

효과적 정보전달을 위한 영상정보의 3D 뷰 및 음성정보와의 융합 연구

정희원 신 준 철*, 종신회원 이 종 수*^o

A Study on 3D View Design of Images and Voices Integration for Effective Information Transfer

C. H. Shin* *Regular Member*, J. S. Lee*^o *Lifelong Member*,

요 약

본 논문에서는 컴퓨터를 사용한 효과적 정보전달을 위해 2D 영상정보를 3D 가상공간에 배치하고 음성 정보와의 유기적 연결기능을 갖는 3D 뷰 설계를 제안한다. 3D 공간에 배치된 영상정보의 사용자 접속을 어느 시점 또는 어떤 각도에서도 가능하도록 하여, 시각정보 전달효과를 제고하고 있으며, 음성정보의 첨부가 용이하도록 설계하였다.

3D공간에서의 영상 및 음성정보 배치는 단순하지만 효과적인 접속기능을 포함하고 있으며, 이들의 탐색 및 시청을 사용자접속 관점에서 설계하여 정보전달 효과를 제고하였다.

Key Words : 3D View, DirectX, Rotation Matrix, Interface

ABSTRACT

In this paper, we propose a 3D view design scheme which arranges 2D information in a 3D virtual space with a flexible interface and voice information. The scheme allows the user interface of the 2D image in 3D virtual space anytime from any view point. Voice information can be easily attached.

It is this simple and efficient image and voice information arrangement in 3D virtual space that improves information transfer.

1. 서 론

영상, 음성, 그래픽 등의 복합정보를 다루는 다양한 소프트웨어가 설계, 제시되고 있으며, 복합정보의 다양성에 따라 이의 접속방법도 다양하게 제시되고 있다. 그 중에는 강연방식이나 발표형식에 맞추어져 제작된 복합정보 표현 기술들이 제안되고 있다.

복합정보를 다루는 소프트웨어 기술과 함께 3D를 다루는 다양한 기술들이 현재 연구 개발되고 있다. 다양한 방면에서 3D 융합의 시도가 연구되고 있으나 실제로 사용되는 주 분야는 게임이나 디자인이다. 하

드웨어 기술의 지속적 발전과 3D 융합기술 연구 결과가 최근에 들어 운영체제¹⁾나 각종 미디어 탐색기에 응용되고 있다.

3D 기술을 사용자접속에 도입함으로써 복잡정보를 다양한 방식으로 표현할 수 있으며, 3가지 공간축이 존재해 더욱 자유로운 탐색이 가능하다. 이는 3D공간에서 활동하는 사람들의 공간감각을 모두 활용할 수 있도록 하기 때문이다.

강연이나 발표에 맞추어 제작된 복합정보가 3D 형태를 취함으로써 정보전달 효과가 제고되고 있으나, 3D 자료를 제작한다는 것은 어렵다. 각종 3D 디자

* 울산대학교 컴퓨터정보통신공학부(jsoolee@ulsan.ac.kr) (° : 교신저자)

논문번호 : KICS2009-07-328, 접수일자 : 2009년 7월 30일, 최종논문접수일자 : 2009년 12월 21일

인 소프트웨어들은 사용법을 익히는데 많은 시간과 노력을 요구하고 있으며, 일부 활용에서는 고수준의 수학적 지식도 요구하고 있다.

최근에 들어 발표나 교육 자료를 위한 소프트웨어들은 요약 내용들을 시각적으로 화려하게 작성하는 기법들의 사용자 인터페이스를 용이하도록 하고 있다. 마이크로소프트사에 파워포인트는 물론이고 유사 소프트웨어들도 발표 자료들을 페이지별로 화려하게 작성하고, 이들을 순차적으로 보여주는 다양한 기능을 가지고 있으나, 3D 사용자 접속 기술을 포함하고 있지 않다.

