
통합된 시스템에서의 얼굴검출과 인식기법

박동희* · 배철수*

An Integrated Face Detection and Recognition System

Dong-hee Park* · Cheol-soo Bae*

요 약

본 논문에서는 임의의 장면에도 얼굴 인식에 영향을 받지 않는 통합된 얼굴 인식 방법을 제안한다. 크기 정규화는 피부 색 분할과 log-polar 매핑 절차의 새로운 조합을 통하여 얻어지고, 주요 얼굴 구성 요소 분석은 자세 변화들을 처리하기 위하여 제안된 멀티 뷰 접근을 통해 이루어진다.

주어진 컬러 입력 이미지로부터 검출기는 얼굴을 원형 경계 안에 둘러싸고 코의 위치를 표시하며 다음 인식을 위해, 원형 경계 내에 배치하는 방사형 격자는 특징 벡터 코 중심에 두었다. 컬러로 분할된 영역의 폭으로서 얼굴의 크기를 평가하고, 추출된 특징 벡터는 평가된 크기에 의하여 정규화 된 크기이다. 특징 벡터는 얼굴 인식을 위해 훈련된 신경망 분류자에게 입력된다. 시스템은 서로 다른 복합적인 배경에서 다양한 크기와 자세를 가진 20명의 얼굴 데이터 베이스를 사용하여 실험한 결과 얼굴 인식기의 수행능력은 매우 작은 크기의 얼굴 이미지 외에는 87%에서 92%의 평균 인식율을 얻을 수 있었다.

ABSTRACT

This paper presents an integrated approach to unconstrained face recognition in arbitrary scenes. The front end of the system comprises of a scale and pose tolerant face detector. Scale normalization is achieved through novel combination of a skin color segmentation and log-polar mapping procedure. Principal component analysis is used with the multi-view approach proposed in[10] to handle the pose variations. For a given color input image, the detector encloses a face in a complex scene within a circular boundary and indicates the position of the nose. Next, for recognition, a radial grid mapping centered on the nose yields a feature vector within the circular boundary. As the width of the color segmented region provides an estimated size for the face, the extracted feature vector is scale normalized by the estimated size. The feature vector is input to a trained neural network classifier for face identification. The system was evaluated using a database of 20 person's faces with varying scale and pose obtained on different complex backgrounds. The performance of the face recognizer was also quite good except for sensitivity to small scale face images. The integrated system achieved average recognition rates of 87% to 92%.

키워드

log-polar, Principal component analysis, circular boundary

1. 서 론

얼굴 인식에 있어 해결 방법은 얼굴의 특징(나이, 성, 인종과 관련), 기하학(거리 및 관찰 위치),

이미지의 질(해상도, 조명도, 코의 비율에 따른 표시) 및 이미지 내용(배경)[1]등을 확인함으로써 고유의 변화성을 바탕으로 하는 것이다. 이러한 복잡성 때문에 대부분의 얼굴인식시스템[2]은 통제

* 관동대학교

접수일자 : 2003. 06. 30

된 환경을 기본으로 하고 근거리의 정면 얼굴을 식별하는 것이다.

최근, 정해지지 않은 환경에서 자유로운 얼굴의 표정 인식을 논의해 왔으며 조명, 포즈, 얼굴 표현에 따른 얼굴의 다양한 표정을 조절하기 위한 강건한 인식 흐름을 따름으로서 복잡한 구성에서 나타나는 얼굴 패턴을 구별하기 위한 얼굴검출 방법을 논의하였다.[3,5,6]

본 논문의 구성은 다음과 같다. 2절에서는 얼굴 검출 방법에 대해 기술하고, 3절에서는 얼굴 인식 방법에 대해 기술하며, 4절에서는 실험 결과를 분석하였으며, 5절에서 결론을 맺었다.

II. 얼굴 검출

본 논문에서는 Log-Polar 매핑의 고유 스페이스를 기반으로 얼굴검출방법에 대하여 논의하였다. 얼굴 검출기가 코의 위치를 찾은 다음 코를 중심으로 원형 그리드 매핑을 이용해 특징벡터를 추출한 후 특징 벡터는 얼굴 인식을 위해 훈련된 신경망에 입력되어진다.

2-1. 피부색 분류

피부색 분류는 20명의 피부에서 측정된 색을 기반으로 색채 또는 순수한 색채공간에서 수행되어지며, 피부색 참조 범위내의 색채성과 관련되지 않은 픽셀은 기타 배경으로 인해 분류된다. 얼굴의 면적은 크기를 의미하고 이는 크기 정규화 목적의 Log-Polar 매핑에 의해 이용되어진다.

