)

Ⅰ. 서 론

디지털 멀티미디어 기술의 발전과 함께 동영상 코딩 기술은 지난 20년간 급속한 발전을 이루었으며 각 동영상 코딩 기술이 통신, 방송, 저장 매체 등 다양한 응용(Application)에 적용되어 고해상도, 고화질의 동영상 정보를 사용자에게 서비스할 수 있게 되었다.

동영상 코딩 기술은 주로 ISO/IEC의 MPEG(Moving Picture Experts Group)과 ITU-T의 VCEG(Video Coding Experts Group)에서 지속적으로 연구 발전되어 왔다. 1990년 저비트율(Low-rate)을 갖는 전송환경에서 실시간 동영상 통신을 위한 ITU-T의 H.261 표준이 완성된 후 1.5Mbps의 전송률을 갖는 저장매체를 위한 ISO/IEC의 MPEG-1 표준(1993년), 수 Mbps에서 수십 Mbps까지의 전송률을 갖는 방송과 저장 매체를 위한 ISO/IEC 와 ITU-T의 MPEG-2/H.262 표준(1994년), 저비트율을 갖는 실시간 통신 환경에서 H.261보다 높은 코딩 효율을 갖는 ITU-T의 H.263 표준(1995년), 객체 기반 코딩과 무선망 같은 에러가 많이 발생하는 환경에서의 효율적인 동영상 데이터 전달을 목표로 하는 ISO/IEC의 MPEG-4 표준(1999년)이 완성되었다[1]-[5]. 그 후에 ISO/IEC의 MPEG과 ITU-T의 VCEG이 함께 방송, 통신 융합의 흐름 속에 공동 표준의 필요성을 인식하여 JVT(Joint Video Team)을 구성하여 2003년에 방송, 통신 및 저장매체에 모두 적용이 가능하며 높은 코딩 효율을 갖는 H.264(MPEG-4 Part 10 Advanced Video Coding) 표준을 완성하였다[6][7].

동영상 코딩 기술의 발전을 통해 사용자가 고화질, 고해상도의 동영상 정보를 접할 수 있게 되었지만 이들은 단지 2차원 평면 디스플레이 장치를 지원하기 위한 방법일 뿐이며 사용자에게 입체감을 주거나 시점(View)을 자유롭게 선택할 수 있도록 하는 실감형 서비스(Realistic service)를 고려하지 않았다. 그리고 앞으로는 방송/통신의 융합 및 유/무선의 융합 등의 과정을 통해 다양한 전송 환경 속에서 다양한 종류의 단말들이 혼재하는 유비쿼터스(Ubiquitous) 환경에서 동영상 정보를 효율적으로 전달할 수 있어야 한다. 하지만 이들 동영상 코딩 방법은 각각의 응용에서 특정한 전송 환경과 특정한 단말을 대상으로 한정된 코딩을 수행하도록 설계가 되어 있기 때문에 효율적인 처리가 불가능하다. 만약 기존의 동영상 코딩 방법을 이용하여 하나의 동영상 정보를 서비스한다고 가정하면 다양한 전송환경과 단말들을 지원하기 위해 반복적으로 인코딩 과정을 수행해야 하는 복잡하고 어려운 작업이 될 것임을 짐작할 수 있다.

이러한 환경의 변화 속에서 앞으로의 동영상 코딩은 스테레오스코픽(Stereoscopic) 영상 또는 다시점(Multi-view) 영상을 통해 사용자가 입체감을 느낄 수 있도록 하는 실감형 서비스를 지원할 수 있어야 함은 물론 다양한 전송 환경과 단말들을 통합적으로 지원하기 위한 동영상 정보를 생성하고 효율적으로 전달, 상호 교환할 수 있는 방법이 함께 실현되어야 한다.

실제로 이러한 동영상 코딩 기술의 발전 방향과 관련된 국제 표준으로 다시점 비디오 코딩(Multi-view Video Coding; MVC)과 스케일러블 비디오 코딩(Scalable Video Coding; SVC) 방법을 예로 들 수 있다.

다시점 비디오 코딩 방법은 ISO/IEC의 MPEG에서 2001년부터 관련된 연구를 지속적으로 수행하였고[8]-[11], 2006년 7월부터 본격적으로 JVT에서 H.264/AVC 동영상 코딩 표준을 기반으로 한 표준화를 진행하였으며 2008년 7월 표준화를 완료 하였다[12]-[14]. 다시점 비디오 코딩 방법은 다양한 배열 안에서 일정한 간격으로 떨어져 있는 복수의 카메라로부터 입력 받은 복수 개의 시점을 효율적으로 코딩하는 방법으로써 3차원 TV(3DTV)나 FTV(Free view-point TV) 등의 실감형 디스플레이를 지원한다.

스케일러블 비디오 코딩 방법은 ISO/IEC의 MPEG에서 2004년 3월에 표준화를 시작하였고 H.264/AVC 기술을 기반으로 표준화를 수행하기로 결정된 후 2005년 1월부터 JVT에서 표준화를 진행하였으며 2007년 7월 표준화를 완료하였다. 스케일러블 비디오 코딩 방법은 다양한 종류의 단말들과 다양한 전송 환경에서 동영상 정보를 통합적으로 취급하기 위한 기술로써 다양한 공간적 해상도, 프레임율, 화질을 지원 가능한 하나의 통합된 데이터를 생성하여 다양한 전송 환경과 다양한 단말들에게 데이터를 효율적으로 전송할 수 있도록 지원하는 방법이다[6][15]-[18].

실감형 디스플레이 장치들을 지원하기 위해 다시점 영상을 효율적으로 코딩하는 방법은 다시점 비디오 코딩 방법

과 다양한 전송 환경과 다양한 종류의 단말들을 통합적으로 지원할 수 있는 스케일러블 비디오 코딩 방법은 존재하지만, 현재 유비쿼터스 환경에서 실감형 서비스를 지원하면서 동시에 다양한 환경을 지원할 수 있는 통합된 동영상 정보를 생성하는 방법은 존재하지 않는다. 현재 실감형 콘텐트에 대한 사용자의 관심은 영화 산업을 중심으로 급격하게 높아지고 있으며 사용자의 욕구 또한 증가되어 현재 개인용 스테레오스코픽 디스플레이 장치나 다시점 영상 디스플레이 장치와 같은 실감형 디스플레이 장치들이 다양한 플랫폼에서 개발되어 상용화되거나 최우선적으로 추진되고 있다. 이와 같은 추세라면 머지않아 실감형 동영상 콘텐트를 다양한 전송 환경과 다양한 단말들에게 효율적으로 전달할 수 있는 방법이 반드시 필요하게 될 것이다. 따라서 본 논문에서는 유비쿼터스 환경에서 실감형 단말들과 기존 2차원 단말들에게 동영상 정보를 통합적으로 처리하여 손쉽게 전송할 수 있는 방법으로 실감형 다시점 스케일러블 비디오 코딩(Realistic Multi-view Scalable Video Coding; RMSVC) 방법을 제안한다. 먼저 실감형 다시점 스케일러블 비디오 코딩의 개념과 동작을 살펴보고, 스케일러블 비디오 코딩과 다시점 비디오 코딩의 기능을 통합 구현한 후 성능 평가를 통해 실제 사용에 있어서의 가능성을 살펴보며, 실감형 다시점 스케일러블 비디오 코딩에서 시점의 랜덤 액세스의 성능을 향상시키는 계층별 예측 구조를 제안한다.

본 논문의 구성은 다음과 같다. Ⅱ장에서는 실감형 다시점 스케일러블 비디오 코딩의 개념과 동작에 대해 살펴본다. Ⅲ장에서는 실제 실감형 다시점 스케일러블 비디오 코딩의 구조를 알아본 뒤, Ⅳ장에서는 실감형 다시점 스케일러블 비디오 코딩의 계층별 예측 구조를 제안한다. Ⅴ장에서는 제안하는 실감형 다시점 스케일러블 비디오 코딩 구조와 계층별 예측 구조의 성능을 평가하고, Ⅵ장에서는 결론을 기술한다.

Ⅱ. 실감형 다시점 스케일러블 비디오 코딩의 개념 및 동작

실감형 다시점 스케일러블 비디오 코딩은 그림 1과 같이 다양한 단말과 다양한 전송환경을 지원하고 또한 실감형 서비스를 동시에 지원할 수 있도록 하는 방법이다. 다양한 단말, 다양한 전송환경, 실감형 서비스를 지원하기 위해서는 다양한 시점, 다양한 화면 크기, 다양한 화질, 다양한 시간적 해상도를 함께 지원할 수 있어야 한다.

그림 1. 실감형 다시점 스케일러블 비디오 코딩의 개념.
Fig. 1. Concept of Realistic Multi-view Scalable Video Coding.

기존의 2차원 형태의 디스플레이 장치와 양안 시차에 의해 입체를 경험할 수 있도록 하는 스테레오스코픽 디스플레이 장치, 여러 개의 시점을 동시에 출력함으로써 다양한 각도에서 각 시점을 2차원 혹은 스테레오스코픽 영상을 감상할 수 있도록 하는 다시점 디스플레이 장치, 다양한 시점 중에 사용자가 원하는 시점을 선택하여 감상할 수 있는 시점 선택형 디스플레이 장치, 3차원 영상의 발전된 형태의 디스플레이 장치에서 시점의 합성 과정을 통해 자유로운 시점 탐색을 가능하게 하는 FTV 등 다양한 단말들을 통합적으로 지원하기 위해서는 기본적으로 복수개의 시점을 효율적으로 전송할 수 있어야 하며 또한 각 단말에서 시점을 다양한 조합으로 자유롭게 선택할 수 있어야 한다.

