SOFM신경망을 이용한 수화 형상 인식

김경호*, 김종민*, 정재영**, 이웅기*
*조선대학교 전산통계학과
**서강정보대학 물리치료과
e-mail:cdma1234@hanmail.net

Sign Language Shape Recognition Using SOFM Neural Network

Kyoung-Ho Kim*, Jong-Min Kim*, Jea-Young Jeong**, Woong-Ki Lee*
*Computer Science and Statistics Department. Chosun. Univ.
**Dept of Physical Therapy Seo-Kang College.

요 약

본 논문은 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 인식함으로서 수화인식을 위한 보다 안정적이며 강인한 인식 시스템을 구현하고자 한다.

1. 서론

본 논문은 단일 카메라 환경에서 손 형상을 입력정보로 사용하고 살색정보[1]를 이용하여 손 영역만을 분할한 후 SOFM(Self Organized Feature Map) 신경망[2] 알고리즘을 이용하여 손 형상을 인식함으로서 수화인식을 위한 보다 안정적이며 강인한 인식 시스템을 구현하고자 한다.

2. 피부색 정보를 이용한 손 영역 검출

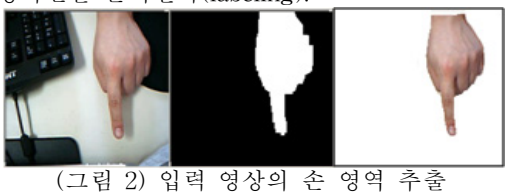
2-1. Skin Color Filtering

RGB 데이터 영상이 입력되면 피부색은 RED 값이 다른 GREEN이나 BLUE보다 크다고 가정하고, 식(1)과 같다.

$$I_2(x,y) = \sum_{x,y} \begin{cases} 1 & (RED >= BLUE) \,\&\&\, (RED >= GREEN) \\ 0 & otherwise \end{cases} \quad 식(1)$$

2-2. 손 영역 검출

본 연구는 칼라 필터가 적용되고 남아 있는 피부색 정보의 픽셀들은 대부분 손에 속해 있기 때문에 횡 좌표 X 축으로 피부색 픽셀들의 수를 계산하여 배경 영역으로부터 손 영역만을 분리한다(labeling).

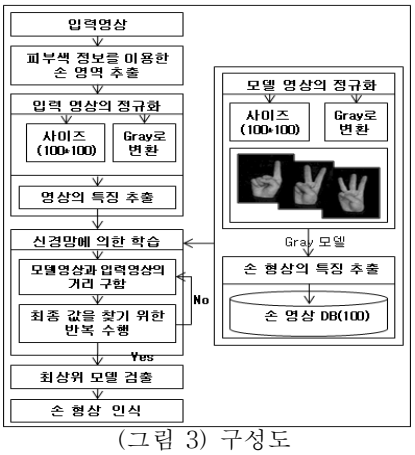


(그림 2) 입력 영상의 손 영역 추출

3. 신경망을 이용한 손 형상 인식

3-1. 제안한 신경망 알고리즘

본 논문에서 제안하는 SOFM은 학습단계에서 다차원 공간상의 특징 벡터로 표현되는 객체들이 경쟁층에서 스스로 유사한 패턴끼리 2차원의 특징지도를 조직화하여 영역지도를 형성하고, 인식단계에서는 학습단계에서 훈련된 연결 가중치 하에서 미지의 특징벡터에 대하여 반응이 일어나는 위치를 통하여 해당 클래스를 인식하게 된다. (그림 3)은 시스템 구성도이다.

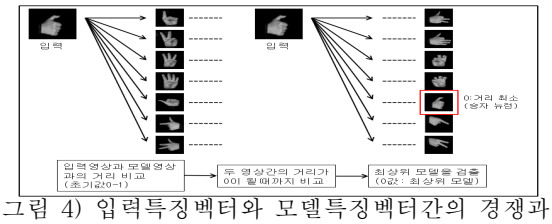


(그림 3) 구성도

3-2. SOFM의 학습과정

3-2-1. 경쟁(Competitive)과정

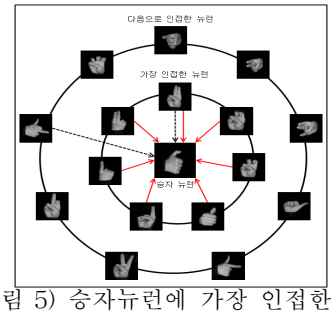
각 뉴런은 입력공간을 구성하는 N차원의 특징벡터와 연결강도에 의하여 연결된다(연결강도 0~1로 초기화). 입력패턴이 주어지면 모든 뉴런의 연결강도와 거리를 계산하여 거리가 최소가 되는 뉴런이 승자가 된다.(그림 4)