2-2. Log-Polar 매핑

Log-Polar 매핑[4]은 스페이스 다양화 추출 그리드를 이용한다. 추출 포인트는 중심 원 N_r 과 원형선 N_θ 가 교차하는 지점에 위치하며, 원이 $\log r$ 에 위치하는 원형선은 동일한 각도상에 위치한다. Log-Polar 특징 벡터는 Log-Polar 이미지에서 열 또는 픽셀의 정렬에 의해 형성되며 이미지의 공간적 크기 또는 사물의 면적은 카메라로부터

떨어진 거리에 따라 다양하다. 크기가 정규화된 특징 벡터를 산출하려면 Log-Polar 그리드에서 가장 큰 원이 전체 사물을 덮을 수 있도록 조절되어야 한다. 그림 1은 얼굴 데이터베이스에서 나타난 3종류의 다른 얼굴 크기 및 얼굴 크기의 정규화된 Log-Polar 이미지를 보여주는 것이다. 얼굴 크기의 다양성에도 불구하고 Log-Polar 얼굴 이미지는 유사한 결과를 나타낸다.

2-3. 멀티뷰 얼굴 검출

얼굴 이미지들로부터 검출된 특징벡터는 m Log-Polar 집합으로 주어진다. 이 스페이스를 측정하는 기본 벡터는 PCA 과정에서 산출될 수 있다. 3차원 축소 및 합축 재현은 $k < m$ 고유 벡터를 유지함으로써 이루어질 수 있다. 저차원적 공간으로 투영되고 \vec{a} 로 재구성된 특징 벡터 \vec{a}_r 은 결과적으로 재구성 오류를 낳는다.

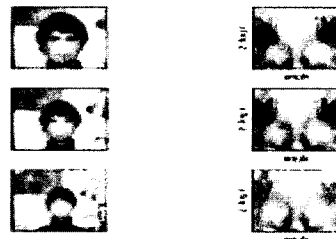


그림 4. 얼굴 이미지와 정규화된 Log-Polar 이미지
Figure 2. Face image and Normalization Log-polar image

자유로운 표정의 얼굴은 포즈의 다양성을 나타낼 수 있는데 이를 조절하기 위해 본 논문에서는 멀티뷰 접근 방식[7]을 사용하였다. 얼굴 데이터베이스를 바탕으로 각각의 다른 포즈에 따른 7개의 고유 스페이스를 구성하였다. 즉, 위 쳐다보기, 아래 쳐다보기, 우측 쳐다보기, 좌측 쳐다보기, 좌측으로 기울이기(평형 회전), 우측으로 기울이기, 정면으로 기울이기 등이다. 정면뷰와 관련한 이런 다양한 포즈는 $\pm 5^\circ \sim \pm 25^\circ$ 사이의 평형 기울기 회전을 전제로 하였다.

고유 스페이스를 구성하기 위해 설정된 뷰 각

도와 함께 각 배경에서 얼굴 이미지를 분류하였다. 얼굴 이미지에 대해서는 코 부분의 추출 그리드에 수동적으로 중점을 두고 얼굴을 덮을만한 정도로 크기를 조절함으로써 322 크기의 Log-Polar 벡터가 추출되었다. 데이터 에너지의 98%에 이르는 주요 구성요소가 유지되었으며 다양한 고유 스페이스는 1~4의 고유 벡터 범위 내에서 조절되었다.

2-4. 임계값 선정

Log-Polar 얼굴 특징 벡터로 형성된 고유 스페이스에서 나타나는 테스트 벡터의 재구성 오류는 고유 스페이스에 대한 테스트 벡터의 유사성을 의미하는 것이므로 임계값은 얼굴의 재현으로서 테스트 벡터를 수락 또는 거부하는 재구성 오류로 사용될 수 있는 것이다.

우선, 코 중심 부위의 10x10 이웃화소를 특정화한다. 재구성 오류의 최소화 및 그 오류가 임계값보다 적다는 가정하에 이 이웃의 모든 픽셀은 참된 검출의 신호가 될 수 있는 것이다. 위의 두 조건이 코 주위 10x10 밖에서 조건화된다면 거짓 검출이 발생하게 된다. 결과적으로 코 주위 내에서 최소화가 일어난다면 거짓 거부가 일어나게 되겠지만 이는 임계값보다 상위에 존재한다. 마지막으로, 최소화가 코 근방 외에서 발생할 경우에는 참 거부가 일어나게 되고 임계값을 초과하게 된다. 오류의 최소 가능성을 산출하는 임계값은 식-1과 같다.

$$p(d_e | \vec{a} \in \text{face}) P(\vec{a} \in \text{face}) = p(d_e | \vec{a} \notin \text{face}) P(\vec{a} \notin \text{face}) \quad (\text{식-1})$$

즉, d_e 는 재구성 오류이고 \vec{a} 는 Log-Polar 테스트 벡터를 의미한다.

