

논문 2004-41CI-6-9

스테레오 영상에서 임베디드 데이터를 이용한 거리에 따른 얼굴인식률 비교

(Face recognition rate comparison with distance change using embedded data in stereo images)

박 장 한*, 남궁 재 찬*

(Chang-Han Park and Jae-Chan Namkung)

요 약

본 논문에서는 스테레오 영상에서 좌·우측 영상을 입력 받아 거리변화와 임베디드 데이터를 이용해 얼굴인식률을 PCA 알고리즘으로 비교한다. 제안된 방법에서는 RGB컬러공간에서 YCbCr컬러공간으로 변환하여 얼굴영역을 검출한다. 또한 거리변화에 따라 추출된 얼굴영상의 확대 및 축소하여 보다 강건한 얼굴영역을 추출한다. 실험을 통하여 제안된 방법은 30cm~200cm 정도의 거리에서 기준 거리(100cm)를 설정하고, 스케일 변화에 따른 평균적인 인식결과로 99.05%(100cm)의 인식률을 얻을 수 있었다. 정규화된 크기(92×112)에서 특정영역인 슈퍼 상태를 정의하고, 각각 정의된 슈퍼 상태의 내부요소인 임베디드 데이터만을 추출하여 PCA 알고리즘을 통하여 얼굴인식을 수행하였다. 원본영상을 모두 학습하는 것이 아니라 임베디드 데이터만을 학습시키기 때문에 제한된 영상의 크기(92×112)에서 특정 데이터를 받아들일 수 있으며, 평균적으로 92×112크기의 영상에서는 99.05%, 실험1은 99.05%, 실험2는 98.93%, 실험3은 98.54%, 실험4는 97.85%의 얼굴인식률을 보였다. 따라서 실험을 통하여 제안된 방법은 거리변화율을 적용하면 높은 인식률을 얻을 수 있음을 보였으며, 얼굴정보를 축소할 뿐만 아니라 처리속도도 향상되었다.

Abstract

In this paper, we compare face recognition rate by PCA algorithm using distance change and embedded data being input left side and right side image in stereo images. The proposed method detects face region from RGB color space to YCbCr color space. Also, The extracted face image's scale up/down according to distance change and extracts more robust face region. The proposed method through an experiment could establish standard distance (100cm) in distance about 30~200cm, and get 99.05% (100cm) as an average recognition result by scale change. The definition of super state is specification region in normalized size (92×112), and the embedded data extracts the inner factor of defined super state, achieved face recognition through PCA algorithm. The original images can receive specification data in limited image's size (92×112) because embedded data to do learning not that do all learning, in image of 92×112 size averagely 99.05%, shows face recognition rate of test 1 99.05%, test 2 98.93%, test 3 98.54%, test 4 97.85%. Therefore, the proposed method through an experiment showed that if apply distance change rate could get high recognition rate, and the processing speed improved as well as reduce face information.

Keywords : stereo images, disparity, embedded data, Principal Component Analysis, face feature extraction, face recognition

I. 서 론

최근 몇 년간 인터넷을 통한 정보가 공유되면서 공유

할 수 없는 개인정보에 대한 가치가 상승하고 있다. 더불어 개인정보 보호 문제가 대두되면서 생체정보를 이용한 보안기술이 주목받고 있으며, 생체정보를 이용해 개인을 인식하는 기술에는 지문, 홍채, 서명검증, 정맥, 음성, 얼굴인식 등이 있다^[1]. 얼굴인식은 다른 생체인식 기술에 비해 경제성과 사용자 편리성이 높다. 또한 고가의 장비 대신 PC카메라를 이용할 수 있으며, 사용자

* 정회원, 광운대학교 컴퓨터공학과
(Department of Computer Engineering, Kangwoon University)
접수일자: 2004년8월18일, 수정완료일: 2004년11월15일

는 기기에 직접 접촉을 하지 않아도 인식이 이루어질 수 있다. 실세계에서 눈으로 얻는 정보는 2차원 정보가 아닌 입체영상이며, 입체영상은 2차원 평면 영상과는 달리 사물에 대한 원근감, 즉 깊이(depth) 정보를 지니고 있기 때문에 이를 원격 작업에 이용하면 작업자는 마치 현장에 직접 투입되어 작업을 수행하는 것과 같은 현장감을 얻게 되어 원격 작업을 쉽게 수행할 수 있다.

얼굴검출 및 특징 추출에서도 2차원을 이용한 방법의 한계를 극복하기 위하여 3차원 얼굴 데이터를 이용한 방법이 연구되고 있다^[2].

얼굴영상의 특징을 획득하는 방법은 경계선, 등고선, 영상 영역의 시각적 특징, 푸리에변환 등의 변환 계수적 특징, 주성분 분석법^[3], 특이값 분해법 등의 대수적 방법이 있다. 또한 NN(Neural Networks)^[3], SVM(Support Vector Machines)^[3], HMM(Hidden Markov Models)^[4] 등이 있으며, 이 중에서 자율학습 방법을 이용하는 PCA(Principal Component Analysis)/LDA(Linear Discriminant Analysis)방법이 많이 사용되고 있다^[3]. 얼굴이 정면으로 구성된 데이터베이스의 경우는 인식률이 상당히 높다고 알려져 있지만, 전체적인 얼굴영상 정보에 기반하여 얼굴을 모델링 하기 때문에 얼굴의 포즈, 조명, 표정변화 등의 영상정보가 취득되지 않으면 인식률이 매우 저하되는 단점을 가지고 있다.

