
ART2 알고리즘을 이용한 수화 인식

김광백* · 우영운**

Sign Language Recognition Using ART2 Algorithm

Kwang-Baek Kim* · Young Woon Woo**

요 약

수화는 청각 장애인에게 있어 중요한 의사소통 수단이며, 청각 장애인들은 수화를 통해 대인관계를 넓히며 또한 불편함 없는 일상생활이 가능하다. 그러나 최근 인터넷 통신의 발전으로 증가하고 있는 많은 화상 채팅 및 화상 통화 서비스에서는 건청인과 청각 장애인 사이에 통역을 하는 시스템이 없어 청각 장애인들이 불편을 겪고 있다. 본 논문에서는 이러한 문제점을 해결하기 위하여 수화 인식 기법을 제안하였다. 제안된 기법에서는 화상 카메라를 통해 얻어진 영상에서 각각 RGB, YUV, HSI 색상 정보를 이용하여 두 손의 위치를 추적하여 잡음을 제거한 후, 두 손의 영역을 추출한다. 추출된 손의 영역은 잡음과 훼손에 강한 ART2 알고리즘을 이용하여 학습한 후 인식한다. 본 논문에서 제안된 수화 인식 방법을 실험한 결과, 수화에 사용된 지 숫자 1부터 10을 효율적으로 인식하는 것을 확인하였다.

ABSTRACT

People who have hearing difficulties use sign language as the most important communication method, and they can broaden personal relations and manage their everyday lives without inconvenience through sign language. But they suffer from absence of interpretation between normal people and people who have hearing difficulties in increasing video chatting or video communication services by recent growth of internet communication. In this paper, we proposed a sign language recognition method in order to solve such a problem. In the proposed method, regions of two hands are extracted by tracking of two hands using RGB, YUV and HSI color information from a sign language image acquired from a video camera and by removing noise in the segmented images. The extracted regions of two hands are learned and recognized by ART2 algorithm that is robust for noise and damage. In the experiment by the proposed method and images of finger number from 1 to 10, we verified the proposed method recognize the numbers efficiently.

키워드

수화 인식, 색상 정보, 지 숫자 Sign language recognition, Color information, Finger number

I. 서 론

수화는 청각 장애인이 사물을 이해하고 의사전달을 위해 청각 장애인에 의해 고안된 언어이다. 수화는 청각

장애인의 손 모양과 손의 위치, 손의 운동을 통해 언어적 의미에 대한 시각적 이해를 제공하는 의사소통 수단이며[1], 수화는 회화적이기 때문에 감정의 변화를 정확히 밝혀야 하며, 얼굴 표정과 몸짓 하나가 대화 내용에 아

* 신라대학교 컴퓨터정보공학부

** 동의대학교 멀티미디어공학과

주 중요한 역할을 한다. 최근 인터넷의 발전으로 인한 화상 채팅과 화상 통화 서비스가 늘고 있지만 청각 장애인과 비장애인의 의사소통을 위한 서비스는 잘 이루어지지 않고 있다. 본 논문에서는 청각장애인과 비장애인의 원활한 의사전달이 가능하도록 하기 위해 ART2를 이용한 수화인식 방법을 제안한다.

본 논문에서는 화상 카메라를 통해 얻어진 청각장애인의 두 손을 추적 및 추출하기 위해 RGB, YUV, HSI 색상 정보를 이용하며, 추출된 두 손의 영상을 잡음과 훼손에 강한 ART2 알고리즘[2]을 적용하여 학습한 후 인식한다.

II. 색상 정보를 이용한 손 영역 추적 및 추출

수화는 손의 움직임으로 이루어진 언어이기 때문에 수화의 자동 인식을 위해서 화상 카메라를 통해 얻어진 청각 장애인의 손을 지속적으로 추적해야 한다. 본 논문에서는 전체 영상에서 사람의 피부색만을 추출하기 위해 RGB, YUV, HSI 색상 정보를 이용하여 피부색을 가장 잘 나타낸 영역들을 각각의 색상 정보에서 선택하여 피부색 추출에 적용한다. 하지만 추출된 피부색 영역 중에서 피부색과 유사한 배경 영역도 같이 추출되는 경우가 발생하기 때문에 이러한 영역들은 잡음으로 간주하여 제거한다.

