# 일상적 행동양식을 통한 인터페이스의 구현

안종윤[1],이경원[2]
아주대학교 미디어학부[1 2]
{amaranth[1],kwlee[2]}@ajou.ac.kr

## Implementing user interface through everyday gesture

Jong Yoon Ahn[1], Kyung Won Lee[2]
Division of media, Ajou University[1 2]

**요 약**

컴퓨터와 인간사이의 원활한 의사소통 및 인터랙션을 위해 기존의 키보드, 마우스를 대체할 수 있는 다양한 입력장치들이 개발되고 있다. 하지만 정보를 탐색, 접근하는 데에 있어서 기존의 장치들은 클릭과 같은 제한적인 동작만을 입력 값으로 받아들이므로 이러한 방식에 익숙하지 않은 사용자의 입장에서는 부자연스러움을 느끼는 요인이 된다. 사용자의 제스처를 인식할 수 있는 인터페이스를 통해 일상에서 사물을 사용할 때의 행동양식을 그대로 가져올 수 있다면, 디지털 콘텐츠에 접근하는데 있어 보다 직관적이고 편리하게 컴퓨터와 의사소통 될 수 있다. 제스처는 동작의 자율성이 높고 때로 그 의미를 파악하기 모호하기 때문에 동작들을 정확히 인식하여 구분할 필요가 있다. 본 논문에서는 이를 바탕으로 효과적인 제스처 인터페이스의 구현을 위해 필요한 점들을 살펴보고, 기술적 구현을 통해 디지털 콘텐츠와의 인터랙션을 보여주고자 한다. 정보 접근에 있어 가장 익숙하고 전통적이라 할 수 있는 책의 메타포를 통해 페이지를 넘기는 행동양식을 인식할 수 있는 인터페이스를 개발하고 이를 입력장치로 사용한다. 사용자의 동작을 인식, 파악하여 책을 앞뒤로 넘기거나 탐색하며 원하는 정보에 접근할 수 있도록 유도하고 손 동작을 통한 인터페이스를 수단으로 컴퓨터와의 유연한 의사소통이 가능하도록 구현한다.

Keyword : HCI, Gesture Interface, Interaction, Reading

## 1. 서론

기술의 발달에 의한 소통양식의 변화로 사람과 컴퓨터 사이의 다양한 의사소통 방법이 연구되며, 보다 현실적이고 자연스러운 인터랙션을 위한 새로운 패러다임이 형성되고 있다. 기존의 컴퓨터를 이용한 인터랙션 디자인에 있어서는 키보드나 마우스를 이용하여 사용자가 모니터를 통해 인터페이스를 조작하는 제한적 형태가 일반적으로 사용되었다. 그러나 이러한 상호작용 방식은 실제 우리가 살고 있는 일상의 물리적 환경이나 행동양식과는 차이점을 보이고 있다.

이를 극복하기 위해 HCI 분야를 중심으로 다른 형태의 입력방식을 만들고 구현하는 인터페이스에 대한 연구가 진행되고 있으며, 나아가 디자인 및 예술 분야에서도 사용자가 직접 참여하여 적극적으로 기기를 작동시켜야만 예술적 표현이 완성되는 형식의 작품들도 선보이고 있다.

일상생활에서 자연적으로 발생되는 형태의 인터페이스는 정보접근과 커뮤니케이션의 이해력을 높일 수 있다. 특히 인류가 물리적인 도구를 만들고 다루며 손짓과 몸동작 등의 제스처를 통해 풍부한 의사소통 문화를 이룬 점을 돌이켜볼 때 이러한 행동양식을 적용한 인터페이스는 가장 자연스러운 형태라 할 수 있다.

일반적으로 제스처란 “생각이나 감정, 태도 등을 표현 혹은 강조하기 위해 몸이나 팔다리를 사용한 움직임”이라 정의되고 있다[1]. 제스처는 말을 하거나 지시 동작을 취하는 과정에서 자연스럽

게 배어 나오며, 그 의미를 쉽고 빠르게 이해할 수 있도록 의도적인 동작을 취할 때 나타나기도 한다. 사람 사이의 의사소통에 있어서 음성 언어가 겨우 7%만 사용되는 반면, 38%가 제스처나 목소리 톤과 같은 준 언어적 방법을 통해 전달된다는 점을 보면 우리의 생활에서 의사전달을 위해 제스처가 얼마나 많이 사용되는지 알 수 있다[2].

