

파노라마 영상에서 효율적인 시점탐색을 위한 시공간 비디오 스트림 예측 필터 설계 방법에 관한 연구

The Design of Spatial-Temporal Prediction Filter for saving resources on the view navigation of a panoramic video service

석주명*, 조용우**

Joo-Myoung Seok*, Yong-Woo Cho**

요 약

사람의 시야각 이상으로 넓게 영상을 채워줌으로써 현장감을 제공하는 파노라마 영상은 넓은 해상도로 인하여 제한된 재생환경과 대역폭의 한계로 파노라마 영상의 넓은 시점을 한 번에 보기 어렵기 때문에 사용자가 시점을 선택하여 시청하는 방법이 필요로 한다. 그러나 시청시점을 선택하기 위해 빈번하게 사용하는 시점탐색 시에 재생하지 않는 시점의 영상데이터까지 전송함에 따라 발생하는 자원낭비를 최소화하기 위하여 본 논문에서는 시점탐색 시 필요한 영상데이터만 전송할 수 있도록 시점탐색 방향과 속도에 기반한 시공간예측 필터방법을 제안한다. 본 제안방법으로 높은 대역폭 사용이 불가피한 인터랙티브 파노라마 영상 스트리밍 서비스에 있어 기존 대비 6%에서 37% 정도의 대역폭 사용을 줄일 수 있다.

Abstract

A panoramic video which supports to make viewers feel an immersion through fitting to a wide field of view (FOV) larger than the human visual angle needs an interactive viewing method such as selecting targeted view point among widely viewing points of a panoramic video because it difficult to simultaneously view a whole panoramic video due to a limited viewing environment and bandwidth. When a user officially uses a view navigation in order to select a view point, it happens waste of resources such as bandwidth owing to the transmitted video data of unnecessary view points. Therefore, this paper proposes the spatial-temporal prediction filter (STPF) which is based on the direction and velocity of the view navigation for transmitting only the necessary video data. As a result of simulation, STPF reduces bitrate saving rates by from 6% to 37% compared to conventional methods in the interactive panoramic video streaming service required high bandwidth

Keywords : Panoramic video, Streaming, Visual perception, SVC

I. 서 론

대형 TV의 일반화, 3DTV의 등장으로 좀 더 사실적이고 실감 있는 방송에 대한 소비자들의 욕구가 증

* 한국전자통신연구원(Realistic broadcasting research Dept., Electronics and Telecommunication Research Institute(ETRI))

** 경희대학교 (Kyunghee University (KHUI))

· 제1저자 (First Author) : 석주명(Joo Myoung Seok, tel : +82-42-860-1216, email : jmseok@etri.re.kr)

· 접수일자 : 2013년 9월 27일 · 심사(수정)일자 : 2013년 9월 27일 (수정일자 : 2013년 12월 13일) · 게재일자 : 2013년 12월 30일
<http://dx.doi.org/10.12673/jkoni.2013.17.6.757>

대 되고 있는 가운데, 최근 실감 미디어 연구 분야는 홀로그래피 (holography), 양안시점 시스템 (stereoscopic system), 다시점 비디오 (multi-view video) 등과 몰입감 증대와 관련한 비디오 기술에 대한 연구가 활발히 진행되고 있다[1]. 특히 사람이 보는 시야각만큼 넓은 화각에 고화질 영상을 채워줌으로써 실제 현장 속에 있는 것 같은 느낌을 주는 파노라마 영상은 입체감을 제공하는 3D와 고해상도로 사실감을 제공하는 UHD (Ultra High Definition)와는 또 다른 현장감을 제공하는 특징을 가진다. 또한 디스플레이 기술 발전으로 일반적 디스플레이 재생환경인 단일화면을 넘어 멀티화면기반의 대형 스크린을 통해 넓은 화각을 한 번에 볼 수 있는 시청환경이 도래할 것으로 예측된다[2].

한편 그림 1과 같이 에어버스가 제시한 2050년 미래 비행기 콘셉트를 살펴보면 투명 동체와 투명 와이드 화면에 활용방안에 대하여 설명하고 있다[3]. 만약 파노라마 영상을 비행기 창문이나 동체 내벽에 외부 전경을 투사한다면, 투명 동체의 효과를 볼 수 있을 만큼 파노라마 영상에 대한 활용이 클 수 있을 것으로 예측된다. 그러나 아직 시청 재생환경과 서비스 대역폭도 HD 환경에 국한되어 있기 때문에 대화각, 대용량 특징을 갖는 파노라마 영상을 시청하기 위해서는 넓은 시청시점을 사용자가 자유롭게 선택하며 볼 수 있는 인터랙션 서비스 방법이 필요하다. 이러한 인터랙티브 서비스 방법은 기존 IPTV, 인터넷 등 커뮤니케이션 네트워크를 기반으로 사용자에게 랜덤 액세스, 고속재생, 고속감기 등 시간적 트릭 모드에 대하여 주로 연구되어 왔다[4]. 그러나, 파노라마 영상과 같이 광화각 영상은 제한된 시청 재생환경을 고려하여 시점을 선택하는 방법[5, 6]에 대해서는 일부 연구되어 왔으나 시점탐색, 시점확대 등의 공간트릭모드에 대하여 끊임 없는 인터랙션 (spatial interaction) 제공에 관한 연구는 미미한 상황이다.