2.1 RGB, YUV, HSI 색상 정보를 이용한 피부색 추출

본 논문에서는 화상 카메라로 획득한 영상에서 피부색 영역만을 추출하기 위해 RGB, YUV, HSI와 같은 3가지 색상 정보[3]를 이용한다. RGB는 색상이 적색, 녹색, 청색의 기본 요소로 구성되어 있으며 색상 정보와 밝기 정보가 혼합되어 있는 색상 모델이다. YUV는 휘도 신호(Y)와 적색 성분의 차(U), 휘도 신호와 청색 성분의 차(V)의 3가지 정보로 색을 나타내는 색상 모델이며 HSI는 인간의 색인지에 기반을 둔 색상 모델로써, H는 색상, S는 채도, I는 명도를 나타낸다. 식 (1)은 RGB 정보를 이용해 YUV 색상 모델로 변화하는 수식이고, 식 (2)는 RGB 정보를 이용해 HSI 색상 모델로 변환하는 수식이다.

$$Y = 0.3R + 0.59G + 0.11B$$

$$U = (B - Y) \times 0.493 \quad (1)$$

$$V = (R - Y) \times 0.877$$

$$H = \cos^{-1} \left[\frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B) + (G - B)}} \right] \quad (2)$$

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)]$$

$$I = \frac{1}{3}(R + G + B)$$

RGB 색상 값에 식 (1)과 (2)를 적용하여 YUV와 HSI로 변환한 후, 사람의 피부색을 가장 잘 나타내는 영역을 선정하여 피부색을 추출한다. 그림 1은 RGB, YUV, HSI 색상 정보를 이용하여 추출한 피부 영역이다.

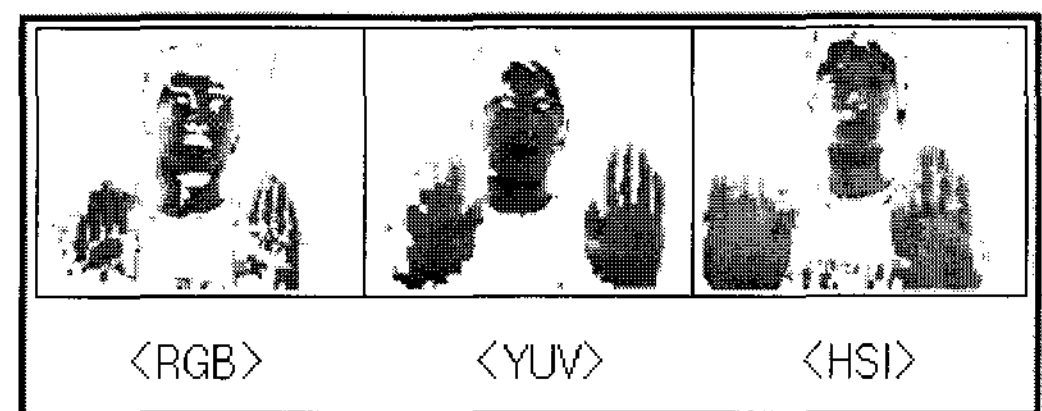


그림 1. 색상정보를 이용한 피부 영역 추출
Fig. 1. Extraction of skin regions using color information

2.2 손 영역 추적 및 추출

RGB, YUV, HSI 정보를 이용하여 추출된 피부색 영역들에 대해 레이블링(Labelling) 기법인 GrassFire 알고리즘을 적용하여 각 영역들을 레이블링 한 후, 100 픽셀(pixel) 이하인 영역들은 잡음으로 간주하여 제거한다. 잡음 제거 후, 상위에 있는 영역 중 가장 큰 영역을 얼굴 영역으로 선택하여 얼굴 영역을 중심으로 좌, 우를 구분한 후, 오른손과 왼손을 각각 구분하여 추적한다. 그림 2는 색상 정보를 이용하여 손 영역을 추적하는 과정이다.

