

자기 조직화 특징 지도(SOFM)와 주성분 분석을 이용한 손 형상 검출 및 인식

Hand Shape Detection and Recognition using Self Organized Feature Map(SOFM) and Principal Component Analysis

김경호*, 이기준**

조선대학교 전산통계학과*, 광주보건대학교 병원전산관리과**

Kyoung-Ho Kim(cdma1234@hanmail.net)*, Kee-Jun Lee(leekj@ghu.ac.kr)**

요약

본 논문은 손 형상 인식을 위한 보다 안정적이며 조명 변화와 회전에 강인하게 손 영역을 검출하며, 계산의 효율성과 검출 성능을 동시에 만족시키는 강인한 검출 알고리즘에 대해 제안한다. 제안한 알고리즘은 단일 카메라 환경에서 손 형상을 입력정보로 사용하여 전처리 과정을 거쳐 손 영역만을 분할한 후 자기 조직화 특징 지도(SOFM: Self Organized Feature Map) 알고리즘을 이용하여 손 형상을 인식하게 된다. 그러나 조명 변화에 민감하고 자유도가 큰 손 영역을 정확히 인식하기란 쉽지 않으며 오차 범위가 크기 때문에 본 논문에서는 인식률을 높이기 위해 각각의 손 형상에 대한 회전 정보를 데이터베이스화 한 후 주성분 분석을 적용하여 군집화 함으로서 인식오차를 줄였다. 또한 차원 축소로 인해 많은 계산 량이 요구되지 않기 때문에 실시간 인식 시간도 줄일 수 있었다.

■ **중심어** : | 자기 조직화 특징 지도 | 주성분 분석 | 손 형상 인식 |

Abstract

This study proposed a robust detection algorithm. It detects hands more stably with respect to changes in light and rotation for the identification of a hand shape. Also it satisfies both efficiency of calculation and the function of detection. The algorithm proposed segmented the hand area through pre-processing using a hand shape as input information in an environment with a single camera and then identified the shape using a Self Organized Feature Map(SOFM). However, as it is not easy to exactly recognize a hand area which is sensitive to light, it has a large degree of freedom, and there is a large error bound, to enhance the identification rate, rotation information on the hand shape was made into a database and then a principal component analysis was conducted. Also, as there were fewer calculations due to the fewer dimensions, the time for real-time identification could be decreased.

■ **keyword** : | Self Organized Feature Map(SOFM) | Principal Component Analysis | Recognition of Hand Shape |

1. 서론

정보통신의 급속한 발전으로 인간과 컴퓨터간의 원

활한 상호작용을 위해 인간의 행동을 분석하는 연구가 계속 진행되고 있다[1]. 인간은 정보 전달을 위해 다양한 수단을 사용하여 상대방에게 자신이 경우 흔히 언어

접수일자 : 2013년 09월 23일

수정일자 : 2013년 10월 17일

심사완료일 : 2013년 10월 22일

교신저자 : 이기준, e-mail : leekj@ghu.ac.kr

적 수단만을 고려하기 쉬우나 이 외에도 몸짓, 손짓, 얼굴의 표정 등과 같은 비언어적 수단이 보다 많은 정보를 전달한다는 연구 결과가 보고되어 있다[2]. 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 연구 테마로 다루어져 왔다. 그러나 복잡한 환경으로부터 사람의 손 형상만을 정확히 추출하고 인식하여 의미 있는 입력정보로 사용하는 것은 어려운 일이다. 그러므로 많은 트레이닝 이미지가 필요하게 되고 영상의 크기를 모두 일정하게 맞추는 정규화 과정이 필요하게 된다[3]. 컴퓨터를 통한 손 형상 인식은 출입 통제 시스템, 보안 시스템, 컴퓨터 시각에 의한 시스템 자동화 등의 많은 수의 데이터베이스에 적용할 수 있으며, 높은 계산량이 요구되어진다. 대표적인 손 형상 인식은 템플릿 매칭 접근, 통계적 분류 접근, 구문적 접근 및 신경망 접근 방법의 4가지 방법으로 분류할 수 있다[4]. 이 중에서 템플릿 매칭 접근에서 Elastic Matching 방법과 통계적 분류방법에서 LDA(Linear Discriminant Analysis), PCA(Principal Component Analysis) 방법[5][6]이 손 형상 인식에서 사용되고 있으며, 이 중에서도 계산량이 가장 적은 통계적 분류방법이 손 형상 인식에서 주류를 이루고 있다.

