

## 인간친화형 인터페이스를 위한 사용자 얼굴에서의 효과적인 의도 파악

김 대 진, 송 원 경, \*김 종 성, 변 중 남  
한국과학기술원(KAIST) 전자전산학과, \*한국전자통신연구원(ETRI) 가상현실 연구센터  
전화 : 042-869-8019 / 핸드폰 : 011-9017-6689

### An Effective Intention Reading from User Face for Human-Friendly Interface

Dae-Jin Kim, Won-Kyung Song, \*Jong-Sung Kim, Zeungnam Bien  
Dept. of Electrical Engineering and Computer Science, KAIST  
\*VR Center, ETRI  
E-mail : djkim@ctrsys.kaist.ac.kr

#### Abstract

In this paper, an effective intention reading scheme is proposed for human-friendly interface. Soft computing techniques such as fuzzy logic and artificial neural networks are used for this. And Gabor filter based feature(GG feature) is also proposed to deal with local activity in the human face. It is based on human visual system and Gabor filter based approach is very popular in these days.

The proposed scheme is adopted for human-friendly interface for rehabilitation service robotic system KARES II.

#### I. 서론

사용자 얼굴에 대한 인식 및 그 응용은 인간친화형 인터페이스로서의 효용가치를 가진다. 사용자 식별, 침입자 감시 장치, 인간-로봇 상호작용 기술 등 그 응용 분야 또한 다양하다. 특히, 최근 각광받고 있는 애완용 로봇이라든가 미래형 로봇, 기타 인간과 생활을 함께 영위할 서비스 로봇에서 사용자 얼굴 인식 및 의도 파악 기술은 무한한 가능성을 갖고 있으며 많은 연구들이 행해지고 있다. 사용자 얼굴에서의 의도 파악 문제

는 가장 어려운 문제 중에 하나로 인식되어지며, 아직도 해결해야 할 문제가 많은 것으로 알려져 있다.

본 논문은 인간친화형 인터페이스를 위한 사용자 얼굴에서의 감정 인식 문제를 다룬다. 특별히, 인간-로봇 상호작용 기술의 일환으로 사용자 얼굴에서 입 부분의 특징을 바탕으로 로봇의 서비스 수준에 대한 만족도를 인식함으로써 효과적인 감정 인식을 수행하는 문제를 다룬다. 감정 인식 과정에 사용된 주요 개념은 퍼지 논리에 기반한 의사결정 시스템이며, 특징 추출을 위해서는 인간의 시각 시스템에 기반한 영상 처리 연산자의 하나인 가버(Gabor) 필터를 이용한다. 구축된 시스템을 기반으로 인간-로봇 상호작용 기술을 구현, 실험 결과를 제시하고 그 유용성 및 가능성을 기술한다.

#### II. 재활 서비스 로봇과 인간친화형 인터페이스

##### 2.1 재활 서비스 로봇

재활 서비스 로봇은 재활 공학과 서비스 기술과의 결합으로 탄생한 서비스 로봇의 일종이다. 미국이나 유럽, 일본 등지에서는 이미 90년대 초반부터 많은 연구가 이루어졌으며, 지금은 상당한 수준에 이르고 있

다[1]. 국내의 경우에도 많은 수는 아니지만 다양한 연구가 이루어지고 있다[2][3]. 한국과학기술원에서 개발 중인 KARES(KAIST Rehabilitation Engineering Service robot system) II 의 경우가 대표적인 예라 하겠다(그림 1 참조). KARES II 는 기본적으로 휠체어 장착형 로봇팔을 주축으로 여러가지 인간친화형 인터페이스들이 탑재되어 있다. KARES II 에서 가장 중시되는 개념은 '자율성'과 '안전성'이다. 자율성을 부여하기 위해 KARES II 에서는 지능형 비주얼 서보잉(Visual Servoing) 기법을 채용하고, 안전성을 위해서 소프트 로봇팔, 인간친화 공동작업 기술 등이 포함되어 있다. 이들과의 효율적인 연계성을 위하여 다양한 인간친화형 인터페이스의 도입은 필수적이다.

