

SOFM 신경망을 이용한 수화 형상 인식

박경우[†]

Sign Language Shape Recognition Using SOFM Neural Network

Kyung-Woo Park[†]

Abstract

인간은 정보전달을 위하여 언어 이외에 동작, 표정과 같은 비언어적인 수단을 이용한다. 이러한 비언어적인 수단을 정확히 분석 할 수 있다면 인간과 컴퓨터간의 자연스럽고 지적인 인터페이스를 구축할 수 있게 된다. 본 논문은 별도의 센서를 부착하지 않은 단일 카메라 환경에서 손 형상을 입력정보로 사용하여 손 영역만을 분할한 후 자기 조직화 특징 지도(SOFM: Self Organized Feature Map) 신경망 알고리즘을 이용하여 손 형상을 인식함으로써 수화인식을 위한 보다 안정적이며 강인한 인식 시스템을 구현하고자 한다. 제안 방법으로는 피부색 정보를 이용하여 배경으로부터 손 영역만을 추출한 후 추출된 손 영역의 형상을 인식한다(전처리과정으로 모델이미지의 사이즈와 압축 및 컬러에 대한 정보를 정규화 시켰다). 또한 인식 효율을 높이기 위해 SOFM 신경망 알고리즘을 적용함으로써 보다 안정적으로 손 형상을 인식할 수 있게 되었으며, 손 형상 인식률에 대한 안전성과 정확성을 향상시킬 수 있었다. 그리고 인식된 손 형상의 의미를 텍스트로 보여줌으로서 사용자의 의사를 정확하게 전달할 수 있다.

Key words : Hand Recognition, SOFM, Sign Language Recognition, Neural Network

1. 서 론

최근 사용자 친화형 유저 인터페이스 구현 환경이 대두되면서 인간과 컴퓨터간의 원활한 상호작용의 필요성이 커짐에 따라 인간의 행동을 분석하는 연구가 많이 수행되고 있다.^[1] 인간은 정보전달을 위하여 언어 이외에 동작, 표정과 같은 비언어적인 수단을 이용한다. 인간과 의사소통이 자유스럽고 보다 지적인 시스템을 제작하기 위해서는 무엇보다도 이러한 비언어적 대화 수단을 자동화함이 필요하다.^[2,3] 그 중에서도 손은 인간에게 있어서 가장 중요한 도구의 하나로 그 동작과 모양이 의미하는 바를 인식하는 것은 정보 시스템과 인간 사이의 자연스러운 인터페이스를 실현할 수 있다는 점에서 오랜 기간 동안 많은 사람들의 연구테마로 다루어져 왔다. 그러나 복잡한 환경으로부터 사람의 손 형상만을 정확히 추출하고 인식하여 의미 있는 입력정보로 사용하는 것은 어려운 일이다. 기본적으로 정확한

손 영역 분할을 위한 알고리즘이 필수이며 손 형상의 분석과 인식을 위한 빠른 처리 속도 또한 고려해야할 요소이다.

본 논문은 별도의 센서^[4]를 부착하지 않은 단일 카메라 환경에서 손 형상을 입력정보로 사용하여 손 영역만을 분할한 후 SOFM(Self Organized Feature Map) 신경망 알고리즘을 이용하여 손 형상을 인식함으로써 수화인식을 위한 보다 안정적이며 강인한 인식 시스템을 구현하고자 한다. 이는 비장애인과 장애인간 상호 편리하게 대화할 수 있는 인터페이스를 제공하려 하며, 제안하는 방법으로는 카메라로 들어오는 영상을 입력 영상으로 하여 피부색 정보를 이용해 배경으로부터 손 영역만을 추출한 후 추출된 손 영역의 형상을 인식한다. 또한 인식 효율을 높이기 위해 SOFM 신경망^[5] 알고리즘을 적용함으로써 보다 안정적으로 손 형상을 인식할 수 있게 되었으며, 손 형상 인식률에 대한 안전성과 정확성을 향상시킬 수 있었다. 그리고 인식된 손 형상의 의미를 텍스트로 보여줌으로서 사용자의 의사를 정확하게 전달할 수 있다.

