
생물학적 특징을 이용한 사용자 인증시스템 구현

문용선* · 정택준*

A study on the implementation of
user identification system using bioinformatics

Yong-seon Moon*, Taeg-jun Jung*

이 논문은 순천대학교 공과대학 학술재단 2000년도 연구비를 지원받았음.

요 약

이 연구는 인식의 정확성을 향상시키기 위하여 단일생체 인식 대신에 얼굴, 입술, 음성을 이용하는 다중생체 인식방법을 제안한다. 각 생체 특징은 다음과 같은 방법으로 찾는다. 얼굴의 특징은 웨이블릿 다중분해와 주성분 분석방법으로 계산하였고, 입술의 경우는 입술의 경계를 구현후 최소 자승법을 이용한 방정식의 계수를 구하였으며, 음성은 멜 주파수에 의한 MFCC를 사용하였으며, 역전파 학습 알고리즘으로 분류하여 실험하였다. 실험을 통해 본 방법의 유효성을 확인하였다.

ABSTRACT

This study will offer multimodal recognition instead of an existing monomodal bioinformatics by using face, lips, voice to improve the accuracy of recognition. Each bioinformatics vector can be found by the following ways. For a face, the feature is calculated by principal component analysis with wavelet multiresolution. For a lip, a filter is used to find out an equation to calculate the edges of the lips first. Then by using a thinning image and least square method, an equation factor can be drawn. A voice recognition is found with MFCC by using mel frequency. We've sorted backpropagation neural network and experimented with the inputs used above. Based on the experimental results we discuss the advantage and efficiency.

키워드

biometrics, recognition, identification, multimodal

1. 서론

현대사회는 정보화라는 미명하에 아주 급격히 발전하고 있다. 정보는 지식이나 경험을 담고있는 데이터베이스 일수도 있고, 또한 어느 회사의 중요한 시스템 또는 장비일수도 있으며, 가장 개인적으로는 개인 전용 컴퓨터에 들어있는 자료를 일컫는다. 외부의 유출 혹은 간섭 없이 보관되어져야 할 정보는 여러 가지 이유로 유실 혹은 도난 당해질 수 있다. 이러한 이유로 보안에 관한 역할이 강조되어지고 있다.

보안에 관한 부분 중 신원을 확인하는 인증에 대한 분야가 강조되고 있다. 사람에게 있어서 본인임을 판단할수 있는 기준은 세 가지, 즉 알고 있는 것, 가지고 있는것, 지니고 있는 것으로 나눌 수 있는데 알고 있는 것의 예를 든다면 패스워드 등 자기의 기억에서 찾아 낼수 있는 것을 일컬으며, 가지고 있는 것은 마그네틱 카드, 스마트 카드 등이 좋은 예가 될 것이다. 세 번째의 지니고 있는 것은 지문이나 홍채등 몸에 지니고 있는것을 일컫는다. 이 중에서 가장 높은 보안등급을 갖는 것은 몸에 지니고 있는 것, 즉 생물학적인 특징을 이용한 인식방법이다. 사람은 그 개개인이 다른 이와는 전적으로 구분되는 생물학적인 특징을 가지고 있다. 얼굴을 비롯해 지문, 눈동자, 손, 혈관 인식은 이미 실용화된 상태이고, 음성, 서명, 망막, DNA, 체온, 귀, 냄새에 이르기까지 인식 기술이 다양하게 확대되고 있다. 이러한 개개인의 생물학적 고유성을 이용하여 자동으로 그 사람의 신원을 인식하는 시스템은 생물학적인 특징의 높은 신뢰도 때문에 오랜 기간 많은 연구가 진행되어 왔다. 그러나 개별적인 생체 인식은 각 생체 인식의 한계 때문에 인식율이 완벽하게 보장되지 않는다는 단점이다.

본 논문에서는 단일 생체인식 방법의 단점을 보완하기 위해 다중 생체인식의 방법, 즉 얼굴 특징, 입술 특징, 음성 특징을 이용하여 역전파 신경망으로 분류, 인식시스템을 구현하였으며 실험을 통하여 유효성과 인식율에 대하여 검토하였다.

본 논문의 순서는 제1장은 최근의 생체인식 기술의 흐름과 장·단점과 인증 시스템에 대하여 기술하고, 제2장은 시스템의 전체 구조와 각 생체 인식의 동향을 살펴보고, 제3장은 시스템의 각 생체 인식별 알고리즘과 특성을 언급하며 제4장은 시스템을 구성하여 실험

한 결과를 제시하며 5장에서는 본 시스템이 사용될 분야와 결론을 언급하였다.

