

論文98-35C-5-7

폐쇄회로 카메라를 이용한 신분 확인용 실물 얼굴인식시스템의 설계

(Design of a Face Recognition System for Person Identification Using a CCTV Camera)

李銓佑*, 成孝慶**, 金星玩***, 崔興文**

(Jeun-Woo Lee, Hyo-Kyung Sung, Sung-Oan Kim, and Heung-Moon Choi)

요 약

본 논문에서는 제한구역에 대한 효율적인 접근제어를 위하여 얼굴의 대칭조건을 이용한 얼굴영역 검출기(face region detector)와 감각 연결세기(sensory synapses)와 표현 연결세기(descriptive synapses)를 갖는 확장된 자기 조직화 지도 신경회로망(extended self-organizing maps: *ESOM*)을 동시에 사용한 신분 확인용 실물 얼굴 인식시스템을 제안하였다. 제안한 방법에서는 복잡한 배경을 가진 입력 영상에서 일반화 Hough 변환으로 추출된 타원 영역에 수직 및 수평 투영을 하여 얼굴의 근사적 타원 대칭성을 이용함으로써 모든 얼굴영역을 추출 및 판별하였다. 또한 주성분 분석 및 탄성적인 유사도 정합이 가능한 *ESOM*을 제안하고, 이를 얼굴의 소속원 여부를 인증하기 위한 얼굴 영역인식에 적용하였다. 한편, 인증된 소속원 얼굴에 대해서는 인증 당시의 그 얼굴에 대한 표현 데이터 베이스(facial description database)를 *ESOM*의 추가 학습 기능을 이용하여 온라인 구축함으로써 소속원 얼굴의 표정 변화, 안경 착용 유무 등에 대처할 수 있도록 하였다. 실험을 통하여 제안된 방법이 제한구역에 대한 출입제어에 효과적임을 확인하였다.

Abstract

We propose an efficient face recognition system for controlling the access to the restricted zone using both the face region detectors based on facial symmetry and the extended self-organizing maps (*ESOM*) which have sensory synapses and descriptive synapses. Based on the visual cues of the facial symmetry, we apply horizontal and vertical projections on elliptic regions detected by GHT(Generalized Hough Transform) to identify all the face regions from the complex background. And we propose an *ESOM* which can exploit principal components and imitate an elastic similarity matching, to authenticate faces of the enlisted member. In order to cope with changes of facial expression or glasses wearing, etc, the facial descriptions of each member at the time of authentication are simultaneously updated on the descriptive synapses online using the incremental learning of the proposed *ESOM*. Experimental results prove the feasibility of our approach.

I. 서론

접근 제어에 있어서 보안등급에 따라서는 개인별로 유일 불변의 신체 특징 중의 하나인 홍채(iris) 또는 지문(fingerprint)의 특징을 근접 또는 접촉 검출하여 신분을 확인할 수 있으나, 일반적으로 단순한 출입 통제에 경우에는 신분증 소지자의 진위 여부와 그에 따른 출입허가 여부를 판별하고, 신분증의 불법 취득 또는 위조를 통한 불법 출입을 막기 위한 간편한 방법으로 얼굴 인식(face recognition)을 통한 인증(au-

* 正會員, 韓國電子通信研究所

(Electronics and Telecommunications Research Institute)

** 正會員, 慶北大 電子電氣工學部

(School of Electronics and Electrical Engineering, Kyungpook National University)

*** 正會員, 水原 專門大 事務自動化科

(Dep. Office Automation, SuWon Junior College)

接受日:1998年2月25日, 수정완료일:1998年5月4日

thentication)이 널리 이용되고 있다^[1,2,5,6]. 이 때 인간의 시각으로는 복잡한 배경 내에 존재하는 얼굴도 쉽게 찾아내고 구분·인식도 용이하지만, 컴퓨터 시각을 이용할 경우에는 사람의 얼굴이 복잡한 패턴이기 때문에 얼굴의 자동 확인은 어려운 문제 중의 하나이다. 따라서 최근 보안 시스템(security systems), 범인 확인(criminal identification), 기계와의 상호작용(man-machine interaction) 등 여러 응용 분야에서 얼굴 인식 문제를 해결하려는 노력이 활발히 이루어지고 있다^[1-2].

지금까지 템플릿 정합(template matching), KL확장(Karhunen-Loeve expansion), 고유얼굴(eigen-faces), 등밀도 지도(isodensity maps), 그리고 신경 회로망(neural networks) 등 얼굴 인식에 대한 많은 방법과 수단들이 제시되어 제한된 영상획득 환경하에서는 상당히 좋은 인식 결과를 보였다^[1-2]. 이들 방법에서는 획득된 영상내에 존재하는 배경으로부터 얼굴의 위치를 찾아내고, 얼굴의 세부 구성요소를 분리·인식하는 과정을 거쳤다. 일반적으로 영상내에서 얼굴의 위치 및 크기는 모두 미지의 인수이므로 그 위치 및 구성 요소를 추출하기는 쉽지 않다. 또한 얼굴 인식에 대한 일부 연구는 단순 또는 일정 배경하에서 영상을 획득하거나 또는 획득된 영상내에 인식할 얼굴이 하나인 경우로 한정하는 경우가 많다^[1-2]. 한편, Govindaraju과 Lee 등의 연구에서는 획득된 미지의 영상 내에 있는 얼굴의 위치를 정확하게 자동 검출하기 위하여 얼굴의 황금비(golden ratio)^[3], 움직임(motion)과 얼굴 색(color)정보^[4] 등의 사전정보(apriori information)를 이용하고 있으며 Yang 등의 연구에서는 복잡한 흑백(black-and-white) 영상에서 여러 해상도의 모자이크(mosaic) 영상과 얼굴 구성요소에 대해 미리 정의된 법칙(rule)들과의 상호작용(interactive)을 이용하여 얼굴의 위치를 정확히 찾아내려 하였다^[10].