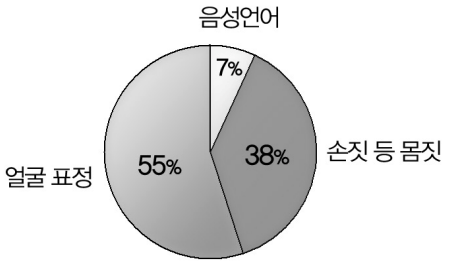


그림 1. 의사 소통과 정보 전달

따라서 본 논문에서는 보다 직관적이고 자유로운 인터랙션을 통해 원하는 정보에 다가갈 수 있도록 제스처를 파악하여 컴퓨터와 의사소통 할 수 있는 방안에 대해 연구하고자 한다.

일반적으로 제스처는 한 번에 여러 가지 동작의 동시 수행이 가능하고, 끊임없이 동작들을 연결시켜 표현할 수 있기 때문에 이를 기반으로 한 인터페이스를 제작할 때 어려움이 있다. 즉 제스처의 높은 자유도(degree of freedom)와 그 불명확함으로 인해 사용자의 의도와는 전혀 다른 결과를 야기할 수 있으므로 각각의 동작을 잘 파악하고 명확히 구분 지을 수 있는 시스템을 통해 가치를 높여야 한다[1].

이러한 문제점들을 해결하기 위해 먼저 제스처의 특성을 파악한 후, 관련 사례를 통해 개선해 나아가야 할 부분들을 알아본다. 그리고 정보 접근에 있어 가장 익숙하고 전통적이라 할 수 있는 책을 읽는 행동양식을 통해서 페이지를 넘기는 제스처로 디지털 콘텐츠에 접근할 수 있는 인터페이스를 개발하고 이를 입력장치로 활용하는 방안을 제안한다.

## 2. 제스처 인터페이스

### 2.1. 제스처의 분류

자연스러운 인터랙션을 위한 제스처 인터페이스는 반복되는 특정 동작들을 인식하여 그에 맞는 기능을 수행함으로써 가능하다. 따라서 실생활에서 다양한 목적의 의사소통 수단으로 사용되는 제스처를 분류하여 정의할 필요가 있다.

제스처는 먼저 그 의도성에 따라 의도가 있는 제스처(gestures)와 의도가 없는 제스처(unintentional movements)로 분류할 수 있으며, 의도가 없는 제스처는 어떤 의미 있는 정보도 전달하지 않는 몸이나 팔다리의 동작을 의미한다. 의도가 있는 제스처는 의사소통을 위한 제스처(Communicative)와 물체를 조작하기 위한 제스처(Manipulative)로 나뉘며, 의사소통을 위한 제스처에는 수화와 같은 인공적인 제스처와 상징적 기호 그리고 자연발생적인 제스처가 있으며 이는 공간상의 어떠한 위치, 물체를 가리키는 지시(Deictic), 행위나 형태를 묘사적으로 표현하는 아이콘(iconic), 반복의 표현을 위해 손을 계속 돌리는 행동처럼 추상적인 동작을 나타내는 은유(Metaphoric), 정해진 형태는 없지만 의미를 지니고 있는 두드림(beat) 등으로 구분된다. 한편 물체를 조작하기 위한 제스처는 대상이 되는 현실세계의 물체를 이동, 변형시키는 행위 등으로써, 가리키기(point), 누르기(push), 움직이기(move), 잡기(grab), 비틀기(twist), 돌리기(rotate) 등으로 분류할 수 있다[3].