본 논문은 손 형상 인식을 위한 보다 안정적이며 조명 변화와 회전에 강인하게 손 영역을 검출하며, 계산의 효율성과 검출 성능을 동시에 만족시키는 강인한 검출 알고리즘에 대해 제안한다. 먼저 전처리과정으로 모델영상의 사이즈와 압축 및 컬러에 대한 정보를 정규화시켰다. 카메라로 들어온 영상을 입력영상으로 하여 손 영역만을 추출한 후 추출된 손 영역의 형상을 인식하기 위해 자기 조직화 특징 지도(Self Organized Feature Map: 이 후 SOFM이라 함) 신경망[7][8] 알고리즘을 적용하여 손 형상을 인식한다.

또한 손 형상의 인식률을 높이기 위해 각각의 손 형상을 5°씩 360° 회전을 가진 데이터베이스를 생성하고 기존의 모델 영상내의 각각의 손 형상에 대한 대표 값을 만든 후에 PCA를 적용하여 실험 영상을 고유 공간

에 투영시켜서 나온 성분과 대표 값의 거리를 비교하여 인식 오차를 줄임으로서 보다 안정적으로 손 형상을 검출할 수 있게 되었으며, 손 형상 인식률에 대한 안전성과 정확성을 향상시킬 수 있었다. 또한 차원 축소로 인해 많은 계산량이 요구되지 않기 때문에 실시간 인식 시간도 줄일 수 있었다.

II. 전체 시스템 구성도

본 논문에서는 손 형상을 정확히 분석하고 인식하기 위해 PCA를 적용하여 각각의 손 영상을 고유 공간에 투영시킨 후 SOFM 신경망 알고리즘을 사용하여 손 형상 인식을 수행하였다. [그림 1]은 전체 시스템 구성도이다.

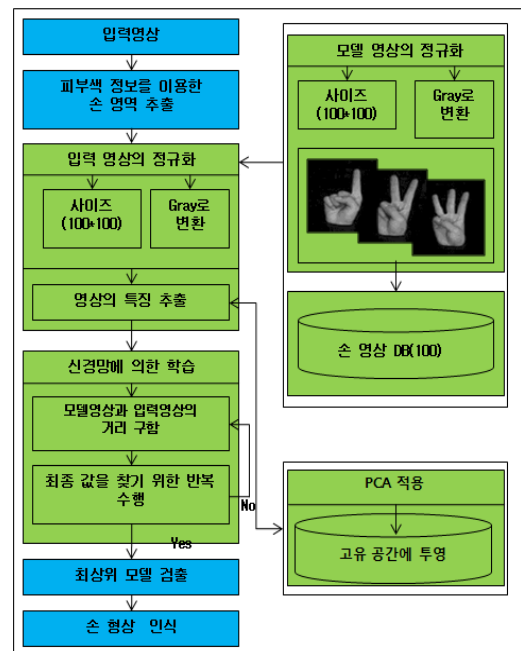


그림 1. 알고리즘 전체 구성도

III. 배경제거

카메라를 통하여 얻은 영상 시퀀스는 단순 배경에서

취득한 것이며, 일반 환경에서 취득한 영상에는 손 형상 인식에 필요 없는 많은 배경들이 포함되어 있다. 그러나 손 형상 인식에 필요한 것은 손 영역(전경)이므로 우선 배경과 손 영역을 분리하는 작업이 필요하고 이를 위해서는 먼저 배경 모델을 생성해야 한다. 그러나 조명의 밝기가 일정하지 않고 수시로 변하기 때문에 같은 카메라로 일정 시간 동안 똑같은 배경을 촬영한다고 할 지라도 모두 동일하지 않아 안정적인 배경 모델을 얻는데 어려움이 따른다.