그러나 얼굴 인증을 통한 접근제어(출입통제)의 경우 제한구역의 입구, 통로 또는 로비(lobby) 등 폐쇄 회로 카메라(CCTV)가 설치된 현장의 배경 영상이 일정하고 조명이 비록 고정되어 있더라도, 접근해오는 얼굴의 흔들림(shaking)과 회전(rotating)에 따라 얼굴에 비치는 조명의 국부적인 밝기 및 각도가 변하여 얼굴과 배경사이의 낮은 대비(contrast), 불규칙한 밝기 확산(intensity diffusion) 또는 얼굴 주위에 원하

지 않는 그림자 등 영상 획득 환경이 수시로 변화하므로 입력된 영상에서 얼굴의 위치를 찾는 그 자체부터가 쉽지 않다^[3-6]. 또한 증명 사진이나 명함판 사진 영상과는 달리 영상을 획득하는 과정에서 얼굴 표정의 변화, 얼굴의 크기 변화, 안경의 착용 유무, 머리 모양의 변화 등 소속된 얼굴 자체의 변화 등이 있을 수 있다.

따라서 제한구역에 대한 출입통제를 효과적으로 수행하기 위해서는 복잡 배경으로부터의 얼굴 위치의 효과적인 추출방법 뿐만 아니라 현장 조명 조건이나 확인 대상 얼굴의 표정, 위치 및 각도 등이 변화하는 환경이나 조건하에서의 신뢰성 있는 얼굴 인증방법에 대한 연구가 중요하다.

본 논문에서는 제한구역에 대한 효율적인 접근제어를 위하여 얼굴의 대칭조건을 이용한 얼굴 영역 검출기와 감각 연결세기와 표현 연결세기를 갖는 확장된 자기 조직화 지도 신경회로망(extended self-organizing maps: *ESOM*)을 동시에 사용한 신분 확인용 실물 얼굴 인식(얼굴 인증)시스템을 제안하였다. 제안한 방법에서는 복잡한 배경을 가진 입력 영상에서 모든 얼굴 영역의 추출 및 판별을 위하여 일반화 Hough 변환^[7]으로 추출된 타원 영역에 얼굴의 타원 대칭성, 수직 및 수평 투영을 사용하였다. 또한 소속원의 얼굴인지를 인증하기 위하여 주성분 분석 및 탄성적인 유사도 정합이 가능한 *ESOM*을 제안하고, 얼굴 영역에 이를 적용하였다. 한편, 소속원 얼굴의 표정 변화, 안경의 착용 유무 등에도 대처할 수 있도록 *ESOM*의 추가 학습기능을 이용하여 얼굴 인증과 동시에 인증된 얼굴들에 한해서 당시의 얼굴 표현에 관한 정보를 *ESOM* 자체에 재구축하였다. 조명의 변화, 표정의 변화, 얼굴의 크기, 얼굴의 방향(좌우 또는 상하 움직임) 등이 다양하게 변화된 얼굴들에 대해 실험하여 제안된 출입 통제 방법의 타당성을 확인하였다.

II. 접근 제어를 위한 얼굴 인증

얼굴 인증에 의한 접근제어에서는 소수의 출입이 허가된 소속원들의 신분정보 데이터와 얼굴 데이터베이스를 구축해 놓고, 출입구로 접근해 오는 사람의 얼굴 영상을 카메라로 획득·인식하여 소속원의 얼굴 여부를 인증한다. 본 연구에서 제안한 얼굴 인식에 의한 접근제어 시스템은 그림 1과 같이 얼굴 영역 추출부와

ESOM에 의한 특징 추출 및 얼굴 인증부로 나누어지며, 얼굴 특징 추출부에서는 얼굴이 근사적 타원 대칭성을 갖는다^[3]는 점을 이용하여 입력된 영상에서 일반화 Hough변환을 이용하여 얼굴 후보 영역을 추출하고, 얼굴에서 나타나는 기울기 대칭성 및 얼굴 후보 영역의 수직 및 수평 투영을 이용하여 추출된 후보 영역이 얼굴 영역인지를 확인하였다. 또한 ESOM의 감각 연결세기를 사용하여 추출된 얼굴영역에 대하여 주 성분 분석 및 탄성적인 유사도 정합을 통하여 얼굴을 인증하였으며, 동시에 인증된 소속원 얼굴들에 대한 표현을 ESOM의 감각 및 표현 연결세기에 구축함으로써 별도의 과정없이 입력되는 얼굴들 중 기존 소속원 얼굴들의 변화 또는 새로 추가되는 소속원 얼굴들에 대하여 추가적 학습이 가능하도록 하였다.

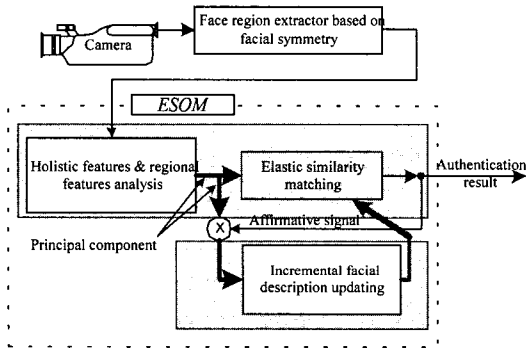


그림 1. 접근 제어를 위해 제안한 얼굴 인증시스템
Fig. 1. Schematic of the proposed face authentication system for access control.

