색상 정보와 퍼지 클러스터링 알고리즘을 이용한 실시간 수화 인식

강효주* · 이동균* · 김광백*

*신라대학교 컴퓨터정보공학부

Real Time Recognition of Finger-Language Using Color Information and Fuzzy Clustering Algorithm

Hyo-joo Kang* · Dong-gyun Lee* · Kwang-baek Kim*

* Division of Computer and Information Engineering, Silla University
E-mail : mc_hyo@hanmail.net, finht@dremwiz.com, gbkim@silla.ac.kr

요 약

사람의 손동작은 오랫동안 하나의 언어역할을 하는 통신 수단으로 사용되어 왔다. 이러한 손동작 중에서 가장 체계를 갖춘 수화는 청각장애인이 일반인과 일상 대화를 할 수 있도록 도와주는 주요한 통신 수단이다. 하지만 건청인들의 대부분이 습득하고 있지 않아 청각장애인들과 의사소통이 거의 불가능 한 것이 현실이다. 따라서 본 논문에서는 건청인과 청각장애인들 간의 의사소통을 원활하게 하기 위해 색상 정보와 퍼지 클러스터링 알고리즘을 이용한 실시간 수화 인식 방법을 제안한다. 제안된 방법은 화상 카메라를 통해 얻어진 실시간 영상에서 YCbCr 컬러 공간에서 색차 정보에 해당하는 Cb, Cr 정보를 각각 추출한 후, 이진화한 영상과 원본 영상에서 마스크를 통한 에지를 추출한 이진화 영상에 대해 논리연산을 통해 두 손의 위치와 외곽을 추출한다. 추출된 각 정보를 조합하여 8 방향 윤곽선 추적 알고리즘을 적용하여 객체의 위치를 추적한다. 그리고 추적한 객체의 영역에 대해 형태학적 정보를 이용하여 잡음을 제거한 후, 최종적으로 두 손의 영역을 추출한다. 추출된 손의 영역은 퍼지 클러스터링 기법 중의 FCM 알고리즘을 적용하여 수화의 특징들을 분류하고 인식한다. 제안된 방법의 성능을 평가하기 위해 화상카메라를 통해 얻어진 실시간 영상을 대상으로 실험한 결과, 제안된 방법이 두 손 영역의 추출에 효과적이고 수화 인식에 있어서 가능성을 확인하였다.

키워드

수화, 화상카메라, YCbCr, FCM 알고리즘

I. 서 론

수화는 청각장애인과 건청인간의 원활한 의사소통을 돕기 위하여 고안된 언어이다[1]. 수화는 양손의 모양과 위치 움직임으로 구성되고 언어적 체계가 갖추어져 있는 몸짓 언어로써 인간의 손을 이용한 여러 가지 신호가 상호간의 정보를 교환하기 위한 수단으로 사용되어 왔다. 특히, 청각 장애인들은 손동작을 이용한 수화를 사용하여 대다수의 의사소통을 하고 있다. 하지만 건청인들이 대부분은 수화를 습득하고 있지 않아서 청각 장애인들과 의사소통이 거의 불가능 한 것이 현실이다.

기존의 손 영역 추출방법에서는 HSI, YCbCr 컬러 정보만을 이용하여 손의 후보영역을 찾으므로 인해서 손의 후보 영역들을 뚜렷하게 추출하지 못하는 단점이 있다[2]. 또 다른 방법으로 영상의 거리 값을 이용하여 손의 영역을 구분하게 되는데, 영상에서 손의 영역이 조금만 멀리 있게 될 경우, 손의 영역을 배경으로 인식

하여 제외시키는 단점이 있다[3]. 이러한 단점을 보완하고자 본 논문에서는 건청인들과 청각장애인들 간에 원활한 의사소통을 하기 위한 전단계로 색상 정보와 FCM 알고리즘을 이용한 실시간 수화 인식 방법을 제안한다. 제안된 실시간 수화 인식 방법은 화상 카메라를 통해서 실시간으로 획득한 영상에서 손 영역을 추출하기 위해 피부색상 특징, 손의 위치와 크기정보를 이용하여 추출하며, 추출된 손 영역은 FCM 알고리즘을 적용하여 수화의 특징을 분류하고 인식한다.