따라서 제한된 스크린 환경에서 효율적인 공간 시청이 가능하도록 시청시점을 선택할 수 있는 시점탐색 방법이 필요하지만, 기존 파노라마 이미지 서비스 방법처럼 전체 영상을 보내고 단말과 사용자간에 시점탐색 방법을 제공하면 동영상 특성상 많은 대역폭과 단말기의 자원 낭비가 발생하는 문제가 발생한다[7].



그림 1. 가변적 시점탐색 속도에 따른 시공간 예측 시청환경

Fig. 1. Concept of Airbus 2050: future cabin

또 하나의 방법으로는 썸네일처럼 저화질 탐색용 영상을 추가 제공할 수 있으나, 불필요한 데이터 사용으로 대역폭 낭비가 발생하고, 탐색전용 영상의 화질이 낮아 품질저하를 느끼는 문제가 있으므로 원래 영상을 필요한 영역별로 전송하면서 자원낭비는 최소화하고 체감품질을 유지할 수 방법에 대한 연구가 필요하다[8]. 이러한 문제점을 해결하기 위하여 Seok[5]은 시점변화 속도에 따라 달라지는 시선인식 능력을 고려하여 인식가능한 계층품질만 전송 재생하는 방법으로 비디오 영상의 체감품질은 동일하게 유지하면서 대역폭 사용량은 최소화할 수 있는 시각 인식 중심의 인터랙티브 파노라마 탐색 방법을 제안하였다. 즉, 급격한 영상 변화로 인한 시공간 정보변화를 모두 인식할 수 없으므로 다중 품질을 갖는 계층화 코딩방법을 활용하여 인식하지 못하는 고품질 데이터는 전송하지 않도록 하는 효율적인 자원사용에 관하여 연구하였다. 그러나 이전 연구에서는 탐색 속도에 따른 품질 인식변화만을 고려하고 시점탐색 방향과 속도, 탐색시간 등을 고려하지 않아서 재생할 필요 없는 영상 데이터까지도 전송하게 되어 여전히 자원낭비가 나타나는 문제점을 가지고 있다[9].

이에 본 논문에서는 시점 탐색 시 품질인식 능력 뿐만 아니라 탐색진행방향과 속도를 따른 시공간적 변화를 고려하여 필요한 영상 데이터만 선택 전송하여 자원 사용의 효율을 높이는 시공간 예측 필터 방법을 제안하고자 한다.

이를 위한 본 논문의 2장에서는 파노라마 비디오의 특성 및 기존 시각인식기반 탐색 방법에 대하여 설명을 하고 3장에서 제안하는 영상 데이터의 시공간 예측 필터에 대하여 설명한다. 4장에서는 제안하

는 방법에 대한 실험을 통해 성능을 분석하며 5장에서 결론을 맺는 것으로 구성한다.

II. 배경

2-1 시각 능력

사람의 눈의 특성으로 빛을 감지하는 간상체(rods)와 색을 감지하는 원추체 (cones) 로 구성된 망막(Retina) 과 중심와 (fovea) 에 대한 이해가 필요하다. 초점에 중요한 역할을 하는 중심와는 상세도가 높은 영상이 생성되는 영역으로 미세 인식 능력의 한계는 시력 (Visual acuity) $1/120^\circ$ 이하는 인식하기 힘들며, 최상의 화질을 만들어내는 시야각은 $4\sim 5^\circ$, 적정한 시청거리에서 투영되는 이미지 평면 상 1° 이하 움직임의 변화는 감지하기 어렵다고 한다[10, 11].

본 논문에서도 이전 연구결과[5]와 같이 선명도를 중심으로 선명도 대비 민감도 (Contrast sensitivity function, CSF) 에 따라 진행하고자 한다[10, 11]. 대비 민감도 함수는 1° 시야각에서 시각이 인식할 수 있는 한 주기의 수로 표현하며 cpd (cycles per degree) 단위를 가진다. 선명도 대비 민감도는 현재 HVS (Human Visual System) 의 기초 모델로 많이 활용되고 있으며, 그림 2와 같이 움직임의 속도가 증가하면 인식할 수 있는 인식능력은 점차 떨어지는 결과를 볼 수 있다.