II. 생물학적 특징의 구조

2.1 시스템 구조

생물학적인 특징을 이용하는 인식 시스템은 단일 생체 인식과 다중 생체 인식으로 나누어지는데 현재는 멀티미디어 환경에서 시스템이나 정보의 접근에 허가를 주고 인식율을 높이기 위한 다중 생체 인식 방법의 연구가 많이 행해지고 있으며 M2VTS(MultiModal Verification for Teleservices and Security application)에 관한 연구사례는 그 대표적인 예로 볼 수 있다^[1].

다중생체인식을 위한 BIOD(BIOinformatics IDentification)의 시스템 구조는 그림1 과 같이 각 생물학적인 특징의 입력, 특징추출, 인식의 과정을 거치는 것을 기본으로 한다.

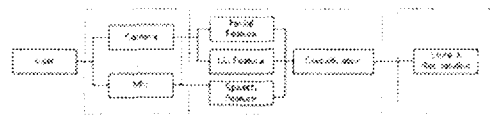


그림 1. 생체인식 시스템
Fig. 1 Biometric identification system

그림 1에서는 입력을 받을 수 있는 장치를 카메라와 마이크, 타블렛, 지문을 감지할 수 있는 입력 장치 등 생물학적 특징을 획득할 수 있는 장치는 모두 입력 부분에 속한다. 인간이 가지고 있는 모든 생물학적 특징에는 각각의 고유한 특성을 가지고 있어서 얼굴 부분은 웨이블릿에 기반한 주성분 분석방법, 윤곽선 추출등에 의해서, 입술 부분은 입술의 윤곽에 의한 방정식의 계수 또는 입술의 화소값을 이용하는 방법 등을 통해서, 음성은 LPC(Linear Predict Coefficient), MFCC(Mel Frequency Cepstral Coefficient)를 통해 각각의 생물학적 특징을 구하게 되며, 특징을 통합하여 분류하는 과정은 유클리드 거리, 신경망 등을 거쳐서 그 자료를 저장하고 저장용 마친 데이터들은 다음에 입력되는 생물학적 특징의 데이터와 비교를 통하여

시스템에 대한 인증의 유·무를 결정하게 된다.

2.2 생체 특징 추출

얼굴 특징을 추출하여 자동으로 얼굴 인식을 하는 것은 많은 지식을 요구하는 꽤 어려운 일이다. 얼굴 특징 추출의 궁극적인 목적으로는 높은 정확성과 실시간 접근을 목적으로 한다. 얼굴 특징을 추출하기 위해서는 얼굴의 특징 파라미터 요소 예를 들면 눈썹, 눈, 코, 입 등을 이용하는 방법과 얼굴의 농도, 색채 변화를 이용한 화소값을 이용하는 방법이 있다.

본 연구에서는 얼굴 영상의 특징 추출을 위한 방법으로 화소값을 이용한 웨이블릿을 기반으로 한 주성분 분석 방법을 이용하는데 웨이블릿 기술은 최근에 신호 및 영상 압축, 다해상도 분석, 음성 동기, 컴퓨터 비전, 그래픽등을 포함한 영상처리 및 신호처리 분야에서 각광을 받고 있다.

정합 방법인 주성분 분석 방법의 기본적인 개념은 더 낮은 차원에서의 특징 벡터를 가진 입력 데이터를 새로운 공간에서 표현하는 것이다. 이러한 주성분 분석 방법은 성분들이 최상의 복원을 제공, 성분들이 서로 상관관계에 있지 않으면 변환이 선형적이어서 얼굴 특징 추출에 많이 사용되어지고 있다^{[2],[3]}.

얼굴의 특징 중에는 눈썹, 눈, 코, 입술 등 여러 가지 요소가 존재하며 이러한 특징 요소들 중 입술의 특징에 근거한 연구가 상당수 진행되어 오고 있는데, 입술의 특징을 추출하기 위한 접근법은 히스토그램이나 화소값에 기반한 영상 변환 기반 접근법과 에지 정보를 이용하는 입술 윤곽선 기반 방식, 이 두 가지로 나눌 수 있는데 영상변환 기반 접근법은 그 인식율이나 입술 특징 파라미터 추출에 있어서 입술 윤곽선 기반 방식보다 안정적이고 우수하다고 알려져 있다. 그러나 입술 윤곽선 기반 방식은 입술의 특징 파라미터를 처리하므로 파라미터 추출에서 인식하는 때까지 비교적 적은 데이터 처리량을 갖는 장점을 가지고 있다^[4].

본 논문에서는 입술 윤곽선 기반 방식을 이용, 입술 영상을 세선화시켰을 때 나오는 입술의 윤곽이 방정식의 특성을 가진 그래프로 구해질 수 있는 점을 착안, 추정치를 이용하여 방정식으로 근사화시킬 수 있는 최소자승법을 이용하여 입술의 방정식 계수를 유도해 낸다.