본 논문에서는 스테레오 영상에서 좌·우측 영상의 정규화된 정면 얼굴영상(92×112)에 내포되어있는 눈을 검출하여 머리, 이마, 눈, 귀, 코, 입, 턱의 특정영역을 분할하고, PCA 알고리즘을 이용한 거리변화에 따라 인식률을 비교한다. 제안된 방법에서는 스테레오 영상의 시차(disparity) 보정처리 단계가 선행되어야 한다. 효율적인 얼굴영역 검출을 위해 RGB컬러영상을 YCbCr컬러공간으로 변환하여 얼굴색상 정보를 획득하고, Cb, Cr성분을 이용해 일정한 범위 내에 있는 객체들을 제거(opening) 연산^[5]를 통해 추출된 후보객체들 중에서 가장 큰 객체를 얼굴후보 영역으로 선정한다. 얼굴의 분할된 특정영역을 “슈퍼 상태(super states)”라고 정의하며, 각 슈퍼 상태는 위에서 아래로 정보를 받아들인다. 각 슈퍼 상태의 내부에 있는 정보를 “임베디드 상태(embedded states)”라고 하며 왼쪽에서 오른쪽으로 정보를 읽어 들인다. 이와 같은 방법으로 전체의 영상에서 특정영역만의 정보를 받아들이고 거리변화에 따라 추출된 얼굴영역을 확대 및 축소하여 PCA 알고리즘을 통해 얼굴인식을 수행한다.

본 논문의 구성은 다음과 같다. II장에서는 스테레오

영상, PCA 알고리즘과 얼굴인식 방법을 소개하고, III장에서는 스테레오 영상의 보정처리 및 임베디드 데이터의 얼굴인식 방법에 대한 알고리즘을 제안한다. IV장에서 실험 결과를 분석하고, V장에서 결론을 맺는다.

II. 관련 연구

1. 스테레오 영상

스테레오 영상의 기술은 2차원 영상으로부터 3차원 정보를 추출하는 과정으로 사물에 대한 원근감, 위치, 속도, 깊이 등의 정보를 2대의 카메라로 획득하는 방법이다^[2,6]. 입체영상은 최소 좌·우측으로 구성된 두 장 이상의 평면영상이 필요하며, 일반적으로 2대 이상의 영상센서를 이용하여 동시에 촬영하여 얻는다. 스테레오 영상을 이용한 입체카메라가 사람의 눈에서와 같이 자연스럽게 선명한 양질의 입체영상을 획득하기 위해서는 인간의 시각기능 중 최소한 세 가지 기능을 가지고 있어야 한다. 즉 양안시차(binocular parallax)기능, 초점제어기능, 주시각 제어기능 등을 갖추어야 한다. 또한 사람 눈에 해당하는 카메라를 설치하는 방법에 따라 크게 순차식과 동시식으로 나눌 수 있다. 순차식은 좌·우측의 카메라를 동시에 존재시키지 않고 하나의 카메라를 이용하여 순차적으로 영상을 입력하는 방식이다. 동시식이란 좌·우측 카메라를 모두 설치하여 동시에 영상을 취득하는 것으로 현재의 스테레오비전 시스템에서 많이 사용하고 있다. 동시식 카메라 구성에는 카메라 광축의 구성법 따라 평행식과 교차식으로 분류한다. 평행식은 두 카메라의 광축이 수직, 수평으로 평행한 구성법이며, 교차식은 두 카메라의 광축이 일치하는 곳 즉, 주시점을 만드는 방법으로 사람의 눈 구조와 매우 가깝다고 할 수 있다. 평행식 카메라 구성법은 수평 시차만 존재하며 수직 시차는 없기 때문에 취급하기에 용이하여 수학적 모델링이 쉽다. 또한 거리 측정에는 수학적으로 쉽게 나타낼 수 있기 때문에 간단한 방법으로 거리를 측정할 수 있다. 교차식 구성방법은 광축을 일치시키는 문제가 있으며, 수직 시차의 변화에 따른 문제가 있기 때문에 수학적 모델링이 어려운 문제가 있다. 본 논문에서는 그림 1과 같이 평행식 카메라를 설치하여 거리정보를 취득하고, 좌·우 영상으로부터 일치점의 변위인 시차(disparity)를 구하여 삼각 측량법 등을 통해 거리정보를 추출한다. 실험에 사용된 카메라의 구성요소는 표 1과 같다.

표 1. 카메라 구성요소
Table 1. Camera component elements.

항 목	특 성	
카메라 설치 방법	평행식	
카메라 설치 간격	65(mm)	
카메라 초점 거리(f)	3.6(mm)	
크 기	1픽셀	7.2×5.6(μm)
	전체	3.6×2.7(mm)
해상도	가로	512(dots)
	세로	492(lines)

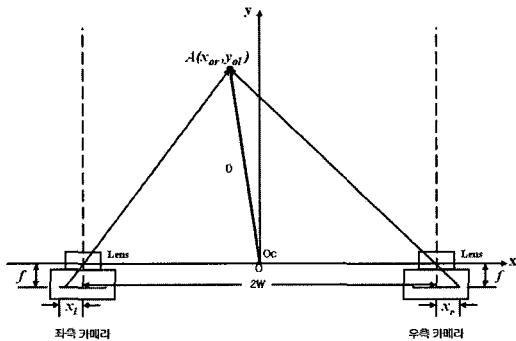


그림 1. 평행식 카메라의 구조
Fig. 1. Structure of parallel camera.

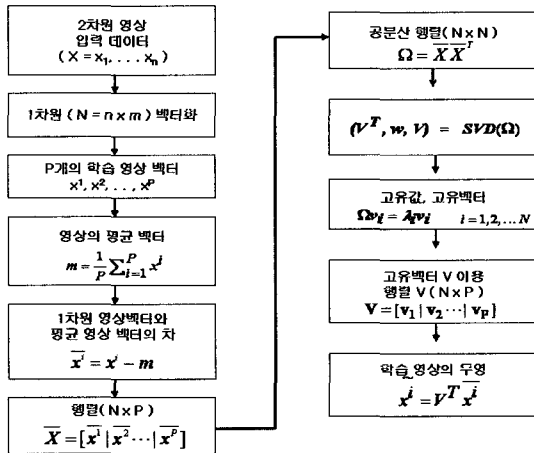


그림 2. PCA 알고리즘 블록도
Fig. 2. PCA algorithm block diagram.

