

가버 웨이블릿 신경망 기반 적응 표정인식 시스템

Adaptive Facial Expression Recognition System based on Gabor Wavelet Neural Network

이상완^{*} · 김대진^{**} · 김용수^{***} · 변증남^{*}

Sang Wan Lee, Dae-Jin Kim, Yong Soo Kim and Zeungnam Bien

^{*} 한국과학기술원 전자전산학과

^{**} 한국과학기술원 인간친화 복지 로봇 시스템 연구센터

^{***} 대전대학교 컴퓨터 공학부

요 약

본 논문에서는 6개의 특징점을 이용하는 가버 웨이블릿 신경망 기반 적응 표정인식 시스템을 제안한다. 특징 추출부를 포함하는 초기 네트워크의 구성은 Levenberg-Marquardt 기반의 학습방법이 사용되며, 따라서 특징 추출부 결정에 있어서 경험적 요소를 배제시킬 수 있다. 또한 새로운 사용자에게 적응 네트워크를 구성하기 위해서 개선된 보상함수를 가지는 Q-학습과, 비지도 퍼지 신경망 모델을 사용하였다. Q-학습을 통해서 개인 사용자에게 분리도가 좋은 특징벡터를 얻을 수 있는 가버 필터 세트를 얻을 수 있으며, 퍼지 신경망을 통해서 사용자의 얼굴변화에 맞게 인식기를 변화시킬 수 있다. 따라서 제안된 시스템은 사용자의 얼굴변화를 따라갈 수 있는 좋은 적응 성능을 보이고 있다.

Abstract

In this paper, adaptive Facial Emotional Recognition system based on Gabor Wavelet Neural Network, considering six feature points in face image to extract specific features of facial expression, is proposed. Levenberg-Marquardt-based training methodology is used to formulate initial network, including feature extraction stage. Therefore, heuristics in determining feature extraction process can be excluded. Moreover, to make an adaptive network for new user, Q-learning which has enhanced reward function and unsupervised fuzzy neural network model are used. Q-learning enables the system to get optimal Gabor filters' sets which are capable of obtaining separable features, and Fuzzy Neural Network enables it to adapt to the user's change. Therefore, proposed system has a good on-line adaptation capability, meaning that it can trace the change of user's face continuously.

Key words : 적응 표정 인식, 가버 웨이블릿 신경망, 특징 추출부 학습, Q 학습, 퍼지 신경망 모델

1. 서 론

인간과 로봇, 인간과 컴퓨터 간의 효율적인 정보 소통에 있어서 좀 더 편안하고, 좀 더 인간에게 친숙한 환경을 제공하기 위한 사용자의 감정 인식 기술은 그 중심에 있다고 하겠다. 이러한 감정을 나타내기 위해 가장 많이 사용되는 수단 중 가장 큰 부분은 차지하는 것이 바로 얼굴 표정이다. 심리학자 Mehrabian은 1971년 보고서에서 글로서 전달되는 메시지는 7%에 불과하며, 35%가 음성이나 제스처에 의해 전달되며, 무려 55%의 메시지는 표정을 통해 전달된다고 밝히고 있다. 그러므로, 표정인식을 통하여 인간과 교류하는 것은 가장 근본적이며, 가장 효율적인 방법이라 할 수 있다.

얼굴 표정은 다차원성, 특징간의 상호 연관성, 애매성, 주관성 등의 특징을 가지고 있다[1]. 표정에 관련된 수많은 특징들이 존재한다는 것과 특징들 간의 복잡성이라는 표정 인식에서의 난제에 대한 해결점은, 그것들 사이의 연관성을 효율적으로 분석해야 한다는 것에서 시작된다. 이러한 복잡한 연관성을 적용하기 위해서 다양하고 복잡한 구조의 관련 연구가 진행되어 왔다[2-6]. 그러나 특징을 더 많이 쓸수록 네트워크 내부의 복잡도가 같이 증가하는 'curse of dimensionality'의 구조적 한계를 벗어나지는 못한다.

더구나, 이러한 표정인식 시스템이 실제 사용자에게 적용될 경우, 개인의 표정 특징 차이로 인해 만족할 만한 성능을 내지 못한다. 따라서 표정인식에 있어서 또다른 문제점은 개인 사용자의 변화에 얼마나 적응해 나갈 수 있는가라고 할 수 있다. 적응에 있어서는 개인화가 기본 전제가 되며, 개인화의 문제는 이미 [6]을 통하여 연구가 진행되어 왔다. 개인화가 각 사용자에게 대한 모델의 갱신과 추가에 대한 문제라고 한다면, 본 연구에서는 각 사용자 모델의 생성 이후 적응의 문제를 다루고 있다고 할 수 있다.

본 연구에서는 새로운 연산 과정과 Levenberg - Marquardt 학습 방법 기반의 가버 웨이블릿 신경망 기반의

접수일자 : 2005년 10월 12일

완료일자 : 2006년 2월 2일

감사의 글 : 본 연구는 과학기술부/한국과학재단 우수 연구센터육성사업(R11-1999-008) 및 산업자원부 지원으로 수행하는 21세기 프론티어 연구개발사업(인간기능 생활지원 지능로봇 기술개발사업)의 재정지원을 받아 수행되었다.

