

# 제한적 네트워크 환경 하에서 효율적인 파노라마식 멀티미디어 콘텐츠 분할 전송 방법

김 병 철<sup>a)</sup>, 이 건 희<sup>a)</sup>, 이 인 재<sup>b)</sup>, 김 규 현<sup>a)†</sup>

## An Efficient Transmission Method of Panoramic Multimedia Contents in a Limited Bandwidth Environment

Byungchul Kim<sup>a)</sup>, Gunhee Lee<sup>a)</sup>, Injae Lee<sup>b)</sup>, and Kyuheon Kim<sup>a)†</sup>

### 요 약

본 논문에서는 파노라마 영상으로 구성된 멀티미디어 콘텐츠를 효율적으로 전송하는 방법을 제안한다. 파노라마 영상은 넓은 시야와 다양한 시점을 사용자에게 제공한다. 하지만 현재의 파노라마식 멀티미디어 콘텐츠 전송 서비스에는 몇 가지 제한조건이 존재한다. 먼저 파노라마 영상이 갖는 데이터양이 기존의 멀티미디어 영상보다 크기 때문에, 제한적인 전송 대역폭을 갖고 있는 통신환경에서는 초기 지연 시간이 길어지는 문제점이 있다. 또한 사용자가 제한된 디스플레이 장치(모바일 폰, PC 모니터 등)에서 전송된 영상 중 일부 시점의 영상만을 재생하게 되어, 콘텐츠의 사용효율이 떨어지는 문제점을 가지고 있다. 따라서 본 논문에서는 이러한 문제점을 해결하기 위한 방안으로 MPEG-4 시스템의 장면 구성 기술을 활용한 파노라마식 멀티미디어 콘텐츠 전송을 제안한다. 또한 본 논문에서는 장면 구성 기술을 활용하여, 단일 파노라마식 멀티미디어 콘텐츠 전송과 분할 파노라마식 멀티미디어 콘텐츠 전송을 비교함으로써 전송의 효율성을 증명하였다.

### Abstract

This paper proposes an efficient transmission method for the panoramic multimedia contents. The panoramic video provides wide sight and various view-point to the user. The traditional methods of the panoramic multimedia content transmission has several limitations, as follow; A client suffers a long initial delay time to play a panoramic video when it is transmitted through a limited bandwidth network, because the panoramic video has larger data size than a general video. And if a client's display device has limited resolution, such as mobile phone, laptop PC monitor, etc. it can not display the entire panoramic video that has a wide view video sequence. So, in order to overcome the obstacles, this paper proposes an efficient transmission of panoramic multimedia contents. This method will increase the transmission efficiency throughout the technique of the scene description in MPEG-4 system. Also we demonstrated the efficiency of the proposed method by comparison with existing methods.

Keyword : MPEG-4 Stream, Panoramic multimedia contents

a) 경희대학교 전자정보학부

College of Electronics and Information, Kyung Hee University

b) 한국전자통신연구원 방통융합미디어연구부 융합서비스연구팀

Convergence Service Research Team, Broadcasting and Telecommunications Media Research Department, ETRI

† 교신저자 : 김규현(kyuheonkim@khu.ac.kr)

※ “본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음” (NIPA-2011-(C1090-1111-0001))

※ “본 연구는 정보통신부의 “지능형 통합정보방송 기술 개발 (SmarTV)” 사업의 일환으로 수행된 연구 결과임”

· 접수일(2011년3월31일), 수정일(1차:2011년7월6일, 2차:9월6일), 게재확정일(2011년9월6일)

## 1. 서론

현재 정보통신기술분야 중 인터넷을 통한 데이터 전송 기술의 발전은, 다양한 멀티미디어 콘텐츠(multimedia contents)의 이용을 가능하게 하는 유희유 역할을 하고 있고, 사용자들은 영상처리 및 편집 기술의 발전으로 보다 다양한 형태의 콘텐츠, 즉 다시점(multiview) 영상<sup>[1]</sup>과 파노라마(panoramic)[2]영상의 사용을 원하게 되었다. 이 중 파노라마 영상은 사용자가 원하는 화면을 보다 쉽고 효율적으로 활용하기 위해, 카메라에서 획득되는 파노라마 영상을 기하학적인 교정과 공간적인 합성 등을 통하여 자유로운 시점과 보다 넓은 시야를 제공하고 있다. 이는 기존의 영상 콘텐츠를 서비스 제작자가 제공하는 시점의 한계를 벗어나 사용자들에게 보다 다양한 시점의 영상을 제공하여, 영상 콘텐츠 재생에 시점의 자유를 증가시킬 수 있다. 이러한 파노라마 영상 콘텐츠 기술은 항공사진의 해석, 영상 변화 감지, 비디오 압축에서부터 간단한 영상의 편집에 이르기까지 다양한 분야에서 응용되고 있다.

이러한 파노라마 영상을 구성하는 방법은 여러 가지가 있다. 예를 들어, 여러 방향에서 획득한 영상을 구 모양의 공간에 투사하거나, 원기둥, 정사각형 모형의 공간에 투사하여, 이 도형들을 사각형 형태로 펼친 2차원에 각각의 영상을 정렬하여 파노라마 영상을 구성할 수 있다<sup>[3]</sup>. 이 중 원기둥 형태는 구 형태 또는 정사각형 형태에 비해 위, 아래 방향에 대한 시야 제한이 있지만, 영상을 획득하는 과정이 수월하기 때문에 가장 널리 사용된다. 실 예로 'Point Grey Research'사에서 출시한 'Ladybug2/3' 카메라<sup>[4]</sup>에서 파노라마 영상을 획득하는 방법 또한 원기둥 형태의 방법을 사용한다. 카메라에서 얻어지는 파노라마 영상을 시퀀스 이미지(sequence image)로 나타낼 때는, 일반적으로 획득된 영상을 세로로 분할하고, 이를 각각의 타일(tile) 단위로 부호화(encoding)한다. 파노라마 영상을 타일로 분할하는 이유는, 일반적인 영상에 비해 많은 양의 데이터를 포함하고 있으므로, 파노라마 전체 영상을 한꺼번에 부호화 하는 방법 보다는, 부분적으로 부호화하는 경우가 부호기(encoder)에서 처리해야 하는 데이터의 양이 줄어들기 때문이다.

마찬가지로 하나의 카메라를 통해 획득한 파노라마 멀티미디어 콘텐츠를 전송하는 시스템은 파노라마 영상이 기존의 멀티미디어 영상보다 많은 데이터를 갖고 있기 때문에, 전송 시에 순간적으로 매우 많은 데이터가 처리되어야 하고, 이에 따라 제한적인 전송 대역폭(bandwidth)을 갖고 있는 통신환경에서는 처리되어야 하는 버퍼링(buffering) 시간이 길어져, 초기 지연 시간(initial delay time)이 발생하는 문제점을 갖고 있다. 또한 사용자가 모바일, 노트북 등과 같이 한정된 디스플레이 장치에서 파노라마 영상을 재생하고자 할 때, 파노라마 영상의 일부 시점에 해당하는 영상만을 재생하게 되어, 즉 소비되지 않는 영상을 전송하는 경우가 발생하여, 파노라마 영상 콘텐츠의 사용효율이 떨어지는 문제점을 가지고 있다.