그리고 이러한 다양한 형태의 디스플레이 장치를 포함하는 단말기에서는 사용의 용도와 성능에 따라서 QVGA 크기의 해상도부터 Full HD(High definition) 해상도 혹은 그 이상의 해상도까지 다양한 화면 크기를 지원할 수 있어야 하며, 모바일 환경과 같은 환경에서의 5Hz 정도의 낮은 시간적 해상도부터 Full HD급을 디스플레이할 수 있는 환경에서의 60Hz 정도의 높은 시간적 해상도 혹은 그 이상의 시간적 해상도까지 지원할 수 있어야 한다. 또한 다양한 시점과 다양한 해상도, 다양한 시간적 해상도를 갖는 영상 데이터는 인터넷, 방송망, 저장매체 등 다양한 전송환경 등에

서 각각의 대역폭에 적합한 화질을 유연하게 유지할 수 있어야 한다.

앞서 설명한 개념에 따라 실감형 다시점 스케일러블 비디오 코딩의 동작 구조를 그림 2와 같이 표현할 수 있다. HD급 해상도의 다시점 영상을 실감형 다시점 스케일러블 비디오 코딩을 수행한다고 가정하였을 때, 한 번의 인코딩 과정을 통해 다양한 시점, 다양한 화면 크기, 다양한 화질, 다양한 시간적 해상도를 포함하는 하나의 RMSVC 비트스트림을 생성한 후, RMSVC 비트스트림에서 다양한 종류의 단말과 각각의 전송 환경에 적합한 비트스트림 데이터를 추출하여 각 단말에 전송하고 각 단말에서 전송 받은 데이터를 디코딩하여 해당하는 영상을 출력할 수 있다.

그림 2의 실감형 다시점 스케일러블 비디오 코딩의 동작 구조에서 몇 가지 응용의 예를 몇 가지로 정리할 수 있다. 먼저 HD급 화질의 다시점 디스플레이를 통해 영상을 보기 원할 경우에는 인코딩된 RMSVC 비트스트림을 그대로 전송하고 RMSVC 또는 MVC 디코더를 통해 디코딩을 수행하여 영상을 감상할 수 있다. 그리고 HD급 화질의 스테레오스코픽 디스플레이를 통해 영상을 보기 원할 경우에는 인코딩된 RMSVC 비트스트림에서 해당 비트스트림을 추출하여 전송하고 RMSVC 또는 MVC 디코더를 통해 디코딩을 수행하여 영상을 감상할 수 있다. 그리고 HD급 또는 SD(Standard definition)급 화질의 2차원 디스플레이에서 영상을 보기 원할 경우에는 인코딩된 RMSVC 비트스트림에서 해당 비트스트림을 추출하여 전송하고 SVC 또는 H.264/AVC 디코더를 통해 디코딩을 수행하여 영상을 감상할 수 있다. 그리고 스테레오스코픽 디스플레이를 지원하는 모바일 단말을 통해서 영상을 보기 원할 경우에는 인코딩된 RMSVC 비트스트림에서 해당 비트스트림을 추출하여 전송하고 MVC 디코더를 통해 디코딩을 수행하여 영상을 감상할 수 있다. 마지막으로 일반적인 2차원 디스플레이를 지원하는 모바일 단말을 통해서 영상을 보기 원할 경우에는 인코딩된 RMSVC 비트스트림에서 단일 시점의 가장 기본적인 화질에 해당하는 비트스트림을 추출하여 전송하고 H.264/AVC 디코더를 통해 디코딩을 수행하여 영상을 감상할 수 있다.

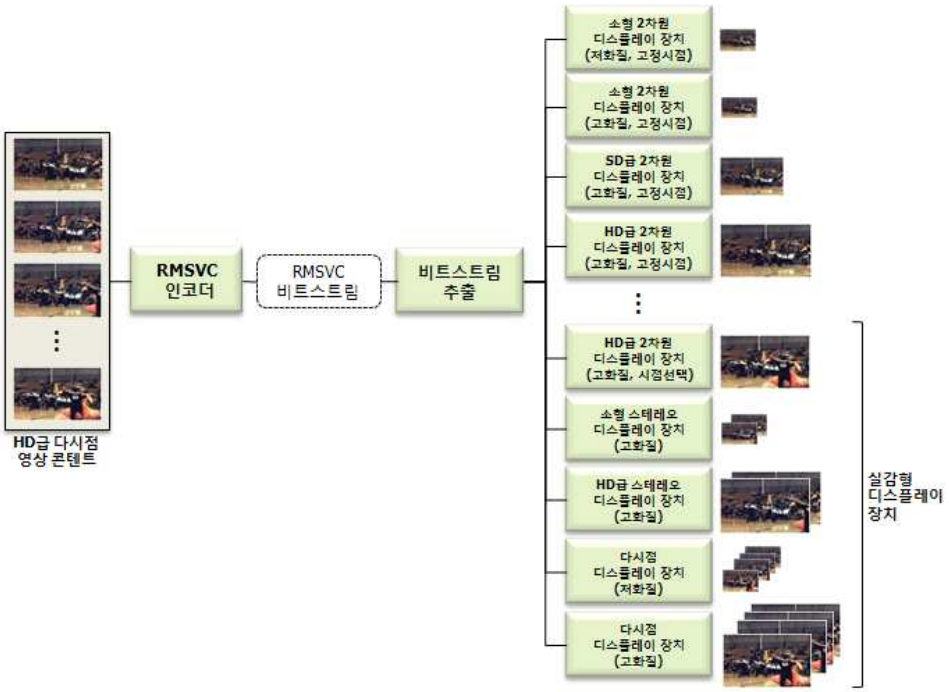


그림 2. 실감형 다시점 스케일러블 비디오 코딩(RMSVC)의 동작 구조의 예
Fig. 2. Example of operating structure of Realistic Multi-view Scalable Video Coding (RMSVC)

Ⅲ. 실감형 다시점 스케일러블 비디오 코딩 구조

본 장에서는 실감형 다시점 스케일러블 비디오 코딩을 실현하기 위해 먼저 다시점 비디오 코딩과 스케일러블 비디오 코딩 구조를 결합한 형태의 실감형 다시점 스케일러블 비디오 코딩 구조(Realistic Multi-view Scalable Video Coding; RMSVC)를 예를 들어 설명한다. 본 논문에서 제안하는 실감형 다시점 스케일러블 비디오 코딩은 다수의 시점을 효율적으로 코딩하여 스테레오스코픽 영상이나 복수의 시점 영상을 통해 사용자에게 실감을 주거나 시점을 자유롭게 선택할 수 있는 장치를 지원하는 기존의 다시점 비디오 코딩 방법에 스케일러블 비디오 코딩 방법에서 지원하는 화질적, 시간적, 공간적 스케일러빌리티(Scalability) 기능을 추가한 형태의 코딩 구조이다.

그림 3은 4개의 시점과 각 시점마다 2개의 공간적 스케일러블 계층을 갖는 실감형 다시점 스케일러블 비디오 코딩 구조의 예를 보여준다. 각 시점에서 입력받은 영상을 기본 계층의 해상도에 맞게 다운샘플링(Down-sampling)을 수행한 후 각각의 영상을 각 공간적 해상도 계층의 인코더로 입력한다. 모든 시점에서 공간적 기본 계층(계층 0)에서는 기존의 다시점 비디오 코딩의 방법과 같이 시점 내(Intra-view) 예측 코딩과 더불어서 시점 간(Inter-view) 예측 코딩을 포함하는 코딩을 수행한다. 이 때 시간적 스케일러빌리티를 지원하기 위해서 계층적인 B-픽쳐 구조(Hierarchical B-pictures structure)와 같은 시간 방향 예측

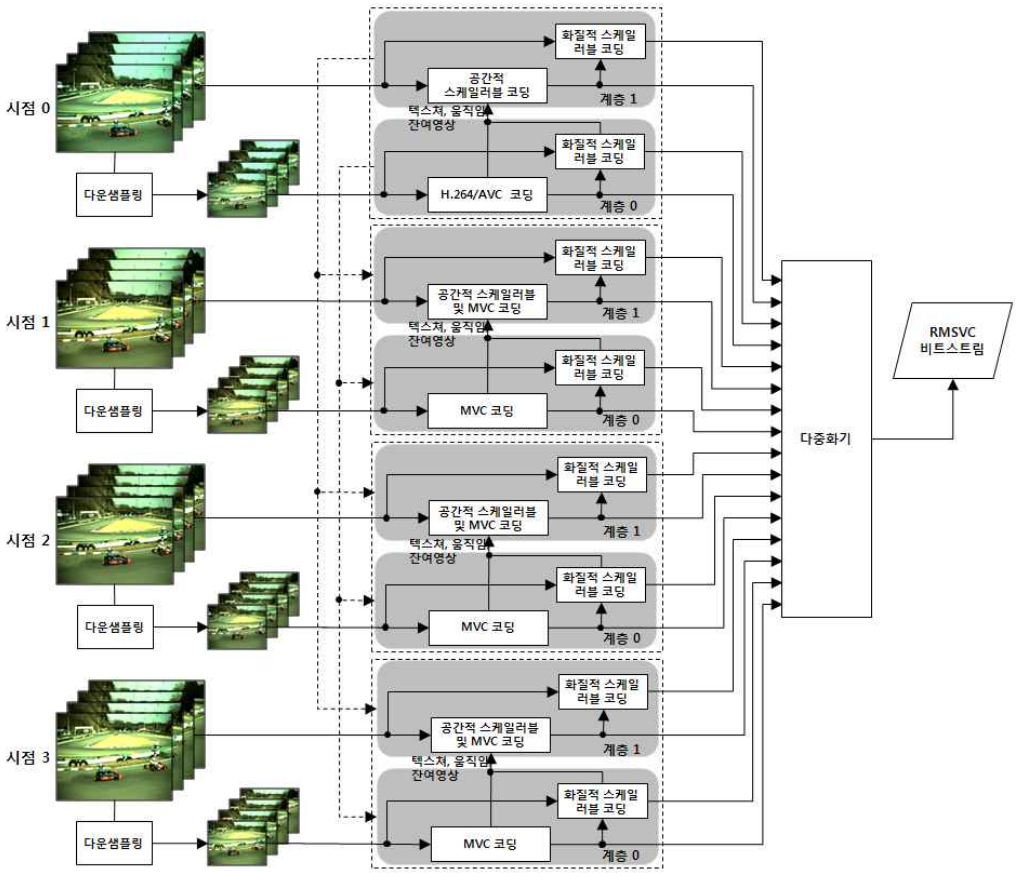


그림 3. 실감형 다시점 스케일러블 비디오 코딩 구조의 예.
Fig. 3. Example of Realistic Multi-view Scalable Video Coding Structure.