선행 가능성 및 d_e 분배를 실험 환경에 주어진 사람들의 177개 이미지를 사용하여 측정할 수 있었다. 각 이미지에 대해 얼굴은 미리 정해진 800 픽셀 범위 안에서 검출된다. 10x10의 코 주변내 픽셀은 얼굴로 간주되며 $P(\text{face}) = \frac{100}{800}$ 과

$P(\overline{\text{face}}) = 1 - P(\text{face})$ 이다. $p(d_e | \vec{a} \in \text{face}) P(\vec{a} \in \text{face})$ 와 $p(d_e | \vec{a} \notin \text{face}) P(\vec{a} \notin \text{face})$ 를 산출하는 픽셀은 모든 이미지에서 형성되는 검출 범위를 통해 계산화되는 d_e 의 분포표를 통해 측정된다. 검출된 얼굴에 관한 몇몇의 예가 그림 2에서 보여지는데 이는 얼굴 검출이 정확히 설명될 수 있는 크기 및 포즈의 현상을 자세히 나타내준다.



그림 2. 얼굴 인식(얼굴영역과 비얼굴영역)
Figure 2. Face Recognition
(Face area and Non Face Area)

III. 얼굴 인식

얼굴인식방법은 원형 그리드 매핑[8]으로부터 생성된 특징벡터를 이용하며 특징벡터는 다음과 같이 형성된다.

1. 부드럽게 기울어진 정도를 이미지에서 추출한다.
2. 원형 그리드의 구성요소는 일정하게 공간을 차지하는 일련의 동심원들로 구성되어 있다. 이 동심원들은 원형형태로 각도에 따라 일정하게 스페이스를 차지하는 조그마한 조각으로 나뉘어진다. 그리드는 얼굴식별모듈이 예측하는 코의 위치에 초점을 맞춘다. 그리고 가장 큰 원이 인식기가 제공하는 크기에 맞게 조정해 준다.
3. 각 세그먼트의 경사도의 평균치는 벡터 요소를 구성하는데 이용된다.

3-1. 자기조직 맵을 이용한 분류

SOM[4]은 신경망과 같은 것이며 신경세포단위 뉴런이 미감독(unsupervised) 경쟁 배움의 프로세스를 통해 여러 가지 입력 시그널 유형에 따라 조절을 하게 된다. 분류를 위한 SOM맵을 사용하기

위해서는 레벨 조작 벡터를 가진 majority vote scheme를 이용하여 측정하게 된다. 맵 사이즈가 클 때에는 가능하며 조작된 벡터의 수가 적을 때에는 뉴런이 조작 벡터를 원하지 않는 경우에는 뉴런은 라벨을 만들어 내지 않는다. 시험용 벡터가 가장 가까이 있는 뉴런의 레벨로 분류되기 때문에 인식단계에서는 인식되지 않을 수가 있다. 우리는 2단계 분류체계를 실험해 보았는데 하나의 벡터가 가장 인접한 뉴런체계를 이용하여 인식을 못하는 경우라면 5-근접 뉴런체계를 이용하는 제2차 시도를 하게 된다.

3-2. 방사형기저함수 신경망에 의한 분류

SOM분류법에 RBF 신경조직을 이용하는 또 다른 분류방식을 사용했다. 실행하는 과정에서 조작된 SOM에 있는 뉴런의 중량을 은닉층(hidden layer)에 있는 가우시언 원형 기반 기능의 센터로 활용한다. 가우시언 기능의 확산은 SOM 뉴런사에 있는 최소한의 유클리드 기하학의 거리에 기초하여 결정된다. 은닉 노드(node)의 수 즉, m 은 SOM에 있는 뉴런의 수와 동일하다. 출력 노드의 수는 인식하고자 하는 데이터 베이스에 있는 사람들의 수와 동일하다. 은닉 노드를 출력 노드로 연결하는 중량은 각 출력 노드가 어느 특수한 사람에 조율을 맞추는 식으로 조절되도록 한다.