II. 컬러 공간과 캐니마스크 이용한 손 영역 검출

손 영역 검출 단계의 목적은 YCbCr 컬러 공간을 이용하여 피부 영역을 검출한 후 실험을 통해 얻은 임계치를 이용하여 이진화 한 영상과 원본 영상에 캐니 마스크를 이용하여 에지를 추출 한 영상에 대해 논리 연산을 적용하여 객체의 외곽을 추출한 후 형태학적 특징을 이용하여 손 영역을 추출한다.

2.1 컬러 공간과 마스크 이용한 피부 색상 분할

본 논문에서는 화상 카메라로 획득된 영상의 피부 영역을 검출하기 위하여 피부색 영역의 픽셀과 피부색영역 외의 픽셀로 구분한다. 피부색상 분할은 빛에 민감하게 반응하므로 적당한 컬러 공간을 사용하는 것이 중요하다. 컬러공간으로는 RGB, HSV, YCbCr 공간이 있으나 본 논문에서 빛과 색을 분리하여 빛의 영향를 최소화하는 YCbCr 컬러공간을 이용한다.

YCbCr 컬러공간은 RGB 컬러공간에서 광도(Y)와 푸른 정보(Cb), 붉은 정보(Cr)를 분리한 컬러공간으로 RGB 컬러 공간에서 YCbCr로 변환하는 식(1)과 같이 계산한다[4,5].

$$\begin{aligned} Y &= 0.3R + 0.59G + 0.11B \\ Cb &= 0.6R - 0.28G - 0.32B + 128 \\ Cr &= 0.21R - 0.52G + 0.31B + 128 \end{aligned} \quad (1)$$

그림 1은 YCbCr 컬러공간을 이용하여 추출한 피부색 영역과 그 영상을 실험을 통하여 얻은 임계치를 이용한 이진화 한 영상이다. 그림 1(a)의 YCbCr 컬러 공간 영상은 피부색과 비슷한 색을 가진 손 후보영역들이 같이 추출되는 결과를 보인다. 그림 1(b)에서와 같이 YCbCr 컬러공간을 이용하여 이진화 한 영상은 비슷한 색을 가진 손 후보영역들이 합쳐지는 결과를 보인다[6,7]. 따라서 본 논문에서는 원 영상에서 캐니마스크를 이용하여 객체의 외곽선을 추출한 영상과 YCbCr 컬러 공간을 이용하여 이진화 한 영상과의 논리 연산을 통해 비슷한 색을 가진 손 후보영역 들을 분리한다. 그림 1의 (d)는 논리 연산을 한 결과 영상이다.

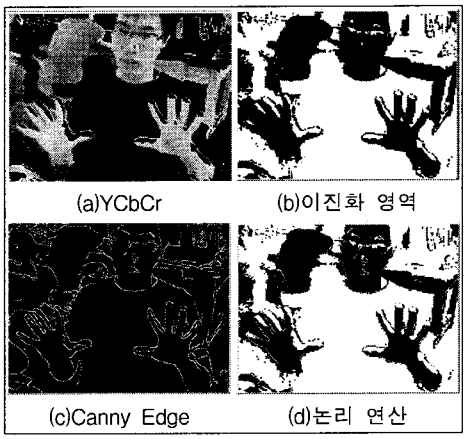
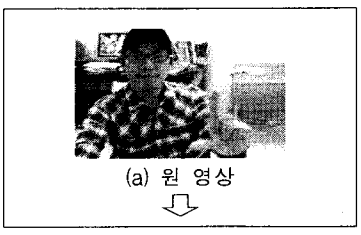


그림 1. 컬러공간과 캐니 마스크를 이용한 피부색 영역 검출

2.2 손 영역 추적 및 추출

YCbCr 컬러 공간을 이용하여 이진화 한 영상과 캐니마스크를 적용하여 에지를 추출한 영상을 이용하여 추출된 손의 후보 영역에서 8방향 라벨링 추적 알고리즘을 적용한 후, 손 후보 영역을 개별 정보화 시킨다. 정보화 된 객체들에 크기를 비교하여 가장 큰 객체 두 개를 얼굴과 손 영역으로 간주하고 객체의 위치 정보를 이용하여 손 영역을 추출한다. 그림 2는 컬러공간과 형태학적 특징을 이용하여 원 영상에서 손 영역을 추출한 과정을 나타내었다.