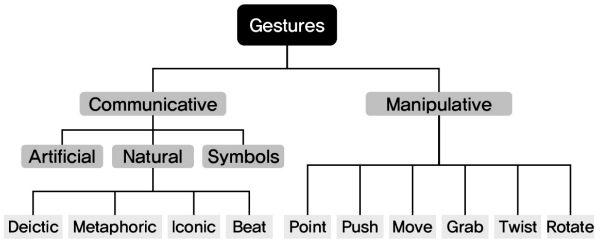


그림 2. 제스처의 분류

### 2.2. 제스처 인터페이스

이러한 제스처를 인식, 해석하여 사용자의 행동 의도와 목적 대상물 사이의 원활한 소통을 가능하게 해주는 것을 제스처 인터페이스라 할 수 있다. 특히 사용자의 자유로운 손 동작을 통한 인터랙션은 간결하면서도 강력한 표현을 보여준다. 가상현실 시스템에서는 사용자의 손 동작과 동기화된 물

체의 움직임을 통한 조작이 가능하고, 수화에서는 손 제스처를 통해 사람과 사람간의 원활한 의사소통을 할 수 있다. 이러한 손 동작을 통해 제스처 인터페이스에서는 가리키기, 잡기 등 물체를 조작하는 제스처들을 사용자의 의사전달 수단으로 이용하여 직접적인 인터랙션을 구현한다.

### 2.3. 관련작업

Bolt 의 'put-that-there' 시스템은 자유로운 공간에서의 손 동작 제스처와 사람의 언어를 결합하여 물체를 조작하는 방법에 대한 연구를 보여주었다[4]. 한 손을 사용해 물체를 가리키는 동작을 취하면 손목에 부착된 공간 감지 센서가 이를 파악하여 반응을 보여주는 형태이다. 이러한 형식은 Thomas Baudel 의 'Charade'에서 보다 구체적으로 볼 수 있다[5]. 사용자는 손의 위치와 손가락의 구부림을 파악할 수 있는 장갑을 착용하고 스크린 앞에서 특정한 상황에 알맞게 주어진 동작을 취하여 컴퓨터와 인터랙션 할 수 있다.

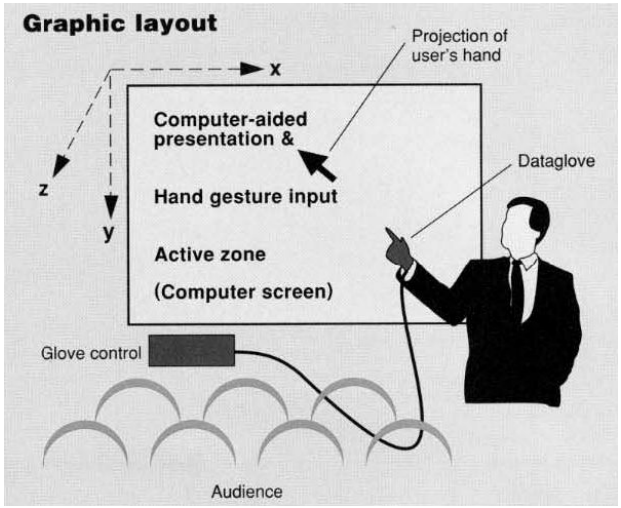


그림 3. Charade, Thomas Baudel, 1993

그러나 이러한 시스템에서는 인터랙션을 위해서는 항상 부가적인 장비를 착용하거나 몸에 마커를 부착하고, 특정한 영역 안에서만 동작을 취해야 하는 불편함이 따른다. 또한 명령어를 전달하기 위해서는 10 여 가지 이상의 주어진 동작들을 미리 암기하여 수행해야 하는 또 다른 노력이 필요하다.

Latoschik 의 연구에서는 3 차원 가상현실 속에서 손 동작과 언어를 통해 물체를 조작하는 시스템을 볼 수 있다[6]. 공간에서의 손 움직임을 인식하고 그 속도와 머문 시간을 파악하여 주어진 알고리즘에 따라 물체를 선택하고, 움직이고, 돌리는 동작 등을 할 수 있다. 그러나 사용자가 선택할 수 있는 동작들은 한계가 있으며, 다양한 표현의 제스처들을 모두 구분할 수 없다는 한계가 있다. 즉 기계가 인식할 수 없는 동작들을 취했을 경우에 대한 대안이나 모호한 동작에 대한 구분을 위한 장치가 마련되어 있지 않다.

한편 MIT 의 Responsive Environments Group 에서는 악기와 연주자 사이의 물리적 접촉 없이 제스처를 감지하여 음악을 연주할 수 있는 'Termenova' 라는 시스템을 보여주었다[7]. 사용자가 붉은 색의 레이저 빛 위로 손을 움직여 흐름을 가로 막으면 기계는 그 거리를 측정하여 음조와 볼륨을 조절하고 다른 효과들을 첨가하여 소리를 만들어 낸다.