본 논문에서는 조명 변화로 인한 배경의 밝기 변화를 측정하여 시간 요소(t)를 고려해서 일정 시간 T_i 동안 배경 영상 I_t 를 취득한 다음 영상 영역 R 내에 있는 각 픽셀(x)들을 분석하여 조명이 가장 밝았을 때의 화소값 $P_{\max}(x)$ 와 가장 어두울 때의 화소 $P_{\min}(x)$ 을 얻는다. 이 두 화소 값의 차이 $D(x)$ 는 조명의 변화로 나타날 수 있는 밝기의 임계치이다. 이 3가지의 요소를 이용해 배경 모델(Background Model)을 구성한다. 이 방법은 임의의 한 특정한 밝기 값을 가진 영상을 이용했을 때보다 목표 영역에 대한 접근 및 검출을 더욱 정확하게 할 수 있게 된다. 이와 같은 내용은 식(1-4)와 같다.

$$BM = \{P_{\max}(x), P_{\min}(x), D(x)\}_{x \in R} \quad (1)$$

$$P_{\max}(x) = \text{Max} I_t(x), (1 \leq t \leq T_i) \quad (2)$$

$$P_{\min}(x) = \text{Min} I_t(x), (1 \leq t \leq T_i) \quad (3)$$

$$D(x) = P_{\max}(x) - P_{\min}(x) \quad (4)$$

일단 배경 모델이 만들어지면 이진 영상 $B(x)$ 는 식(5)서 보여주는 것처럼 입력 영상 $I(x)$ 가 가장 밝은 화소값 $P_{\max}(x)$ 와 가장 어두운 화소값 $P_{\min}(x)$ 의 차분 연산을 통해 얻은 차이 값이 임계치 $D(x)$ 보다 크면 최대값(255)을 그렇지 않으면 최소값(0)의 화소값을 갖게 한다.

$$B(x) = \begin{cases} 255 & \text{if } |P_{\max}(x) - I(x)| \text{ or } |P_{\min}(x) - I(x)| > D(x) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

식(5)는 조명으로 인해 생길 수 있는 밝기 차이는 무시하고 손 형상 변화의 움직임의 차이를 갖는 영역만 분리하는 기준이 된다. 식(2)의 결과로 얻어진 이진 영상에는 배경 모델에서 설정한 밝기 값의 임계치를 벗어나는 조명의 변화로 인해 배경임에도 불구하고 전경 영역으로 분리되어 1픽셀의 잡음 점들이 포함될 수 있다.

따라서 이 잡음을 없애기 위해 모폴로지(Morphology) 연산을 통해서 잡음을 제거하였다. 한 번의 침식 연산을 수행하고 이때 객체의 줄어드는 현상이 발생하기에 다시 팽창 연산을 사용하여 복원하였다.

IV. 주성분 분석을 이용한 손 형상 인식

본 논문에서는 손의 방향성이 틀어지는 경우에도 같은 형상으로 인식하기 위해 PCA를 이용해 파라메트릭 공간을 생성하여 손의 방향성 문제를 해결하는 방법을 제시한다. 이는 입력 영상의 자유도를 크게 하기 위해 각 손 형상에 대해 5°씩 틀어 360°의 형상을 만들어 데이터베이스화함으로써 입력 영상의 방향이 틀어져도 같은 형상으로 인식할 수 있게 된다.

PCA란 각 손 형상 영상의 데이터들의 공간적 위치 값이 주로 각 영상에서 어디에 분포하는가를 계산하여 확률 빈도가 높은 벡터 값들을 고유치 값에 비례하여 재구성하는 방법이다. [그림 2]는 손 형상을 5°씩 틀어서 획득한 영상의 일부이다(1개 포즈 = 72 image).

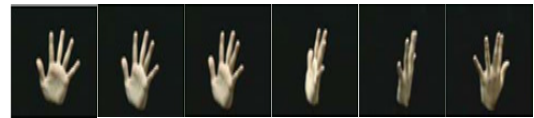


그림 2. 손 형상의 틀어지는 경우의 일부 영상

1. 주성분 분석을 이용한 고유공간 구성

고유벡터를 계산하기 위해서는 먼저 모든 영상의 평균 데이터를 구하여 각 영상들의 데이터의 차를 구한다. 평균 데이터 C 와 새로운 영상 데이터 집합 X 를 식(6)과 식(7)과 같이 나타낸다.

$$C = (1/N) \sum_{i=1}^N x_i \quad (6)$$

$$X = \left\{ x_{1,1}^{(1)} - c, x_{2,1}^{(1)} - c, \dots, x_{R,1}^{(1)} - c, \dots, x_{R,L}^{(P)} - c \right\} \quad (7)$$

고유공간을 구하기 위해서는 $M * N$ 의 크기를 지닌 영상 집합 X 를 식(8)과 같이 계산하고 식(9)를 만족하는 고유벡터를 구하면 된다. 즉, 공분산 행렬 Q 에 대한 고유치 λ 와 고유벡터 e 를 구한다.

$$Q = XX^T \quad (8)$$

$$\lambda_i e_i = Q e_i \quad (9)$$

여기서 M은 한 영상의 데이터(100x100)이고 N은 전체 영상의 개수를 나타내는 정수이다.