표정인식 시스템을 제안한다. 새로운 사용자를 위한 개인화 적응 네트워크 구성을 위해서는 특징 추출부의 학습 방법으로 Q-학습기반의 방법[7]을 이용함으로써 경험적인 방법에 의존하였던 특징 추출의 한계를 근본적으로 극복한다. 그리고 인식기로는 학습되지 않은 새로운 사용자에 적응하며, 그 사용자의 계속된 변화에도 적응해 나갈 수 있도록 하기 위하여, 비지도 퍼지 신경망 모델[8]을 사용하여 사용자의 변화에 계속적으로 적응하도록 하였다.

본 논문의 구성은 다음과 같다. 2장에서는 가버 웨이블릿 신경망 모델과 초기 네트워크 구성을 위한 학습방법이 소개되며, 3장에서는 개인 적응 모델 형성을 위한 Q-학습 기반의 특징 추출부 학습 방법과, 인식기로서의 퍼지 신경망 모델이 소개된다. 4장에서는 실험 및 결과를, 마지막 5장에서는 결론을 제시하고 향후 과제에 대한 언급을 하기로 한다.

2. 가버 웨이블릿 신경망

2.1 가버 웨이블릿 신경망의 구조

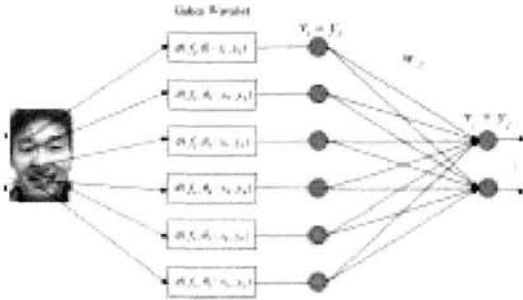


그림 1. 제안된 가버 웨이블릿 신경망 구조
Fig. 1. Proposed Gabor Wavelet Neural Network (GWNN)

전체 시스템의 구조는 그림 1과 같다. 기본적으로 가버 노드와 가중치단으로 구성된 신경망 구조를 가지며, 모든 노드는 활성함수로 선형함수를 사용한다. 얼굴의 특징점은 MPEG-4에 규정된 특징점 중 [8]에서 제시된 중요도 순서의 1-6번째까지의 점을 선정하였다. 적용되는 가버 웨이블릿 함수는 even과 odd를 모두 사용하며 함수는 다음과 같이 주어진다. 이때, 영상에서의 (x_θ, y_θ) 는 점 (x, y) 가 θ 만큼 회전된 점을 의미하며, σ_x, σ_y 는 가버 필터의 각 방향으로의 분산을 의미한다.

$$\begin{aligned} \phi^{odd}(x, y) &= g^{odd}(x, y) = \exp\left\{-\frac{1}{2}\left(\frac{x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right)\right\} \sin(2\pi x_\theta f) \\ \phi^{even}(x, y) &= g^{even}(x, y) \\ &= \exp\left\{-\frac{1}{2}\left(\frac{x_\theta^2}{\sigma_x^2} + \frac{y_\theta^2}{\sigma_y^2}\right)\right\} \{\cos(2\pi x_\theta f) - \exp\left(-\frac{\sigma_x \sigma_y}{2}\right)\} \end{aligned} \quad (1)$$

그리고, 사용자로부터 들어오는 i 번째 입력 영상은 $I_i = I(x_i, y_i)$ 라 두고, 각 kernel과 이미지와 필터의 곱 연산은 다음과 같이 정의한다. 이때의 특징점은 각 필터가 적용되는 중심점이며, 동시에 각 필터가 가지는 f, θ 는 미리 학습된 파라미터 이다.

$$I_i \otimes g_i^{even/odd} = \sum_{k=1}^{K_i} \sum_{l=1}^{L_i} I_i(K_i - k, L_i - l) g_i^{even/odd}(k, l) \quad (2)$$

이때, $g_i^{even/odd}$ 는 (1)식의 even 또는 odd 함수를 의미하고, K_i, L_i 는 2차원 필터의 가로, 세로크기를 나타낸다.

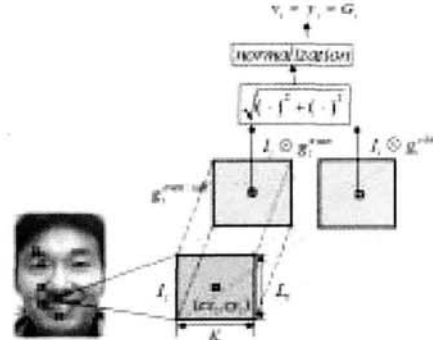


그림 2. 가버 웨이블릿 신경망의 특징 추출
Fig. 2. Feature extraction of GWNN

이제, 영상과 커널의 연산결과는 다음과 같다.

$$G_i^{even} = I_i \otimes g_i^{even} \quad G_i^{odd} = I_i \otimes g_i^{odd} \quad (3)$$

이를 이용하여 i 번째 Gabor 필터 출력의 크기를 계산할 수 있다. Even과 odd 커널의 출력을 이용하여 정규화된 값을 구하게 되는데, 이때 특징점 적용의 부정확성을 반영하는 위상은 무시하고, 크기만을 취한다.

$$v_i = y_i = \frac{\sqrt{|G_i^{even}|^2 + |G_i^{odd}|^2}}{G_i^{MAX}} \quad (4)$$

위 식에서 가버 필터가 낼 수 있는 최대치는 다음과 같이 구할 수 있다.