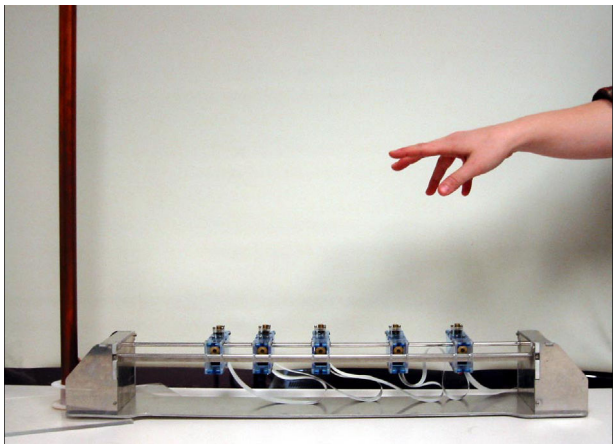

그림 4. The Termenova, MIT Media Lab, 2002

지속적인 제스처 인식과 응답을 위해서 많은 채널 센서들이 동시에 작동하고 있으며 그에 알맞은 소리를 내기 위해서는 측정 값과 반응 값의 효과적인 매핑이 필요하다. 이때 제스처에 대한 대응 값이 잘 못 적용되거나 그 자유도가 높아지게 되면 이 시스템은 악기로써 음악을 연주하는 기능을 상실하고 효과적인 소리를 재생할 수 없게 된다. 따라서 적당한 값으로 자유도를 한정시켜 주는 것이 바람직하며, 레이저의 방향을 다양하게 하거나 페달을 첨가하여 기능을 분화시키는 방안도 생각

할 수 있다.

## 3. 효과적인 제스처 인식과 정보탐색

인간과 컴퓨터 혹은 미디어와의 상호작용은 사용자 인터페이스라 할 수 있다. 책이라는 매체를 통해서 우리는 정보를 전달받고 이해하였으며 그러한 경험을 바탕으로 디지털 미디어 시대에서 시각적 커뮤니케이션을 발전시켜 왔다. 여기서는 책을 읽을 때의 동작을 인식하여 일상생활 속에서 책을 읽는 것처럼 디지털 콘텐츠에 접근할 수 있는 인터페이스에 대해 알아보고 이를 제작해본다. 'Book_ing' 라 칭한 이 프로젝트는 스크린 앞에 책 모양의 작은 상자를 두고 이를 중심으로 손을 움직여 스크린 속의 콘텐츠를 이동, 탐색할 수 있다.

### 3.1. 효율적 인식을 위한 인터페이스

기존의 제스처 인터페이스 작업에서는 손에 마커를 붙이거나 센서가 부착된 장갑을 착용해야 하는 불편함이 있었다. 이러한 문제점의 해결을 위해 카메라로 동작을 촬영하여 분석하는 시스템이나, 거리 센서로 위치를 파악하여 작동하는 방식이 시도되었다.

본 프로젝트에서는 적외선 거리 감지 센서를 사용하여 허공에서 움직이는 사용자의 두 손을 파악하고자 했다. 이러한 시스템을 구성한 것은 첫 째, 동작의 의도를 정확히 파악하기 위해서 이다. 책을 읽을 때에는 책을 펼친 상태에서 양쪽의 모서리부분을 잡고 앞, 뒤로 페이지를 넘기게 된다. 책 내용을 훑어볼 때에는 많은 페이지를 한 번에 넘기면서 표지부터 마지막 장까지 이동하고, 그 후 정독할 때에는 한 장씩 차례대로 넘기는 것이 일반적인 독서의 형태이다. 'Book_ing' 에서는 손의 가로 방향 위치에 따라서 책장이 넘어가도록 인터랙션에 제한을 두고, 그 위치에 따라 넘어가는 페이지의 양을 조절하도록 한다. 즉 제스처 행동에 제약을 두고 자유도를 한정시켜서 사용자가 하려는 행동의 의도를 정확히 파악한다.