2. 주성분분석 방법

주성분분석 방법(Principal Component Analysis)의 알고리즘은 그림 2와 같으며, 서로 관련 있는 변수들의 데이터 집합이 관측 되었을 때, 변수들이 가지고 있는 원래 정보들을 최대한 반영할 수 있는 보다 적은 수의 새로운 변수들을 생성하는 방법이다. 또한 원래의 변수들이 가지고 있는 정보의 양을 가장 많이 확보하는 순서대로 변수들의 선형 결합을 이용하여 새로운 변수를 만들며, 고차원 데이터 집합을 저차원 데이터 집합으로 생성하는 방법이다^[3]. 입력 영상을 1차원 벡터로 변환한 후에 전체 입력 영상들의 직교정규화된 공분산 행렬

(covariance matrix)을 계산하고, 고유값(eigenvalue)과 고유벡터(eigenvector)를 구하여 입력된 얼굴과 기준 얼굴의 고유벡터를 벡터 공간상에서 비교하여 가장 가까운 거리의 벡터를 지닌 얼굴이 인식되도록 하는 방법이다. 본 논문에서는 얼굴영역 전체를 학습시키는 것이 아니라 분할된 특정영역의 슈퍼 상태에서 임베디드 데이터를 추출하여 특정영역의 데이터를 학습하기 위한 PCA 알고리즘을 사용한다.

3. 얼굴특징 추출과 인식방법

얼굴특징의 추출은 얼굴영상 및 얼굴을 포함한 영상 내의 분리된 얼굴영역에서 얼굴특징들, 즉 눈, 코, 입 등을 추출하는 것으로 다양한 방법에 의해 연구되어 왔다. 수염이 없는 얼굴 영상에 대해 KL변환을 적용하여 얼굴을 표현하는 방법^[7], 얼굴특징 추출을 위해 Hough 변환을 사용한 방법^[8], 가변 형판을 사용하여 얼굴의 특징을 추출한 방법^[9], Gabor 웨이블릿 분해와 지역적 크기 변환의 상호 작용을 사용하여 곡률이 최대인 점에서 얼굴특징을 추출하는 방법^[10] 등이 있다. 최근에 웨이블릿 해석을 사용한 방법과 템플릿과 같은 영상 필터를 사용한 기술들이 선보이고 있다^[11].

얼굴인식 방법은 컬러정보^[12], 모양^[13], 특징^[13], 유형^[14] 등의 기반으로 하는 방법이 있다. 또 다른 유형을 사용하는 방법으로는 입력된 영상을 이미 학습된 얼굴 영상들 간의 중요한 변화를 나타내는 공간으로 투영시키는 방법과 3차원 영상을 사용하여 얼굴을 인식하는 방법들이 있다. 3차원 영상은 물체의 깊이(depth) 구조를 포함하고 있으며, 이런 데이터들이 대부분의 응용에서는 이용 불가능하지만, 2차원 영상에서는 얻을 수 없는 매우 유용한 정보를 가지고 있어 얼굴특징 추출 및 인식에 사용되기도 한다.

본 논문에서는 얼굴특징인 눈의 위치를 검출하여 머리, 이마, 눈, 귀, 코, 입, 턱의 분할된 영역을 검출하고, 추출된 얼굴영역에서 분할된 특정영역의 특징요소인 임베디드 데이터를 이용한 얼굴인식을 한다. 따라서 얼굴 전체를 대상으로 인식하는 것이 아니라 분할된 특정영역 정보만 이용하여 학습 초기의 데이터를 축소하여 PCA 알고리즘으로 학습한다.

III. 스테레오 영상의 보정처리 및 얼굴인식 방법

1. 제안된 알고리즘 블록도

본 논문에서 제안하는 스테레오 영상의 보정처리 및

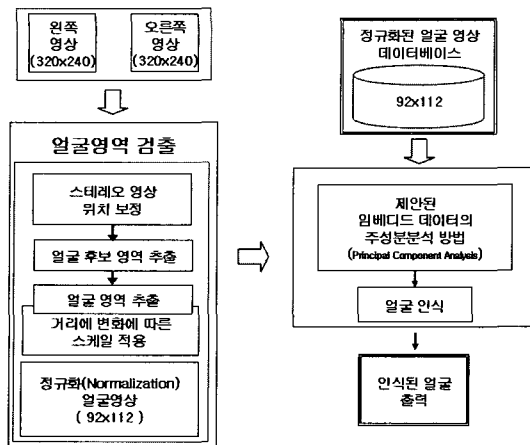


그림 3. 제한된 알고리즘 블록도
Fig. 3. Proposed algorithm block diagram.

얼굴인식의 블록도는 그림 3과 같다. 제안된 블록도는 카메라에서 획득된 스테레오 영상의 영상위치 보정, 정면 얼굴후보 영역 추출, 정면 얼굴영역 추출 단계의 정규화된 얼굴영상의 정보를 거리 변화(30cm, 50cm, 100cm, 120cm, 150cm, 200cm)에 따라 확대 및 축소의 정규화 과정을 거치며, 정규화된 영상 크기의 92×112에서 임베디드 데이터를 추출한다. 또한 데이터베이스도 정규화된 임베디드 데이터의 정보를 저장한다. 얼굴인식은 입력 영상을 PCA 알고리즘으로 학습하여 저장된 데이터베이스에서 가장 근접한 얼굴영상을 출력한다.

