

인터랙티브 다시점 입체 영상 시스템 개발

김익재, 송문재, 송창준, 김진우, 권용부
한국과학기술연구원 영상미디어연구센터

Development of Interactive Multiview Stereo System

Ig Jae Kim, Mun-Jae Song, Chang-Jun Song, Jin-Woo Kim, Yong-Moo Kwon
Imaging Media Research Center, Korea Institute of Science and Technology

ABSTRACT : 최근 대두되고 있는 입체 TV는 보는 사람의 위치가 바뀌어도 동일한 영상을 보여주는 방식을 가지고 있어 입체감은 느끼지만 현실감과 생동감이 떨어지게 된다. 따라서 인터랙티브 다시점 입체 영상 시스템에서는 이러한 문제점을 개선하기 위해 유한개의 획득 영상으로부터 생성된 입체 영상을 사람의 위치에 따라 다른 영상을 보여줌으로써 보다 자연스러운 입체감을 주도록 한다. 본 시스템은 이를 위해 로봇암을 사용하여 영상 데이터베이스를 만들고 이로부터 영상기반 3차원 모델링 방법을 모델링을 한 후 안경식과 무안경식 방법으로 입체 영상을 표시하게 하였다. 또한 이 시스템은 생성된 3D 모델의 가상환경과의 결합을 가능케 하고, 또한 가상환경내에서의 3D 모델을 원격 제어할 수 있도록 하였다.

1. 서 론

고선명 텔레비전 (HDTV) 이후 차세대 디스플레이로서 입체 텔레비전이 널리 연구되고 있다. 입체 텔레비전을 구현하는 방법으로서 양안 시차를 이용하는 여러 방법들이 제시되고 있는데, 이러한 방법들이 제시하는 시스템은 관찰자로 하여금 어느 정도의 입체감을 느낄 수 있도록 하지만, 사람이 실제 어떤 사물을 볼 때와 같은 생동감·현실감을 위해서는 관찰자의 위치에 따라 인식되는 사물의 모양도 조금씩 다르게 보이도록 하는 다시점 영상 표시기술을 필요로 한다[1,2]. 이를 위해서는 디스플레이 할 대상의 주위에 무한개의 카메라를 설치하고, 각 위치에서 영상을 획득·저장하여 디스플레이단

에서는 관찰자의 위치를 추적하고 그 위치에 해당하는 영상을 표시함으로써 구현될 수 있다. 그러나 이러한 방식은 현실적으로 해결하기 어려운 많은 한계를 가지고 있기 때문에 실제 시스템들은 미리 정해진 위치에 유한개의 카메라를 설치하여 영상을 얻고, 카메라를 설치하지 않은 위치에 해당하는 영상들은 실제 카메라에서 얻은 영상들로부터 보간법에 의해 합성한다. 또한 관찰자가 원하는 임의의 시점을 관찰자와 시스템간의 상호작용에 의해 표시할 수 있는 기법이 요구된다.

따라서, KIST에서는 인터랙티브 다시점 입체 영상 시스템(Interactive Multiview Stereo System)을 개발하였는데, 이 시스템에서는 다시점 영상 입력 및 분석 기술, 영상 기반 모델링 기술, 입체 영상 저작 기술, 입체 영상 표시 기술, 사용자 인터랙션 기술 등을 포함하고 있어 유한개의 카메라로부터 획득된 영상으로부터 생성된 입체 영상을 보는 사람의 위치에 따라 다른 영상을 보여줌으로써 보다 자연스러운 입체감을 주도록 하며, 생성된 3D 모델의 가상환경과의 결합을 가능케 하고, 또한 가상환경내에서의 3D 모델을 원격 제어할 수 있도록 하였다.

2. 시스템 구성도

본 장에서는 인터랙티브 다시점 입체 영상 시스템인 IMSS에 대한 전체적인 설명을 하도

북 한다. 그림 1은 KIST에서 구축한 IMSS의 전체적인 구성도이다.

