

동작인식기반 Digital TV인터페이스를 위한  
지시동작에 관한 연구

A Study on Vision Based Gesture Recognition Interface Design for Digital TV

주저자 : 김현석

홍익대학교 미술대학 디자인학부 교수

**Kim, Hyun suk**

Hongik University, School of Design

공동저자 : 황성원

홍익대학교 대학원 영상학과

**Hwang, Sung won**

Hongik University, Graduate School

공동저자 : 문현정

홍익대학교 대학원 시각디자인과

**Moon, Hyun jung**

Hongik University, Graduate School

본 연구는 전략산업 혁신 클러스터 육성 지원사업의 연구진흥비로 연구되었습니다.  
Seoul Research and Business Development Program(10555)

## 1. 연구 배경 및 목적

## 2. 지시동작 인식기반 인터페이스

- 2-1. 지시동작 인식프로세스
- 2-2. 화상인식을 위한 지시동작의 분류
- 2-3. 지시동작의 구성

## 3. 디지털 TV 기능 및 메뉴 분석

- 3-1. 디지털 TV본체의 버튼 입력방식
- 3-2. 리모컨을 통한 입력방식
- 3-3. 디지털 TV의 기능 분석

## 4. 디지털 TV 인터페이스를 위한 지시동작 개발

- 4-1. 실험개요
- 4-2. 연구결과분석

## 5. 결론 및 제언

## 참고문헌

## (要約)

인간 대 컴퓨터 인터페이스의 발전은 기술의 발전에 크게 의존되어 왔다. 그러나 인위적 환경을 통한 커뮤니케이션 방법인 키보드, 마우스나 OMR 등의 비직관적 인터페이스를 넘어선 직관적 인터페이스의 연구가 활발히 진행되고 있다. 이와 더불어 일반 가정에서도 대형화면 TV의 급속한 보급과 디지털환경을 통한 다양한 서비스의 제공이 이뤄지고 있다. 이러한 인터페이스와 TV환경의 변화는 인간 대 TV의 인터페이스에 획기적 변화를 요구하게 되었다. 기존의 인터페이스에서 동작인식을 기반으로 하는 인터페이스로의 전환은 보다 직관적인 인터페이스로서 제시될 수 있을 것이다. 본 연구에서는 지시동작인식이론과 디지털 TV 인터페이스 분석에 대한 이론을 바탕으로, 이에 요구되는 최소기능 추출 후 인터뷰를 통해 추출된 지시동작을 분석하는 기초적 연구이다. 이를 통해 디지털 TV에서 요구되어지는 특정 기능에 대한 동작모형을 제시하여 이에 적합한 인터페이스 환경을 디자인하는데 대한 기초자료로 활용하고자 한다.

## (주제어)

디지털 TV, 인터페이스, 동작인식, 인터렉션

## (Abstract)

The development of Human Computer Interface has been relied on the development of technology. Mice and keyboards are the most popular HCI devices for personal computing. However, device-based interfaces are quite different from human to human interaction and very artificial. To develop more intuitive interfaces which mimic human to human interface has been a major research topic among HCI researchers and engineers. Also, technology in the TV industry has rapidly developed and the market penetration rate for big size screen TVs has increased rapidly. The HDTV and digital TV broadcasting are being tested. These TV environment changes require changes of Human to TV interface. A gesture recognition-based interface with a computer vision system can replace the remote control-based interface because of its immediacy and intuitiveness. This research focuses on how people use their hands or arms for command gestures. A set of gestures are sampled to control TV set up by focus group interviews and surveys. The result of this paper can be used as a reference to design a computer vision based TV interface.

## (Keyword)

Digital TV, interface, Gesture Recognition, interaction

## 1. 연구 배경 및 목적

컴퓨터와 가전기기의 통합 및 네트워크의 발전은 우리의 생활을 매우 크게 변화시키고 있다. 이러한 변화의 가장 밀접한 부분 중의 하나가 일상생활에서 자주 사용되는 사물/기기의 인터페이스 변화이다. 많은 종류의 기기들이 디지털화 되었고 또 많은 종류의 디지털 기기들이 네트워크화 되었다. 냉장고, 세탁기 등과 같은 백색가전제품은 컴퓨터 기능 및 네트워크 기능이 내장되었고, TV 또한 디지털TV나 IPTV 등으로 진화되어가고 있다. 그러나 급속한 컴퓨터의 발전에 비해 컴퓨터와 인간을 이어주는 인터페이스 환경의 변화는 그다지 급속하게 변하고 있지 않다. 대표적인 HCI 모드는 키보드와 마우스로 이루어진 입력시스템이다. 이러한 입력기기는 우리에게 매우 익숙한 방식이지만 인간과 컴퓨터간의 소통을 위해 인공적으로 만들어졌다는 점 때문에 그 사용에 있어 자연스러움(naturalness)이나 속도(Speed)같은 부분의 태생적 한계를 가지고 있다. 이러한 한계를 극복하고자 새로운 방식의 HCI 모드 개발에 많은 연구가 집중되어지고 있으며, 그 중에서도 인간 대 인간 인터페이스와 가장 유사한 형식의 인터페이스인 음성인식과 동작인식을 통한 입력모드개발에 대한 연구가 급속한 진전이 이뤄지고 있다.

LCD 및 PDP TV의 급속한 기술발전과 함께 화면 크기의 확장은 이러한 동작인식을 기반으로 하는 인터페이스 적용에 좋은 기회를 제공하고 있다. 컴퓨터와 비교하여 비교적 거리를 두고 시청하는 TV는 구분 가능한 동작을 취하기 위한 충분한 공간을 제공하고 있으며 큰 화면은 이러한 입력신호의 Feedback을 주기에 적당하다. 또한 컴퓨터 및 다른 생활가전 기기에 비해 비교적 단순한 명령구조 및 체계는 초기 동작인식 인터페이스 개발에 대한 기술적 한계를 최소화 시켜주고 있다. 하지만 이와 관련된 연구는 인터페이스 디자인적 관점에서의 연구에 비해 동작인식 기술에 관한 연구에 집중되고 있는 것이 현실이다.

