

---

# 인터랙티브 3D 플로팅 영상 포인팅 장치의 구현

신동학\* · 김은수\*\*

Implementation of interactive 3D floating image pointing device

Dong-Hak Shin\* · Eun-Soo Kim\*\*

---

본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 육성·지원사업  
(IITA-2008-C1090-0801-0018)에 의하여 수행되었음

---

## 요 약

본 논문에서는 3D 환경에서의 사용을 위한 새로운 형태의 인터랙티브 3D 플로팅 영상 포인팅 장치를 제안한다. 이 장치는 플로팅 렌즈 배열을 사용한 3D 플로팅 영상 재생 시스템과 실시간 손가락 검출 기반의 버튼형 인터페이스로 구성된다. 제안하는 장치에서는 사용자가 손가락을 이용하여 3D 공간에 플로팅된 영상 배열에서 원하는 영역을 지시하고, 2대의 카메라를 이용하여 사용자의 손가락 인식하여 원하는 이벤트를 발생하는 인터랙션 기능을 효과적으로 수행한다. 제안하는 시스템의 유용함을 보이기 위해서, 실험적인 장치 구현을 수행하고 기초적인 실험 결과를 보고한다.

## ABSTRACT

In this paper, we propose a novel interactive 3D floating image pointing device for the use of 3D environment. The proposed system consists of the 3D floating image generation system by use of a floating lens array and the a user interface based on real-time finger detection. In the proposed system, a user selects single image among the floating images so that the interaction function are performed effectively by pressing the button event through the finger recognition using two cameras. To show the usefulness of the proposed system, we carry out the experiment and the preliminary results are presented.

## 키워드

3D display, Human computer interaction, Lens array, Floating image array

## I. 서론

산업의 발전과 함께 컴퓨터는 급속도로 발전하여 왔고 현 시대에 컴퓨터는 생활에 꼭 필요한 필수품으로 자리매김하고 있다. 따라서 생활에서 컴퓨터를 사용하는

시간의 비중이 늘어나고 최근 유비쿼터스(ubiquitous)라는 개념이 사회적으로 이슈가 됨에 따라 키보드와 마우스만을 이용한 단조로운 인터랙션(interaction)으로부터 탈피하고자 하는 인간의 욕구로 인해 기존의 많은 장비에 유비쿼터스 개념을 도입한 새로운 장비들이 속속들

---

\* 동서대학교 영상콘텐츠학과  
\*\* 광운대학교 차세대 3D 디스플레이 연구센터

이 개발되어 지고 있다 [1-3].

현재까지의 컴퓨터는 텍스트(text)기반 아래 명령어 위주의 인터페이스에서부터 GUI에 이르면서까지 컴퓨터와의 인터랙션을 위한 여러 가지 디바이스가 생겨났다. 현재 우리가 가장 일반적으로 사용하는 포인팅 디바이스 (pointing device)인 마우스를 비롯하여 트랙볼, 타블렛, 펜 마우스, 터치패드, 터치스크린, 자이로마우스 등 많은 포인팅 디바이스들이 개발되어 왔다. 그러나 이러한 디바이스 들은 모두 터치(touch)형식의 단조로운 디바이스들이었으며 이는 디바이스의 외형에만 약간의 변화를 주었을 뿐, 포인팅 인식방법이나, 지역적 제한으로 인한 사용상의 불편함을 가지고 있을 뿐만 아니라, 유비쿼터스 환경 구축을 위한 현실에도 뒤쳐지는 결과를 초래하고 있다.

