

# PC 카메라에서 추출한 이미지를 이용한 수화인식

## Recognition of Finger Language using Image from PC Camera

이 병 환\*, 이 기 성\*\*

(Byoung Hwan Lee, Keeseong Lee )

\*홍익대학교전기정보제어공학과(전화:(02)325-7514, 팩스:(02)320-1110, E-mail : byounghwan76@hotmail.co.kr)

\*\*홍익대학교전기정보제어공학과(전화:(02)320-1670, 팩스:(02)320-1110, E-mail : leeksyh@yahoo.com)

**Abstract** - Finger language is a typical tool for deaf persons. But learning the finger language for non-handicapped persons is very difficult. To overcome these difficulties, a new communication method using visual function is developed recently. Even though the developed system uses the visual function, it needs expensive equipments such as camera and computer. To be used in the real environments, the cost of equipments is a critical factor. If the recognition system for the finger language can be developed with low price equipments, the system can be used in the notebook or cellular phone. The image captured by PC camera was processed by preprocessing algorithm. To recognize the finger language, the resulting image was divide into 5x5 sections. The recognition system uses a similarity method and position information. The simulation results shows the effectiveness of the proposed algorithm.

**Keywords** : Recognition of Finger Language, Visual Function, PC Camera.

### 1. 서 론

청각 장애를 가지고 있는 사람들을 위해 기존에 개발되어 있는 통역도구들은 일반인이 장애인과 같이 사용하려면 새로 배워야하는 어려움이 있었다. 이를 극복하기 위해서 근래에 가장 많이 개발되고 보강되고 있는 것은 비주얼기능 이용한 의사교류에 대한 연구이다.

본 논문에서는 비주얼 기능을 이용하여 장애를 가진 사람과 장애를 가지지 않은 사람과의 의사소통에 도움을 줄 수 있는 수화인식 시스템을 개발하였다. 이 시스템은 카메라를 이용하여 물체의 영상을 받아들이고 그 받아들인 영상을 제한하는 인식방법을 이용하여 해석하여 기존에 저장되어 있는 몇가지의 데이터와 비교하여 단어를 인식하게 된다.

수화는 흔히 삼차원적인 언어라고 이야기 되어진다. 손 모양, 손위치, 그리고 손의 이동방향등 이 세가지가 크게 수화의 의미를 만들어내는 중요한 요소이다[3,4,6]. 이 세가지 요소가 결합되어 만들어지는 수화는 수화단어와 지문자로 나누어진다. 수화단어란 우리가 흔히 길에서 볼 수 있는 명사의미를 나타내는 것을 흔히 말하고 기본단어, 유사단어, 복합단어로 나눌수 있다. 기본단어는 하나의 의미를 정의한 기초단어이고 유사단어는 수화제스처등과 동일하게 쓰이는 의미가 유사한 단어로 정의된 기본단어이다. 마지막으로 복합단어는 두가지 이상의 단어를 결합하여 하나의 뜻을 만드는 단어로 '작다'란 단어와 '아버지'란 단어가 합쳐져 '작은 아버지'란 식으로 표현되는것들을 말한다. 본 논문에서 자료로 사용할 지단어는 손모양으로 표현하는 기본단어중 명사를 골라낸것이다[3,5].

이러한 수화의 동작성과 3차원적 움직임은 수화인식을

어렵게 하는 큰 장애요소가 되었다. 여러 가지 아이디어와 기술이 개발되어 수화인식 시스템이 개발되었으나 고가의 컴퓨터와 장비 그리고 장소의 제한성 때문에 크게 환영받지는 못하고 있는 실정이다.

개인용 컴퓨터와 저사양의 PC 카메라를 가지고도 수화인식을 성공한다면 좀 더 작은 노트북이나 혹은 카메라 핸드폰을 이용해서도 수화의 의미를 전달받을 수 있게 된다.

### 2. 영상 처리

영상을 받아들이고 알고리즘을 사용하여 전처리하는 과정은 Microsoft사에서 개발한 Visual C++.net을 기반으로 하여 개발하였다. 여러 사람이 쉽게 사용할 수 있도록 펜티엄4급의 하드웨어를 가진 컴퓨터와 PC용 카메라를 사용하고 Window XP를 기본 OS로 사용 하였다.