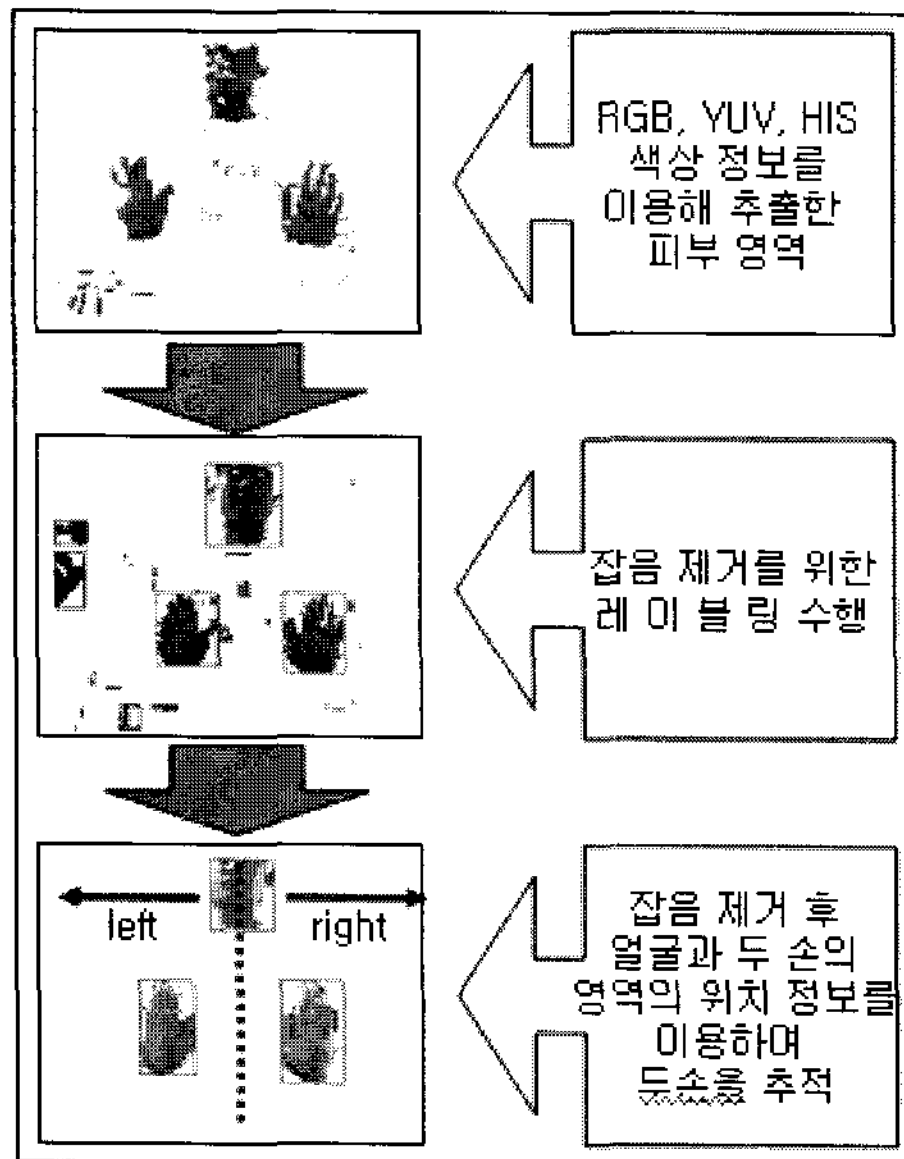


그림 2. 색상정보를 이용한 손 영역 추적
Fig. 2. Tracking of hand regions using color information

손 영역의 추적이 시작되면 손 영역의 추출을 한다. 손 영역 추출은 화상 카메라로 들어오는 영상에서 프레임들을 무작위로 선택하여 손 영역을 추출하는 것이 아니라 현재 프레임과 이전 프레임에서의 손 영역을 템플릿(template) 정합[4]을 이용하여 비교하며 두 영역의 오차가 5% 미만일 경우 현재 프레임에서의 손 영역을 추출하게 된다. 두 프레임간의 손 영역 오차는 식 (3)을 적용하여 계산한다.

$$r = \frac{a \cdot b}{|a| \cdot |b|} = \frac{\sum_{i=0}^M \sum_{j=0}^N a(i,j)b(i,j)}{\sqrt{\sum_{i=0}^M \sum_{j=0}^N a(i,j)^2 \sum_{i=0}^M \sum_{j=0}^N b(i,j)^2}} \quad (3)$$

식 (3)에서 i 와 j 는 각 영역의 가로와 세로의 크기를 나타내며, M 과 N 은 비교할 영상의 크기를 나타낸 것이다. 템플릿 정합을 하면 두 영상이 유사할수록 0에 가까워지고 차이가 커질수록 1에 가까워진다. 본 논문에서는 템플릿 정합을 통해 구해진 결과를 백분율로 환산하여 추출에 적용한다.