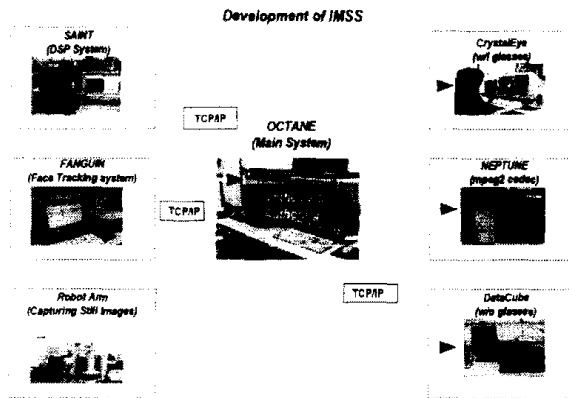


그림 1. IMSS의 전체 구성도

IMSS는 크게 디지털 입체 영상 입력부, 3차원 모델링 생성 및 저작부, 입체 영상 표시부와 마지막으로 인터랙티브 입체 영상 조작부로 나뉘어진다. 디지털 입체 영상 입력부에서는 온라인 입력과 오프라인 입력으로 나뉘어지는데 전자의 경우는 스테레오 카메라로부터 Matrox Genesis DSP보드로 입력된 영상을 사용하였고, 후자인 경우는 로봇암을 이용하여 평행 광축 환경하에서 영상을 획득하여 사용하였다. 3차원 모델링 생성 및 저작부에서는 5개의 시점으로부터 입력된 영상을 사용하여 MR-MBS기법[3]에 의해 기준 영상에 대한 깊이 정보를 추출하여 메쉬 생성 기법을 사용하여 wire frame 모델을 구한 후 원영상의 텍스처를 사용하여 실사 기반 3차원 모델을 생성하였고[4], 또한 생성된 3차원 모델을 사용자 조작에 따라 실시간으로 각각의 오브젝트 단위로 XYZ축에 대한 이동 및 회전이 가능하고 오브젝트 축척도 자연스럽게 변화 시킬수 있도록 하였다. 입체 영상 표시부에서는 크게 무안경식과 안경식으로 나뉘어진다. 무안경식으로는 디지털 마스크 기법에 의한 렌더큐라를 사용하여 입체 영상을 표시하였고[5], 안경식으로는 StereoGraphic사의 CrystalEye를 사용하여 입체 영상을 볼 수 있도록 하였다. 마지막으로 인터랙티브 입체 영상 조작부에서는 크게 두가지로 나뉘어진다.

DSP를 사용하여 실시간으로 거리 정보를 측정하여 3차원 모델의 원격 제어를 가능케 하

였고, 또한 Head-tracking 기법을 이용한 시점 대응 영상 렌더링을 가능케하여 사람의 위치에 따라 다른 뷰를 보여줌으로써 자연스러운 입체감을 주도록 하였다.

3. 디지털 입체 영상 입력부

영상 입력은 두가지 방법 즉, 로봇 암 또는 스테레오 카메라를 이용한 입력을 사용하였다.

(1) 로봇 암을 사용한 입력

정밀한 카메라 위치 제어 기능을 갖는 로봇 암을 사용하여 다시점 영상 데이터베이스를 구축한다. 본 연구실에 구축한 로봇 암을 사용하여 여러 가지 상황에서의 실험을 위한 영상 데이터베이스를 구축하였다. 또한 일본 쓰쿠바 대학에서 제작한 9 x 9 행렬 형태의 위치에서 촬영한 평행 광축 환경하에서 영상 데이터베이스를 구축하였다.

(2) 스테레오 카메라를 이용한 입력

스테레오 정합을 위해서는 epipolar 구속조건을 만족하도록 카메라 배치를 하여야 한다. 그러나 일반 스탠드를 사용한 상황에서의 카메라 배치는 epipolar 구속 조건을 만족하지 않는다. 이 문제를 해결하는 방법으로는 카메라 캘리브레이션을 통해 스테레오 카메라 각각에 대해 카메라의 내부 및 외부 변수를 구하고 스테레오 카메라로부터 구한 영상에 대해 영상 보정 처리를 함에 의해 epipolar 구속 조건을 만족하는 영상을 구할수 있다. 본 연구에서는 카메라 파라미터 추출을 위한 카메라 캘리브레이션 및 영상 보정 기법을 구현하였으며 실험을 통해 그 유효성을 확인하였다.