본 논문에서 고사양의 컴퓨터와 고사양의 카메라를 사용하지 않은 이유는 기기와 장소의 제한을 받지 않고 인식을 하기 위해서이다. 기기의 제한이 없으면 좀 더 이동성이 높은 인식장치를 개발하여 PC 카메라와 노트북만 가지고 있으면 수화를 별도로 익히지 않고도 의사소통을 할 수 있게 된다.

기존 논문들은 수화의 기본요소를 4가지로 나누고 있다. 그 4가지 요소는 손 운동, 손 모양, 손 위치, 손 방향이다. 본 논문에서는 손 방향은 손 모양과 손 운동에서 검출한 위치정보로 움직이는 방향과 손 방향을 알아낼 수 있으므로 따로 분류하지 않았다. 영상을 이용한다는 것은 기본적으로 기존의 칼라에 기초한 인식의 확장으로 볼 수 있으나 발달된 알고리즘을 이용하여 좀 더 간단하고 명료한 영상 처리를 할 수 있게 된다[9].

## 2.1 영상의 전처리

카메라에서 받아들인 영상의 크기는 320\*240으로 정했다. 영상의 저장형식은 영상의 처리를 원활히 하기 위해 bmp 방식을 사용하였다.

카메라가 받아들인 영상은 우선 전처리과정을 거치게 된다. 영상을 전처리 하는 이유는 받아들인 영상 안의 배경값을 보정하기 위해서이다. 여기서 선택된 디지털 영상 처리 방법은 RGB 모델 추출법과 HSI 모델 추출법이 사용되었다.

### 2.1.1 RGB 모델 추출법

RGB값이란 컬러 CRT 모니터와 컴퓨터 그래픽스에서 사용하는 직각 좌표계이다. 빛의 삼원색을 이용하여 원하는 색을 만드는 방법인데 예를 들어 RGB컬러모델에서 검정색은 좌표계 (0,0,0)의 값을 가지고 (1,1,1)는 흰색을 가진다. 그림 1은 RGB 컬러 육면체를 나타낸다.

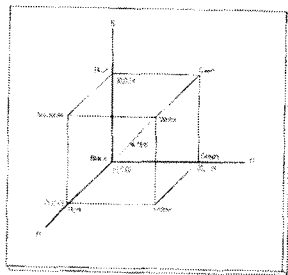


그림 1. RGB 컬러 육면체

### 2.1.2 HSI 모델 추출법

앞에서 말한 RGB 모델 추출법이 시스템이나 하드웨어에서의 사용을 위해 만들어진 색상모형이라면 HSI모델은 인간의 색인지에 기반을 둔 사용자 지향성의 색상모형이다. H는 색상(Hue), S는 채도(Saturation), I는 명도(Intensity)를 나타내며 이모형은 구체적인 컬러를 만들기 위해 색을 조합할 필요가 없다. 그림 2는 HSI 컬러모형을 나타내고 그림 3은 프로그램상에서 HSI 모델 추출법을 적용하여 영상 전처리를 보여주고 있다[1,10].

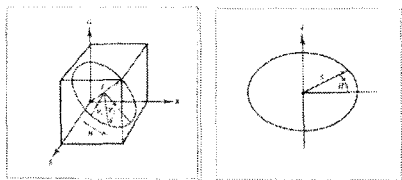


그림 2. HSI 컬러 모형

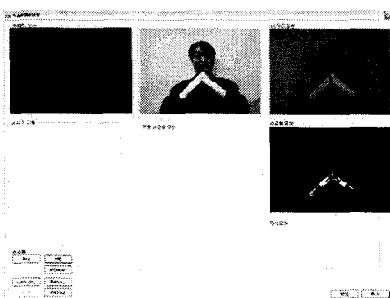


그림 3. 영상 전처리 (예시 1)

## 2.2 영상의 위치정보

수화는 손의 위치와 손의 운동 방향이라는 삼차원적 움직임을 가지고 있는 3차원적 언어이기 때문에 화상처리나 인식을 위해서 고사양의 카메라와 고사양의 컴퓨터가 필요한 것이다.

제안하는 알고리즘을 사용하면 저사양의 컴퓨터와 저사양의 카메라를 이용하여도 수화를 인식 할 수 있다. 카메라에서 받아들인 영상 그림4의 모양처럼 다음과 같이 분할하게 된다[1,10].