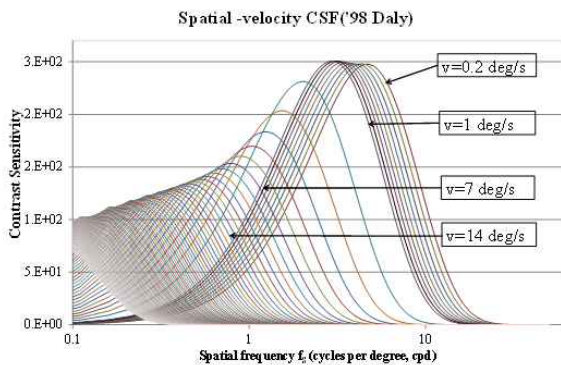


그림 2. 속도변화에 따른 CSF 결과
Fig. 2. Variance of velocity based CSF curve

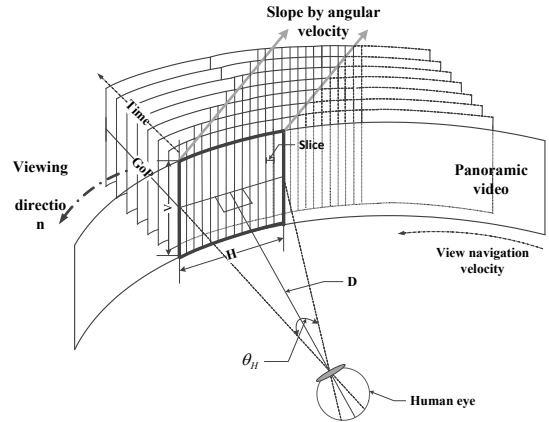


그림 3. 파노라마 영상 시청 환경
Fig. 3. Viewing environment for watching panoramic video

2-2 파노라마

파노라마의 사전적 의미를 살펴보면, 하나의 이상의 카메라로 촬영된 영상을 하나의 카메라로 찍은 것처럼 시공간적으로 이어 붙여 만든 영상으로 관람자가 마치 실경을 보는 것과 같은 느낌을 주는 영상 기술이다. 한 카메라로는 넓은 화각을 촬영하기 어렵기 때문에 멀티 카메라로 동시에 촬영하여 획득한 인접 영상간 중복 촬영 영역으로부터 공통 특징점 (Feature point) 을 추출하고, 이로부터 영상간의 기하학적 관계인 호모그래피 (Homography) 산출하여 이어 붙인다. 그 다음 단계로는 이미지 포개짐 때문에 발생하는 이질감을 완화하기 위하여 블렌딩 (Blending) 과정을 거쳐 스티칭이 완성된다[12]. 최상의 스티칭 정보를 산출하여 파노라마 영상을 생성하는 것은 중요한 연구 분야이지만, 본 논문에서는 파노라마 비디오 스트리밍 서비스 제공을 목적하므로 본 논문에서는 제외하도록 한다.

차후 wall TV 등 시청 환경이 진화를 하면 파노라마 영상 화각만큼 동시에 시청할 수 있지만 그림 3과 같이 파노라마 영상 콘텐츠를 일반적 시청환경에서 보는 시청형태가 우선적으로 도입될 것으로 예상된다. 따라서 제한된 스크린에서 파노라마 영상을 시청하기 위해서는 사용자가 원하는 시청시점을 선택 시청할 수 있도록 시점탐색 방법이 필요로 하며, 이 기능의 발생 빈도가 높다는 특징을 갖는다. 이러한 시점탐색은 탐색방향에 따라 파노라마 영상의 영역이

새롭게 나타나거나 재생되고 있는 영상 영역이 사라지는 상황이 발생하게 되는데, 전체 영상을 모두 보내는 자원 낭비를 최소화하기 위해서는 넓은 파노라마 영상을 의미가 있는 영역단위로 분할하여 부복호화 및 전송하는 방법이 필요하다. 예를 들어 전방위 탐색 방법을 제공하기 위해서는 파노라마 영상을 바둑판 형태로 나누어 분할 코딩을 한 후, 재생에 필요 부분만을 골라 전송할 수 있다면, 재생 스크린에 필요한 영역만 조합하므로 자원 낭비를 최소화할 수 있을 것이다. 그러나 시점탐색은 필요 영역이 시점탐색 속도와 방향에 따라 지속적으로 변하는 특징이 있으므로, 공간적 측면뿐만 아니라 공간적인 변화에 따른 시간적인 측면에서도 필요한 영상 데이터를 선택하여야 자원 낭비를 최소화 할 수 있다.