최근 보다 현실감 있는 컴퓨터와의 인터랙션을 위하여 3D 입체 영상 디스플레이 장치와 연동한 인터랙션 디바이스에 대한 연구가 활발히 진행되고 있다 [3-6]. 이는 기존의 2D 평면 디스플레이 환경에서 벗어나 실제의 3D 입체 공간 환경을 인간에게 제공하고자 하는 것이다. 따라서 이러한 환경 구성을 위하여 컴퓨터 영상처리 기술인 3D 비전 분야의 기술을 많이 활용하고 있다. 이 기술은 경제적이면서도, 정밀도 또한 높은 인터랙션 시스템 환경 구축이 가능하다는 장점이 있다. 이러한 3D 입체 영상을 이용한 인터랙션 기술은 가상현실 및 유비쿼터스 환경 구성과 같은 여러 다양한 분야에서 적용되어지고 있다.

최근 3D 입체 디스플레이는 가상현실에서 많은 관심을 가지고 있다. 가상현실에 사용하기 위한 3D 입체 디스플레이 장비들에 대한 많은 연구들이 진행되고 있다 [3-6]. 그러나 지금까지의 대부분의 연구는 깊이(depth)를 느끼게 하기 위해서 특별한 안경을 사용하는 스테레오 디스플레이에 관한 연구들이었다. 이 방식은 시스템적으로 간단하다 할지라도 인간에게 어지러움을 주는 근본적인 문제점을 가지고 있어서 장시간의 사용이 불가능하다. 따라서 특별한 장치의 사용 없이 영상을 공간에 입체적으로 표시하는 방식에 대해서 많은 연구가 진행되고 있다.

이들 기술 중에서 하나 이상의 렌즈나 거울을 이용하여 공간에 실제 깊이를 가지는 3D 입체 영상을 디스플레이하는 플로팅 (floating) 영상 디스플레이 방식이 있다 [7,8]. 이 방식은 단순한 광학 장비들을 이용하여 영상을

일정공간에 생성시키고 자연스러운 영상을 관측할 수 있는 방법이며, 다양한 형태의 시스템이 활발히 연구되고 있다. 실제 깊이의 넓은 시야각의 3D 입체 영상 디스플레이 장치로써 최근 100인치 급의 Fresnel 렌즈를 사용한 3D 플로팅 영상 디스플레이가 보고되었다 [7].

본 논문에서는 3D 환경에서의 사용을 위한 새로운 형태의 인터랙티브 (interactive) 3D 플로팅 영상 포인팅 장치를 제안한다. 이 장치는 플로팅 렌즈 배열을 사용한 3D 플로팅 영상 재생 시스템과 실시간 손가락 포인팅 알고리즘을 이용한 유저 인터페이스로 구성된다. 제안하는 장치에서는 사용자가 손가락을 이용하여 3D 공간에 플로팅된 영상 배열에서 원하는 영역을 지시하고, 2대의 카메라를 이용하여 사용자의 손가락 인식하는 인터랙션 기능을 효과적으로 수행한다. 제안하는 시스템의 유용함을 보이기 위해서, 실험적으로 시스템을 구현하고 기초적인 실험을 수행하였다.

## II. 제안하는 장치의 구조

그림 1은 본 논문에서 제안하는 인터랙티브 3D 플로팅 영상 포인팅 장치의 개략도를 보여준다. 이 장치는 크게 두 개의 시스템으로 구성되어 있다. 첫 번째 시스템은 플로팅 렌즈 배열을 이용한 3D 입체 영상 디스플레이 시스템이고, 두 번째 시스템은 사용자의 손가락을 인식하는 포인팅을 수행하는 유저 인터페이스이다. 3D 입체 영상 디스플레이 시스템은 저가의 Fresnel 렌즈 배열을 이용하여 3D 영상 배열을 일정한 깊이의 공간에 생성하게 한다. 이에 반해 포인팅 유저 인터페이스는 사용자의 손가락 검출을 위하여 2대의 카메라로 구성되어 있으며, 검출의 효율을 높이기 위하여 검출라인을 사용한다. 그리고 검출된 손가락의 정보를 이용하여 컴퓨터와의 인터랙션 기능을 수행한다.