1. 얼굴 영역 추출

얼굴 확인에 의한 효율적인 접근제어를 위하여 현장에 설치된 카메라로부터 획득한 영상내에 존재하는 여러 얼굴들의 확인 및 복잡한 배경을 가진 밝기값(gray level) 영상에서 색정보나 움직임에 대한 정보 없이 빠른 시간에 얼굴들을 추출하기 위하여 타원 대칭조건을 가지고 얼굴 영역을 추출하였다. 본 논문에서는 일반화 Hough 변환^[7]을 이용하여 획득된 영상내에 존재하는 모든 타원 물체들을 추출하고, 이 때 파생되는 밝기값 기울기(intensity gradient)영상에서 얼굴의 타원 대칭성 및 수직 및 수평 투영을 사용하여 추출된 후보 영역이 얼굴 영역인지를 판별하였다.

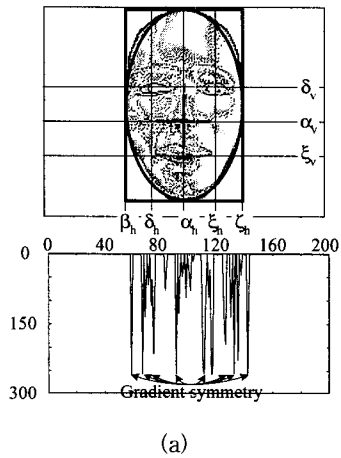
임의 화소의 밝기값 기울기 크기를 $g_k = \nabla I_k = \|\partial I_k / \partial x, \partial I_k / \partial y\|$ 라 하고, 그 방향을 $a_k = \alpha(\partial I_k / \partial x, \partial I_k / \partial y)$ 라고 정의하면, 얼굴과 정합되는 타원의 장축

을 기준으로 동일한 각 α 만큼 떨어진 두 화소 i 및 j 사이에 존재하는 화소의 밝기값 기울기 대칭 정도 $s_{2\alpha}(i, j)$ 및 국부적인 기울기 대칭정도 $s_\alpha(i, j)$ 는 각각 다음과 같다.

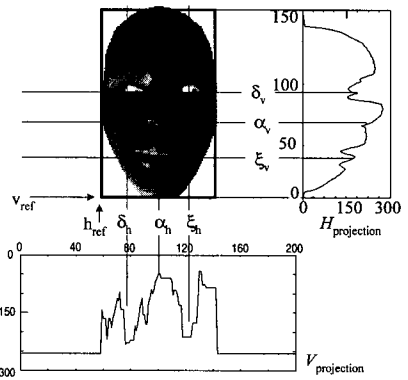
$$s_{2\alpha}(i, j) = g_i * \tan(\alpha_i) + g_j * \tan([\pi - \alpha_j])$$

$$s_\alpha(i, j) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{\|s_{2\alpha}(i, j)\|^2}{2\sigma^2}\right)$$

여기서 $s_\alpha(i, j)$ 는 $[0, 1]$ 이며, σ 는 대칭성을 추출하기 위한 관심영역의 크기로서, 추출된 타원에 최소 외접 사각형을 씌워 관심영역을 설정하였다.



(a)



(b)

그림 2. 얼굴 영역 확인을 위한 특징 (a) 두 눈을 가로지르는 선상의 국부적인 기울기 대칭성과 (b) 얼굴 영역의 수평 및 수직 투영

Fig. 2. Features for face verification. (a) Local gradient symmetry along the eye line and (b) the horizontal and the vertical projections in a face region

얼굴의 여러 대칭성질 중 추출된 얼굴의 두 눈을

가로지르는 선을 따라 구한 국부적인 기울기 대칭성을 그림 2(a)에 도시하였다. 또한 일반적으로 얼굴의 중방향으로는 눈과 입 주위의 밝기 기울기 변화가 심하게 나타나며, 눈과 코사이의 간격이 코와 입사이의 간격과 대략 같거나 크고, 횡방향으로는 두 눈 주위의 기울기 변화가 심하게 나타난다. 이러한 얼굴의 성질을 이용하여 그림 2(b)에 도시한 바와 같이 추출된 타원 영역에 수평 및 수직 투영을 행하여 얼굴 영역을 쉽게 판정할 수 있도록 하였다. 따라서 그림 2(a)에 나타난 것과 같이 추출된 타원의 중심을 지나는 장축을 기준으로 그 대칭 정도가 $s_c(i, j) \geq 0.8$ 이고, 그림 2(b)와 같은 수평 및 수직 투영에서 대략 $\alpha_v \approx (\delta_v + \xi_v)/2$, $\alpha_h \approx (\delta_h + \xi_h)/2$ 이 되는 경우 추출된 후보 영역을 얼굴 영역으로 판정하였다. 여기서 눈, 코 및 입의 수직 위치 δ_v, α_v , 및 ξ_v 는 v_{ref} 를 기준으로 구한 수직 거리를 나타내며, 두 눈과 코의 수평위치 δ_h, α_h 및 ξ_h 는 h_{ref} 를 기준으로 구한 수평거리를 나타낸다.