$$\begin{aligned} G_i^{MAX} &= \sqrt{(\text{Max}|G_i^{even}|)^2 + (\text{Max}|G_i^{odd}|)^2} \\ \text{Max}|G_i^{even/odd}| &\triangleq \sum_{k=1}^{K_i} \sum_{l=1}^{L_i} \left[255 \times \frac{1 + \text{sgn}\{g_i^{even/odd}(k, l)\}}{2} \right] \end{aligned} \quad (5)$$

각 노드의 출력에 각각의 웨이트가 곱해지면, 최종적인 전체 네트워크의 출력은 다음과 같다.

$$y_j = \overline{w_j^T} \cdot \overline{v} = \sum_{i=1}^6 w_{ji} y_i = \sum_{i=1}^6 w_{ji} \frac{\sqrt{|G_i^{even}|^2 + |G_i^{odd}|^2}}{G_i^{MAX}} \quad (6)$$

(4)식의 특징 벡터는 (6)과 같은 형태 외에도 여러 가지 인식기로의 적용이 가능하다. 적응 인식을 위한 비지도 학습 인식기로의 적용은 3장에서 설명될 것이다.

2.2 가버 웨이블릿 신경망의 초기 학습

본 논문에서는 미분 기반의 학습 중, 실제 상황에서 가장 좋은 성능을 보이는 것으로 알려진 Levenberg - Marquardt 방법을 사용한다. 최종 출력층의 j 번째 노드의 출력을 y_j ,

그때의 올바른 출력을 d_j 라 두면, 출력 노드의 에러는 다음과 같다.

$$e_j(n) = d_j(n) - y_j(n), E(n) = \frac{1}{2} \sum_{j=1}^J e_j^2(n), E_{total} = \frac{1}{2} \sum_{n=1}^N \sum_{j=1}^J e_j^2(n) \quad (7)$$

이로부터 기존의 오차 역전파 방법과 동일한 방법으로 최종 출력단 가중치의 갱신을 위한 전체에러의 가중치에 대한 부분 미분을 구할 수 있으므로 수식은 생략한다.

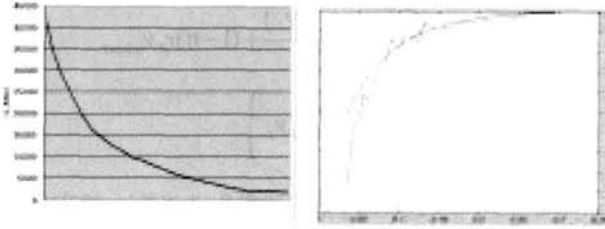


그림 3. 주파수에 대한 필터의 최대값과 그것의 미분함수 근사화 그래프

Fig. 3. Maximum value of Gabor filters and its derivative's approximation

두 번째 학습은 특징 추출부인 가버 웨이블릿 파라미터 f, θ 를 학습시키는 것이다. 자코비안 행렬을 위해서는 $(\partial E / \partial v_i)$ 와 $(\partial v_i / \partial f, \partial v_i / \partial \theta)$ 를 구해야 한다. 전체오차에 대한 가버 노드의 부분 미분을 구하면 다음과 같다.

$$\frac{\partial E_{total}(n)}{\partial f_i(n)} = \sum_{n=1}^N \frac{\partial E(n)}{\partial v_i(n)} \frac{\partial v_i(n)}{\partial f_i(n)}, \frac{\partial E(n)}{\partial v_i(n)} = - \sum_{j=1}^J e_j(n) w_{ji}(n) \quad (8)$$

이제, 위 식에서 가버 노드의 출력에 대한 특징 주파수와 특징 각도의 부분 미분 $(\partial v_i / \partial f, \partial v_i / \partial \theta)$ 을 구하기 전에 먼저 필터의 최대값에 대한 주파수의 부분미분을 (4), (5)를 이용하여 구해야 한다. 그러나, 미분식의 전개가 어려우므로 실제 데이터를 바탕으로 미분 함수를 다음과 같이 Gradient descent 방법으로 근사화하여 사용한다. 근사화를 위한 최대값 데이터와 근사화 그래프는 그림3과 같고, 근사화 식은 아래와 같다.

$$\frac{dG_i^{MAX}(f)}{df} \cong \beta(f) = -9.1306 \cdot 10^5 \cdot e^{16.3272f} - 10.0174 \cdot e^{0.99970f} \quad (9)$$

이제 주파수에 대한 첫번째 층의 출력의 부분 미분값을 다음과 같이 정리할 수 있다.

$$\frac{\partial v_i(n)}{\partial f_i(n)} = \frac{1}{G_i^{MAX}(n) \sqrt{G_i^{even}(f_i(n), \theta_i(n))^2 + G_i^{odd}(f_i(n), \theta_i(n))^2}} \cdot \left\{ G_i^{even}(f_i(n), \theta_i(n)) \frac{\partial G_i^{even}(f_i(n), \theta_i(n))}{\partial f_i(n)} + G_i^{odd}(f_i(n), \theta_i(n)) \frac{\partial G_i^{odd}(f_i(n), \theta_i(n))}{\partial f_i(n)} \right\} - \frac{1}{G_i^{MAX}(n)^2} \frac{\partial G_i^{MAX}(n)}{\partial f_i(n)} \sqrt{G_i^{even}(f_i(n), \theta_i(n))^2 + G_i^{odd}(f_i(n), \theta_i(n))^2}$$

$$\cong \frac{1}{(G_i^{MAX})^2(n) \cdot v_i(n)} \left\{ G_i^{even}(f_i(n), \theta_i(n)) \frac{\partial G_i^{even}(f_i(n), \theta_i(n))}{\partial f_i(n)} + G_i^{odd}(f_i(n), \theta_i(n)) \frac{\partial G_i^{odd}(f_i(n), \theta_i(n))}{\partial f_i(n)} \right\} - \beta(f) \frac{v_i(n)}{G_i^{MAX}(n)} \quad (10)$$

$$\left(\because \beta(f) \cong \frac{\partial G_i^{MAX}(n)}{\partial f_i(n)}, v_i(n) = \frac{\sqrt{G_i^{even}(n)^2 + G_i^{odd}(n)^2}}{G_i^{MAX}(n)} \right)$$

