고품질 휴먼융합형 파노라마 서비스

이인재 | 조용주 | 기명석 | 임성용 | 이희경 | 차지훈
한국전자통신연구원

요 약

근래 다양한 형태의 멀티미디어 콘텐츠가 광범위하게 소비됨에 따라 단방향의 평면적 콘텐츠가 제공하는 시점의 한계에서 벗어나 사용자에게 보다 다양한 시점의 영상을 제공함으로써 사용자가 원하는 시점을 선택할 수 있는 실감형 콘텐츠에 대한 욕구가 증가하고 있다. 이러한 실감형 콘텐츠 중 하나인 파노라마 영상은 다수의 카메라를 통해 획득된 영상들을 하나의 영상으로 이어 붙여 넓은 시야의 영상을 제공함으로써 증대된 현장감과 몰입감을 제공하며 사용자 시점의 자유도를 증가시킬 수 있다. 사용자에게 파노라마 서비스를 제공할 때 마우스, 리모콘과 같은 정형적인 인터페이스가 아니라 사용자의 선호도에 따라 손쉽게 파노라마 공간을 제어할 수 있는 사용자 친화적인 인터페이스 기술이 함께 제공된다면 사용자가 느낄 수 있는 현실감과 몰입감은 배가될 것이다.

따라서 본 고에서는 현존하는 파노라마 기술 및 사용자 인터페이스 기술의 동향에 대해 알아보고, 상기 기술들의 결합을 통해 보다 사실적인 미디어 소비 환경을 제공하기 위한 사용자 친화적인 고품질의 휴먼융합형 파노라마 서비스에 대해 소개하고자 한다.

Ⅰ. 서 론

방송통신망의 고도화에 따라 HD, 3D, UHD 등 뉴미디어를 이용한 대화면 고품질 서비스에 대한 관심이 고조되고 있다. 이러한 변화의 일환으로 대형 TV의 일반화 및 3D TV의 등장과 함께 좀더 사실적이고 실감 있는 방송에 대한 소비자들의 욕구가 증대되고 있다.

파노라마 영상은 다수의 카메라를 통해 획득된 영상들을 기하학적으로 교정하고 공간적인 합성을 통해 하나의 고해상도 영상으로 변환하여 사용자에게 여러 방향의 다양한 시점을 제공하는 실감영상 기술 중 한 분야이다. 지금까지 파노라마와 같은 실감형 콘텐츠는 전시관이나 체험관 또는 영화관 등 대중을 상대로 하는 특수 분야의 시장에서 제한적으로 적용되어 왔다. 그러나 근래에는 lean-back 형태의 수동적인 소비 환경이 감성이 융합된 고품질의 미디어를 시청자와 함께 상호작용하는 능동적인 소비 환경으로 진화하고 있다. 따라서 파노라마 서비스는 인간이 볼 수 있는 시야각보다 넓은 뷰를 제공하는 파노라마 AV 기술과 함께 직관적이고 조작이 쉬운 사용자 친화적인 인터페이스 기술을 결합할 경우 3D 의 입체감, UHD의 사실감과는 또 다른 극대화된 현실감이라는 새로운 실감 요소를 제공할 수 있을 것이다.

본 고에서는 사용자의 몰입감과 현장감을 증대시켜 주기 위하여 보다 사실적인 미디어 소비 환경을 제공하기 위한 사용자 친화적인 고품질의 휴먼융합형 파노라마 기술에 대해 소개한다. 이를 위해 II장에서는 파노라마 비디오 획득, 합성 및 재생을 위한 요소기술에 대해 기술하고, III장과 IV장에서는 현존하는 파노라마 기술 및 사용자 친화적 인터페이스 기술의 동향에 대해 살펴본다. V장에서는 한국전자통신연구원에서 개발하고 있는 휴먼융합형 파노라마 시스템

을 간단히 소개하고 이를 이용한 서비스 시나리오에 대해 살펴보고자 하며, VI장에서는 결론을 맺는다.

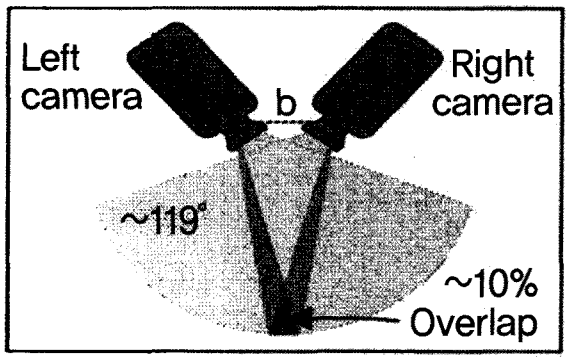
Ⅱ. 파노라마의 핵심 요소기술

파노라마 서비스를 제공하기 위해서는 크게 파노라마 영상 획득 기술, 생성 기술 및 재생 기술이 요구된다. 즉, 넓은 시야각 (Field of View) 확보를 위해서는 다수의 카메라 (또는 한 대의 카메라의 회전)를 이용한 영상획득이 선행되어야 하며, 획득된 영상은 합성(mosaicing) 기술을 이용하여 하나의 끊김 없는(seamless) 영상으로 생성되어야 한다. 이렇게 생성된 파노라마 영상은 다양한 방법으로 소비될 수 있는데, 본 고에서는 멀티 프로젝터를 이용한 대화면 재생 기술에 대해 알아본다. 멀티 프로젝터를 이용한 재생 기술은 프로젝터의 개수, 위치, 특성 및 스크린의 모양을 고려한 재생 기술이 필요하다. 따라서 본 장에서는 상기와 같은 파노라마의 세가지 핵심 요소기술에 대하여 살펴보고자 한다.