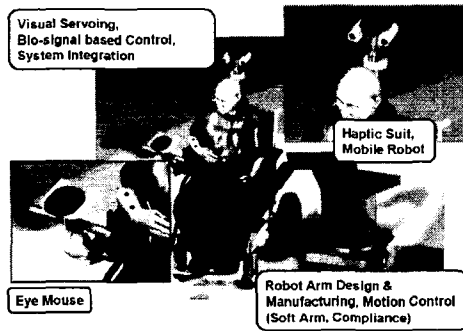


그림 1 휠체어 기반 소프트웨어 로봇팔

## 2.2 인간친화형 인터페이스

KARES II 에서는 인간친화형 인터페이스로써 아이마우스(Eye-mouse), 햅틱수트(Haptic Suit), 생체신호(EMG) 기반 제어기술 등이 이용되고 있다[4][5][6]. 이러한 독립적인 인터페이스를 위한 기술 이외에도 지능형 비주얼 서보잉을 위한 인터페이스로 연구 중인 것이 사용자 얼굴에서의 의도 파악 기술이다.

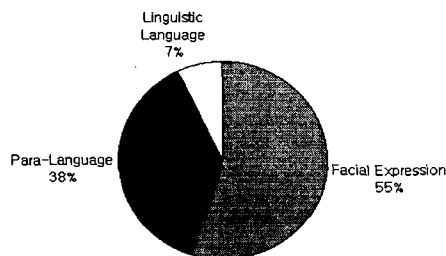


그림 2 인간의 의사소통시 정보량 분포 [7]

인간은 상대방의 의도 파악을 위해서 여러 가지 정보를 이용한다. 언어적인 수단보다는 비언어적인 수단, 즉 몸짓이라든가 표정에 의해서 전달되는 정보량이 더 많은 사실은 익히 알려져 있다. 그 중에서도 얼굴 표정은 약 55%에 해당하는 정보량을 전달한다(그림 2 참조). KARES II 의 경우, 휠체어 사용자를 대상으로 한 작업이 대부분이므로 사용자와의 효과적인 상호작용을 위해서는 얼굴 표정에 의한 의도 파악 기술이 매우 중요하게 된다.

## III. 효과적인 의도 파악

사용자 얼굴에서의 의도 파악은 사용자의 입 부위 추출, 입 부위에 대한 영상 특징 추출, 추출된 특징을 기반으로 한 언어 변수의 값 설정, 퍼지 논리에 의한 의도 파악의 순으로 진행된다.

### 3.1 얼굴 및 입 부위 추출

본 논문에서는 얼굴 영역 추출을 위해 I1I2I3 컬러 공간을 이용하였다[8]. 다양한 컬러 공간과의 비교를 통해 I1I2I3 컬러 공간의 유용성이 검증되었다. I1I2I3 컬러 공간은 기존의 RGB 컬러 공간에 대하여 식 (1)과 같이 표현된다.

$$\begin{aligned} I_1 &= (R + G + B)/3 \\ I_2 &= R - B \\ I_3 &= (2G - (R + B))/2 \end{aligned} \quad (1)$$

추출된 얼굴 영역에서 입은 얼굴 하단에 위치하므로, 적분투영(Integral Projection) 과정을 거쳐서 입 부위만을 추출할 수 있다.

### 3.2 의도 파악을 위한 특징

본 논문에서는 의도 파악을 위한 특징으로 전역적인(Global) 특징과 국부적인(Local) 특징을 모두 사용하였다. 이것은 실제로 사람이 상대방의 얼굴 부위를 관찰할 때도 사용되는 개념이며, 얼굴 인식 분야 혹은 여타 영상 처리 분야에서도 중요한 개념이다.

#### (1) 얼굴 부위에 대한 면적비 및 높이비

전역적인 특징으로는 입 부위의 얼굴 부위에 대한 면적비(Area Ratio) 및 높이비(Height Ratio)가 선정되었다. 이것은 3.1 절에서 설명한 얼굴 및 입 부위 추출

과정을 통해 쉽게 얻어질 수 있다. 면적비와 높이비의 경우 영상의 크기와 사용자 얼굴의 영상 내에서의 크기에 무관하다는 특징을 갖는다.

(2) 가버(Gabor) 필터 기반 특징

가버 필터는 인간의 시각 시스템을 모방한 것으로써 이미 그 유용성에 대한 연구가 활발히 이루어지고 있다[9]. 본 논문에서는 3.1 절에서 추출된 입 부위에 해당하는 흑백(Grey) 영상에 대한 가버 필터 계수를 이용한 특징을 제안하고 이를 입 부위에 대한 국부적인 특징으로 이용한다.