음성 특징 추출은 음성 신호로부터 신호의 특징을 나타내 주는 파라미터를 추출하는 과정이다. FFT(Fast

Fourier Transform)를 비롯한 여러 분석 방법 중 주파수의 잡음을 효과적으로 감소시키는 MFCCC(Mel Frequency Cepstral Coefficient)를 이용하는 방법은 음성 특징 추출의 방법으로 도입되면서 많은 연구가 진행되어져 왔다. MFCC는 음성 주파수를 mel scale을 통하여 적절한 계수를 구하여 인지 가능한 주파수 크기로 변환하는데 그 목적이 있는데 다른 접근방법에 비해서 통계적 특성이 좋고, 간단한 유클리드 거리만으로도 두 신호의 스펙트럼 왜곡을 측정할 수 있는 등의 장점 때문에 음성인식에서 주로 사용하는 파라미터가 되어 있으며, 실제로 인식율면에서도 보다 나은 성능을 보이고 있다^[5].

본 논문에서는 음성 특징 추출을 위한 방법으로 mel-scale을 이용하는 MFCC를 구하여 음성의 특징을 추출한다.

여러 생체 인식 시스템은 효율적인 분류를 위해 여러 방법을 사용하여 왔는데 그 중 순환 신경망의 일종인 역전과 학습 알고리즘은 복잡한 비선형 함수를 모델링하는 기술로서 함수를 나타내는 식을 모를 때에도 사용가능하기 때문에 현재 신경망 알고리즘 중에서 가장 많이 쓰이고 있는 알고리즘이다. 역전과 학습 알고리즘은 출력층 오차신호를 이용하여 은닉층과 출력층 간의 연결강도를 변경하는 방법으로서 학습에는 시간이 오래 걸리지만 일단 학습이 끝나고 응용단계에서는 결과가 빠르게 나온다^[6].

본 논문에서는 역전과 알고리즘을 이용하여 분류 시스템을 구현한다.

III. Bioid 알고리즘

3.1 얼굴 특징 추출

푸리에변환이나 DCT (Discrete Cosine Transform)와는 달리 주파수와 시간 정보를 동시에 가지고 있는 웨이블릿의 다해상도 분해를 이용하여 영상을 변환하게 되면 보다 많은 영상의 정보를 가질 수 있게 된다.

본 논문에서는 웨이블릿 다해상도 분해방법을 이용하여 영상을 처리하고 주성분 분석 방법을 이용해 얼굴 영상의 특징을 추출해 내게 되는데 이는 주성분 분석 방법의 한계인 이미지의 데이터가 증가할 때 인식의 정확성은 증가하나 많은 계산식이 요구되는 것을

웨이블렛의 특징인 이미지 데이터를 축소시켜도 원래 이미지의 정보가 많이 손실하지 않는 것을 이용하여 주성분 분석 방법의 한계를 극복하는 방법을 얼굴 특징 추출 알고리즘으로 이용하였으며^[7] 이를 위한 알고리즘 구성도는 그림 2와 같다.

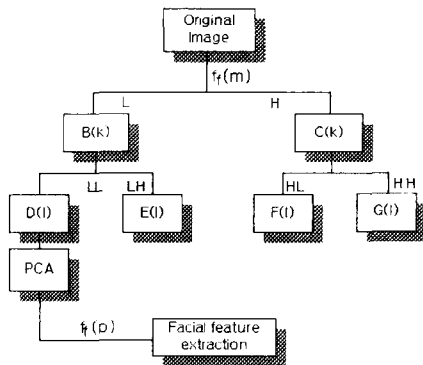


그림 2 . 얼굴 특징 추출 알고리즘 구성도
Fig. 2 Facial feature extraction algorithm

이산 웨이블렛 변환은 크기 와 위치를 연속으로 보여주지 않고 이산적인 방법으로 보여주는 변환으로 식 (1)과 같이 표현된다.

$$\psi_{j,k}(t) = \frac{1}{\sqrt{s_0^j}} \psi\left(\frac{t - k\tau_0 s_0^j}{s_0^2}\right) \quad (1)$$

식 (1)에서 j, k 는 정수, scaling factor $s_0 > 1$, τ_0 는 translation factor 이다.

식 (1)을 근거로 하여 크기 스펙트럼과 위치 스펙트럼에 웨이블렛 스펙트럼을 덧붙이면 다해상도 분해를 한 후 함수 $f(t)$ 의 j번째 항을 구하게 되면 식 (2)와 같이 표현할 수 있다.

$$\begin{aligned} \phi(t) &= \sum_k \lambda_j(k) \varphi(2^j t - k) \\ \psi(t) &= \sum_k \gamma_j(k) \varphi(2^j t - k) \end{aligned} \quad (2)$$

$\lambda_j(k)$, $\gamma_j(k)$ 는 haar, daubechies 계수로 나타낼 수 있으며 식 (2)에서 0번째 항에 대하여 정리하면

식 (3)으로 표현될 수 있다.

$$\begin{aligned} \phi(t) &= \sum_k L \varphi(2t - k) \\ \psi(t) &= \sum_k H \varphi(2t - k) \end{aligned} \quad (3)$$

식 (3)에서 k,l은 1부터 n까지를 나타내며, L,H는 haar, daubechies 계수를 나타낸다.