따라서 본 논문에서는 이러한 문제점을 해결하기 위한 방안으로, 파노라마식 멀티미디어 콘텐츠 전송 서비스를 구성함에 있어, 하나의 카메라에서 얻은 파노라마 영상을 다수의 영상으로 분할하고, 이를 MPEG-4 시스템(MPEG-4 systems)<sup>[5]</sup>의 장면 구성(scene description)기술을 활용하여 효율적인 전송을 이끌어 낼 수 있는, 분할 파노라마식 멀티미디어 콘텐츠 전송 방안을 제안한다.

MPEG-4 시스템은 비디오, 오디오, 이미지 및 텍스트와 같은 멀티미디어 콘텐츠를 각각의 객체(object)로 취급하고, 각 객체에 대해서 독립적인 기능을 제공하는 표준이다. 이 중 장면 구성 기술은 MPEG-4 시스템의 객체에 대한 시공간적 위치 정보를 표현하고, 사용자 대화형 서비스(user interactivity service)를 제공하기 위한 목적으로 제정된 표준이다.

본 논문에서 제안하는 방법은 분할된 파노라마 영상에 대하여 장면 구성 기술을 기반으로 개별적인 객체로 정의하고, 전체 파노라마식 멀티미디어 콘텐츠를 해당 객체들로 구성된 장면으로 설정한다. 분할된 영상의 전송에서는 사용자가 선택한 시점을 기준으로 해당 시점에 해당하는 영상과 이에 인접한 영상만을 전송함으로써, 사용자에게 제공할 필요가 없는 시점의 영상은 전송하지 않는 방법을 사용한다. 따라서 제안된 방안은 파노라마 영상 중 불필요한 시점에 대한 영상 정보를 전송하지 않기 때문에, 제한된 대역폭에서도 효율적으로 이용할 수 있다. 또한 사용자가

선택하는 시점이 변경되어 다른 시점의 콘텐츠를 재생할 경우에는 MPEG-4 시스템 규격의 객체 서술자 명령어를 사용하여 재생 화면을 재구성 할 수 있다.

본 논문의 2장에서는 MPEG-4 시스템의 장면 구성 기술을 소개하고, 3장에서는 제안된 효율적 파노라마식 멀티미디어 콘텐츠 분할 전송 방법을 소개하고, 효율성을 증명하기 위한 실험 환경 및 결과를 4장의 결론으로 끝맺는다.

## II. MPEG-4 시스템 장면 구성(Scene Description) 기술 개요

본 논문에서는 파노라마식 멀티미디어 콘텐츠 분할 전송 시스템을 구성하는 데 있어, 사용자의 요구에 따른 시점 이동을 수행하기 위해, MPEG-4 시스템의 장면 구성 기술을 사용한다. MPEG-4 시스템은 비디오, 오디오, 정지영상 및 2D 그래픽 객체 등 멀티미디어 콘텐츠를 구성하는 개별의 미디어를 각각 독립적인 객체로 취급하여 콘텐츠를 제작할 수 있는 방법을 제공하는데, 이와 같이 객체를 기반으로 각 미디어를 구성함으로써 사용자 대화형 서비스를 제공할 수 있는 특징을 갖고 있다. 이러한 특징을 제공하기 위하여 하나의 콘텐츠를 구성하고 있는 비디오, 오디오, 정지영상 등의 미디어 객체들의 정보를 객체 기술(object description)로 구성하고, 이를 장면 구성(scene description) 기술을 사용하여 나타낸다.

장면 구성은 BIFS(Binary Format for Scene Description)라는 규격으로 표준화되어 있으며, 각 미디어 객체들의 콘텐츠 내에서의 시간적, 공간적인 배치 및 사용자 입력이나 시간에 따른 객체의 동작을 나타내는 정보를 설명한다. 그 외의 MPEG-4의 장면을 구성하기 위해서는 초기 객체 서술자(IOD: Initial Object Descriptor), 객체 서술자(OD: Object Descriptor), 장면을 구성하는 오디오, 비디오, 미디어 데이터(audio, visual, media data)가 필요하다.

초기 객체 서술자는 MPEG-4 장면을 구성하기 위해 가정 먼저 복호되어야 하는 정보로, 각 미디어의 프로파일(profile)과 레벨(level)을 설명하고, BIFS 스트림과 객체 서술자

스트림을 포인팅 하기 위한 기초 스트림 서술자(Elementary stream descriptor)를 포함한다<sup>6)</sup>.

BIFS는 장면의 구조를 기술하는 노드(node)들의 계층적인 구조를 가지는 집합으로 구성되며, 실제로 위치하게 되는 객체와는 독립적으로 부호화된다. BIFS에서는 장면을 구성하는 객체간의 대화적 기능(interactivity)과 행동을 표현하기 위해서 이벤트(event) 구조를 사용하고, 이는 센서(sensor) 노드의 사용과 노드 내의 속성을 나타내는 필드(field)와 노드간의 연결을 정의하는 라우트(route)의 개념으로 기술된다.

센서 노드는 사용자의 요구나 장면의 변화에 기초한 화면의 전환이나, 객체 표면의 변화 등을 일으키는 이벤트를 생성해 내는 노드이며, 라우트는 이벤트 발생 노드와 이벤트 수신 노드 사이의 연결로서, 노드에서 사용되는 필드 값을 변경하여 노드를 직접적으로 변화시키기 위한 이벤트 전달에 사용된다.

객체 서술자는 장면을 구성하는 각 미디어 객체의 기초 스트림에 대한 정보를 기술하는 서술자들의 집합으로서 각 미디어 객체의 기초 스트림과 장면 구성과의 연결을 제공한다<sup>7)</sup>.

그림 1은 객체 서술자 스트림과 BIFS 스트림, 미디어 스트림의 관계에 대한 설명을 담고 있다. 그림 1에서 보는 바와 같이 (A)부분의 초기 객체 서술자는 장면 구성 스트림과 객체 서술자 스트림을 참조할 수 있는 정보를 갖고 있고, ES\_ID를 통해 각 스트림을 참조한다. 장면 구성 스트림은 (B)부분과 같이 노드와 필드를 이용하여 장면을 구성하는 내용을 담고 있고, 이를 이진화(binary)한 스트림을 BIFS 스트림이라 한다. 장면 구성 스트림이 표현하는 장면 구성 정보의 각 노드 중 Audio source 또는 Movie Texture 같은 일부 노드에서 (C)부분과 같이 ObjectDescriptorID 정보를 통해 객체 서술자 스트림의 객체 서술자를 지정하고, 객체 서술자는 객체 서술자에서 기술된 ES\_ID를 통해 오디오 비주얼 미디어의 한 부분을 지정하여 전체적인 장면을 구성한다. 따라서 위와 같은 구조에 의해 장면 구성과 객체 서술자와 실제 나타나는 미디어 스트림이 각각 따로 전송되지만, 참조관계를 통해 전체적인 장면을 구성하게 된다.