필터링된  $W \times H$ 의 크기를 갖는 영상  $I_G$ 에 대하여, 수직투영(Vertical Projection)을 행한 다음 수직투영된 결과의 오프셋(Offset)의 영향을 줄이기 위하여 미분된 결과의 절대값을 취한다. 추출된 입 부위의 크기와 영상의 밝기값에 무관한 특징치를 얻기 위하여 가우시안(Gaussian) 가중치를 이용하여 식 (2)와 같은 특징을 제안하였다. 이 특징은 가버 필터와 가우시안 가중치에 기반한 것에서 가버-가우시안(Gabor-Gaussian; 이하 GG) 특징으로 명명하였다.

$$f_G = \frac{\sum_{i=1}^{H-1} w_G(i) dy_{proj}(i)}{\sum_{i=1}^{H-1} w_G(i)} \quad (2)$$

3.3 퍼지 논리에 의한 의도 파악

퍼지 논리는 기존의 논리 구조에 비하여 애매하고 구체적으로 정량화하기 어려운 문제를 효과적으로 다룰 수 있다는 장점을 갖는다. 또한, 사람이 가진 기존의 지식을 언어 변수를 통해 표현함으로써 그 논리 구조의 변경이 용이하고 전문가의 지식을 표현하는데도 용이하다. 퍼지 논리에 기반한 의도 파악 역시 이러한 장점을 갖는다[10].

(1) 퍼지룰의 구성

사용자 입 부위에 대한 퍼지룰은 비교적 간단하게 구성된다. 우선, 언어 변수로는 사용자의 입 부위에 대한 개폐정도(Mouth Openness)와 사용자의 만족도를 긍정/부정(Positive/Negative)으로 구분하여 선정하였다. 각 언어변수는 2단계의 언어 변수값(High, Low)을 가질 수 있으며, 전체 퍼지룰은 다음과 같이 구성된다(표 1 참조).

표 1 의도 파악을 위한 퍼지룰

IF(전건부)	THEN(후건부)
Mouth Openness is High	Positive is High
Mouth Openness is Low	Positive is Low
Mouth Openness is High	Negative is Low
Mouth Openness is Low	Negative is High

(2) 소속도 함수의 결정

각 언어 변수에 대한 소속도 함수의 결정을 위하여 퍼지 관찰기가 사용되었다[10]. 퍼지 관찰기는 영상 특징치를 이용하여 간접적으로 언어 변수를 측정하는 방법이다. 3.2 절에서 추출한 3가지 특징값을 학습 데이터로 이용하여, 긍정/부정에 대한 2가지 언어 변수값을 결정한다. 이 값은 최종적으로 퍼지룰의 추론 과정에 사용된다.

IV. 재활 서비스 로봇에 대한 적용 및 토의

4.1 실험 환경 및 시나리오

본 논문에서 제안한 의도 파악 방법은 실제 재활 서비스 로봇인 KARES II 에서 사용자에 대한 식사 작업을 보조하기 위한 인터페이스로 활용되었다.

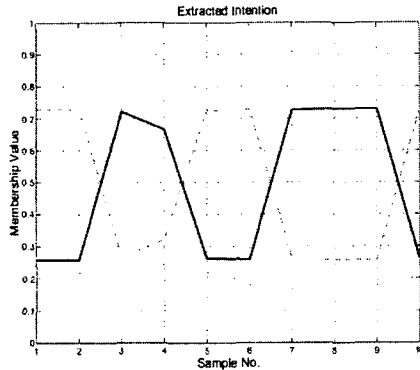
실험 시나리오는 다음과 같다. 식사 작업 중에 사용자에게 컵에 담긴 마실 것을 제공하는 경우, 사용자가 거부 반응(부정적인 반응)을 보이면 로봇팔은 서비스 작업을 중단하기 위하여 컵을 사용자의 입 근처로부터 멀어지게 움직일 것이고 긍정적인 반응을 보이면 컵을 입 근처로 접근시키게 된다(그림 3 참조). 의도 파악은 사용자의 안전성을 위하여 한 가지 작업(접근 혹은 멀어짐)이 완료된 시점에서 진행된다.