1. 영상 획득 기술

사람의 시야각은 약 120°(깊이 정보를 인지할 수 있는 수평각)이며, 사람의 시야각보다 제공되는 영상이 클 경우 현실감을 증대시킬 수 있다. 현재 카메라로 촬영할 수 있는 시야각은 인간의 시야각보다 작으며 이를 해결하기 위해 어안렌즈를 이용하여 시야각 이상의 영상을 촬영하는 방법을 사용하기도 하나 이는 영상에 큰 왜곡을 발생 시킨다. 따라서, 현실감을 증대시킬 수 있는 넓은 영상 시야각을 확보하기 위해서는 한 대의 카메라를 축(axis)을 중심으로 회전시켜 영상을 획득하거나, 다수의 카메라를 활용하여 동시에 다수의 영상을 획득하는 방법의 고화질의 파노라마 영상을 획득할 수 있는 방법이 있다.

(그림1)과 같이 다수의 카메라로 영상을 획득할 경우, 각각의 카메라에서 촬영된 영상간의 공간적 끊김이 없도록 영상의 일부분이 겹치도록 카메라를 배치하여야 한다. 이때, 카메라들은 동일한 프로젝션 센터를 가질 수 없기 때문에 시차(parallax)가 발생하게 된다. 시차는 영상합성을 통하여 해결할 수 없으므로 영상획득 시 이를 최소화 할 수 있도록 카메라 간의 간격을 줄여 시차를 최소화 하여야 한다.



(그림 1) 듀얼 카메라 셋업 예 [1]

2. 영상 생성 기술

한 대 이상의 카메라에서 획득된 다수의 영상은 영상 합성 절차를 걸쳐 끊김 없는 하나의 파노라마 영상으로 생성한다. 파노라마 영상 생성 기술은 영상간의 기하학 보정(geometric correction)을 위한 스티칭 (stitching) 기술과 컬러 보정(photometric correction)을 위한 블렌딩(blending) 기술로 구성된다.

1) 스티칭

파노라마 영상 획득을 위해 사용된 카메라들은 위치와 방향이 서로 다르기 때문에 영상들 간의 중첩부분을 찾고 기하학 보정을 해야 끊김 없는 하나의 파노라마 영상을 만들 수 있다. 이 작업을 스티칭이라 하며 본 절에서는 두 가지 기하학 보정 (위상관계 접근방법, 호모그래피)에 대하여 살펴본다. 호모그래피(homography) 기하학 보정을 위해서는 중첩된 영상에서 동일한 오브젝트를 찾는 작업(feature extraction)이 선행되어야 하고 이를 위해 Scale Invariant Feature Transformation(SIFT)이 널리 사용되고 있으며 자세한 알고리즘 설명은 [3]을 참조하면 된다.

가. 위상관계 접근 방법(Phase correlation approach)

위상관계 접근방법은 주파수 도메인(frequency domain)에서의 영상 속성을 기반으로 한다. Shift theorem에 따르면 주파수 도메인에서 두 영상의 cross power spectrum 위상은 영상간의 변화(displacement)와 동일하다. 즉, 두 영상의 2D

Fourier transform 결과를 기반으로 cross power spectrum 위상을 계산하고, 이 위상 값을 이용하여 영상간의 변화량(이동, 회전, 규모)를 계산한다[4]. 이 변화량은 두 영상의 합성에 사용된다.

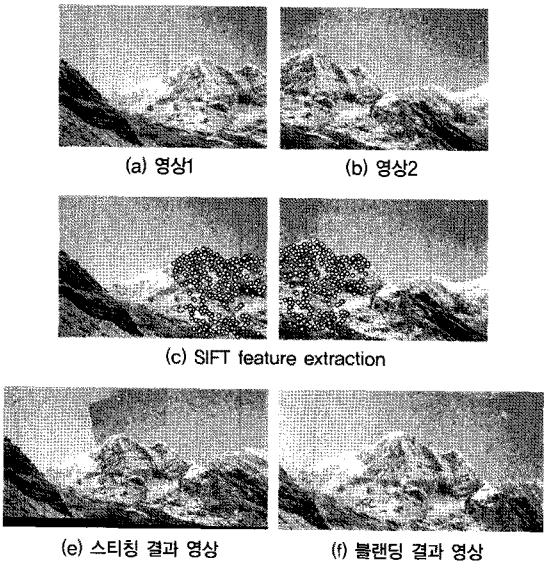
나. 호모그래피 (Homography)

동일한 광학중심(optical center)을 갖는 두 대의 카메라를 통해 획득된 영상은 이차원 투영변환(2D projective transformation)의 관계를 가지며, 이를 호모그래피라 한다. 즉, 호모그래피는 한 영상(I)의 한 점(x,y)을 다른 이미지(I')의 한 점(x' ,y')으로 사상(mapping) 한다. 두 영상간의 이동, 회전, 규모변화는 3x3 호모그래피 행렬 H로 표현되며 (1)과 같다.

$$H = \begin{pmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & m_8 \end{pmatrix} \quad (1)$$

호모그래피 행렬 H를 uniform하게 만들어 주기 위해 $m_8 = 1$이라는 제약을 적용하면, 호모그래피는 (2)와 같이 표현할 수 있다.

$$x' = \frac{m_0 x + m_1 y + m_2}{m_6 x + m_7 y + 1} \quad y' = \frac{m_3 x + m_4 y + m_5}{m_6 x + m_7 y + 1} \quad (2)$$



(그림 2) 파노라마 영상합성 절차 및 결과 예 [2]

(2)의 m_0부터 m_7은 영상 내 최소 4개 이상의 동일한 점(correspondence pairs)을 이용하여 구할 수 있다[5]. 실제로는 에러를 줄이기 위해 많은 점들을 이용하여 호모그래피를 구한다.