2. ESOM에 의한 얼굴 인증 및 얼굴 표현 구축

ESOM은 그림 3과 같이 입력층(input layer), 실행층(action layer) 및 표현층(description layer)으로 구성되며, 입력층과 실행층사이에는 주성분 추출을 위한 감각 연결세기로 연결되고, 실행 층과 표현 층사이에는 출입이 허가된 얼굴들에 대한 세부적인 표현 정보를 저장하는 표현 연결세기로 연결되어 있다. 감각 연결세기는 Kohonen이 제안한 자기 조직화 지도(self-organizing maps)의 학습^[8]을 이용하여 주성분 분석(principal component analysis) 및 유사도 정합(elastic similarity matching)이 가능하도록 하였으며, 표현 연결세기는 시각 운동계(oculomotor system)의 단속적인 보정운동(corrective saccade)개념^[9]을 도입하여 세부적인 주성분 분석 및 유사도 정합이 가능하도록 하였고, 유사도 정합단계에서 호의적인 결과를 보여 인증된 얼굴들에 한해서 그 입력된 얼굴표현에 관한 정보를 ESOM 자체에 재구축하였다.

ESOM에서는 입력 얼굴과 감각 연결세기에서 구축된 인증 얼굴에 대한 주성분과 비교하여 인증 여부를 결정하며, 인증된 얼굴 영상에 대하여 표현 연결세기를 갱신한다. 한편 거절된 영상에 대하여 표현 연결세기에 구축된 보정 주성분과의 비교를 통하여 인증 여부를 다시 한번 검증하며 이 때 인증 신호가 발생하면

ESOM의 연결세기가 갱신된다. 즉, 입력된 얼굴은 각각 연결세기 및 표현 연결세기와의 두 번 비교를 통하여 인증 및 거절이 최종 결정된다. 이 때 ESOM은 망막의 정량적인 응답 특성^[9]을 모방하여 그림 4와 같이 와(fovea) 영역을 나타내는 중심부의 확률 밀도가 0으로 움푹 들어간 $\sigma_r = 4$ 인 변형된 정상분포(modified normal distribution)의 확률로 인증 또는 학습을 수행하도록 하였다. 얼굴 인증을 위한 와영역의 반경 A_{fovea} 는 1.0, 얼굴 표현에 대한 학습 종료를 위하여 와 반경 L_{fovea} 를 $0.5 * A_{fovea}$ 로 하였다.

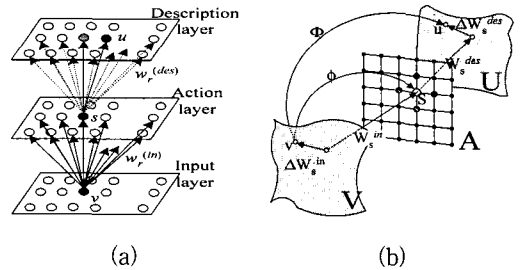


그림 3. ESOM의 (a) 구조와 (b) 위상 모델
Fig. 3. ESOM. (a) architecture and (b) topological model

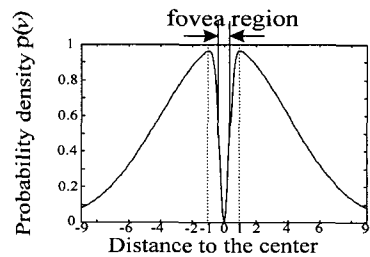


그림 4. 유사도 정합을 위한 ESOM의 입력 확률 밀도
Fig. 4. Input probability density for similarity matching in ESOM.

그림 3(a)의 입력층 V 에 있는 와 영역 중심에서 영상점으로서의 입력 벡터 v에 대하여 최대 유사도 M_{max} 를

$$M_{max} = \|v - w_s^{(in)}\| \leq \|v - w_r^{(in)}\|, \quad \forall r \in A \quad (2)$$

로 정의하고, 이 때 실행층 A의 뉴런 s를 승리 뉴런으로 선택하며, $w_s^{(in)}$ 및 $w_r^{(in)}$ 은 입력층과 실행층사이 연결된 승리 뉴런의 연결 세기 및 기타 뉴런의 연결세기를 각각 나타낸다.

최대 유사도가 $M_{max} \leq A_{fovea}$ 되면 입력된 얼굴의 출입을 허가하고, 불허되면 실행층의 뉴런 s에 연결된

표현 연결세기 $w_s^{(des)}$ 에 2차 인증을 행하도록 하였다. 이 때 $\|v + w_s^{(des)}\| \leq A_{fovea}$ 이면 출력을 허가하고 그렇지 않으면 출력을 최종 불허한다. 이 때 출력이 허가된 얼굴에 대하여 얼굴의 표정 변화, 안경의 착용 유무 등에 적응적으로 대처할 수 있도록 *ESOM*의 각 연결 세기에 대하여 다음과 같은 얼굴 표현 갱신 과정을 수행하도록 하였다.

영상 점 v 에 대한 실행층 A 상의 승리 뉴런을 s , 학습율을 ϵ , 학습 반경을 h_{rs} 라 하면, 감각 연결세기 w 를

$$w_r^{(in, new)} = w_r^{(in, old)} + \epsilon h_{rs}(v - w_r^{(in, old)}) \quad (3)$$

와 같이 갱신한다.

ESOM 학습 알고리즘

0. 감각 연결세기 $w_r^{(in)}$ 및 표현 연결세기 $w_r^{(des)}$ 을 임의의 값으로 초기화한다.
1. 입력에 대한 확률 밀도 $p(v)$ 에 따라 새로운 입력 v 를 선택한다.
2. 입력 층 V 의 영상 점 v 에 대한 실행 층 A 의 승리 뉴런 s 를 식 (2)에 정의된 M_{max} 조건을 만족하게 선택한다.