2. 스테레오 영상의 시차 보정 처리

스테레오 영상에서 시차를 추출하는 방법에서 고정된 크기의 블록정합 알고리즘을 사용한다. 블록정합에 사용되는 기저함수는 MAD(Mean Absolute Difference)와 MSD(Mean Square Difference)를 주로 사용하는데, 계산량을 줄이기 위해 SAD(Sum of Absolute Difference)를 기저함수로 사용한다. 공간상의 한 점이 좌·우 영상에 투영되었을 때 두 점을 연결하는 가상의 선을 에피폴라선이라 한다^[15]. 우측영상의 대응 블록은 좌측영상과 동일한 Y축의 좌표를 가지는 에피폴라선상에서 정합된다. 본 논문에서는 Y축의 오차범위를 상·하 5픽셀 이내로 보정하여 실험한다. 식(1)은 스테레오 영상의 시차추출을 위해 사용된 SAD를 나타낸다.

$$SAD = \sum_i^{N_x} \sum_j^{N_y} |I_L(i, j) - I_R(i + k, j)| \quad (1)$$

여기서 N_x, N_y 는 블록의 크기를 나타낸다. I_L 은 좌측 영상의 블록을 나타낸다. 또한 I_R 은 I_L 에 대응되는 우측 영상의 블록을 나타내며, k 는 왼쪽 영상에 대응되는 오

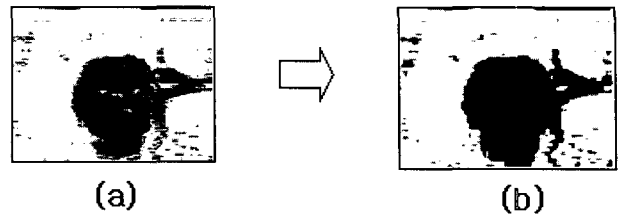


그림 4. 색상 영역을 이용한 얼굴영역 추출
(a) 색상영역 분할 후의 영상
(b) 제거 연산을 적용한 후의 영상
Fig. 4. Face region extraction by using color model.

른쪽 영상의 시차이다^[2].

3. 얼굴영상의 거리변화에 따른 스케일 처리

카메라로부터 획득된 입력 영상은 320×240 크기로 얼굴을 포함하는 RGB컬러영상이다. 먼저 획득된 얼굴 후보 영역 영상에서 색상 정보만을 이용하여 조명에 대한 영향을 줄이기 위해 주어진 영상을 RGB컬러공간으로부터 YCbCr컬러공간으로 변환하게 된다. 입력영상의 YCbCr 정보 중에서 Cb, Cr성분에 대해 동양인에 맞게 얼굴 피부색상 범위를 $R_{Cb}=[77 \ 127]$, $R_{Cr}=[133 \ 173]$ 으로 색상 영역분할과정을 거쳐 피부색 값을 갖는 영역만을 추출한다^[16]. 식(2)을 이용하여 R_{Cb}, R_{Cr} 의 영역을 선택한다.

$$S(x, y) = \begin{cases} 1, & \text{if} [(Cb(x, y) \in R_{Cb}) \cap (Cr(x, y) \in R_{Cr})] \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

그림 4의 (a)는 입력된 영상을 색상영역 분할처리를 거친 결과이다. (a)의 영상을 보면 세밀한 잡음 형태가 존재하는 것을 볼 수 있는데, 이런 잡음의 형태를 제거하여 얼굴 영역만을 분리해 내기 위하여 형태학적 필터(morphological filter) 중에서 제거(opening) 연산을 수행하였다. 제거 연산은 원 영상을 침식 연산으로 영상의 최외각을 한 픽셀씩 없애고 다시 팽창 연산으로 최외각을 한 픽셀씩 확장시키는 방법이다. 따라서 제거 연산은 미세한 잡음을 제거하고, 다시 팽창 연산을 수행함으로써 원 영상의 크기를 유지할 수 있다. 본 논문에서는 자율학습의 입력 차원을 결정하기 위해 92×112 크기의 영상에서 임베디드 데이터의 얼굴 데이터베이스를 구축한다.

본 논문에서는 실제 거리를 측정하기 위해 식(3)을 이용하여 거리를 측정한다^[17]. 거리를 계산하기 위해서는 표 1에서 제시한 카메라 특성을 이용하여 카메라의 한 픽셀당 실제 유효거리를 계산하며 산출된 한 픽셀의

거리는 $86.80 \times 10^3 [m]$ 이다.

$$D = \frac{bf}{x_l - x_r} \times 86.80 \times 10^3 [m] \quad (3)$$

단, 여기에서 b 는 카메라간의 설치 거리, f 는 카메라의 초점거리, x_l 은 좌측 영상의 x 거리 및 x_r 은 우측영상의 x 거리이다. 입력 영상의 320×240 에서 가로 픽셀의 시차 ($x_l - x_r$)의 최대 측정거리는 1이며, 최소 측정거리는 320이다. 또한 거리 변화에 따른 스케일치리는 식(4)과 같이 처리한다^[17]. 여기에서 x' , y' 는 스케일 적용후의 변화된 위치이며, s_x, s_y 는 스케일 변화에 적용하는 값이고, x, y 는 현재위치이다.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

4. 임베디드 데이터 집합

취득된 얼굴영역에서 두 눈을 검출하여 특징요소인 머리, 이마, 눈, 귀, 코, 입, 턱의 분할된 특정영역을 블록 데이터 집합으로 표기하는 방법을 ‘슈퍼 상태’라고 하고, 각 분할된 특정영역에서 내부의 데이터를 임베디드 ‘데이터 집합’이라고 정의한다. 그림 5의 슈퍼 상태는 정면얼굴에서 얼굴의 특징요소의 분할된 영역이고, 상위에서 하위로 정보를 읽고, 임베디드 상태는 각 슈퍼 상태에서 좌측에서 우측으로 픽셀정보를 읽는 방법이다. 본 논문에서는 좌·우측 영상에서 각각 92×112 의 얼굴영상을 취득한 후 얼굴영역 내에 있는 특정영역의 분할된 특징요소인 임베디드 데이터를 입력값으로 사용한다.

5. 주성분분석 방법에서 임베디드 데이터 학습방법
본 논문에서는 입력 데이터를 전체영역으로 처리하

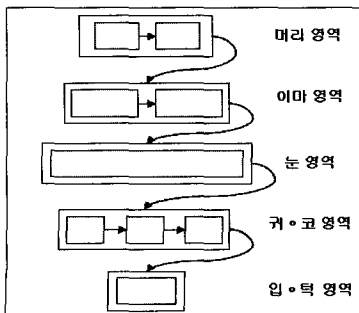


그림 5. 임베디드 데이터 집합
Fig. 5. A set of embedded states.