크게 동작인식을 기반으로 하는 인터페이스 개발은 가상마우스로서의 포인팅 디바이스적 측면으로서 접근과 지시동작 인식적 측면으로서의 접근으로 나뉘어 연구 되고 있다. VR에서의 동작인식 인터페이스는 포인팅디바이스로서의 동작인식에 관한 연구가 주류를 이루고 있으며 스크린을 기반으로 하는 연구에서는 지시동작(제스처)을 중심으로 한 연구가 많이 이루어지고 있다.

본 연구에서는 스크린 기반 동작인식에서 기존 디지털 TV의 인터페이스로 사용되고 있는 리모컨 대응으

로의 인터페이스 측면에서 지시동작인식기반 인터페이스로 연구의 범위를 한정하였으며 이를 위한 사용자 직관적인 지시동작 추출에 연구의 목적이 있다. 디지털 TV의 메뉴구조와 관련된 선행연구를 바탕으로 최소한의 지시명령을 추출하고 이를 인식하기 위한 지시동작을 개발하기 위하여 사용자 설문조사와 Think Aloud 기법을 사용하였다.

## 2. 지시동작 인식기반 인터페이스

인간 대 인간의 커뮤니케이션 방식에 있어서 음성 커뮤니케이션 다음으로 중요한 부분을 차지하고 있는 것이 제스처를 이용한 커뮤니케이션 방식이다. 인간은 일상생활에서 매우 많은 부분 제스처의 의존한 커뮤니케이션을 통해 의사를 전달한다. 상하로 머리를 움직이는 동작을 통해 '예'라는 대답을, 좌우로 움직이는 동작을 통해 '아니오'라는 대답을 전달한다. 이외에도 수 없이 많은 일상적 대화 형태로서의 제스처가 인간 대 인간 커뮤니케이션에서 사용되고 있다.

인간 대 컴퓨터 커뮤니케이션 방식으로서의 인터페이스, HCI에서도 이러한 지시동작(제스처)을 기반으로 한 연구가 보다 폭 넓은 인간 커뮤니케이션을 모방하기 위한 측면으로서 연구되어왔다. 가상현실(Virtual Reality) 및 증강현실(Augmented Reality) 기술의 발전은 키보드와 마우스를 대체할 수 있는 새로운 방식의 커뮤니케이션 기술을 필요로 하였고 이에 따른 동작인식 인터페이스가 새로운 멀티모달 인터페이스로서 주목을 받고 있다.

지시동작의 데이터를 컴퓨터로 입력시키기 위하여 다양한 방식의 입력장치가 연구되어 왔는데 그 중 가장 활발히 연구되어지고 있는 입력장치로는 데이터글로브(Data Glove)와 시각 입력장치가 있다. 특히 카메라를 통한 컴퓨터의 시각인식방식은 가장 인간 대 인간 인터페이스와 유사한 방식으로 멀티모달 HCI 방식으로 지속적 연구가 이루어지고 있다.

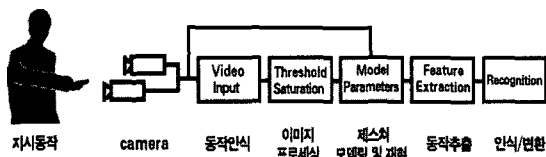
### 2.1 지시동작 인식 프로세스

군스(Hatice Gunes, 2004)는 지시동작(제스처) 인식은 프로세스는 '동작감지', '이미지프로세싱', '제스처 모델링 및 재현', '동작추출', '인식'의 5단계로 구성될 수 있다고 하였다.

- (1) 동작감지 (Sensing Human Position): 카메라를 통해 연속된 이미지의 입력을 통해 인간의 동작을 입력 가능한 데이터로 변환
- (2) 이미지 프로세싱 (Image Processing): 입력된 이미지

- 를 분석가능 하도록 변환, 증폭, 처리하는 과정
- (3) 제스처 모델링 및 재현 (Gesture Modeling and Representation): 입력된 이미지 데이터의 변환을 통해 제스처의 특성을 파악하고 이를 통해 인식 가능한 지시동작으로 재현
- (4) 제스처 분석을 통한 동작 추출(Feature Extraction and Gesture Analysis): 인식되어진 제스처의 분석을 통해 해당 변수(parameters)를 전달
- (5) 인식된 제스처를 지정된 명령으로 변환(Gesture Recognition and Classification): 데이터베이스에 입력되어있는 명령과 인식된 제스처와의 매치를 통해 실질적 명령실행

[그림 1] 화상기반 지시동작 인식 시스템



## 2.2 화상인식을 위한 지시동작의 분류

### (1) 동작형태에 따른 구분

#### a. 머리와 얼굴을 사용하는 지시동작

인간 대 인간의 커뮤니케이션 방식에서 가장 활발히 사용되는 지시동작으로 음성언어와 함께 정보를 전달하기 위하여 사용되어진다. 고개 젓기, 끄덕이기, 눈썹 찌푸리기, 윙크 등과 같은 동작이 이에 해당된다.

#### b. 손과 팔을 사용하는 지시동작

손과 팔을 몸통으로부터 떨어뜨려 동작하는 행위로 구성되어있다. 특히 손가락 및 팔을 활용한 지시동작은 직접적으로 음성언어와 관계되어 있는 경우가 많으며 방향의 지시 및 다양한 의사전달 도구로서 사용되어지고 있다. 또한 손과 팔을 사용하는 지시동작은 다른 지시동작에 비해 국가 간, 문화 간 차이를 내포하고 있는 경우들이 있다. 이와 더불어 수화의 사용과 같이 직접적 음성언어를 대체하는 경우도 있어 매우 표현적이고 자유롭게 커뮤니케이션을 수행한다.