CCD 카메라를 이용하여 캡처 한 얼굴 이미지의 신호를 $f_f(m)$ 이라고 하면 위의 식 (3)을 이용하여 식 (4)와 같은 다해상도 분해식을 사용할 수 있다.

$$\begin{aligned} B(k) &= \sum_m L(2k - m) f_f(m) \\ C(k) &= \sum_m H(2k - m) f_f(m) \end{aligned} \quad (4)$$

식 (4)를 적용시킨 함수 $f_f(m)$ 을 저주파 부밴드 대역(LL)을 선택하여 주성분 분석 방법으로 특징을 추출하는 식은 (5)과 같다.

$$f_{p(d)} = A(X - M_x) \quad (5)$$

식 (5)에서 X는 $f_f(m)$ 의 성분이며, M_x 는 $f_f(m)$ 의 평균, A는 고유값이다.

얼굴특징을 추출하는 최종식 (5)에서 나온 값을 신경망을 통해 분류를 하기 위해 데이터 베이스에 저장한다.

3.2 입술 특징 추출

입술을 추출한 후 대각선방향으로 특성이 좋은 소벨 연산자를 이용하여 윤곽선 검출을 실시하고 잡음을 최소화 시키기 위하여 이진화 방법을 이용한다. 실험치에서 정형화된 방정식을 구하기 위해서는 추정치를 최대화시킬 수 있는 최소자승법을 이용하여야 한다. 입술의 특징을 추출하기 위한 알고리즘 구성도는 그림 3과 같이 표현되어진다.

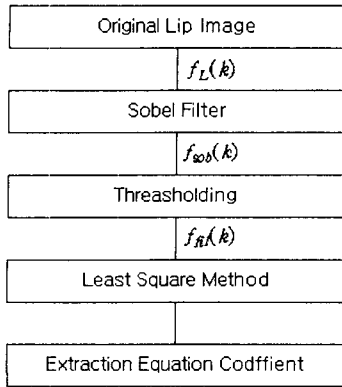


그림 3. 입술 특징 추출 알고리즘 구성도
Fig 3 . Lip-line feature extraction algorithm

입술 영역의 이미지 신호 $f_L(k)$ 을 소벨 연산자와 이진화 기법을 이용하여 필터링을 한 신호는 $f_{th}(k)$ 이다. 여기서 아랫입술을 구하기 위해 최소자승법을 이용하는데 먼저 입술 경계선에 대한 좌표값을 구하여 식 (6)의 x_j, y_j 에 대입한 후 식 (7)에서 보이는 것처럼 미분을 취한다.

$$s = \sum_{j=1}^n (y_j - a - b x_j - c x_j^2)^2 \quad (6)$$

$$\frac{\partial s}{\partial p} \Big|_{p=a, b, c} = 0 \quad (7)$$

식 (7)를 이용해 미분을 취한 값은 가우스-조던 소거법을 이용하여 아랫 입술 방정식의 계수를 구할 수 있다.

윗 입술의 방정식의 계수를 구하기 위하여 역시 최소 자승법의 확장을 이용하였으며 방법은 먼저 식 (6)을 사차 방정식으로 확장하며 그 확장식은 식(8)과 같다.

$$s = \sum_{j=1}^n (y_j - a - b x_j - \dots - e x_j^4)^2 \quad (8)$$

식(8)을 이용해 아랫입술을 구하는 방법과 마찬가지로 식 (9)에서 보이는 것처럼 미분을 취한다.

$$\frac{\partial s}{\partial p} \Big|_{p=a, b, c, d, e} = 0 \quad (9)$$

식 (9)를 이용해 미분을 취한 값은 가우스-조던 소거법을 이용하여 윗 입술의 방정식의 계수를 구할 수 있다.

아랫 입술과 윗 입술의 계수를 구하는 최종식 (7), (9)에서 나온 값을 입술 특징점으로 사용하기 위해 데이터 베이스에 저장한다.

3.3 음성 특징 추출

음성 데이터로부터 음성의 특징을 추출하기 위해서는 잡음에 대하여 주파수의 특성이 강한 MFCC를 이용하였으며 MFCC를 위한 알고리즘 구성도는 그림 4와 같다.