본 논문에서는 특이치 분해(Singular Value Decomposition)를 이용하여 고유 공간을 구하였다. 특이치 분해를 이용하여 식(10)에서처럼 행렬이 분해되는데, 영상 집합 X의 공분산 행렬에 대한 고유 벡터를 쉽게 얻을 수 있다.

$$\begin{bmatrix} X \end{bmatrix} = \begin{bmatrix} V \end{bmatrix} \cdot \begin{bmatrix} \sigma^1 & & \\ & \sigma^{21} & \\ & & \sigma^{31} \end{bmatrix} \cdot \begin{bmatrix} U^T \end{bmatrix} \quad (10)$$

행렬은 식(11)과 같이 분해가 가능하며 원 식은 식(15)와 같이 나타낼 수 있다.

$$X = \sum_{i=1}^r \sigma_i V_i U_i^T \quad (11)$$

$$V = [v_1, \dots, v_r] \quad (12)$$

$$U = [u_1, \dots, u_r] \quad (13)$$

$$\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r) \quad (14)$$

$$X = V \Sigma U^T \quad (15)$$

V와 U 행렬은 각각 직교 행렬이기 때문에 식(16)에서 나타낸 것처럼 표현된다.

$$[V^T] \cdot [V] = [U^T] \cdot [U] = [I] \quad (16)$$

이런 과정으로 얻어진 행렬 중 공분산의 고유 벡터로 이용되는 행렬은 공분산 행렬 X와 크기가 일치하는 U이다. 특이치 분해 과정에서 나온 고유벡터를 고유치가 큰 순서대로 재구성한다. 각 고유벡터가 지닌 고유치의 크기는 그 고유벡터의 중요도를 의미하므로 그 고유공간을 규정하는 중요 고유벡터를 식(17)을 이용하여 선택한다. 따라서 모든 고유벡터를 고유 공간 구성에 사용하지 않고 많은 영상을 대표할 수 있는 주성분의 벡터만을 이용할 수 있다.

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^N \lambda_i} \geq T_1 \quad (17)$$

여기서 T₁은 고유벡터의 개수를 조정하는 문턱치이며, 인식과 포즈 평가 시 이용되는 고유벡터는 저차원 공간을 구성하기 위해 실험에서는 K=3을 이용했다. 이렇게 구성된 공간을 손 형상의 고유 공간으로 표현한다.

2. 손 형상 공간에서 상관관계와 거리

앞 절에서 얻어진 손 형상의 고유 공간에서 평균 영상 C에서 뺀 영상 집합 x를 모두 식(18)을 이용하여 투영시킨다.

$$f_j = [e_1, e_2, e_3, \dots, e_k]^T (x_n - c) \quad (18)$$

구해진 f_j 손 형상의 고유 공간상에서 점들로 표현되는데 이러한 점들을 손 형상 포즈 특징 심볼을 구한 것과 같이 입력 특징 심볼을 구성하게 된다. 투영시킨 결과는 이산적인 점들로 표현되며, 이들 각 점은 입력 손 형상 하나 하나를 의미한다. 고유공간의 특성상 비슷한 특징 값을 가지는 벡터는 고유공간에서도 가까운 곳에 투영된다. 따라서 같은 손 형상의 경우 서로 비슷한 특징 벡터 값을 가지게 되고, 고유공간에서 비슷한 영역에 분포하게 된다. 6개의 손 형상(회전한 손 형상)을 본 고유공간에 투영시킴으로서 모델 동작을 모델링한다. 고유 공간에 투영된 점이 가까울수록 영상들의 높은 상관관계를 가진다. 각 손 형상 포즈에 대한 고유 공간내의 분포는 [그림 3]과 같다.