#### c. 몸을 사용하는 지시동작

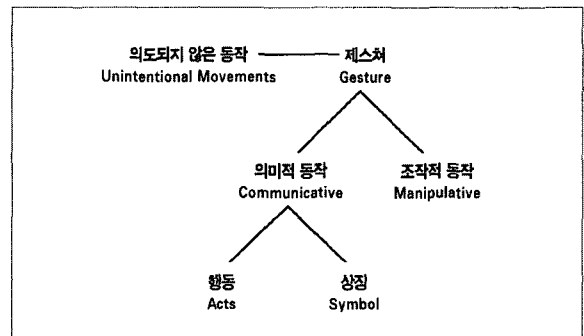
몸동작은 전체 또는 부분적인 움직임을 포함하고 있다. 예를 들어 허리를 돌리는 동작이나 어깨를 꾸부리는 동작 등과 같은 몸체의 부분적 움직임과 몸체 자체의 움직임(좌/우, 앞/뒤 등)으로 구성된다. 이러한 지시동작 중 손과 팔을 사용하는 지시동작은 그 표현의 다양성과 편리성으로 인하여 가장 활발하게 활용되어지고 있다.

### (2) 의미체계에 따른 구분

인간의 제스처를 분류하는 방식에 대한 연구는 다양하게 진행되어져 왔다. 켄돈 (A. Kendon, 1986)은 음성언어와 독립적으로 발생하는 "자동기술동작"(Autonomous gestures)을 "지시동작/몸짓"(gesticulation)과 구분하였고, 맥네일과 레비(McNeill & Levy)는 지시동작을 "아이콘 형식"(Iconic Gesture), "은유적 형식"(Metaphoric Gesture), "비트 형식"(Beat)의 3가지 그룹으로 구분하고 있다. 블라드미르와 라지브 (Vladimir & Rajeev, 1997)는 아래의 그림에서 보여주는 것처럼 "제스처"와 "의도되지 않은 동작"의 2개의 주요 클래스로 구분하고 있다.

'의도되지 않은 동작'은 손 또는 팔의 움직임 중 어떠한 의미정보를 가지고 있지 않은 것을 의미하며 '제스처'는 다시 '의미적 동작'(Communicative)와 '조작적'(Manipulative)동작으로 나누어진다. '조작적 동작'이 사물의 움직임이나 회전 등을 묘사하기 위한 동작이라면 '의미적 동작'은 고유의 의사전달을 위한 의미작용으로서의 동작을 말한다. 의미적 동작은 '행동'(Act)과 '상징'(Symbol)으로 나누어진다.

[그림 2] 손과 팔을 사용하는 지시동작의 분류



### (3) 기호학적 분류에 따른 구분

기호라는 것은 어떤 대상을 대신하여 메시지를 전달하는 것으로, 동작 인식기반 디지털TV의 인터페이스를 위한 지시동작 역시 동작 기호라고 볼 수 있다. 이를 유형화 하는데 있어 지시동작의 의미를 분석하고 적절한 활용과 동시에 발전방향을 모색하기 위하여 퍼스의 기호학을 바탕으로 분석하고자 한다.

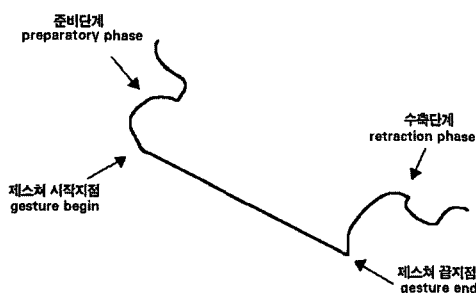
퍼스는 기호의 유형으로서 도상, 지표, 상징의 세 유형으로 분류하였다. 도상적 기호(iconic sign)란 기호와 지시대상 사이에 존재하는 어떤 유사성에 착안한 것이다. 사진이 실물과 유사한 도상, 분자모형은 물질의 구조적 특징과 유사한 시각적 도상, 그리고 한국에서는 '멍멍'이 개 짖는 소리를 나타내는 음성적 도상으로 통용되는 것처럼 일반적으로 기호와 그 지시

대상 사이의 유사성에 근거하여 의미작용을 가능하게 하는 기호를 도상(icon)이라고 한다. 상징적 기호(symbolic sign)란 도상이 유사성을 기반으로 하고, 지표가 인과성에 근거하여 의미작용을 일으키는 기호라면, 상징적 기호는 자의성이 그 특징이다. 상징적 기호와 그 지시대상 사이에는 어떤 유사성이나 인과적 관계가 있는 것이 아니다. 동작기호에서 가장 많이 사용되어지고 있는 수화는 대표적 상징기호이다. 두 손을 모아 하트를 만드는 손동작은 그 자체로서 사랑이라는 의미를 내포하고 있지 않지만 이러한 동작 기호자체가 자의성을 가지고 관행적으로 사용되면서 기호로서 역할을 수행한다. 지표적 기호(indexical sign)란 기호와 그 지시대상 사이에 인과관계나 서열 관계(sequential relation) 같은 구체적이고 실제적인 관계가 있는 경우를 지표적 기호라고 한다. 그러나 모든 기호는 도상, 지표 및 상징 세 가지 유형의 특성 중 하나만으로 의미 작용하는 것은 아니다. 이러한 유형의 특성은 어디까지나 개념적 구분이며 상호 배타적이라고 볼 수는 없다. 즉, 실제적으로는 유기적으로 조합되어 의미작용을 하는 것이다.

### 2.3 지시동작의 구성

맥 널(Mc Neill, 1985)은 이러한 다양한 제스처를 기반으로 하는 지시동작은 '사전 동작'(Pre-Stroke), '동작'(Stroke), '사후동작'(Post-Stroke)의 3단계로 구성되어져야 한다고 정의하고 있다. 어떤 지시동작은 정적(Static)인 요소와 동적(Dynamic)인 요소를 모두 가지고 있는 경우가 있는데 이러한 경우 2개 이상의 연속동작과의 구분이 반드시 이루어져야 한다. 예를 들어 수화에서와 같이 하나의 지시동작이 다른 지시동작과 연속적으로 행해졌을 경우 각각의 지시동작이 구분되어져서 인식되어야 하므로 지시동작과 지시동작을 구분할 수 있는 사전, 사후 동작을 필요로 한다.

