DBpia

논문 99-24-8B-17 한국통신학회논문지 '99-8 Vol.24 No.8B

차영상과 DCT 기반 특징 추출과 신경망을 이용한 얼굴 인식에 관한 연구

정회원 임 춘 환*, 고 낙 용**, 박 종 안*

A Study on Face Recognition using Neural Networks and Characteristics Extraction based on Differential Image and DCT

Chun-Hwan Lim*, Nak Yong Ko**, Jong-An Park* *Regular Members*

요 약

이 논문에서는 잡음에 대해 유연성이 있는 신경망과 차영상법-DCT를 이용한 얼굴인식 알고리즘을 제안한다. 동일환경(조도의 세기, 얼굴에서 카메라 까지의 거리)에서 연속적으로 두 개의 영상을 캡쳐했다. 이 때 한 영상은 얼굴을 포함하지 않고 다른 영상은 얼굴을 포함하게된다. 차영상 방법을 이용하여 두 개의 이미지로부터 얼굴영상과 배경영상을 분리하고 그 다음에 분리된 얼굴영역에서 사각영역을 추출하여 이 영역을 얼굴의 특징영역으로 이용하였다. 이 사각 영역은 눈, 코, 입, 눈썹 등이 포함된다. 다음으로 이 영역에 대해 DCT 연산을 수행한 후 특징 벡터를 추출하였고 추출된 특징벡터는 정규화되어 신경망의 입력벡터로 사용되었다. 시뮬레이션 결과 학습된 얼굴영상에 대해서는 100% 인식률을 보였고 학습되지 않는 얼굴 영상에 대해서는 92.25%의 인식률을 보였다.

ABSTRACT

In this paper, we propose a face recognition algorithm based on the differential image method-DCT. This algorithm uses neural networks which is flexible for noise. Using the same condition (same luminous intensity and same distance from the fixed CCD camera to human face), we have captured two images. One doesn't contain human face. The other contains human face. Differential image method is used to separate the second image into face region and background region.

After that, we have extracted square area from the face region, which is based on the edge distribution. This square region is used as the characteristics region of human face. It contains the eye bows, the eyes, the nose, and the mouth. After executing DCT for this square region, we have extracted the feature vectors. The feature vectors were normalized and used as the input vectors of the neural network.

Simulation results show 100% recognition rate when face images were learned and 92.25% recognition rate when face images weren't learned for 30 persons.

I. 서 론

최근 5년 동안 정면 얼굴의 영상을 인식하는 방법중 입력 영상에서 얼굴영역을 분리한 후 눈, 코, 입 등과 같은 특징을 추출하는 분야에서는 많은 발전이 있었다. 또한 Karhunen-Loeve(KL)변환에 근거한 통계적 방법이나 신경망을 사용한 얼굴 영상인식 분야도 많은 발전이 있었다[1,2].

측면 얼굴 영상을 이용한 얼굴 인식 방법에서는 전형적으로 측면 얼굴의 기준점 간의 거리를 특징으로 사용한다. 이러한 인식 방법은 특징 추출이 용이하다는 장점이 있으나 사진 촬영 시에 제약조건

* 조선대학교 전자정보통신공학부(japark@chosun.ac.kr) ** 조선대학교 제어계측공학과
논문번호 : 98552-1228, 접수일자 : 1998년 12월 28일

www.dbpia.co.kr

이 가해지는 단점이 있고 영상의 인식률을 높이기 위해서는 얼굴과 배경 부분을 정확히 분리해야 한다는 등의 많은 어려움이 따르기 때문에 현재까지는 연구가 많이 진행되지 않았으나, 일반적으로 정면 얼굴 영상을 사용한 방법보다 정확도가 높기 때문에 죄인들의 사진들 중에서 범인을 찾아내는 등의 문제에 주로 이용되었다.

입력된 영상에서 얼굴을 분할한 후 특징을 추출하는 방법은 영상의 윤곽, 특징점간의 거리와 같은 부분적인 특징을 추출하는 부분적 특징 추출법이 있다. 그리고 주어진 영상을 블록화한 후 특징추출 알고리즘에 의해 특징 파라미터를 추출해 내는 전역적 특징 추출 알고리즘이 있다. 특히 얼굴 부분의 특징추출은 전처리과정에서 얼굴영상을 적당히 변형시킨 후 얼굴부위에서 특징 위치를 찾는다.

특징 검출방법에는 자기 조직화 특징 지도를 이용하여 얼굴 특징을 검출하는 방법, snakelets를 이용하여 얼굴 후보 에지들을 찾아 얼굴을 검출하는 방법, 기하학적인 대칭에 의한 방법, 눈·코·입 등의 특징 템플리트와 영상의 상호상관을 이용하는 방법 그리고 영상을 푸리에 변환할 때 발생하는 주파수 특성과 DCT를 한 후 영상의 주파수 정보를 이용하여 특징을 추출하는 방법 등이 있다.