(그림 4) 입력특징벡터와 모델특징벡터간의 경쟁과정

3-2-2. 협동(Cooperation)과정

경쟁과정에서 거리가 가장 최소가 되는 승자만이 출력을 낼 수 있으며, 승자의 이웃들만이 연결 강도를 조정할 수 있도록 한다.(그림 5)



(그림 5) 승자뉴런에 가장 인접한 뉴런

3-2-3. 적응(Adaptation)과정

승자뉴런과 그와 이웃한 뉴런들이 특정한 입력 값에 더욱 민감해지도록 활성함수를 적응시켜 연결강도를 갱신한다. 가까운 뉴런이 멀리 있는 뉴런보다 더욱 강하게 적응이 된다. (그림 5)의 점선과 실선부분이 그것을 나타낸다.

3-3. SOFM의 학습 3단계

1단계: 연결강도 초기화

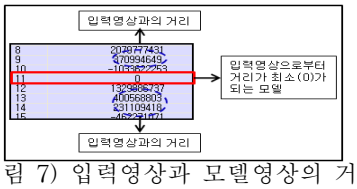
N개의 입력으로부터 M개의 출력뉴런 사이의 연결강도를 작은 값의 임의 수로 초기화한다. 연결강도는 $h_{ik}(t)$에 따라 점차 줄어든다.

2단계 입력벡터가 제시되고 경쟁

입력 벡터와 모든 뉴런들 간의 거리를 계산한다. 입력과 출력뉴런 K사이의 거리는 식(2)와 같이 계산되고 그 중에서 가장 최소 값이 선택된다.(그림 7)

$$i(X^{(\in}) = \mathrm{argmin}_j \| X^{(\in} - W_j \| \qquad \text{식(2)}$$

$X^{(\in}$는 시간 t에서 i번째 입력벡터, W_j는 시간 t에서 i번째 입력벡터와 j번째 출력뉴런 사이의 연결 강도이다.



(그림 7) 입력영상과 모델영상의 거리

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

연결강도의 갱신은 식(3)에 의해 갱신된다.

$$W_k = W_k + \eta(t) \cdot h_{ik}(t) \cdot \left(X^{(\in} - W_j\right) \qquad \text{식(3)}$$

$\eta(t) \cdot h_{ik}(t)$는 0과 1사이의 값을 가지는 이득항(Gain Term)인데 시간이 경과함에 따라 점차 작아진다. 학습률 감소 함수를 사용하며, 식(4)와 같다.

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

η_0: 초기학습 이득 ,t: 학습반복 회수, T_1: 학습률 시상수

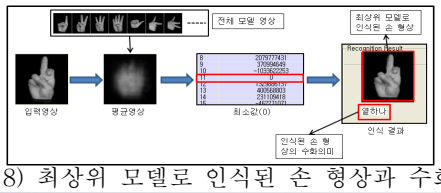
본 논문에서는 거리에 따라 다른 이득으로 학습되는 방법을 적용하였으며, 식(6)과 같다.

$$w_{ij}(t+1) = w_{ij}(t) + \eta(t) m_{ij}(x_i(t) - w_{ij}(t)) \qquad \text{식(6)}$$

m_{ij}: 이웃 뉴런의 이득

4. 실험결과

입력 영상이 들어오면 영상의 특징벡터 사이의 거리를 경쟁과정, 학습과정, 적응과정을 통해 비교하고, 거리가 가장 최소값(0)이 되는 모델을 찾는다.



(그림 8) 최상위 모델로 인식된 손 형상과 수화 의미

(그림 10) 수화인식을 위한 손 형상 인식 시스템

5. 결론

본 실험에서는 모델에 속하는 영상에서는 높은 인식률을 보이지만 속하지 않을 때는 최상위 모델에 가장 인접한 모델을 최상위 모델로 선택하였기 때문에 모델 영상과 입력영상 간 일치하지 않으므로 인식률은 0%이다. 향후에는 복잡한 배경과 조명변화에서도 정교한 분할하며 더 높은 인식결과를 낼 수 있는 알고리즘 개발이 필요하다.

참고문헌

[1] 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

[2] T. Kohonen, "Self-Organizing Maps," Springer Series in Information Science. Vol. 30, Springer, Berlin, Heidelberg, New York, 1995, 1997, 2001, Third Extended Edition