본 논문은 효과적 정보전달이 필요한 교육이나 발표를 위해 3D 기술을 복합정보의 사용자접속에 융합하는 기법을 제안하고 이를 구현하는 실제 프로그램의 개발과 설계 방법을 설명한다. 제안하는 프로그램은 간단하지만 효율적이고 효과적인 방법으로 가상의 3D 공간에 영상정보를 배치하고, 각 영상정보와 관련된 음성정보와의 접목을 용이하도록 한다. 정보들을 탐색하기 위해서 가상 카메라 기법을 도입하고, 이의 원활한 움직임을 제어하기 위한 키보드와 마우스의 사용자 인터페이스를 설계하였다. 빠르고 매끄러운 화면 출력을 위해 GPU를 사용하여 실시간으로 화면을 갱신한다.

II. 3D 뷰 설계

2.1 연구 동향

컴퓨터 프로세서의 성능이 개선되면서 화면처리 전용 GPU가 등장하였다. GPU는 주로 3D화면 처리를 위해 백터와 행렬 연산이나 색상처리 등을 대신하고 나아가서는 물리연산을 대신한다. 이런 3D기술이 등장하면서 게임과 산업이 크게 성장하고 있으며, 현재는 운영체제나 미디어 탐색기에 3D기술을 다양한 방향으로 활용하면서 가능한 많은 것을 효과적으로 표현하려는 기술이 연구되고 있다^{[4], [7], [8]}. 수많은 인터넷 웹 페이지들에 구조를 표현하는 연구^[7]에서는 다양한 페이지들을 한눈에 볼 수 있도록 3D에 배치하고 출력한다. 교육에서도 3D 활용 연구가 진행되고 있다^{[2], [3]}. 교육 등과 같은 효과적인 의사 전달이 필요한 환경에서, 발표에 사용되는 소프트웨어의 효율적 활용에 관한 연구^[2]와 가상 환경 기반 구현 연구^{[1], [9]} 등의 공통점은 다양한 정보의 효과적인 전달을 추구한다는 것이다. 이에 따라 사용자와 상호작용하는 많은 소프트웨어들이 개발되고 있으며, 이러한 연구는 향후 몇 년간 지속될 것으로 전망된다.

2.2 컨셉 설계

기존의 뷰들은 그림 1과 같이 평면 배치를 통해 영상정보를 배치한다. 이런 방식은 쉽고 효율적으로 구현이 가능하나, 수많은 영상정보를 탐색하기 위해서는 스크롤을 가로와 세로 2방향으로만 사용이 가능하며, 사용자로 하여금 평면적인 사고를 유도한다.

3D에서는 가로와 세로뿐 아니라 제 3의 공간 축을 사용하여 자료를 배치할 수 있다. 평면에서도 자료의 배치는 다양한 방법이 가능하나 선형배치나 그림 1처럼 단순한 배치가 대다수의 시스템에서 사용되고 있다. 3D환경에서도 단순한 축의 추가에 따라서, 자료의 배치 방법을 그림 3처럼 확장 시킬 수 있다. 상대적으로 뒤에 출력되는 자료들은 앞의 자료들에 의해 일부 가려지게 되나, 배치 형태에 따라 보이는 부분을 제어할 수 있으며, 상대적으로 뒤나 앞

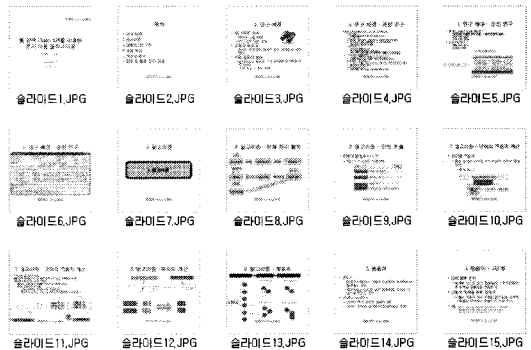


그림 1. 기존의 2D 뷰

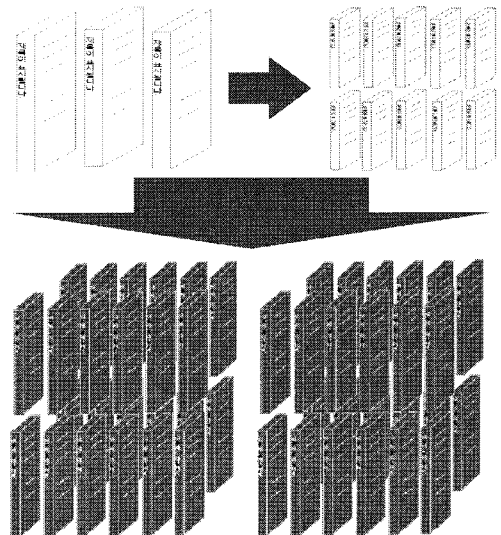


그림 2. 공간 축 확장에 따른 나열 방법 변경

에 있다는 위치가 정렬 상태나 폴더의 깊이 등을 나타내는 정보가 될 수 있다.