### 2.1 3D 플로팅 영상 배열의 생성

제안하는 인터랙티브 3D 플로팅 영상 포인팅 장치를 구현하기 위해서 먼저 본 논문에서는 3D 플로팅 영상 배열을 생성해야 한다. 이 영상 배열은 얻기 위해서는 그림 1에서 보여지듯이 디스플레이 장치와 렌즈배열이 필요하다. 일반적으로 플로팅 영상 디스플레이 장치는 공간에 실제 깊이의 영상을 표시하며 관측자는 특별한 보조

장치 없이 직접적으로 3D 영상을 볼 수 있는 장점이 있다. 이 플로팅 영상 디스플레이 시스템은 영상 표시장치로 LCD와 같은 기존의 2D 영상 장치를 사용한다. 따라서 기존의 잘 알려진 소프트웨어 및 비디오 카마라를 그대로 사용할 수 있다. 그리고 렌즈 배열은 저가의 Fresnel 렌즈 배열을 사용할 수 있다.

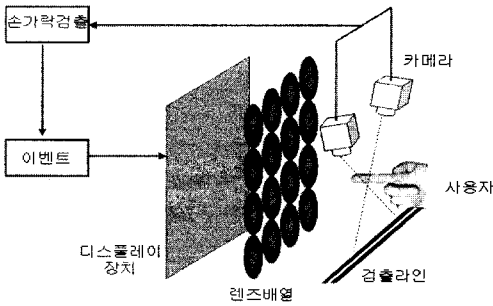


그림 1. 제안하는 인터랙티브 3D 플로팅 영상 포인팅 장치  
Fig. 1. Proposed interactive 3D floating image pointing device

플로팅 영상 디스플레이 시스템의 자세한 이해를 위해서 기존의 방법을 설명하다. 그림 2는 기존의 플로팅 영상 디스플레이 시스템의 동작 원리를 보여준다. 그림 2에서 보여지듯이 양의 초점거리를 가지는 Fresnel 렌즈를 LCD 패널 앞에 위치시켜서 시스템을 구현한다. Fresnel 렌즈는 디스플레이 장치에 표시된 영상을 관측자의 앞 공간에 실제 깊이 3D 영상으로 투영시켜주는 역할을 한다. 그림 2에서 디스플레이 패널과 Fresnel 렌즈 사이의 거리를  $h$  라고 하고, Fresnel 렌즈의 초점거리를  $f$ 라고 하자. 그러면 플로팅되는 3D 영상은 식 (1)에 표현되는 거리에 나타난다.

$$H = \frac{hf}{h-f} \quad (1)$$

기존의 플로팅 영상 디스플레이 방식은 하나의 깊이 평면을 공간에 만들기 때문에 실제 깊이감이 다른 여러 영상을 표시할 수 없다. 따라서 이러한 문제점을 보완하기 위한 다양한 시스템들이 연구되고 있다 [6].

제안하는 장치에서의 3D 플로팅 영상 배열은 그림 3과 같이 Fresnel 렌즈 배열을 이용하여 얻어진다. 각각의 렌즈는 대응하는 2D 영상을 공간에 서로 다른 위치에 포

시한다. 그림 3은 4x4의 3D 플로팅 영상을 생성하는 예를 보여주고 있다. 각각의 영상은 대응하는 Fresnel 렌즈를 통해서 볼 수 있다. 이때 관측되는 영상들은 2D 영상의 역상으로 관측이 된다.