광주보건대학 병원전산과 (Dept. of Hospital Information Management Gwangju Health College)

[†]Corresponding author: kwpark@ghc.ac.kr

(Received : January 28, 2010, Accepted : March 20, 2010)

2. 피부색 정보를 이용한 손 영역 검출

칼라 공간에서 사람의 살색 분포는 자연계의 다른 물체와 구분할 수 있는 뚜렷한 특징을 가지고 있으므로 색차 성분을 이용하면 쉽게 살색 영역을 추출할 수 있다.^[6]

2.1. Skin Color Filtering

본 논문에서는 칼라 정보만을 이용하여 손 영역을 검출하기 위해 다음과 같이 제안한다. 이 방법은 간단하고 빠르게 RGB공간 안에서 피부색을 추출 할 수 있다. RGB 데이터의 영상이 입력되면 피부색은 RED 값이 다른 GREEN이나 BLUE보다 크다고 가정했다. (식 1)과 같이 피부색에 해당하는 부분은 그 픽셀을 1로 할당하여 주고 이 이외에 해당하는 픽셀의 위치에서는 살색이 아닌 영역으로 판정하여 0의 값으로 할당을 하게 된다.

$$I_z(x,y) = \sum_{x,y} \begin{cases} 1 (RED \geq BLUE) (RED \geq GREEN) \\ 0 \text{ Otherwise} \end{cases} \quad (\text{식 1})$$

2.2. 손 영역 검출

입력 영상에서 피부색 정보를 이용하여 손 영역을 추출할 때 칼라 필터를 사용하더라도 손만을 정확히 추출하기란 쉽지 않다. 본 연구는 칼라 필터가 적용되고 남아 있는 피부색 정보의 픽셀들은 대부분 손에 속



그림 1. 살색정보를 이용한 손 영역 추출.
Fig. 1. Hand area detection using skin color.



(a)입력 손 영상 (b)손 영역 분할 (c) 손 영역 추출

그림 2. 입력 영상의 손 영역 추출.
Fig. 2. Hand area detection in input image.

해 있기 때문에 횡 좌표 X축으로 피부색 픽셀들의 수를 계산하여 배경 영역으로부터 손 영역만을 분리한다 (labeling). (그림 2)는 입력 영상에서의 손 영역을 추출한 것이다.

3. 신경망을 이용한 손 형상 인식

본 논문에서는 손 형상을 정확히 분석하고 인식하기 위해서 자기 조직화 특징 지도(SOFM:Self Organized Feature Map)란 신경망 알고리즘^[7]을 사용하였다. 본 논문에서 제안한 알고리즘 구성은 (그림 3)과 같다.

3.1. 제안한 신경망 알고리즘

본 논문에서는 손 형상을 정확히 분석하고 인식하기 위해서 SOFM 신경망 알고리즘을 사용한다. SOFM은 입력뉴런들, 출력뉴런들, 그리고 입·출력 뉴런 간에 연결된 시냅스들로 구성되어 있으며 학습단계에서 다차원 공간상의 특징 벡터로 표현되는 객체들이 경쟁층에서 스스로 유사한 패턴끼리 2차원의 특징 지도를 조직화하여 영역 지도를 형성한다. 그리고 인식단계에서는 이미 학습단계에서 훈련된 연결 가중치 하에서 미지의 특징벡터에 대하여 경쟁층에서 반응이 일어나는 위치를 통하여 해당 클래스를 인식하게 된다. 또한, 입력패턴이 주어졌을 때 출력 뉴런들은 경쟁을 한 후 단 한 개의 출력 뉴런만이 승리하게 된다.