유사도 정합 단계(similarity matching phase)

- 3.a. $M_{max} < A_{fovea}$ 일 경우
출력을 허가함과 동시에 얼굴표현 갱신과정을 수행한다.
- 3.b. $M_{max} \geq A_{fovea}$ 일 경우
 - 3.b.1 $M_{max} = \|v + w_s^{(des)}\| < A_{fovea}$ 이면 출력 허용하고, 얼굴표현 갱신단계를 수행한다.
 - 3.b.2 그렇지 않으면 단계 1을 수행한다.

얼굴표현 갱신단계(facial description updating phase)

4. 승리 뉴런 s 에 대하여 식(3)으로 정의된 학습 방법으로 감각연결세기를 학습한다.
5. 단속적인 학습 $w_s^{(des)}$ 를 행하여 입력 영상 점 v 를 $v' = v + w_s^{(des)}$ 로 변경한다.
- 6.a. $\|v'\| < L_{fovea}$ 가 되면 단계 1을 수행한다.
- 6.b. $\|v'\| \geq L_{fovea}$ 되면 실행단내의 모든 뉴런들에 대하여 $\|v' - w_s^{(des)}\| \leq \|v' - w_r\|$ 관계를 가지는 뉴런들 중에서 입력 벡터 v' 와 새로운 승리 뉴런 s' 를 선택한 후 보정학습 $w_s^{(des)}$ 를 행한다
 - 6.b.1 $\|v' - w_s^{(des)}\| \leq \|v'\|$ 이 되면 식(4)로 정의된 보정학습을 행한 후 단계 1을 수행한다.
 - 6.b.2 $\|v' - w_s^{(des)}\| > \|v'\|$ 되어 보정 효과가 없으면 곧 바로 단계 1을 수행한다.

그림 5. *ESOM* 학습 알고리즘
Fig. 5. *ESOM* learning algorithm.

이 때 $\|v - w_s^{(in)}\| \geq L_{fovea}$ 이면 *ESOM*은 영상 벡터 v 에 대한 승리 뉴런 s 에 대한 보정 학습을 위하여 표현 연결세기를 $w_s^{(des)}$ 를 도입하여 *ESOM*을 강화 학습시킨다. 표현 연결세기 $w_s^{(des)}$ 에 의해 영상벡터 v 는 새로운 영상벡터 $v' = v + w_s^{(des)}$ 로 변경된다. 이 때, $\|v'\| \leq L_{fovea}$ 되면 $w_s^{(des)}$ 을 영상 벡터 v 에 대한 표현 연결 세기로 결정하며, 그렇지 않으면 $\|v + w_s^{(des)} + w_s^{(des, mf)}\| < \|v + w_s^{(des)}\|$ 관계가 성립하는 다른 표현 연결세기 $w_s^{(des, mf)}$ 를 선택한 후 다음과 같이 학습한다.

$$u = w_s^{(des, old)} + w_s^{(des, mf)} \quad (4)$$

$$w_r^{(des, new)} = w_r^{(des, old)} + \epsilon' h'_{rs}(u - w_r^{(des, old)})$$

여기서, u 는 $w_s^{(des)}$ 의 추정 보정량을 나타내며, h'_{rs} 는 학습 반경을 나타내고, ϵ' 는 학습률을 나타낸다. *ESOM*의 학습 알고리즘은 그림 5와 같다.

*ESOM*은 비교사 학습(unsupervised learning)방법으로 위상 보존 사상관계(topology-conserving mapping)를 유추하는 감각 연결세기 $w_r^{(in)}$ 과 보정 학습을 위한 표현 연결세기 $w_r^{(des)}$ 를 사용함으로써 최적의 얼굴 표현 데이터 추출을 위한 보정 학습 정도를 뉴런 공간내에서 자체적으로 유추할 수 있을 뿐만 아니라, 표현 연결세기 갱신에 학습 반경개념을 도입함으로써 *ESOM*을 안정되게 수렴시킬 수 있도록 하였다.

III. 실험 및 고찰

제안한 방법의 타당성을 확인하기 위하여 본 연구실에서 카메라를 이용하여 얼굴의 표정, 방향, 크기 및 조도를 달리하면서 만든 실물 얼굴 데이터 베이스 *KPU1*), 매사추세츠 대학의 미디어 연구실에서 얼굴 방향, 조도 및 얼굴의 크기를 달리하면서 만든 얼굴 데이터 베이스 *MIT* 및 올리베티 연구소에서 조도 변화 및 얼굴의 표정을 달리하면서 만든 얼굴 데이터 베이스 *ORL*의 세가지에 대하여 얼굴 영역 추출 실험 및 얼굴 인증실험을 행하였다.

그림 6 및 그림 7은 복잡한 배경속에 얼굴이 하나만 있는 영상과 얼굴이 두 개 이상 있는 영상에 대한 얼굴 영역 추출 실험 결과이다. 그림 6에서(a)는 복잡

1) 이 데이터 베이스는 "http://eeg.kyungpook.ac.kr/~pdp"에서 구할 수 있다.

한 배경에 얼굴이 하나만 있는 얼굴 영상을, (b)는 타원추출을 통하여 추출된 얼굴 후보영역을, (c)는 얼굴의 대칭성을 이용하여 추출된 얼굴 영역, 그리고 (d)는 얼굴 영역 확인을 위한 수직 및 수평 투영을 도시한 것이다. 그림 7은 여러 개의 얼굴이 있는 영상에서 얼굴 영역을 추출한 결과이다. 그림 6(d)와 그림 7(c)에 표시된 '+' 기호들은 추출된 얼굴 영역에서 구한 두 눈, 코, 입 영역의 중심위치들이다.