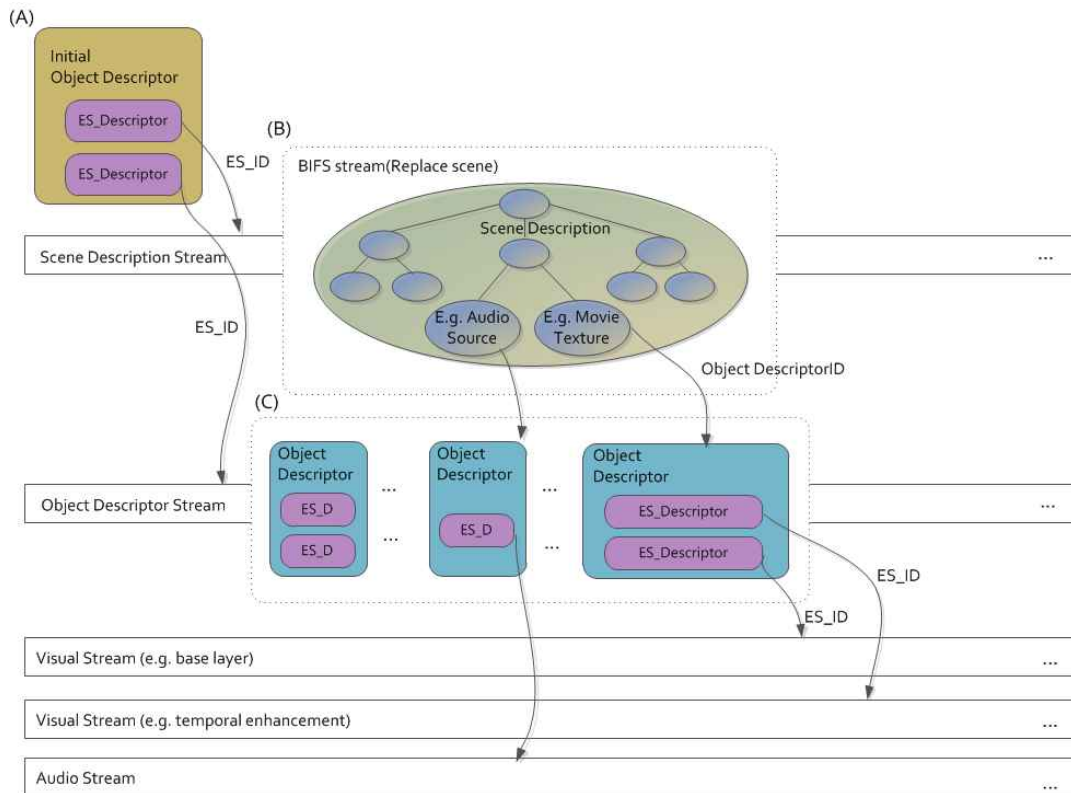


그림 1. MPEG-4 시스템의 구조도  
 Fig. 1. Structure figure of MPEG-4 System

### III. 효율적 파노라마식 멀티미디어 콘텐츠 분할 전송

앞서 설명한 바와 같이 파노라마 영상의 전송은 많은 데이터를 갖고 있기에 초기 지연 시간의 발생과, 한정된 디스플레이에 일부 시점에 해당하는 영상만을 재생하게 되는 문제점이 있다.

본 논문에서는 이러한 문제점을 해결하고자 MPEG-4 시스템을 기반으로 분할된 파노라마 영상을 각각 하나의 시점으로 정의하고, 각 시점간의 시점 이동을 위해 장면 구성 기술을 사용하여, 파노라마식 멀티미디어 콘텐츠를 효율적으로 전송하는 방법을 제안한다.

그림 2는 본 논문이 제안하는 파노라마식 멀티미디어 콘텐츠의 효율적인 분할 전송을 위한 방법을 설명하는 그림이다. 그림 2의 (A)에서 보는 바와 같이 대용량의 파노라마

식 콘텐츠 영상을 임의의 크기의  $N \times M$ 개의 부분 영상으로 분할하도록 한다. 대용량 파노라마식 멀티미디어 콘텐츠를 전송하기 위하여, 그림 2의 (B)에서 나타난 바와 같이 (A)에서 분할된 영상을 각각 부호기를 통해 압축하였으며, 각 분할된 영상간의 공간 정보(spatial information)는 MPEG-4 시스템 기반의 장면 구성 정보로 장면 생성기를 통해 생성한다. 장면생성기는 전체 화면을 구성하는  $N \times M$ 개의 부분 영상을 독립적인 객체로 인식하는 객체 서술자로 기술하고, 부호화된 대용량 파노라마식 멀티미디어 콘텐츠를 전송하는 방법을 나타낸 것으로, 장면 구성 방법에 따라 (A)부분에서 얻어진 각 시점의 부호화 된 영상을 처리하는 역할을 수행한다. 또한 장면생성기에서는 전체 화면을  $N \times M$ 개의 객체 서술자로 구성된 화면으로 정의한 후, 이를 MPEG-4 시스템에 따라 전송되는 영상의 복호화 정보, 즉 영상 압축방법 및 프로파일, 레벨 등의 정보를 나타