이러한 특징 값들을 이용하여 얼굴을 인식하는 방법 중 유클라디안 거리를 이용하는 방법은 시스템 구현이 용이하다는 장점이 있으나 데이터가 많을 경우 인식률이 떨어지는 단점이 있다. 또한 신경회로망을 이용하여 정면 얼굴 인식을 할 경우 입력 패턴의 잡음이나 왜곡, 크기의 다양성, 위치의 변화 등에 유연하게 적응하여 얼굴을 인식할 수 있는 장점이 있다.

신경망의 입력 수가 너무 많으면 많은 입력 노드가 필요하고 망의 연결 개수가 증가하기 때문에 얼굴 인식 시스템 구현이 어렵다. 그러나 입력 영상의 종류에 따라 공간 영역 또는 주파수 영역으로 변환한 후 물체에 대한 특징 파라미터를 추출하여 신경회로망의 입력으로 이용하면 소수의 입력벡터를 사용하기 때문에 노드 수와 연결선의 수를 줄일 수 있어 처리 시간 단축 및 시스템 구현이 용이하다.
본 논문에서는 일정한 조도 하에서 CCD 카메라로부터 취득된 얼굴영상을 차영상 기법을 이용하여 배경과 분리한 후 분리된 얼굴 영상의 에지 분포를 이용하여 얼굴영역을 사각영역으로 분리한다. 이 사각 영역에는 얼굴영상에서 가장 많은 정보를 가지고 있는 눈썹, 눈, 코, 입, 뺨 등이 포함된다. 이 영역을 DCT를 이용하여 연산한 후 특징파라메터를 추출한다. 추출된 각 특징파라메터를 정규화하여 다층신경회로망의 입력 벡터로 사용하여 학습을 수행한다. 이를 위하여 제 Ⅱ 장에서는 제안된 얼굴영상 인식 알고리즘에 대해 살펴보고 제 III 장에서는 제안된 방법을 이용한 시뮬레이션 결과들이 소개된다. 그리고 IV장에서는 그 결과를 분석하고 결론을 맺는다.

Ⅱ. 얼굴 영상 인식 알고리즘

이장에서는 현재까지 연구되었던 얼굴과 배경영역 분리에 대하여 몇 가지를 고찰하고 제안된 영상 인식 알고리즘을 소개한다.

2.1 얼굴과 배경영역 분리 연구 동향

주어진 영상에서 얼굴의 존재 유무를 가려내는 초기의 방법에서는 주로 에지(Edge) 영상을 사용하였다. Sakai는 입력 영상에서 추출된 에지 지도에 타원형의 형판을 적용하여 대략적인 머리 영역을 설정하고, 설정된 머리 영역 내에서 눈, 입의 에지 영상을 검사함으로써 최종적인 머리 영역을 추출하는 방법을 제시하였다[3]. 이 방법은 조명의 방향에 크게 영향을 받는 단점을 갖는다. 또한 Kelly는 입력 영상으로부터 자동적으로 머리와 몸의 윤곽을 추출하고, 계속해서 눈, 코, 입의 위치를 추출하는 하향식 영상 해석 방법을 소개하였다[4].

Craw 등은 주어진 영상에서 계층적인 크기의 영상과 형판을 사용하여 머리 영역을 추출하는 방법을 제시하였다[5]. Govindaraju는 복잡한 배경의 영상에서 얼굴의 위치를 찾기 위해 에지 영상으로 이루어진 머리 윤곽을 형판으로 사용하였다[6].

Sirohey는 배경이 있는 영상으로부터 Canny의 에지 탐색기로 추출된 에지 영상과 명도 영상을 함께 사용하여 얼굴을 분리하였다[7]. 이 방법은 제약을 가하지 않은 48개의 영상에 대해 약 80%의 정확도를 보여주었다.

2.2 차영상과 마스크를 이용한 얼굴분리

2.2.1 전처리

본 연구에서는 일정한 조도상태에서 차영상법을 이용하여 얼굴 영상과 배경영상을 분리한다. 분리를 위하여 먼저 CCD카메라로부터 동일거리에 있는 배경영상을 256×256 크기의 그레이 스케일(Gray Scale) 256 레벨로 취득하여 가우시안(Gaussian) 필

www.dbpia.co.kr

터를 이용하여 영상 내에 존재하는 잡음을 제거한다. 다음으로 얼굴영상을 입력 받는다.
정면 얼굴영상 및 배경영상 취득시 다음과 같은 제약조건을 부여한다.

1. 정면 얼굴영상 취득 및 배경영상 취득시 조도가 일정하여야 한다.
2. 턱수염이나 콧수염을 많이 기르지 않은 사람의 영상이어야 한다.
3. 영상 내에 얼굴의 위치는 상관없으나, 얼굴 전체가 포함되도록 취득되어야 한다.
4. 카메라로부터 얼굴까지의 거리는 일정하여야 한다.