이제 앞 식에서 even과 odd 각 가버 출력에 대한 부분 미분값을 구하면 다음과 같이 정리된다.

$$\frac{\partial G_i^{even}(f_i, \theta_i)}{\partial f_i} = \frac{\partial \{ I_i \otimes g_i^{even}(f_i, \theta_i) \}}{\partial f_i} = \frac{\partial \left\{ \sum_{k=1}^{K_i} \sum_{l=1}^{L_i} I_i(K_i - k, L_i - l) g_i^{even}(k, l | f_i, \theta_i) \right\}}{\partial f_i} = \sum_{k=1}^{K_i} \sum_{l=1}^{L_i} I_i(K_i - k, L_i - l) \frac{\partial g_i^{even}(k, l | f_i, \theta_i)}{\partial f_i} = I_i \otimes \frac{\partial g_i^{even}(f_i, \theta_i)}{\partial f_i} \quad (11)$$

odd 부분 역시 마찬가지로 구할 수 있다.

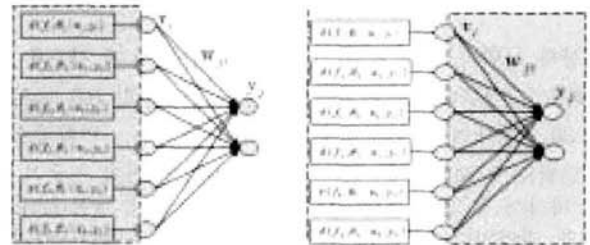
이로서, (8)식의 전체 미분에 대한 각 f 의 부분 미분을 계산할 수 있으며, θ 의 경우 역시 마찬가지로 방법으로 구할 수 있다.

이제, Levenberg - Marquardt 학습을 위한 자코비안 행렬을 구성하기 위하여 본 문제에서 학습시키려는 파라미터 세트를 다음과 같이 구한다.

$$\phi \triangleq (f_1, \dots, f_6, \theta_1, \dots, \theta_6, w_{11}, w_{12}, \dots, w_{66}) \quad (12)$$

이제 전체 오차로부터 기울기 벡터를 정의할 수 있으며, 자코비안 행렬을 구성하여 다음과 같은 갱신식에 적용한다. 학습의 원리는 기본적인 가우스-뉴턴 방향으로 진행하다가, 에러가 증가하면 작은 학습률의 최대 기울기 방향으로 전환하여 안정화를 꾀하게 된다.

3. 개인화를 위한 적응 네트워크 구성



(가) 특징 추출부 조정 (나) 비지도 분류기 구성

그림 4. 개인화를 위한 적응 네트워크 구성 과정
Fig. 4. Adaptive network for personalization

초기 학습에서는 앞서 2장에서 설명하였듯이, 가버 필터의 파라미터와 가중치를 동시에 학습시킨다. 그러나, 초기 학습된 인식기를 학습 데이터와는 다른 성향의 특정 사용자가 사

용하게 되는 경우, 인식률이 떨어질 수 있다. 그 원인은 사용자의 얼굴 특징 자체가 달라지기 때문일 수도 있으며, 감정을 표현하기 위한 방법이 다르기 때문일 수도 있다. 전자의 경우는 특징 추출이 제대로 이루어지지 않기 때문이라고 할 수 있으며, 후자의 경우는 두 번째 층인 가중치가 달라지기 때문이라고 할 수 있다. 따라서 개인 사용자에게 인식기가 적용하기 위해서는 특징 추출부와 그 조합을 설명하는 가중치에 대한 학습이 모두 이루어져야 한다. 단, 적용 단계에서는 개인 사용자가 학습 데이터에 일일이 해당 표정의 결과, 즉 인덱스를 붙이기 번거로우므로, 사실상 지도 학습은 지양해야 한다.

3.1 적응 1단계 : 특징 추출부 조정

적응 1단계인 Q-학습은 학습되지 않은 새로운 사용자가 적응을 필요로 할 때, 더 나은 가버 필터의 특징 주파수와 각도를 찾기 위해 수행된다. 한 번의 입력이 들어올 때마다 상태 이동이 일어나며, 일정량의 데이터가 축적되면 반복학습을 수행한다.