코딩을 수행하고 사람의 눈이 인지하기에 적절한 화질을 지원하도록 코딩을 수행한다. 그 후에 다양한 화질을 지원하기 위한 화질적 스케일러블 코딩을 수행한다. 각 시점의 공간적 기본 계층의 코딩이 끝난 후에는 각 시점의 공간적 향상 계층(계층 1)의 코딩을 공간적 기본 계층에서 수행한 코딩의 방법으로 코딩을 그대로 수행하되 공간적 기본 계층의 움직임 정보(Motion information), 텍스쳐(Texture) 정보, 잔여 영상(Residual) 등을 예측 데이터로 사용하여 코딩을 수행한다. 이와 같이 수행된 데이터를 다중화기를 통해 통합하여 하나의 RMSVC 비트스트림을 생성한다.

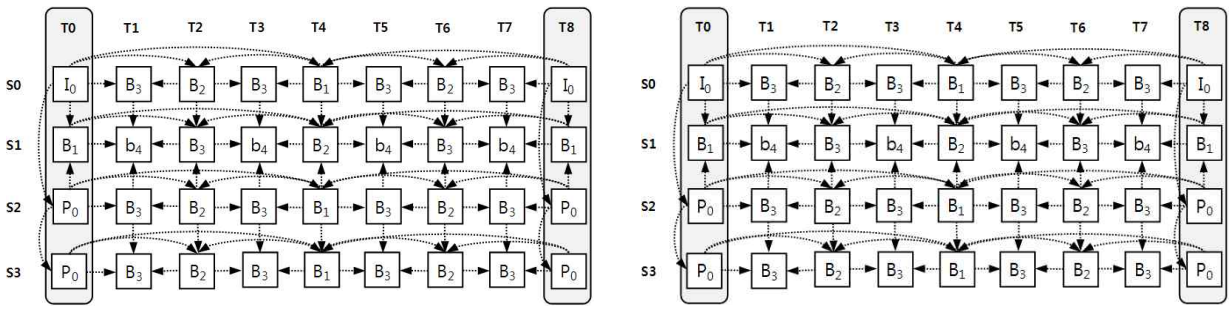
Ⅳ. 실감형 다시점 스케일러블 코딩의 계층별 예측 구조

본 장에서는 실감형 다시점 스케일러블 비디오 코딩의 구조에서 실제 실감형 응용분야를 더욱 잘 지원하기 위해 기존의 단순한 예측 구조를 계층별로 예측할 수 있는 구조를 제안한다.

앞 장에서 설명한 실감형 다시점 스케일러블 비디오 코딩의 구조에서는 그림 4의 예와 같이 모든 계층에서 동일한 예측 구조를 갖는다. 동일한 예측 구조를 갖는 이유는 보통 공간적 향상 계층에서는 하위 계층을 모두 디코딩하여 텍스쳐 정보를 예측하는 것이 아니라 하위 계층의 예측 구조에 따른 움직임 정보와 해당 움직임 정보를 사용하여 예측된 잔여 데이터 정보를 주로 예측하여 코딩을 수행하기 때문에 공간적 향상 계층에서는 기본적으로 공간적 향상 계층의 코딩 구조를 그대로 따르게 되는 것이다. 하지만 그림 4와 같이 모든 공간적 계층에서 동일한 예측 구조를 갖게 된다면 다시점 비디오 코딩에서와 같이 시점 방향 예측 코딩으로 인한 랜덤 액세스 성능은 떨어지게 된다. 랜덤 액세스 기능은 실제 응용에 있어서 상당히 중요한데 그림 4와 같이 4개의 시점의 영상을 전송한다고 가정하였을 때 사용자가 항상 모든 시점의 영상을 필요로 하는 것이 아니라 임의의 원하는 시점 또는 시점들을 필요로 할 수 있기 때문이다. 예를 들어 만약 하나의 시점을 지원하는 단말을 소유하고 있는 사용자일 경우에는 임의의 원하는 시점을 선택할 수 있어야 하고 스테레오를 지원하는 단말을 소유하고 있는 사용자일 경우에는 임의의 두 시점을 선택할 수 있어야 한다.

만약 그림 4의 공간적 향상 계층의 S1시점을 액세스하기 원한다면, 먼저 기본적으로 공간적 기본 계층의 같은 S1시점에 대해 움직임 보상을 제외한 디코딩을 수행해야 하고, 동일한 공간적 향상 계층에서는 S1시점이 S0시점과 S2시점을 참조하기 때문에 S0, S1, S2의 모든 픽쳐에 대해 디코딩을 수행해야 하는 복잡한 코딩 과정이 필요한 것을 확인할 수 있다.

시점 랜덤 액세스의 복잡도를 정량적으로 알아보기 위해 적절한 표현 방법을 정의할 필요가 있다. 그 방법으로 현재 액세스하고자 하는 시점의 픽쳐의 수를 C라고 하였을 때, 기본 계층을 포함하여 디코딩을 수행해야 하는 픽쳐의 수(M)로 복잡도를 표현할 수 있는데 이를 수식 (1)과 같이 표



(a) 공간적 기본 계층의 예측 구조 (b) 공간적 향상 계층의 예측 구조

그림 4. 실감형 다시점 스케일러블 비디오 코딩에서 각 공간적 계층의 예측 구조
Fig. 4. Prediction structure of each spatial layer in Realistic Multi-view Scalable Video Coding

현할 수 있다.

$$M = \left\{ \left(1 + \frac{h_{base} \cdot w_{base} \cdot r_{base}}{h_{enh} \cdot w_{enh} \cdot r_{enh}} \right) \cdot \left(1 + \frac{\alpha + \beta \cdot (GOP_size - 1)}{GOP_size} \right) \right\} \cdot C \quad (1)$$

hbase와 wbase는 각각 공간적 기본 계층의 각 픽쳐의 높이와 넓이를 의미하고, henh와 wenh는 각각 공간적 향상 계층의 각 픽쳐의 높이와 넓이를 의미한다. rbase는 공간적 기본 계층의 프레임율을 의미하고, renh는 공간적 향상 계층의 프레임율을 의미한다. GOP_size는 각 시점의 시간 방향의 GOP의 크기를 의미하고 α는 현재 시점의 앵커픽쳐(Anchor picture; 그림 4에서 음영에 해당하는 픽쳐)의 연속적으로 참조하는 시점의 수를 의미하고 β는 현재 시점의 비앵커픽쳐(Non-anchor picture)의 연속적으로 참조하는 시점의 수를 의미한다.

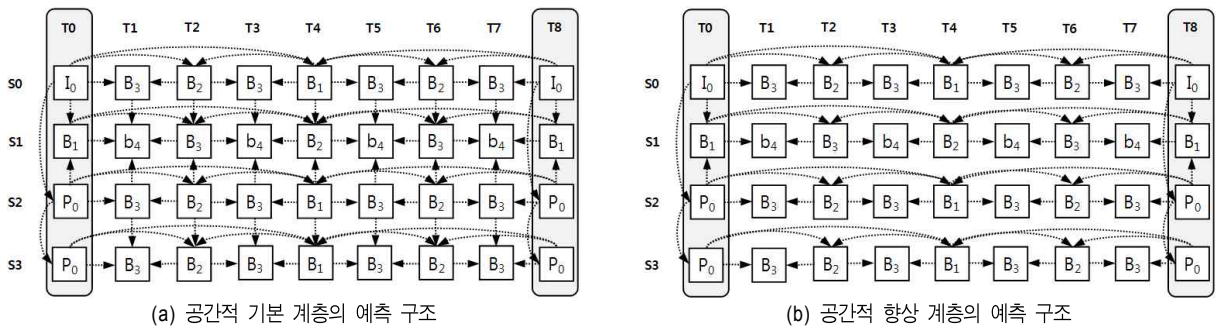
그림 4의 공간적 향상 계층의 영상 크기가 공간적 기본 계층의 영상 크기보다 4배 일 경우에 각 시점의 랜덤 액세스를 위해 디코딩이 필요한 픽쳐의 수(M)를 표 1에 나타내었다. S0시점은 앵커픽쳐와 비앵커픽쳐 모두 참조하는 시점이 존재하지 않기 때문에 1.25C의 값을 갖고, S1시점은 앵커픽쳐와 비앵커픽쳐 모두 2개의 연속적으로 참조하는 시점이 존재하기 때문에 3.75C의 값을 갖는다. S2시점은 앵커픽쳐에서만 1개의 연속적으로 참조하는 시점이 존재하기 때문에 1.41C의 값을 갖고, S3시점은 앵커픽쳐에서는 S2시점을 참조하나 S2시점이 S0시점을 참조하기 때문에 2개의 연속적으로 참조하는 시점이 존재하고 비앵커픽쳐에서는 1개의 연속적으로 참조하는 시점이 존재하기 때문에 2.66C의 값을 갖는다. 결과를 통해 시점 간 예측을 가장 많이 수행하는 S1시점의 M이 3.75C로 가장 큰 것을 확인할 수 있으며 시점 랜덤 액세스를 위해 상당히 많은 연산을 필요로 함을 확인할 수 있다.