III. 수화 인식

본 논문에서 수화 인식은 지 숫자(finger number) 1부터 10까지를 대상으로 한다. 숫자 1~10을 나타내는 동작 10 개를 학습 및 인식하기 위해 본 논문에서는 신경망 알고리즘 중에 잡음과 훼손에 강한 ART2 알고리즘을 적용하여 지 숫자를 인식한다. 그림 3은 1~10의 지 숫자의 손 모양을 나타낸 것이다.

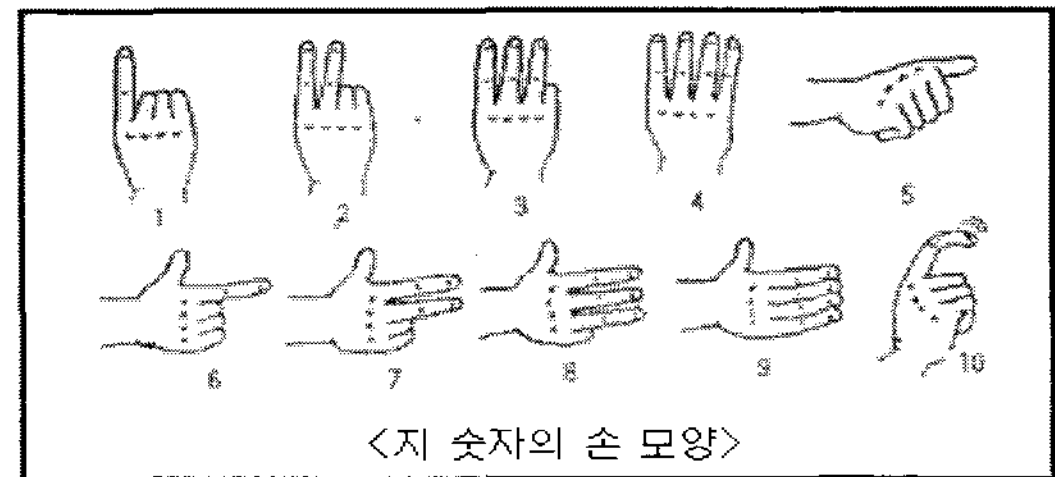


그림 3. 지 숫자 1~10
Fig. 3. Finger numbers from 1 to 10

3.1 ART2 알고리즘

ART2 신경망 모델[2][4]은 임의의 입력 패턴에 대해 이미 학습된 패턴을 잊지 않고 새로운 학습 패턴을 학습할 수 있는 안정성과 적응성을 가지면서 실시간 학습이 가능하여 저속 및 고속 학습을 지원할 뿐만 아니라 지역 최소화 문제가 발생하지 않는 장점을 갖는다. 따라서 ART2 알고리즘은 조명과 주위 환경에 의해 생기는 훼손과 잡음이 심한 화상 영상 인식에 효율적이다. 색상 정보를 이용하여 추출된 손 영역을 인식하기 위한 ART2 학습 알고리즘[5]은 다음과 같다.

Step 1. k 번째 입력 벡터를 x_k , 신경회로망의 i 번째 클러스터의 중심 벡터를 w_i 라 정한다.

Step 2. 새로운 입력 벡터 x_k 에 대해 최소 거리(minimum distance)를 가지는 클러스터 j^* 을 승자 클러스터로 선택한다. 일반적으로 입력 벡터와 클러스터 중심 벡터 사이의 거리는 유클리디안 거리로 계산한다.

$$\|x_k - w_{j^*}\| = \min \|x_k - w_i\| \quad (4)$$

Step 3. 입력 벡터에 대한 Vigilance Test를 수행한다. 만

약 입력 벡터와 승자 클러스터의 중심 벡터 사이의 거리가 반경 ρ (vigilance parameter) 이내에 들어오면 이 입력 패턴은 승자 클러스터와 유사한 패턴임을 의미하여 이 입력 벡터를 승자 클러스터에 포함시키고 그 클러스터의 중심 벡터를 수정한다.