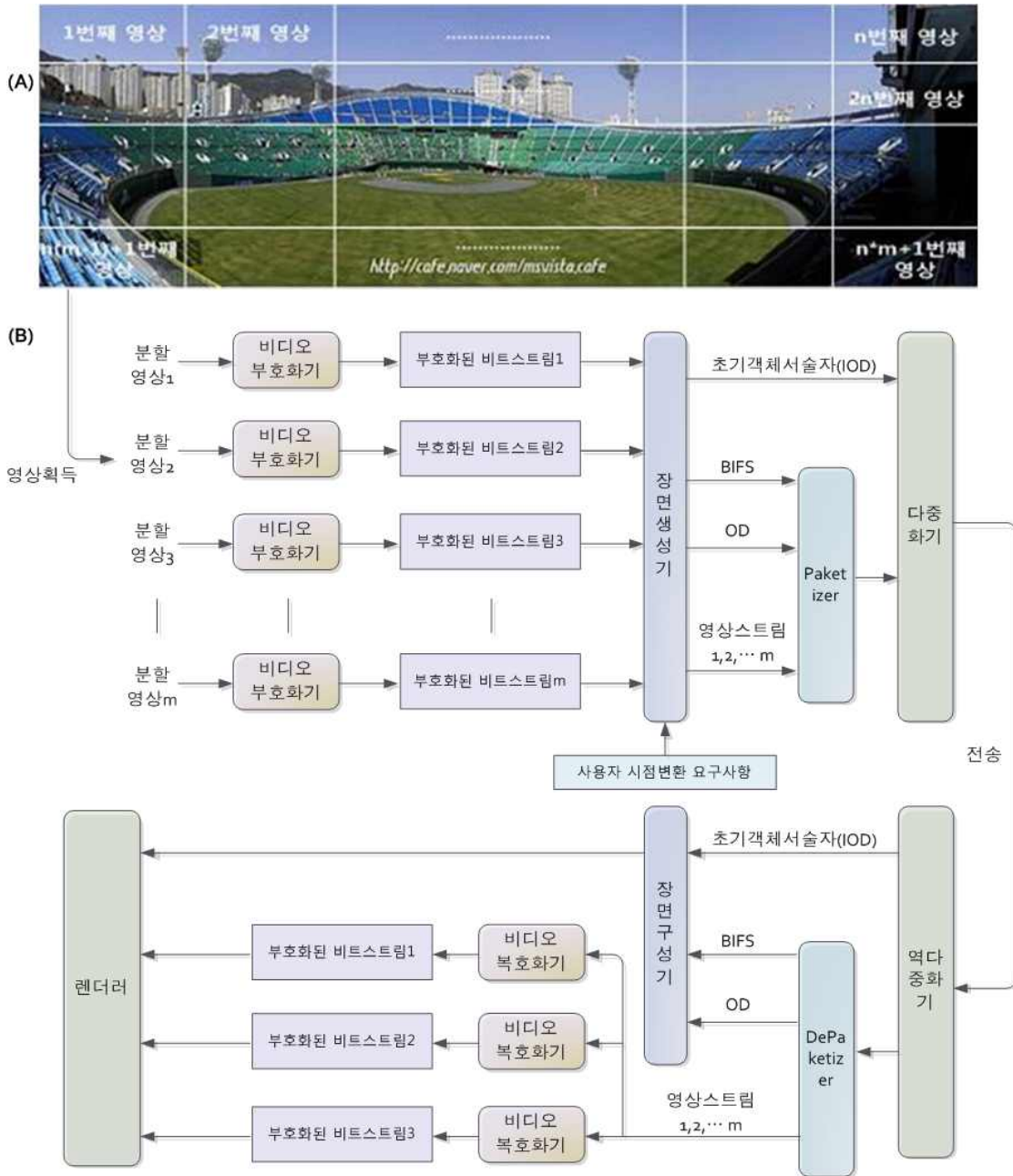


그림 2 다시점 파노라마식 콘텐츠 전송 방법의 도식  
 Fig. 2 Method for multiview panoramic contents transmission

내는 초기 객체 서술자와 장면 구성 스트림, 객체 서술자 스트림과 각 분할된 영상의 기초 스트림으로 장면을 구성한다.

이 중 장면 구성 스트림인 BIFS는 전달하려고 하는 부분 영상의 위치 정보, 즉 k번째 스트림을 중앙으로, k-1과 k+1 번째는 좌우에, k-(n+1)은 중앙 윗부분, k-(n+2)는 왼쪽 상

단 대각선 위치에, k-n은 우측 상단 대각선에 위치한 영상 정보를 나타내고, 객체 서술자 정보는 각각의 분할된 영상 (1, 2, ..., k, ..., n\*m)을 가리키는 포인터 정보를 통해 장면 구성 정보를 나타낸다.

장면생성기에서 위의 역할을 수행한 이후의 처리는 먼저 수신단의 장면 재구성을 위한 기초 정보인 초기 객체 서술자가 다중화기(multiplexer)에 직접 전달되고, 그 외 정보인 장면 구성 스트림, 객체, 전송하고자 하는 특정 분할 장면 스트림은 패킷화기(packetizer)를 통해 패킷화를 수행한 후, 다중화기를 통해 비트스트림으로 변환한 후 단말로 전송한다.

이 후 수신 단에서는 송신된 비트스트림을 역다중화기(de-multiplexer)를 통해 IOD, BIFS, OD, ES 스트림으로 분리하고, IOD를 제외한 스트림은 역패킷화(de-packetizer)과정을 수행한다. 그림2에서 보는 바와 같이 IOD, OD, BIFS는 장면 구성기를 통해 화면 구성 정보로 변환되고, 각각의 영상 스트림은 비디오 복호화기를 통해 재생 가능한 형태의 영상 정보로 복호된다. 최종적으로 렌더러에서는 장면 구성기를 통해 얻은 정보를 바탕으로 복호된 영상 정보를 화면에 재생한다.

또한 사용자가 시점 변환을 원할 경우, 재생하고자 하는 시점 정보를 객체 서술자 명령어(OD Command)를 통해 송신단에 전달하게 되고, 송신 단은 장면 생성기에서 객체 서술자 명령어를 받아 추가로 요청되는 분할 화면의 정보를 전송함으로써 실시간으로 재생되어지는 화면을 갱신(update)한다.

본 논문에서 제안하는 파노라마식 멀티미디어 영상 콘텐츠의 효율적인 분할 전송 방법은 전송하고자 하는 영상 콘텐츠를 임의의 크기로 분할하여 화면을 구성하고, 사용자에게 재생되는 일정 크기의 영상과 이에 이웃한 분할된 영상을 전송하고, 수신 단에서는 전달된 분할 영상만으로 화면을 구성하여 재생한다. 또한 사용자의 시점 변화에 따른 대화형(interaction)기능을 사용하여, 파노라마 영상에서 소비되지 않은 불필요한 영상을 전송하지 않음으로서, 전송 효율을 높이고 시점 변화에 따른 사용자의 요구사항을 실시간으로 반영할 수 있는 장점이 있다.

#### IV. 파노라마식 멀티미디어 콘텐츠 분할 전송의 실험 결과

##### 1. 파노라마식 멀티미디어 콘텐츠 실험 영상

본 실험은 파노라마식 멀티미디어 콘텐츠 송수신 시스템을 구현하고, 전송의 효율성을 증명하는 것을 목적으로 하고 있다. 이는 수신 단말의 디스플레이 해상도와 전송 대역폭을 고려하여, 파노라마식 멀티미디어 영상의 일부만을 전송하게 됨으로써, 불필요한 부분을 전송하는 데 소요되는 부하를 줄이는 것으로 문제를 해결하고자 한다.

이를 구체적으로 설명하면, 사용자의 송신 단말에서 전송하고자 하는 영상을 임의의 N개로 분할하여 저장한 후, 이중 n(n<N)개만을 수신 단말로 전송하여 소비되도록 하는 것이다. 실험은 송신단(Server)에서 하나의 객체 서술자로 이루어진 단일 파노라마식 멀티미디어 콘텐츠와 여러 개의 객체 서술자로 이루어진 분할 파노라마식 멀티미디어 콘텐츠를 수신단(client)으로 전송하고, 재생할 때의 대역폭의 사용량을 측정하여, 이를 비교하는 방식으로 이루어진다.