표 1. 그림 4의 공간적 향상 계층 각 시점의 랜덤 액세스를 위해 디코딩이 필요한 픽쳐의 수
Table 1. Number of pictures for random access of each view in the spatial enhancement layer (in case of Fig. 4)

시점	$\frac{h_{base} \cdot w_{base} \cdot r_{base}}{h_{enh} \cdot w_{enh} \cdot r_{enh}}$	GOP_size	α	β	M
S0	0.25	8	0	0	1.25C
S1			2	2	3.75C
S2			1	0	1.41C
S3			2	1	2.66C

이러한 단점을 극복하기 위해 본 논문에서는 계층별 예측 구조를 제안한다. 이는 계층별로 시점 간 예측 구조를 각각 적응적으로 설정하여 랜덤 액세스의 성능을 높일 수도 있고 다양한 실감형 응용 분야에 적용이 실제적으로 가능하도록 허용하는 방법이다. 계층별로 시점 간 예측 구조를 상당히 다양한 조합으로 생성할 수 있지만 본 논문에서는 공간적 향상 계층의 예측 구조를 변경시킨 두 가지 경우를 설명한다.

공간적 향상 계층의 예측 구조를 변경시킨 첫 번째 예로 그림 5와 같이 그림 4의 예측 구조에서 공간적 향상 계층의 비앵커픽쳐의 시점 간 예측을 수행하지 않는 형태의 예측 구조를 가질 수 있다.



(a) 공간적 기본 계층의 예측 구조 (b) 공간적 향상 계층의 예측 구조

그림 5. 계층별 예측 구조의 첫 번째 예
Fig. 5. Example 1 of the layer-based prediction structure

표 2. 그림 5의 공간적 향상 계층 각 시점의 랜덤 액세스를 위해 디코딩이 필요한 픽쳐의 수
Table 2. Number of pictures for random access of each view in the spatial enhancement layer (in case of Fig. 5)

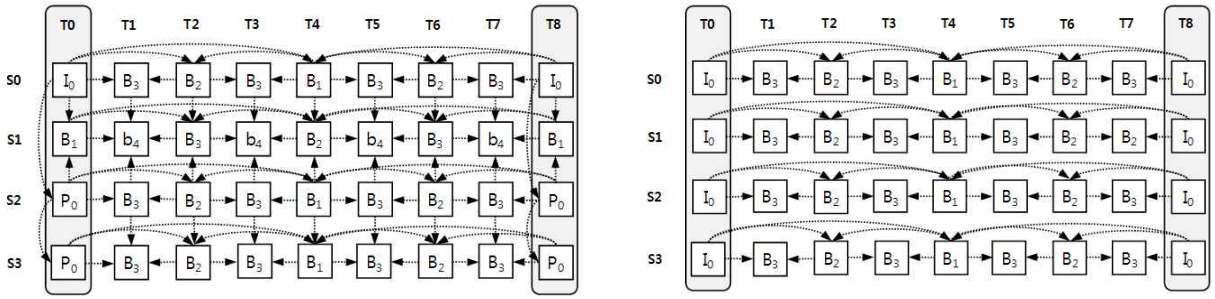
시점	$\frac{h_{base} \cdot w_{base} \cdot r_{base}}{h_{enh} \cdot w_{enh} \cdot r_{enh}}$	GOP_size	α	β	M
S0	0.25	8	0	0	1.25C
S1			2	2	1.56C
S2			1	0	1.41C
S3			2	1	1.56C

그림 5의 예측 구조에서는 비앵커픽쳐에서 시점 간 예측을 수행하지 않기 때문에 공간적 향상 계층의 랜덤 액세스 성능 측면에서 향상이 있을 것을 예상할 수 있다. 실제 각 시점의 랜덤 액세스를 위해 디코딩이 필요한 픽쳐의 수(M)를 표 2에 나타내었다. S0시점은 앵커픽쳐와 비앵커픽쳐 모두 참조하는 시점이 존재하지 않기 때문에 1.25C의 값을 갖고, S1시점은 앵커픽쳐에서만 2개의 연속적으로 참조하는 시점이 존재하기 때문에 1.56C의 값을 갖는다. S2시점은 앵커픽쳐에서만 1개의 연속적으로 참조하는 시점이 존재하기 때문에 1.41C의 값을 갖고, S3시점은 앵커픽쳐에서만 2개의 연속적으로 참조하는 시점이 존재하기 때문에 1.56C의 값을 갖는다. 이로써 랜덤 액세스의 성능이 향상이 되었음을 확인할 수 있는데, 그림 4의 공간적 향상 계층의 S1시점을 액세스 할 경우와 그림 5의 공간적 향상 계층의 S1시점을 액세스 할 경우를 비교해 보면, 디코딩을 수행해야 하는 픽쳐의 수가 약 2.4배 감소되었음을 확인할 수 있다.

공간적 향상 계층의 예측 구조를 변경시킨 두 번째 예로 그림 6과 같이 그림 4의 예측 구조에서 공간적 향상 계층의 모든 시점 간 예측을 수행하지 않는 형태의 예측 구조를 가질 수 있다. 그림 6의 예측 구조에서는 시점 간 예측을 수행하지 않기 때문에 랜덤 액세스 성능 측면에서는 가장 성능이 좋다고 할 수 있다. 실제 각 시점의 랜덤 액세스를 위해 디코딩이 필요한 픽쳐의 수(M)는 모든 시점에서 참조하는 시점이 존재하지 않아 α와 β가 0이 되기 때문에 1.25C로 모두 동일하다. 그림 6의 예측 구조를 살펴보면 공간적 기본 계층의 앵커픽쳐와 공간적 향상 계층의 앵커픽쳐의 픽쳐의 타입이 다르게 되는 경우를 보게 되는데 이 경우에는 향상 계층에서 기본 계층의 움직임 정보와 잔여 데이터 정보를 예측 데이터로 사용하지 못하기 때문에 예측 코딩 효율의 저하가 예상이 된다. 하지만 기본 계층의 성능은 그대로 유지하면서 향상 계층에서는 스케일러빌리티 기능을 가지면서 동시에 랜덤 액세스 기능을 완벽하게 지원할 수 있는 것이 장점이라고 할 수 있다.

V. 성능 평가

본 장에서는 실감형 다시점 스케일러블 비디오 코딩의 성능을 살펴보기 위해 다시점 비디오 코딩(MVC), 스케일러블 비디오 코딩(SVC) 방법과 비교 실험을 수행하였다. 다시점 비디오 코딩 방법은 참조 소프트웨어인 JMVM



(a) 공간적 기본 계층의 예측 구조 (b) 공간적 향상 계층의 예측 구조

그림 6. 계층별 예측 구조의 두 번째 예.
Fig. 6. Example 2 of the layer-based prediction structure.

8.0[19]을 사용하였고 스케일러블 비디오 코딩 방법은 참조 소프트웨어인 JSVM 9.12[20]를 사용하였다. 그리고 실감형 다시점 스케일러블 비디오 코딩 방법은 JMVM 8.0과 JSVM 9.12를 통합하여, 본 논문의 제안 방법에 적합하도록 수정하여 구현된 소프트웨어를 사용하였다. 실제 응용 환경에 적합한 성능 평가를 위하여, 그림 7에 도시한 바와 같이, 응용 가능한 가상 시나리오를 설정하여 해당 시나리오에 맞게 실험환경을 조성하였다.

그림 7의 가상 시나리오에서는, 입력되는 콘텐트는 HD급의 다시점 영상이며, 입력 받은 HD급의 다시점 영상을 RMSVC 인코더로 인코딩하여 다양한 방송망(지상파, 위성, 케이블)을 통해 전송한다. 그리고 방송망 안에서는 일반 2차원 디스플레이 장치, 스테레오스코픽 디스플레이 장치, 다시점 디스플레이 장치, 시점 선택형 2차원 디스플레이 장치, 시점 선택형 스테레오스코픽 디스플레이 장치 등 다양한 디스플레이 장치가 존재할 수 있는데 그림 7의 위에서부터 차례대로 각 장치의 동작을 다음과 같이 살펴볼 수 있다.

H.264 디코더를 지원하는 2차원 SDTV의 경우에는 전송 받은 RMSVC 동영상 정보에서 H.264에 해당하는 동영상 데이터만을 수신하여 디코딩함으로써 영상을 감상할 수 있다.

MVC 디코더를 지원하는 2차원 시점 선택형 SDTV의 경우에는 전송받은 RMSVC 동영상 정보에서 MVC에 해당하는 동영상 데이터만을 수신하여 디코딩하고 디코딩한 시점들 중 멀티플렉서를 통해 원하는 시점을 선택함으로써 해당하는 영상을 감상할 수 있다.

SVC 디코더를 지원하는 2차원 HDTV의 경우에는 전송받은 RMSVC 동영상 정보에서 고화질 혹은 고해상도의 단일시점 동영상 데이터만을 수신하여 디코딩함으로써 고화질, 고해상도의 영상을 감상할 수 있다.