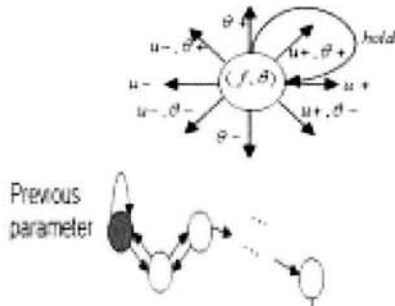


그림 5. Q-학습의 행동 정의와 상태 이동
Fig. 5. Action and state transition of Q-learning

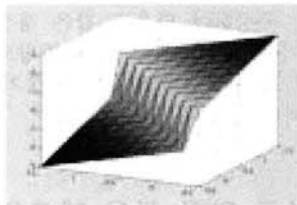


그림 6. 제안된 보상 함수 그래프
Fig. 6. Proposed reward function

상태 그림과 같이 특징 주파수의 역수인 u 와 각도 θ 의 변화로 구성된다. 먼저, $u = 1/f$ 는 +, - 변화량 0.5 단위로 4~30 범위내의 54가지 경우가 있으며, θ 는 +, - 변화량으로 3° 단위로 $\theta - 90^\circ \sim \theta + 90^\circ$ 범위 내에서 60가지 경우가 있다. 따라서 총 3240 개의 상태가 존재하게 된다. 행동(action)은 두 파라미터의 1단위 증가, 정지, 감소의 조합인 9가지를 그림과 같이 가지게 된다. 이러한 학습은 하나의 필터씩 따로 계산된다.

먼저, 사용자로부터의 보상을 다음과 같이 정의한다.

$$r_h : +5 \text{ (맞음)} / -5 \text{ (틀림)} \quad (13)$$

두 번째는, 노드간의 상대적인 비교를 통해 평가하는 부분으로써 다음과 같이 정해진다.

$$r_h \sum_j \frac{(y_{winner} - y_j)}{N} \quad (14)$$

세 번째는, 가장 높은 출력을 낸 노드만을 평가하는 부분으로 다음과 같다.

$$r_h (y_{winner} - 0) = r_h y_{winner} \quad (15)$$

세 가지를 모두 합친 전체 보상은 다음과 같다.

$$r = r_h + \eta r_h \sum_j \frac{(y_{winner} - y_j)}{N} + (1 - \eta) r_h y_{winner} \\ = r_h \left(1 + y_{winner} - \frac{\eta}{N} \sum_j y_j \right) \quad (16)$$

이때, η 는 세 가지 보상간의 균형을 조절하는 파라미터이다.

3.2 적응 2단계 : 적응 분류기 구성

적응 2단계에서는 퍼지 신경망 모델[7]이 사용되며, Q-학습의 결과로 조정된 가버 필터 층으로부터의 특징벡터가 입력으로 사용된다. 학습 순서는 그림 7과 같다.

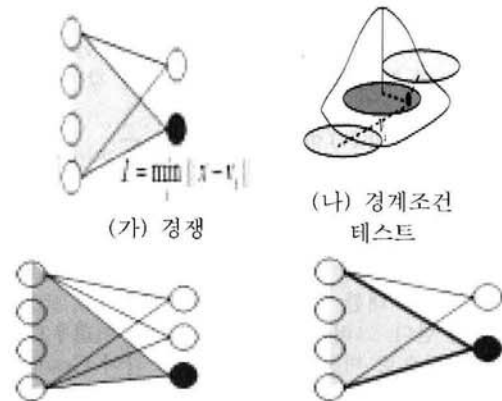


그림 7. 퍼지 신경망 모델의 학습
Fig. 7. Training of Fuzzy Neural Network Model

먼저, 입력 벡터에 대해 가장 가까운 클래스의 중심값을 찾는다.

$$winning\ neuron = \arg \min_j \|v - w_j\| \quad (17)$$

이제, 입력 벡터와 결정된 클래스의 중심값을 이용해서 경계조건 테스트를 수행하게 된다. 이는 입력된 데이터의 위치 정보가 기존의 클래스에 관련이 있는 것인지를 판별하기 위함이다. 이때, u_i 는 클래스들이 모여있는 곳일수록 엄격한 조건을 적용하는 역할을 한다. T는 클래스의 유효범위를 조정해 줄 수 있는 파라미터이다[7].

$$e^{-\gamma u_i} \|x - v_i\| \leq T$$

$$u_i = \left[\frac{1}{\|x - v_i\|^2} \right]^{\frac{1}{m-1}} / \sum_{j=1}^n \left[\frac{1}{\|x - v_j\|^2} \right]^{\frac{1}{m-1}} \quad (18)$$

단, 경계 조건을 넘어서면 새 클래스를 생성하고, 넘어서지 못하면 (19)와 같이 기존의 클래스들의 중심값을 학습시키게 된다[7]. 학습에는 누수 학습 규칙이 사용된다.

$$v_i(n+1) = v_i(n) + f(l)u_i^2(n)\Pi(x, v_i(n), T)(x - v_i(n)) \quad (19)$$

이때, $f(l) = 1/[k(l-1)+1]$ (k 는 상수, l 은 반복횟수),

$$\Pi(x, v_i(n), T) = \begin{cases} 1 - 2\left(\frac{\|x - v_i(n)\|}{T}\right)^2 & 0 \leq \|x - v_i(n)\| \leq \frac{T}{2} \\ 2\left(1 - \frac{\|x - v_i(n)\|}{T}\right)^2 & \frac{T}{2} \leq \|x - v_i(n)\| \leq T \\ 0 & \|x - v_i(n)\| \geq T \end{cases}$$

로 정의된다.

사용되는 퍼지 신경망 모델은 Self-Organizing Feature Map보다 빨리 수렴하고, 초기화 문제가 없으므로, 일정시간의 초기 수렴과정이 끝나면 각 클래스의 중심값으로 수렴함을 뜻하며, 이후 입력되는 데이터마다 온라인 적응 학습이 계속 진행될 수 있다. 온라인 적응시 학습 강도의 조절은 1을 이용하여 조절한다.