$$w_{j*}^{t+1} = \frac{x_k + w_{j*}^t \times n_{j*}}{n_{j*} + 1} \quad (5)$$

여기서, n_{j*} 는 j 번째 클러스터에 포함된 입력 벡터의 개수이다. 만약 입력 벡터와 승자 클러스터의 중심 벡터 사이의 거리가 반경 ρ 보다 크면 이 입력 패턴은 기존의 클러스터와는 상이한 패턴임을 의미하며 이 입력 벡터로 새로운 클러스터를 생성한다.

Step 4. 모든 입력이 제시될 때까지 Step 1에서부터 Step 3의 과정을 반복 수행한다.

Step 5. 지정된 횟수의 학습을 반복 수행하거나 신경망의 클러스터 중심 벡터가 각각 변함이 없으면 학습을 종료한다. 그림 4는 학습에 사용된 1부터 10까지의 학습 패턴을 나타낸 것이다.

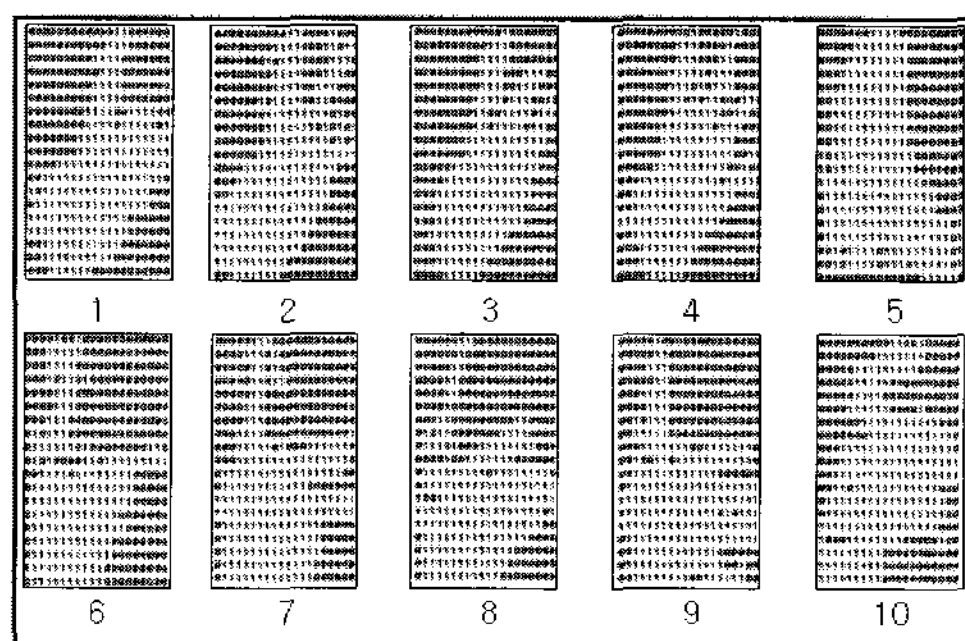


그림 4. 학습에 사용된 패턴
Fig. 4. Patterns used for learning

3.2 손 영역 인식

본 논문에서는 화상 카메라를 통해 들어오는 영상에서 색상 정보를 이용해 손 영역을 찾는 전 처리 과정이 있기 때문에 초당 24 프레임을 전부 사용하지 않고 초당 10 프레임의 영상을 사용한다. 24 프레임 영상을 사용할 경우 부드러운 영상을 얻을 수는 있지만 정확한 전처리 과정을 거치기 위해서는 10 프레임을 사용하는 것이 효

율적이다. 손 영역의 인식 과정은 전 처리를 통해서 구해진 손 영역에 대해 현재 프레임에서의 손 영역과 이전 프레임에서의 손 영역을 템플릿 매칭을 통해 비교하여 오차가 5% 미만일 경우에만 인식을 수행한다. 화상 카메라로 획득되는 영상은 정지 영상이 아니기 때문에 매 프레임마다 조금씩 다른 영상이 들어오므로 현재 프레임의 손 영역과 이전 프레임의 손 영역을 비교하여 인식함으로써 인식률의 정확도를 높일 수 있다. 그림 5는 손 영역 인식 과정을 나타낸 것이다.