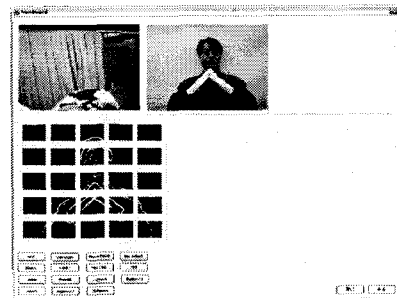


그림 4. 영상의 분할처리(예시 2)

좌측 최상단부터 (64\*48)의 크기로 오른쪽으로 순서대로 번호를 정한다. 그러면 5\*5의 25개의 분할면을 구할 수 있고 1번 - 25번 까지 번호를 지정한 후 분할된 영상을 관찰해보면 8번은 머리 12번은 오른쪽 어깨, 13번은 턱, 14번은 왼쪽 어깨, 17번은 오른쪽 몸통, 18번은 가슴, 19번은 왼쪽 몸통, 22번은 오른쪽 허리, 23번은 복부, 24번은 왼쪽 허리를 나타내는것을 볼 수 있다. 이위치는 카메라와 대상과의 거리를 1m로 한정했을 때 나오는 영상의 위치이다. 이는 대상이 키와 체격에 상관없이 적용되고 이 위치정보는 손의 위치를 판별하는데 쓰이게 된다.

손의 모양과 위치는 HSI값을 검출하여 찾아낸다. 좀 더 정확한 판별을 위해 손에 흔히 구할 수 있는 하얀색 면장갑을 착용한다. 하얀색 면장갑의 HSI값을 구한 후 그 값만 화면에 출력하게 되면 그림5에서처럼 출력하게 된다. 최초 검출된 영상 값을 이용하여 영상을 필요한 부분만 검출해 낼 수 있다. 이 검출 방법은 손의 위치가 어느방향으로 움직이게 되어도 손모양을 정확히 찾아오게 때문에 손모양에 대한 인식률을 높일수 있다.

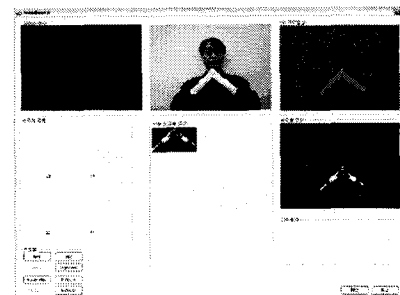


그림 5. 손모양의 검출

이렇게 출력된 손모양은 미리 저장해 놓은 이미지 데이터 베이스에 유사도법을 이용하여 비교된후 수화를 인식하

게 된다.

본 논문에서 인식 과정에 보통 많이 사용하는 Template Matching 법을 사용하지 않는 이유는 Template Matching 법은 두 가지 물체가 완전히 같아야 같은 물체라고 인식하게 된다. 하지만 이 인식될 영상은 항상 같은 환경하에서 촬영할 수 없고 주변 사물에 의한 영향을 무시할 수 없다. 그래서 유사도법을 사용하여 기존의 데이터와 비교하여 손 모양을 찾게 되고 손 모양을 찾게 되면 이미 전에 기억되어있는 영상 전체의 위치에 적용하여 인식을 하게된다[8].

Template Matching법은 비교영상의 픽셀을 직접비교하여 픽셀에 필요한 값이 들어있는지를 판별하지만 유사도법은 한단계 더 나아가 그 픽셀을 수치화해서 기억한후 영상 전체의 픽셀 크기와 비교 영상의 픽셀크기를 비교해 그 퍼센트와 일치도를 이용해 물체를 판별한다.

$$t = \frac{s}{x \times y} \times 100 \quad (1)$$

위의 식에서 s는 비교시 값을 가지는 픽셀의 개수이고  $x \times y$  는 픽셀의 전체 크기 이다. 그렇게 나오게 되는 결과 값 t가 비교영상이 가지는 퍼센트가 된다.

이러한 비교법을 가짐으로써 비교시 주변 사물과 기타 환경에의 영향을 무시할 수 있게 되어 좀 더 확장된 인식 환경을 가질 수 있게 되었다.

25분할된 이미지에서 손의 위치와 인체의 위치를 숫자화해서 기억하게 된다. 예를 든다면 '집'라는 단어는 손의 위치는 가슴이고 저장된 손 모양과 새로 받아들이는 이미지의 모양이 유사하면 분할된 위치 18번에 손모양이 검출되어 있으면 이 프로그램은 그 이미지를 '집'라는 단어로 인식하게 된다[5,7].