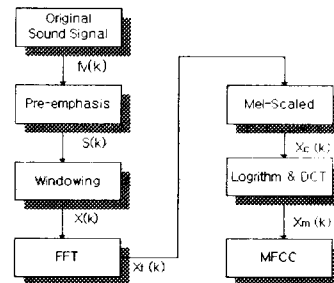


그림 4. 음성 특징 추출 알고리즘 구성도
Fig. 4 Sound feature extraction algorithm

화자가 말한 것을 마이크로폰을 통해 입력받은 신호는 $f_v(k)$ 이다. 음성의 고주파 영역을 강조하기 위해 pre-emphasis과정을 거치게 되는데 수식은 식(10)과 같으며, α 값은 $0.9 \leq \alpha \leq 0.97$ 사이의 값을 갖는데 아래의 α 값은 실험치중에서 가장 적당한 값을 선택한 것이다.

$$S(k) = f_v(k) - \alpha f_v(k-1) \quad (10)$$

식 (10)에서 $\alpha = 0.95$ 이다.

전처리 과정을 거친 신호 $S(k)$ 는 윈도우링 과정을 거치게 되는데 여기서 윈도우는 해밍 윈도우 알고리즘을 도입하였으며 수식은 (11)과 같다.

$$X(k) = S(k)(0.54 + 0.46 \cos n \frac{\pi}{M}) \quad (11)$$

식 (11)에서 M은 윈도우의 길이를 나타내고, n은 1 부터 M 까지의 정수이다.

윈도잉과정을 거친 신호 X(k)는 FFT를 통해 Xf(k) 신호를 출력시키게 된다.

신호 X(K)에 대한 FFT 과정을 거친 식은 식(12)로 표현할 수 있다.

$$\begin{aligned} X_f(k) &= \sum_{n=0}^{N-1} X_n W_N^{kn} \\ &= \sum_{n=even} X_n W_N^{kn} + \sum_{n=odd} X_n W_N^{kn} \\ &= \sum_{r=0}^{\frac{N}{2}-1} X_{2r} W_{\frac{N}{2}}^{kr} + W_N^k \sum_{r=0}^{\frac{N}{2}-1} X_{2r+1} W_{\frac{N}{2}}^{kr} \end{aligned} \quad (12)$$

식(12)에서 k와 n은 1부터 N-1 까지의 정수 값이고, $X_n = X(k)$ 이며

$$W_n = e^{-\frac{j2\pi}{N}n}, W_{\frac{N}{2}} = W_{\frac{N}{2}}, W_N^k = e^{(-\frac{j2\pi}{N})kn}$$

으로 나타낸다.

식(12)를 통해 나온 신호 Xf(k)를 mel-scale 형태로 변환하기 위한 식은 (13)과 같이 표현 할 수 있다.

$$X_m(k) = 2595 \log_{10} \left(1 + \frac{X_f(k)}{700} \right) \quad (13)$$

마지막으로 mel-scale 형태로 변환한 신호 Xm(k)를 DCT알고리즘을 적용한 식은 (14)과 같다.

$$X_c(k) = \frac{1}{N} \sum_{n=0}^{N-1} X_m(k) \cos\left(\frac{k2\pi n}{N}\right) \quad (14)$$

음성특징인 MFCC를 구하는 최종식 (14)에서 나온 값을 음성 특징점으로 사용하여 분류를 하기 위해 데이터 베이스에 저장한다.

4.4 분류

Mcculloch와 Pitts에 의하여 인공 신경망이 제안된 이후 많은 종류의 신경망들이 여러 형태의 문제에 적

용되고 있으며 특히 공학적으로는 영상 인식, 음성 인식 등의 패턴 인식을 비롯하여 여러 인식 시스템과 특징 벡터의 분류 등에 사용되어 지고 있다^[8].

신경망을 이용한 인식 과정이나 분류 과정은 모든 인식 과정의 마무리 단계로서 중요한 역할을 담당하는데 본 논문에서는 여러 제안된 신경망 알고리즘 중 시그모이드(Sigmoid)를 전달함수로 갖는 역전파 알고리즘을 사용하게 되는데 역전파 알고리즘은 델타 학습 법칙의 일종으로 훈련 패턴 각각에 대해 정확한 목표 출력값을 학습에 이용하는 방법인데 미분가능한 전달 함수(시그모이드 함수)를 가진 은닉층에 LMS(Least Mean Square) 방법을 확장하여 적용시키는 알고리즘이다. 역전파 신경망을 이용한 생물학적 특징 분류의 구성도는 그림 5와 같다.