[그림 3] 지시동작 생성단계



인간은 동작언어를 이해하는데 많은 훈련이 되어있고 이러한 지시동작(제스처)의 해독에 있어서 정보와 의

미 없는 동작인 노이즈를 구분해 내는 훈련을 음성언어의 인식과정과 더불어 수행함으로써 분명한 지시동작의 의미를 파악하는데 적응이 되어왔다. 하지만 컴퓨터 시스템이 이러한 지시동작의 의미를 파악하기 위하여 노이즈로부터 정보를 분리하기 위해서는 분명한 지시동작의 인식신호(Identifier)가 존재해야 한다.

### 3. 디지털 TV의 기능 및 메뉴 분석

디지털 TV의 입력방식은 TV본체 입력방식과 리모트 컨트롤을 통한 입력방식으로 나뉜다. 이러한 입력방식과 디지털 TV의 일반적인 기능을 분석, 비교하여 기본 조작 단위를 추출하고자 한다.

#### 3.1 디지털 TV본체의 버튼 입력방식

본체에서 직접 조작하는 경우의 입력기능은 일반적인 TV 시청태도를 고려하여 다소 제한적으로 설정되어 있다. 제조사 및 브랜드 별 차이를 바탕으로 본체 입력방식 설정기능을 다음과 같이 분류 할 수 있다. (김현석 외, 2006)

[표 1] 본체 입력방식 기능의 분류

구분	기능
필수입력 기능	전원, 음량조정, 채널조정, 외부입력선택, 메뉴보기
선택 입력기능	확인, 입력선택, 전원대기

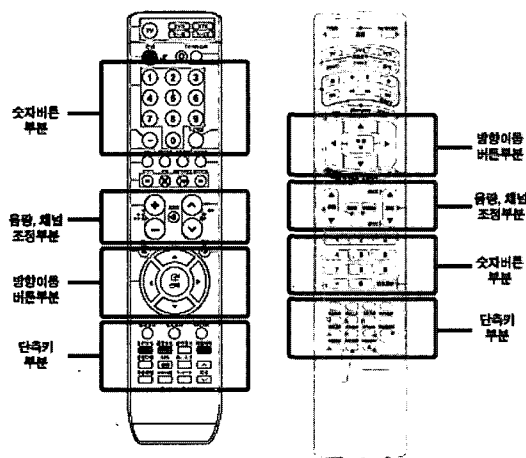
각 제조사 및 브랜드 별 본체 입력방식에서는 전원과 음량, 채널조정 그리고 외부입력선택 버튼은 동일하며 메뉴선택 이후의 '확인'이나 '선택'버튼 그리고 전원 대기 버튼의 유무에 차이가 있음을 알 수 있다.

#### 3.2 리모컨을 통한 입력방식

TV를 제어하는 일반적 입력방식인 리모컨의 경우에는 본체입력방식보다 훨씬 다양하고 복잡하게 구성되어 있으며 이러한 기능의 복잡성과 브랜드별 기능의 차이, 그리고 모델별 특징과 장점의 부가를 위한 고려 등으로 인해 제조사 및 브랜드별로도 차이가 있다. 그러나 리모컨의 외부 구성은 크게 네 부분으로 나눌 수 있는데 이는 숫자 버튼부분, 방향이동 버튼부분, 음량/채널 조정부분 그리고 단축키부분이다. 단축키 부분은 사용자들이 선호 기능들을 독립적으로 구성하여 브랜드별 모델의 특, 장점을 살릴 수 있는

이점과 함께 사용자들이 기능을 쉽고 빠르게 쓸 수 있도록 배려한 것이다. 그러나 이러한 단축키들이 오히려 시각적으로 사용자의 선택을 방해함으로써 중요한 버튼이 무엇인지 분별하기 어렵도록 발전되어 왔으며 디지털 TV의 기능이 다양해지고 발전함에 따라 더욱 복잡해지는 양상을 띠고 있다. 우선 기본적인 조작 방법으로 제어되는 기능을 중심으로 기본 조작 단위를 정리하고 단축키부분에서 제외 될 수 없는 필수적 버튼을 추가하는 방법으로 최소 기본 조작단위를 추출하였다.

[그림 4] 삼성전자<sup>1)</sup>와 LG전자<sup>2)</sup> 리모컨으로 한 외형적 구분



위의 견본모델로 구분해 본바와 같이 리모컨의 구성과 위치는 브랜드별로 차이가 있으나 외형적으로는 네 부분으로 구분된다. 따라서 리모컨으로 제어하는 기능을 중심으로 다음과 같이 크게 묶을 수 있다.

[표 2] 리모컨으로 제어되는 기능의 분류

구분	기능
전원	TV on/off
메뉴	메뉴 들어가기와 메뉴종료
숫자 버튼부분	채널 선택
음량, 채널조정부분	음량조정, 채널조정
방향이동 버튼부분	설정기능에서의 상하좌우 이동 과 확인, 취소
단축키 부분	화면 크기 조정, 화면 모드 설정, 음성 모드 설정, 음향 모드 설정, 음소거, 동시화면, 취침예약, 입력전환, 화면정지, 위치변경, 화면크기, 주부전환, 이전채널, PC/HDMI 모드 전환, DTV/TV 전환, 입력선택, 외부입력, 방송안내, 신호채널, 키잠금, -버튼, 절전모드, 설정-비디오

1) 삼성 LN32R71BD, LN32R71WD, LN32T71BD사용설명서  
2) LG 37LB1DR, 42LB1DR, 47LB1DR 사용설명서

DVD 등/초기화, Anynet (애니넷), SRS 버튼-입체음, 조명버튼, 타임머신, XSTUDIO, 녹화목록, XD 영상, 자막방송, MEMORY 기능