### 3. 실험 및 결과

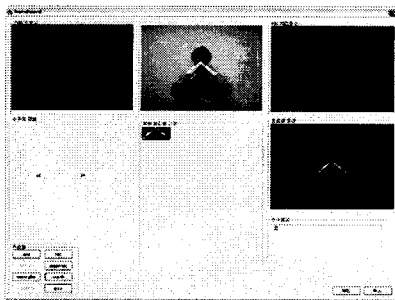


그림 5. 수화인식 결과

그림5는 수화인식결과를 보여준다. 오른쪽 상단부에는 검출된 RGB값을 적용한 이미지가 출력되고 그 아래에는 HSI처리를 거쳐 조정된 결과값이 적용된 이미지가 출력된다. 중앙부에 조그마하게 나온 사진은 결과값에서 손부분만 잘라 출력한 부분이고 왼쪽 중간부분은 영상에서 손이 검출된 위치를 나타내게 된다. 위의 결과에서 볼수있듯이 손의 모양을 정확하게 검출 해내기 때문에 손의 위치가 변하게 되어도 인식률은 낮아지지 않는다.

위의 결과로 손의 모양과 위치를 알수있으며 이 데이터를 기본으로 하여 영상처리한 결과값이 어떤단어를 나타내는가를 오른쪽 하단부에 출력하게 된다. 위의 그림에서 볼

수있듯이 대상의 키와는 상관없이 일정한 위치에 대상의 얼굴과 흉부가 위치하기 때문에 어느 누구에게나 사용할수 있는 범용성을 가질수 있고 좀더 높은 인식률을 가질수 있다.

	대상1	대상2	대상3	대상4	대상5
대상1 (나)	95%	95%	87%	55%	88%
대상2 (동쪽)	97%	90%	88%	40%	90%
대상3 (북쪽)	98%	88%	90%	58%	90%
대상4 (뺨)	98%	94%	90%	65%	98%
대상5 (서쪽)	98%	94%	90%	50%	98%
대상6 (소)	95%	95%	97%	50%	90%
대상7 (집)	90%	98%	97%	35%	98%

표1. 실험대상별 단어 인식률

위 표1의 단어들은 40여 단어에서 임의로 선별된 나, 동쪽, 북쪽, 뺨, 서쪽, 소, 집이란 단어로서 인식률의 정확성을 비교하기 위해 비슷한 손모양의 단어를 모아놓은것이다.

### 4. 결 론

본 연구에서는 영상을 인식하는데 절차와 프로그램의 처리속도를 높이기 위해 1차적으로 영상 전처리를 이용하여 처리하려는 용량을 줄였고 영상의 위치를 인식하기 위해 2차적으로 화면을 분할하여 저장하였다. 마지막으로 유사도법을 이용함으로써 각각 사람마다 다른 특징과 기타환경에서의 인식률을 높였다.

이로서 과거 고사양을 요구하던 수화인식의 하드웨어를 좀 더 저사양으로 만들 수 있었고 이로 인해 수화인식프로그램에 대한 저변확대와 보급을 원활히 함으로서 장애인과 비장애인의 의사소통을 원활하게 하였다. 하지만 아직까지 손 모양과 그리고 손의 이동경로를 탐색하여 동작을 필요로 하는 수화에서는 약한 부분을 보였고 손모양의 인식에서 많은 보완을 필요로 했다.

### 참 고 문 헌

- [1] 장동혁, "디지털 영상처리의 구현", 정보게이트, 2002.
- [2] Gonzalez & Wood *Digital Image Processing Second Edition* Prentice Hall, 2002.
- [3] 김기훈, "청각 장애인과 정보통신", 청각 장애인 정보통신 재할 심포지엄, 1996.
- [4] 이찬수, "지문자를 포함한 한글 수화의 실시간 인식시스템 구현에 관한 연구", 한국 과학기술원석사 학위논문, 1977
- [5] "특수학교(청각 장애) 중학부 수화 I", 교육부, 1992
- [6] 김승국, "한국 표준 수화 연구", 단국 대학교 특수교육과, 1994
- [7] "표준 수화 사전"金玉學術文化財團, 1982
- [8] Milan sonka, Vaclav Hlavac and Roger Boyle, *Image Processing, Analysis, and Machine Vision*, PWS Publishing, 1998
- [9] 김정배, "칼라 손 추적기를 이용한 실시간 한글 수화 인식 시스템의 확장" 한국 과학기술원 석사학위논문, 1998
- [10] 강동중, Visual C++을 이용한 디지털 영상처리, 사이텍 미디어, 2003