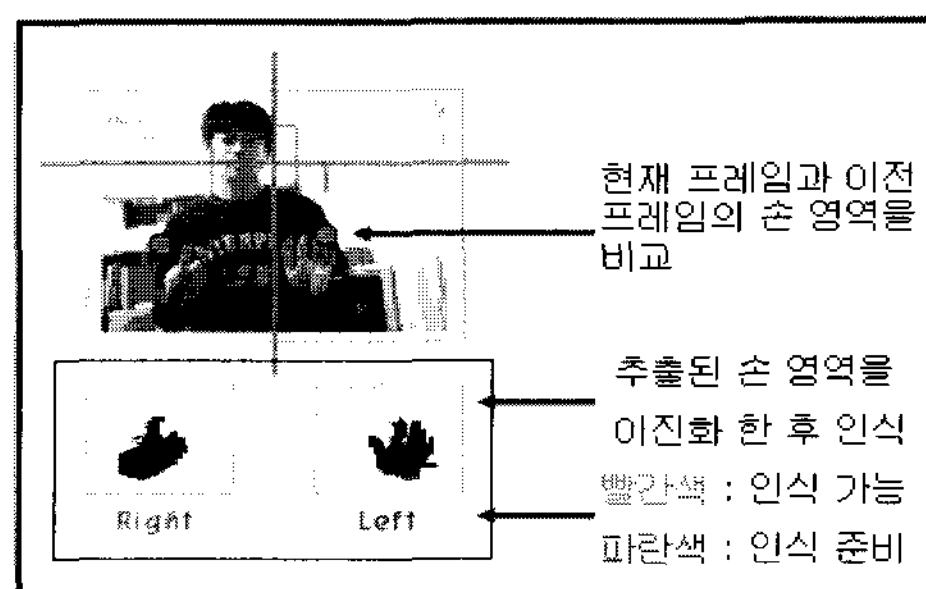


그림 5. 손 영역 인식 과정
Fig. 5. Recognition process of hand regions

IV. 실험 및 결과 분석

본 논문에서는 제안한 ART2를 이용한 수화 인식의 성능 평가를 위해 블루맥스 DUO CAP 캠(Cam)을 사용하였으며, 실험 환경은 Intel Pentium-IV 2GHz 환경에서 VC++ 6.0으로 구현하였다. 1~10의 10개의 지 숫자를 인식하기 위해 각 숫자에 대해 10개의 패턴을 선택하여 학습에 사용하였으며 인식 실험은 640X480 해상도에 초당 프레임의 수를 10으로 설정하여 1~10의 지 숫자를 각각 1번씩 반복해서 사용자 10명을 대상으로 인식 실험을 하였다. 표 1은 지 숫자 인식 결과를 나타내었고 그림 6은 제안된 수화 인식의 결과 화면을 나타내었다.

표 1. 지 숫자 인식 결과
Table 1. Results of finger number recognition

지 숫자	1	2	3	4	5
인식 개수	10/10	10/10	10/10	10/10	10/10
지 숫자	6	7	8	9	10
인식 개수	10/10	10/10	10/10	10/10	10/10