IV. 실험결과 및 고찰

통합된 인식시스템을 평가하기 위해서 20명의 얼굴 데이터 베이스를 사용했으며, 2가지 형태의 복합 배경 및 하나의 보통 배경에 대해서 실험하였다. 이미지는 130x120 이미지 샘플로 짧은 비디오 계열에서 나온 것이며, 다양한 회전과 크기를 가진 5,000여개의 이미지로 구성되었다.

사람 포즈의 변화를 6가지 유형으로 나누었다. 즉, 머리를 좌우로 회전(lr), 머리를 좌우로 비스듬히 기울인 형태(tlr), 머리를 위아래로(ud), 정면을 크게(ff), 정면을 중간 크기로(ms), 그리고 정면을 작은 크기로(ss)의 6가지 유형이다. 각 유형에는

각각 25개의 이미지가 있다. 장면 1과 복합 장면(장면 2와 3)으로 본 모든 대상의 검출률은 다음과 같다.

표 1. 각기 다른 유형의 검출 결과
Table1. Face detection or true detection rates

	lr	tlr	ff	ud	ms	ss	Ave
SCENE 1	99.4	99.8	100	98.0	99.8	85.4	97.1
SCENE 2	92.8	96.4	98.4	96.0	99.8	86.8	95.0
SCENE 3	94.4	98.2	99.4	96.0	98.4	87.4	95.6

표1은 각기 다른 장면에 대해 평균얼굴검출률이 95%와 97% 사이로 나타난 것을 보여주고 있다. 그러나 검출기 성능은 두 경우의 상황에서 다소간 차이가 있음을 확인할 수 있었다.

첫째, 유형 lr의 경우 이미지의 검출률은 복합장면의 경우에는 더욱 현저하게 감소하고 있다. 이는 머리가 좌우로 회전할 때 그림 3에 나와 있듯이 복합장면은 검출된 얼굴에 포함되기 때문이다. 이는 로그-폴라 벡터를 넣게 한 나머지 검출의 정확도를 떨어뜨린다.

둘째, 크기가 작은 얼굴의 경우(ss)는 보다 검출률이 떨어져 그 수치는 85%와 87% 사이로 나타났다. 이 수치는 중간 크기의 얼굴의 경우(ms)보다 11%에서 14% 떨어진 수치이다. 이는 얼굴 검출기에 대한 크기의 오차범위가 어떠한가를 나타낸다.



그림 3. 원형 로그-폴라 매핑
Figure 3. Circular log-polor mapping

통합얼굴인식분류를 평가하기 위해 얼굴검출기 결과는 분류를 위한 얼굴그림의 벡터를 이끌어 내는데 사용되었다. 원형 그리드 매핑과 SOM분류기용으로 사용되는 파라미터는[8]와 원형 그리드는 15x15 반경 및 각도이고 30x30의 SOM 사이즈는 적절한 것으로 나타났으며, 이 동일한 수치를

이용했다. RBF 네트워크에 대해서는 입력 벡터는 225였다. 은닉층은 900노드(30x30 SOM 뉴런)로 구성되어 있고 출력층은 20노드를 갖고 있다.

평가받은 분류기들 가운데는 RBF성능이 시종 일관 훨씬 나은 RBF가 있는가 하면, 다른 것은 평균인식률이 87%와 92% 사이인 것도 있다.

유형 lr에서 얼굴이미지를 위한 인식률이 약간 떨어지는 것은 그림 3에 설명된 바와 같이 원형형 그리드 매핑에서 복합배경이 포함되었기 때문이다. 인식률은 중간 크기 얼굴(ms)에서 작은 크기(ss)얼굴로 옮겨 갈수록 급격히 떨어지게 된다. 이와는 달리 큰 크기(ff)에서 중간 크기(ms)로 옮겨 갈수록 덜 심해진다.

표2. 20명에 대한 통합 얼굴인식률(%)
Table2. Integrated face recognition rates for 20 persons

	lr	tlr	ff	ud	ms	ss	Ave
Scene 1							
NNR	99.2	99.8	100	98.0	87.2	51.6	84.3
SOM1	85.0	90.8	92.9	92.9	49.6	20.4	62.1
SOM2	95.8	98.3	99.2	98.0	74.4	46.2	74.5
RBF	99.4	99.8	100	98.0	92.2	57.2	87.1
Scene 2							
NNR	84.4	96.4	96.6	96.0	81.8	59.4	85.8
SOM1	64.7	95.3	90.3	96.0	52.2	22.7	70.2
SOM2	81.6	96.4	97.6	96.0	77.4	56.6	84.3
RBF	92.8	96.4	98.4	96.0	89.8	69.9	90.6
Scene 3							
NNR	87.8	98.2	99.4	96.0	84.6	59.4	87.6
SOM1	59.3	88.3	81.7	91.3	47.2	23.1	65.2
SOM2	80.8	98.2	95.4	96.0	73.6	54.0	83.0
RBF	94.4	98.2	99.4	96.0	90.4	73.6	92.0