이에 본 논문에서 전송에 필요한 파노라마 영상을 골라내기 위하여 시공간예측필터 (Spatio-Temporal Prediction Filter, STPF) 를 제안하고자 한다. 제안방법의 이해와 성능분석 편의를 위하여, 가장 쉽게 접할 수 있는 파노라마 영상인 하나 이상의 표준크기 영상을 수평으로 이어 붙인 파노라마 영상을 기준으로 설명하고자 한다. 그러므로 시점탐색 방향은 수평이동에 한하여 발생한다고 한정할 수 있다.

제안방법의 설명에 앞서 자원 낭비 여부를 결정하기 위해서는 사용자의 시청환경을 정하여야 좀 더 정확하고 사용자가 만족할 수 있는 기준을 가지고 실험할 수 있기 때문에 식 (1)을 이용하여 최적의 시청환경을 도출한다. 스크린 수평 크기(H), 시청거리는 스크린 수직 크기(V) 의 3배 거리(d) 로 한다[9]. 참고로 HD의 경우의 시청각은 30~33° 정도이다.

시점탐색방향은 좌우이동에 한정하였으므로 그림 3과 같이 부복호화 방법은 시점탐색 방향과 직교된 수직방향으로 영상을 분할조합하여 선택영역을 조합 재생할 수 있도록 정의하였다.

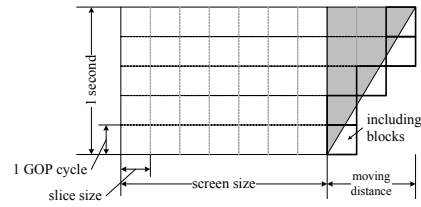
$$\theta_H = \left[\theta_{H,radians} = 2\arctan\left(\frac{H}{2dV}\right) = 2\arctan\left(\frac{H}{2D}\right) \frac{180}{\pi} \right] \quad (1)$$

III. 제안방법

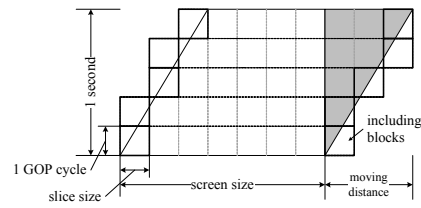
그림 4에서 볼 수 있듯이 한정된 스크린에 자원 낭

비 없이 파노라마 영상을 탐색하면서 시청하기 위해서는 그림 3 와 같이 시점탐색 속도 v_H 에 따라 시청 시점 공간영역이 변하게 되는데 이용자가 시청하는데 불편함이 없이 전송 대역폭을 절감하기 위해서는 시점탐색 방향 및 속도에 따라 전송받아야 할 영상의 시공간 영역을 예측하여 필요한 영상 데이터만 전송하고 재생할 수 있어야 한다.

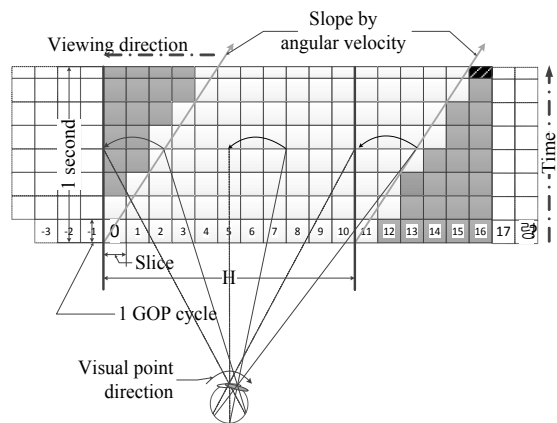
이를 위하여 제안한 STPF는 그림 4와 같이 Half-STPF와 Full-STPF 모드를 가진다. 설명을 위하여 시점탐색 방향이 오른쪽 방향으로 이동한다고 가정한다면, 그림 4에서 보는 바와 같이 오른쪽의 회색 영역은 시점탐색으로 인해 새롭게 나타나는 시청시점 영역을 의미한다.



(a) Half-STPF mode



(b) Full-STPF mode



(c) the design of STPF

그림 4. 시공간 예측 필터 (STPF) 개념
Fig. 4. Basic concept of SPTF

이전 연구에서는 시점탐색 시 시선 인식 능력에 따라 인식 가능한 품질 계층만을 선택하기 때문에 시점 방향과 속도를 고려하지 않았으나, 제안한 STPF는, 그림 4 (a) 와 (b) 에서는 시선인식을 고려할 뿐만 아니라 시점탐색 속도와 방향에 따라 재생할 필요가 없는 영상 데이터까지도 배제하여 스트리밍할 수 있도록 한다. Half-STPF와 Full-STPF의 차이점은 탐색 방향에 따른 탐색시간 및 속도에 따라 불필요할 것으로 예측되는 영상데이터를 제거함에 있어, 새롭게 등장하는 입력영상영역만을 고려하여 전송할지, 사라지는 영상의 시공간영역까지도 고려하여 전송하는 것인지가 다르다.