4. 3차원 모델링 생성 및 저작부

실제 영상으로부터 3차원 모델을 이끌어 내기 위해서는 조밀하고(dense), 정확한 깊이 정보를 얻는 것이 중요하며 텍스처를 올바르게 모델에 적용하는 문제를 고려해야 한다. 모델링 과정에서 입력되

는 영상은 배경과 오브젝트가 어느 정도 거리를 두고 떨어져게 하였으며 이러한 조건에서 구하여진 깊이 맵을 히스토그램 방식을 이용하여 배경과 오브젝트를 분리해 낸다. 오브젝트 영역을 구하게 되면 깊이 정보를 이용하여 일정한 간격으로 vertex를 선택하고, 선택된 vertex를 연결하여 삼각형을 만든후 그 삼각형으로써 오브젝트의 표면을 표현한다. 최종적으로 입력 영상 중 깊이 정보와 동일한 위치인 영상을 사용하여 텍스처 매핑을 한다.

오브젝트의 깊이 맵 추출은 본 연구에서 개발된 MR-MBS (Multi Resolution-Multiple-Baseline Stereo) 기법을 사용하였다. MR-MBS 기법은 적응적 폐색 영역 처리 기능을 가지며 다시점 영상으로부터 기준 영상에 대한 깊이 값을 추출해낸다.

그림 2는 본 연구에서 사용한 영상 기반 3차원 모델링의 기본 개념 및 예를 나타낸 것이다. 그림 2에 나타낸 바와 같이 5개의 시점으로부터 입력된 영상을 사용하여 MR-MBS 기법에 의해 기준영상 (중앙의 영상) 에 대한 깊이 정보를 추출한다. 그 다음 메쉬 생성 기법을 사용하여 wire frame 모델을 구한후 원 영상의 텍스처를 사용하여 실사 기반 3차원 모델을 생성하게 된다.

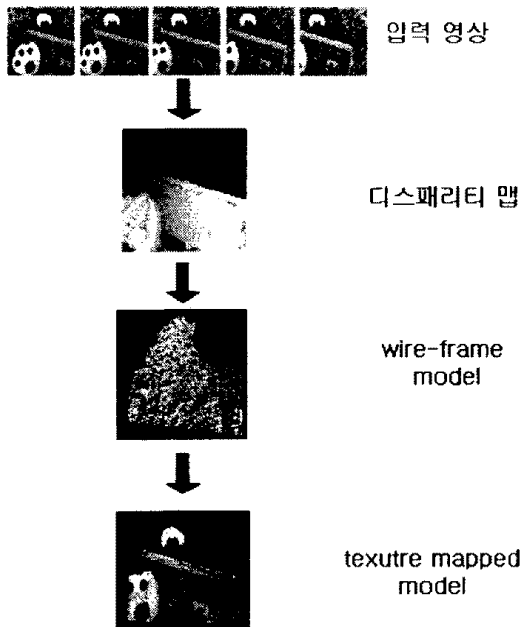


그림 2. 영상 기반 3차원 모델링 과정

z-key 방법은 픽셀 대 픽셀 단위로 실제의 오브젝트와 배경의 깊이 정보를 이용함으로써 오브젝트와 배경간의 전후 관계를 자연스럽게 합성할 수 있다. 구현된 시스템은 사용자 조작에 따라 실시간으로 각각의 오브젝트 단위로 XYZ축에 대한 이동 및 회전이 가능하고 또한 오브젝트 크기를 자연스럽게 변화시킬 수 있다. 그림 3은 저작 과정을 보여준다.