2.2.2 차영상을 이용한 얼굴 분리

배경과 얼굴영역을 분리하기 위해서 잡음이 제거된 배경 영상과 얼굴이 존재하는 영상의 차를 구한다. 차영상법을 적용할 때 실제 환경에서는 동일한 카메라로 동일한 영역을 취득하였다 하더라도 같은 영역에 대해 일정한 픽셀값을 얻기가 어려웠고, 조도상태도 미소한 변화가 있었다. 따라서, 얼굴이 포함된 영상에서 얼굴영역 밖의 픽셀 값과 얼굴이 포함되지 않는 배경영상의 픽셀값이 정확히 일치하지 않으므로, 차영상을 구할 때에 다음과 같이 threshold값을 주어 동일 카메라로 얻어진 영상의 픽셀 값이 미소하게 변화하여도 잘 적응하도록 하였다.

$$\text{if } |Image1(x,y) - Image2(x,y)| < threshold \text{ then } Differ_Imgae(x,y) = 0$$
$$\text{else } Differ_Imgae(x,y) = |Image1(x,y) - Image2(x,y)| \quad (1)$$

단, *Image1* : 배경 영상,
Image2 : 얼굴이 존재하는 영상 ,
Differ_Image: 차영상

식(1)에서 배경의 픽셀 값과 얼굴의 픽셀 값이 비슷한 경우, 차영상법을 적용하면 얼굴 영역이 배경 픽셀 값에 의해 제거되고, 또한 배경 영역에서 threshold 값보다 크게 변화되는 부분이 차영상 내에서 잡음으로 발생하게되어 이를 직접 마스크로 사용할 수가 없다. 하지만 약간의 조도 변화와 빛의 반사에 의하여 생기는 잡음은 발생 확률이 적어 영상 내에서 미소 면적을 차지하게 된다. 따라서, 차영상을 이진영상으로 변환시킨 후, 메디안 필터를 이용하여 배경에서 나타나는 잡음을 제거하였다. 다음으로, 배경과 얼굴 영역의 픽셀값이 비슷하여 지워진 얼굴 경계면에 생기는 왜곡을 보정하기 위해 8×5 마스크를 이용하여 이진영상을 확장시킨다.

이진영상에서 확장은 물체의 경계면에서 0의 픽셀 값을 1로 바꾸는 방법으로, 지정된 이웃 픽셀 값들을 이용하여 픽셀을 추가시킴으로써 이루어진다. 출력 이미지에서 지정된 픽셀의 상태(0 또는 1)는 입력 이미지에서 부합되는 이웃 픽셀에 대한 규칙성을 적용시켜 결정한다. 즉, 입력 픽셀의 이웃 픽셀에서 어떤 픽셀값이 1이면, 그에 대한 출력 픽셀 값은 1이고 그렇지 않으면 0의 값을 가진다. 이웃 픽셀들은 임의로 정할 수 있는 형태와 크기로 0과 1의 값을 가진 행렬로 표시되는 행렬 구성 요소에 의해 대치되어진다. 따라서, 이진영상의 경계면에서 0의 값이 1로 확장되어진다. 팽창된 이진영상 좌, 우에서 픽셀 값을 조사하여 마스크를 생성한 후, 마스크를 얼굴이 존재하는 원 영상에 투영하여 배경과 얼굴 영역을 분리해 내었다.

2.2.3 얼굴영상에서 사각영역 분리

사람의 얼굴에서 대부분의 특징정보는 눈썹, 눈, 코, 입, 뺨에 존재하게된다. 실제 얼굴을 인식할 때 이들의 정보만 이용하여도 양호한 인식률을 얻을 수 있다. 따라서 분리된 얼굴영상에서 특징추출을 하기 전에 이들 영역만이 포함되도록 얼굴영상을 다시 일정한 사각영역으로 분할한다. 여기서 눈썹, 눈, 코, 입 등을 각각 분리할 수 있지만 이들을 전부 포함한 사각영역으로 분할한 이유는 DCT 변환을 이용하여 특징추출을 하기 위함이다.

사각 영역 분할을 위해 먼저 Sobel 연산자를 이용하여 에지를 검출한다. 검출된 에지 성분들의 가로 방향 히스토그램을 보면 눈 영역에서 가장 많은 에지 분포가 나타난다. 이 특성으로 눈썹의 상한 위치를 찾는다.

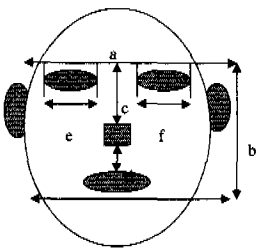


그림 1. 정규화된 거리

또한 카메라로부터 일정한 거리에 있는 사람의 얼굴의 크기는 비슷하므로, 그림 1과 같이 눈썹, 눈,

www.dbpia.co.kr

코, 입, 뺨이 모두 포함되도록 얼굴의 중심선을 기준으로 상단부는 눈썹의 근처에서 하단부는 입술근처에서 동일한 크기로 분리한다.