그림 4 (c) 를 통해 좀 더 자세하게 설명하면, 그림 5의 한 블록은 재생 시 독립된 복호화가 가능하도록 슬라이스와 1 GOP (Group of Pictures) 크기인 [S, G] 단위로 구성하도록 하며 필요한 [S, G]의 총 개수는 식 (2)에 의하여 산출된다.

$$\begin{aligned} [S_m \times G_n] &= m \times n; \\ m &= 1, 2, \dots, h; n = 1, 2, \dots, [G_{\max}]; \\ G_{\max} &= \frac{x \text{ fps} (\text{frame per second})}{y \text{ fpg} (\text{frame per GOP})} \end{aligned} \quad (2)$$

파노라마 영상은 시점탐색 방향에 직교하는 수직 방향으로 나누어지며, 각각으로 나누어진 분할 영상을 슬라이스라고 정한다. 만약 시청화면을 7개의 슬라이스 (m)로 나누고, 1초 동안 필요한 프레임이 30 프레임이면, GOP 구성은 5 프레임주기를 갖는 슬라이스 당 $G_{\max} = 6 \text{ GPS}$ (GOP per second) 인 n 를 갖게 되므로 $[S_m \times G_n] = 42$ 개임을 알 수 있다. G_{\max} 산출 시 GOP 주기에 따라 나누어 떨어지지 않는 경우에는 G_{\max} 는 소숫점으로 명시하고, 그 이전에는 G_n 은 1 씩 증가하는 것으로 정하도록 한다.

앞서 설명한 설명한 바와 같이 Half-STPF 방법이 Full-STPF 보다 움직이는 방향만을 예측하여 불필요한 영상 데이터를 제거하기 때문에 전송할 대상 영상 데이터가 Full-STPF보다 크다는 단점이 있지만 시점탐색속도 v_H 가 급속히 감소하는 경우에 사라지는 영상의 영역 데이터가 충분히 존재하기 때문에 인터랙티브 지연으로 인하여 재생이 끊기는 현상은 최소화할 수 있다는 장점을 가진다. 따라서 STPF의 종류

는 전송 및 재생환경을 고려하여 선정할 수 있도록 하였다.

따라서 본 논문에서 제안하는 STPF 는 그림 4 (c) 와 같이 v_H 에 의하여 새로운 시점이 입력되는 입력시점 영역과 이미 재생되어 사라지는 출력시점 영역을 GOP ($G_0 < G_0 \leq G_{\max}$) 주기별로 재생에 필요한 블록 범위를 기준으로 필터링한다. GOP 주기별 블록의 범위가 정해지면 재생에 필요한 필수 블록들을 계산할 수 있다. 이와 같이 [S, G] 집합을 IPB (Indispensable play blocks)라 정의한다.

식 (3)는 V_H 에 의하여 새로운 시점이 입력되는 입력시점 영역만을 필터링하는 half-SPTF 를 설명한다.

$$\begin{aligned} STPF_{half} &= \sum_{n=1}^{G_{\max}} IPB_n [S_{begin}, S_{end}] \left\{ \begin{aligned} begin &= 0 \\ end &= \frac{H}{H_s} + \left\lfloor \frac{MD_G \times n}{H_s} \right\rfloor \end{aligned} \right. \quad (3) \end{aligned}$$

$$\begin{aligned} STPF_{full} &= \sum_{n=1}^{G_{\max}} IPB_n [S_{begin}, S_{end}] \left\{ \begin{aligned} begin &= 0 + \left\lfloor \frac{MD_G \times n}{H_s} \right\rfloor \\ end &= \frac{H}{H_s} + \left\lfloor \frac{MD_G \times n}{H_s} \right\rfloor \end{aligned} \right. \quad (4) \end{aligned}$$

여기서 H는 스크린 크기 (픽셀단위) 을 말하여 H_s 는 슬라이스 크기 (픽셀단위)를 의미한다. $IPB_n [S_{begin}, S_{end}]$ 의 S_{begin} 와 S_{end} 는 n 번째 GOP 주기에서 재생에 사용되는 블록의 시작위치 블록의 끝 위치를 나타낸다. 여기서 MD_G (moving distance per GOP) 는 1 초당 움직인 픽셀거리를 의미하며, G_{\max} 까지 이동 예측거리만큼 슬라이스 범위를 정할 수 있다. $\lfloor \rfloor$ 는 반 내림 (round down) 을 의미한다. 만약 슬라이스 번호 시작이 S_1 인 경우라면 반올림 (roundup) 으로 처리하도록 한다. 이러한 반올림과 반내림이 필요한 이유는 해당 IPB 블록이 재생하는데 계속 사용해야 한다는 의미이므로 필요한 블록으로 처리하기 위함이다. 여기서 $IPB_n [S_{begin}, S_{end}]$ 에 해당하는 예측 시간의 범위는 송수신, 복호화, 버퍼링 등 일정지연시간 (propagation delay) 고려하여 정하도록 한다.