둘 째, 제스처를 분할해서 인식하기 위함이다. 제스처는 그 특성상 연속적으로 끊임없이 이어져서 발생하므로 상황에 맞게 동작을 구분 지어서 시작과 끝을 파악하고 이어지는 연속동작을 따로 인식하도록 구성해야 한다. 따라서 'Book_ing' 에서는 사용자가 손을 움직여 책장을 넘기다가 센서가 감지할 수 있는 범위를 벗어날 경우, 그 시점에서 페이지가 넘어간 정도에 따라 이전 혹은 다음 페이지로 구분하여 자동적으로 넘어간다. 센서가 감지하는 영역 안에서의 손 동작을 사용자의 의도에 맞게 구분하고 그에 대응하는 스크린의 변화를 보여준다.

### 3.2. 화면구성과 정보탐색

우리가 평소에 접하는 제본되어 있는 책은 페이지의 순서대로 읽게 되어있다. 일반적으로 넘겨서 읽는 페이지들은 각 페이지가 하나의 단위로 인식되고 계산되어 책을 구성하게 되며 우리는 페이지를 단위로 해서 읽는다. 세부적인 정보를 찾고자 할 때 이러한 페이지 인터랙션 디자인은 기능적으로 가장 빠른 시간에 정보를 찾을 수 있게 해준다. 이는 페이지 인터랙션 형태가 사용자들에게 가장 익숙하여 심리적 안정감을 줄 수 있고, 검색 시에는 페이지 단위 별로 이동하여 정보를 찾을 수 있기 때문이다[8].

따라서 디지털 콘텐츠에서도 페이지 인터랙션 형태의 구성을 통해 효과적으로 정보를 전달할 수 있고, 스크롤이나 팝업 형태의 인터랙션보다 직관적인 구성이 가능하다.

## 4. 프로젝트의 구현

'Book_ing' 프로젝트는 사용자의 동작을 인식, 파악하여 디지털 콘텐츠에 접근하는 제스처 인터페이스로 실제의 책을 읽는 것처럼 유연한 상호작용이 가능하도록 구현하였다.

### 4.1. 동작감지 및 데이터 전송

사용자의 동작을 감지하기 위해서 적외선 반사각 측정방식의 거리감지 센서인 sharp 전자의 GP2D12 가 사용되었다. 책 형태의 직육면체 상자 속에 적외선 거리감지 센서 2 개를 좌우 양쪽 방향으로 설치하고 이를 통해 사용자가 양 손을 사용해 허공

| 범주 | 구분 | 동작표현 | 동작(Book_ing) |
|---|---|---|---|
| 시작 | 책을 펼친다 | 오른손으로 표지를 잡고 왼쪽에서 오른쪽으로 넘긴다. | - |
| 정보 획득 | 책을 읽는다 | 책을 읽는 동안 손은 가만히 두거나 책을 살짝 잡고 있는다. | - |
| 이동 | 다음 페이지로 책장을 넘긴다 | 오른손의 엄지와 검지로 책장을 잡고 오른쪽에서 왼쪽방향으로 옮긴 후 손을 편다. | 오른손을 상자의 오른쪽에서 왼쪽으로 움직인다 |
| 이동 | 이전 페이지로 책장을 넘긴다 | 왼손의 엄지와 검지로 책장을 잡고 왼쪽에서 오른쪽방향으로 옮긴 후 손을 편다. | 왼손을 상자의 왼쪽에서 오른쪽으로 움직인다. |
| 검색 | 원하는 내용을 찾는다 | 왼손으로 책의 가운데 부분을 잡고 오른손으로 책을 약간 휘면서 감싸듯이 잡고 엄지손가락을 이용해 책을 빠르게 넘긴다. | 오른손/왼손을 상자에 가까이 가져가서 일정 시간 이상 동안 멈춘다 |
| 종료 | 책을 덮는다 | 오른손/왼손을 사용해 책의 겉 표지에 손바닥을 대고 반대편으로 책 전체를 덮는다. | - |

표 1. Book_ing 의 사용자 제스처 분석

에서 책을 넘기는 동작을 취할 때 그 거리를 감지한다. 일상적으로 책을 볼 때 페이지의 한 쪽 모서리를 잡고 말아 넘기는 것처럼 동작을 취하면 센서는 그 거리 값을 측정하여 전압의 변화로써 그 양을 나타낸다.