RMSVC 디코더를 지원하는 시점 선택형 스테레오스코픽 디스플레이 장치의 경우에는 전송받은 RMSVC 동영상 정보에서 고해상도, 고화질의 원하는 두 시점의 데이터를 비트스트림 단위에서 선택하여 디코딩함으로써 고화질, 고

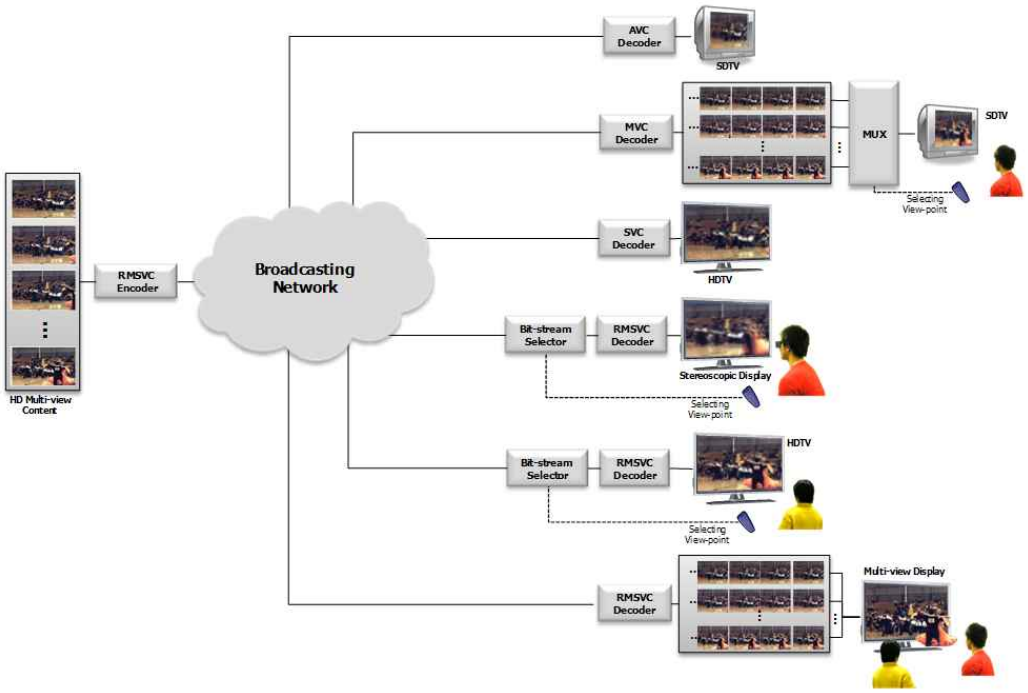


그림 7. 응용 시나리오: 방송 환경에서 실감형 다시점 스케일러블 비디오 코딩
Fig. 7. Application scenario: .Realistic Multi-view Scalable Video Coding in Broadcasting

화질의 스테레오스코픽 영상을 감상할 수 있다.

RMSVC 디코더를 지원하는 2차원 시점선택형 HDTV의 경우에는 전송 받은 RMSVC 동영상 정보에서 고해상도, 고화질의 원하는 시점의 데이터를 비트스트림 단위에서 선택하여 디코딩함으로써 해당하는 시점의 고해상도, 고화질의 영상을 감상할 수 있다.

RMSVC 디코더를 지원하는 다시점 영상 디스플레이 장치의 경우에는 전송받은 RMSVC 동영상 정보에서 기본해상도의 고화질의 다시점 데이터를 선택하여 디코딩함으로써 기본해상도의 고화질의 다시점 영상을 감상할 수 있다.

본 논문에서는, 위의 가상 시나리오 환경을 전부 적용하면 매우 복잡한 실험결과가 도출되며 또한 실험 가능한 다시점 비디오 시퀀스가 다양하게 확보되어 있지 않은 관계로, MPEG 및 ITU-T 국제표준화 회의에서 다시점 비디오 코딩 표준을 제정하기 위한 참조 영상 (reference image)으로 이용되는 3개의 다시점 비디오 시퀀스를 이용하여 실험결과를 도출하도록 한다. 본 논문에서는 공간적 해상도 및 시간적 해상도는 각각 2개의 계층 (공간적 해상도 및 프레임율 측면에서)으로 나누고, 시점은 8개의 시점과 각각의 시점을 전부 각각 코딩한 경우로 한정하여 실험결과를 도출하였다.

실험을 위해 다시점 비디오 코딩 표준의 참조 비디오 시퀀스인 Race1 VGA 30Hz 시퀀스와 Ballroom VGA 25Hz 시퀀스, Akko&Kayo VGA 30Hz 시퀀스를 사용하였다. 두 개의 공간적 계층을 지원하기 위해서 각각의 시퀀스를 1/2로 MPEG 다운샘플링 필터를 사용하여 QVGA 영상을 생성하였으며, QVGA 영상에서는 VGA 영상의 1/2의 프레임율을 갖도록 조정하였다. 또한 각 실험에서는 계층적인 B-픽쳐 구조를 사용하였고 계층적인 B-픽쳐 구조의 각 레벨마다의 QP(Quantization parameter)값은 동일한 비교를 위해 스케일러블 비디오 코딩에서 사용한 방법을 사용하여 동일하게 적용하였다[16]. 그리고 Akko&Kayo 시퀀스는 2차원 배열 형태의 시퀀스이므로 가로 방향으로 4개의 시점과 세로 방향으로 2개의 시점의 코딩을 수행하였다.

1. 실감형 다시점 스케일러블 비디오 코딩 구조의 성능 평가

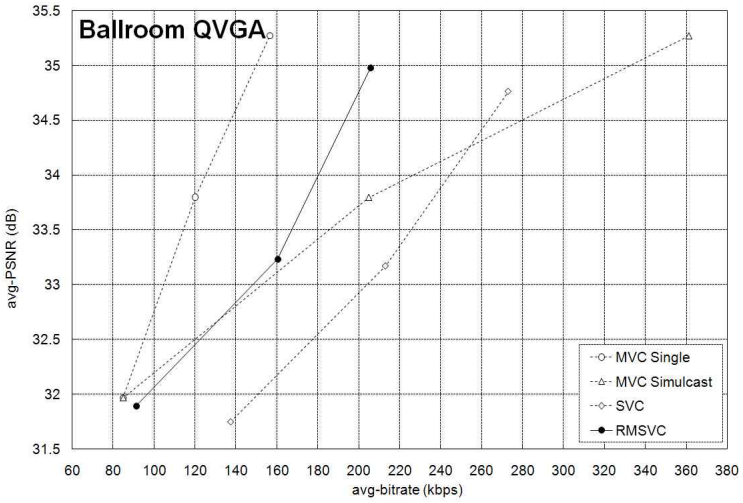
본 절에서는 실감형 다시점 스케일러블 비디오 코딩 구조의 성능 평가를 실시하였다. 다양한 실감형 단말들을 지원하기 위한 기본적인 구조를 설정하여 실험을 수행하였으며 성능 평가를 위한 실험 조건은 표 3과 같다. 그림 8, 그림 9, 그림 10은 각각 Ballroom, Race1, Akko&Kayo 시퀀스에 대한 RD(Rate-distortion)-Curve 결과를 보여준다. 다시점 비디오 코딩 방법의 실험에서 각각의 QP를 사용하여 코딩한 결과를 MVC Single로 표시하였으며 다시점 비디오 코딩 방법의 실험에서 하위 계층들의 정보를 모두 포함하고 있는 경우를 가정하였을 때의 결과를 MVC Simulcast로 표

표 3. 실감형 다시점 스케일러블 비디오 코딩의 비교를 위한 실험 조건
Table 3. Test condition for Realistic Multi-view Scalable Video Coding

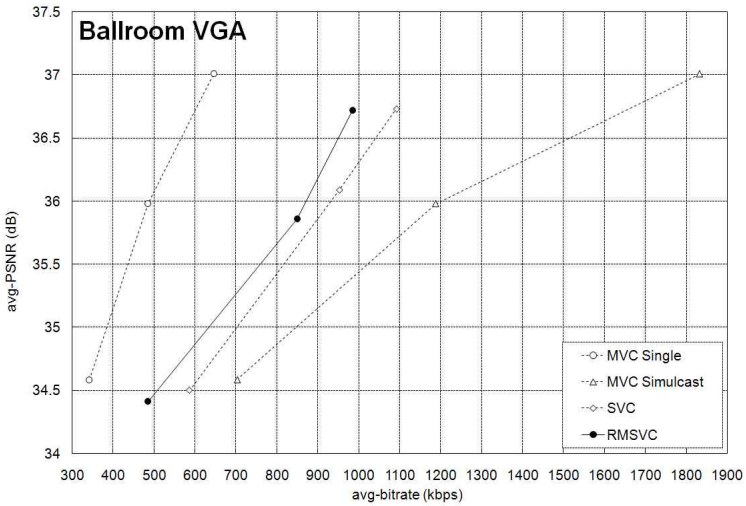
	RMSVC	MVC	SVC
공간적 기본 계층 (QVGA)	- GOP 크기: 8 - 앵커 픽쳐 간격: 8 - QP: 37 - MGS 계층 QP: 32 - MGS Vector: 3, 13 - 시점의 수: 8 - View간 예측 수행	- GOP 크기: 8 - 앵커 픽쳐 간격: 8 - QP: 37, 34,5, 32 - 시점의 수: 8 - View간 예측 수행	- GOP 크기: 8 - 인트라픽쳐 간격: 8 - QP: 37 - MGS 계층 QP: 32 - MGS Vector: 3, 13 - 시점의 수: 8
공간적 향상 계층 (VGA)	- GOP 크기: 16 - 앵커 픽쳐 간격: 16 - QP: 37 - MGS 계층 QP: 32 - MGS Vector: 3, 13 - 시점의 수: 8 - View간 예측 수행	- GOP 크기: 16 - 앵커 픽쳐 간격: 16 - QP: 37, 34,5, 32 - 시점의 수: 8 - View간 예측 수행	- GOP 크기: 16 - 인트라픽쳐 간격: 16 - QP: 37 - MGS 계층 QP: 32 - MGS Vector: 3, 13 - 시점의 수: 8

시하였다. 스케일러블 비디오 코딩 방법의 결과는 SVC로 표시하였고, 실감형 다시점 스케일러블 비디오 코딩 방법의 결과는 RMSVC로 표시하였다.