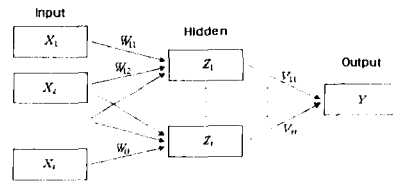


그림 5. 생물학적 특징 분류 알고리즘 구성도
Fig. 5 Bioinformatics classification algorithm

입력층의 입력 벡터를 $X = [x_1, x_2 \dots x_n]$, 은닉층 벡터를 $Z = [z_1, z_2 \dots z_n]$, 출력층 벡터를 Y, 입력층과 은닉층의 연결강도를 W, 은닉층과 출력층과의 연결강도를 V 라고 하면 학습을 위한 은닉층은 식 (15)로 표현 할 수 있다.

$$Z = \frac{1}{1 + e^{-XW}} \quad (15)$$

식 (15)에서 XW 는 벡터의 곱을 나타낸다. 출력치를 구하기 위해 식 (16)을 이용한다.

$$Y = \frac{1}{1 + e^{-ZV}} \quad (16)$$

연결강도의 변화량을 구하기 위한 출력층의 오차는 식 (17)에 의해서 구해진다.

$$\delta_y = (d - y)y(1 - y) \quad (17)$$

식 (17)에서 d는 목표치, y는 출력치 이며 은닉층의 오차는 식 (18) 에 의해서 구해진다.

$$\delta_z = z(1 - z) \sum_{j=1}^m \delta_j v \quad (18)$$

식 (17)과 식 (18)을 이용해 k번째 항에서의 은닉층과 출력층간의 연결강도 변화량 Δv^k , 입력층과 은닉층간의 연결강도 변화량 Δw^k 를 구하면 식 (19), 식(20) 과 같다.

$$\Delta v^k = \alpha \delta_y z \quad (19)$$

$$\Delta w^k = \alpha \delta_y x \quad (20)$$

식 (19), (20)에서 α 는 학습률이며, k+1 번째의 연결강도 변화량과 k 번째의 연결강도 변화량이 똑같아질때까지 학습을 한다.

식(19)과 식(20)을 이용하여 나온 값을 인식을 위한 비교 요소로 참고하기 위해 저장한다.

N. 시스템의 구현 및 결과 고찰

4.1 시스템 구성

생물학적인 특징의 샘플은 카메라와 마이크를 통해 각 샘플을 획득하게 된다. 카메라를 통해 획득한 이미지 샘플은 얼굴 인식을 위한 얼굴 추출과 입술 영역 추출, 두가지 프로세싱을 위해 각 영역이 추출되어지며, 마이크를 통해 획득된 음성 샘플은 음성 특징 추출을 위해 사용되어진다. 이후 각 특징은 역전파 신경망에 의해 분류 되어져 저장되고, 다시 인식을 위해 입력 받은 생물학적 특징들은 위와 똑같은 과정을 거쳐 특징이 추출되어 저장값과 비교하여 결과를 추출해 내게 된다. 영상 데이터를 저장 받기 위해서 DT3155 이미지 보드를 이용하여 영상을 획득 한 후 PentiumIII 450Mhz ,128M RAM을 이용하여 처리하였다. 또한 음성 데이터는 야마하 DS-XG 음성 획득 장치로부터 8-kHz sampling rate , 8-bit resolution으로 입력받아 동일한 시스템을 이용하여 처리하였다. 이러한 처리

과정을 거친 샘플들은 분류를 거쳐 데이터베이스에 저장된다. 각 저장된 값들은 분류 과정을 거쳐 원래의 값과 비교 후 결과를 처리하게 된다. 인증 구현 시스템은 그림 6과 같이 나타낼 수 있다.

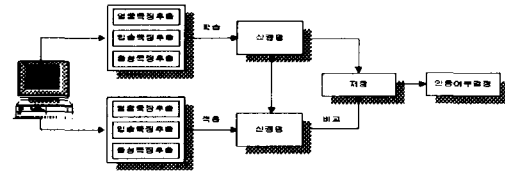


그림 6. 인증 구현 시스템의 전체 구성도
Fig. 6 Identification system

4.2 얼굴 특징 추출

얼굴 특징 추출의 실험 결과는 실험 대상자중 한 명의 데이터 하나를 토대로 작성하였으며 OS는 Windows 2000, CCD 카메라, DT-3155 이미지 보드 그리고 C++를 이용하여 시스템을 설계하고 구현하였으며 히스토그램 비교를 위해 matlab을 이용하였다.

DT 3155 image board는 640×480과 같은 3 : 2의 크기 비율로 이미지를 캡처하기 때문에 먼저 영상을 640×480의 크기로 영상을 캡처한 후 마우스로 256×256 크기의 얼굴영상을 선택하여 추출하게 된다.

그림 7은 CCD 카메라와 DT 3155 보드를 이용하여 캡처한 영상중 얼굴 특징 추출을 위하여 256×256 크기로 얼굴 영역을 선택하여 나온 데이터를 웨이블릿 변환의 다해상도 분해에 관한 식 (4)를 적용시키고 계수는 haar 계수를 이용하여 얻은 결과이다.