2.2.4. DCT에 의한 특징 파라메터 추출

이산 코사인 변환(DCT; Discrete Cosine Transform)은 많은 영상 압축 알고리즘의 기초가 된다. 이산 푸리에 변환에 비해 DCT의 장점은 복소수 계산을 하지 않아도 된다는 것이다[8]. 또한 DCT는 화상 데이터의 분포에 잘 맞기 때문에 자주 이용되고 있다. DCT는 식(2)로 표현된다.

$$F(u) = \frac{2}{N} C(u) \sum_{x=0}^{N-1} f(x) \cos \frac{(2x-1)u\pi}{2N}$$

$$f(x) = \frac{2}{N} \sum_{x=0}^{N-1} C(u) F(u) \cos \frac{(2x+1)u\pi}{2N}$$

$$u = 0 \text{인 경우 } C(u) = 1/\sqrt{2}$$

$$u \neq 0 \text{인 경우 } C(u) = 1 \qquad (2)$$

DCT를 화상에 이용하는 경우에는 m화소×n화소를 1구획으로서 2차원 DCT를 행하고있으며 8화소×8화소 또는 16화소×16화소의 구획이 자주 사용되고 있다. M×N 화소의 2차원 DCT와 IDCT는 아래의 식 (3)와 (4)와 같으며 영상에 대해 DCT를 수행하면 영상에 대한 대부분의 정보는 일반적으로 저주파수 영역에 집중한다.

$$F(u,v) = \frac{2}{\sqrt{MN}} C(u)C(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \cdot \cos \frac{(2x+1)u\pi}{2M} \cos \frac{(2y+1)v\pi}{2N} \qquad (3)$$

$$f(x,y) = \frac{2}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} C(u)C(v)F(u,v) \cdot \cos \frac{(2x+1)u\pi}{2M} \cos \frac{(2y+1)v\pi}{2N} \qquad (4)$$

$$C(\gamma) = 1/\sqrt{2}, \ \gamma = 0 \text{ 일 때}$$

$$1 \ \gamma > 0 \text{ 일 때}$$

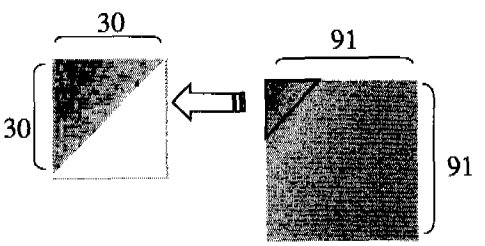


그림 2. 대각방향 30×30 특징추출

256×256 영상에서 검출된 얼굴의 사각 영역은 91×91의 크기를 갖는다. 식 (3)을 이용하여 사각영역에 대해 DCT를 수행한 후 91×91의 계수중 저주파수 영역에서 얼굴인식을 위한 특징벡터를 추출한다. 즉, 그림 2와 같이 30×30의 좌상삼각행렬을 추출해 내어 같은 주파수대(대각방향)의 DCT 계수에 대한 절대치 합을 구하여 특징 벡터를 추출한다.

2.2.5 신경망을 이용한 얼굴인식

얼굴 영상 인식 분야에서 신경망은 성별 분류, 얼굴 인식, 얼굴 표정의 인식 등에 주로 사용되어왔다. 초기에 신경망을 얼굴 인식에 적용한 대표적인 사람으로는 Kohonen을 들 수 있으며, 그는 소규모의 얼굴 영상 데이터베이스를 사용하여, 입력 영상에 잡영이 많거나 영상의 일부가 손실된 경우에도 정확한 재현이 가능함을 보였다[9]. Brunelli 등은 눈썹의 두께, 코와 입의 넓이, 6개의 턱 반경 등과 같은 수치적 특징 16가지를 사용하여 성별 분류를 위한 신경망을 개발하였다[10]. 각각의 성별 분류를 위해서 2개의 HyperBF 신경망이 사용되었고, 입력 영상들에 대해 눈의 위치를 기준으로 크기와 방향 정규화를 수행하였다. 신경망의 입력값으로는 위에서 소개된 16개의 특징이 사용되었으며, 두 개의 신경망 중에서 결과 값이 높은 쪽을 최종적인 분류 결과로 선택하였다. 21명의 남성과 21의 여성에 대해서 실험한 결과 학습 데이터에 대해서는 92.5%, 실험 데이터에 대해서는 87.5%의 인식률을 보였다. 박래홍 등은 눈썹의 길이, 눈썹 사이의 길이, 눈에서 입까지의 거리, 코와 입사이의 거리의 각각 비로 4개의 특징 벡터를 추출하여 정규화를 수행한 후 신경망 입력으로 하였다. 이 실험에서는 남자 14명과 여자 6명에 대해 각각 4장의 영상을 이용하여 학습데이터 및 실험 데이터에 대해 100%의 인식률을 보였다[11].

본 연구에서는 분리된 사각영역에 대해 DCT 연산을 수행한 후 특징 벡터를 추출한다. 추출된 특징벡터들을 신경망에 입력시키기 위해서는 최소값을 -1, 최대값을 1로 정규화 시켜야한다. 그런데 DCT 계수의 대각 합을 살펴보면, 일반적으로 첫 항과 두 번째 항의 값이 다른 값들과 비교하여 상당히 크다는 것을 알 수 있다. 따라서 선형적으로 정규화를 시키면 다른 값들이 매우 작은 값을 가지며, 서로 비슷한 값들만을 가지게된다. 선형적으로 정규화를 수행할 경우, 최대값 및 최소값과 중간에 분포하는 값들의 차가 매우 크므로 정규화된 값들은 최대값