는 것이 아니라 분할된 특정영역에서 데이터를 학습시키면 원래 정보를 최대한 반영하면서 새로운 변수들을 생성할 수 있는 특정영역을 학습시킬 수 있는 PCA 방법을 제한한다. 먼저 임베디드 데이터를 취득하는 방법은 주어진 자료를 X 라 놓고, 얼굴영역의 특정영역에서 P 개의 관찰된 표본이 있다면 식(5-1)~식(5-7)과 같이 정의할 수 있다. 이때 P 의 각 표본 $x^{embed\ i}$ 는 해당 표본을 구성하는 식(5-1)~식(5-7)과 같이 p 개의 데이터로 구성되어 있다. 만약 얼굴영상의 경우라면 이때의 p 의 값은 해당 얼굴의 픽셀(pixel) 수가 될 것이며 1차원 벡터로 나타낼 수 있다. 임베디드 데이터의 학습방법은 과정1~과정12까지 이루어진다.

과정1. 2차원 영상($m \times n$)을 1차원 ($N = m_p \times n_p$) 벡터로 차원을 변경한다.

과정2. 얼굴영역에서 두 눈을 검출하여 분할된 특정영역인 머리, 이마, 눈, 귀, 코, 입, 턱 등의 슈퍼 상태들에 대한 P 개 학습영상 벡터표기는 식(5-1)~식(5-7)과 같다. 여기에서 $P_{hair}, P_{forehead}, P_{eyes}, P_{ears}, P_{nose}, P_{mouth}, P_{chin}$ 는 각각 머리, 이마, 눈, 귀, 코, 입, 턱 등의 영역이다. 또한 머리(x^a), 이마(x^b), 눈(x^c), 귀(x^d), 코(x^e), 입(x^f), 턱(x^g)은 각 슈퍼 상태에서 내포되어있는 임베디드 데이터이다.

$$P_{hair} = x^{a_1} x^{a_2} x^{a_3} \dots x^{a_p} \quad (5-1)$$

$$P_{forehead} = x^{b_1} x^{b_2} x^{b_3} \dots x^{b_p} \quad (5-2)$$

$$P_{eyes} = x^{c_1} x^{c_2} x^{c_3} \dots x^{c_p} \quad (5-3)$$

$$P_{ears} = x^{d_1} x^{d_2} x^{d_3} \dots x^{d_p} \quad (5-4)$$

$$P_{nose} = x^{e_1} x^{e_2} x^{e_3} \dots x^{e_p} \quad (5-5)$$

$$P_{mouth} = x^{f_1} x^{f_2} x^{f_3} \dots x^{f_p} \quad (5-6)$$

$$P_{chin} = x^{g_1} x^{g_2} x^{g_3} \dots x^{g_p} \quad (5-7)$$

과정3. 영상에서 임베디드 데이터의 평균벡터 표기는 식(6)과 같다.

$$m_m = \frac{1}{7} \sum_{i=1}^7 \left(\frac{1}{p} \sum_{j=1}^p x^{a_j} + \frac{1}{p} \sum_{j=1}^p x^{b_j} + \frac{1}{p} \sum_{j=1}^p x^{c_j} + \frac{1}{p} \sum_{j=1}^p x^{d_j} + \frac{1}{p} \sum_{j=1}^p x^{e_j} + \frac{1}{p} \sum_{j=1}^p x^{f_j} + \frac{1}{p} \sum_{j=1}^p x^{g_j} \right) \quad (6)$$

과정4. 1차원 임베디드 영상벡터와 평균벡터의 차 값은 식(7)과 같이 구한다. 단, 여기에서 x^{ed} 는 임베디드 집합으로 각 특정영역에서 머리(x^a), 이마(x^b), 눈(x^c), 귀(x^d), 코(x^e), 입(x^f), 턱(x^g)이다.

$$\overline{x^{ed_i}} = x^{ed_i} - m_m, (i=1,2,3,\dots,p) \quad (7)$$

과정5. 과정 4의 식(7)을 이용하여 임베디드 데이터의 $N_{ed} \times P_{ed}$ 의 행렬을 식(8)과 같이 구한다.

$$\overline{X^{ed}} = [\overline{x^{ed_1}} | \overline{x^{ed_2}} | \overline{x^{ed_3}} \dots \overline{x^{ed_p}}] \quad (8)$$

과정6. 식(8)을 공분산 행렬(covariance matrix) $N_{ed} \times N_{ed}$ 로 정의하면 식(9)과 같다.

$$\Omega_{ed} = \overline{X^{ed} X^{edT}} \quad (9)$$

과정7. 공분산 행렬에 대해서 식(10)과 같이 고유값(λ_{ed})와 고유벡터(v_{ed})를 구한다.

$$\Omega_{ed} v_{ed_i} = \lambda_{ed_i} v_{ed_i}, (i=1,2,3,\dots,N_{ed}) \quad (10)$$

과정8. 고유벡터(v_{ed})를 이용하여 ($N_{ed} \times P_{ed}$)행렬 V_{ed} 을 식(11)과 같이 구한다.

$$V_{ed} = [v_{ed_1} | v_{ed_2} | v_{ed_3} \dots | v_{ed_p}] \quad (11)$$

과정9. 학습영상의 투영은 식(11)의 V_{ed} 값을 이용하여 식(12)과 같이 학습영상에 대한 특징벡터 \tilde{x}^{ed_i} 을 구한다.

$$\tilde{x}^{ed_i} = V_{ed}^T \overline{x^{ed_i}} \quad (12)$$

과정10. 새로 입력되는 인식대상의 영상에 대해 평균영상값을 빼주는 식은 식(13)과 같다.