식 (4)은 v_H 에 의하여 새로운 시점이 입력되는 입력시점 영역과 현재 시점이 사라지는 출력시점 영역을 모두 고려한 Full-SPTF 방법에 대하여 수식화 하였다.

IV. 실험 및 성능분석

본 논문에서 사용하는 파노라마 실험영상은 좀 더 사실적인 파노라마 영상을 실험하기 위하여 1024x512 해상도를 가진 파노라마 영상인 랜드투어와 좀 더 움직임이 많은 뉴욕시티투어[13, 14] 영상을 수평방향으로 반복적으로 스티칭을 하여 그림 5와 같이 4096x512 해상도로 확장하여 만들었다. 실험환경은 그림 3를 고려하여 시청환경은 실험영상을 고려하여 스크린 크기를 $H=704$ pixels, $V=512$ pixels로 영상의 수직높이와 동일하도록 가정하였으며, 수평해상도는 수직해상도의 비율에 맞게 설정하였다. 또한, 시청거리는 기존 연구와 비교분석을 위하여 수직해상도 높이의 3배 ($d=3$) 로 정함에 따라 시야각은 식 (1)에 의해 약 27° 정도로 산출된다. 파노라마 영상의 인코딩 방법도 이전 연구와의 비교 분석을 위하여 동일한 조건인 3계층 공간 스케일러빌리티 (spatial scalability, $L=0:4096 \times 512$, $L=1:2048 \times 256$, $L=2:1024 \times 128$) 로 인코딩을 하였다. 인코딩 시 GOP 구성은 IPPP 구조를 가지며 주기별 성능을 분석하기 위하여 GOP=4 프레임 (frames) 과 일반적으로 영상 실험시 정의하는 GOP=8, 현재 IPTV에서 사용하는 구조인 GOP=15 프레임 형태를 기준으로 인코딩하였다. 인코딩 시 양방향을 참조하여 압축률을 높이는 B 프레임 구조를 배제하는 이유는 디코딩 순서와 재생 순서가 바뀌게 되어 채널변경 지연 (zapping delay time) 와 버퍼 크기가 증가하는 원인이 되고, 실제 현장에서도 많이 사용하지 않으므로 본 논문에서도 제외하기로 한다. 슬라이스의 크기는 인코딩 최소단위인 매크로 블록의 최소 크기와 3 계층의 연관성을 고려하여 64 픽셀로 정하였다. 실험영상은 10 초 (300 프레임), 고품질 (QP=26) 로 정하여 JSVM 9.11[15]으로 인코딩 하였다.

그림 5는 실험 파노라마 영상 중 뉴욕 시티 투어 영상에 대하여 앞서 정의한 여러 인코딩 방법에 따라 인코딩된 비트율 변화를 보여준다. 시점탐색으로 인하여 시각인식능력에 적응적으로 대체하기 위하여 사용되는 계층화 인코딩 SVC[15]는 계층별로 AVC 사이멀캐스트 (simulcast) 방식으로 인코딩한 비트율보다는 7~12% 정도로 적기 때문에 본 논문에서는 SVC를 기준으로 제안 방법의 성능을 분석한다. 또한, GOP 주기

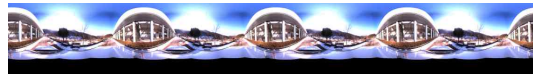
와 슬라이스 크기가 커질수록 참조관계가 많아져서 인코딩 비트율이 낮아지는데, GOP 주기가 8, 15 프레임 일 때는 GOP 주기가 4일 때 보다 각각 약 10.5%, 15.5% 정도 줄어들고, 64 픽셀인 경우가 스크린 크기인 704 픽셀보다는 약 4% 정도 인코딩 비트율이 증가한다. 슬라이스 단위는 독립적으로 디코딩이 되어야 하므로 참조관계가 범위가 줄어서 압축율이 떨어지게 된다.

이와 같이 실험 시나리오는 이전 연구 결과와의 비교를 위하여 $1^\circ/s \leq v_H \leq 26^\circ/s$ 로 10초 중 [0~4]초는 완만한 가속, [5~7] 초는 구간은 등속, [8~10]초는 급가속 구간으로 동일하게 정하였다. 이전 연구 방법과 같이 시점탐색 속도가 3%/s 이상이 되면 인식 가능한 영상품질이 줄어들기 때문에 L=1 비디오 품질을 제공하여도 사용자는 품질변화를 인식하지 못하는 연구결과 내에서 효율을 비교하기 위하여 비트절약율 (bit saving rate, BSR) 로 설명하였다.