그림 8의 Ballroom 시퀀스 결과를 비교하였을 때 공간적 기본 계층(QVGA)의 비교 결과에서는 RMSVC가 MVC Single 보다 PSNR 측면에서 0.4 dB ~ 2.1 dB 성능이 감소함을 확인할 수 있다. 하지만 MVC Simulcast와 비교하였을 경우에는 상위 계층(Higher Layer)으로 갈수록 RMSVC의 성능이 크게 향상되는 것을 확인할 수 있으며, 공간적 기본 계층의 최상위 화질적 향상 계층의 결과는 MVC Simulcast보다 비트율 측면에서 약 37% 데이터를 절약할 수 있음을 확인할 수 있다. SVC와의 비교 결과에서는 RMSVC의 성능이 PSNR 측면에서 1 dB ~ 1.9 dB 성능이 향상되었음을 확인할 수 있다. 공간적 향상 계층(VGA)의 비교 결과



(a) 공간적 기본 계층(QVGA)

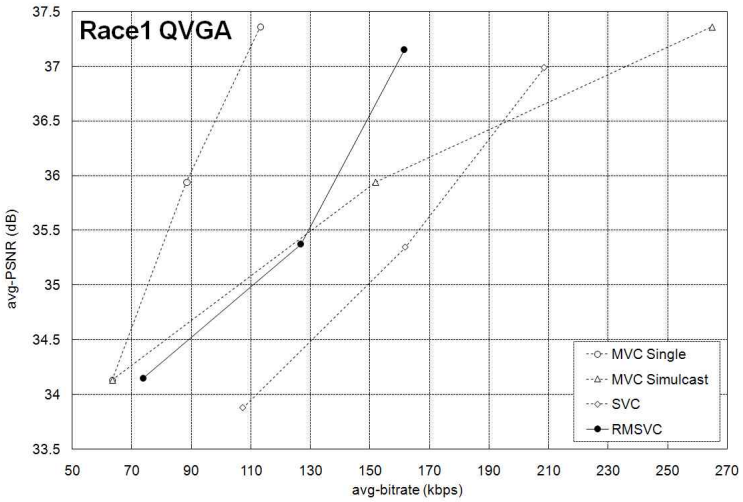


(b) 공간적 향상 계층(VGA)

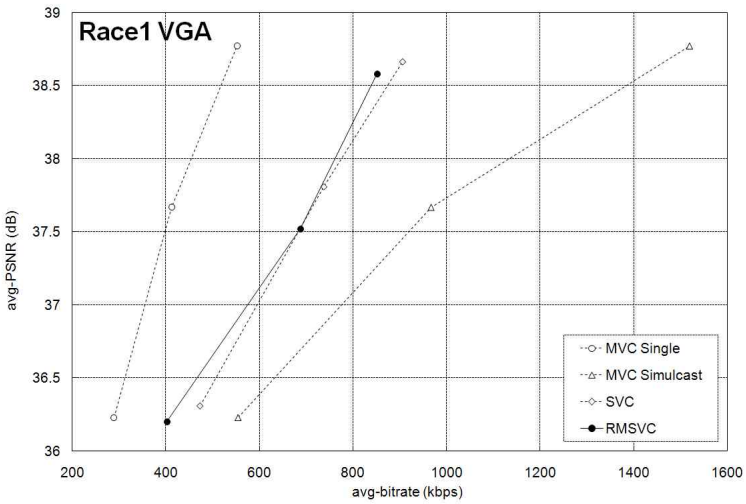
그림 8. Ballroom 시퀀스의 PSNR 결과 비교
Fig. 8. Comparison of PSNR results in Ballroom sequence

에서는 RMSVC의 성능이 MVC Single의 성능보다 PSNR 측면에서 약 1.6 dB ~ 1.9 dB 감소하였음을 확인할 수 있다. 하지만 MVC Simulcast와 비교하였을 경우에는 비트율 측면에서 약 24% ~ 41% 데이터를 절약할 수 있음을 확인할 수 있다. SVC와의 비교 결과에서는 RMSVC의 성능이 PSNR 측면에서 0.2 dB ~ 0.5 dB 성능이 향상되었음을 확인할 수 있다.

그림 9의 Race1 시퀀스 결과를 비교하였을 때 공간적 기본 계층(QVGA)의 비교 결과에서는 RMSVC가 MVC Single 보다 PSNR 측면에서 약 0.7 dB ~ 2.3 dB 성능이 감소하였음을 확인할 수 있다. 하지만 MVC Simulcast와 비교하였을 경우에는 상위 계층으로 갈수록 RMSVC의 성능이 크게 향상되는 것을 확인할 수 있으며, 공간적 기본 계층의 최상위 화질적 향상 계층의 결과는 MVC Simulcast



(a) 공간적 기본 계층(QVGA)

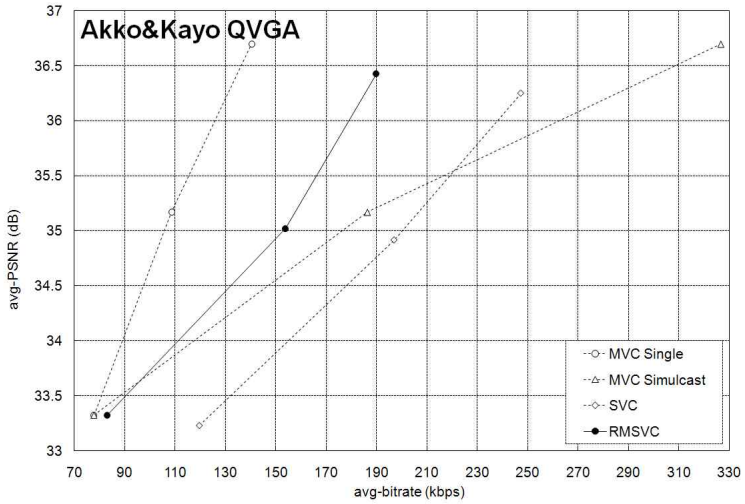


(b) 공간적 향상 계층(VGA)

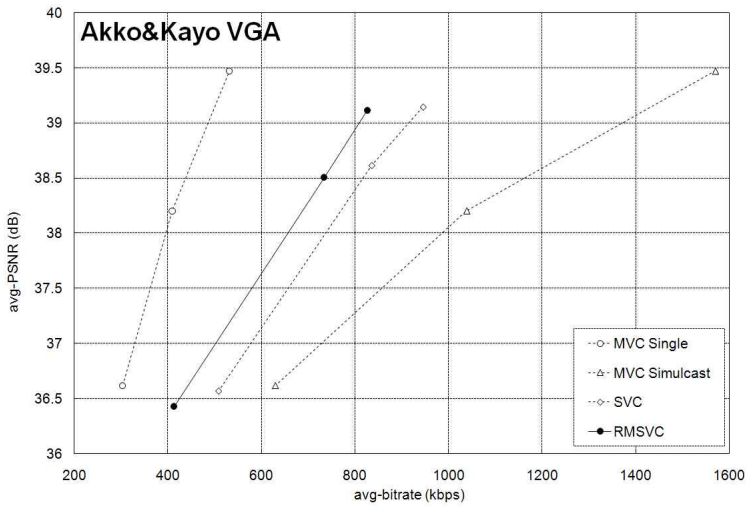
그림 9. Race1 시퀀스의 PSNR 결과 비교
Fig. 9. Comparison of PSNR results in Race1 sequence

보다 비트율 측면에서 약 35% 데이터를 절약할 수 있음을 확인할 수 있다. SVC와의 비교 결과에서는 RMSVC의 성능이 PSNR 측면에서 1 dB ~ 1.8 dB 향상되었음을 확인할 수 있다. 공간적 향상 계층(VGA)의 비교 결과에서는 RMSVC의 성능이 MVC Single의 성능보다 PSNR 측면에서 약 1.3 dB ~ 1.8 dB 감소하였음을 확인할 수 있다. 하지만 MVC Simulcast와 비교하였을 경우에는 비트율 측면에서 약 27% ~ 40% 데이터를 절약할 수 있음을 확인할 수 있다. SVC와의 비교 결과에서는 RMSVC의 성능이 PSNR 측면에서 0 dB ~ 0.3 dB 향상되었음을 확인할 수 있다.

그림 10의 Akko&Kayo 시퀀스 결과를 비교하였을 때 공간적 기본 계층(QVGA)의 비교 결과에서는 RMSVC가 MVC Single 보다 PSNR 측면에서 약 0.3 dB ~ 2.1 dB 성능이 감소하였음을 확인할 수 있다. 하지만 MVC Simulcast



(a) 공간적 기본 계층(QVGA)



(b) 공간적 향상 계층(VGA)

그림 10. Akko&Kayo 시퀀스의 PSNR 결과 비교
Fig. 10. Comparison of PSNR results in Akko&Kayo sequence

와 비교하였을 경우에는 상위 계층으로 갈수록 RMSVC의 성능이 크게 향상되는 것을 확인할 수 있으며, 공간적 기본 계층의 최상위 화질적 향상 계층의 결과는 MVC Simulcast보다 비트율 측면에서 약 37% 데이터를 절약할 수 있음을 확인할 수 있다. SVC와의 비교 결과에서는 RMSVC의 성능이 PSNR 측면에서 1 dB ~ 1.6 dB 향상되었음을 확인할 수 있다. 공간적 향상 계층(VGA)의 비교 결과에서는 RMSVC의 성능이 MVC Single의 성능보다 PSNR 측면에서 약 1.6 dB ~ 2.3 dB 감소하였음을 확인할 수 있다. 하지만 MVC Simulcast와 비교하였을 경우에는 비트율 측면에서 약 30% ~ 40% 데이터를 절약할 수 있음을 확인할 수 있다. SVC와의 비교 결과에서는 RMSVC의 성능이 PSNR 측면에서 0.4 dB ~ 0.5 dB 향상되었음을 확인할 수 있다.