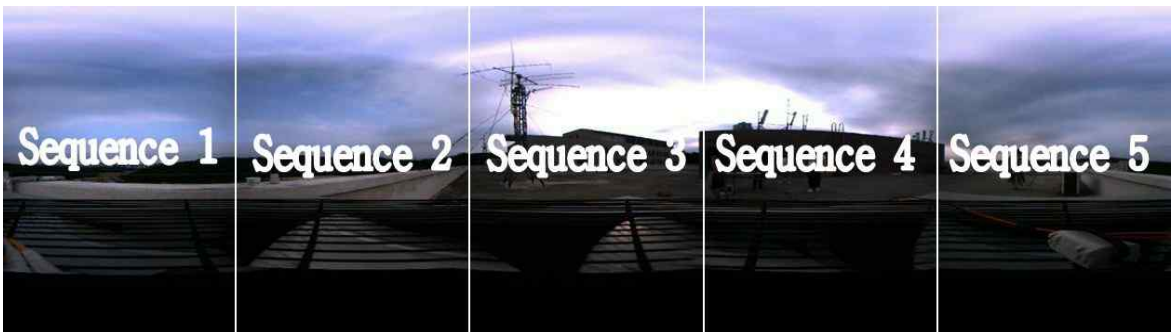
이를 위한 실험 시스템은 BIFS기술로 작성된 xmt 파일<sup>[8]</sup>을 이진 코드로 부호화하여, MPEG-4의 저장포맷 중 하나인 MP4 파일 포맷<sup>[9]</sup>에 독립된 트랙으로 저장을 한다.

본 xmt 파일에서는 하나 또는 5 개의 영상을 지정하고 있다. 실험의 명료함을 위해서 대용량 파노라마식 멀티미디어 영상은 5 개의 화면으로 분할하였으며, 이를 MP4 파일 포맷(file format)내의 각각 트랙으로 생성하여 저장하였다.

먼저 송신 단말의 콘텐츠 중 첫 번째 단일 파노라마식 콘텐츠는 그림 3의 (가)에서 나타난 것처럼 가로, 세로 해상도가 2048\*512인 동영상 콘텐츠로 3D 파노라마식 카메라에서 얻은 영상 시퀀스(sequence)를 MP4 파일 포맷에 하나의 트랙으로 저장하였다. 두 번째 송신 단말의 콘텐츠는 해당 단일 파노라마식 영상을 그림 3의 (나)에서 보는 바와 같이 가로 5 조각으로 분할하였으며, 이렇게 분할된 개별의 영상을 MPEG-4 시스템에서 설명된 객체 서술자의 개념으로 각각 표현하였고, 객체 서술자 명령어를 통해 접근이 가



(가) 단일 파노라마식 동영상 콘텐츠 (2048\*512, 25 fps, mp4 file)



(나) 분할 파노라마식 동영상 콘텐츠

그림 3. 파노라마식 멀티미디어 실험 콘텐츠  
Fig. 3. Panoramic multimedia test contents

능한 형태로 구성하였다.

또한 본 연구에서는 콘텐츠 전송 시스템(Server-Client system)을 사용하여, MP4 파일 포맷 형태의 영상 콘텐츠를 송/수신단을 이용하여 전송하고, 전달된 콘텐츠를 재생하여 소비되는 전체 과정에 대하여 실험을 수행하였다.

콘텐츠 전송 시스템이 시작되면 수신단은 가운데 화면인 시퀀스3을 중심으로, 좌측 시퀀스2, 우측 시퀀스4 만을 전송 받게 된다. 이 상태에서 사용자가 좌측 또는 우측의 원하는 장면을 보기 위해 약속된 대화형 기능을 발생시키면, 객체 서술자 명령어의 접근을 통해 좌측 또는 우측의 화면으로 전환을 수행하게 되어, 좌측이면 시퀀스1을 추가로 전송하고, 시퀀스4는 전송을 하지 않게 된다. 또한 반대로 우측으로의 전환의 경우에는 시퀀스5를 추가로 전송하고, 시퀀스2를 전송하지 않게 된다.

따라서 총 5 개로 분할된 영상은 항상 단말기에 3개의

시퀀스만을 전달하게 되어, 수신 단말기에서는 디스플레이의 해상도에 맞는 영상만을 전달받게 되고, 전송 대역폭을 효율적으로 사용하게 된다. 또한 사용자의 요구사항을 반영한 대화형 기능을 수반하기에, 전체 파노라마 영상을 사용자가 원하는 시점에 맞추어 소비할 수 있는 기능을 제공한다.

## 2. 파노라마식 멀티미디어 콘텐츠를 위한 객체 서술자

표 1에서는 본 실험의 콘텐츠에 대한 메타데이터의 주요 부분을 기술하였다.

가장 먼저 표 1에서 보는 바와 같이 XMT는 Header와 Body부분으로 나누어진다. 문서 상단의 생략된 <Header> (㉠) 부분에는 초기 객체 서술자의 프로파일(profile) 정보 및 복호화 정보가 기술되고, <Body>(㉡) 부분은 MPEG-4