4. 실험 및 결과

테스트는 행복, 슬픔, 분노의 3가지 표정에 대해 이루어졌다. 초기 학습을 위해서 42장의 EKMAN DB와 45장의 JAFFE DB를 사용하였다. 그러나, 이 두 DB는 많은 사람들을 포함하는 표정의 수집을 목적으로 만들어진 것이므로, 개인 사용자에게 대한 데이터 개수가 부족하다.

표 1. 학습 및 테스트 DB
Table 1. Training and test DB

	학습DB			테스트DB
	EKMAN	JAFFE	BSCL	BSCL
데이터수	42	90	300	300
사람수	14	10	5	5

따라서 6일 동안 5명의 개인에 대해 PC-CAM을 이용하여 수집된 600장의 BSCL DB를 개인화 성능 테스트를 위해 사용하였다. 처음 3일간의 데이터는 개인화 인식기 구성을 위한 학습에 사용되었고, 이후 3일간의 데이터는 적응 인식기의 성능 테스트를 위해 사용되었다.

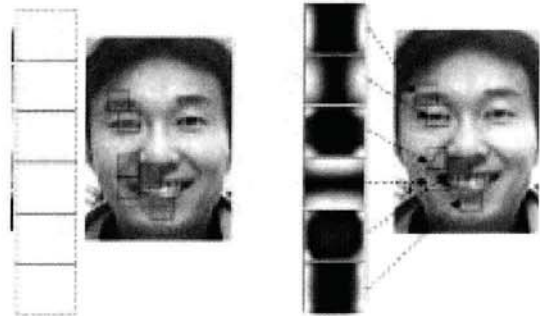
초기 일반적인 인식기 형성방법으로 EKMAN DB에 대해서는 74%, JAFFE DB에 대해서는 89%의 인식결과를 얻었다. 학습된 초기 인식기로 BSCL DB를 측정된 결과는 표 2의 첫 번째 열과 같다(초기 인식기). 보이는 바와 같이 학습에 사용된 개인과 다른 환경, 다른 성향의 입력에 대해서는 대부분의 경우가 50% 이하의 인식률로 EKMAN DB나 BSCL DB보다 현저히 낮은 인식 결과를 보이고 있다.

다시 BSCL DB에 대해서 3절에서 소개하였던 적응 네트워크 구성방법을 적용하여 개인화 인식기를 구성하였다. 먼저 그림 8과 같이 Q-학습을 통해 특징 추출부를 조정하고,

퍼지 신경망 모델의 초기 수렴과정을 통한 클러스터링 결과는 표2의 두 번째 열(Q학습+퍼지신경망 모델을 이용한 클러스터링)과 같다.

표 2. 인식 결과
Table 1. Recognition rate

	초기인식기	Q학습+퍼지신경망 모델을 이용한 클러스터링	이후 테스트 데이터에 대한 적응 인식 결과
사용자 1	33.3	100.0	100.0
사용자 2	25.0	51.7	28.3
사용자 3	58.3	93.3	78.3
사용자 4	33.3	90.0	70.0
사용자 5	33.3	91.7	96.7



(가) 초기 상태 (나) 학습으로 특징 추출부 조정 후
그림 8. Q-학습에 따른 특징 추출부 학습

Fig. 8. Adjustment of feature extractor using Q-learning

3번째 열은 학습에 사용되지 않은 이후 3일간의 데이터를 하나씩 입력시키면서 온라인 적응을 수행한 결과이다. 사용자2의 경우에는 시간에 따른 표정 변화 추적이 실패하였지만, 다른 경우는 비교적 만족스러운 성능을 보이고 있다.

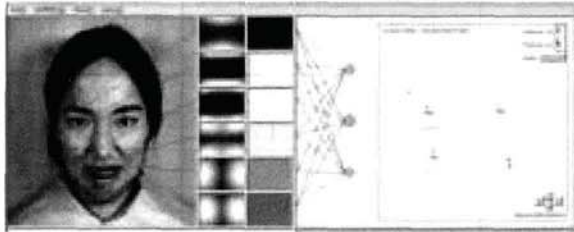
정리하자면, 인식기가 존재할 수 있는 전체 공간을 제대로 포함하지 못하는 작은 양의 학습데이터로 학습시킬 경우, 학습되지 않은 성향의 입력에 대해서는 표2의 첫 번째 열과 같이 제대로 된 인식결과를 내지 못하게 된다. 이때, Q-학습을 이용하여 특징 추출부를 조정하게 되면, 새로운 성향의 입력을 기존 학습기의 입력 공간으로 끌어 올 수 있게 된다. 그리고 인식기는 학습할 수 있는 공간으로 끌려 들어온 데이터를 온라인 학습해 나가게 된다.

학습된 표정인식기를 PC-CAM을 이용하여 PC에서 제작하였다. 표정인식을 위한 얼굴추출은 adaboost 기반 얼굴 추출기[9][10]를 이용하였다. 그림 9의 (가)에서는 입력된 얼굴 영상과 인식 결과 및, 6 차원 특징 벡터의 선택된 두 축으로의 사영을 볼 수 있다. 그림 9의 (나)에서는 얼굴 추출기와 제안된 표정인식 시스템이 연결된 모습을 볼 수 있다.

또한 그림 10과 그림 11에서 볼 수 있듯이, 개인화 단말기(PDA)로도 제작하였다. 이 시스템은 제한된 매개 인터페이스에서도 표정인식이 수행될 수 있으며, 개인화라는 관점에서 바람직하다는 취지를 가진다. 현재까지는 6개의 특징점을 PDA 스틱으로 찍어주어야 한다. Window Mobile 2003에서 구현되었다.