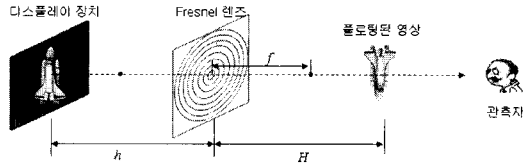


그림 2. 플로팅 영상 디스플레이의 원리  
Fig. 2. Principle of floating image display

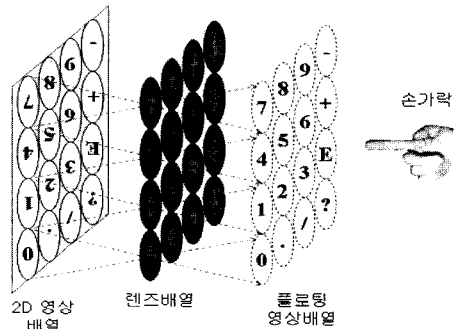


그림 3. 2D 영상 배열과 3D 플로팅 영상 배열과의 관계  
Fig. 3. Relation between 2D image array and 3D floating image array

## 2.2. 손가락 검출 기반의 버튼형 인터페이스

본 논문에서 제안하는 3D 플로팅 영상 포인팅 장치의 올바른 동작을 위해서는 3D 플로팅 영상 배열과 사용자 사이의 인터랙션을 위한 버튼형 인터페이스를 사용한다. 이 인터페이스 방식을 구현하기 위해서 작업 순서도는 그림 4와 같다.

본 논문에서 사용하는 인터페이스의 첫 과정은 2대의 카메라를 이용하여 사용자의 손가락 영상을 캡처하는 것이다. 이때 2대의 카메라는 그림 1에서 보여지듯이 렌즈 배열 앞의 위쪽으로 위치한다. 이때 2대의 카메라로부터 손가락의 위치를 검출하는 개략도를 그림 5에 나타내었다. 2대의 카메라는 렌즈 배열의 특정한 기준점을 향해서 모이도록 설치한다. 여기서 3D 플로팅 영상 배열이 놓이는 위치가 고정되어 있기 때문에 그 위치로 손가락 검출이 용이하도록 흰색과 검은색의 검출라인을 사용한다.

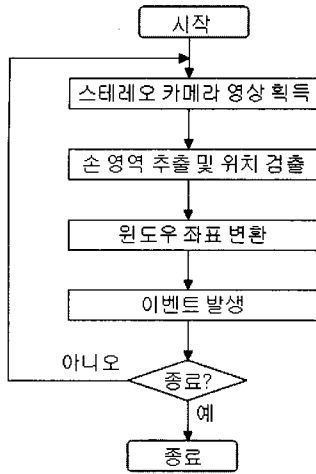


그림 4. 인터페이스 방식을 구현하기 위해서 작업 순서도  
 Fig. 4. Operational flowchart to implement the user interface

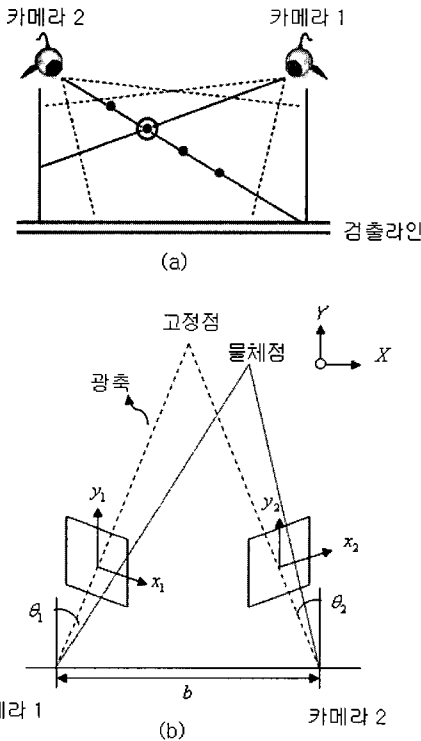


그림 5. (a) 인터페이스의 구조 (b) 교차시각 스테레오 카메라 모델  
 Fig. 5. (a) Structure of user interface (b) convergent stereo camera model

다음으로 캡처된 두 영상에 대해서 손가락의 영역을 추출한다. 이때 검출라인을 사용하기 때문에 쉽게 손의 영역을 구분해 낼 수 있다. 손가락의 경우에는 특정한 색상영역을 가지기 때문에 주어진 영상에 대해서 손 생성 영역에 속하는 픽셀과 속하지 않은 픽셀로 역치화한다. 그리고 추출된 영역에 대해서 평균값을 구하여 두 영상에 대한 손가락 영역의 대응하는 두 카메라의 영상 좌표계를 얻었다. 즉, 카메라 1에 대해서는  $(x_1, y_1)$ , 카메라 2에 대해서는  $(x_2, y_2)$ 의 좌표 값을 실험적으로 얻는다.