본 논문에서 제시하는 이런 단순한 설계에 따라 구현된 프로그램이 그림 3에 나타난다. 많은 영상정보가 한 화면에 출력되고 있으며, 순서상 뒤의 자료들이 화면상으로도 뒤에 출력 되고 있으며 일부 영상이 앞의 자료에 가려져 있고 크기가 작아 입체감을 느끼게 한다.

그러나 너무 많은 영상정보의 출력으로 인해 개별 정보의 해상도가 낮아지고 화면이 작아지는 문제가 발생한다. 이를 해결하기 위해 다양한 조합의 배치 방법으로 2D에서는 그림 4와 같은 방법이 사용되곤 하며 3D에서는 그림 5와 같은 배치 연구가 있었다.

본 논문에서는 다량의 영상정보의 출력으로 인한 이런 문제를 해결하기 위해, 3D GPU의 실시간 영상처리 기술을 활용하여 영상정보 사이를 실시간으로

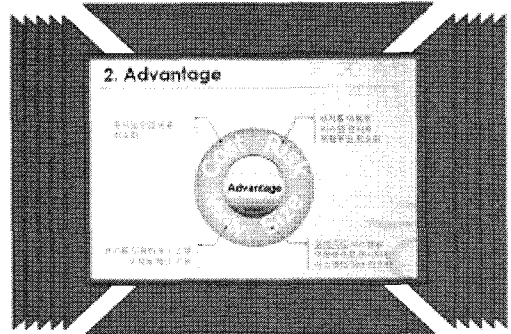


그림 5. 3D에서의 나열 방법 조합

탐색할 수 있도록 설계하였다. 즉, 사용자는 화면상의 자료를 이동시키거나 선택하기 위해 스크롤을 이동하는 것이 아니라, 자신(카메라)이 직접 영상정보들 사이를 이동하고, 원하는 정보에 가까이 다가감으로 탐색과 함께 화면 확대 효과를 동시에 얻게 된다. 그림 6처럼 카메라를 이동하면 그림 7처럼 관심 있는 영상정보를 더 자세하고 크게 볼 수 있으며, 관련된