실험 결과를 통해 RMSVC의 공간적 기본 계층의 결과는 SVC 성능에서 각 시점 간의 예측을 통해 시점 간에 존재하는 중복성을 제거함으로써 PSNR 측면에서 기본 계층에서 1 dB ~ 1.9 dB, 향상 계층에서 0.5 dB 까지 성능이 향상되었음을 확인할 수 있었다. 또한 MVC Simulcast와의 비교에서 비트율 측면에서 20% ~ 40% 절약 효과가 있었다. 따라서 실제 응용에 있어서 다양한 전송환경과 다양한 단말들을 지원함에 있어서 고화질의 서비스가 가능하며 저장 공간이나 전송 대역폭 절약에 효과가 있을 것을 예상할 수 있다.

2. 계층별 예측 구조의 성능 평가

본 절에서는 계층별 예측 구조의 객관적인 화질의 비교와 실감형 응용분야에서 중요한 요소가 되는 시점 랜덤 액세스의 성능 평가를 수행하였다. 실감형 다시점 스케일러블 비디오 코딩 소프트웨어를 계층별로 시점 간 예측 구조를 변경할 수 있도록 구현하였으며, 공간적 향상 계층(VGA)에서 앵커픽쳐에서만 시점 간 예측을 허용하는 방법(RMSVC Proposed-1)과 공간적 향상 계층(VGA)에서 시점 간 예측을 허용하지 않는 방법(RMSVC Proposed-2)에 대한 실험을 수행하였다. 비교를 수행하는 다시점 비디오 코딩(MVC), 스케일러블 비디오 코딩(SVC), 실감형 다시점 스케일러블 비디오 코딩(RMSVC) 방법의 실험 조건은 표 3과 같으며, RMSVC Proposed-1과 RMSVC Proposed-2에 대한 실험 조건은 표 4와 같다.

실제 객관적인 성능평가에서 제안하는 방법들의 공간적 기본 계층의 결과는 RMSVC의 결과와 동일하기 때문에 생략하며, 공간적 향상 계층의 결과의 비교 결과만을 보여준다. 그림 11, 그림 12, 그림 13은 각각 Ballroom 시퀀스, Race1 시퀀스, Akko&Kayo 시퀀스에 대한 RD-Curve 결과를 보여준다. 다시점 비디오 코딩 방법의 실험에서 각각의 QP를 사용하여 코딩한 결과를 MVC Single로 표시하였으며, 다시점 비디오 코딩 방법의 실험에서 하위 계층들의 정보를 모두 포함하고 있는 경우를 가정하였을 때의 결과를

표 4. 계층별 예측 구조의 실험 조건
Table 4. Test condition for layer-based prediction structure

	RMSVC Proposed-1	RMSVC Proposed-1
공간적 기본 계층 (QVGA)	- GOP 크기: 8 - 앵커 픽쳐 간격: 8 - QP: 37 - MGS 계층 QP: 32 - MGS Vector: 3, 13 - 시점의 수: 8 - View간 예측: 수행	- GOP 크기: 8 - 앵커 픽쳐 간격: 8 - QP: 37 - MGS 계층 QP: 32 - MGS Vector: 3, 13 - 시점의 수: 8 - View간 예측: 수행
공간적 향상 계층 (VGA)	- GOP 크기: 16 - 앵커 픽쳐 간격: 16 - QP: 37 - MGS 계층 QP: 32 - MGS Vector: 3, 13 - 시점의 수: 8 - View간 예측: 앵커픽쳐에서만 수행	- GOP 크기: 16 - 앵커 픽쳐 간격: 16 - QP: 37 - MGS 계층 QP: 32 - MGS Vector: 3, 13 - 시점의 수: 8 - View간 예측: 수행하지 않음

MVC Simulcast로 표시하였다. 스케일러블 비디오 코딩 방법의 결과는 SVC로 표시하였고, 실감형 다시점 스케일러블 비디오 코딩 방법의 결과는 RMSVC로 표시하였다. 그리고 공간적 향상 계층에서 앵커픽쳐에서만 시점 간 예측을 허용하는 방법을 RMSVC Proposed-1로 표시하였고, 공간적 향상 계층에서 시점 간 예측을 허용하지 않는 방법을 RMSVC Proposed-2로 표시하였다.

그림 11의 Ballroom 시퀀스의 공간적 향상 계층의 결과에서 RMSVC Proposed-1을 RMSVC와 비교하였을 때 PSNR 측면에서 약 0.1 dB ~ 0.2 dB 성능이 감소된 것을 확인할 수 있는데, 이것은 공간적 향상 계층에서 비앵커픽쳐의 공간적 계층 간 예측 코딩 시 기본 계층의 시점 간 예측 움직임 정보를 예측하지 못했기 때문이라고 할 수 있다. RMSVC Proposed-2의 결과는 SVC의 결과보다 PSNR

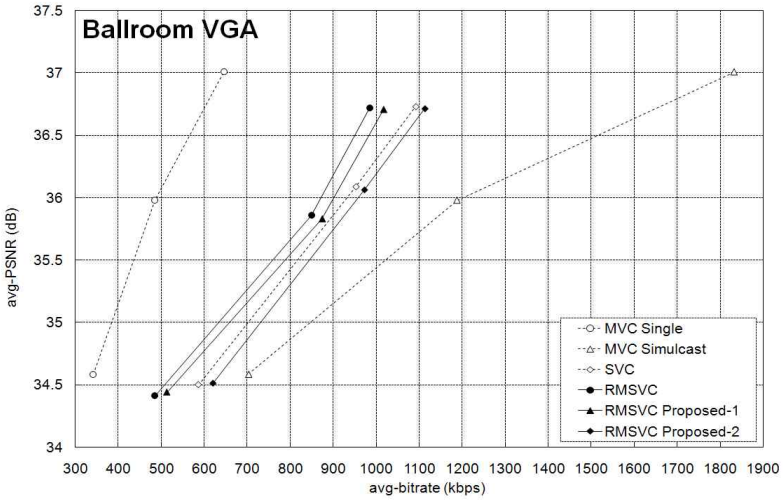


그림 11. Ballroom 시퀀스의 공간적 향상 계층의 PSNR 결과 비교
Fig. 11. Comparison of PSNR results of the spatial enhancement layer in Ballroom sequence

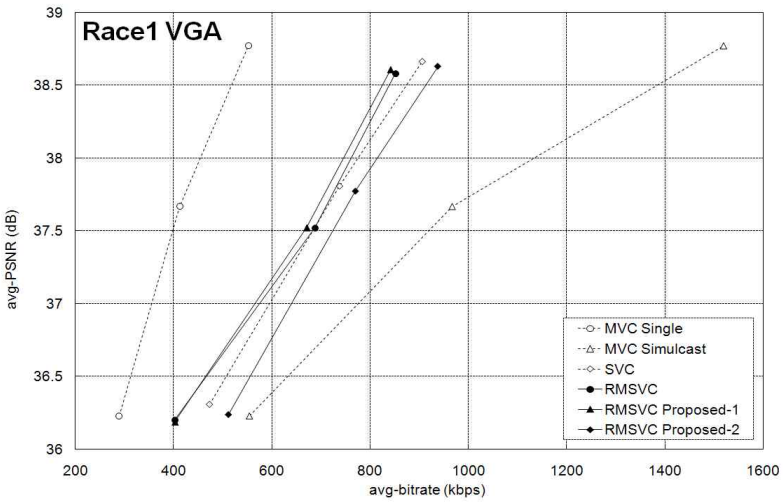


그림 12. Race1 시퀀스의 공간적 향상 계층의 PSNR 결과 비교
Fig. 12. Comparison of PSNR results of the spatial enhancement layer in Race1 sequence

측면에서 약 0.15 dB 감소가 된 것을 확인할 수 있다. 이러한 감소 현상은 앞에서 설명한 것과 같이 공간적 향상 계층에서 기본 계층의 움직임 정보와 잔여 데이터 정보를 예측 데이터로 사용하지 못하기 때문에 오히려 SVC의 결과 보다 성능이 감소된 것이라 할 수 있다.

그림 12의 Race1 시퀀스의 공간적 향상 계층의 결과에서 RMSVC Proposed-1을 RMSVC와 비교하였을 때 PSNR 측면에서 약 0 dB ~ 0.1 dB 성능이 오히려 향상된 것을 확인할 수 있는데, 이것은 공간적 향상 계층에서 비앵커픽쳐의 공간적 계층 간 예측 코딩 시 기본 계층의 시점 간 예측 움직임 정보를 예측하지 못했지만 시간 방향 예측 효율이 더욱 좋기 때문에 향상이 된 것이라 볼 수 있다. RMSVC Proposed-2의 결과는 SVC의 결과보다 PSNR 측면에서 약 0.17 dB ~ 0.25 dB 감소가 된 것을 확인할 수 있다.

그림 13의 Akko&Kayo 시퀀스의 공간적 향상 계층의 결과에서 RMSVC Proposed-1을 RMSVC와 비교하였을 때 성능이 거의 동일함을 확인할 수 있다. RMSVC Proposed-2의 결과는 SVC의 결과보다 PSNR 측면에서 약 0.1 ~ 0.15 dB 감소가 된 것을 확인할 수 있다.

실제 본 실험의 RMSVC, RMSVC Proposed-1, RMSVC Proposed-2 각각의 예측 구조에서 공간적 향상 계층의 각 시점의 액세스하기 위해 디코딩이 필요한 픽쳐의 수(M)를 구함으로써 랜덤 액세스의 성능을 비교할 수 있다.

RMSVC의 공간적 향상 계층의 각 시점의 액세스하기 위해 디코딩이 필요한 픽쳐의 수는 모든 시점에서 평균적으로 2.21C의 값을 가지며 최대로 S5시점에서 3.52C의 값을 갖는다.

RMSVC Proposed-1의 공간적 향상 계층의 각 시점의 액세스하기 위해 디코딩이 필요한 픽쳐의 수는 모든 시점에서 평균적으로 1.29C의 값을 가지고, 최대로 S5와 S7시점에서 1.41C의 값을 갖는다. RMSVC Proposed-1의 디코딩을 수행해야 하는 픽쳐의 수가 RMSVC에 비해 평균적으로 약 1.7배 감소되었고 S5시점의 경우에는 약 2.5배 감소되었다. 결과를 통해 랜덤 액세스의 성능이 매우 향상되었음을 확인할 수 있다.