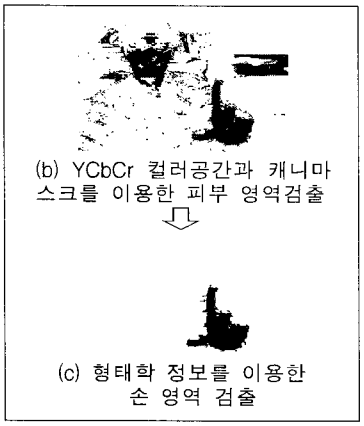


그림 2. 손 영역 추출과정

III. FCM 알고리즘을 이용한 수화 인식

본 논문에서는 수화 동작 중 지화와 지숫자를 인식한다. 수화의 지화는 기역(ㄱ)에서 히읗(ㅎ)까지의 자음을 수화로 나타낸 동작이고, 지숫자는 숫자를 수화로 나타낸 동작이다. 본 논문에서는 자음 기역(ㄱ)에서 히읗(ㅎ)까지와 숫자 일(1)에서 십(10)까지의 손 모양을 인식한다. 그림 3은 자음 기역(ㄱ)에서 히읗(ㅎ)까지 14개의 지화 손 모양과 일(1)에서 십(10)까지 10개의 지숫자 손 모양을 나타낸 것이다.

그림 3. 지화의 손 모양

FCM 알고리즘은 하나의 클러스터에 속해져 있는 각각의 데이터들을 클러스터에 대한 데이터의 소속정도에 따라 분류하는 알고리즘이다. n개의 벡터 집합을 c개의 퍼지그룹으로 분리하고 각 벡터들은 소속정도가 최대가 되는 그룹으로 분류되며 각각의 그룹 안에서 클러스터의 중심을 찾는다[8].

FCM 알고리즘은 0과 1사이의 소속정도에 의해서 나타난 소속 정도를 주어진 데이터 점이 몇 개의 그룹에 속할 수 있다는 퍼지분할을 사용한다. 즉, 퍼지 분할을 적용하기 위해서 소속함수 U는 0과 1사이 값을 가지는 요소들을 가지며 데이터 집합에 대한 소속감 정도의 합은 식 2와 같이 항상 1이다[2].

$$\sum_{i-1}^{c} u_{ik} = 1, \nabla k = 1, \cdots, n \qquad (2)$$
$$0 < \sum_{k=1}^{n} u_{ik < n}$$

FCM 알고리즘의 목적함수(J)는 식 3과 같다.

$$J(u_{ik}, v_i) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m}(d_{ik})^2 \qquad (3)$$
$$d_{ik} = d(x_k - v_i) = [\sum_{j=1}^{1}(x_{kj} - v_{ij})^2]^{1/2}$$
$$v_i = v_{i1}, v_{i2}, \ldots, v_{ij}, \ldots, v_{iL}$$

여기서, x_k는 입력 데이터이고 u_{ik}는 x_k가 클러스터 i에 속한 정도, 그리고 v_i는 i번째의 클러스터 중심이다. FCM 알고리즘은 다음과 같다.

[단계 1] 클러스터 개수 $c(2 \le c \le n)$와 초기 소속 함수 $U^{(0)}$를 초기화한다. 지수의 가중치 m을 결정하고, 알고리즘의 반복횟수 r을 표시한다.

[단계 2] 데이터 x와 소속 함수 $U^{(0)}$를 이용하여 퍼지 클러스터의 중심 $\{v_i | i = 1, 2, \ldots, c\}$을 식 4와 같이 계산한다.

$$v_{ij} = \frac{\sum_{k=1}^{n}(u_{ik})^m x_{kj}}{\sum_{k=1}^{n}(u_{ik})^m} \qquad (4)$$

[단계 3] 단계 2에서 계산된 퍼지 클러스터의 중심값과 데이터 x와의 거리값 d에 의하여 새로운 소속 함수 $U^{(r+1)}$을 식 5와 같이 계산한다.

$$u_{ik}^{(r+1)} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ik}^{r}}{d_{jk}^{r}}\right)^{2/m-1}} \quad \text{for } I_k = 0 \qquad (5)$$

또는 $u_{ik}^{(r+1)} = 0$ for all classes i

여기서, $i \in \overline{I_k}$, $I_k = \{i | 2 \le c < n; d_{ik}^{(r)} = 0\}$ 이고 $\overline{I_k} = 1,2,...,c$이다. 그리고 $\sum_{i \in I_k} u_{ik}^{(r+1)} = 1$ 이다. 또한, d_{jk}는 j번째 데이터 x_j와 k번째 퍼지 클러스터의 중심값 v_k와의 유클리디안 거리값을 의미하며 식 6과 같이 계산한다.