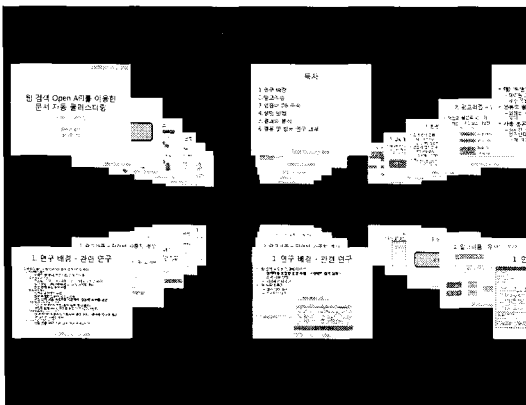


그림 3. 설계에 따라 구현된 3D 뷰

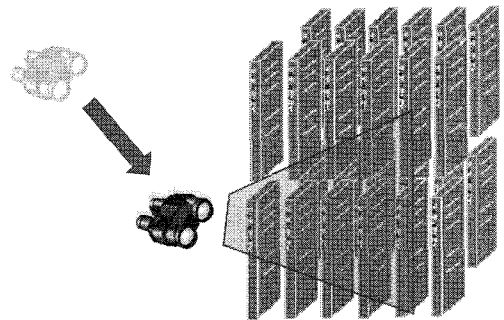


그림 6. 카메라 이동 기법

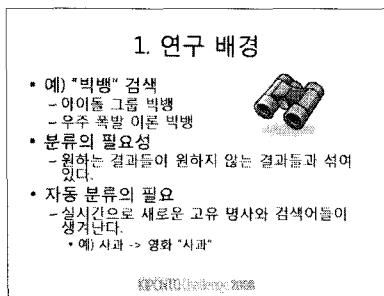


그림 4. 나열 방법의 조합 예

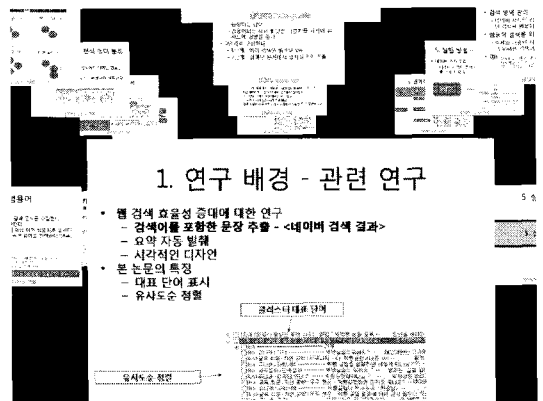


그림 7. 카메라 이동의 결과

주변의 정보들이 주변에 일부만 보이게 된다. 그림 6 과 그림 7은 원하는 정보가 있는 “1. 연구배경” 페이지를 더 자세히 보기 위해, 해당 페이지에 가까이 이동한 것이다.

카메라를 이용해 영상정보들 사이를 탐색하면서 사용자는 영상정보와 연결된 음성정보들을 들을 수 있다. 즉, 사용자는 복합정보들로 구성된 가상 3D 공간을 탐색하게 된다.

2.3 요구 기능 정의

영상정보는 3개의 축 방향으로 배치되며, 사용자는 카메라를 사용해 이들 사이를 자유롭게 이동하며 탐색할 수 있다. 탐색 과정은 실시간으로 영상 처리 되어 화면에 나타난다.

카메라의 이동은 키보드를 사용한다. 직관적인 조작을 위해 방향키와 PageUp/Down 키를 사용한다. 예를 들어 앞뒤 방향키로 카메라를 앞뒤로 이동시킬 수 있다. 또는 스페이스나 다른 키를 사용하여 영상 자료를 차례대로 선택할 수 있다.

영상정보들 사이에서는 배경 음성정보가 출력되며, 각 영상정보 별로 지정된 음성정보를 선택하여 들을 수 있다. 이것은 키보드를 사용하여 영상정보를 선택할 경우에 나타난다.

카메라는 이동과 함께 회전도 가능하다. 회전에는 다양한 기법이 존재하는데, 키보드와 마우스를 사용하는 일반 PC 환경에 경우 3D게임에서 많이 연구되고 있다. 일반적으로 마우스를 사용하여 회전시키며, 회전방식은 인체구조에 유사하게 한다. 즉, 턱을 들거나 고개를 좌우로 회전하듯이 카메라를 회전시키는 데, 좌우로 돌릴 때의 회전축은 고개를 들더라도 변하지 않으나, 고개를 좌우로 돌린 이후에는 턱을 들 때의 회전축은 처음과는 다르다. 사람을 모델로 하여 그림 8처럼 고개를 움직이더라도 몸의 방향은 변하지 않도록 한다. 몸의 방향이 변하지 않으면, 이동할 때의 방향 축들은 변하지 않는다. 즉 고개를 좌우로 돌리더라도 앞으로 걷는 것이지, 보는 방향으로 걷는 것이 아니다. 때문에 이동 방향은 변하지 않는다. 그림 8의 가장 우측의 시점과 같이 우측 위로 이동 후, 영상정보들을 내려다보았을 때의 화면 상태가 그림 9에 나타나있다.

시선 방향에 관계없이 걸어가는 방향은 유지되기 때문에 그림 9와 같은 방향으로 영상정보들을 바라보면서 앞으로 걸어가면 차례대로 모든 영상정보들을 탐색할 수 있다.