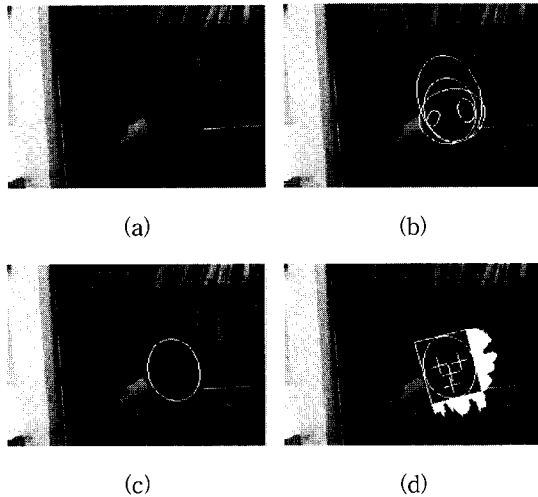


그림 6. 복잡한 배경에서 얼굴 영역 추출: (a) 촬영한 실 영상 (b) HT로 추출한 타원 영역들, (c) 추출된 얼굴 영역, 그리고 (d) 얼굴 영역 확인을 위한 수평수직 투영

Fig. 6. Face region extraction from the complex background. (a) captured image, (b) elliptical regions by HT, (c) extracted face region, and (d) horizontal and vertical projections

카메라를 이용하여 얼굴 형태, 머리 모양, 나이, 성별, 인종 등 다양한 변화속에 구축된 실물 얼굴 데이터 베이스 KPU, MIT, 그리고 ORL에 대하여 얼굴 영역 추출 실험을 수행한 결과 데이터 베이스 KPU와 ORL에서는 추출 거부된 영상이 없었으나, 데이터 베이스 MIT 전체 432개 얼굴 영상 중 배경에 비하여 얼굴 영역의 크기가 상대적으로 작은 14개의 얼굴 영상, 즉 약 3.2%의 얼굴에 대해서 얼굴 영역 추출 거부가 발생하였다. 한편, 언급한 3가지 얼굴 데이터 베이스에 대하여 얼굴 영역 추출과 얼굴 인증을 실험하였다. 이 때 ESOM에 소속원의 얼굴로 미리 학습된 얼굴의 수는 KPU, MIT, ORL 데이터 베이스에서 임의로 각각 6, 8 및 6인의 정면 얼굴, 좌우 및 상

하로 향하는 얼굴별로 각각 1개씩을 사용하였다. 한편 ESOM을 학습할 때 학습 인수는, 감각 연결세기 학습에서는 $\epsilon(t)=1/(1+t/t_{max})$ 및 $\sigma(t)=10 \cdot \exp(-t/t_{max})$ 로, 표현 연결세기 학습에서는 $\epsilon'(t)$ 및 $\sigma'(t)$ 를 $\exp(-(t/t_{max})^2)$ 로 하였으며, t_{max} 는 200,000으로 설정하였다.

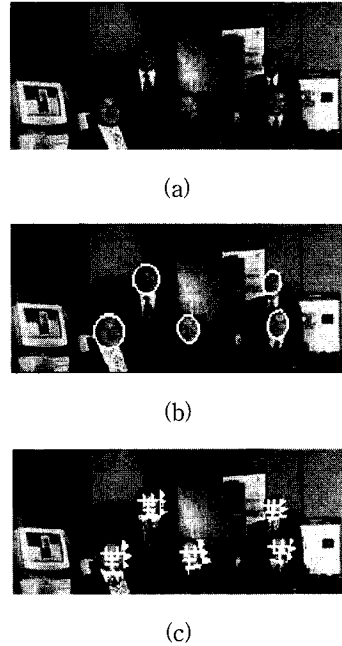


그림 7. 복잡한 배경 영상에서 다수의 얼굴영역 추출 (a) 실 영상 (b) 추출된 얼굴영역들, 그리고 (c) 각 얼굴 영역 확인을 위한 수직 및 수평 투영

Fig. 7. Multiple face regions extraction from the image with complex background. (a) captured image, (b) extracted face regions, and (c) horizontal and vertical projections

그림 8은 얼굴 데이터 베이스 KPU 중 ESOM에 미리 학습된 소속원들의 정면 얼굴이며, 그림 9는 인증을 위한 실험에 사용한 KPU 중의 얼굴영상 예들이다. 또한, 입력 영상의 주성분(principal component) 추출을 위해 구성되는 기저함수는 ESOM의 가중치 공간상에 존재하며, 그림 10은 A_{fovea} 를 1.0, L_{fovea} 0.5인 ESOM을 사용하여 모든 KPU 얼굴들에 대해 실험한 후 형성된 얼굴 영역과 얼굴 구성 요소 영역들이 이러한 기저 함수를 통과한 후 ESOM이 구성하는 2 차원 특징지도상으로의 투영(projection)된 결과를 도시한 것이다.



그림 8. 출입허가를 가진 소속원의 정면 얼굴영상 예
Fig. 8. Examples of frontal face images of the access-allowed members.