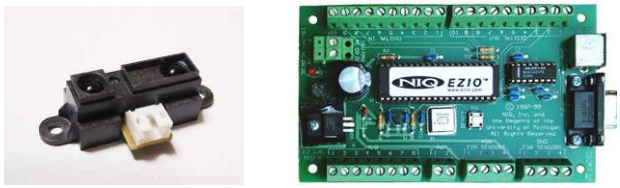

그림 5. 적외선 센서와 EZIO 보드

이렇게 측정된 아날로그 값은 마이크로 컨트롤러를 통해 디지털 값으로 변화시켜 주어야 한다. 이를 위해 Michael Rodemer 가 개발한 EZIO 보드가 쓰였다. EZIO 보드는 센서를 연결하고 전원을 공급시켜 주는 것만으로 간단하게 구현시킬 수 있도록 만들어진 장비로 이를 통해 디자이너들은 전자 통신이나 프로그래밍에 관한 자세한 지식 없이도 쉽게 인터랙션 인터페이스를 구체화 시킬 수 있다.

아날로그로 측정된 값은 보드를 통해 디지털로 변화되고, 이 데이터는 RS232 시리얼 케이블을 통해 컴퓨터로 전송된다. 시리얼 통신이나 미디 통신은 하드웨어 표준방식으로 컴퓨터와 마이크로 컨트롤러의 상호 연결을 위해 필요하다.

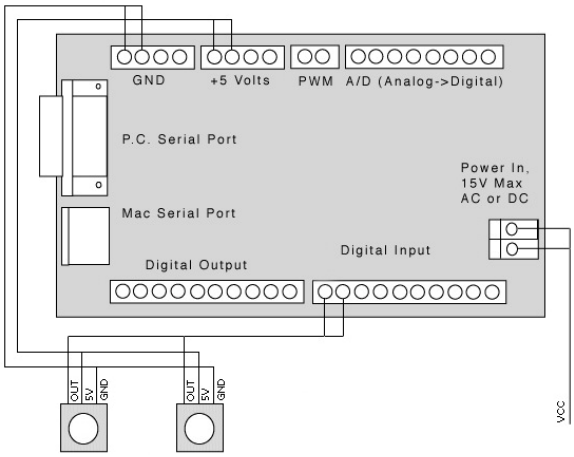


그림 6. EZIO 보드와 센서의 연결

4.2. 소프트웨어의 연결

시리얼 케이블을 통해 들어오는 입력 값을 받기 위해 cycling' 74 의 Max/MSP 가 사용되었다. Max/MSP 는 멀티미디어 제작을 위한 그래픽 환경의 소프트웨어로 복잡한 텍스트가 아니라 오브젝트 형태로 제공되는 명령어들을 연결하여 쉽게 프로그래밍이 가능하다. 또한 여러 플러그인들이 개발되어 있어 사운드, 비디오, 네트워킹 등을 구현할 수 있다.

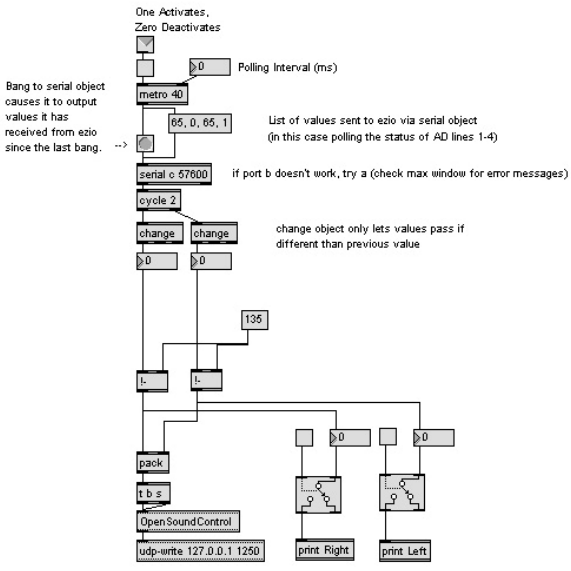


그림 7. Max/MSP 의 패치파일 구성도

시리얼 포트를 지정해 센서로부터 값을 실시간

으로 전달받아 원하는 값을 취한 후에 이를 매크로미디어사의 플래시로 전송해야 한다. 그러나 Max/MSP 에서는 곧바로 플래시로 데이터를 전송할 수 있는 오브젝트가 제공되지 않기 때문에 Ben Chun 이 개발한 FLOSC(flash open sound control)를 사용한다. OSC(Open Sound Control)는 일종의 프로토콜로써 컴퓨터와 멀티미디어 장비들간의 커뮤니케이션을 위해 사용되며 FLOSC 에서는 데이터를 Java XML socket 형태로 변환시킨 후 전송한다.