깊이 정보를 이용한 영상 합성 기법은 관찰자가 원하는 시점에서의 영상을 합성하기 위해서 사용될 수 있으며 이는 최근 유럽 PANORAMA 프로젝트에서 추구하는 실감 통신에서 look-around를 가능하게 하는 기반 기술이다. 한편 영상 합성 기술은 최근 주목받고 있는 디지털 방송분야, 전자 레크레이션분야, 교육분야, 의료분야, 멀티미디어 콘텐츠분야등에 핵심 기술로 사용될 수 있으며 그 한 예로서 다른 깊이 값을 갖는 두 영상을 깊이 정보 (depth-key) 에 의한 합성이 가능하다.

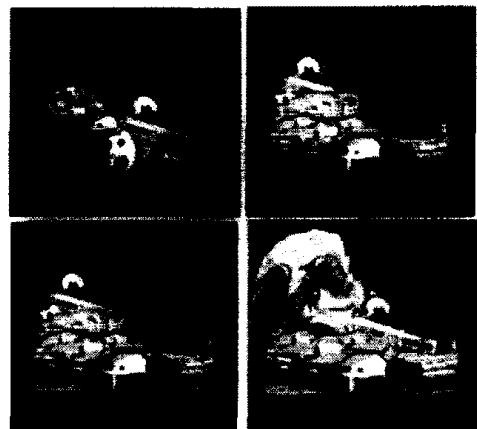


그림 3. 입체 영상 저작

5. 입체 영상 표시부

본 시스템에서는 입체 영상을 표시하는 방법으로 안경식과 무안경식으로 나눌 수 있다. 우선 무안경식 입체 표시방법으로 본 연구에서는 입체 영상 시스템 구현의 새로운 접근 방식인 디지털 마스킹 기법에 의한 렌더큐라 입체 영상 표시 기술을 개발하였다. 본 연구에서는 디지털 방식에 의한 영상 처리, 압축, 저장, 검색,

전송 등의 표준화 연구 방향을 고려하여 파이프라인/병렬 디지털 영상 신호 처리 기술을 이용한 렌티큐라 입체 영상 표시 기술을 개발하였다. 렌티큐라 입체 영상 표시 시스템 구현을 위해 10.4 인치 TFT LCD 모니터에 적합한 렌티큐라 슈트를 설계 제작하였으며 스테레오 카메라, DataCube 영상 처리 시스템 및 렌티큐라 슈트를 부착한 10.4inch LCD 모니터로 구성되는 렌티큐라 입체 영상 시스템을 구현하였다. 또한 Octane 시스템과 TCP/IP 통신을 통해 3차원 모델로 렌더링 되는 영상을 렌티큐라 시스템으로서 표시할수 있는 환경을 구현하였다.

안경식으로는 OCTANE상에서 입체영상을 렌더링한 후 Stereographic사의 CrystalEye를 사용하여 시점에 따른 뷰를 볼 수 있도록 하였다. 그림 4는 렌티큐라 입체 영상 시스템이며 그림 5는 안경식 및 무안경식 표시 개념을 나타낸 것이다.

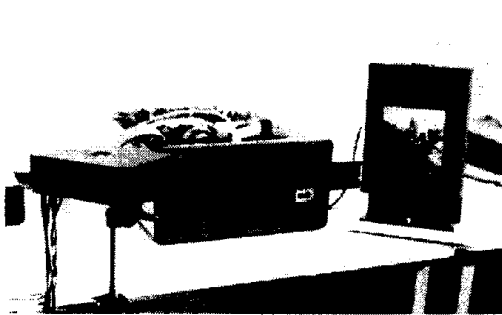


그림 4. 구현된 10.4" 렌티큐라 입체 영상 시스템

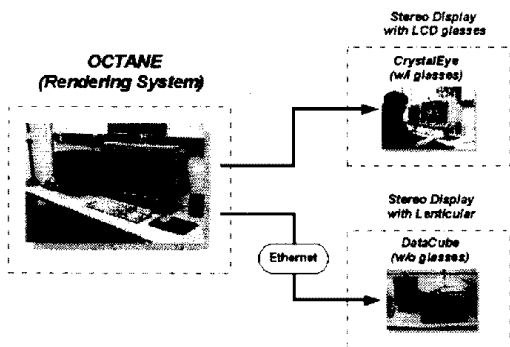


그림 5. 입체 영상 표시

6. 인터랙티브 입체 영상 조작 및 표시

(1) 3차원 모델 조작

스테레오 카메라를 이용하여 움직이는 오브젝트의 3차원 공간상에서의 위치 정보를 구해 이를 이용해 3차원 가상 공간상에서 3차원 모델의 위치를 제어하는 기법을 구현하였다.