조사를 위해 선정한 표본제품은 국내 시장점유율<sup>3)</sup>을 기본으로 선정하였다. 국내 LCD TV시장은 삼성전자와 LG전자가 83%를 차지하고 있고, 중소브랜드인 이레전자가 15%, 외산브랜드인 소니, 아이와, 샤프, HP가 2%를 차지하고 있으므로 현 시장을 양분하고 있는 LG전자, 삼성전자 그리고 대우일렉트로닉스와 중소기업인 이레전자의 브랜드 중 현재 각 제조사에서 시장선점을 위해 노력하고 있는 제품모델을 선택하였다. 이러한 제품들의 리모컨을 비교해 각 부분별로 기능을 구분하였다. 그러나 제품별 특정보다는 일반적인 리모컨의 기능에 중점을 두고 보면 단축키부분을 제외한 부분은 크게 다르지 않음을 확인 할 수 있고, 앞서 말한 바와 같이 각 제조사나 브랜드 별 차이는 그 구성과 위치 그리고 단축키 부분에서 두드러짐을 알 수 있다. 사실상 단축키는 설정메뉴의 하위 구조임을 감안 할 때 공통된 기본 구조 안에서 기본 조작단위를 추출 가능하다. 이러한 단축키의 다양성은 위치 및 구조 이외의 리모컨 외형에서 시각적으로 상당히 복잡하게 발전되어 왔는데 이로 인해 간단한 기능의 제어를 위한 직관적 이해도가 낮다.

### 3.3 디지털 TV의 기능분석

디지털 TV의 일반적 기능은 현재 제품들에서 사용되고 있는 '입력, 채널, 화면(영상), 음향(음성), 시간, 설정, 방송안내, 일반, 기능, 시각, DVR'과 같은 Control Setup의 Main메뉴를 중심으로 살펴 볼 수 있다. 이를 제조사 및 브랜드 별 기능의 차이를 고려하여 '입력, 채널, 화면, 음향, 시간, 설정'의 여섯 가지로 나누어 살펴본다.

[표 3] 디지털 TV의 기능 분석

기본조작단위	단위별 기능 예시
입력	TV/외부입력목록, 외부입력편집, Anynet-TV방송, 셋톱박스, DVD/VCR 보기, 주화면 입력, 부화면 입력, 외부입력연결,
채널	DTV 채널선택, 자동채널, 수동채널, 신호채널, 채널편집, 채널이름편집, 채널미세조정, 디지털 신호세기, 수신감도증폭, 신호세기, 채널자동설정, 안테나 설정,

3) 2005년 4분기 한국DTV업체 출하동향 분석, 디스플레이뱅크

화면	화면모드, 화면크기, 위치선택, 화면잡음제거, Dnle, 필름모드, 청색비탕화면, 동시화면, 주부전환, 색온도, XD모드, 부가기능, 영상초기화조정, PC/HDTV, TV자막방송, 밝기, 명암, 색농도, 색상, 선명도
음향	음향모드, 이퀄라이저, SRS TruSurround XT, 음성설정, 선호언어, 음성다중, 자동음량, 내부앰프, 주부화면 음성선택, 초기화, 음성언어, TV스피커, BBE조정, 프로로직, 디지털 음성출력, 디지털 음향, 스테레오 음성출력/입체 음향, 음균형(좌,우)
시간	초기설정, 현재시각, 취침예약, 켜짐시각, 꺼짐시각, 켜짐음량, 자동꺼짐, 시간모드설정, 켜짐채널
설정	게임모드, 배경색상 투명도, 메뉴 투명도, 멜로디, 절전모드, PC설정, 초기화, 기능도움말, 청색비탕화면, 방송안내(현재 채널방송안내, 전체채널방송안내, 시청예약목록)

위와 같이 디지털 TV의 기능은 Main메뉴를 중심으로 여러 단계의 하위메뉴로 구성되어 있는데 이는 일반적인 Tree구조와 같음을 알 수 있다. 이러한 정보 구조를 가진 디지털 TV의 설정기능(Control Setup box)은 어떤 형식으로 시각화 되어 있는지를 조사해보고 이에 따라 기본조작단위를 추출해 보고자 한다.

[그림 5] 제조사별 Control Set-up의 인터페이스

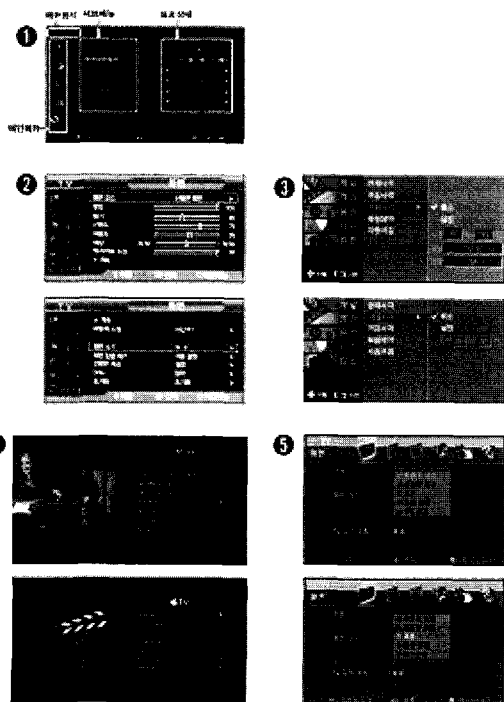


그림 5에서 제시된 샘플은 각각 (1)은 대우전자, (2)는 삼성전자 모델의 인터페이스이며 (3)은 LG전자, 다음은 Apple TV, 마지막(5)는 이레전자의 제품이다. 대부분의 제품들이 그래픽적으로는 달라 보이지만 정보 구성면에서는 트리구조를 채택하고 이를 가장 이해하기 쉽게 나타 낼 수 있는 List-up 형식으로 구성하고 있다. 기본적으로 메인메뉴 부분과 서브메뉴 부분, 그리고 각 설정상태를 나타내어주는 부분으로 구성되어 있으며 부분별 다른 색상 등으로 활성화 되어 있는 메뉴를 알려주고 있다. 이러한 세 단계의 구분 이외에 Apple TV는 일반적으로 사용되어지고 있는 3단구성(메인, 서브, 설정상태를 순차적으로 구성)을 수용하고는 있으나 화면의 왼쪽 부분에 메인메뉴에 해당하는 아이콘을 3D 공간에 재현함으로써 공간감과 몰입감을 확대시키고 있다. List-up 방식이 공통적으로 나타나는 이유는 아직까지 리모트 컨트롤의 상, 하, 좌, 우 방향이동 버튼으로 제어가 간편하기 때문으로 생각된다.