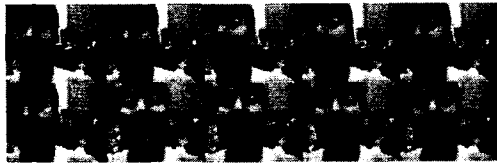
그림 3 사용자에 대한 의도 파악

4.2 실험 결과

실제 실험을 통해 얻어진 의도 파악 결과는 그림 4와 같다. 그림 4(b)의 연속된 10장의 영상에 대하여 그림 4(a)와 같은 의도가 파악되었다. 실선은 긍정에 대한 소속도 함수, 파선은 부정에 대한 소속도 함수를 나타낸다.



(a)



(b)

그림 4 사용자에 대한 의도 파악 실험 결과

V. 결론

본 논문에서는 인간친화형 인터페이스를 위한 사용자 얼굴에서의 의도 파악 문제를 퍼지 논리와 가버 필터를 기반으로 한 영상 특징에 기반하여 다루었다. 제안된 방법은 퍼지 논리를 사용함으로써 감정 인식 문제와 같이 애매하고 구체적으로 정량화하기 어려운 문제를 효과적으로 다룰 수 있다. 특히, 가버 필터를 기반으로 한 영상 특징은 최근 영상 처리 분야에서 전 세계적으로 자주 거론되는 주요 개념으로써 그 응용 분야 역시 상당할 것이다. 제안된 방법은 6축 실험용 로봇팔과의 상호작용에 이용되어 서비스 대상자인 사용자의 의도 파악에 효과적으로 이용되었다.

본 논문은 인간-로봇 상호작용 기술에 있어서 사용자 얼굴 인식에 의한 사용자 의도파악을 위한 근간을 제시하며, 차후 고수준의 사용자 의도 파악 기술 구현을 위한 기초자료로 활용될 수 있을 것으로 사료된다.

참고문헌(또는 Reference)

- [1] K. Kawamura, R.A. Peters II, D.M. Wilkes, W.A. Alford, and T.E. Rogers, "ISAC: Foundations in Human-Humanoid Interaction", to appear in IEEE Intelligent Systems, July/August 2000.
- [2] Z. Bien, and W.-K. Song, "Novel wheelchair-based robotic arm with visual servoing capability for human-robot interaction," Workshop on Service Automation and Robotics, Hong Kong, June, 2000.
- [3] C.-S. Lee, J.-S. Kim, G.-T. Park, W. Jang, and Z. Bien, "Implementation of Real-time Recognition System for Continuous Korean Sign Language(KSL) mixed with Korean Manual Alphabet(KMA)", The Institute of Electronics Engineers of Korea, vol. 35-C, no. 6, pp. 76-87, June 1998.
- [4] D.H. Yoo, Y.J. Lee, G.W. Chu, J.H. Kim, D.H. Kim, and M.J. Chung, "Assistive Robot Control by an Eye-mouse System for the Disabled," 4th Biannual World Automation Congress, 2000.
- [5] K. Lee and D.-S. Kwon, "Application of optical fiber sensors for the wearable master device used by the disabled," 1st Int. Workshop on Human-friendly Welfare Robotic Systems, Taejon, Korea, pp. 65-68, Jan. 2000.
- [6] J.-S. Han, W.-K. Song, J.-S. Kim, W.-C. Bang, H. Lee, Z. Bien, "New EMG Pattern Recognition based on Soft Computing Techniques and Its Application to Control of Rehabilitation Robotic Arm", 6st Int. Conf. on Soft Computing(IIZUKA2000), IIZUKA, Japan, Oct. 1-4, 2000.
- [7] A. Mehrabian, Communication without words, In Joseph A. DeVito, editor, Communication: Concepts and Process, pp.106-114, Prentice-Hall, Inc., 1971.
- [8] Y. Ohta, T. Kanade and T. Sakai, "Color Information for Region Segmentation," Computer Graphics and Image Processing, vol. 13, pp. 222-241, 1980.
- [9] M. Lyons, S. Akamatsu, M. Kamachi and J. Gyoba, "Coding Facial Expressions with Gabor Wavelets," 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition, pp. 200-205, 1998.
- [10] G.-T. Park and Z. Bien, "Neural network-based fuzzy observer with application to facial analysis", Pattern Recognition Letters, vol. 21, pp. 93-105, February, 2000.