www.dbpia.co.kr

과 최소값을 제외하고 0 근처에 분포하게 된다. 따라서, 시그모이드 함수의 특성상 미소한 차이의 입력값에 대해서는 출력이 거의 일정하게 되어 그 특성을 잃어버리게되므로 1과 -1사이에 각 특징 벡터들간 시그모이드 함수의 특성을 살리기 위해 어느 정도 거리가 생기도록 비선형 정규화를 수행하였다. 일반적으로 입력 벡터들이 -1과 1사이에 어느 정도 군집이 되어있을 때 신경망의 학습이 잘되는 경향이 있지만, 다른 패턴의 입력 벡터들간 거리가 매우 짧을 경우에는 오히려 다른 패턴들과의 구별이 잘 안되므로 -1과 1사이에 고르게 분포하도록 정규화를 수행하였다. 정규화된 신경망 입력 벡터를 각각의 인물에 대하여 동일한 구성을 가지는 다층 신경망에 입력하여 학습시키고, 각 학습된 신경망의 가중치 벡터와 출력벡터를 데이터 베이스화하였다.

Ⅲ. 시뮬레이션

본 연구에서는 일정한 조도상태에서 CCD카메라로부터 동일거리에 있는 얼굴영상을 256×256 크기의 그레이 스케일(Gray Scale) 256 레벨로 취득하여 PC상에서 가우시안(Gaussian) 필터를 이용하여 영상 내에 존재하는 잡음을 제거한 후 차영상법을 이용하여 얼굴 영상과 배경영상을 분리하였다. 다음으로 에지 검출을 한 후 에지 특성에 의해 분리된 얼굴영상을 사각영역으로 분리후 DCT를 수행하였다.

DCT 계수 행렬로부터 특징 벡터를 추출하여 각 인물에 대해 다층 신경망에 학습시킨 후 이를 데이터 베이스화하여 입력 영상에 대해 인식을 수행하였다. 이에 대한 얼굴 영상 인식 알고리즘 순서도는 그림 3과 같으며 이를 위한 시스템 구성도를 그림 4에 보였다.

실험을 위해 CCD 카메라로 30인에 대해 각각 8장을 취득하여 전처리를 거쳐 영상의 잡음을 제거하였다. 그림 5는 입력영상인 배경영상과 얼굴영상을 나타내고 있다.

이때, 각 한 장의 얼굴 영상에 대해 배경 영상을 한 장씩 입력하였으며, 배경 영상과 얼굴 영상에 대한 카메라의 위치는 일정하게 하였다.

그림 5의 배경 영상과 얼굴 영상의 차영상을 구하여 이진 영상으로 변환한 영상을 그림 6에 보였다. 그림 6의 이진 영상을 살펴보면 얻고자 하는 얼굴 영역의 경계면이 많이 지워진 것을 볼 수 있으며, 또한 약간의 조도의 변화에 의해 잡음이 생긴 것을 볼 수 있다.

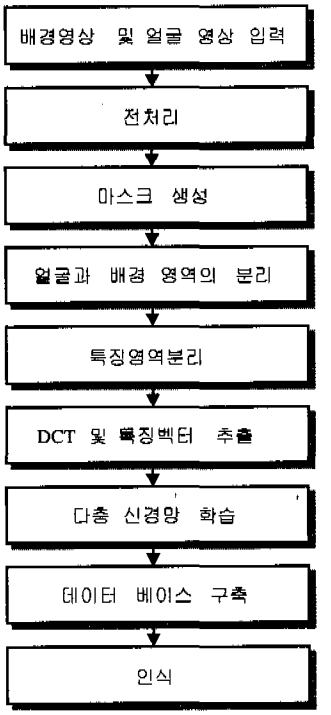


그림 3. 얼굴 인식 알고리즘 순서도

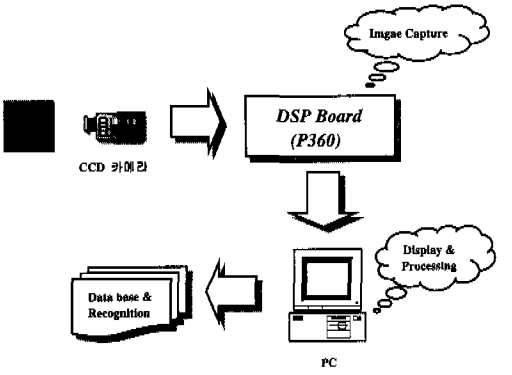
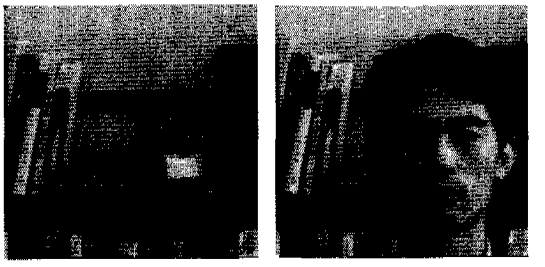