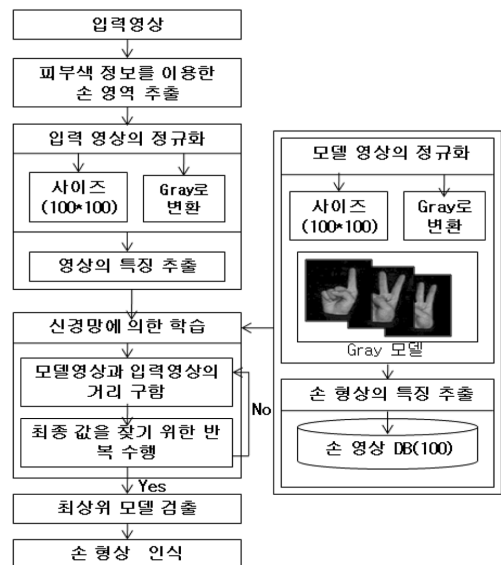


그림 3. 구성도.
Fig. 3. System structure.

3.2. SOFM의 학습과정

SOFM의 학습과정은 경쟁과정, 협동과정, 적응과정의 세가지 과정으로 이루어진다.

3.2.1. 경쟁(Competitive)과정

SOFM에서 경쟁층 상에 있는 각 뉴런은 입력 공간을 구성하는 N차원의 특징 벡터와 연결 강도에 의하여 연결된다. 연결 강도는 0에서 1사이(0~1)로 정규화된 임의의 값으로 초기화되어 있다. 학습 단계에서 어떤 입력 패턴이 주어지면 모든 뉴런의 연결 강도와와의 거리를 계산하여 거리가 최소가 되는 뉴런이 승자가 되는 경쟁의 규칙을 가진다(그림 4).

3.2.2. 협동(Cooperation)과정

협동과정에서는 경쟁 과정에서의 승자 뉴런과 인접한 이웃 뉴런들만이 제시된 입력 벡터에 대하여 학습이 허용된다. 즉, 유사한 특징 패턴에 민감하게 경쟁층에서 지도를 형성하기 위해 승자 뉴런은 이웃한 뉴런

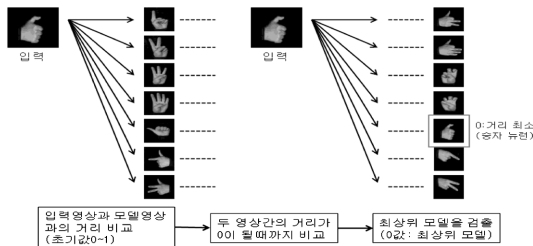


그림 4. 경쟁과정.
Fig. 4. Competitive course.

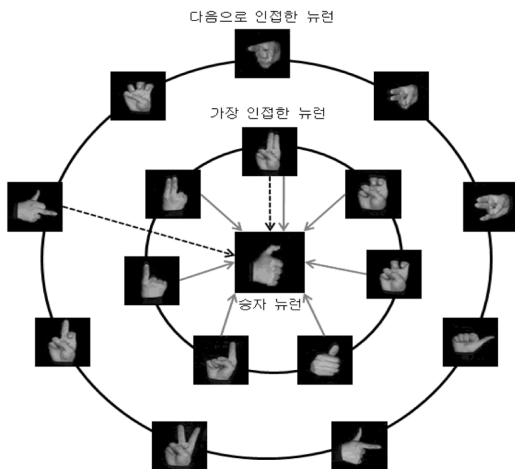


그림 5. 협동과정.
Fig. 5. Cooperation course.

을 정해진 함수에 의하여 결정하고 그 뉴런의 연결 강도를 갱신한다. 승자만이 출력을 낼 수 있으며, 승자와 그의 이웃들만이 연결 강도를 조절할 수 있도록 한다(그림 5).

3.2.3. 적응(Adaptation)과정

적응과정에서는 승자 뉴런과 그와 이웃한 뉴런들이 특정한 입력 값에 더욱 민감해지도록 활성 함수를 적용시켜 연결 강도를 갱신한다. 이 과정을 통하여 승자 뉴런과 인접해 있는 뉴런들은 멀리 떨어져 있는 뉴런보다 더 강하게 적응이 된다(그림 6).