$$d_{jk} = d(x_j, c_k) = \left[\sum_{i=1}^{p}(x_{ji} - v_{ki})^2\right]^{-1} \qquad (6)$$

[단계 4] 다음과 같이 계산한 후, 만약 식 7에서 $\Delta > \varepsilon$이면 알고리즘을 종료하고, 그렇지 않으면 $r = r+1$로 정하고 [단계 2]로 가서 반복 수행한다.(ε : 임계값)

$$\Delta = \| U^{(r+1)} - U^{(r)} \| = \max_{i,k} |u_{ik}^{(r+1)} - u_{ik}^{(r)}| \qquad (7)$$

지화와 지숫자 인식을 위해 적용된 FCM 알고리즘은 그림 4와 같다.

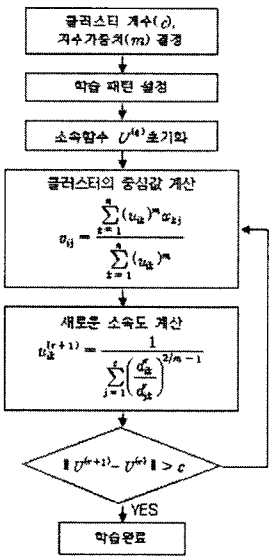


그림 4. FCM 알고리즘

Ⅳ. 실험 및 결과 분석

실험 환경은 Intel Pentium-Ⅳ 3GHz CPU와 1GB RAM이 장착된 IBM 호환 PC상에서 VC++ 6.0으로 구현하였고, 화상 카메라는 Microsoft사의 NX-6000모델을 사용하였다. 화상 카메라에서 초당 15 프레임의 320×240 크기를 가진 영상을 자음 14개와 지숫자 10개를 각각 5회씩 인식하여 성능 평가에 적용하였다.

본 논문에서 제안한 화상 카메라에서 처리된 실시간 영상에서의 손 영역 추출 및 지화 패턴 인식 방법을 적용한 결과, 표 1과 같이 지화 영상 60장, 지숫자 영상 50장의 손 영역이 모두 추출되었다. 표 2는 FCM 알고리즘 학습에 사용된 지수가중치와 임계값을 나타내었다.

표 1. 화상카메라에서 손 영역 추출 결과

	제안된 방법 (추출개수 / 영상개수)
지화(자음)	60 / 60
지숫자	50 / 50

표 2. 학습에 사용된 지수가중치와 임계값

m	ε
100	0.01

지화와 지숫자 패턴 인식을 위해 적용된 FCM 알고리즘의 학습 및 인식 성능을 평가하기 위해 화상 카메라로부터 실시간 획득한 지화 영상 60장과 지숫자 영상 50장에 대해 추출된 지화 손 영역 60개, 지숫자 손 영역 50개를 학습 패턴으로 적용하였다. FCM 알고리즘으로 지화와 지숫자 패턴을 인식한 결과는 각각 표 3, 표 4로 나타내었다.

표 3. 지화 패턴 인식 결과
(인식패턴개수 / 패턴개수)

	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ
인식	5/5	3/5	5/5	4/5	4/5	5/5	5/5

	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
인식	4/5	5/5	5/5	5/5	3/5	3/5	4/5

표 4. 지숫자 패턴 인식 결과
(인식패턴개수 / 패턴개수)

	1	2	3	4	5
인식	5/5	5/5	4/5	4/5	4/5

	6	7	8	9	10
인식	3/5	5/5	5/5	5/5	4/5

표 3과 표 4에서 알 수 있듯이 지화 패턴은 60개 중 50개가 인식되었고, 지숫자 패턴은 50개 중 44개가 인식되었다. FCM 알고리즘으로 지화와 지숫자 패턴 인식에 실패한 경우는 자

음 니은(ㄴ)과 지숫자 육(6)의 손 모양이 같아 같은 패턴으로 인식한 경우, 자음 리을(ㄹ)과 자음 티읕(ㅌ), 자음 미음(ㅁ)과 자음 피읖(ㅍ), 지숫자 삼(3)과 지숫자 사(4)등과 같이 지화와 지숫자의 손 모양이 비슷하여 같은 그룹으로 클러스터링 된 경우, 실시간 영상의 획득과정에서 빛의 반사로 인해 영상이 손실되어 올바른 학습패턴이 적용되지 못한 경우이다. 그림 5는 지화와 지숫자 패턴 인식에 실패한 몇 개의 패턴들을 나타내었다.