위와 같은 디지털 TV의 인터페이스와 일반적 기능의 정보구조 설계 그리고 각 기능별 시나리오를 분석하여 다음과 같은 기본조작단위를 추출할 수 있다. 이러한 기본조작단위는 디지털 TV의 설정기능 및 조절 기능을 제어함에 있어 가장 필수적인 최소단위이면서도 조작단계를 효율적으로 제어 가능한 방법을 기준으로 추출한 것이다. 어떠한 입력방식이든 기능제어를 위한 필수동작을 바탕으로 이를 효율적으로 수행할 수 있는 각 조작단위에 해당하는 동작모형을 추출함으로써 동작인식 제어 인터페이스의 모델을 제시할 수 있을 것이다.

[표 4] 기본조작단위의 추출

기본조작단위	단위별 기능 예시
ON/OFF	TV on/off
상	채널 이동 / 메뉴 이동 등
하	채널 이동 / 메뉴 이동 등
좌	음량 낮춤 / 메뉴 이동 등
우	음량 높임 / 메뉴 이동 등
확인	선택 / 확인 / 클릭 / 누름 / 메뉴 들어가기 등
취소	취소 / 전 단계로 / 메뉴 나가기 등

#### 4. 디지털 TV 인터페이스를 위한 지시동작 개발

##### 4.1 실험개요

본 연구는 디지털 TV의 지시동작인식기반 인터페이스에서 사용자의 직관적인 지시동작 추출 및 개발에 목적이 있다. 따라서 기존의 TV의 조작방법에서 벗어나 직관적이고 자유롭게 발상할 수 있도록 하였다. 또한 이 조사는 정량적 평가가 아닌 정성적 평가로 실험 대상자의 응답을 통계적으로 분석하는 대신 총체적으로 분석하고자 한다. 이로써 실험 대상자들의 Mental Model에 접근하여 디지털 TV 사용자들의 TV 사용과 제어에 대한 직관적 이해도와 방법 그리고 제어과정을 어떻게 인지하는지에 대한 검증하고자 함이다. 이를 위해 Think Aloud 기법을 이용한 인터뷰 방식과 설문 조사를 병행하여 실시하였다.

본 논문 2.1에서 제시한 바와 같이 지시동작 인식 프로세스는 '전원을 켜야겠다'는 사용자의 task를 실행시킬 수 있는 동작을 행하면 controller는 이를 동작 모형으로 변환하여 인식하고 task를 실행하는 커뮤니케이션 단계를 밟게 된다. 따라서 전문가 집단에서 동작모형으로 변환 가능한 지시동작을 다양하게 추출하고 일반인들을 대상으로 직관적 이해도가 높은 동작을 선정하는 방식으로 평가를 진행하도록 한다.

##### (1) 연구대상자 선정

1차 실험 집단의 전문가는 총 20명으로 인터렉션 디자인 전공 박사 4명과 석사 16명으로 구성되었다.

2차 실험 집단은 일반인으로 구성되었으며 총 20명을 대상으로 평가하였다.

##### (2) 실험 방법

1차 실험에서는 디지털 TV의 메뉴구조와 관련된 선행연구를 바탕으로 추출된 기본적 조작단위-ON/OFF, 상, 하, 좌, 우, 확인, 취소-에 대한 지시동작을 직관적이고 자유롭게 발상할 수 있도록 Think Aloud 기법을 이용하여 지시동작들을 추출하였다.

2차 실험에서는 1차 실험에서 추출된 지시동작들을 바탕으로 각 조작단위별 가장 적합한 지시동작을 선택하도록 설문조사를 실시하였다.

#### 4.2 연구결과분석

##### (1) 사용자의 Task선정

디지털 TV의 사용자 Task와 현재의 제어방식을 참조하여 추출된 기본조작단위의 Task시나리오를 작성할 수 있다. 이러한 사용자 Task시나리오는 다음과 같다.

[표 5] 사용자의 task시나리오

task	기존 입력방식	
TV를 켜다	전원 버튼	
control setup 비활성 상태	볼륨을 조정한다	리모컨의 좌, 우 방향의 이동 버튼
	채널을 조정한다	리모컨의 상, 하 방향의 이동 버튼
	control setup을 본다	메뉴 들어가기 버튼
control setup 활성화	메뉴list에서 이동한다	방향이동 버튼
	메뉴list에서 현 상태를 선택한다	확인/ok 버튼
	메뉴list에서 선택을 취소한다	취소 버튼
	메뉴list에서 전단계로 이동한다	취소 버튼

##### (2) 1, 2차 평가에서 추출된 지시동작의 분석

전문가들에게 실시한 1차 실험에서 기술적인 문제로 인하여 동작의 방향성을 인식하기 어려운 손가락만을 사용하는 동작, Z축을 사용하는 동작, 몸동작 등 오류 확률이 높은 동작은 2차 실험에서 제외시킨 후 일반인을 대상으로 설문을 하였다.

##### a. 'ON/OFF'의 경우

[그림 6] 실험에서 추출된 ON/OFF 지시동작





1, 2차 실험으로 통해 최종적으로 두 손을 박수치듯 마주치는 5번 동작이 'ON/OFF'의 지시동작으로 가장 적합한 것으로 나타났다. 이 동작은 도상적이거나 지표적인 지시동작 기호가 아닌 자의적으로 만들어진 지시동작이라는 것이 그 특징이다. 물론 1차 전문가 평가에서 한손으로 원을 그리는 1번, V자를 그리는 듯한 6번, OK를 표현하는 듯한 7번과 같은 행위의 시작을 알리는 긍정적인 표현이나 확인의 의미로 쓰이고 있는 기호를 표상하는 지표적인 기호도 포함되었으나 2차 실험에서는 피험자 대부분이 가장 간단하고 손쉽게 할 수 있는 지시동작인 5번을 선택하였다. 이는 자의적으로 만들어진 상징적 지시동작에 속한다.