RMSVC Proposed-2의 공간적 향상 계층의 각 시점의 액세스하기 위해 디코딩이 필요한 픽쳐의 수는 모든 시점에서 참조하는 시점이 존재하지 않기 때문에 1.13C로 모두 동일하다.

그림 11, 그림 12, 그림 13의 RD-Curve의 결과를 통해서에서 RMSVC Proposed-1의 결과는 RMSVC의 결과와 거의 유사함을 확인할 수 있다. 랜덤 액세스 성능에서는 제안

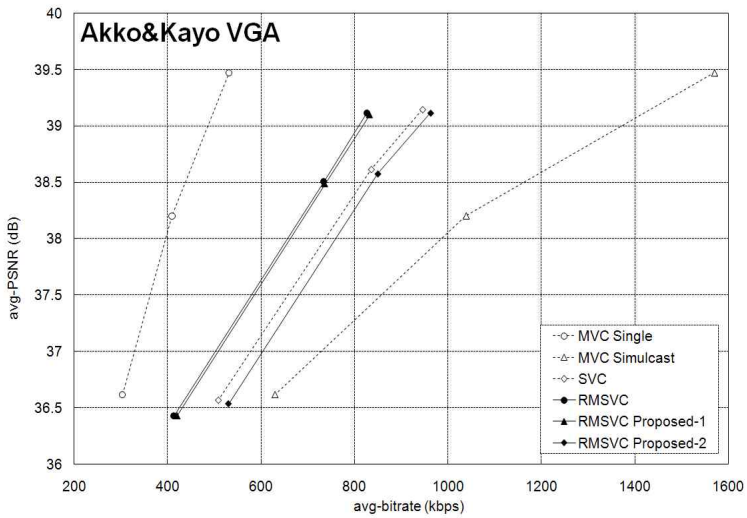


그림 13. Akko&Kayo 시퀀스의 공간적 향상 계층의 PSNR 결과 비교.
Fig. 13. Comparison of PSNR results of the spatial enhancement layer in Akko&kayo sequence.

하는 RMSVC Proposed-1의 방법이 RMSVC의 방법보다 매우 우수함을 확인할 수 있고, 랜덤 액세스 성능이 가장 좋은 RMSVC Proposed-2의 방법과도 큰 차이를 보이지 않음을 확인할 수 있다. 따라서 RMSVC Proposed-1의 방법 즉 공간적 향상 계층의 시점 간 예측을 앵커픽쳐에서만 수행하도록 하는 구조가 실제 실감형 응용을 가능하게 하는 적합한 방법이라 할 수 있다.

RMSVC Proposed-2의 방법은 공간적 향상 계층에서 오히려 SVC의 성능보다 0.1 dB ~ 0.25 dB 성능이 감소한 것을 확인할 수 있다. 하지만 앞 절의 공간적 기본 계층에서의 성능 평가에서 확인한 것과 같이 RMSVC의 성능은 SVC의 성능 보다 매우 우수하며 MVC Simulcast와 비교하였을 경우에는 항상 우수한 성능을 가진다는 장점이 있다. 또한 앞에서 언급하바와 같이 시점의 랜덤 액세스 성능이 가장 뛰어난 장점이 있다.

실제 시점 랜덤 액세스 성능 향상을 통해 많은 숫자의 시점에서 고화질의 특정한 시점의 선택하여 전송할 때 적은 양의 시점만을 선택하여 전송하여 전송 대역폭을 절약할 수 있으며, 다시점 영상의 디코딩 시에 항상 1.5C 이하의 값을 갖기 때문에 디코딩 복잡도를 낮출 수 있다.

Ⅵ. 결 론

본 논문에서는 실감형 다시점 스케일러블 비디오 코딩을 실현하기 위한 방법에 대한 연구를 수행하였다. 현재 실감형 응용 분야에 대한 관심이 확대되고 있으며 방송/통신의 융합과 유/무선의 융합이 활발히 진행되고 있다. 이와 같이 다양하고 복잡한 환경에서 동영상 정보를 효율적으로 취급하기 위한 방법으로 실감형 다시점 스케일러블 비디오 코딩의 개념 및 동작을 알아보았다. 그리고 기존의 다시점 비디오 코딩 방법의 기능과 스케일러블 비디오 코딩의 기능을 통합한 형태의 실감형 다시점 스케일러블 비디오 코딩 구조를 설계하였고, 해당 구조를 기존의 다시점 비디오 코딩 방법과 스케일러블 비디오 코딩 방법과의 객관적 비교 실험을 통해 실제 적용 가능성에 대해 살펴보았다. 그리고 실제 응용에서 중요한 시점 간의 랜덤 액세스 성능을 향상시켜주는 계층별 예측 구조를 제안하였다. 실제 실험 결과를 통해서 제안 하는 방법이 랜덤 액세스의 성능을 향상시켜주는 것을 확인하였으며 이를 통해 실제 좀 더 효율적이고 유연한 응용이 가능함을 확인할 수 있었다.

참 고 문 헌

[1] "Generic Coding of Moving Pictures and Associated Audio Information – Part 2: Video," ITU-T Recommendation H.262 and ISO/IEC 13818-2 (MPEG-2 Video), November 1994.

[2] "Video Codec for Audiovisual Services at p×64 kbit/s," ITU-T Recommendation H.261, Version 1: November 1990, Version 2: March 1993.

[3] "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s – Part 2: Video," ISO/IEC 11172-2 (MPEG-1 Video), March 1993.

[4] "Video Coding for Low Bit Rate Communication," ITU-T Recommendation H.263, Version 1: November 1995, Version 2: January 1998, Version 3: November 2000.

[5] "Coding of Audio-Visual Objects – Part 2: Visual," ISO/IEC 14496-2 (MPEG-4 Visual), Version 1: April 1999, Version 2: February 2000, Version: May 2004.

[6] "Advanced Video Coding for Generic Audiovisual Services," ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 Part 10 AVC), Version 1: March 2003, Version 2: May 2004, Version 3: March 2005, Version 4: September 2005, Version 5 and Version 6: June 2006, Version 7: April 2007, Version 8: July 2007.

[7] T. Wiegand, G. Sullivan, G. Bjøntegaard, A. Luthra, "Overview of the H.264/AVC Video Coding Standard," IEEE Transactions on circuits and systems for video technology, vol. 13, no. 7, pp. 560-576, July 2003.

[8] "AHG on 3D Video coding in MPEG," ISO/IEC JTC1/SC29/WG11, N4524, Thailand, December 2001.

[9] "Application and Requirements for 3DAV," ISO/IEC JTC1/SC29/WG11, N5877, Norway, July 2003.

[10] "Report on 3DAV Exploration," ISO/IEC JTC1/SC29/WG11, N5878, Norway, July 2003.

[11] "Survey of Algorithms used for Multi-view Video Coding (MVC)," ISO/IEC JTC1/SC29/WG11, N6909, Hong Kong, January 2005.

[12] A. Vetro, P. Pandit, H. Kimata, A. Smolic, "Joint Multiview Video Model (JMVM) 1.0," Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-T208, Austria, July 2006.

[13] A. Vetro, P. Pandit, H. Kimata, A. Smolic, Y. Wang, "Joint Multiview Video Model (JMVM) 8.0," Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-AA207, Switzerland, April 2008.

[14] A. Vetro, P. Pandit, H. Kimata, A. Smolic, Y. Wang, "Joint Draft

7.0 on Multiview Video Coding," Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-AA209, Switzerland, April 2008.

[15] T. Wiegand, G. Sullivan, J. Reichel, H. Schwarz, M. Wien, "Joint Video ITU-T Rec. H.264 | ISO/IEC 14496-10 / Amd.3 Scalable Video Coding," Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-X201, Switzerland, June-July 2007.

[16] J. Reichel, H. Schwarz, M. Wien, "Joint Scalable Video Model JSVM-11," Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-X202, Switzerland, June-July 2007.

[17] H. Schwarz, D. Marpe, T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," IEEE Transactions on circuits and systems for video technology, vol. 17, no. 9, pp. 1103-1120, September 2007.

[18] T. Wigend, "Overview paper and presentation on SVC," Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-W132, California, USA, April 2007.

[19] P. Pandit, A. Vetro, Y. Chen, "JMVM 8 software," Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-AA208, Switzerland, April 2008.

[20] "JSVM Software," Joint Video Team (JVT) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, CVS repository: jvtuser@garcon.ient.rwth-aachen.de, tag: JSVM_9_12.

저 자 소 개

박 민 우

- 2003년 2월 : 경희대학교 컴퓨터공학과 학사
- 2005년 2월 : 경희대학교 컴퓨터공학과 석사
- 2005년 3월 ~ 2009년 2월 : 경희대학교 컴퓨터공학과 박사
- 2009년 3월 ~ 현재 : 삼성전자 DMC 연구소 선임연구원
- 주관심분야 : 비디오 신호처리, 멀티미디어, 영상처리

박 광 훈

- 1985년 2월 : 연세대학교 전자공학과 학사
- 1987년 7월 : 연세대학교 전자공학과 석사
- 1991년 1월 : Case Western Reserve University, Dept. of EEAP 석사
- 1995년 1월 : Case Western Reserve University, Dept. of EEAP 박사
- 1995년 3월 ~ 1997년 2월 : 현대전자 멀티미디어연구소 책임연구원
- 1997년 3월 ~ 2001년 2월 : 연세대학교 문리대학 전산학과 부교수
- 2001년 3월 ~ 현재 : 경희대학교 컴퓨터공학과 교수
- 주관심분야 : 멀티미디어, 비디오 신호처리, 패턴인식, 영상처리, 계산지능