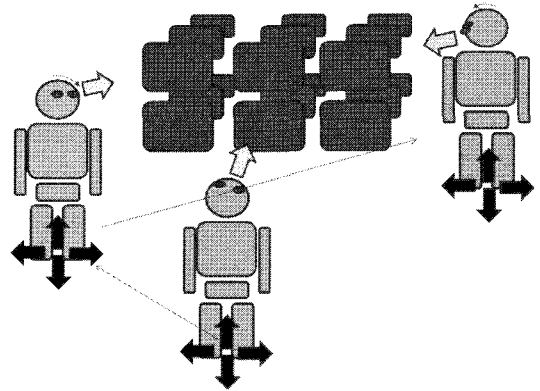


그림 8. 카메라 이동과 회전 모델

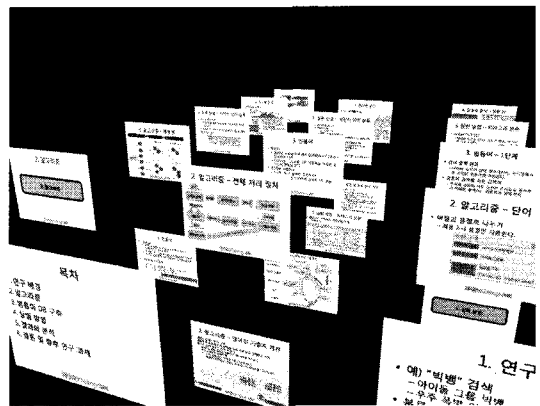


그림 9. 카메라 이동과 회전 결과

III. 3D 뷰 구현 및 실험

3.1 DirectX 3D & DirectX Input

빠른 실시간 화면 처리를 지원하는 대표적인 2가지 개발 툴로 DirectX와 OpenGL이 있다. 둘 사이에는 상당한 차이가 있고 서로 장단점이 있다. OpenGL은 3D 입문자에게 편리하고 다양한 플랫폼을 지원한다는 장점이 있고, DirectX는 Window만을 지원하지만 다양한 장점을 가지고 있다. 둘 사이는 기본적으로 좌표계부터 지원하는 함수들의 인터페이스까지 많은 부분에서 차이를 보인다. 본 논문에서 제시하는 설계를 구현하기 위해서는 둘 다 선택 가능하며, 3D개발에 주로 사용되는 C++과 Window에 맞추어 DirectX를 사용하여 프로그램을 작성하였다.

각 영상정보들은 직각사각형으로 나타나지만, 실제로 3D 기술에서는 모두 삼각형의 조합으로 처리한다. 따라서 2개의 삼각형을 사용하여 하나의 직각사

각형을 표현한다.

직각사각형의 좌표를 지정한 뒤에, 여기에 출력될 텍스처(영상정보)를 지정할 수 있다. 대부분의 이미지 파일 포맷을 지원하며, 내부적으로 로딩된 이미지는 저수준 포맷을 사용한다. 이미지를 사각형 틀에 맞추기 위해 몇 가지 세부적인 설정을 마치면 하나의 영상정보를 출력할 준비가 끝난다. 이런 작업을 모든 영상정보별로 반복한다. 즉, 3개의 축에 따라 영상정보 직각사각형의 좌표들을 계산한다.

모든 준비가 완료되면 DirectX는 GPU를 사용하여 화면을 빠른 속도로 구성한다.

카메라 이동 효과를 위해서는 화면에 출력되는 모든 좌표들을, 가상의 카메라 좌표에 맞추어 이동시켜야 한다. 이를 위해서 세계행렬(World Matrix)이 필요하며, 카메라의 좌표와 방향을 통해 세계행렬이 계산된다. 세계행렬을 통해 모든 삼각형들의 좌표를 실시간으로 계산한다.

$$\begin{aligned}
 W &= \text{World Matrix} \\
 Y &= Y\text{Axis Rotation Matrix} \\
 X &= X\text{Axis Rotation Matrix} \\
 T &= \text{Translation Matrix} \\
 W &= TYX
 \end{aligned}
 \tag{1}$$