V. 결론

본 논문에서 제안한 방법은 복합장면에서도 영향을 받지 않은 자유로운 얼굴인식을 위한 통합접근 방법에 대해 논의하였다. 얼굴검출기는 피부색 구성, 로그-폴라 매핑, PCA로 구성되는데 다양한 크기와 포즈에 대한 얼굴 검출을 위한 멀티-뷰 접근을 하였다.

실험을 통해 본 얼굴검출율은 95%이상의 효과를 얻을 수 있었으며, 얼굴인식체계는 하나의 색다른 원형 그리드 방식으로 검출된 얼굴 이미지를 가지고 분류를 위해 신경망에 계속 입력하는 것이다.

얼굴에서 검출된 코의 위치에 초점을 맞춘 후 크기를 정규화 한 원형 그리드 매핑에서 검출한 얼굴 벡터는 분류용으로 사용되며 SOM과 RBF신경망을 비교하였다. RBF는 87%와 92%사이의 가장 좋은 인식률을 나타내었다. 데이터 베이스는 20명의 여러 가지 형태의 크기, 포즈 그리고 복합 장면을 담은 5,000개의 이미지의 20명분이 들어 있다. 인식율은 얼굴 사이즈가 너무 작을 때는 정규화 하는데 다소 민감한 부문이 있었다.

참고 문헌

- [1] B. Takacs and H. Wechsler. Detection of faces and facial landmarks using iconic filter banks. *Pattern Recognition*, 30(1):1623-1636, 1997.
- [2] R. Chellappa, C. Wilson, and S. Sirohey. Human and machine recognition of faces : A survey. *Proceedings of IEEE*, 83:705-740, May 1995.
- [3] M. Doi, K. Sato, and K. Chihara. A robust face identification against lighting fluctuation for lock control. *Proc. of 3rd Int'l Conf. on Face and Gesture Recognition*, pages 42-47, 1998.
- [4] H. Araujo and J. Dias. An introduction to the log-polar mapping. *Proc. of the Workshop on Cybernetic Vision 1996*, pages 139-144, 1997.
- [5] S. McKenna, S. Gong, and Y. Raja. Face recognition in dynamic scenes. *Proc. British Machine Vision Conference*, 1997.
- [6] A. Nefian and M. Hayes. Face detection and recognition using hidden markov models. *Proc. of IEEE Conf. on Image Processing*, 1:141-145, 1998.
- [7] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. *Proc. of IEEE*

Conf. on Computer Vision and Pattern Recognition, pages 84-91, June 1994.

[8] S. Ranganath, M. Lee, Y. Venkatesh, A. Kassim, and S. Ong. Face recognition with a radial grid representation and the self organizing map. Proc. of 5th Int'l Conf. on Control, Automation, Robotics and Vision, pages 1387-1391, December 1998.

저자 소개



박동희(Dong-Hee Park)

1996. 2 관동대학교 전자계산공학과 졸업(공학사)
1999. 2 관동대학교 대학원 전자통신공학과 졸업(공학석사)

2003.3~ 현재 관동대학교 대학원 전자통신공학과 박사과정 재학중
※ 관심분야 : 영상처리, 신호처리시스템, 영상압축



배철수(Cheol-Soo Bae)

1979. 2 명지대학교 전자공학 학과 졸업(공학사)
1981. 2 명지대학교 대학원 전자공학 학과 졸업(공학석사)

1988. 8 명지대학교 대학원 전자공학과 졸업(공학박사)
1991. 3~현재 산업자원부 공업기술기반 평가위원 및 심의위원
1998. 3~현재 강원도 정보화위원회 위원
1999. 3~2001. 5 관동대학교 이공대학 학장
2000. 3~2002. 2 관동대학교 양양캠퍼스 창업보육센터 소장
2001. 3~현재 해양정보통신학회 강원지부장
2002. 5~현재 중소기업 정보화 혁신 컨소시엄 추진 위원
2002. 6~현재 관동대학교 평생교육원 원장
1991~현재 관동대학교 정보통신공학부 교수
※ 관심분야 : 영상처리, 신호처리시스템, 영상압축