표 1. 다시점 파노라마식 멀티미디어 실험 콘텐츠 생성을 위한 XMT  
 Table. 1. XMT for generating of panoramic multimedia experiment contents

```

<XMT-A>
<Header> ㉑
- Header 자세한 기술 생략 -
</Header>
<Body> ㉒
<Replace> ㉓
<Scene>
<OrderedGroup> ㉔
<children>
<Transform2D translation="-200 0"> ㉕
- 자세한 기술 생략 -
<MovieTexture DEF="MovieTexture_80000" URL="od4" loop="true" speed="1.0000000" repeatS="true"
repeatT="true"
stopTime="-1.0000000" startTime="0.0000000"/> ㉖
</Transform2D>
<Transform2D translation="200 0">
- 자세한 기술 생략 -
<MovieTexture DEF="MovieTexture_80000" URL="od5" loop="true" speed="1.0000000" repeatS="true"
repeatT="true"
stopTime="-1.0000000" startTime="0.0000000"/>
- 반복되는 기술 생략 -
</Transform2D>
<Transform2D translation="-300 -320">
- 자세한 기술 생략 -
<ImageTexture DEF="ImageTexture_60000" URL="od8" repeatS="true" repeatT="true"/>
</Transform2D>
<Transform2D translation="-150 -320">
- 반복되는 기술 생략 -
</Transform2D>
<Transform2D translation="150 -320">
- 반복되는 기술 생략 -
</Transform2D>
<Transform2D translation="300 -320">
- 반복되는 기술 생략 -
</Transform2D>

<TouchSensor DEF="TS17"/> ㉗
<TouchSensor DEF="TS18"/>
<TouchSensor DEF="TS19"/>
<TouchSensor DEF="TS20"/>

<Conditional DEF="C_1"> ㉘
<buffer>
<Replace atNode="VIEWleft" atField="URL" value="od3"/>
<Replace atNode="VIEWright" atField="URL" value="od4"/>
</buffer>
</Conditional>
<Conditional DEF="C_2"> - 반복되는 기술 생략 -
<Conditional DEF="C_3">
<Conditional DEF="C_4">
</children>
</OrderedGroup>
<ROUTE fromNode="TS17" fromField="isActive" toNode="C_1" toField="activate"/> ㉙
<ROUTE fromNode="TS18" fromField="isActive" toNode="C_2" toField="activate"/>
<ROUTE fromNode="TS19" fromField="isActive" toNode="C_3" toField="activate"/>
<ROUTE fromNode="TS20" fromField="isActive" toNode="C_4" toField="activate"/>
</Scene>
</Replace>
<Par begin="0.0000000">
<ObjectDescriptorUpdate> ㉚
<OD>
<ObjectDescriptor ObjectDescriptorID="od3">
<Descr>
<esDescr>
    
```



```

<ES_Descriptor ES_ID="es4">
  <StreamSource URL="sequence1.m4v">
    <EncodingHints streamFormat="MPEG-4 Visual" />
  </StreamSource>
  <decConfigDescr>
    <DecoderConfigDescriptor streamType="4" bufferSizeDB="4600000"
      objectTypeIndication="108">
      </DecoderConfigDescriptor>
    </decConfigDescr>
    <slConfigDescr>
    <SLConfigDescriptor/>
    </slConfigDescr>
  </ES_Descriptor>
</esDescr>
</Descr>
</ObjectDescriptor>
</OD>
<OD>
  <ObjectDescriptor ObjectDescriptorID="od4">
    <Descr>
      <esDescr>
        <ES_Descriptor ES_ID="es5">
          <StreamSource URL="sequence2.m4v">
            <EncodingHints streamFormat="MPEG-4 Visual" />
          </StreamSource>
          <decConfigDescr>
            <DecoderConfigDescriptor streamType="4" bufferSizeDB="4600000"
              objectTypeIndication="108">
              </DecoderConfigDescriptor>
            </decConfigDescr>
            <slConfigDescr>
            <SLConfigDescriptor/>
            </slConfigDescr>
          </ES_Descriptor>
        </esDescr>
      </Descr>
    </ObjectDescriptor>
  </OD>
  <OD> - od3과 od4와 유사함 -
    <ObjectDescriptor ObjectDescriptorID="od5">
    </ObjectDescriptor>
  </OD>
  <OD>
    <ObjectDescriptor ObjectDescriptorID="od6">
    </ObjectDescriptor>
  </OD>
  <OD>
    <ObjectDescriptor ObjectDescriptorID="od7">
    </ObjectDescriptor>
  </OD>
  <OD>
    - 반복되는 기술 생략 -
  </OD>
  <OD>
    - 반복되는 기술 생략 -
  </ObjectDescriptorUpdate>
</Par>
</Body>
</XMT-A>

```

시스템을 이용한 장면 구성 정보를 담고 있다.

장면 구성 정보의 처음은 command로 이루어지며, 해당 xmt 파일은 ReplaceScene command를(㉔) 이용해 장면을 새로 구성하고, 장면 구성의 처음을 구성될 수 있는 4가지 노드 중 하나인 <OrderedGroup>(㉕)을 시작으로 필요한 노드와 각 노드의 필드(field)를 이용해 장면의 구성을 완성하

였다. <Transform2D>(㉖) 노드는 각 이미지 및 동영상 콘텐츠가 재생기(Renderer)에서 보이는 위치를 정해주는 노드이고, <MovieTexture>(㉗) 노드는 실제 보이는 동영상 콘텐츠를 직접 지정할 수 있는 노드로 동영상 콘텐츠의 다양한 조건을 다룰 수 있다.

표 1에서 <Transform2D>(㉖)에서 표현한 것은 그림 3

(나)에서 보이는 분할된 파노라마식 동영상 콘텐츠를 표현하는 것으로 <MovieTexture>(㉑)를 사용해 초기 화면 재생을 기술한 것이다. 또한 장면의 구성이 끝남을 뜻하는 </Replace>아래에 쓰인 <ObjectDescription Update>(㉒)를 이용하여 장면 구성에서 쓰였던 5 개로 분할된 장면에 대해 실제 각각의 영상이 존재하는 URL 값을 기술하여 각 영상을 객체 서술자로 지정하였음을 알 수 있다.

<MovieTexture>(㉑)에서 쓰인 URL="od4" 와 <ObjectDescriptor ObjectDescriptorID= "od4">를 통해서 분할된 영상이 화면에 표시되는 과정을 알 수 있다.

사용자 대화형 기능으로 시점 이동을 수행하기 위하여, <conditional>(㉓) 노드를 이용해 객체 서술자 명령어 중 하나인 <Replace>의 접근을 정하고, <touchsensor>(㉔) 노드를 사용하여 마우스 클릭 동작을 처리하며, 이후 <Route>(㉕)를 이용해 <touchsensor> 노드와 <conditional> 노드를 연결 시켜, 마우스 클릭이라는 사용자 대화형 기능을 받았을 때, 객체 서술자 명령어의 <Replace>를 통하여 화면 전환이 일어나는 것을 구현하였다.

본 논문에서는 화면 전환이 일어났을 때 필요로 하는 영상을 서버 단에 요청을 하여야 하지만, 본 MPEG-4 system에서 제공하는 송수신 시스템은 이미 수신된 영상들에 대해서 명령을 수행하는 것으로 제작되어 있다. 즉 <Replace>를 사용하여 영상의 URL 값을 변경하여도, 서버 단에 저장된 영상이 아닌 이미 수신된 영상을 사용하게 된다. 따라서 본 실험에서는 송수신 시스템의 <Replace>명령어에 임의

적으로 수신 기능을 추가하여, 사용자가 요청을 하였을 때 영상을 전송하도록 구현을 하였다.

### 3. 실험 및 결과

본 논문에서 제안한 대용량 파노라마식 멀티미디어 콘텐츠 전송방안의 효율성을 증명하기 위한 실험은 다음과 같이 이루어졌다. 실험을 위해 사용된 파노라마 영상은 2048\*512의 해상도를 가지고 25 fps의 프레임 율(frame rate)로 30 초의 재생시간을 가지며, MPEG-4 part2: Visual<sup>[10]</sup>의 기술로 압축되었다. 그림 3의 (가) 영상은 상기에서 설명된 영상이며, 하나의 객체 서술자를 가지는 단일 대용량 파노라마식 멀티미디어 콘텐츠 영상이다. 이를 각각 400\*512의 해상도를 가지는 5 개의 영상으로 분할 한 뒤, 각각의 객체 서술자로 지정하고, 표1에서 설명된 것과 같이 장면 구성 및 사용자의 시점 이동 요구사항인 대화형 기능을 포함하는 xmt 문서를 통해, 분할된 영상을 mp4 file로 구성한다. 이는 분할된 대용량 파노라마식 멀티미디어 콘텐츠 영상을 나타내며, 사용자 대화형 기능 처리를 위한 화살표 모양의 정지 영상(JPEG 영상)을 추가로 포함한다. 각 영상 또한 25 fps의 프레임 율로 30 초의 재생시간을 가지며, 단일 파노라마 영상과 같은 MPEG-4 part2:Visual의 기술로 압축되었으며, 재생은 3장의 영상을 소비하여 1200\*512의 해상도로 나타나도록 하였다.