이렇게 구해진 호모그래피는 여전히 영상합성 시 에러를 포함할 수 있다. 따라서, (3)과 같이 영상간의 에러를 최소화하는 호모그래피를 찾는 절차가 필요하다.

$$E = \sum [I'(x',y') - I(x,y)]^2 \quad (3)$$

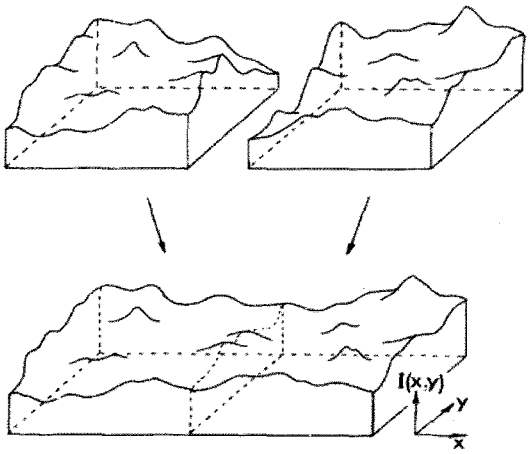
(그림 2(e))는 호모그래피를 기반으로 한 스티칭 결과를 보여주고 있다.

2) 블랜딩

두 영상 간의 기하학 보정을 통해 스티칭된 결과는 카메라 특성(조리개, 노출시간, 비네팅, 시차 등)의 변화에 의해 그림 2(e)와 같이 영상간의 계조(grey level)차가 존재하게 된다. 계조차는 소량으로도 사람의 눈에 확연하게 보여지므로 (그림 3)과 같이 블랜딩기법을 통해 계조차를 최소화 한다.

블랜딩기법으로 선형(linear) 블랜딩 또는 멀티밴드(Multi-Band) 블랜딩[2]이 사용된다.

선형 블랜딩은 일차원 무게함수(weight function)를 사용하여 두 영상이 겹치는 부분의 계조를 매끄럽게 한다. 소량의 스티칭 에러가 있을 경우, 선형 블랜딩은 고주파대의 블러링(blurring) 현상을 초래할 수 있다.



(그림 3) 두 영상의 블랜딩의 예 [6]

멀티밴드 블랜딩은 넓은 공간에 저주파 블랜딩과 좁은 공간에 고주파 블랜딩을 수행함으로써 계조차를 최소화 한다. (그림 2(f))은 멀티밴드 블랜딩이 적용된 영상합성 결과를 보여주고 있다.

3. 멀티 프로젝션 기반 재생 기술

합성된 파노라마 영상은 프로젝터를 통해 대화면 스크린에 디스플레이 된다. 두 대 이상의 프로젝터로 디스플레이를 할 경우, 프로젝션 된 영상간의 기하학 보정 및 프로젝터 간의 컬러 보정은 필수적이다.

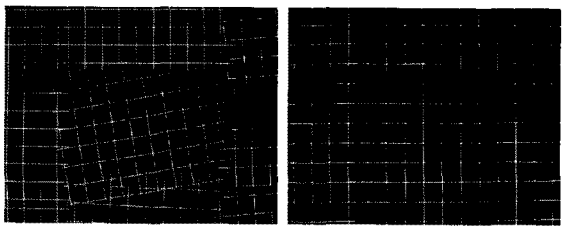
1) 멀티 프로젝션 기하학 보정

넓은 시야각을 확보한 파노라마 영상을 디스플레이 하기 위해서는 다수의 프로젝트 사용이 필수적이다. 하지만, 스크린의 모양, 프로젝터의 방향, 위치 등에 따라 파노라마 영상의 부조화 또는 왜곡이 발생하기 때문에 (그림 4)와 같이 이를 위한 보정이 선행되어야 한다. 보정방법은 호모그래피의 원리와 유사하다.

2) 멀티 프로젝터 컬러 보정

동일한 제조사, 모델, 램프 사용시간을 가지는 프로젝터들도 서로 다른 밝기와 색상을 출력한다. 특히, 멀티 프로젝터로부터 겹치는 영상부분은 겹치지 않는 부분보다 밝게 보여질 수 밖에 없다.

이러한 문제는 프로젝션된 영상들의 컬러 변화(variation)을 파악하고, 그래픽카드의 감마커브(gamma curve)를 조절하여 컬러 보정을 할 수 있다[7].



(a) 보정(calibration) 전 (b) 보정 후

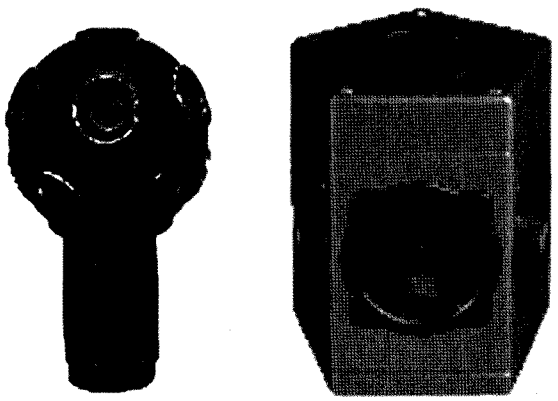
(그림 4) 평면 스크린에서 멀티 프로젝션 기하학 보정의 예 [6]

Ⅲ. 파노라마 기술 동향

파노라마 기술은 한 대 이상의 카메라를 방산형 또는 평행으로 배치시켜 획득된 다수의 영상을 하나의 고해상도 영상으로 변환하여 실경을 전망하는 것과 같은 느낌을 줄 수 있도록 넓은 시야각의 영상을 생성하는 기법이나 장치를 의미한다. 본 장에서는 국내외에서 진행되고 있는 파노라마와 관련된 영상 획득 기술, 디스플레이 기술, 서비스 기술 및 표준화 동향에 대해 살펴보고자 한다.