본 논문에서 사용하는 2대의 카메라가 이루는 구조는 교차 스테레오 카메라 모델 (convergent stereo camera model)과 동일하다 [9,10]. 따라서 실험적으로 얻은 두 영상 좌표계를 이용하여 손가락의 기준 좌표계  $(X, Y, Z)$ 의 값을 유도할 수 있다. 따라서 교차시각 스테레오 카메라 모델에 대해서 좀 더 자세히 설명하기 위해서 그림 5(b)의 카메라 모델을 고려하자. 그림 5(b)에서 보여지듯이 광축이 평행하지 않고, 카메라 1의 광축은 오른쪽으로  $\theta_1$ 만큼, 카메라 2는 왼쪽으로  $\theta_2$ 만큼 기울어진 구조를 갖고 있어서 한 고정점 (fixation point)에서 만나는 구조이다. 그리고 초점거리  $f$ 와 두 렌즈 중심사이의 거리는  $b$ 이다. 따라서 이 구조에서 기록된 두 영상에 대해서 한 물체(손가락) 대응하는 점들의 다음과 같이 유도될 수 있다. 카메라 1에 대해서는

$$x_1 = f \frac{(X+b/2)\cos\theta_1 - Z\sin\theta_1}{(X+b/2)\cos\theta_1 + Z\sin\theta_1},$$

$$y_1 = f \frac{Y}{(X+b/2)\cos\theta_1 + Z\sin\theta_1} \quad (2)$$

이고, 카메라 2에 대해서는

$$x_2 = f \frac{(X-b/2)\cos\theta_2 + Z\sin\theta_2}{-(X+b/2)\cos\theta_2 + Z\sin\theta_2}$$

$$y_2 = f \frac{Y}{-(X+b/2)\cos\theta_2 + Z\sin\theta_2} \quad (3)$$

가 된다. 식 (2)와 식 (3)을 이용하여 영상좌표계와 카메라 변수들로 기준 좌표계  $(X, Y, Z)$ 는 다음과 같이 유도된다.

$$X = \frac{b(x_1 \cos \theta_1 + f \sin \theta_1)(x_2 \cos \theta_2 + f \sin \theta_2)}{T} - \frac{b}{2} \quad (4)$$

$$Y = \frac{by_1(x_2 \sin \theta_2 + f \cos \theta_2)}{T} \quad (5)$$

$$Z = \frac{b(-x_1 \sin \theta_1 + f \cos \theta_1)(x_2 \sin \theta_2 + f \cos \theta_2)}{T} \quad (6)$$

여기서

$$T = f(x_1 - x_2) \cos(\theta_1 + \theta_2) + x_1 x_2 \sin(\theta_1 + \theta_2) + f^2 \sin(\theta_1 + \theta_2)$$

이다. 이 수식에서  $\theta_1 = \theta_2 = 0$ 이면 평행 스테레오 모델이 얻어진다.

따라서 그림 5(a)의 실험으로부터 얻은 두 대응 영상 좌표계  $(x_1, y_1)$ 와  $(x_2, y_2)$ 를 이용하여 손가락의 좌표계를 식(4), (5),와 (6)을 이용하여 구한다. 특히 제안하는 방식에서  $Y$ 의 값보다는  $X$ 와  $Z$ 의 값이 손가락의 정확한 위치를 측정하는데 주요한 변수가 된다.