### 4.3. Book_ing 프로젝트의 표현

인터랙션 표현을 위해 동적이면서도 실시간으로 표현이 가능한 Macromedia 사의 Flash 를 사용하였다. 전달받은 데이터 값을 통해 사용자의 손의 위치와 연속동작들을 구분하여 그에 맞는 페이지 동작을 표현하며 손 동작에 따라 실시간으로 화면 속의 책 페이지는 앞뒤로 넘겨지게 된다.

그림 8. Book_ing 의 구현 모습

그림 9. 스크린 변화 과정 시퀀스

실제로 책을 읽을 때 페이지를 옮겨가며 원하는 부분을 읽는 것처럼 사용자는 손을 움직여서 스크린 속의 디지털 콘텐츠를 제어할 수 있게 된다.

## 5. 결론

본 연구에서는 효과적인 제스처 인터페이스의 구현을 위해 필요한 점들을 살펴보고, Book_ing 프로젝트를 통해 디지털 콘텐츠와의 인터랙션을 보여주고자 했다. Book_ing 는 평소 작업의 형태를 변경할 필요 없이 그대로 유지하여 사용자에게 편리함을 주고는 있으나, 실제 책 읽는 과정의 부분적인 동작만을 인식하는 한계점을 가지고 있다. 추후 개선을 통해 원하는 페이지로 빠르게 이동하기 위한 책갈피나, 사용자가 메모를 남길 수 있는 기능 등이 추가되면 보다 유연한 인터랙션을 보여줄 수 있을 것이다.

다가오는 유비쿼터스 환경에서는 기술적 발전에 의한 변화뿐 아니라 인간과 제품 사이의 네트워크가 강화된 감성적 인터페이스로 발전 할 것이다. 즉 인간과 컴퓨터 사이의 물리적이 인터페이스가 현재의 데스크탑 패러다임보다는 인간이 실제 상호작용 하는 방식에 더 가까워 진다는 것을 의미한다. 말을 하거나, 몸짓을 하고, 필기도구를 이용해서 커뮤니케이션하는 등의 자연적인 행동들을 명시적인 또는 암묵적인 입력으로서 활용할 수 있는 가능성을 가지고 있다. 이러한 기술을 통해 가상현실과 게임, 미디어 아트 등 다양한 분야에서의 활용은 물론, 일상의 환경 속에서 자연스러운 인터랙션이 이루어지도록 더 많은 연구가 필요하다.

## 6. 참고자료

[1] Roger Dannenberg, Dale Amon, “A Gesture Based User Interface prototyping System”, in Proceedings of the 2nd annual ACM SIGGRAPH symposium on User interface software and technology, 1989

[2] A. Mehrabian, “Communication without words”, communication: Concepts and Process, pp.106-114, 1971

[3] Jacob Eisenstein, Randall Davis, “Visual and Linguistic Information in Gesture Classification”, in

Proceedings of ICMI, 2004
[4] Richard A. Bolt, Edward Herranz, “Two handed Gesture in Multi-Modal Natural Dialog”, in Proceedings of UIST, 1992
[5] Thomas Baudel, Michael Beaudouin-Lafon, “Charade: Remote Control of Objects Using Free-hand Gestures”, Communications of the ACM, 1993.
[6] Ed Kaiser, Alex Olwal, David McGee, Hrvoje Benko, Andrea Corradini, Xiaoguang Li, “Mutual disambiguation of 3D Multimodal Interaction in Augmented and Virtual Reality”, in Proceedings of ICMI, 2003
[7] Lelia Hasan, Nicholas Yu, Joseph A. Paradiso, “The Termenova: A Hybrid Free-Gesture Interface”, in Proceeding of the 2002 Conference on New Instruments for Musical Expression (NIME-02), 2002
[8] 이승윤, “디지털북 내용구조의 분석적 표현을 위한 시각인터랙션디자인 연구 – 그리스신화제작 사례를 중심으로”, 서울대학교 디자인학부 석사학위 청구 논문, 2003.