그림 6은 사람이 이동함에 따라 3차원 모델이 장애물을 피해 이동하는 예를 나타낸 것이다. 스테레오 영상으로부터 실시간으로 깊이 정보와 위치 정보를 얻고[6,9], 이들 정보를 서버인 OCTANE으로 전달해 주는 클라이언트로 사용된 시스템은 IBM-PC이며 Intel Pentium Pro 233MHz의 Dual CPU와 메인 메모리가 128MB의 사양을 갖추고 있다. 스테레오 영상을 얻는데 사용된 카메라는 Pulnix TM-7EX monochrome 카메라이며 배이스라인은 100mm로 하고 입력 영상은 640x480의 크기이며

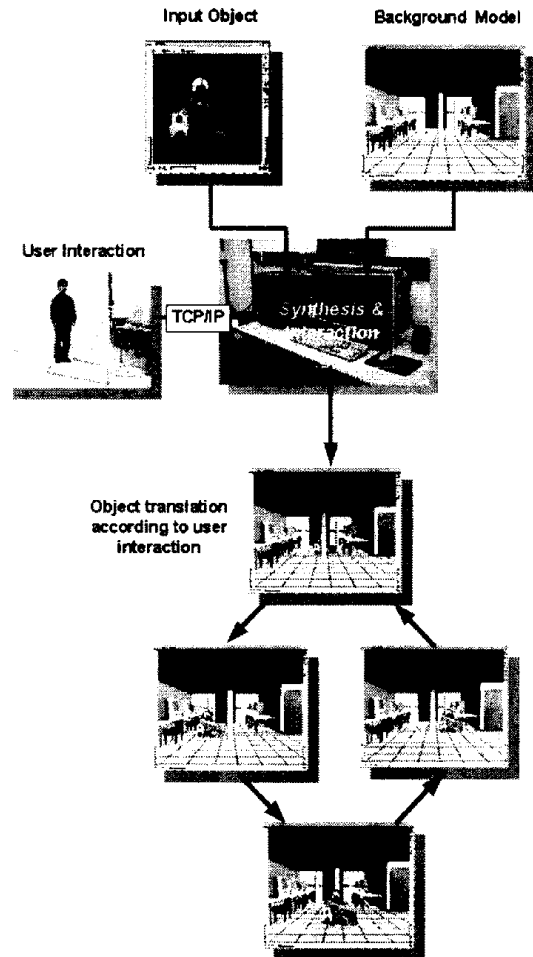


그림 6. 움직이는 오브젝트의 위치정보 추출 및 이를 이용한 3차원 모델 원경 제어

처리 해상도는 다운 샘플링한 후 160x120의 해상도에서 처리하였다. 움직이는 오브젝트의 위치 정보와 깊이 정보의 연산 처리 속도는 초당 25프레임의 속도로 실시간으로 처리 되어진다. 클라이언트에서 구해진 움직이는 오브젝트의 위치 정보와 깊이 정보는 TCP/IP를 통해 IMSS의 메인 시스템인 OCTANE에 전달되어 지고 이에 따라 3차원 가상 공간내에서 3차원 물체의 이동을 원격으로 제어 할 수 있도록 하였다. 배경으로 사용되어진 3차원 모델은 SGI의 Open Inventor를 사용하여 구현하였으며[8], 입력 오브젝트는 IMSS의 모델링부분인 다시점 영상을 기반으로 한 3차원 모델이다.

사용자는 실제 실험실에서 가상 공간상에서 장애물의 위치에 해당되는 지점을 미리 정의해 두고 이 위치를 피해감으로써 입력된 3차원 객체가 가상공간상에 존재하는 장애물을 피해감도록 하는 원격 제어가 가능하도록 하였다.