3.3. SOFM의 학습 3단계

1단계: 연결강도 초기화

N개의 입력으로부터 M개의 출력 뉴런 사이의 연결 강도를 작은 값의 임의의 수로 초기화한다. 초기의 이웃 반경은 모든 뉴런들이 포함될 수 있도록 충분히 크게 잡았다가 함수 $h_{ik}(t)$ 에 따라 점차 줄어든다.

2단계: 입력벡터가 제시되고 경쟁

입력 벡터와 모든 뉴런들 간의 거리를 계산한다. 입력과 출력 뉴런 K사이의 거리는 (식2)와 같이 계산되고 그 중에서 가장 최소값이 선택된다. (그림 7)은 (그림 4)의 입력 영상과 모델 영상사이의 거리를 계산하여 가장 최소값(0)이 되는 모델(최상위 모델)을 찾는 것이다.

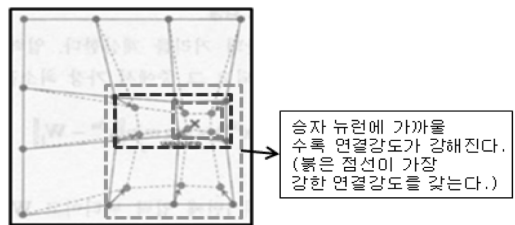


그림 6. 적응과정.
Fig. 6. Adaptation course.

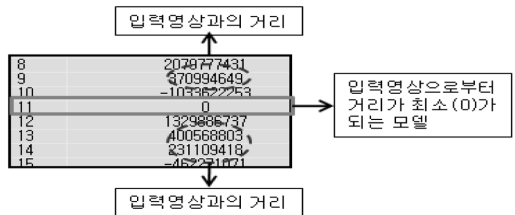


그림 7. 입력영상과 모델영상의 거리.
Fig. 7. Input and model of distance.

$$i(X^{(e)}) = \operatorname{argmin}_i \|X^{(e)} - W_i\| \quad (식 2)$$

여기서 $X^{(e)}$ 는 시간 t 에서의 i 번째 입력 벡터이고 W_j 는 시간 t 에서의 j 번째 입력 벡터와 j 번째 출력 뉴런 사이의 연결 강도이다. 모든 연결 강도와와의 거리를 계산하여 최소 거리에 있는 출력 뉴런을 선택한다(그림7).

3단계: 선택된 뉴런과 그 이웃 반경 내의 뉴런들의 연결강도를 갱신(협동과정과 적응과정)

연결강도의 갱신은 (식3)에 의해 갱신된다.

$$W_k = W_k + \eta(t) \cdot h_{ik}(t) \cdot (X^{(e)} - W_j) \quad (식 3)$$

$\eta(t) \cdot h_{ik}(t)$ 는 0과 1사이의 값을 가지는 이득항(Gain Term)인데 시간이 경과함에 따라 점차 작아진다. 학습 이득 항에는 지수 함수적으로 학습 이득을 감소시키는 함수(학습률 감소 함수)를 사용하며, (식4)와 같다.

$$\eta(t) = \eta_0 \exp\left(-\frac{t}{T_1}\right) \quad (식 4)$$

η_0 : 초기학습 이득, t : 학습반복 회수

T_1 : 학습률 시상수

출력뉴런의 편중학습과 미사용 문제를 보완하기 위하여 승리뉴런과 함께 이웃뉴런을 학습하는데 이때 학습의 범위인 $\alpha(t)$ 도 역시 학습이 진행됨에 따라 학습 이득항과 같이 감소시키는 함수(이웃 감소 함수)를 사용하며, (식5)와 같다.