표 1는 제안한 STPF에 의한 결과에 대한 이해를 돕기 위하여 v_H 가 12%/s일 때 초당 필요한 영상 데이터를 보여준다.



(a)뉴욕시티투어파노라마 영상



(b)랜드투어 파노라마 영상
그림 5. 파노라마 실험 영상

Fig. 5. test sequence of panoramic video

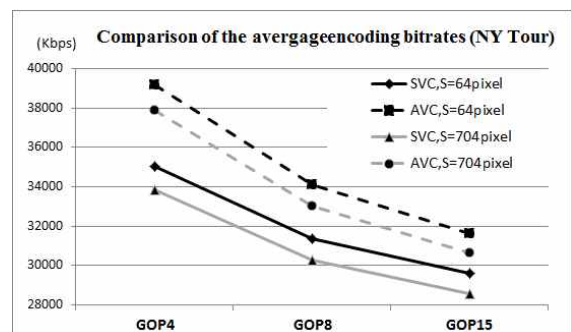


그림 6. 뉴욕시티투어 영상 인코딩 비트율 비교

Fig. 6. Comparison of the average encoding bitrates of New York city tour sequence

표 1. $v_H = 12^\circ/s$ 시 STPF 실험 환경 및 결과

Table 1. Experiment result of STPF at $v_H = 12^\circ/s$

H (pixels)	H_S (pixels)	MD_P (pixels/deg)	v_H (deg/s)		
704	64	27.27116	12		
	GOP4	GOP8	GOP15		
MD_G	43.63386	87.26771	163.627		
G_{max}	7.5	3.75	2		
구분	종류	Half-STPF		Full-STPF	
	G_n	S_{begin}	S_{end}	S_{begin}	S_{end}
GOP4	7.5	0	16	4	16
	7	0	15	4	15
	6	0	15	3	15
	5	0	14	2	14
	4	0	13	2	13
	3	0	13	1	13
	2	0	12	0	12
	1	0	11	0	11
GOP8	3.75	0	16	4	16
	3	0	15	2	15
	2	0	13	1	13
	1	0	12	0	12
GOP15	2	0	16	2	16
	1	0	13	0	13

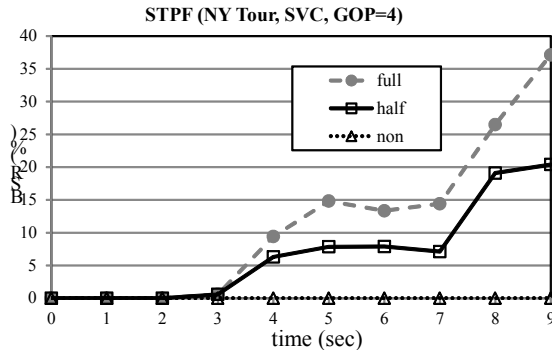


그림 7. mode 별 STPF 성능 분석

Fig. 7. STPF performance for each modes

그림 7과 같이 제안하는 STPF를 사용한 경우는 사용하지 않았을 때보다 6~15% (Half-STPF), 10~37% (Full-STPF) 정도의 리소스 낭비를 줄일 수 있다. 참고로 v_H 가 $7^\circ/s$ 이상이 되면 비디오 인식 품질은 L=2로 떨어지게 되므로 $12^\circ/s$ 와 $26^\circ/s$ 일 때의 비디오 품질은 L=2 인 환경이며, v_H 가 높을수록 동일 슬라이스 위치에 필요한 G_n 가 적기 때문에 제안방법의 비트절약율은 더욱 커지는 결과를 가지게 된다. 또한, GOP 주기가 클수록 GOP 내 프레임간의 참조관계로

인하여 GOP 전체가 필요하므로 비트절약율이 떨어지는 결과를 가진다. 이는 시점탐색속도에 변화에 따른 시간적 관계가 발생하기 때문이다. 부복호화 방법에 따라 STPF에 의한 비트절약율의 정도는 다른 결과를 보일 수 있으나, 앞선 실험결과와 같이 기존 결과 대비 자원낭비가 최소화될 수 있다. 또한 파노라마 영상은 특성상 대용량 영상이므로 적은 비트절약율이라도 대역폭 절감의 실제량은 클 수 있으므로 시청환경과 파노라마 영상에 따라 효과는 더욱 커질 수도 있다.