그림 4. 얼굴 인식을 위한 시스템 구성도



(a) 배경 영상 (b)얼굴 영상

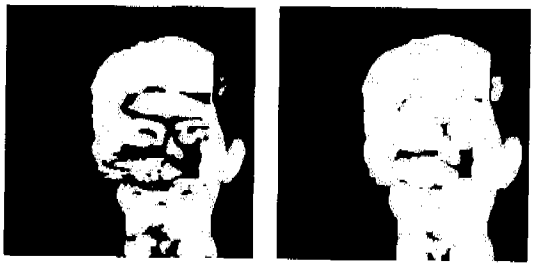
그림 5. 입력 영상

www.dbpia.co.kr

(a) 차영상 (b) 이진 영상

그림 6. 차영상과 이진영상

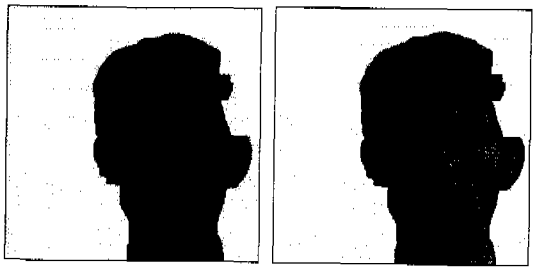
따라서 메디안 필터를 이용하여 잡음을 제거한 후 좌우로 픽셀값을 조사하는 방법을 이용하여 마스크 영상을 생성시킨 후 얼굴 영상의 경계면을 확장시켰다. 그림 7에 메디안 필터를 통과시킨 영상과 경계면에 대해 확장시킨 결과를 보였다.



(a) 필터링된영상 (b) 확장된 영상

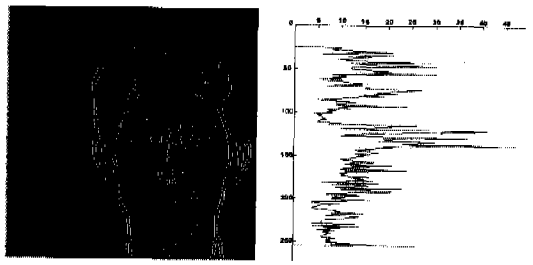
그림 7. 필터링된 영상과 확장된 영상

확장된 이진 영상은 얼굴 내의 픽셀들이 배경 픽셀에 의해 제거되어 이를 그대로 마스크로 사용할 수가 없으므로 마스크를 생성하였다. 마스크 영상은 타원방정식 또는 이웃 픽셀들을 조사하여 생성할 수도 있지만, 본 논문에서는 알고리즘을 간단하게 하기 위해 좌우에서 픽셀값을 조사하는 방법으로 생성하였으며, 이렇게 얻어진 마스크 영상을 얼굴이 존재하는 원영상에 투영하여 배경과 얼굴을 분리하였다.



(a) 마스크 영상 (b) 분리된 얼굴 영역

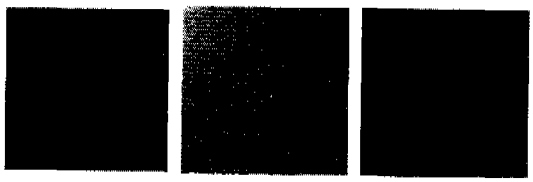
그림 8. 마스크 영상과 분리된 얼굴 영역



(a) 에지 영상 (b) 에지의 합

그림 9. 분리된 얼굴영역에 대한 에지 영상 및 에지합

분리된 얼굴 영역에서 에지 영상을 구하여 가로 방향으로 에지 개수 합을 보면 눈의 위치에서 가장 큰 값을 가진다는 것을 확인할 수 있다. 눈의 위치에서 픽셀값을 조사하여 눈썹의 상한선을 찾은 후, 눈썹으로부터 일정 크기의 윈도우로 DCT를 수행하고자 하는 영역만을 분리하였다. 분리된 영역에 대해 DCT를 수행하여 주파수 영역별로 DCT계수들의 합을 구하여 30개의 특징 벡터를 추출하였으며, 추출된 특징 벡터들을 비선형으로 정규화 시킨 후 신경회로망의 입력벡터로 하였다. 그림 10에 특징 영역과 특징 영역에 대한 DCT영상 및 추출된 특징 벡터 30×30 좌상삼각행렬에 대한 IDCT영상을 보였다. 그림 10의 (c)에서 추출된 특징 벡터들만으로 복원된 영상을 보면, 저주파수만을 이용한 관계로 고주파 성분, 즉 에지 성분들이 약해졌지만 육안으로도 원래영상과 IDCT영상이 동일 인물임을 알 수 있었다.