### b. 상, 하, 좌, 우 지시동작

[그림 7] 실험에서 추출된 상/하/좌/우 지시동작



각각의 1, 2번 동작 순으로 가장 적합하다고 답하고 있으나 1, 2번의 동작의 경우 손가락을 모두 편 상태인지 아니면 무엇을 가리키듯 두 번째 손가락만 편 채로 지시하느냐에 대한 차이일 뿐 '상, 하, 좌, 우'라는 방향을 향하여 팔 동작을 움직이는 행위로 그 방향성은 동일하였다. 1차 전문가 실험에서도 ON/OFF, 확인, 취소의 경우 다양한 동작이 추출 된 것과 달리 '상, 하, 좌, 우'의 경우 대부분 사회적으로 약속이 되어있는 '상, 하, 좌, 우'라는 그 방향성에 따라 팔을 뻗는 동작을 가장 선호하는 것으로 나타났다. 단, 1차 전문가 실험에서는 각 동작별 1번 동작을 추출한 경우가 대부분이었으나 2차 일반인 실험에서는 1, 2번

동작을 선택한 비율이 거의 비슷하였다. 따라서 '상, 하, 좌, 우'에 대한 지시동작에서 손동작에 대한 차이는 크게 작용하지 않았으며, 자신을 좌표의 기준으로 하여 '상, 하, 좌, 우'라는 사회적으로 이미 약속된 방향을 지시하는 동작을 선호하는 것으로 나타났다. 이는 방향에 대한 구체적이고 실제적인 관계가 있는 지표적 기호이다. 이 경우 특히 지표적 기호 성향이 강해 도상적 표현이나 상징적 표현의 지시동작은 볼 수 없었다.

### c. 확인 지시동작

[그림 8] 실험에서 추출된 확인 지시동작



'확인'과 '취소'의 지시동작 추출 실험에서는 손, 팔 동작만 사용하는 것이 아닌 머리카 몸 사용하거나 동시에 사용하는 동작도 있었다. 먼저 '확인'을 나타내는 지시동작으로는 손을 360도 돌리는 1번, 팔을 시계방향 또는 반시계방향으로 360도 돌리는 3, 4번 지시동작은 O자인 동그라미 형태를 나타내는 지시동작을 표현하였다. 8번 동작은 OK라는 의미를 표현하는 손동작이다. 2, 9번 동작의 경우는 버튼을 눌러 선택하는 행위를 표현하는 팔 동작이었으며 6, 7번 동작 역시 버튼을 누르거나 마우스로 더블클릭하는 행위를 표현하는 것으로 선택한다는 의미를 담고 있는 손동작이었다. 5번 동작은 고개를 끄덕이는 동작으로 긍정의 의사를 담고 있는 몸동작이다. 이렇게 '확인'의 경우 긍정 및 선택이라는 의미를 내포하고 있는 동그라미의 O자, OK, 끄덕임, 버튼의 누름, 더블클릭 등의 행위 및 형태의 유사성에 근거하고 있는 지시동

작들로 지표적 기호에 가깝다고 볼 수 있다. 이 중 2차 실험에서는 동그라미를 표현한 3번의 팔 동작을 가장 선호하였다. 따라서 일반적으로 '확인'의 의미를 내포하고 있는 여러 표현들 중에서 동그라미 O자를 가장 먼저 떠올리는 것으로 생각할 수 있으므로 이 동작이 직관적으로 이해도가 높다고 평가할 수 있다.

#### d. 취소 지시동작

[그림 9] 실험에서 추출된 취소 지시동작



'취소'의 경우 부정 또는 아니라는 의미를 표현하는 손을 앞으로 내밀어 좌우로 흔드는 1, 7번의 손동작과 좌우로 고개를 젓는 4번의 몸동작, X자의 형태를 그리는 듯한 8번 동작, V자의 형태를 그리는 듯한 5번 동작들로 부정의 의미를 담고 있는 행위이거나 부정의 의미로 사용되는 아이콘의 형태를 나타내는 지시행동들이 1차 실험에서 많이 추출되었다. 2, 3번 동작의 경우는 '확인'에서 보여준 동작의 반대 방향으로 움직임으로써 '확인'의 반대인 취소라는 의미로 사용되고 있다. 이렇게 '취소'는 '확인'의 경우와 같이 지표적 기호의 성격을 가지고 있는 지시동작들이 대부분이었으며 이 중 2차 실험 결과 취소의 한 손으로 X자를 그리는 8번 동작을 대다수가 선호하고 있음을 알 수 있었다. 이는 '취소'의 지표적 표현으로 X자를 가장 먼저 떠올린다는 것이며 따라서 '취소'에 있어 이 형태를 그리는 듯한 8번의 동작이 가장 직관적인 동작이라 할 수 있다.

#### (3) 선정된 지시동작의 분석과 분류

디지털 TV의 기본 조작단위를 수행하기 위한 지시동작을 1, 2차 실험을 통해 추출하였다. 추출된 동작 및 동작을 기반으로 한 동작모형은 아래와 같다.

[표 6] 디지털TV 기능 수행을 위한 지시동작

task	지시동작	동작모형	동작분류
TV를 켜다/ 끈다			상징적 지시동작
메뉴list에서 위로 이동한다/상위 채널로 이동한다			
메뉴list에서 아래로 이동한다/하위 채널로 이동한다			
메뉴list에서 왼쪽으로 이동한다/ 볼륨을 낮춘다			지표적 지시동작
메뉴list에서 오른쪽으로 이동한다/ 볼륨을 높인다			
메뉴list에서 현 상태를 선택한다			지표적 지시동작
메뉴list에서 선택을 취소 /전단계 이동한다			

#### 5. 결론 및 제언

동작인식을 바탕으로 한 디지털 TV의 입력모델 개발에서 중요한 것은 사용자가 직관적으로 이해할 수 있는 지시동작의 디자인이다. 이는 기능제어의 효율성 뿐 아니라 사용성 측면에서도 매우 중요하다. 본 논문에서는 이러한 동작 모형 개발을 위해 디지털 TV의 기능 분석을 통한 기본조작단위 추출 후 전문가 대상의 1차 실험과, 일반인을 대상으로 한 2차 실험