인식과정은 그림 10과 같다. 먼저 원하는 얼굴영상을 읽어 들인 후, 스틱으로 특징점을 찍어주면 학습된 가버 노드를

통과한 출력이 화면에 표시된다. 그 후, 인식결과를 내어주도록 아이콘을 클릭하면 퍼지 신경망 모델의 평가방법에 의해 인식결과가 나오게 된다. 입력된 영상이 각 표정 클래스의 중심에 가까운 순서대로 진한 색으로 표시하도록 하였다. 동시에 연산 속도가 화면에 표시된다.



(가) 가버 웨이블릿 네트워크 인터페이스



(나) 실시간 표정인식 시스템

그림 9. PC에서의 표정인식 시스템

Fig. 9. Facial expression recognition system in PC

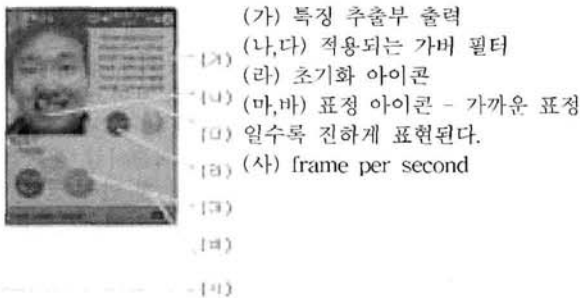


그림 10. 표정인식 시스템의 인터페이스 (PDA)

Fig. 10. Interface of facial expression recognition system in PDA

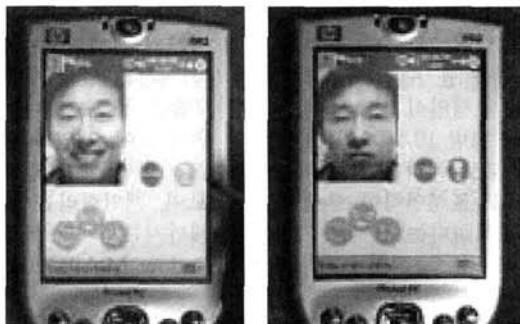


그림 11. PDA 표정인식 시스템

Fig. 11. Facial expression recognition system (PDA)

PDA상에서 영상을 읽고, 인식 결과까지 내어주는 데 소요되는 시간은 약 0.1초로써 PC상에서 소요되는 시간인 약 0.01초와 비교했을 때, 약 10배정도의 속도차이가 난다고 할 수 있다. 그러나, 표정인식 모듈만으로 10fps이므로 빠른 얼굴 검출 시스템과 연결되었을 때, 실시간 검출이 가능하다고 할 수 있다.

5. 결론 및 추후 연구과제

가버 웨이블릿 신경망 표정인식 시스템과 Q-학습과 퍼지 신경망 모델을 이용한 적응 방법이 제안되었다. 기존의 특징 추출을 위한 가버 필터단의 파라미터까지 학습시킴으로써 특징 추출에 있어 사람의 주관성을 배제하고자 하는 목적에서 출발한 것으로서, 학습 과정을 통해 어떠한 특징을 어떻게 뽑아낼 것인지가 모두 결정되는 자동화로서의 첫 단계라 할 수 있다. 현재 기쁨, 슬픔, 놀람의 세 가지 감정에 대한 실험 결과를 보이고 있지만, 인식기의 구조상 적용되는 시스템의 특성과 필요에 따라 원하는 표정의 노드만 늘려주는 것으로 더 많은 감정에 대한 인식기를 구현할 수 있다. 세 가지 표정에 대해서는 EKMAN과 JAFFE DB에 대해 각각 74%와 89%의 일반화 성능을 얻었다. 개인 사용자에 대해서는 특징 추출부의 조정을 통하여 성공적인 인식률을 얻어낼 수 있었다.

이전 연구들이 얼굴의 특징들 사이의 연관성을 고려하였지만, 가장 기본적인 얼굴 특징 추출에 대해서는 심각하게 고려하지 않았다. 그러나 만일, 인식 시스템이 다양한 사용자에 대해 좋은 특징들을 잘 추출해 낼 수 있다면, 또한 새로운 개인 사용자에 적응하기 위해 자동적으로 사용자에 대한 학습이 가능한다면, 인식기의 불필요한 부분을 최대한 줄이면서 최적의 인식기 구현이 가능할 것이다. 실제로 더 많은 특징점을 사용하면, 더 좋은 성능이 예상되지만 본 논문에서는 특징점 학습의 유용성을 보이기 위해 중요도 상위 6개의 특징점만을 선정하였다.

이러한 측면에서 제안된 방법은 세 가지 장점을 가진다. 첫째로 가버 웨이블릿신경망 구조에 기반하여 특징 추출부와 인식기 부분을 학습에서 인식 단계까지 하나로 통합하였다. 또한 특징 추출부로서 가버 웨이블릿 필터를 사용함으로써 사물과 같이 정확한 모서리 정보를 뽑아내기 애매하고 복잡한 얼굴로부터 특징을 효율적으로 뽑아낼 수 있었다.

두 번째로, 전체의 컨벌루션이 없는 간단한 신경망 구조를 가짐으로써, 연산에 따른 부담을 최소화시켰다. 세 번째로, 새 사용자 적용시 비지도 학습이 가능한 퍼지 신경망 모델의 구성이 가능해짐으로써, 개인 사용자에 대하여 자유롭게 적용할 수 있고 사용자의 계속된 성향 변화도 따라갈 수 있도록 하였다.