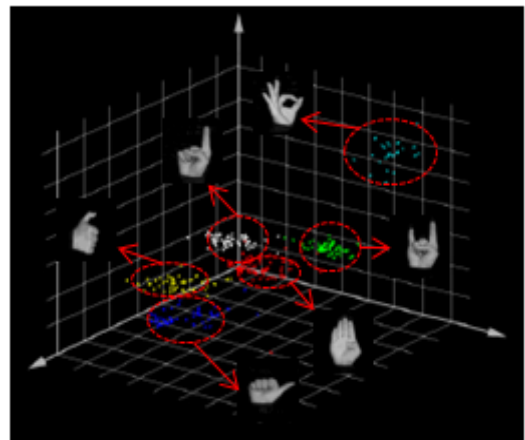


그림 3. 고유공간에서의 손 형상 분포

V. SOFM에 의한 손 형상 인식

SOFM은 입·출력 뉴런 간에 연결된 시냅스들로 구성되어 있으며 학습단계에서 다차원 공간상의 특징 벡터로 표현되는 객체들이 경쟁 층에서 스스로 유사한 패턴끼리 2차원의 특징 지도를 조직화하여 영역 지도를 형성한다. [그림 4]는 SOFM 신경망의 구조이다.

X_i 는 입력 벡터, W_{ik} 는 시냅스의 Weight를 나타낸다. 이와 같은 형태로 구성된 신경망 모델은 출력 뉴런들이 승리하기 위하여 서로 경쟁학습(Competitive Learning)을 하게 되며 단 한 개의 출력 뉴런만이 승리하게 된다. 입력 영상의 특징 벡터와 모델 영상의 특징 벡터 사이의 거리를 비교하기 위하여 앞의 식(6)과 식(7)을 이용한다.

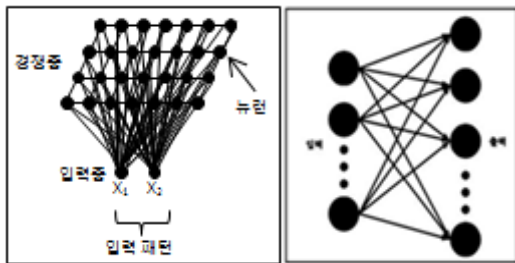


그림 4. SOFM 신경망의 구조

[그림 5]는 전체 모델 영상에 대한 평균 영상이며, 모델 영상은 일부만 나타낸다. 앞의 식에서 N 은 손 영상의 개수이며 x_i 는 입력 영상이다.

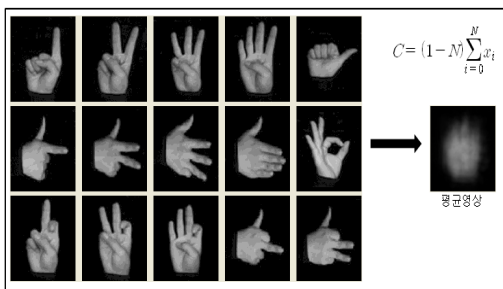


그림 5. 전체 모델 영상의 평균

1. SOFM의 학습과정

1.1 경쟁(Competitive)과정

각 뉴런은 입력 공간을 구성하는 N차원의 특징 벡터와 연결 강도에 의하여 연결되며 연결 강도는 0에서 1 사이(0~1)로 정규화 된 임의의 값으로 초기화되어 있다. [그림 6]은 최상의 모델인 승자 뉴런을 찾는 경쟁과정을 보여준다.

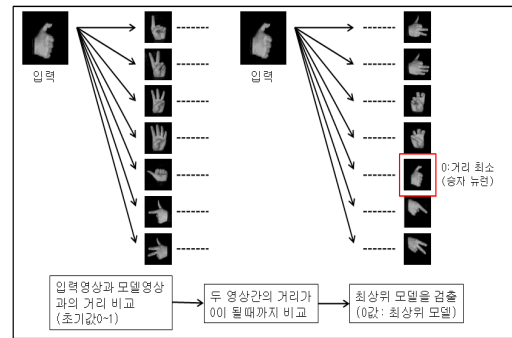


그림 6. 입력특징벡터와 모델특징벡터간의 경쟁과정

1.2 협동(Cooperation)과정

경쟁 과정에서의 승자 뉴런과 인접한 이웃 뉴런들만이 제시된 입력 벡터에 대하여 학습이 허용된다. 즉, 승자만이 출력을 낼 수 있으며, 승자와 그의 이웃들만이 연결 강도를 조정할 수 있도록 한다. [그림 7]은 승자 뉴런에 가장 인접한 이웃 뉴런을 나타낸다.