다음 단계로 손가락의 위치에 대응하는 이벤트를 발생시키기 위한 윈도우 좌표축으로 변환이 이루어져야 한다. 이를 위하여 그림 6의 두 종류의 좌표축을 고려하자. 하나는 카메라에서 얻어진  $(X, Y, Z)$  좌표축이고 다른 하나는 윈도우(window) 좌표축이다. 그림 5(a)에서 보여 지듯이  $X$ 와  $Z$ 의 축만을 고려하여 윈도우 좌표축을 계산할 수 있다. 그러면 그림 6에서와 같이 카메라의 좌표축의 한 점  $P(X, Z)$ 를 윈도우 좌표축에서 한 점  $p(u, v)$ 로 계산한다. 그림 6(a)의  $(X, Z)$ 의 위치에 손가락 위치를 정하고, 최대 유효 범위를  $(1, X_{max})$ ,  $(1, Z_{max})$ 라고 하면, 윈도우 좌표축은 다음과 같이 간단히 구해진다.

$$\{u, v\} = \left\{ w \left( \frac{X}{X_{max}} \right), w \left( \frac{Z}{Z_{max}} \right) \right\} \quad (7)$$

구해진 윈도우의 좌표축에 대해서 적절한 이벤트를 수행함으로써 손가락검출 기반의 버턴형 인터페이스를 완성한다.

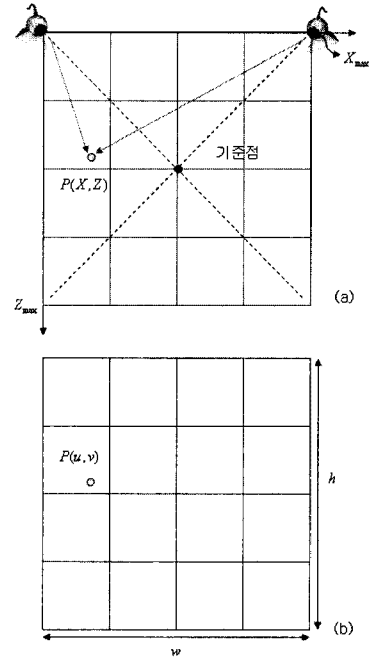


그림 6. (a) 카메라 좌표축 (b) 윈도우 좌표축  
Fig. 6 (a) Camera coordinate (b) Window coordinate.

### III. 실험 및 결과

제안하는 인터랙티브 3D 플로팅 영상 포인팅 장치의 유용함을 보이기 위해서 기초적인 실험을 수행하였다. 전체 구현된 시스템은 그림 7에 나타내었다. 실험에 사용한 Fresnel 렌즈 배열은 4x4의 렌즈로 구성되어 있고 각 렌즈의 초점거리는  $f=20\text{ mm}$ 이고 직경은  $20\text{ mm}$ 이다. 그림 8은 3D 플로팅 영상 배열을 생성하기 위해서 사용하는 2D 영상을 나타내었다. 각각의 렌즈에 대응되도록 4x4의 아이콘 영상들로 구성되어 있다. 각각의 2D 아이콘 영상은 대응하는 Fresnel 렌즈를 통하여 역상으로 보이기 때문에 이를 고려하여 제작되었다. 그림 8(b)는 실제 실험으로 얻어진 광학적 영상을 보여주고 있다. 렌즈 배열과 2D 영상 사이의 거리는  $h=40\text{ mm}$ 로 하였다. 그러면 식 (1)에 의해서 3D 플로팅 영상은 렌즈 배열로부터  $H=40\text{ mm}$ 에 위치하게 된다.