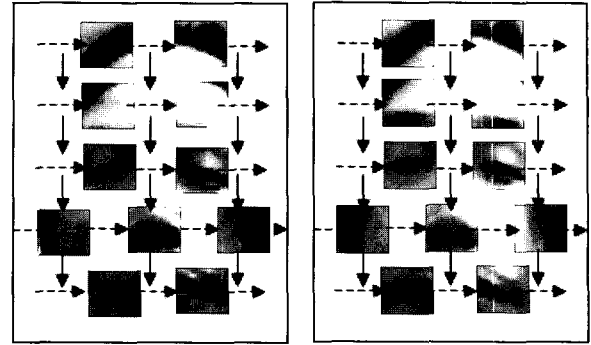
$$\begin{aligned} \overline{y^{ed_i}} &= y^{ed_i} - m_m, \\ \left(m_m &= \frac{1}{7} \left(\frac{1}{p} \sum_{i=1}^p x^{a_i} + \frac{1}{p} \sum_{i=1}^p x^{b_i} + \frac{1}{p} \sum_{i=1}^p x^{c_i} + \frac{1}{p} \sum_{i=1}^p x^{d_i} \right. \right. \\ &\quad \left. \left. + \frac{1}{p} \sum_{i=1}^p x^{e_i} + \frac{1}{p} \sum_{i=1}^p x^{f_i} + \frac{1}{p} \sum_{i=1}^p x^{g_i} \right) \right) \end{aligned} \quad (13)$$

과정11. 인식대상의 영상 투영은 식(14)과 같이 구한다.

$$\tilde{y}^{ed_i} = V_{ed}^T \overline{y^{ed_i}} \quad (14)$$

과정12. 식(14)의 투영된 \tilde{y}^{ed_i} 와 가장 가까운 \tilde{x}^{ed_i} 을 구하여 인식값으로 취하고, 인식대상으로 결과를 출력한다.

본 논문에서는 좌·우측 영상에서 과정1~과정12 단계를 거쳐 그림 6과 같이 임베디드 데이터를 추출하고, 거리변화에 따른 얼굴인식률을 비교한다.



(a) 좌측 영상

(b) 우측 영상

그림 6. 임베디드 데이터 추출과정

Fig. 6. Extraction process of embedded data.

6. 얼굴 인식

입력 얼굴영상의 전처리 과정을 거친 후, 그림 5와 같이 PCA 알고리즘에 의해 고유값과 고유벡터를 구한다. 또한 정규화된 학습 영상 데이터베이스도 같은 과정을 거쳐 부호화된다. 얼굴인식은 학습된 얼굴 PCA 패턴과 정규화된 학습영상 패턴을 비교하여 가장 비슷한 얼굴을 찾아내는 과정이며, 가장 단순한 유클리안(euclidean)거리를 이용한 최근접 이웃(nearest neighbor)방법을 이용하여 얼굴을 인식한다.

입력 얼굴 코드 $R=(r_1, \dots, r_m)$, 임의의 학습 얼굴의 코드 $T=(t_1, \dots, t_m)$ 이라고 할 때, 유클리안 거리 d 는 식(15)과 같다. 따라서 본 논문에서는 모든 학습된 얼굴에 대하여 유클리안 거리를 구하고, 그 중에서 거리가 가장 작은 값을 갖는 영상을 출력으로 하는 방법이다.

$$d = \sum_{i=1}^m (r_i - t_i)^2 \quad (15)$$

IV. 실험 및 고찰

본 논문에서는 스테레오 320×240영상에서 얼굴영역을 취득한 후 얼굴영상의 크기(92×112) 영상 50~400개를 실험영상으로 사용하였다. 또한 실험에 사용된 컴퓨터는 Pentium-IV 2.8/512MB이고, VC++ 6.0으로 구현하였다. 제안된 실험은 왼쪽영상에서 상단좌표(145,100), 하단좌표(175,140)의 30×40블록을 사용해서 오른쪽 영상의 상단좌표(70,70), 하단좌표(250,170)에서 가장 유사한 블록을 찾아서 시차정보를 추출하여 스테레오 정합을 한다. 그림 7~9는 거리 변화에 따른 얼굴영상을 취득한 후에 좌측영상에서 우측영상의 가장 유사한 블록



그림 7. 거리 50cm에서 측정
Fig. 7. Measurement in distance 50cm.



그림 13. 제안된 방법의 인식 결과
Fig. 13. Recognition result of proposed method.



그림 9. 거리 150cm에서 측정
Fig. 9. Measurement in distance 150cm.



그림 10. 보정된 영상
Fig. 10. Compensated image.



(a) 좌측 영상(30cm,50cm,100cm,120cm,150cm,200cm)



(b) 우측 영상(30cm,50cm,100cm,120cm,150cm,200cm)

그림 11. 스케일을 적용한 결과로 취득된 영상
Fig. 11. Image is acquired by result that apply scale.



그림 12. 제안된 방법의 인식 결과
Fig. 12. Recognition result of proposed method.

을 찾은 결과이다. 또한 그림 10은 SAD를 이용하여 좌측영상에서 우측영상의 가장 유사한 블록을 찾고 좌·우측 영상을 정합한 결과이다.

그림 11은 그림 7~9에서 거리를 측정한 후 거리변화에 따른 스케일 변화를 적용하여 취득된 영상이다. 320×240영상에서 92×112크기의 영상을 대한 확대/축소 비율은 기준 거리 100(cm)을 기준으로 거리에 따라 확대 비율은 120(cm)은 1.2, 150(cm)은 1.5, 200(cm)은 1.8이며, 축소비율은 30(cm)은 0.4, 50(cm)은 0.5이다. 이것은 실험에 의해 결정된 비율이다.

그림 12는 거리변화에 따라 얼굴영상을 확대/축소하

표 2. 제안된 방법의 인식률 비교
Table 2. Recognition rate comparison of proposed method.

영상의 수 (좌측/우측)	제안된 방법으로 PCA 실험					
	30(cm)	50(cm)	100(cm)	120(cm)	150(cm)	200(cm)
25/25	99.30	99.50	99.70	99.50	98.60	98.20
50/50	99.20	99.40	99.40	99.30	98.20	98.00
75/75	99.00	99.20	99.20	99.10	97.70	97.50
100/100	98.80	99.00	99.10	99.00	97.20	96.80
125/125	98.50	98.80	99.00	98.90	97.00	96.40
150/150	98.20	98.70	98.80	98.80	96.40	96.00
175/175	98.00	98.50	98.70	98.50	97.00	95.80
200/200	97.90	98.20	98.50	98.10	96.40	95.00
평균	98.61	98.91	99.05	98.90	97.31	96.71

표 3. 제안된 방법의 거리 따른 인식률 비교
Table 3. Recognition rate comparison by distance of proposed method.