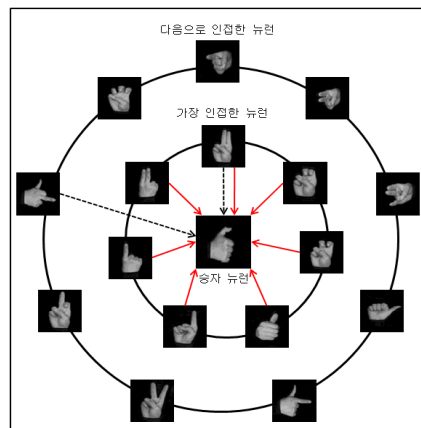


그림 7. 협동과정에서의 승자뉴런에 가장 인접한 뉴런

1.3 적응(Adaptation)과정

승자 뉴런과 그와 이웃한 뉴런들이 특정한 입력 값에 더욱 민감해지도록 활성 함수를 적용시켜 연결 강도를 갱신한다. [그림 7]에서의 붉은 점선이 가장 인접한 이웃뉴런들이며 다른 이웃뉴런들 보다 연결강도가 높다는 것을 나타낸다.

2. SOFM의 학습 3단계

2.1 1단계: 연결강도 초기화

N개의 입력으로부터 M개의 출력 뉴런 사이의 연결 강도를 작은 값의 임의의 수로 초기화하며, 함수 $h_{ik}(t)$ 에 따라 점차 줄어든다.

2.2 2단계: 입력벡터가 제시되고 경쟁

입력 벡터와 모든 뉴런들 간의 거리를 계산한다. 입력과 출력 뉴런 K사이의 거리는 식(19)과 같이 계산되고, 그 중에서 가장 최소값이 선택된다. [그림 8]은 입력 영상과 모델 영상 사이의 거리를 계산하여 가장 최소값(0)이 되는 최상위 모델을 찾은 것이다.

$$i(X^{(e)}) = \operatorname{argmin}_j \| X^{(e)} - W_j \| \quad (19)$$

여기서 $X^{(e)}$ 는 시간 t 에서의 i 번째 입력 벡터이고, W_j 는 시간 t 에서의 i 번째 입력 벡터와 j 번째 출력 뉴런 사이의 연결 강도이다.

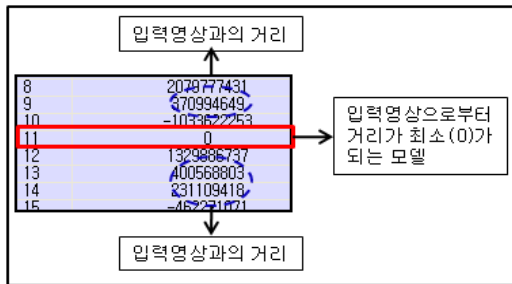


그림 8. 입력영상과 모델영상의 거리

2.3 3단계: 연결강도를 갱신

선택된 뉴런과 그 이웃 환경 내의 뉴런들의 연결강도는 식(20)에 의해 갱신된다.

$$W_k = W_k + \eta(t) \cdot h_{ik}(t) \cdot (X^{(e)} - W_j) \quad (20)$$

$\eta(t) \cdot h_{ik}(t)$ 는 0과 1사이의 값을 가지는 이득항 (Gain Term)인데 학습률 감소 함수인 식(21)과 이웃 감소 함수인 식(22)에 의해 시간이 경과함에 따라 점차 작아진다.

$$\eta(t) = \eta_0 \exp\left(-\frac{t}{T_1}\right) \quad (21)$$

η_0 : 초기학습 이득, t : 학습 반복 회수, T_1 : 학습률 시상수

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{T_2}\right) \quad (22)$$

σ_0 : 초기학습 이웃의 범위, t : 학습 반복 회수

T_2 : 이웃뉴런 학습의 시상수

학습 양의 결정은 거리에 따라 학습되는 이득을 다르게 적용하였으며, 시냅스 Weight 학습 함수는 식(23)과 같다.

$$w_{ij}(t+1) = w_{ij}(t) + \eta(t) m_{ij}(x_i(t) - w_{ij}(t)) \quad (23)$$

m_{ij} : 이웃 뉴런의 이득

VI. 실험결과

손 형상 인식 과정은 첫째, 카메라로부터 촬영된 640*480의 입력 영상을 전처리과정으로 100*100으로 정규화 하였으며, 연산속도를 줄이기 위해 모델 영상을 Gray Level로 변환하였다.