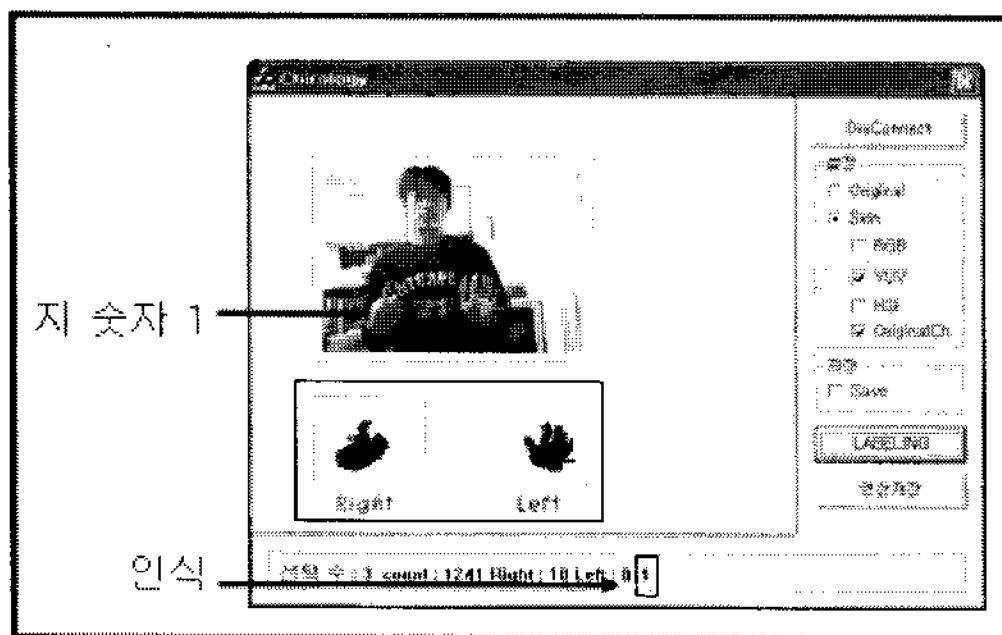


그림 6. 인식 화면

Fig. 6. A sample screen of recognition results

V. 결 론

인터넷의 발전으로 인해 화상 채팅과 화상 통화 서비스가 많이 활성화 되어있지만 청각 장애인과 비장애인의 의사소통을 위한 서비스는 아직 미흡하다. 따라서 본 논문에서는 청각 장애인과 비장애인 간의 의사소통을 돕기 위해 ART2를 이용한 수화 인식 방법을 제안하였다. 본 논문에서는 화상 카메라를 통해 들어오는 영상에서 RGB, YUV, HSI 색상 정보를 각각 이용하여 피부 영역을 추출하였다. 추출된 피부 영역에서 잡음으로 간주되는 영역을 제거한 후, 두 손 영역을 추적 및 추출하였다. 추출된 손 영역의 인식은 훼손과 잡음에 강한 ART2 알고리즘을 적용하였다.

향후 연구 과제는 청각 장애인과 비장애인의 원활한 의사소통을 돕기 위해 모든 지 숫자와 한글지화 인식에 대해 연구할 것이다.

참고문헌

- [1] 석동일, 한국수화의 언어학적 분석, 박사학위논문, 대구대학교 대학원, 1989.
- [2] Gail A Carpenter and Stephen Grossberg, "ART2: Self-Organization of Stable Category Recognition Codes for Analog Input patterns," *Applied Optics*, Vol.26, No.23, pp.4919-4930, 1987.
- [3] 문규형, 최유주, "조명 변화에 강건한 피부색 영역 검출을 위한 혼합 색상 모델," *한국정보과학회논문지*, 제33권, 제2호, pp.98-101, 2006.

- [4] Milan Sonka, Vaclav Hlavac, and Roger Boyle, *Image Processing, Analysis, and Machine Vision*, CENGAGE-Engineering, 3 edition, 2007.
- [5] K. B. Kim, M. Kim, and Y. W. Woo, "Recognition of Shipping Container Identifiers Using ART2-Based Quantization and a Refined RBF Network," *Lecture Notes in Computer Science*, LNCS 4432, pp.572-581, 2007.

저자 소개



김광백(Kwang-Baek Kim)

1999년 부산대학교 전자계산학과
(이학박사)
1997년~현재 : 신라대학교
컴퓨터정보공학부 교수

2005년~ 현재 : 한국멀티미디어학회 이사 및 논문지 편집위원

2005년~ 현재 : 한국해양정보통신학회 이사 및 논문지 편집부위원장

※ 관심분야 : Image Processing, Fuzzy Logic, Neural Networks, Medical Imaging and Biomedical System, Support Vector Machines



우영운(Young Woon Woo)

1989년 2월 연세대학교 전자공학과
(공학사)

1991년 8월 연세대학교 본대학원
전자공학과(공학석사)

1997년 8월 연세대학교 본대학원 전자공학과(공학박사)

1997년 9월 ~ 현재 : 동의대학교 멀티미디어공학과 교수

2007년 ~ 현재 : 한국해양정보통신학회 국제이사

※ 관심분야 : 인공지능, 패턴인식, 퍼지이론, 의료정보