카메라 회전 효과를 위해서도 세계행렬을 사용한다. 회전행렬을 구하여 행렬곱을 연산함으로써 화면이 회전하게 된다. 이 때 그림 8과 같은 모델로 구현하기 위해서는 회전 행렬을 계산할 때 좌우로 고개를 돌리기 위해 연산된 Y축 회전행렬을 먼저 계산한다. 그리고 고개를 들거나 낮추는 X축 회전행렬을 계산하여 행렬곱을 연산한다. 이렇게 구해진 회전행렬을 이동행렬에 곱함으로써 수식 1과 같이 세계행

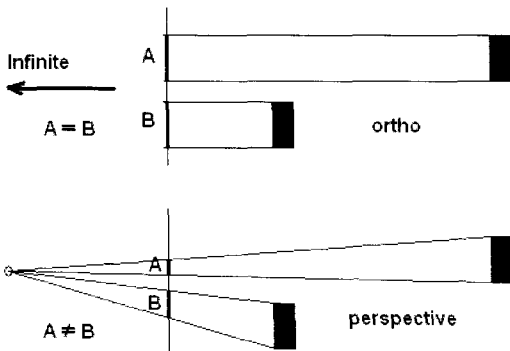


그림 10. 직각투영 및 원근투영의 차이

렬이 계산된다. 특정 축 회전행렬 계산과 이동행렬 계산은 DirectX에서 제공하는 함수를 사용한다.

투사법그림 10에 따라 멀리 있는 물체가 작게 보이는 효과를 얻을 수 있다. 이런 투사법에는 다양한 수치 설정이 필요한데, 가장 중요한 것은 시야각으로 본 논문에서는 45도로 설정한다. 이 경우 영상정보의 가로 길이를 a로 가정하면, 영상자료로부터 1.2a만큼 떨어질 때 전체 화면이 해당 영상자료로 가득 차게 된다.

음성정보를 처리하기 위해서 DirectShow를 사용한다. DirectX와 함께 Window 플랫폼에서 작동하며, mp3포맷을 지원한다.

3.2 프로그램 모듈들

로더: 사용자는 특정 폴더를 선택하여, 폴더 내부의 모든 영상정보(이미지 파일)를 로드할 수 있다. 또한 영상정보의 파일명에 따라 해당하는 음성정보(음악 파일)를 자동으로 로드하여 연결한다.

카메라: 키보드를 통해서 카메라를 제어함으로써 3D 가상공간 내부를 이동할 수 있다. 카메라의 가상 좌표와 방향을 통해 세계행렬을 계산하고, 이 행렬을 통해 모든 영상정보들의 좌표가 계산된다.

3D 슬라이드: 실제로 화면에 출력되는 각 영상정보들을 처리한다. 카메라로 계산되는 세계행렬을 통해 현재 영상자료의 현재좌표를 계산하고 출력한다. 카메라가 가까워지면 영상자료와 연관되는 음성정보를 재생한다.

배경 음성정보: 선택된 영상정보가 없을 경우에 출력된다. 전체 영상정보들을 대표하는 내용을 담는다. 프로그램이 처음 실행되었을 때이나 특정 키보드 입력에 의해 출력된다.

3.3 프로그램 인터페이스

키보드 방향키를 통해 앞뒤 좌우로 이동할 수 있는 직관적인 인터페이스를 사용한다. 마우스를 좌우로 이동하면 화면이 좌우로 회전하며, 앞뒤로 움직이면 위나 아래를 보도록 회전한다. 화면이 회전하더라도 이동 방향에 영향을 주지는 않는다.

특정 영상정보를 선택하기 위해서는 해당 영상 자료에 가까이 이동한 뒤 키보드에 Enter키를 누른다. 선택된 영상정보가 화면에 가득 차도록 카메라가 자동으로 이동하게 되며, 영상정보와 연결된 음성정보가 재생된다.

IV. 결론 및 향후 연구 방안

2D에 의한 정보접속은 3D에 비해 표현이 제한적이지만 비교적 제어가 쉬운 편이다. 반면에 3D에 의한 정보접속은 표현 제한이 없으며 정보전달 효과 면에서 상대적으로 높다. 그러나 3D 접속의 구현 방법은 다양한 분야의 기술이 필요로 하며 구현 기술도 다양하다.