실험에서 사용되는 파노라마 영상 콘텐츠의 크기는 표

표 2. 실험 데이터 크기  
Table 2. experiment data size

	영상 크기	
	MPEG-4 visual 압축 영상	MP4 file
단일 파노라마 영상	18,822,828 (Bytes)	
분할 파노라마 영상	sequence 1	3,615,718 (Bytes)
	sequence 2	3,571,610 (Bytes)
	sequence 3	4,953,457 (Bytes)
	sequence 4	3,842,325 (Bytes)
	sequence 5	3,062,824 (Bytes)
	19,162,171(Bytes)	

2에서 나타난 바와 같으며, 단일 파노라마 영상은 하나의 영상이 하나의 객체 서술자로 기술되는 장면 구성을 통해 mp4 파일로 생성한 것이다.

분할 파노라마 영상은 본 논문에서 제안한 방법으로 5개의 영상을 5개의 각각의 객체 서술자로 기술하여 mp4 파일로 생성된 것이다. 표 2에 따르면 제안된 시스템의 분할 파노라마 영상 콘텐츠의 총 크기가 단일 파노라마 영상 콘텐츠 크기보다 303,541 Byte 만큼 크다. 그 이유는 제안된 시스템에서는 표 1에서 기술된 바와 같이 사용자의 시점 이동을 장면 구성 기술로 수행하기 위하여, BIFS 스트림을 추가하기 때문이다. 추가된 스트림의 크기는 전체 영상의 약 1.6% 정도를 차지하지만, 재생시간이 1시간가량이 되는 일반적인 콘텐츠를 고려하면 본 실험에서의 영상 데이터가 비교적 매우 작은 데이터 크기이며, 영상의 데이터가 많아 지더라도 BIFS 스트림의 크기는 변화가 없기 때문에, 단일 파노라마 영상과 제안된 시스템의 영상은 크기 차이가 거의 없다고 할 수 있다.

전송 효율을 측정하기 위해서, 본 논문에서는 상기에 설명된 분할 파노라마 영상 콘텐츠와 단일 파노라마 영상 콘텐츠를 MPEG-4 시스템을 기반으로 하는 전송 시스템 (server client system)을 통해서 전송하고, 두 파노라마 영상 콘텐츠의 전송 대역폭 사용량을 비교하여 실험하였다.

단일 파노라마 영상 콘텐츠와 분할 파노라마 영상 콘텐츠의 전송 시스템의 전송 대역폭 사용량의 비교 결과는 그림 4를 통해 알 수 있다. 그림 4는 대역폭 사용량의 비교를 통해서 전송의 효율성을 증명한 그래프로 제안된 시스템과 단일 파노라마 영상의 전송에서 전송이 일어난 시간 동안 사용한 대역폭을 기록하고, 반복된 실험을 통해 얻어진 값을 기록한 것이다. 단일 파노라마 영상 콘텐츠의 전송을 위해서는 약 700Kbytes/s 정도의 전송 대역폭이 필요하였고, 제안된 영상 콘텐츠의 전송에서는 초기의 데이터를 제외하고는 약 380Kbytes/s 정도의 전송 대역폭이 필요하여, 제안된 방안이 절반정도의 대역폭만 사용하기에 전송에서의 효율성을 알 수 있다. 제안된 시스템에서 초기에 대역폭 사용량이 커진 이유는 MPEG-4 시스템의 장면 구성을 표현하기 위해 초기에 많은 데이터를 사용하였기 때문이고, 이는 시간이 지나면서 사용되는 대역폭이 줄어들음을 알 수 있다.

이 실험을 통해 단일 파노라마 영상 콘텐츠의 전송보다 본 논문에서 제안된 방식의 영상인 분할 파노라마 영상 콘텐츠의 전송이 대역폭을 적게 점유한다는 것을 알 수 있다. 이는 본 논문에서 제안한 방안이 대용량 파노라마 멀티미디어 영상 콘텐츠의 전송에서 효율성을 가진다는 것을 증명한다.

표 3에서는 전송이 일어난 시간동안의 평균 대역폭 사용

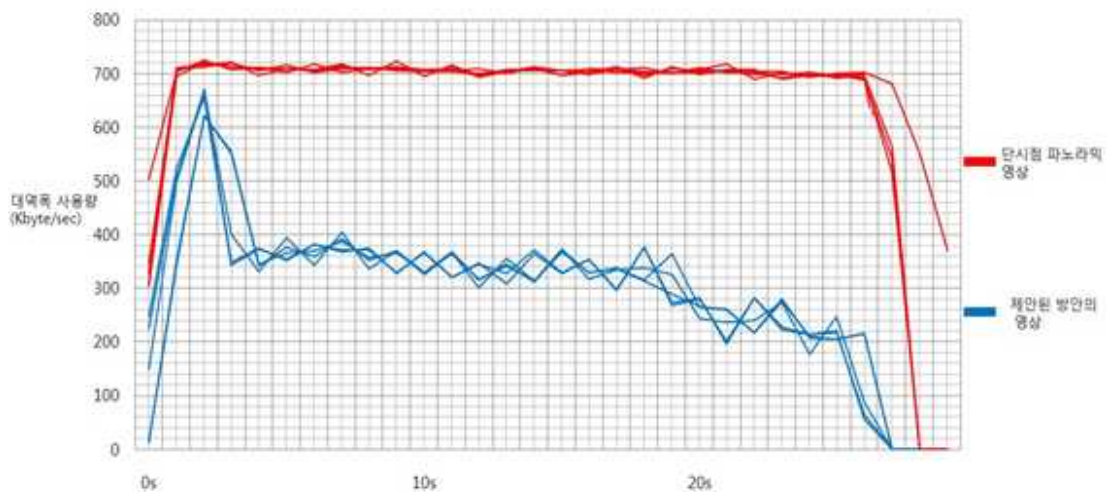


그림 4. 파노라마 영상 콘텐츠 전송 대역폭 그래프  
 Fig. 4. Graph of panoramic multimedia contents transmission bandwidth

표 3. 평균 대역폭 사용량  
Table 3. average bandwidth usage

	1회차	2회차	3회차	4회차	5회차	평균
단일 파노라마 영상 전송	675.93 (Kbyte/s)	664.73 (Kbyte/s)	640 (Kbyte/s)	640.41 (Kbyte/s)	640.85 (Kbyte/s)	684.53 (Kbyte/s)
제한된 영상 전송 (분할 파노라마 영상)	292.18 (Kbyte/s)	292.18 (Kbyte/s)	292 (Kbyte/s)	289.67 (Kbyte/s)	292.62 (Kbyte/s)	269.21 (Kbyte/s)

량을 비교해서 보여준다. 단일 파노라마 영상의 전송보다 제안된 방법으로 전송한 영상이 적은 평균 대역폭 사용을 보여, 마찬가지로 제안된 시스템의 전송 효율성을 증명한다.