(2) 시점 적용 입체 영상 표시

관찰자의 머리 위치를 추적함에 의해 관찰자에 대응되는 시점을 추정하고 이를 바탕으로 현재 시점에 대응되는 3차원 모델을 렌더링 해 주는 기능을 구현하였다. 기본 원리는 얼굴의 살색 정보를 이용하고 있으며 안경을 사용한 경우 및 사용하지 않은 경우 모두에 적용된다. 시점 추적은 IBM-PC에서 구현하였으며 TCP/IP를 통해 Octane 시스템에 전송되며 이에 따라 시점에 대응되는 뷰를 렌더링해준다.

Head-tracking을 이용한 시점 대응 영상 렌더링을 구현하기 위해 본 연구에서는 색상 움직임을 이용하여 얼굴 영역을 추출하고[7], 얼굴 영역의 중심이 입력 영상에서 자리 잡은 위치를 네트워크를 통하여 화면 표시를 담당하는 장치에게 알려주어 사용자에게 시점 대응 영상을 보여 줄 수 있도록 하였다. 얼굴 영역을 추적하는 알고리즘을 그림 7에서 나타내었다. 일반적인 얼굴과 스테레오 영상을 보기 위해 LCD 안경을 착용한 상태로 나누어 두가지 방법으로 얼굴 영역을 추적하게 된다.

안경을 착용하지 않은 경우의 얼굴 추적은 GFCD[7]를 통하여 살색 영역을 찾아 각 영역을 레이블링하고 그 중 가장 큰 세 개의 영역

을 얼굴 후보 영역으로 선택한다. 얼굴 후보 영역에서 BWCD[7]를 통하여 검은색인 눈썹과 눈동자, 콧구멍이 존재하는 영역을 얼굴영역으로 결정한다. 최종적으로 얼굴의 중심점은 BWCD를 통하여 구하여진 눈썹과 눈동자, 콧구멍을 포함하는 사각형의 중심을 얼굴의 중심으로 결정한다. 안경을 착용한 경우에는 눈썹과 눈동자를 구할 수 없고, 얼굴 영역이 안경에 의하여 나누어진다는 제약 조건 때문에 큰 살색 영역이 상하로 존재하는 경우를 얼굴 영역으로 결정하고, 두 개의 살색 영역의 중심점을 얼굴의 중심점으로 결정한다.