1. 파노라마 영상 획득 기술

파노라마 영상 획득 기술은 Immersive Media, Point Grey, Fraunhofer HHI 등을 중심으로 파노라마 카메라 에 대한 연구가 활발히 진행되고 있다. 캐나다 IT 업체인 Immersive Media는 11개의 카메라를 이용한 파노라마 카메라 시스템인 dodeca 2360(2400X1200 pixels/frame, 30fps)를 개발하였으며, Point Grey사에서는 최대 5400 X 2700(15fps, JPEG)의 고화질 파노라믹 영상을 획득할 수 있는 spherical 형태의 파노라마 동영상 획득 시스템인 Ladybug3를 개발하였다. 상기 획득 장비들은 주로 웹 기반의 서비스를 위한 것으로 줌인/줌아웃 기능 등을 지원하지 않는다. 그밖에 HHI에서는 실시간 양안식 파노라마 영상 생성을 위한 3D 파노라마 카메라 시스템을 개발하고 있다.

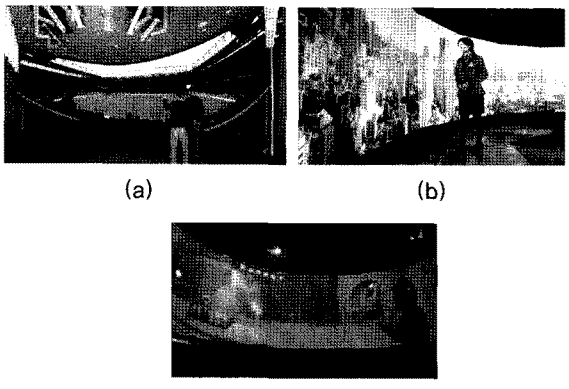


(a)Dodeca2360 (b)Ladybug3

(그림 5) 대표적인 파노라마 비디오 획득 장치

2. 파노라마 디스플레이 기술

파노라마 디스플레이 기술은 Fraunhofer HHI, 미쓰비시 등 다양한 업체에서 개발하고 있다. Fraunhofer HHI time lab에서는 immersive cinema라는 명칭으로 노르웨이 projectiondesign사의 F32 DLP 프로젝터 7대를 연동하여 7K 크기의 180도 파노라마 디스플레이 기술을 개발하고 있으며, 일본 미쓰비시 사에서는 67인치 패널 17개를 이용하여 높이 2m, 직경 7.5m짜리 대형 파노라마 디스플레이를 개발하고 있다. 또한 호주 IT업체인 ZKM에서는 4~6개의 프로젝터를 이용하여 높이 2.8m, 직경 8m 크기의 360도 실린더형 디스플레이인 PanoramaScreen을 개발하여 해당 디스플레이를 통해 1:49 비율의 Quake III arena 360 panoramic VR 게임을 시연하였다. 그밖에도 Immersive media에서는 360도 파노라마 영상을 볼 수 있는 HMD(head mounted display)를 개발 중이다.



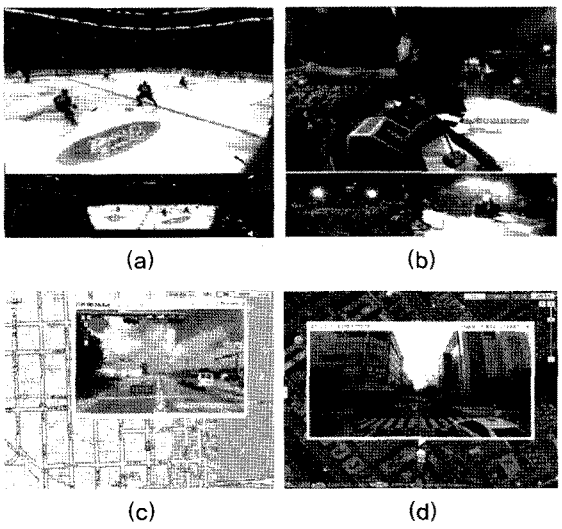
(그림 6) 파노라마 디스플레이 기술
(a) HHI의 immersive cinema
(b) 미쓰비시 사의 대형 파노라마 디스플레이
(c) ZKM의 PanoramaScreen

3. 파노라마 서비스 기술 동향

파노라마 서비스 기술은 파노라마 정지영상을 이용한 위치 기반 서비스, 경기나 공연 중계 영상 제공 서비스 등이 있으며, 현재까지는 웹 기반 서비스를 중심으로 이루어지고 있다.

캐나다 CBC(Canadian Broadcasting Corporation)의 하키 관련 프로그램인 "Hockey Night in Canada" 에서는 2009-2010 시즌에 하키 경기장 내에 360° 카메라를 설치하여 이를 웹에서 제공하고, 사용자는 뷰어를 이용해 자유롭게 시점을 선택 시청할 수 있도록 하였다.

음악관련 전문 회사인 MuchMusic 에서는 360° 카메라를 공연장에 설치하여 인터넷에서 실시간 방송을 제공하였다. 구글은 구글맵과 파노라마 정지영상 연계하여 웹 기반 스트리트뷰 서비스를 제공하고 있다. 국내에서도 다음, 네이버 등의 포털 사이트를 중심으로 웹 기반에서 지도와 연동하는 파노라마 서비스를 제공하고 있다.



(그림 7) 파노라마 서비스 일례
(a)캐나다 CBC 하키 프로그램 (b)캐나다 muchMusic 음악 프로그램 (c)구글 스트리트뷰 (d)다음 스트리트뷰

4. 표준화 기술 동향

ISO/IEC 산하의 동영상 부호화 기술 전문가 그룹인 MPEG에서는 2004년 8월부터 다수의 카메라를 통해 촬영된 영상물을 기하학적으로 교정하고 공간적으로 합성하여 여러 방향의 시점을 사용자에게 제공하는 다시점 비디오 부호화(Multi-view Video Coding, MVC)에 대한 표준화 작업이 진행되었으며, ISO와 ITU가 공동으로 JVT를 통해 표준화를 진행한 결과, MPEG-4/AVC Part 10 또는 H.264 Amendment 4 표준으로 마무리되었다.