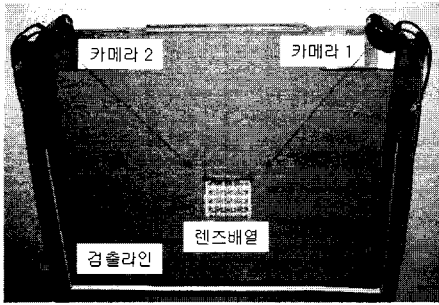
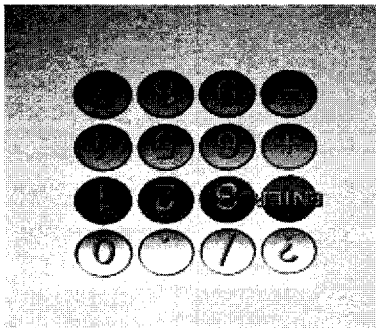
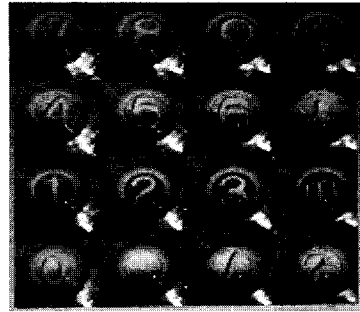


그림 7. 제작된 시스템 구조  
Fig. 7. System structure

생성된 3D 플로팅 영상 배열에 대해서 2.2장에서 설명한 2대의 카메라를 이용하는 버튼형 인터페이스를 구성하여 인터랙션 기능을 구현하였고 사용자의 손가락 검출에 대한 실험을 수행하였다. 그림 7은 본 논문에서 구현한 시스템의 개략도를 보여준다. 2대의 카메라는  $b=56\text{cm}$ 만큼 떨어져 있으며, 각각 렌즈 배열의 중앙점을 향하도록 하였고, 따라서 교차시각 카메라 모델에서의 두 각도인  $\theta_1$ 과  $\theta_2$ 는 모두  $45^\circ$ 이었다. 실험에 사용한 2대의 카메라는  $320 \times 240$  화소의 영상을 초당 30 프레임의 실시간으로 손가락 영상을 검출한다. 영상의 잡음을 최소화하기 위해서 카메라는 흑백 영상으로 감지하였다. 결과적으로 그림 9(b)와 같이 상당히 깨끗한 손가락 검출 결과를 얻을 수 있었다. 그림 9는 두 카메라의 촬영된 영상들의 예를 보여준다. 손가락의 움직임이 있는 경우와 없는 경우를 보여주고 있다. 검출라인의 사용은 보다 정확하고 빠른 손가락의 위치 정보를 제공하였다. 최종적으로 제작한 시스템을 이용하여 그림 4의 작업순서도를 이용하여 사용자에게 의한 이벤트 발생에 대한 실험을 성공적으로 수행할 수 있었다.



(a)



(b)

그림 8. (a) LCD에 표현되는 2D 영상 (b) 실험으로 얻어진 3D 플로팅 영상 배열  
Fig. 8. (a) 2D image displayed in LCD (b) 3D floating image array captured from experimental setup

#### IV. 결 론

현재 3D 환경에서의 대화형 인터랙션을 위한 수단으로 다양한 방법들이 연구되고 있지만 지역적이며, 제한적인 환경을 탈피하지 못할 뿐만 아니라, 고가의 장비만을 활용한 대처방안 등은 사용자의 욕구를 충족시키지 못하고 있다. 본 연구에서는 인터랙션 3D 디스플레이의 응용 분야를 위한 새로운 시도으로써 3D 환경에서의 사용을 위한 새로운 형태의 인터랙티브 3D 플로팅 영상 포인팅 장치를 제안하고 실험하였다. 이 장치는 플로팅 렌즈 배열을 사용한 3D 플로팅 영상 재생 시스템과 실시간 손가락 포인팅 알고리즘을 이용한 유저 인터페이스로 구성되어 있으며, 사용자가 손가락을 이용하여 3D 공간에 플로팅된 영상 배열에서 원하는 영역을 지시하고, 2대의 카메라를 이용하여 사용자의 손가락 인식하는 인터랙션 기능을 효과적으로 수행함을 실험적으로 확인하였다. 현재 시스템은 간단한 성능 테스트형의 초보적인 제작되었지만, 향후 몇몇 성능의 개선 보완한다면 단지 영화에서처럼 허공에 손짓만으로 컴퓨터에 사용자가 원하는 인터랙션을 전달 할 수 있는 시스템을 연구개발함으로써 유비쿼터스 시스템에 좀 더 다가서는 시스템으로 발전할 것으로 기대한다.