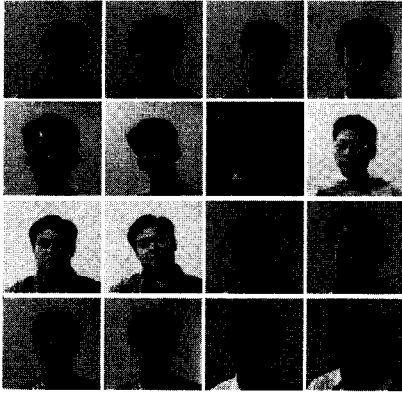


그림 9. 접근제어 환경에서 인위적으로 획득한 실험영상의 예(얼굴 데이터베이스 KPU)
Fig. 9. Examples of experimental face images artificially acquired from the access control environments. (KPU face database)

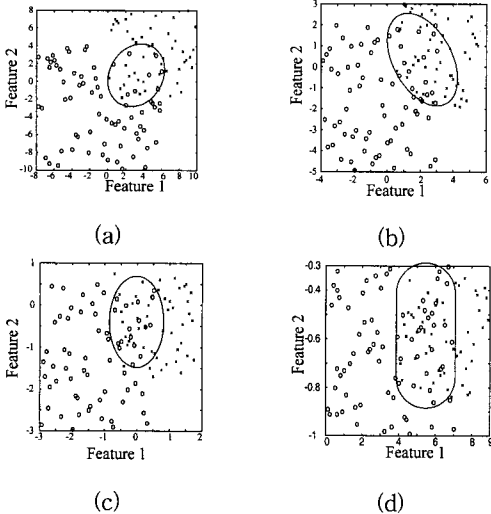


그림 10. 실험 영상 데이터베이스들의 확인에 사용한 주성분들 중 (a) 얼굴 전체, (b) 눈, (c) 입, 그리고 (d) 코에 대한 첫 번째 및 두 번째 주성분 분포

Fig. 10. Distribution of the first and the second principal components among the principal components of images from the experimental face database: Distribution of features in (a) face itself, (b) eyes, (c) mouth, and (d) nose.

그림 10에서 feature 1과 feature 2는 ESOM 특징지도상의 좌표축을 의미한다. 그림에서 'o'표시는 출입이 허가된 경우이고 'x'는 출입이 거절되는 경우를 표시하며, 가운데 영역 표시는 오 인증이 발생하는 영역을 표시한 것이다. 그림에서 알 수 있는 바와 같이 코에 의한 얼굴 변별력이 가장 떨어지며, 얼굴 전체에 의한 변별력이 비교적 우수하게 나타났다. 표 1은 각 데이터 베이스에 대하여 A_{fovea} 를 바꾸면서 실험한 결과이다. 이 때 ESOM에서 M_{max} 가 $A_{fovea} \leq M_{max} \leq \min(1.0, 2 * A_{fovea})$ 범위일 때는 인식 거절로 판단하였다.

표 1에서 전체 소속원과 비소속원들의 얼굴들에 대하여 Right항은 올바르게 인증된 얼굴들에 대한 결과를, False항은 잘못 인증된 얼굴에 대한 결과를, Reject항은 인증이 거부된 얼굴들에 대한 결과를 각각 나타낸다. R_{Ra} 는 소속원의 총 얼굴 수에 대하여 호의적으로 인증된 소속원 얼굴의 백분율을, R_{Rd} 는 비소속원의 총 얼굴 수에 대하여 출입이 거절된 비소속원 얼굴의 백분율을 의미한다. 또, R_{Fa} 는 비소속원의 얼굴수에 대한 출입이 허가된 비소속원 얼굴의 백분율(false acceptance rate)로, R_{Fd} 은 소속원의 얼굴수 대 출입이 거절된 소속원 얼굴의 백분율(false denial rate)을 나타낸다. 한편, R_{Ri} 는 인증이 거부된 소속원 얼굴수의 전체 소속원의 얼굴 수에 대한 백분율(rejection rate for member)로, R_{Ro} 은 인증이 거절된 비소속원의 얼굴수의 전체 비소속원 얼굴 수에 대한 백분율(rejection rate for non-member)을 나타낸다.

표에서 올바르게 인증되는 Right항의 값이 높아야 되겠지만, 접근제어에서는 무엇보다도, 비소속원임에도 불구하고 출입이 허가된 비소속원에 대한 결과인 R_{Fa} 와 인증이 거절되는 R_{Ri} 및 R_{Ro} 의 수치가 낮아야 한다. A_{fovea} 의 크기가 줄어들면, 표 1에서 알 수 있는 바와 같이 카메라를 이용하여 얼굴 표정, 얼굴 방향 및 얼굴의 크기 등을 달리하면서 획득한 데이터베이스 KPU, MIT, ORL 등에 대한 실험에서 A_{fovea} 가 1.0 일 때에는 인증이 거절되는 R_{Ri} 및 R_{Ro} 가 발생하지 않는 반면에 약 4~5%정도의 비소속원이 오인증 받아 출입이 허가되는 결과를 보였으며, 아울러 소속원이 약 4.0~6.9%정도 출입이 거절되었다. 본 연구에서는 폐쇄회로 화면으로부터 입력되는 표정 변화가 심하고 15° 기울어진 얼굴까지도 포함하여 실험하였기

표 1. 각 얼굴 데이터 베이스에 대한 얼굴 인증률(백분율)
Table 1. Face authentication rate for each face database(%).