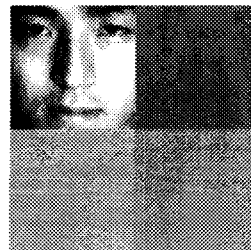


그림 7. 다해상도 분해를 한 얼굴 이미지
Fig. 7 Multiresolution analysis face image

그림 7에서 나온 데이터에서 LL 성분만을 다시 추출하여 웨이블릿 다해상도 분해를 취하여 다섯 번째 모 웨이블릿을 구할 때 까지 위의 과정을 계속 반복하여 8×8 형태의 행렬을 구하고 이것을 이용하여 고유값을 구하여 주성분 분석 방법을 통하여 분석한 행렬은 표 1로 나타낼 수 있으며 이 데이터 값을 역전파 신경망의 입력으로 사용하기 위해 저장한다.

표1. 주성분 분석 후의 행렬
Table 1. Matrix after principal component analysis

-139	-109	-83	31	15	-64	-69	-36
-132	-92	-24	-40	73	-2	-67	-24
-63	-2	-8	-70	47	54	67	70
-2	57	21	-63	61	74	103	109
-43	69	-24	-31	80	70	110	103
-77	20	-8	-21	56	80	108	83
-105	-30	-59	-28	33	21	69	38
-141	-72	-54	-5	41	40	44	-17

4.3 입술 특징 추출

그림 8은 CCD 카메라와 DT 3155 보드를 이용하여 캡처한 영상중 입술 특징 추출을 위하여 90×30 크기로 영역을 선택한 후 에지검출 후 이진화시켜 얻은 영상이다.



그림 8. 원래의 입술영상 과 필터 후의 영상
Fig 8. Original lip image and lip image after filtering

필터를 취한 영상의 입술 라인을 따라 식 (9), (11)에 의하여 연산을 취한 결과 아랫 입술의 계수와 윗 입술의 계수를 구한 후의 방정식은 식(21), (22)와 같다.

$$L_b(x) = 0.003587x^2 + 0.269816x - 3.728390 \quad (21)$$

$$L_t(x) = -0.840386x^4 + 1.135255x^3 + 0.001461x^2 + 0.000027x + 0.000001 \quad (22)$$

식(21), (22)의 방정식 계수는 역전파 신경망의 입력으로 사용하기 위해 저장한다.

4.4 음성 특징 추출

음성 특징 추출을 위한 실험도 얼굴 특징 추출 실험과 마찬가지로 실험 대상자중 한 명의 데이터 하나를 토대로 작성하였으며 OS는 Windows 2000, 녹음장치, 야마하 DS-XG 음성 획득장치를 이용하였으며, C++를 이용하여 프로그램을 구현하였으며 히스토그램 비교를 위해 matlab을 이용하였다. 녹음 장치를 통하여 8kHz, 8bit mono의 형태로 음성을 입력받은 음성신호의 파형은 그림 9로 표현 할 수 있다.

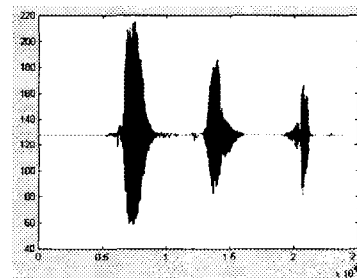


그림 9. 입력된 음성신호
Fig. 9 Input speech signal

그림 9와 같이 입력된 음성 신호를 그림 6에서 언급했던 음성 특징 추출 알고리즘인 MFCC 알고리즘을 이용, 연산을 하여 음성 신호의 특징을 추출해 내는데 그 순서는 그림 4에 나타내어지는 구성도의 순서로 표현되어지며 연산을 마친 신호는 그림 10과 같이 표현 된다.

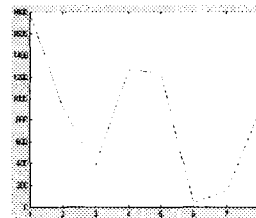


그림 10. MFCC 후의 음성 특징
Fig. 10 Speech feature after MFCC

알고리즘 순서에 의하여 연산을 한 결과는 그림 10에서 보여지듯이 8개의 값으로 나타내어지며 이는 인식을 위한 역전파 신경망의 입력으로 사용하기 위해 저장한다.

4.5 신경망을 이용한 결과 추출

얼굴 특징 추출 과정을 통해 얻어진 결과값인 8×8 형태의 행렬에 입술 특징 추출 과정을 통해 얻은 윗 입술을 표현한 사차 방정식의 계수 5개, 아랫 입술을 표현한 이차 방정식의 계수 3개, 음성 특징 추출과정을 통해 얻은 MFCC 계수 8개 총 80개의 데이터를 입력으로 하여 그림 7과 같이 구성하며 이때 은닉층과 출력층의 전달함수는 단방향 시그모이드 함수를 이용하여 학습한다.