또한 MPEG에서는 FTV/3DV에 대한 표준화가 진행되고

있으며, MVC가 주로 다시점 영상의 색차 정보를 압축하는 데 비해 3DV는 각 시점의 깊이 영상을 포함한 3차원 입체 영상의 획득/처리/전송/재생에 대한 모든 과정을 포함하고 있다. MPEG-A 시스템 표준 Part 11에서 스테레오스코픽 비디오의 저장 및 재생을 위한 파일 포맷(AF)이 표준화되었으며, MPEG-C Part 3에서는 깊이(depth)와 시차(disparity) 데이터를 포함한 3DTV 응용을 위한 부가 비디오 데이터 포맷을 정의하고 있다.

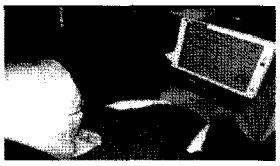
상기와 같이 표준화 기술은 입체영상이나 FTV와 같은 멀티뷰 비디오 압축을 위한 기술 위주로 진행되고 있으며, 향후 파노라마 비디오를 위한 효과적인 표현 및 압축 방식에 대한 기술표준이 필요할 것으로 판단된다.

IV. 사용자 인터페이스 기술 동향

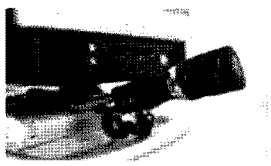
고품질의 파노라마 서비스를 제공함에 있어 기존의 멀티미디어 소비에 사용하는 키보드나 마우스 터치 등의 일반적인 사용자 입력 인터페이스는 몰입감을 증대시켜 주기에 적당하지 않다. 따라서 고품질의 파노라마 비디오 소비에 적합한 사용자 인터랙션을 지원해 주기 위해서는 최근 각광받고 있는 휴먼 팩터에 기반한 사용자 인터페이스 기술의 적용이 필요하며, 그 중 사용자 친화적인 파노라마 영상 소비를 위해 가장 효과적이라고 판단되는 사용자 인터페이스 기술들에 대해 살펴보고자 한다.

1. 시선추적 기술 동향

시선추적 기술과 관련하여 프랑스 IRISA, 이탈리아 ARTS, 독일의 HHI 연구소가 대표적이다[8]. 독일의 HHI 연구소의 경우 시청자의 응시 각도에 따라 다른 3차원 양안 영상을 합성, 생성, 재생하는 디스플레이 장치에 대한 연구를 수행하고 있다. 버지니아대학은 장애자를 위한 통신 수단의 한 방법으로 눈으로 문자를 입력할 수 있는 가상 키보드(virtual keyboard) 장치인 ERICA시스템을 개발하였으며, 싱가포르 NTU(난양 과기대)에서는 복수개의 팬&틸트 카메라와 적외선 조명을 사용한 시선추적 기술을 개발하였다. 국내에서는 한국과학기술원이 이용자의 두 동공중심 간의 거리를 이용하여 3D 시선좌표 계산 기술을 개발한 바 있다. 앞서 나열한 현존하는 시선추적 기술들은 PC 환경에서 사용할 수 있는 근거리 시선추적 기술들이 대부분이다. 반면에 한국전자통신연구원에서는 눈 위치 및 Z거리에 적응적인 패닝/틸팅/포커싱 기능을 포함하여 원거리(1.4m)에서도 추적이 가능한 비착용형 시선추적 기술을 개발하였으며, 이를 IPTV 환경에 적용함으로써 TV 화면을 응시하는 것만으로도 메뉴 조작이 가능한 새로운 사용자 인터페이스를 시연한 바 있다.



(a)ERICA 시스템



(b)Eyetech TM3

(c)비착용형 시선추적 시스템

(그림 8) 시선추적 서비스 일예

2. 제스처 인식 기술 동향

제스처 인식 기술과 관련하여 MIT Media Lab에서는 2009년 2월 실세계에 디지털 정보를 보다 편리하게 연결하여 사용할 수 있는 마커 기반 손가락 제스처 인식기인 'Sixth Sense'를 발표했다. 마이크로소프트(MS) 사에서는 2003년부터 2008년까지 5년간 Surface 프로젝트를 수행하였으며, 해당 프로젝트를 통해 멀티터치와 손동작 제스처 인식 기술을 선보였다. 또한 마이크로소프트 사에서는 2010년 게임용 동작인식 센서 키넥트를 상용화하면서 제스처 인식 기술에 새로운 변화를 가져왔다. EU FP6 CHIL (Computers in the Human Interaction Loop) 프로젝트에서는 스탠포드, CMU, 프라운호퍼, INRIA등 미국과 유럽의 총 9개국 15개 기관이 참여하여 사무실 및 강의실 환경에서의 음성, 제스처 등 멀티모달 정보 인식 시스템을 개발하였다.

3. 음성인식 기술 동향

음성인식 기술은 모바일 산업 및 스마트 TV 시장 장악에 있어 필수적인 기술로 근래 많은 진화를 거듭해 생활 속으로 파고들기 시작했다. 구글은 자사의 '모바일 음성검색' 인식률이 95%에 달한다고 밝혔으며, 마이크로소프트(MS)도 윈도 모바일 스마트폰에서 자체 음성검색 엔진을 만들어 제공한다. IBM · AT&T 등 IT 기업과 MIT, CMU 등의 연구기관에서도 자체 음성 인식 엔진을 개발한 상태다. 현재 음성인식 기술 분야의 선도기업인 미국의 '뉘앙스 커뮤니케이션(Nuance Communication)' 은 2009년 음성인식 기술로 1조원의 매출을 올렸으며, 삼성전자 등 국내 기업들이 유럽으로 수출하는 휴대전화에도 음성인식 기술을 제공하고 있다. 한국전자통신연구원(ETRI)은 2009년 12월 '음성 모바일 지능형 검색기술' 을 발표하고 다음커뮤니케이션과 제휴를 통해 음성인식 검색 서비스를 선보였다. 그러나 국내 음성인식 기술은 아직 키워드 검색만 가능하고 문장 형태의 자연어 검색 기술은 현재 개발 중이다[9].