을 바탕으로 TV 기능 제어에 관한 일반적 Mental Model에 접근, 이를 유형별 분석해 보았다. 실험과 분석을 바탕으로 각 task에 해당하는 기본조작단위별 지시동작 모형을 제시할 수 있었는데, 이때 선정된 지시동작은 지시동작 인식 프로세스에서 본 바와 같이 '동작감지', '이미지프로세싱', '제스처 모델링 및 재현', '동작추출', '인식'의 5단계를 거쳐 수행되므로 단계별 적용에 있어서 z축의 동작이나 얼굴, 몸동작 등의 '동작추출' 및 '인식'의 어려움이 예상되는 기술적 한계에 따른 오류범위 및 사용자의 시정태도를 감안하여 추출한 것이다.

최종적으로 선별된 지시동작모형의 특성을 종합적으로 분석해 보면 도상적 아이콘들을 다양하게 활용해 온 기존의 마우스나 키보드, 리모트 콘트롤 등을 통한 입력방식의 인터페이스와 달리 지시동작 기반의 인터페이스는 상징적, 지표적 특성이 강하게 나타나는 것을 알 수 있다. 따라서 이러한 동작기호를 사용하는 것이 기능수행에 있어 직관적으로 효율적일 것임을 유추할 수 있다. 이는 기존의 시각적 인터페이스에서는 실제 사물의 메타포를 이용한 그래픽적 아이콘이 필요했던 반면, 동작인식을 통한 입력방식에서는 상징적이고 지표적인 메타포가 필요하다는 것을 말해주는 것이다. '확인', '취소' 등과 같은 관념적이고 추상적인 task를 동작기호로 수행하기 위해서는 이제까지의 도상중심적 기호로는 동작 학습 및 유추 면에서 어려울 뿐 아니라 행위에 있어 동작자체를 수행하기 어렵기 때문이다.

또한 앞으로의 실제 사용자가 동작인식 기반의 디지털 TV를 사용함에 있어서는 디지털 TV의 시각적 인터페이스 즉 화면의 GUI와 동작인식 기호의 지표적, 상징적 특성이 결합되어 사용되므로, GUI가 사용자의 사고와 인식을 상호 보완하거나 유도하게 될 것이다. 따라서 동작인식 기반 디지털 TV의 GUI 및 UI는 사용자의 직관적 task 수행을 보완, 보조할 수 있도록 설계해야 할 것으로 기대된다.

## 참고문헌

- 조강현. (1999). 3차원 지시 동작 인식을 위한 휴먼 인터페이스. 울산대학교 공학 연구논문집, 30(1), 265-273
- 권숙경. (2005). 웨어러블 컴퓨터의 인터페이스를 위한 사용자 행동의 의미추출과 적용가능성에 관한 연구. 국민대학교 테크노디자인 대학원 석사학위 논문
- 김현석, 서경자, 서현석, 문현정. (2006). 디지털

TV 설정기능의 메뉴명 및 구조 분석. 기초조형학연구 Vol7. No.4

- A. Kendon. (1986). "Current Issues in the Study of Gesture," The Biological Foundations of Gestures: Motor and Semiotic Aspects. Lawrence. Erlbaum Assoc. pp. 23-47.
- D. McNeill. (1985). So you think gestures are nonverbal? Psychological Review, 92, 350-371
- Antonis A. Argyros and Manolis I.A. Lourakis. (2006). Vision-Based Interpretation of Hand Gestures for Remote Control of a Computer Mouse. Institute of Computer Science, Foundation for Research and Technology
- Darius Burschka, Guangqi Ye, Jason J. Corso, Gregory D. Hager.(2005).A Practical Approach for Integrating Vision-Based Methods into Interactive 2D/3D Applications. Computational Interaction and Robotics Lab Technical Report 05-01.
- Flavia Sparacino.(2001). (Some) computer vision based interfaces for interactive art and entertainment installations. In: INTER\_FACE Body Boundarie, n.2, Paris, France.
- Moritz Storing, Thomas B. Moeslund, Yong Liu, and Erik Granum. (2004). COMPUTER VISION-BASED GESTURE RECOGNITION FOR AN AUGMENTED REALITY INTERFACE. In 4th IASTED International Conference on VISUALIZATION, IMAGING, AND IMAGE PROCESSING, pages 766-771, Marbella, Spain.
- Hideyuki SAWADA, Seiji UTA and Shuji HASHIMOTO. (1999). Gesture Recognition for Human-Friendly Interface in Designer - Consumer Cooperate Design System. Proceedings of the 1999 IEEE International Workshop on Robot and Human Interaction Pisa, Italy
- Matthew Turk, Jeremy Bailenson, Andrew Beall, Jim Blascovich, Rosanna Guadagno(2004.) Multimodal Transformed Social Interaction. Proceedings of the 6th international conference on Multimodal interfaces
- Le Song and Masahiro Takatsuka. (2006). Real-time 3D Finger Pointing for an Augmented Desk. National ICT Australia
- Vladimir I. Pavlovic, Rajeev Sharma, Thomas S. Huang. (1997). Visual Interpretation of Hand Gestures for Human-Computer Interaction. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 7

- Thomas B. Moeslund, Moritz St'orring, Erik Granum. (2000). Vision-Based User Interface for Interacting with a Virtual Environment. Aalborg University, Denmark.
- Hatice Gunes, Massimo Piccardi, Tony Jan. (2004). Face and Body Gesture Recognition for a Vision-Based Multimodal Analyzer. ACM International Conference Proceeding Series; Vol. 100
- Rick Kjeldsen, Jacob Hartman. (2001). Design Issues for Vision-based Computer Interaction Systems. ACM International Conference Proceeding Series; Vol. 15
- Paul Robertson, Robert Laddaga, Max Van Kleek. (2004). Virtual Mouse Vision Based Interface. Proceedings of the 9th international conference on Intelligent user interfaces