A_{fovea}	KPU						MIT						ORL					
	Right		False		Reject		Right		False		Reject		Right		False		Reject	
	R_{Ra}	R_{Rd}	R_{Fd}	R_{Fa}	R_{Ri}	R_{Ro}	R_{Ra}	R_{Rd}	R_{Fd}	R_{Fa}	R_{Ri}	R_{Ro}	R_{Ra}	R_{Rd}	R_{Fd}	R_{Fa}	R_{Ri}	R_{Ro}
1.0	95.7	96.0	4.3	4.0	0	0	93.3	94.7	6.7	5.3	0	0	96.0	95.6	4.0	4.4	0	0
0.5	92.8	92.0	0	0	7.2	8.0	91.4	90.4	0	0	8.6	9.6	93.3	93.6	0	0	6.7	6.4

때문에 이러한 수치는 단순 비교는 어렵지만 Kotropolis 등이 제안한 모포로지컬 격자 정합(Morphological grid matching)^[6]을 이용한 정면 얼굴에 대한 오인증률 7.4% 및 9.35%보다 우수한 결과로 보인다. 한편 A_{fovea} 를 0.5로 인증범위를 강화한 경우에는 잘못 인증되는 일은 발생하지 않았으나, 소속원에 대한 인증 거부률이 약 6.7~8.6으로 증가하였다. 이는 ESOM의 추가 학습기능을 사용하여 소속원에 대한 좀 더 다양한 정보를 구축함으로써 해결 가능하리라 본다.

구에서와 같은 정도의 인증거부가 일부 발생하였으며, 이는 현재 활용하고 있는 ESOM의 추가 학습기능을 사용하여 소속원에 대한 좀더 다양한 정보를 구축함으로써 해결 가능하리라 본다.

감사의 글

※ 본 논문은 한국 학술 진흥재단의 연구비지원(과제 번호 01-E-0035)의 일환으로 이뤄진 연구결과임.

참 고 문 헌

IV. 결 론

본 논문에서는 통제 구역에 대한 효율적인 접근 제어를 위하여 얼굴의 대칭조건을 이용한 얼굴영역 검출기와 확장된 자기 조직화 지도 신경회로망 ESOM을 동시에 사용한 신분 확인용 얼굴인식시스템을 제안하였다. 폐쇄회로 카메라로 획득한 실물 얼굴 데이터 베이스 KPU와 MIT, ORL 얼굴 데이터 베이스에 대하여 실험을 통한 제안된 방법의 타당성을 알아본 결과 오인증을 완전히 제거할 경우 소속원에 대한 올바른 인증률은 약 91~93%정도로써 약 7~9%정도의 인증거절이 발생하였고, 인증 거절률을 0으로할 경우에는 4~5%의 오인증이 발생하였다. 이러한 수치는 본 연구에서 폐쇄회로 카메라로부터의 표정 변화가 있고 최대 15° 기울어진 얼굴까지도 포함하여 실험한 결과이며, 따라서 정면 얼굴에 대하여 모포로지컬 격자 정합방법^[6]을 사용한 경우보다 우수한 결과로 보인다. 제안된 시스템에서는 얼굴의 대칭성과 수평 및 수직 투영을 이용함으로써 비교적 간단하게 계산얼굴 영역을 추출할 수 있었다. 다만, 잘못된 인증을 완전 방지하려 할 때는 정당한 소속원에 대해서도 기존 연

[1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 704-740, May 1995.

[2] 이성환, 이미숙, "얼굴 영상 인식 기술의 연구 현황," *大韓電子工學會誌 第 23卷 第 6號*, 페이지 688-701, 1996 6

[3] V. Govindaraju, "Locating human faces in photographs," *International Journal of Computer Vision*, vol. 19, no. 2, pp. 129-146, 1996.

[4] Choong Hwan Lee, Jun Sung Kim, and Kyu Ho Park, "Automatic human face location in a complex background using motion and color information," *Pattern Recognition*, vol. 29, no. 11, pp. 1877-1889, 1996.

[5] S. O. Kim, H. K. Sung, K. T. Park, and H. M. Choi, "Automatic face recognition for access control using horizontal and vertical feature based on facial symmetries," *International Conference on Neu-*

- ral Information Processing and Intelligent Information Systems*, Dunedin, New Zealand, vol. 2, pp. 1179-1182, Nov. 1997.
- [6] C. Kotropoulos and I. Pitas, "Face authentication based on topological grid matching," *Proceeding on IEEE Conference on Image Processing*, Santa Barbara California vol. 1, pp. 105-108, Oct. 1997.
- [7] E. R. Davies, *Machine Vision: Theory, Algorithms, Practicalities*, 2nd Edition, Academic Press, San Diego, 1997.
- [8] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, pp. 1464-1480, 1990.
- [9] J. T. McIlwain, "Large receptive fields and spatial transformations in the visual system," in J. A. Anderson, A. Pellionisz, and E. Rosenfeld(Ed.) *Neurocomputing2: Directions for Research*, The MIT Press, Cambridge, 1988.
- [10] Guangzheng Yang and Thomas S. Huang, "Human face detection in a complex background," *Pattern Recognition*, vol. 27, no. 1, pp. 53-63, 1994.

 저 자 소 개

李 鎰 佑(正會員)

1961년 6월생. 1985년 2월 경북대학교 일반대학원 전자공학과 졸업(공학석사). 현재 한국 전자통신 연구원 미디어 연구실 선임 연구원. 경북대학교 일반대학원 박사과정 재학중. 관심분야는 신경회로망을 이용한 영상압축, 멀티미디어 시스템설계

金 星 玩(正會員) 第 34卷 S編 第 6號 參照

수원 전문대 사무자동화과 전임강사

成 孝 慶(正會員) 第 34卷 S編 第 6號 參照

경북대학교 공과대학 전자전기공학부 박사과정

崔 興 文(正會員) 第 34卷 S編 第 6號 參照

경북대학교 공과대학 전자전기공학부 교수