4.6 고찰

각각의 특징 벡터 추출을 위한 알고리즘의 실행 정상 여부를 판단하기 위해 먼저 얼굴 영상 특징 벡터를 위한 웨이블릿 알고리즘 으로 나온 연산 결과를 역 웨이블릿으로 확인한 결과 영상의 정상 복원이 되었고, 입술 특징 벡터는 계수를 이용한 방정식을 그래프로 그려본 결과 정상적인 입술 모양이 출력되는 것을 확인할 수 있다. 이러한 데이터를 토대로 실험을 시행한 결과는 다음과 같다.

6명의 각각 10개의 데이터를 토대로, 5개는 인식을 위한 기준을 찾기 위한 데이터로 사용하고, 나머지 5개는 인식의 유·무를 확인하는 데이터로 사용한다. 인식 유·무를 확인하기 위한 기준 데이터의 설정은 실험 결과에 유연성을 가하기 위해서 신경망 실험 결과에 따른 은닉층과 출력층사이의 연결강도 5개를 평균과 표준편차를 구하여 평균에서 표준편차의 차를 최소값으로 정하고 평균과 표준편차의 합을 최대값으로 정하여 비교 데이터 입력시 연결강도의 범위를 위의 최소값과 최대값으로 정한다. 이를 토대로 실험 대상자들의 실험 결과 인식 유무를 확인하면 표 2와 같다.

표 2. 실험 결과
Table 2. Result of experiment

사람 \ 횟수	1	2	3	4	5	Total
A	×	○	○	○	○	80 %
B	○	○	○	×	○	80 %
C	○	○	○	○	○	100 %
D	○	×	○	×	○	60 %
E	○	×	×	○	○	60 %
F	×	○	○	○	×	60 %
Total						73.3 %

V. 결론

본 논문에서는 생물학적인 특징중 얼굴, 입술 그리고 음성을 통한 인증 시스템을 구현하였는데 본 시스템에서는 얼굴 특징은 웨이블릿을 이용한 다해상도 분해를 통하여 특징을 추출한 후 얼굴의 고유값을 구하여 주성분 분석방법으로 추출하며, 입술 특징은 소벨 연산자를 이용한 필터와 문턱값을 이용한 이진화 필터로 입술 윤곽을 추출한 후 최소자승법을 통하여 윗 입술과 아랫 입술의 방정식을 추출하고 음성특징은 MFCC 방법을 통하여 추출하며 역전파 신경망을 이용하여 분류, 학습시킨 결과 인식에 대한 유효성을 확인하였다.

본 논문에서는 얼굴, 입술, 음성을 개별적으로 프로세싱하는 시스템을 구현하였지만, 앞으로는 실제로 많이 사용되어지는 컴퓨터용 카메라와 마이크를 이용하여 PC 환경에서 통합화되고 일반화된 시스템을 구현하고 더 향상된 생물학적인 특징을 이용하여 인식율을 높이는 방법을 연구를 계속 해야 될 것으로 사료된다.

참 고 문 헌

- [1] G.Richard, Y.Mengay, etc "multimodal Verification for Teleservices and Security Application" 0-7659-0253-9/99, IEEE
- [2] Ki Chung Chung, Seok Cheol Kee and Sang Ryong Kim "Face recognition using principal component analysis of gabor filter response" 0-7659-0378-0/99 IEEE

- [3] Vic Brennan and Jose Principe "Multiresolution using principal component analysis" 0-7803-6293-4/00 IEEE 2002년 3월 ~ 현재 순천대학교 전자공학과(박사과정)
※관심분야 : Biometrics, 인터넷 응용
- [4] 김진영, 김진범 "서브워드 단위 음성 인식을 위한 영상 자동 립리딩" 전자통신 기술논문지, Vol.3, No.1, 2000
- [5] R. Vergin, D. O'Shaughness, A. Farhat "Generalized mel frequency cepstral coefficients for large-vocabulary speaker independent continuous-speech recognition" IEEE transactions on speech and audio processing, vol,7, no.5, september 1999
- [6] 김병기, 정낙우 "패턴 인식 성능을 향상시키는 새로운 형태의 순환신경망" 한국정보처리학회 논문지, Vol.4, No.2, 1997
- [7] PC. Yuen, D. Q. Dai and G. C. Feng "Wavlet-based PCA for human face recognition" 0-7803-4876-7-1/98 IEEE
- [8] 원용관, 백용창, 이정수, "목표물의 고속 탐지 및 인식을 위한 효율적인 신경망 구조" 한국정보처리학회 논문지 제4권 제10호(97.10)

저자 소개



문용선(Yong-Seon Moon)

1983년 2월 조선대학교 전자공학과(공학사)
1985년 2월 조선대학교 전자공학과(공학석사)
1989년 2월 조선대학교 전자공학과(공학박사)

1992년 3월 ~ 현재 순천대학교 전자공학과 부교수
※관심분야 : Biometrics, Vision-based control



정택준(Taeg-jun Jung)

2000년 2월 순천대학교 전자공학과(공학사)
2002년 2월 순천대학교 전자공학과(공학석사)