둘째, 조명변화에 덜 민감하게 하기 위해 서로 다른 밝기값을 가진 두 영상의 차를 이용하여 임계치를 적용함으로써 목표 영역에 대한 접근 및 검출 능력을 향상시켰다.

셋째, 손 형상의 인식률을 높이기 위해 각 손 형상에 대해 5°씩 틀어 360°의 회전을 가진 영상을 만들어 데이터베이스를 생성한 후 PCA를 적용하여 실험 영상을 고유공간에 투영시켜서 나온 성분과 대표 값을 비교함으로써 입력 영상의 방향성에 대한 인식 오차를 줄였다.

넷째, SOFM 알고리즘의 경쟁과정, 협동과정, 적응과정을 통해 입력 영상의 특징 벡터와 모델 영상의 특징 벡터의 거리를 계산한 후 연결강도의 가중치를 부여함으로써 입력 영상과 모델 영상간의 거리가 최소값(0)이

되는 최상위 모델을 찾아 손 형상으로 인식한다[그림 9]. [그림 10]은 시스템 상에서 손 형상 인식 결과를 보여준다.

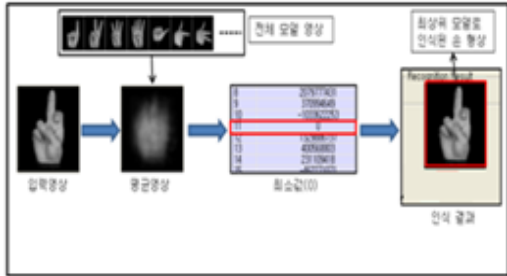


그림 9. 최상위 모델로 찾는 과정

[표 1]은 Pattern Matching을 이용한 인식률, PCA를 이용한 인식률과 본 논문에서 제안하는 PCA와 SOFM 신경망을 이용한 각각의 인식률을 분석한 결과이다. [표 1]의 인식 결과를 보면 알 수 있듯이 본 논문에서 제안하는 알고리즘이 다른 알고리즘의 인식률보다 높은 인식 결과를 보였으며, 특히 모델구성에 속하지 않을 때의 인식률이 많은 개선을 보였다. 그러나 손 형상이 90°가 되었을 경우와 모델영상 내에 입력 영상이 없을 때에는 인식과정에서 잘못된 손 형상으로 오-인식(Error Recognition)하는 경우가 종종 발생하였다.



그림 10. 정상적인 손 형상 인식

[표 1]의 인식결과를 분석해 보면 첫째, 패턴 인식은 기하학적인 정보만을 이용하기 때문에 입력 영상의 방향이 조금만 틀어져도 기하학적인 정보가 매칭이 되지 않기 때문에 인식률은 현저히 떨어진다. 둘째, PCA의 경우 입력 영상의 포즈가 거의 비슷하다면 다른 포즈인데도 불구하고 같은 형상으로 오-인식하는 문제가 자주 발생한다. 셋째, SOFM의 경우 PCA로 획득한 특징 정보를 학습단계에 걸쳐 입력 영상에 대해 연결강도를 강화할 수 있으므로 모델영상에 속하는 포즈에 대해서 다른 포즈를 같은 형상으로 오-인식하는 문제가 많이 개선되어 높은 인식률을 보인다. 하지만 모델영상에 속하지 않을 때는 입력 영상의 특징과 모델영상의 특징을 비교하는 과정에서 거리가 가장 가까운 것을 최상의 모델로 인지하기 때문에 다른 알고리즘보다 실패율이 있어 많은 개선을 보였지만 여전히 오-인식하는 문제가 존재한다.

표 1. 인식 방법에 따른 성능평가 분석표

알고리즘	모델 영상 분류	성공률 (%)	실패율 (%)
Pattern Matching	속하는	78.5	21.5
	속하지 않는	58.4	41.6
PCA	속하는	83.7	16.3
	속하지 않는	63.5	36.5
PCA+SOFM	속하는	90.2	9.8
	속하지 않는	73.1	26.9

[표 2]는 제안한 알고리즘에 대한 조명 변화에 따른 실시간 인식 성능을 분석한 결과로써 조명이 밝았을 때와 어두웠을 때 인식률이 떨어지는 것은 입력 영상에서 손 영역의 검출이 정확히 이루어지지 않았기 때문이다. 인식 시간에서도 단적인 밝기 정보보다 두 영상의 차를 이용한 경우가 약 0.6~0.7초 정도 더 빠른 인식을 보였다.