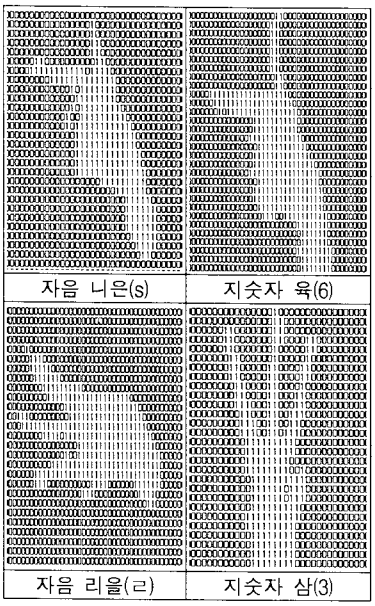


그림 5. 지화 패턴 인식 실패 영상

VI. 결 론

본 논문에서는 FCM 알고리즘을 이용한 실시간 수화 인식 방법을 제안하였다. 화상 카메라에서 실시간으로 획득한 영상을 YCbCr 컬러 공간 중 푸른 정보(Cb), 붉은 정보(Cr)와 캐니 마스크 적용한 후, 추출된 에지 정보를 이용하여 피부색 영역을 추출하였다. 추출된 피부색 영역은 8 방향 라벨링 추적 알고리즘을 적용하여 손 후보 영역을 추출한 후, 추출된 객체의 형태학적 정보를 이용하여 가장 큰 두개의 객체를 선택하고, 선택된 객체들의 위치 정보를 이용하여 최종 손 영역을 추출하였다. 추출된 손 영역을 FCM 알고리즘을 적용하여 손의 특징을 분류한 후, 지화와 지숫자 인식에 적용하였다. 제안된 방법으로 화상카메라에서 실시간으로 처리된 영상에 적용한 결과, 지화와 지숫자의 손 영역이 모두 추출되었다. 추출된 지화와 지숫자 손 영역은 FCM 알고리즘을 적용하여 수화 인식에 적용한 결과, 지화 패턴은 50개, 지숫자 패턴은 44개 각각 인식에 성공하였다. 지화와 지숫자 인식에 실패한 패턴은 손 모양이 거의 흡사하거나 비슷하여 각각의 특징이 정확히 분류되지 않거나, 빛의 반사로 인해 영상이 손실되어 올바른 학습패턴이 적용되지 못한 경우이다.

따라서 향후 연구과제로는 실시간 수화 인식에 있어 보다 처리시간을 단축시키고 빛의 영향을 효과적으로 처리 할 수 있는 방법을 연구하여 지화와 지숫자를 더욱 정확하게 인식할 것이다.

참고문헌

[1] 강옥, 가장 아름다운 언어 수화, 혜성출판사, 2001.

[2] 김광백, 우영운, "FCM 알고리즘을 이용한 지화," 한국해양정보통신학회논문지, 12권, 6호, pp.1101-1106, 2008.

[3] Javed, K. Shaque and M. Shah," A Hierarchical Approach to Robust Background Subtraction using Color and Gradient Information," Proceedings of the Workshop on Motion and Video Computing , pp.22-27, 2002.

[4] Maria P., Panagiota B., Image Processing, John Wiley & Sos Ltd., 1999.

[5] 이준우, 송근원 "피부색 정보를 이용한 적응적 얼굴영역 추출" 한국멀티미디어학회 춘계학술발표논문집, 제 6권 1호, pp.359-361, 2003.

[6] 박호식, 배철수, "수화 인식을 위한 얼굴과 손 추적 알고리즘", 한국통신학회, 한국통신학회논문지, 제31권 제11C호, pp.1071-1076, 2006.

[7] 조오영, 김형곤, 고성제, 안상철, "인터액티브 가상 환경을 위한 손 제스처 인식 시스템", 대한전자공학회, 전자공학회논문지-S 전자공학회논문지 제36권 S편 제4호, pp. 70 ~ 82, 1999.

[8] Robert B., Fuzzy Modeling for Control, Kluwer Academic Publishers, 1998.