4. 표준화 기술 동향

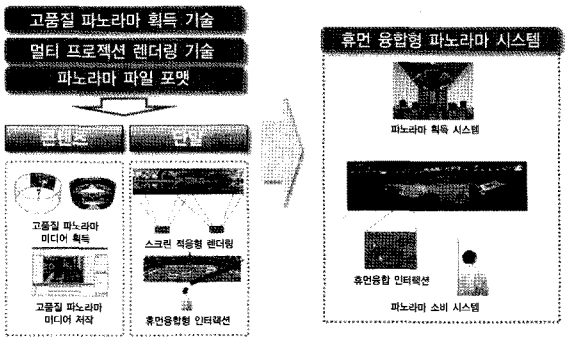
ISO 산하에서 인간 공학을 주요 이슈로 하는 TC159 (Technical Committee 159)에서 인간과 시스템의 상호작용(Ergonimics of human-system interaction)을 다루는 SC4(Sub-Committee)는 사용자 관점의 인터페이스에 대한 다양한 표준화를 수행하고 있다. 특히 WG9 은 햅틱 상호작용에 대한 전반적인 지침(Haptic and tactile interactions)을 중점적으로 표준화 하고 있다. 또한 WG11에서는 현실 세계와 가상 세계 간의 상호 데이터 교환을 위한 표준 인터페이스 포맷을 ISO/IEC 23005 MPEG-V 에서 1차 완료 하였다. 이와 함께 Widget, LASeR, 그리고 BiFS 와 같은 장면 서술 표준을 관리하고, 모션 센서와 같은 고급 사용자 인터페이스 장치를 연동하기 위한 ISO/IEC 23007 MPEG-U 표준이 진행되고 있다.

V. 휴먼융합형 파노라마 시스템

휴먼융합형 파노라마 시스템은 파노라마 기술과 사용자 친화적 인터페이스 기술의 융합을 통해 보다 사실적인 미디어 소비 환경을 제공해 줌으로써 사용자의 몰입감과 현장감을 증대시켜 주는데 그 목적이 있다. 일예로 사용자의 의도에 따라 시선 추적이나 제스처 정보를 분석하여 시점을 변환한다거나 부가 정보를 제공함으로써 보다 직관적이고 손쉽게 파노라마 영상을 소비할 수 있도록 한다.

1. 시스템 구조

고품질의 휴먼융합형 파노라마 시스템의 구조는 크게 파노라마 영상 획득을 위한 고품질 멀티 영상 획득 시스템, 복수 개의 카메라를 통해 획득된 영상들을 이용하여 파노라마 영상 및 파노라마 콘텐츠 파일을 생성하기 위한 고품질 파노라마 영상 생성 시스템, 그리고 멀티 프로젝터를 이용하여 파노라마 영상 콘텐츠를 적응적으로 재생하기 위한 휴먼융합형 파노라마 소비 시스템으로 구성되며 그 구성은 (그림 9)와 같다.



(그림 9) 휴먼융합형 파노라마 시스템의 구성

1) 고품질 멀티 영상 획득 시스템

휴먼융합형 파노라마 서비스를 위한 고품질 멀티 영상 획득 시스템은 파노라마 영상 생성을 위해 사용될 다수의 영상을 획득하는데 그 목적이 있다. 이를 위하여 동기화된 여러 대의 HD 카메라로부터 파노라마 비디오 생성을 위해 영상 간에 중첩 영역을 포함하는 다수의 비디오를 획득하여 저장 장치에 저장하는 기능을 제공한다. 일반적으로 카메라 간의 동기는 양안식 혹은 파노라마와 같은 응용에서 멀티카메라에서 출력되는 프레임간의 동기화를 목적으로 하며 Gen-lock 이라고 하는 동기화 방법을 지원하는 카메라를 사

용하거나, 소리가 발생하는 피크 정보를 바탕으로 프레임들을 정렬시킨다.

일반적으로 카메라 보정은 격자 패턴을 촬영한 이미지로부터 교차점들간의 상관 관계를 이용한 Zhang 방법[10]을 사용한다. 하지만 파노라마와 같이 다수의 카메라를 동시에 사용하는 본 시스템에서는 신속성 및 편의성이 중요하기 때문에 일차원 객체를 3대 이상의 카메라에서 동시에 촬영하여 일차원 객체의 이동 정보를 바탕으로 보정 상수를 획득하는 Svoboda 방법[11] 혹은 Zhang 방법[12]을 사용한다.

2) 고품질 파노라마 영상 생성 시스템

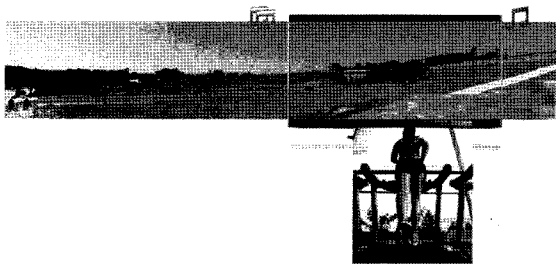
고품질 파노라마 영상 생성 시스템은 획득된 복수 개의 비디오를 스티칭, 블렌딩 등의 영상 처리 과정을 거쳐 하나의 파노라마 비디오로 생성하는 기능을 제공한다. 세부적으로는 상기와 같이 고품질 멀티 영상 카메라로부터 획득된 복수의 비디오 영상과 카메라/영상 파라미터, 동기화 정보 등을 이용하여 UHD 급 해상도의 고품질 파노라마 영상을 생성하는 기능 외에도 생성된 고품질 파노라마 영상을 부호화하기 위한 파노라마 비디오 부호화 기능, 부가정보 및 인터랙션 등의 리치미디어 서비스 제공을 위한 파노라마 장면 생성 기능 등을 포함한다.