(a) 특징 영역 (b) DCT 영상 (c) IDCT 영상

그림 10. 특징 영역과 DCT, IDCT 영상

또한 특징벡터는 구해진 DCT계수의 좌상삼각행렬 30×30 블록내의 계수 중에서 동일 주파수대인 대각방향의 절대치 합을 구해 30개의 특징벡터를 추출한다. 특징 벡터 추출의 당위성을 확인하기 위해 동일인에 대해 4장의 각각 다른 영상을 획득하여 특징영역을 검출한 후 DCT 계수를 구한 결과 특징 벡터간의 RMSE는 +0.002061627로 나타났다. 아래 그림 11에서 알 수 있듯이 동일인에 대해 구해진 특징 벡터의 데이터 곡선은 거의 일정함을 알 수 있다.

www.dbpia.co.kr

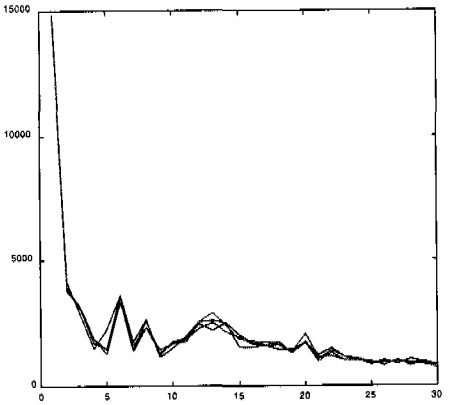


그림 11. 특징 벡터

신경망 구조는 최적의 신경망 구성을 얻기 위해 반복 실험으로 인식률, 학습 속도를 비교하여 2층으로 하였으며, 입력층, 은닉층 및 출력층 뉴런 수를 각각 30, 15, 5개로 하였다. 또한 학습 계수로써 학습률은 0.1, 시그모이드 함수는 0.9로 하였다. 학습 종료는 출력벡터와 목표벡터의 오차가 모두 0.005 보다 적을 때까지 학습하였다. 시뮬레이션 결과 학습시 반복 횟수(Iteration)는 6489~4427이었다. 학습하지 않은 영상, 즉 실험 영상에 대한 출력 벡터 및 목표 벡터와의 오차를 표 1에 보였다. 표 1를 살펴보면 모두 오차 한계 0.005내에 들어옴을 볼 수 있었다. 또한, 다른 사람에 대하여는 오차 한계 0.005 이상이 되는 경우도 있었으며, 오차 한계 안에 들어오는 영상도 있었다. 오차 한계 안에 들어오는 경우에는 그 가중치 벡터에 대한 학습시의 출력 벡터와 가장 유사한 패턴을 찾음으로써 인식을 수행하였다.

표 1. 출력 벡터와 출력 오차 값의 예.

출력 벡터	오차	출력벡터	오차
실험영상 1		실험영상 2	
+0.995584595	+0.004415405	+0.995680071	+0.004319929
-0.995496581	-0.004503419	-0.995733043	-0.004266957
-0.995502969	-0.004497031	-0.995651185	-0.004348815
-0.995520841	-0.004479159	-0.995566135	-0.004433865
+0.996104228	+0.003895772	+0.996023186	+0.003976814
실험영상 3		실험영상 4	
+0.995873922	+0.004126078	+0.995137077	+0.004862923
-0.995891887	-0.004108113	-0.995009101	-0.004990899
-0.995816873	-0.004183127	-0.995188857	-0.004811143
-0.995702012	-0.004297988	-0.995300707	-0.004699293
+0.996101442	+0.003898558	+0.995813701	+0.004186299

다음 그림 12는 학습된 영상을 나타내고 있으며 그림 13는 학습되지 않는 테스트 영상을 보여준다.

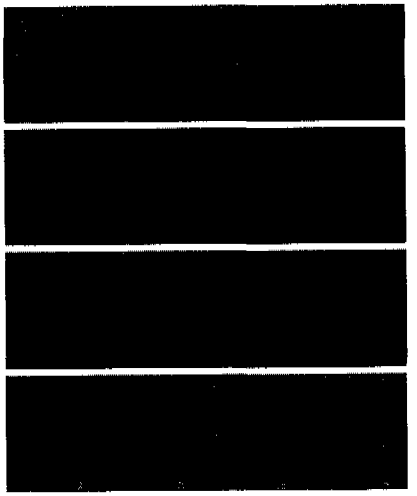


그림 12. 학습된 영상의 예

실험 결과 30인에 대한 학습 영상 120장에 대하여 모두 인식되었으며, 비학습 영상, 즉 실험 영상 120장에 대하여 9장이 오인식 되었다. 오인식된 영상을 살펴보면 표정이 많이 변화되거나 상하좌우 약 10° 이상 회전된 얼굴 영상에 대해 오인식이 일어났음을 알 수 있었다.

그림 13. 테스트 영상의 예

또한, 데이터 베이스화 되어있지 않은 사람을 인식할 경우에는 데이터 베이스에 저장된 다른 사람으로 인식하는 경우도 발생하였다.

www.dbpia.co.kr

Ⅳ. 결 론

본 논문에서는 차영상-DCT기반 특징 추출과 오류 역전파 학습 알고리즘을 이용하여 정면 얼굴 영상을 인식하는 알고리즘을 제안하였다. 차영상법에 의해 마스크를 생성하였고 생성된 마스크를 이용하여 배경영상과 얼굴영상이 쉽게 분리되었다. 시뮬레이션을 통해 모든 입력 영상에 대해 얼굴영상이 100% 분리됨을 확인하였다. 다음으로 얼굴영상의 에지 특성을 이용하여 눈, 눈썹, 코, 입, 뺨이 포함되도록 얼굴 영상을 사각영역으로 분리하여 이를 DCT 변환한 후 특징벡터를 추출하였다. 추출된 특징 벡터를 1과 -1사이의 값으로 정규화한 후 신경망의 입력벡터로 하였다. 이 때 인식률, 학습 속도를 비교하여 층의 수는 2층으로 하였다. 학습 계수로써 학습률은 0.1, 시그모이드 함수는 0.9로 하였고 학습 종료는 출력벡터와 목표벡터의 오차가 모두 0.005 보다 적을 때까지 학습하였다. 시뮬레이션 결과 학습시 반복 횟수는 6489~4427이었다.