$$\alpha(t) = \alpha_0 \exp\left(-\frac{t}{T_2}\right) \quad (식 5)$$

α_0 : 초기학습 이웃의 범위, t : 학습 반복 회수

T_2 : 이웃뉴런 학습의 시상수

학습되는 양을 결정하는 방법에는 거리와 무관하게 동일한 이득항으로 갱신하는 방법과 거리에 따라 학습되는 이득을 다르게 적용하는 방법이 있다. 본 논문에서는 거리에 따라 다른 이득으로 학습되는 방법을 적용하였으며, 학습 알고리즘에서 시냅스 Weight 학습함수는 (식6)과 같다.

$$w_{ij}(t+1) = w_{ij}(t) + h(t)m_{ij}(x_i(t) - w_{ij}(t)) \quad (식 6)$$

m_{ij} : 이웃 뉴런의 이득

4. 실험결과

본 논문에서는 수화인식을 하기 위한 손 형상 인식 시스템을 구현하고자 한다. 먼저 손 형상을 정확하게

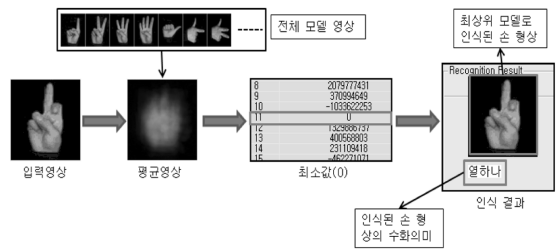


그림 8. 최상위 모델로 인식된 손형상과 수화의미.

Fig. 8. Hand shape recognition of a higher model and meaning sign language.

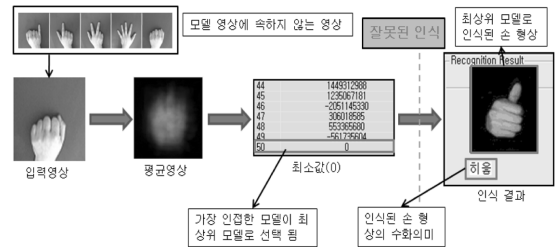


그림 9. 모델 영상에 속하지 않는 영상에 대한 인식결과.

Fig. 9. About a result image that not belong model image.

인식하기 위해 전처리과정으로 모델 영상의 크기를 100*100으로 정규화 했으며, 인식 속도를 높이기 위해 영상을 Gray Level로 변환하였다. 카메라로 들어오는 영상을 입력정보로 사용하였으며, 피부색 정보를 이용하여 영상의 전체 영역에서 손 영역만을 분리하였다. 분리된 손 영역으로부터 손 형상을 정확하고 안정적으로 인식하기 위해 SOFM 신경망 알고리즘을 사용하였다.

4.1. 손 형상 인식 결과

수화인식을 위한 손 형상 인식 시스템은 입력 영상을 전처리 과정에 의해 손 영역을 분할하고, 분할된 손 영역을 손 형상으로 인식한다. 입력 영상이 들어오면 DB안에서 입력 영상의 특징 벡터와 모델 영상의 특징 벡터 사이의 거리를 SOFM 신경망 알고리즘의 경쟁과정(그림 4), 학습과정(그림 5), 적응과정(그림 6)을 통해 비교하고, 거리가 가장 최소값(0)이 되는 모델을 찾는다. 거리가 가장 최소값이 되는 모델이 우리가 찾는 최상위 모델로 인식된다. 또한 이미 부여한 각 영상의 수화 의미도 보여준다. (그림 8)은 최상위 모델의 손 형상 인식과 수화의미 이다.