그러나, 앞으로 추가적인 테스트 수행이 요구되며, 더 많은 개인 사용자의 DB에 대한 폭 넓은 적응성 검증이 필요하다. 현재는 특징 추출부에 BhattaCharyya거리 개념을 고려한 새로운 특징 분리도를 목적 함수로 두고 초기 학습 단계에서부터 비지도 표정 인식을 구현하는 연구를 진행 중이다.

또한 실제로 PC뿐만 아니라, PDA상에서 시스템을 실행함으로써 매개 인터페이스 상에서의 비전기반 인식시스템의 가능성을 확인하였다. 따라서 이러한 장점을 살려 차후 유비쿼터스 환경에서 개인 사용자가 휴대하게 되는 PDA와 같은 작은 장치나, 가정용 로봇 등에 적용할 수 있게 되었다.

참 고 문 헌

[1] Gyu-Tae Park, "A Study on Extraction of Emotion from Facial Image using Soft Computing Techniques", Ph.D Thesis, Dept. of Electrical Engineering and Computer Science, KAIST.1998.

[2] S.B. Gokturk, et al., "Model-Based Face Tracking for View-Independent Facial Expression Recognition", Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 272 - 278.,1998.

[3] A. Kapoor, Q. Yuan, R.W. Picard, "Fully Automatic Upper Facial Action Recognition," IEEE International Analysis and Modeling of Faces and Gestures, pp. 195 - 202., 1998.

[4] L. Franco, A. Treves, "A Neural Network Facial Expression Recognition System using Unsupervised Local Processing," Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis (ISPA 2001), pp. 628 - 632. ,2001.

[5] B. Fasel, "Multiscale Facial Expression Recognition using Convolutional Neural Networks." Proceedings of the third Indian Conference on Computer Vision, Graphics and Image processin, 2002.

[6] B. Fasel, "Robust Facial Analysis using Convolutional Neural Networks," Proceedings of the International Conference on Pattern Recognition, 2002.

[7] Y. S. Kim, C. H. Ham and Y. S. Baek, "A Fuzzy Neural Network Model Solving the Underutilization Problem", Journal of Korea Fuzzy Logic and Intelligent Systems Society, Vol. 11, pp. 354-358. 2001.

[8] Yonsei University et al., Systems for Recognizing and Synthesizing facial Expressions and Gestures, The Report of the Project supported by Ministry of Science and Technology, G17-A-06

저 자 소개



이상완(Sang Wan Lee)

2003년 : 연세대 기계전자공학부 학사 과정 졸업
 2005년 : 한국과학기술원 전자전산학과 석사과정 졸업
 2005년~현재 : 한국과학기술원 전자전산학과 박사과정

관심분야 : 패턴인식, 특징추출, 학습이론
 Phone : +82-42-869-5419
 Fax : +82-42-869-8750
 E-mail : bigbean@ctrsys.kaist.ac.kr



김대진(Dae-Jin Kim)

1997년 : 경북대학교 전자공학과 학사 과정 졸업
 1999년 : 한국과학기술원 전자공학과 석사과정 졸업
 2004년 : 한국과학기술원 전자공학과 박사과정 졸업
 2004년~현재 : 한국과학기술원 인간친화복지로봇연구센터 박사 후 연구원

관심분야 : 패턴인식, 인간-로봇 상호작용, 개인화 시스템, 퍼지 신경망, 지능로봇 시스템 등
 Phone : +82-42-869-8415
 Fax : +82-42-869-8410
 E-mail : djkim@ctrsys.kaist.ac.kr



김용수(Yong Soo Kim)

1981년 : 연세대학교 전기공학과 학사 과정 졸업
 1983년 : 한국과학기술원 전기 및 전자공학과 석사과정 졸업
 1986년 : 삼성전자 종합연구소 주임연구원
 1993년 : Dept. of Electrical Eng., Texas Tech Univ. 박사 과정 졸업

1995년~현재 : 대전대학교 컴퓨터공학부 부교수
 관심분야 : 신경회로망, 퍼지 논리, 패턴인식, 영상처리, 침입 탐지 시스템 등
 Phone : +82-42-280-2547
 Fax : +82-42-284-0109
 E-mail : kystj@dju.ac.kr



변증남(Z. Zenn Bien)

1969년 : 서울대학교 전자공학과 학사과정 졸업
 1972년 : University of Iowa 전자공학과 석사과정 졸업
 1975년 : University of Iowa 수학과 석사과정 졸업
 1975년 : University of Iowa 전자공학과 박사과정 졸업

1976년~1977년 : University of Iowa 조교수
 1981년~1982년 : University of Iowa 객원부교수
 1987년~1988년 : Syracuse University, New York, USA 객원 연구원
 1988년 : 일본 동경공대 객원 교수
 1990년~1995년 : 한국 퍼지 및 지능 시스템 학회장
 2003년~2005년 : 국제퍼지시스템학회(IFSA) 회장
 1977년~현재 : 한국과학기술원 전기및전자공학과 교수
 1999년~현재 : 한국과학기술원 인간친화복지로봇연구센터 소장
 2003년~현재 : 한국로봇공학회 회장
 2005년~현재 : 한국전력 석좌교수

관심분야 : 지능제어, 학습이론, 소프트웨어, 퍼지이론, 서비스로봇
 Phone : +82-42-869-3419
 Fax : +82-42-869-8410
 E-mail : zbien@ee.kaist.ac.kr