30인을 대상으로 각각 학습된 영상 4개와 학습되지 않은 실험 영상 4개에 대해 시뮬레이션을 수행하였다. 실험 결과 30인에 대한 학습된 영상 120장에 대하여 모두 인식되었으며, 비학습 영상, 즉 실험 영상 120장에 대해서는 9장이 오인식 되어 평균 96.25%의 인식률을 보였다.

그러나 학습을 통해 데이터 베이스화 되어있지 않은 사람을 인식할 경우에는 데이터 베이스에 저장된 다른 사람으로 인식하는 경우도 발생하였다. 따라서 보안 시스템 등, 보다 정확한 인식을 요하는 응용에서는 다른 인식 방법과 병행할 필요가 있다.

향후 연구과제는 동영상을 이용하여 거리 정보 및 움직임 정보 등을 이용하여 표정의 변화 및 거리에 관계없이 특징점을 추출하고 인식률도 높일 수 있는 알고리즘을 개발해야 할 것이며 보다 많은 인물에 대한 데이터 베이스를 갖는 얼굴 인식 시스템을 개발해야 할 것이다.

감사의 글

이 논문은 한국과학재단 지정 지역협력연구센터인 조선대학교 수송기계부품 공장자동화 연구센터의 연구비 지원에 의해 연구되었음.

참 고 문 헌

[1] 이성환 : 얼굴영상의 이해를 위한 얼굴 특징의 자동 추출: 한국정보과학회 HCI' 96 학술대회 발표 논문집, 1996, pp.205~219.

[2] R. Chellappa, Charles L. Wilson, and Saad Sirohey : Human and Machine Recognition of Faces : *A Survey, IEEE*, Vol.83, No.5, 1995, pp.704~740

[3] T. Sakai, M. Nagao, and Fukibayashi : Line extraction and pattern recognition in a photograph : *Pattern Recognition*, Vol. 1, pp. 233~248, 1969.

[4] M.D. Kelly : Visual identification of people by computer : Tech. Rep. AI-130, Stanford AI Proj., Stanford, CA, 1970.

[5] I. Craw, H. Ellis, and J. Lishman : Automatic extraction of face feature : *Pattern Recognition Lett.*, Vol. 5, pp. 183~187, 1987.

[6] V. Govindaraju, S. N. Srihari, and D. B. Sher : A computational model for face location : *Proc. 3rd Int. Conf. on Computer Vision*, pp. 718~721, 1990.

[7] S. A. Sirohey : Human face segmentation and identification : Tech. R. CAR-TR-695 : Center for Automation Research, Univ. of Maryland, 1993.

[8] Randy Crane : *A simplified approach to image processing*, New Jersey : Prentice-Hall, 1997.

[9] T. Kohonen, *Self-Organization and Associative Memory* : Berlin, Springer, 1988.

[10] R. Burunelli and T. Poggio : HyperBF networks for gender classification : *Proc. DARPA Image Understanding Workshop*, pp. 311~314, 1992.

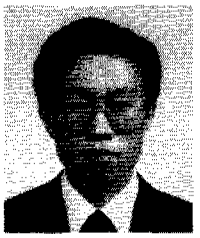
[11] 박래홍, 이상영, 성영국: 지식에 기초한 특징 추출과 역전파 알고리듬에 의한 얼 굴 인식 : 전자 공학회논문지 제31권 B 편 제7호, pp. 119~128, 1994

www.dbpia.co.kr

임 춘 환(Chun-Hwan Lim) 정회원

1991년 2월 : 조선대학교 전자공학과 졸업
1993년 8월 : 조선 대학교 전자공학과(공학석사)
1999년 2월 : 조선대학교 전자공학과(박사수료)
1999년 3월~현재 : 동강대학 전자과 겸임교수
<주관심 분야> 영상통신, 의료영상, 패턴인식, 신경망 및 로보틱스

고 낙 용(Nak Yong Ko) 정회원



1985년 : 서울대학교 제어계측공학과(공학사)
1993년 : 서울대학교 제어계측공학과(공학박사)
현재 : 조선대학교 전기제어계측공학부 부교수

박 종 안(Jong-An Park) 정회원

1975년 2월 : 조선대학교 전자 공학과(공학사)
1978년 2월 : 조선대학교 전자 공학과(공학석사)
1986년 2월 : 조선대학교 전자 공학과(공학박사)
1993~현재 : 조선대학교 전자정보통신공학부 교수
<주관심 분야> 디지털 신호처리, 패턴인식, 디지털 통신.

www.dbpia.co.kr