V. 결 론

최근 실감방송 서비스의 진화가 급속화 되고 있는 상황에서 차후 중요한 실감기술로 자리매김할 것으로 예상하는 파노라마 영상 서비스를 제공함에 있어서 기존의 제한된 스크린 크기와 대역폭의 한계를 극복하기 위하여 사용자가 능동적으로 공간을 선택할 수 있는 인터랙티브 파노라마 영상 스트리밍 방법이 우선적으로 도입될 것으로 예상된다. 그러나 사용자가 공간을 선택하기 위하여 넓은 화각과 해상도를 갖는 파노라마 영상을 효율적으로 스트리밍 하기 위하여 전체 영상을 수신 받지 않고, 재생에 필요한 데이터만을 수신받기 위한 방법이 필요하며, 이를 위하여 본 논문에서는 재생해 필요한 데이터만을 예측하여 골라 전송할 수 있는 시공간 필터방법 (STPF) 을 제안하였다. 이를 통하여 고 대역폭 사용이 불가피한 파노라마 비디오 스트리밍 서비스의 사용 대역폭을 기존 대비 6~37% 정도를 줄일 수 있기 때문에 파노라마 특성상 발행하는 고 대역폭 사용으로 인하여 파노라마 영상 서비스 확산 문제를 해결할 수 있음을 확인 할 수 있었다.

감사의 글

본 논문은 한국전자통신연구원에서 지원하는 “휴먼융합형 파노라마 기술 개발” 과제 결과임.

Reference

- [1] O. Scheer et al., "Ultrahigh-Resolution Panoramic Imaging for Format-Agnostic Video Production," *Proc. IEEE*, vol. 101, no.1, pp. 99-114, Jan. 2013.
- [2] <http://immersivemedia.com/>
- [3] <http://www.airbus.com/innovation/future-by-airbus/concept-planes/the-airbus-concept-cabin/>
- [4] I. Kopilovic, and M. Wagner, "A benchmark for fast channel change in IPTV," in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, pp. 1-7, Apr. 2008.
- [5] H. Kimata, K. Fulazawa, and N. Matsuura, "Partial delivery method with multi-bitrates and resolutions for interactive panoramic video streaming system," *ICCE, 2011 IEEE Int. Conf. on*, vol. 4, no. 1, pp. 891-892, Jan. 2011.
- [6] M. Makar et al., "Real-time video streaming with interactive region of-interest," *Image Processing (IPCP) 2010 17th IEEE Int. Conf. on*, pp.1558-1565, Oct. 2007.
- [7] E. Kurutepe, M.R. Civanlar, and A.M. Tekalp, "Client-driven selective streaming of multi-view video for interactive 3DTV," *IEEE Trans. Circuits System Video Technology*, vol. 17, no. 11, pp. 1558-1565, Oct. 2007.
- [8] C.Y. Lee, C.K. Hong, and K.Y. Lee, "Reducing Channel Zapping Time in IPTV Based on User's Selection Behaviors," *IEEE Trans. on Broadcasting*, vol. 56, no. 3, Sep. 2010, pp. 321-330.
- [9] J.M. Seok et al., "A visual perception based view navigation trick mode in the panoramic video streaming service," *IEICE Trans. on Comm.*, vol. E94-B, no. 12, pp. 3631-3634, Dec. 2011.
- [10] J. You et al., "Visual Contrast Sensitivity Guided Video Quality Assessment," *ICME, 2012 IEEE Int. Conf. on*, Jul. 2012, pp. 824-829.
- [11] S. Daly, "Engineering observations from spatiotemporal and spatiotemporal visual models," *Proc. SPIE Human Vision and Electron. Imaging*, vol. 3299, pp. 180-191, Jan. 1998.
- [12] R. Szeliski, "Image Alignment and Stitching: A Tutorial," *Microsoft Research, Technical Report MSR-TR-2004-92*, <http://research.microsoft.com>, Dec. 2006.
- [13] <http://media.xiph.org/video/derf/>
- [14] Drive the City Streets of NYC, "<http://www.immersivemedia.com/demos/index.php>"
- [15] Text of ISO/IEC 14496-10:2005/FDAM 3 Scalable Video Coding, Joint Video Team (JVT) of ISO-IEC MPEG & ITU-T VCEG, Lausanne, N9197, Sep. 2007.

석 주 명 (Joo-Myoung Seok)



1999년 2월 : 경희대학교 전자
공학과(공학석사)
2011년 8월 : 경희대학교 전자
공학과(공학박사)
1999년 4월~현재 : 한국전자통신
연구원 실감방송미디어연구부
선임연구원 재직

2009년 3월~7월 : 방송통신위원회 방송통신 융합콘텐츠
TF팀 파견

2009년 12월 : 방송통신위원회 위원장 표창

관심분야 : 영상처리, 방송통신미디어, 비디오 스트리밍,
파노라마, QoS

조 용 우 (Yong-Woo Cho)



2010년 2월 : 경희대학교 전자
공학과(공학사)
2012년 2월 : 경희대학교 전자
공학과(공학석사)
2012년 8월~12월 : 한국전자통신
연구원 위촉연구원 재직
2012년 8월~현재: 경희대학교 전자

공학과 박사과정 재학

관심분야 : 영상처리, 무선통신, SVC,, 비디오 스트리밍,
파노라마, QoS