거리	제안된 방법1	제안된 방법2	PCA	HMM
30cm	97.9	96.9	97.5	97.2
50cm	98.2	97.3	96.0	97.0
100cm	98.5	98.2	81.8	91.0
120cm	98.1	97.8	72.2	78.6
150cm	96.4	95.2	65.3	67.9
200cm	95.0	92.4	50.7	54.9
평균	97.35	96.30	77.25	81.10

고, 50~400개의 영상 중에서 원본영상을 입력하여 제안된 방법으로 정확하게 인식된 결과이다. 그림 13은 테스트 데이터를 입력한 결과 인식된 결과이다.

표 2는 제안된 방법의 거리변화에 따른 얼굴인식률을 비교한 결과이며, 스테레오 영상의 좌·우측 영상에서 92×112영상을 각각 취득하여 실험하였다. 취득된 영상을 거리에 따라 실험한 결과에서 기준거리인 100cm의 경우가 가장 인식률이 높았다. 또한 거리변화에 따른 확대 및 축소 영상에 대한 인식률은 떨어지지만 보간 방법을 개선한다면 높은 인식률을 얻을 수 있다.

표 3은 좌·우측 영상(200/200)을 학습하여 120장의 테스트영상으로 제안된 방법에서 거리에 따른 얼굴인식률을 비교한 결과이며, 제안된 방법1의 실험은 정면 얼굴영상에서 인식률을 실험한 결과이다. 또한 제안된 방법2의 실험은 정면 얼굴이 아닌 약간의 포즈변화에 따라 실험한 결과이다. 확대 및 축소를 적용하지 않고 실험한 PCA와 HMM은 거리에 따라 인식률이 떨어진 것을 볼 수 있다. 따라서 제안된 방법으로 얼굴인식을 하면 인식률을 높일 수 있다.

표 4는 제안된 방법의 얼굴인식률을 비교한 결과이

다. 일반적인 PCA에서는 영상의 크기가 92×112의 10304픽셀이지만, 제안된 방법은 92×112영상에서 실험1은 6640픽셀, 실험2는 4524픽셀, 실험3은 3092픽셀, 실험4는 1176픽셀의 크기를 취득한다. 임베디드 데이터를 이용한 얼굴인식의 평균 결과로 단계1에서 약 99.05%의 높은 결과를 얻을 수 있었다. 따라서 제안된 방법을 통하여 얼굴영상의 임베디드 데이터를 이용하여 인식하면 얼굴 데이터베이스의 용량 및 크기를 줄일 수 있다.

그림 14는 표 4에서 제시한 결과에서 실험을 통하여 학습된 시간의 결과이며, 92×112 영상의 경우는 영상의 개수가 많아질수록 학습 시간은 급격히 증가하고, 본 논문에서 제안한 방법의 학습시간에서는 짧은 것을 볼 수 있다.

표 5는 실험에 사용된 5개 영상의 거리측정 데이터를 나타낸 것으로 측정 거리가 멀수록 측정 오차는 적어지

표 4. 제안된 방법의 인식을 비교
Table 4. Recognition rate comparison of proposed method.

영상의 수	PCA	제안된 방법			
		실험1	실험2	실험3	실험4
50	99.70	99.70	99.60	99.20	99.00
100	99.40	99.40	99.40	99.00	98.70
150	99.20	99.20	99.20	98.80	98.60
200	99.10	99.10	99.00	98.70	98.20
250	99.00	99.00	98.80	98.50	97.70
300	98.80	98.80	98.70	98.20	97.20
350	98.70	98.70	98.50	98.00	97.00
400	98.50	98.50	98.20	97.90	96.40
평균	99.05	99.05	98.93	98.54	97.85

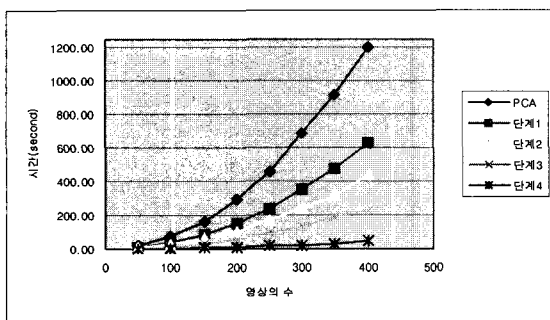


그림 14. 학습 시간 비교
Fig. 14. Comparison of learning time.

표 5. 거리 측정의 비교
Table 5. Comparison of distance measurement.

목표물	실측거리	측정거리	오차율
A	0.5(m)	0.521(m)	4.2(%)
B	0.7(m)	0.719(m)	2.7(%)
C	0.9(m)	0.894(m)	-0.6(%)
D	1.1(m)	1.078(m)	-0.7(%)
E	1.5(m)	1.491(m)	-0.6(%)

며, 근 거리 일수록 측정 오차가 커짐을 알 수 있다. 오차율은 식(16)을 이용해 산출하였다. 여기서 e는 오차율이며, M은 측정값, R은 기준 값이다.