3D접속 기술에서는 다양한 정보표현 방법을 복합적으로 활용할 수 있으며, 이를 활용한 영상과 음성 정보의 융합 설계를 연구하였다. 그 결과 3D 뷰에서 영상정보의 기초적인 배치와 음성정보와의 접목을 용이하도록 하였다. 기초적인 배치에서 발생하는 문제점을 해결하기 위해 가상 카메라 기법을 도입하였고, 카메라의 제어는 인체구조에 유사하도록 설계하였다.

본 논문에서 개발한 프로그램은 강의용 프리젠테이션으로 사용할 수 있으며, 학생 스스로 학습하기 위한 학습 데이터 작성용으로도 사용할 수 있다. 수업에 적용하여 학생들의 반응을 수집하는 방식으로 프로그램의 질적 평가를 할 수 있다.

향후 연구로는 3D 공간에서의 카메라의 현 위치와 각도를 표현하는 기법이다. 이 외에 경과시간 표현 등 다양한 부가 정보가 화면의 적정 위치에 표시될 수 있도록 하는 연구와, 발표자와 시청자가 복합 정보에 용이하게 접근할 수 있는 환경 구축도 연구할 예정이다.

현 카메라의 위치에서 가장 가까운 정보를 선택하는 기법으로 특정 정보를 선택하고 있으나, 보다 직관적이고 편리한 선택 인터페이스를 위해, 현재까지 널리 사용되고 있는 마우스 커서⁹⁾를 도입하거나, 화면상 중앙에 위치하는 정보의 선택기법 등 다양한 인터페이스를 연구하여야 한다.

참 고 문 헌

[1] George Robertson, Maarten van Dantzich, Daniel Robbins, Mary Czerwinski, Ken Hinckley, Kirsten Ridsen, David thiel, and Vadim Gorokhovsky, "The Task Gallery : A 3D Window Manager", CHI Letters, Vol.2, issue 1, pp.494-501, April, 2000.

[2] Russell J. Craig, Joel H. Amernic, "PowerPoint Presentation Technology and the Dynamics of Teaching", Innovative higher education, Vol.31, No.3, pp.147-160, August, 2006.

[3] Colin B. Price, "Unreal PowerPoint : Immersing PowerPoint presentations in a virtual computer game engine world", Computers in human behavior, Vol.24, No.6, pp.2486 - 2495, May, 2008.

[4] 김병수, 강병익, "OpenGL을 이용한 3D 뷰어 개발", 2002년 한국멀티미디어학회 추계학술발표논문집, pp.776 - 779, 2008.

[5] <http://www.opengl.org/sdk/docs/>

[6] 이종수, Dinh Quang Truong, 박진우, 황철희, Nguyen Kiem Hieu, 노창수 효과적 발표를 위한 문서의 3차원 뷰 및 접속, 2008年度 秋季 合同 學術論文發表會 論文集, pp.181 - 184, 2008.

[7] Mehrnaz Sadat Akhavi, 3D Visualization of Hierarchical Clustered Web Search Results, Computer Graphics Imaging and Visualisation (CGIV 2007)

[8] Andrade, B., Almeida, R., Cubaud, P., Dupire, J., Natkin, S., Topol, A.: Experiments towards 3D Immersive Interaction for Digital Libraries. In: Pan, Z., Aylett, R.S., Diener, H., Jin, X., Göbel, S., Li, L. (eds.) Edutainment 2006. LNCS, Vol.3942, pp.1348 - 1357. Springer, Heidelberg (2006)

[9] <http://bumptop.com/>

신 준 철 (C. H. Shin)

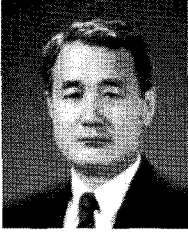
정희원



2007 울산대학교 공학사.
2009 울산대학교 공학석사.
<관심분야> 정보검색, 기계학습, GUI

이 종 수 (J. S. Lee)

종신회원



1973 서울대학교 공학사

1981 미국 VPI & State Univ.
공학석사

1985 미국 VPI & state Univ.
공학박사

현재 울산대학교 교수

<관심분야> 멀티미디어 영어학

습 콘텐츠 물체인식 및 추적기법