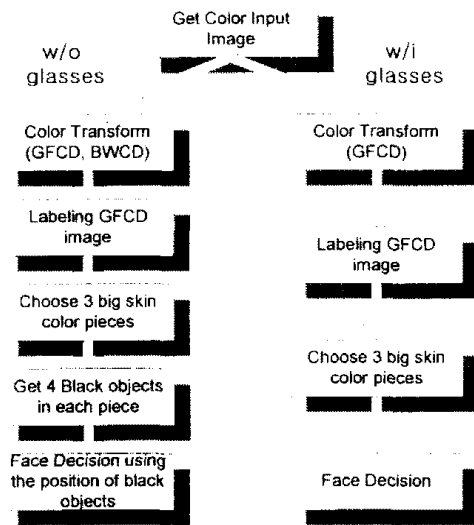


그림 7. 얼굴 영역 추출 알고리즘

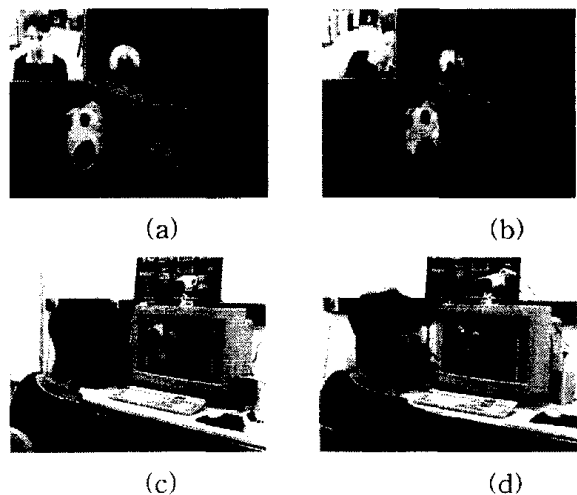


그림 8. 시점 대응 영상 렌더링의 결과

그림 8은 시점이 변화할 때 그에 해당되는 뷰를 렌더링해 주는 결과를 보여주는 결과이다. 그림 (a)와 (b)는 실제로 보는 사람의 움직임에 따라 3D 모델의 뷰가 달라짐을 볼 수 있다. 그림 (c)와 (d)는 사용자가 안경 착용의 유부에 상관없이 시점 대응 영상 렌더링이 된다는 것을 나타내고 있다. 따라서 Head-tracking 기법을 사용함으로써 보는 사람의 움직임에 따라 다른 뷰의 영상을 보여 줌으로써 보다 자연스러운 입체감을 주도록 하였다.

7. 결 론

본 연구는 디지털 입체 영상 시스템 기술 개발을 위한 기반 기술로서 영상 입력 기술, 영상 분석 기술, 3차원 모델링 및 저작 기술, 입체 영상 표시 기술 그리고, 인터랙티브 입체 영상 기술에 관한 것으로 유한개의 획득 영상으로부터 처리 시간 및 깊이 정보 정확도 측면에서 개선된 깊이 정보 분석 기법을 이용하여 시각 기반 3차원 형상 모델을 생성하고, 생성된 모델을 사용하여 새로운 3차원 영상물을 저작할 수 있는 인터랙티브 저작 툴을 개발하였다. 또한 생성된 모델을 입체로 표시하는 방법으로 렌더큐라를 이용한 무안경식과 CrystalEye를 사용한 안경식으로 나누어서 입체로 표시할 수 있게 하였다. Head-tracking 기법을 이용하여 사람의 위치에 따라 다른 영상을 보여줌으로써 보다 자연스러운 입체감을 주도록 하였으며, DSP를 사용하여 실시간으로 움직이는 오브젝트의 위치 정보와 깊이 정보를 구하여 3차원 가상 공간상에서 3D 모델의 원경 제어가 가능하도록 하였다. 향후 본 연구 결과의 성능을 개선하고 또한 새로운 기능을 추가하여 실감 컨트롤 시스템 기술 및 실감 단말기 기술 개발이 요구된다.

참고 문헌

[1] D. B. Diner, D. H. Fender, *Human*

Engineering in Stereoscopic Viewing Devices, Plenum Press, pp. 36-48, 1993.

[2] Toshio Motoki, et al., "Present Status of Three-dimensional Television Research," *Proc. of IEEE*, Vol. 83, No. 7, pp. 1009-1021, July, 1995.

[3] 박남준, 이제호, 권용무, 박상희, "중간 영상 합성을 위한 다해상도 다기선 스테레오 정합 기법," 방송공학회논문지, 제2권, 제2호, 1997년.

[4] 송문재, 권용무, 임성규, "영상기반 3-D 모델링 및 합성 기법 구현," 정보과학회 추계학술발표 논문집, pp. 591-593, 1998년.

[5] 안진호, 장홍엽, 이제호, 권용무, 김상국, 박상희, "무안경 방식의 실시간 입체 영상 표시 시스템 구현," 한국통신학회 하계종합학술대회 논문집, 1996년 7월.

[6] C. J. Song, I. J. Kim, Y. M. Kwon, "Depth extraction of moving object using TMS320C80 DSP," *Proc. of ITC-CSCC '98*, pp. 285-288, July, 1998.

[7] 김남호, 함상진, 안상철, 김형곤, "색상 움직임을 이용한 실시간 얼굴/손 추적 알고리즘," HCI '98 학술대회 논문집, pp. 9-14, 1998년 2월.

[8] J. Wernecke, *The Inventor Mentor : Programming Object-Oriented 3D Graphics with Open Inventor*, Addison-Wesley, 1994.

[9] 김남규, 이광도, 김형곤, 차균현, "저해상도 변위 히스토그램을 이용한 고성능 변위 정보 추출 알고리즘," 전자공학회지, 제 35 권, 제 3 호, pp. 131-143, 1998년 3월