표 2. 조명 변화에 따른 인식 성능평가 분석표

조명	인식률 (%)	실패율 (%)	인식시간 (sec)
밝았을 때	86.3	13.7	1.54
두 영상의 차	90.2	9.8	0.82
어두웠을 때	84.7	15.3	1.43

VII. 결론

본 논문은 손 형상 인식을 위한 보다 안정적이며 조명 변화와 회전에 강인하게 손 영역을 검출하며, 계산의 효율성과 검출 성능을 동시에 만족시키는 강인한 검출 알고리즘에 대해 제안했다.

손 영역의 검출은 밝기값이 서로 다른 두 영상의 차를 이용하였으며, 손 형상을 정확히 인식하기 위해 SOFM의 학습과정을 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 선택하여 손 형상으로 인식하였다.

또한 손 형상의 인식률을 높이기 위해 각각의 손 형상을 5°씩 틀어 360° 회전을 가진 데이터베이스를 생성하고 기존의 모델 영상내의 각각의 손 형상에 대한 대표 값을 만든 후에 PCA를 적용하여 실험 영상을 고유 공간에 투영시켜서 나온 성분과 대표 값의 거리를 비교함으로써 방향성에 대한 인식 오차를 줄임으로서 보다 안정적으로 손 형상을 인식할 수 있게 되었다. 그 결과 기존 다른 알고리즘 보다 인식률이 개선되었으며, 손 형상 인식률에 대한 안전성과 정확성을 향상시킬 수 있었다. 또한 차원 축소로 인해 많은 계산 량이 요구되지 않기 때문에 실시간 인식 시간도 줄일 수 있었다. 그러나 손 형상이 90°가 되었을 경우와 모델구성에 속하지 않을 때에는 잘못된 손 형상으로 인식하는 경우가 종종 발생하였다.

향후에는 이러한 문제점을 해결할 수 있는 보다 안정적인 손 형상 인식 알고리즘을 개발할 계획이다.

참고 문헌

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, 2002.
 [2] J. Ohya, Y. Kitamura, H. Takemura, F. Kishino, and N. Terashima, "Real-Time Reproduction of

3D Human Images in Virtual Space Teleconferencing," in Proc. of '93 IEEE Virtual Reality Annual Int. Symp, pp.408-414, 1993(9).

[3] J. M. Kim and M. K. Song, "Three Dimensional Gesture Recognition Using PCA of Stereo Images and Modified Matching Algorithm," *IEEE Fuzzy Systems and Knowledge Discovery*, Vol.4, pp.116-120, 2008(10).
 [4] R. O. Duda, P. E. Hart, and D. G. Strok, *Pattern Classification*, Second Edition by John Wiley & Sons, Inc, 2001.
 [5] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *PAMI, IEEE Transactions on*, Vol.20, Issue.1, pp.23-38, 1998(1).
 [6] J. Yang and A. Waibel, "A real-time face tracker," In Proc. 3rd Workshop on Appl. of Computer Vision, pp.142-147, 1996(12).
 [7] T. Kohonen, "Self-Organization and Associative Memory," *Springer Series in Information Science*, Vol.8, 1988.
 [8] T. Kohonen, "Self-Organizing Maps," *Springer Series in Information Science*, Vol.30, Berlin, Heidelberg, New York, 1995, 1997, 2001.

저자 소개

김 경 호(Kyoung-Ho Kim)

정회원



- 2006년 8월 : 전남대학교 컴퓨터 공학과(공학석사)
- 2010년 8월 : 조선대학교 전산통계학과(이학박사수료)
- 2009년 3월 ~ 현재 : 조선대학교, 광주대학교 외래교수

<관심분야> : 컴퓨터비전, 신경망, 멀티미디어

이 기 준(Kee-Jun Lee)

정회원



- 1994년 2월 : 조선대학교 전산통계학과(이학사)
 - 1997년 2월 : 조선대학교 전산통계학과(이학석사)
 - 2001년 2월 : 조선대학교 전산통계학과(이학박사)
 - 2002년 3월 ~ 현재 : 광주보건대학교 병원전산관리과 교수
- <관심분야> : 영상처리, 내용기반검색, 멀티미디어