단, 전체 모델 영상에 속하지 않을 때의 인식률은 0%이다. 이는 최상위 모델에 가장 인접한 모델을 최상

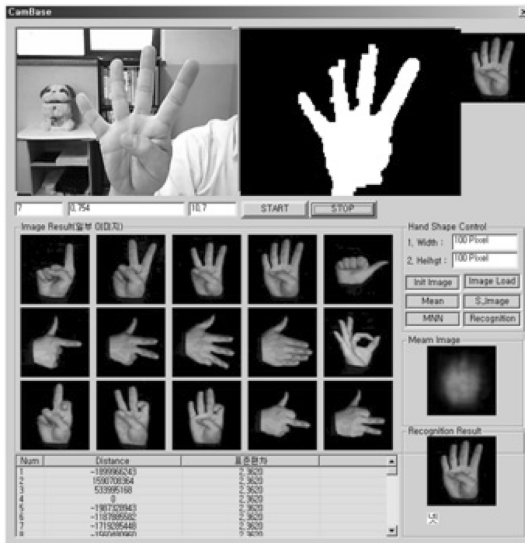


그림 10. 수화인식을 위한 손 형상 인식 시스템.
Fig. 10. Hand shape recognition system for sign language recognition.

위 모델로 선택하였기 때문에 모델 영상과 입력 영상 간 일치하지 않으므로 인식결과가 0%가 된다(그림 9).

5. 결 론

본 논문에서는 장애인과 비장애인 간 원활한 의사 전달을 위하여 수화를 인식하고, 수화의 의미를 전달하는 수화인식 시스템을 구현한다. 수화를 인식하기 위해서는 손 형상을 인식해야 한다. 인식과정은 첫째, 카메라로부터 들어온 입력 영상의 전체 영역에서 피부색 정보를 이용하여 손 영역만을 추출한다(전처리 과정으로 100개의 전체 모델 영상에 대해 사이즈를 100*100으로 정규화 하였으며, 연산속도를 줄이기 위해 모델 영상을 Gray Level(회색조) 정규화 시켰다). 둘째, 전체 영상으로부터 추출된 손 영역을 손 형상으로 정확하고 안정적으로 인식하기 위해 SOFM(Self Organized Feature Map) 신경망 알고리즘을 적용한다. 셋째, SOFM 알고

리즘의 경쟁과정, 학습과정, 적응과정을 통해 입력 영상의 특징 벡터와 모델 영상의 특징 벡터의 거리를 계산한 후 초기 연결강도로부터 점차 연결강도의 가중치를 부여함으로써 입력 영상과 모델 영상간의 거리가 최소값(0)이 되는 모델을 찾고, 찾은 최소값의 모델을 최상위 모델로 선택하며, 최상위 모델로 인식한다(그림 10).

향후에는 복잡한 배경에서의 정교한 분할과 조명변화에 강한 방법과 다른 알고리즘과의 접목으로 손 형상 인식이 더 좋은 결과를 낼 수 있도록 노력을 아끼지 않을 것이다.

참고문헌

- [1] C. Ragael, Gonzalez and E. Richard, Woods "Digital Image Processing", Prentice Hall, 2002.
- [2] E. Y. Kim, S. K. Kang, K. C. Jung, H. J. Kim, "Eye mouse:mouse implementation using eye tracking", IEEE Consumer Electronics, 2005. ICCE. 2005 Digest of Technical Papers. International Conference on 8-12 Jan. 2005.
- [3] J. Segen and S. Kumar, "Shadow Gestures: 3D Hand Pose Estimation Using a Single Camera," CVPR99, vol. 1, pp. 479-485, Fort Collins, Colorado, June, 23-25, 1999.
- [4] K. Imagawa, S. Lu and S. Igi, "Color-Based Hand Tracking System for Sign Language Recognition," Proceedings of International Conference on Automatic Face and Gesture Recognition, Killington, 1996.
- [5] T. Kohonen, "Self-Organization and Associative Memory," Springer Verlag, 1984.
- [6] D. Chi, King N. Ngan "Face Segmentation Using Skin-Color Reference map in Videophone Applications", IEEE Transactions on Circuits and Systems for Vidoe Technology, vol.9, June 1999
- [7] T. Kohonen, "Self-Organizing Maps," Springer Series in Information Science. Vol. 30, Springer, Berlin, Heidelberg, New York, 1995, 1997, 2001, Third Extended Edition.