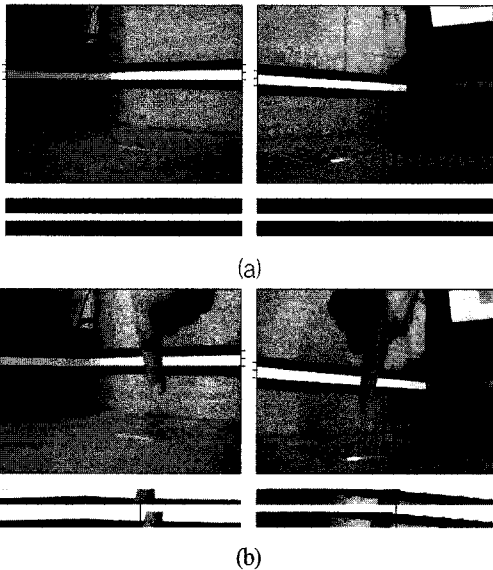


그림 9. (a) 검출라인만 촬영된 영상  
(b) 손가락 움직임을 촬영한 영상

Fig. 9. (a) Captured images only with capture line.  
(b) Captured images when a finger appear.

### 참고문헌

[1] V. I. Pavlovic, R. Sharma, T. S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: a Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 677-695, 1997.

[2] M. Adcock, M. Hutchins, and C. Gunn, "Augmented Reality Haptics: Using ARToolKit for Display of Haptic Application," 2nd International Augmented Reality Toolkit Workshop, Tokyo, Japan. 2003.

[3] D. Ebert, E. Bedwell, S. Maher, L. Smoliar, and E. Downing, "Realizing 3D visualization using crossed-beam volumetric displays," *Communications of the ACM*, vol. 42, no. 8, pp. 101-107, 1999.

[4] R. Balakrishnan, G. Fitzmaurice, and G. Kurtenbach, "User interfaces for volumetric displays," *IEEE Computer*, pp. 37-45. 2001.

[5] S. Pastoor and M. Wopking "3-D displays: A review of current technologies," *Displays*, vol. 17, no 2, pp. 100-110, 1997.

[6] M. Halle, "Autostereoscopic displays and computer graphics," *Computer Graphics, ACM SIGGRAPH*, vol. 31, no. 2, pp. 58-62, 1997.

[7] S.-J. Jang, S.-C. Kim, J.-S. Koo, J.-I. Park and E.-S. Kim, "100-inch 3D real-image rear-projection display system based on Fresnel lens," *Proc. SPIE*, vol. 5618, pp. 204-211, 2004.

[8] S. -W. Min, M. Hahn, J. Kim, and B. Lee, "Three-dimensional electro-floating display system using an integral imaging method," *Opt. Express* vol. 13, pp. 4358-4369, 2005.

[9] E. Krotkov, K. Henriksen, and R. Kories, "Stereo ranging with verging cameras," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-12, no. 12, pp. 1200-1205, 1990.

[10] 이적석, "교차시각 스테레오 카메라 모델 해석" 대한전자공학회 논문지, 제33권 10호, pp. 1536-1548, 1996.

### 저자 소개

신 동 학(Dong-Hak Shin)  
한국해양정보통신학회논문지 제11권 제5호 참조

김 은 수(Eun-Soo Kim)  
한국해양정보통신학회논문지 제11권 제5호 참조