## V. 결론

기존 멀티미디어 영상과 비교하여 보다 자유로운 시점과 넓은 시야를 제공하는 파노라마 영상은 그 장점으로 인하여 많은 양의 데이터를 갖고 있어, 파노라마 영상의 전송에서 기존 영상보다 많은 양의 데이터를 전송해야 한다. 하나의 카메라에서 파노라마 영상을 단일 파노라마 멀티미디어 콘텐츠로 만들어 사용자에게 전송하면, 대용량의 데이터가 전송되기 때문에 제한적인 대역폭을 가진 전송 환경에서는 의도하지 않은 시간 지연으로 버퍼링 문제가 발생한다. 또한 사용자가 모바일, 노트북 등과 같이 제한된 디스플레이 장치에서 단일 파노라마식 멀티미디어 콘텐츠를 재생하고자 하면, 파노라마 영상의 해상도가 가지는 문제로 인해 전체 영상 중 일부분을 재생하기 때문에 전송된 파노라마 영상 콘텐츠의 사용효율이 떨어지는 문제점을 갖고 있다.

본 논문은 상기에 기술한 파노라마 영상의 전송 효율과 디스플레이 장치와의 문제점을 해결하기 위하여 MPEG-4 시스템의 장면 구성 기술을 활용하여, 다수의 카메라에서 얻은 파노라마 영상을 각각의 객체로 정의하며, 분할 파노라마 동영상의 화면을 구성하고, 분할 파노라마식 멀티미디어 콘텐츠의 전송 시스템을 이용한 실험 결과를 통해 효율성을 증명하였다. 이와 같은 효율적 분할 파노라마식 멀티미디어 콘텐츠 전송 시스템은 수신 단말의 디스플레이

해상도와 전송 대역폭을 고려하여 미디어의 일부분을 전송하고, 불필요한 부분을 전송하는데 소요되는 부하를 줄일 수 있기 때문에 단일 파노라마 영상의 제한상황을 해결하기 위한 좋은 방안이라 볼 수 있다.

이는 다양한 각도의 다수 시점을 가지고 시점 이동을 실시하는 다시점 영상 콘텐츠나 파노라마 영상 콘텐츠와 같은 대용량의 멀티미디어 콘텐츠의 전송을 고려하는 서비스를 구성할 때, 본 논문의 제안을 활용하면 효율적인 전송을 이끌어 낼 수 있을 것이다. 추가로 보다 더 다양한 사용자 요구사항을 충족하기 위해, 파노라마 영상 콘텐츠의 전송 시스템에서 장면 구성 기술을 활용하여, 사용자 요구에 대응하는 대화형 기능에 대한 확장 및 보안 연구가 필요할 것으로 생각된다.

## 참고 문헌

- [1] 이인재(Injae Lee) 기명석(Myungseok Ki) 김옥중(Wook-Joong Kim) 김규현(Kyuheon Kim) "MPEG-4 기반 대화형 복수시점 영상콘텐츠 저작 시스템" 한국방송공학회, 한국방송공학회 학술발표대회 논문집 2005년도 한국방송공학회 학술대회, 2005. 11, pp. 209 ~ 212 (4pages)
- [2] King-To Ng, Shing-Chow Chan, Heung-Yeung Shum, IEEE : "Data Compression and Transmission Aspects of Panoramic Videos", IEEE Transaction on Circuit and Systems for Video Technology, VOL. 15, NO. 1, January 2005
- [3] K.H. Jang, S.K.Jung and M. L, "Constructing Cylindrical Panoramic Image using Equidistant Matching," IEE Electronics Letters, IEE, England, Vol. 35, No. 20, pp. 1715-1718, September 1999.5
- [4] <http://www.ptgrey.com/products/ladybug3/index.asp>
- [5] "ISO/IEC 14496-2:2002 - Information technology -- Coding of audio-visual objects -- Part 1: System"
- [6] F. Pereira and T. Ebrahimi, The MPEG-4 Book. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [7] Julien Signès "BINARY FORMAT FOR SCENE (BIFS): COMBINING MPEG-4 MEDIA TO BUILD RICH MULTIMEDIA SERVICES" PROCEEDINGS OF THE SPIE, SPIE, BELLINGHAM,

VA, US, vol. 3653, 25 January 1999 (1999-01-25), pages 1506-1517, XP000984179 ISSN: 0277-786X  
[8] ISO/IEC JTC1/SC29/WG11/N4091, Study of ISO/IEC 14496-1:2001/PDAM2:Extensible MPEG-4 Textual Format(XMT), (March 2001)

[9] "ISO/IEC 14496-14:2003 - Information technology -- Coding of audio-visual objects -- Part 14: MP4 file format".  
[10] "ISO/IEC 14496-2:2004 - Information technology -- Coding of audio-visual objects -- Part 2: Visual"

---

저 자 소 개



김 병 철

- 2008년 2월 : 경희대학교 전자공학과 학사
- 2010년 2월 : 경희대학교 전자전파공학과 석사
- 2010년 3월 ~ 현재 : 경희대학교 전자전파공학과 박사과정
- 주관심분야 : 멀티미디어, MPEG 장면 구성 기술



이 건 희

- 2008년 2월 : 경희대학교 전자공학과 학사
- 2010년 2월 : 경희대학교 전자전파공학과 석사
- 주관심분야 : 멀티미디어, MPEG-2/4 systems



이 인 재

- 1999년 2월 : 성균관대학교 전자공학과 (공학사)
- 2001년 2월 : 성균관대학교 전기전자및컴퓨터공학부 (공학석사)
- 2001년 1월 ~ 현재 : 한국전자통신연구원 선임연구원
- 주관심분야 : 리치미디어, 영상처리, 컴퓨터비전, UI/UX



김 규 현

- 1989년 2월 : 한양대학교 전자공학과 공학사
- 1992년 9월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학석사
- 1996년 7월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학박사
- 1996년 ~ 1997년 : 영국 University of Sheffield, Research Fellow
- 1997년 ~ 2006년 : 한국전자통신연구원 대화형미디어연구팀장
- 2006년 ~ 현재 : 경희대학교 전자정보대학 교수
- 주관심분야 : 영상처리, 멀티미디어통신, 디지털 대화형 방송