고품질 파노라마 비디오는 일반 Full HD 영상에 비해 수배 이상의 해상도를 갖기 때문에 현재의 비디오 부/복호화기로는 영상을 한꺼번에 코딩하기 어렵다. 따라서 고품질 파노라마 영상을 처리하기 위해서는 이를 패치 형태로 분할하여 부/복호화 하는 방법이 필요하며, 각각의 패치는 독립적인 형태가 아니라 인접 패치간 영상간에 밀접한 관계가 존재하므로 각각의 패치들간의 영상 참조가 필요하다. 또한 실린더 및 구 형태의 파노라마 영상은 부/복호화를 위해 사각형 태로 저장할 때 영상의 기하학적 변형이 생기며 해당 영상은 소비 단말에서 렌더링 할 때 영상 처리를 통해 실린더 및 구 형태로 복원된다. 따라서 이러한 변형된 파노라마 영상을 바로 부호화 하고, 이를 그대로 복호화하여 렌더링할 경우 영상의 기하학적 정보가 밀집된 부분이 다른 부분에 비해 영상 열화가 커질 가능성이 존재하므로 이를 고려하여 기하학 정보가 밀집된 영역에 대한 강화된 코딩 방법이 요구된다.

그밖에도 파노라마 서비스에서 실린더, 구 형태 등 다양한 소비 환경에 맞게 파노라마 영상을 매핑하고, 사용자 인터랙션을 지원해 주기 위하여 리치미디어 기술이 필요하며, 이러한 기능을 지원해 주기 위한 리치미디어 기술의 일예로 MPEG-4 BiFS 등이 있다. 기본적인 장면 구성, 사용자 인터랙션 외에 4D 효과 제공 등을 고려해 볼 수 있으며 향후 이를 위한 기술 확장이 필요할 것이다.

3) 휴먼융합형 파노라마 소비 시스템

휴먼융합형 파노라마 소비 시스템은 2대 이상의 멀티프로젝터(UHD)를 이용하여 다양한 소비환경에 적응적으로 고품질의 파노라마 비디오를 재생하고, 특화된 사용자 인터페이스를 이용하여 휴먼융합형 파노라마 콘텐츠를 보다 직관적으로 편리하게 소비하는 기능을 수행한다. 즉, 생성된 고품질 파노라마 콘텐츠를 해석한 후 멀티 프로젝터를 이용하여 렌더링 환경에 따라 적응적으로 파노라마 비디오를 렌더링한다. 그밖에 시선추적과 제스처 인식이 결합된 사용자 친화적인 인터페이스 기술을 적용하여 보다 손쉽게 파노라마 네비게이션을 수행하고 부가정보 등의 리치미디어 데이터를 소비하는 기능을 수행한다.



(a)디지털 산책 서비스

(b)원격 쇼핑 서비스

(그림 10) 휴먼융합형 파노라마 응용 서비스의 일예

2. 응용 서비스 시나리오

파노라마 영상은 하나 이상의 카메라로부터 획득된 여러 장의 영상을 합성하여 넓은 시야각을 가지는 하나의 영상을 만드는 기술로, 좌우 최대 360°, 상하 최대 180° 등 사용자에게 넓은 시야각을 제공함으로써 특정 장소에 가지 않더라도 현장에 와 있는 듯한 현실감을 제공할 수 있다. 이러한 파노라마 영상은 주요 관광지, 유적지, 관공서, 전시장 등 특정 장소나 지역 등을 표현하는데 효과적으로 쓰일 수 있다. 이는 또한 제스처, 음성, 시선 추적 등의 사용자 친화적 인터페이스와의 결합을 통해 보다 사실적인 미디어 소비 환경을 제공함으로써 몰입감과 현장감을 증대시켜 줄 수 있다. 이를 이용한 일 예로 디지털 산책 서비스, 원격 쇼핑, 전시 관람 서비스 등이 있다.

디지털 산책 서비스의 경우 (그림 10(a))에서 보여지는 바와 같이 사용자의 시선이나 동작을 분석하여 움직임에 따른 사용자 의도를 파악하여 파노라마 영상 중 일부분을 사용자의 의도에 맞게 제공해 줄 수 있다. 이와 더불어 바람이나 꽃향기 등의 4D 효과까지 결합된다면 현장감은 더욱 증대될 것이다. 원격 쇼핑 서비스의 경우 (그림 10(b))와 같이 사용자의 제스처와 시선 정보를 파악하여 현장에서 직접 고르는 것처럼 파노라마 화면을 전환해 주고 사용자가 관심을 보이는 상품에 대한 상세 정보를 제공해 주는 등 사용자 친화적 인터페이스 제공을 통해 보다 손쉽게 구매할 수 있는 서비스를 제공할 것이다.