$$e = \frac{M - R}{R} \times 100 (\%) \tag{16}$$

V. 결 론

본 논문에서는 스테레오 영상에서 좌·우측 영상을 입력 받아 거리변화와 얼굴영역에 내포되어있는 임베디드 데이터를 이용해 얼굴인식을 PCA알고리즘으로 비교하였다. 제안된 방법에서는 RGB컬러공간에서 YCbCr 컬러공간으로 변환하여 얼굴영역을 검출한다. 또한 거리 변화에 따른 추출된 얼굴영상의 스케일을 통해 확대 및 축소하여 보다 강건한 얼굴영역을 추출할 수 있었다. 실험을 통하여 제안된 방법은 30cm~200cm 정도의 거리에서 기준 거리(100cm)를 설정하고, 스케일 변화에 따른 평균적인 결과로 98.61%(30cm), 98.91%(50cm), 99.05%(100cm), 98.90%(120cm), 97.31%(150cm), 96.71%(200cm)의 인식률을 얻을 수 있었다. 또한 정규화된 크기(92×112)에서 특정영역인 슈퍼 상태를 정의하고, 각각 정의된 슈퍼 상태의 내부요소인 임베디드 데이터만을 추출하여 PCA알고리즘을 통하여 얼굴인식을 수행하였다. 원본영상을 모두 학습하는 것이 아니라 임베디드 데이터만을 학습시키기 때문에 제한된 영상의 크기(92×112)에서 특정 데이터를 받아들일 수 있으며, 평균적으로 92×112크기의 영상에서는 99.05%, 6640픽셀은 99.05%, 4524픽셀은 98.93%, 3092픽셀은 98.54%, 1176픽셀은 97.85%의 얼굴인식률을 보였다. 따라서 실험을 통하여 제안된 방법은 거리에 따른 변화율을 적용하면 높은 인식률을 얻을 수 있음을 보였으며, 얼굴영상의 정보를 축소할 뿐만 아니라 처리속도도 향상되었다. 향후 얼굴의 방향 및 각도별로 연구를 진행하여 얼굴 인식률을 높일 수 있는 연구가 진행되어야 한다.

참 고 문 헌

[1] Y. Ming-Hsuan, D.J. Kriegman, N. Ahuja, "Detecting faces in images: a survey", *Pattern Analysis and Machine Intelligence, IEEE Transactions*, Vol. 24, no. 1, pp. 34-58, Jan. 2002.
[2] 이학찬외 3인, "스테레오 영상을 이용한 물체 추적 방법", 대한전자공학회, 제39호 SP편 제5호, pp.

522~534, 2002.

[3] Z. Sun, G. Bebis, X. Yuan, S. J. Louis, "Genetic Feature Subset Selection for Gender Classification: A Comparison Study", *Applications of Computer Vision, 2002. (WACV 2002). Proceedings. Sixth IEEE Workshop on*, 2002. pp.165-170. Dec 2002.

[4] F. Samaria and S. Young, "HMM based architecture for face identification", *Image and Vision Computing*, vol. 12, pp. 537-543, 1994.

[5] B.A. McLindin, "Baselining illumination variables for improved facial recognition system performance", *Video/Image Processing and Multimedia Communications, 2003. 4th EURASIP Conference*, Vol. 1, pp. 417-422, 2-5 July 2003.

[6] W.A. IJsselsteijn, H. de Ridder, J. Vliegen, "Subjective evaluation of stereoscopic images: effects of camera parameters and display duration", *Circuits and Systems for Video Technology, IEEE Transactions*, Vol. 10, no. 2, pp. 225-233, March 2000.

[7] Y. Hongxun, L. Mingbao, Z. Lizhuang, "Eigen features technique and its application", *Signal Processing Proceedings, 2000. WCCC-ICSP 2000*, vol. 2, pp. 1153-1158, 21-25 Aug. 2000.

[8] B. A, McLindin, "Baselining illumination variables for improved facial recognition system performance", *Video/Image Processing and Multimedia Communications, 2003. 4th EURASIP Conference*, Vol. 1, pp. 417-422, 2-5 July 2003.

[9] Y. Zhong, J. A.K., D. Jolly, M.-P, "Object tracking using deformable templates", *Pattern Analysis and Machine Intelligence, IEEE Transactions*, Vol. 22, no. 5, pp. 544-549, May. 2000.

[10] L. Chengjun, H. Wechsler, "Independent component analysis of Gabor features for face recognition", *Neural Networks, IEEE Transactions*, vol. 14, no. 4, pp. 919-928, July 2003.

[11] R.S. Feris, R.M. Cesar, "Tracking Facial Features Using Gabor Wavelet Networks", *Proc. 13th Brazilian Symposium on Computer Graphics and Image Processing*, pp. 22-27, 2000.

[12] L. Zhi-fang, Y. Zhi-sheng, A.K. Jain, W. Yun-qiong, "Face detection and facial feature extraction in color image", *Computational Intelligence and Multimedia Applications, 2003. ICCIMA 2003. Proceedings. Fifth International Conference*, pp.126-130, 27-30 Sept. 2003.

[13] M. Rizon, T. Kawaguchi, "Automatic eye detection using intensity and edge information", *TENCON 2000. Proceedings*, vol. 2, pp. 415-420, 24-27 Sept. 2000.

[14] H. Rowley, S. Baluja and T. Kanade, "Neural Network-Based Face Detection", *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 20, no. 1, pp. 203-208, 1998.

[15] C. Liang-hua, L. Wei-Chung, "Visual surface segmentation from stereo", *Image and Vision Computing*, Vol. 15, pp. 95-106, 1997.

[16] Chi, D., Ngan, K.N.: "Face Segmentation Using Skin-Color map in Videophone Applications", *IEEE Trans. Circuits and systems for video technology*, June,1999, 9, (4), pp. 551-564.

[17] O. D. Faugeras, "Three-Dimensional Computer Vision", MIT Press, 2001.

저 자 소 개



박 장 한(정회원)
 1997년 2월 방송통신대학교
 전산계산학과(독학학위)
 졸업
 1999년 8월 광운대학교
 전산대학원 소프트웨어
 공학과 석사

2001년 3월~현재 광운대학교 대학원
 컴퓨터공학과 박사과정
 <주관심분야: 스테레오비전, 3D영상, 영상처리,
 영상인식>



남공재찬(정회원)
 1970년 2월 인하대학교
 전기공학과 졸업
 1976년 8월 인하대학교 대학원
 전자공학과 석사
 1982년 2월 인하대학교 대학원
 전자공학과 박사

1979년 3월~현재 광운대학교 컴퓨터공학과 교수
 <주관심분야: 신경망, 영상처리, 문자인식>