VI. 결 론

본 고에서는 파노라마 기술 및 사용자 인터페이스 기술의 동향에 대해서 알아보고, ETRI에서 개발 중인 고품질의 휴먼융합형 파노라마 시스템에 대해 살펴보았다. 휴먼융합형 파노라마 서비스는 고품질의 파노라마 기술과 사용자 친화적 인터페이스 기술이 접목되어 시야각이 넓은 파노라마 콘텐츠를 직관적이고 손쉽게 조작할 수 있도록 제공함으로써 현장감과 더불어 사용자의 몰입감을 한층 높여줄 것이다. 이처럼 실감미디어 서비스와 인간 친화적인 디지털 휴머니즘 기술과의 결합을 통해 실감형 서비스의 실효성에 대한 기반을 마련할 수 있을 것이며, 이는 IPTV, 스마트 TV와 같은 차세대 방송통신융합서비스의 활성화에 기여할 수 있을 것이다.

앞서 살펴본 바와 같이 파노라마 서비스에 대한 관심은 꾸준히 증가하고 있으며 관련 기술 개발도 지속적으로 이루어지고 있다. 그러나 개발된 기술들은 이렇다 할 국제 표준 또는 사실상 표준(de-facto standard)이 존재하지 않은 채 개발사마다 독자적으로 개발되고 있는 형국이다. 일 예로 국내외적으로 파노라마 AV 코덱이나 파일 포맷 등에 대한 연구개발이 미비하다. 따라서 파노라마 AV 코덱, 파일 포맷 등의 연구를 통해 파노라마 서비스 시장에 표준화된 서비스 방향을 제시할 수 있을 것이다. 또한 향후 시선 추적, 헤드 트랙킹 등을 이용한 다중 사용자 의도 기반의 신개념 파노라마 인터페이스에 대한 개발도 필요할 것이다.

고품질의 휴먼융합형 파노라마 서비스는 시점의 자유도를 증가시키고 넓은 시야각을 제공함으로써 나날이 증가되는 소비자의 다양한 욕구를 충족시킬 수 있을 것이며, 개발될 요소기술을 기반으로 관광, 광고, 사회 안전, 교통 등 다양한 산업분야에 확산 및 기여할 수 있을 것이다.

참 고 문 헌

[1] C. Fehn, C. Weissig, I. Feldmann, M. Mueller, P. Eisert, P. Kauff and H. Bloss, "Creation of high-resolution video panoramas of sport events," in Proc. 8th IEEE Intl. Symp. on Multimedia (ISM' 06), San Diego, CA, USA, pp.291?298, Dec. 2006.

[2] M. Brown and D. G. Lowe, "Recognising Panoramas," In Proceedings of the 9th International Conference on Computer Vision, Nice, France, pages 1218.1227, 2003.

[3] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," cascade filtering approach. IJCV 60 (2004) 91 ? 110

[4] B. S. Reddy and B. N. Chatterji, "An FFT Technique for Translation, Rotation, and Scale-Invariant Image Registration," IEEE Trans. On Image Processing, Vol. 5,

No. 8, Aug 1996.

[5] R. Szeliski, "Image Mosaicing for Tele-Reality Applications," DEC Cambridge Research Lab., Tech. Rep. CRL 94/2, May 1994.

[6] H. Chen, R. Sukthankar, G. Wallace, and T. Cham, "Calibrating scalable multi-projector displays using camera homography trees," In Computer Vision and Pattern Recognition, 2001.

[7] L. P. Soares, R. J. Costa, B. R. Araujo, and J. A. Jorge, "Automatic color calibration for commodity multi-projection display walls," in SVR 2007 ? Proceedings of the IX Symposium on Virtual and Augmented Reality. Porto Alegre, RS, Brazil: SBC, 2007.

[8] 류기정 외, 시선추적 기술 개발 동향, SERI 뉴로마케팅 포럼

[9] 백강녕 외, 질문하면 바로 검색… 말로 통하는 모바일 세상 온다, 조선비즈, 2010.7.

[10] Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, Nov. 2000.

[11] T. Svoboda, D. Martinec, T. Pajdla, "A convenient multicamera self-calibration for virtual environments," PRESENCE:Teleoperators and Virtual Environments, 14(4), August

[12] Z. Zhang, "Camera calibration with one-dimensional objects," IEEE Trans. Pattern Analysis and Machine Intelligence, 26(7):892?899, 2004.

약 력

이 인 재

1999년 성균관대학교 전자공학과 학사
2001년 성균관대학교 전기전자및컴퓨터공학부 석사
2001년 ~ 현재 한국전자통신연구원 선임연구원
관심분야 : Computer vision, Richmedia, UX

약 력

조 용 주

1997년 Iowa State University 공학사
1999년 Iowa State University 공학석사
2009년 Michigan State University 공학박사
2001년 ~ 현재 한국전자통신연구원 방통융합미디어연구부 선임연구원
관심분야 : Wireless video, Computer vision

기 명 석

1999년 전남대학교 컴퓨터공학사
2001년 전남대학교 컴퓨터공학석사
2001년 ~ 현재 한국전자통신연구원 선임연구원
관심분야 : MPEG-4 Systems, 파노라마 영상처리

임 성 용

1999년 한국과학기술원 전기 및 전자공학과 학사
2001년 한국과학기술원 전기 및 전자공학과 석사
2001년 ~ 현재 한국전자통신연구원 선임연구원
관심분야 : 멀티 카메라 보정, 하이브리드 시스템 시뮬레이션

이 희 경

1999년 영남대학교 컴퓨터공학과 학사
2002년 KAIST-ICC 정보통신공학부 석사
2002년 ~ 현재 한국전자통신연구원 선임연구원
관심분야 : HCI, Gaze Tracking, Bi-directional AD

차 지 훈

1992년 명지대학교 전자계산공학사
1996년 Florida Institute of Technology 공학석사 (Computer Science)
2002년 Florida Institute of Technology 공학박사 (Computer Science)
2003년 ~ 현재 한국전자통신연구원 방통융합미디어연구부 융합미디어연구팀장
2008년 ~ 현재 과학기술연합대학원대학교 부교수
관심분야 : Multimedia streaming, Interactive broadcasting system, feature extraction/tracking, Richmedia