

다층신경망을 이용한 얼굴 검색 알고리즘

박철용, 이인범, 이병용, 박수봉

동신대학교 정보통신공학과

The Facial Image Searching Algorithm using Multi-Layer Neural Networks

Chul-Woong Park, In-Bum Lee, Byong-Yong Lee, Soo-Bong Park
Department of Information & Communication Eng., Dongshin University

요약

이 논문에서는 다층 신경망을 이용한 얼굴 인식 알고리즘을 제안한다. 이를 위해 동일한 환경, 즉 일정한 조도 하에서 카메라로부터 동일거리에 있는 영상을 256×256 크기의 그레이 스케일(Gray Scale)로 취득하여 영상 내의 잡음을 가우시안(Gaussian) 필터를 이용하여 제거한다. 인물이 없는 배경영상과 같은 위치에 사람이 포함된 영상을 취득하여 두 입력 영상의 차를 구한다. 추출된 차영상에 대해 팽창과 수축과정을 통해 차영상 내부의 잡음을 완전히 제거한 후 픽셀 값을 조사하여 마스크를 생성한다. 그리고 마스크를 인물이 포함된 최초의 입력 영상에 투영하여 배경부분에서 인물영역만을 분리하고 부호화를 위해 인물영역의 에지성분을 조사하여 코, 눈, 귀, 입을 차례대로 검출한다. 데이터 단순화를 위해 특징값이 밀집된 얼굴영역의 T형 영역을 정의하고 이 영역을 얼굴인식을 위한 특징영역으로 정의한다. 정의된 영역에 대해 DCT를 수행하여 특징벡터를 추출한다. 특징벡터는 정규화되어 다층신경망의 입력벡터로 이용된다. 시뮬레이션을 통해 제안된 알고리즘에 의한 인식률향상과 속도 향상을 입증한다.

1. 서론

사람의 얼굴을 정확히 검출하여 인식하는 검증 시스템은 여러 분야에서 다양하게 사용되어질 수 있다. 현재 은행, 각종 출입관리시스템, 자동차급기, 업무용 컴퓨터 등에 있어서의 보안은 열쇠나 ID-Card, 그리고 비밀번호 등을 통해서 이루어지고 있다. 그러나 이와 같은 시스템은 열쇠나 카드의 분실과 위조, 비밀번호의 유출 등으로 인하여 개인의 신용 및 정보가 악용될 수 있는 위험에 노출되어 있다. 또한 범죄 발생 이후에도 계좌추적이나 흔적수사, 그리고 폐쇄회로 카메라에 찍힌 용의자에 대한 수사에 있어서 정확한 범죄자의 신원확보의 어려움을 가지고 있다.

따라서 사람의 얼굴 인식 시스템을 도입함으로써 분실 등으로 인한 악용의 위험으로부터 보호하고, 수동적인 보안에서 벗어나 범죄자의 신원확인 등 능동적인 접근을 가능하게 한다.

본 논문에서는 보다 집중된 얼굴영역을 추출하여 이를 DCT를 이용하여 정면 얼굴영상의 특징벡터를

추출한다. 이를 정규화한 후 다층 신경망에 입력함으로써 연산시간과 학습 시간을 단축하고 시스템구현을 용이하도록 하였다. 이를 위하여 2장에서는 DCT와 다층 신경망의 제반 이론에 대하여 고찰하고, 3장에서는 차영상 기반 얼굴검색 알고리즘과 DCT를 이용한 특징추출 알고리즘을 제안한다. 4장에서는 시뮬레이션을 실행하고 그 결과를 분석한 후 5장에서 결론을 맺는다.

2. 관련 이론

2.1 신경망

신경회로는 상호 연결된 뉴론에 의해 임의의 입력 N 차원 공간을 출력 M 차원 공간으로 매핑하는 것으로 생각할 수 있으며, 입력 및 출력 공간과 매핑 특성에 따라 연상기억, 필터, 변환, 분류, 최적화 등의 기능을 수행할 수 있다. 그림 1에서와 같이 다층 퍼셉트론은 입력층과 출력층 사이에 하나 이상의 중간층이 존재하는 신경망이다.

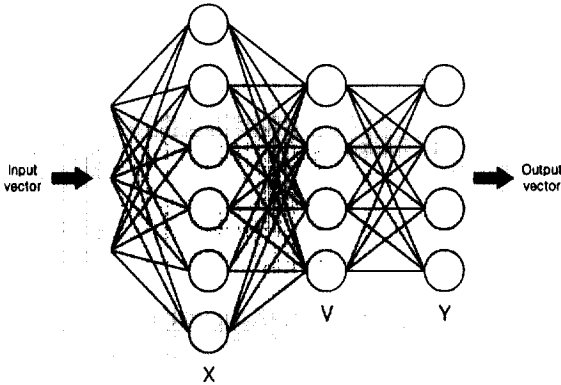


그림 1. 다층 퍼셉트론의 구조

이 때 입력층과 출력층 사이의 중간층을 은닉층(hidden layer)이라 부른다. 네트워크는 입력층, 은닉층, 출력층 방향으로 연결되어 있으며, 각 층내의 연결과 출력층에서 입력층으로의 직접적인 연결은 존재하지 않는 전방향(feedforward)의 네트워크이다. 역전파 학습 알고리즘은 최소자승(least mean square) 알고리즘의 비선형적인 확장이다. 역전파 학습 알고리즘은 미분의 반복규칙(chain-rule)을 여러번 반복 적용하여 확률근사치 프레임워크와 관련 지음으로써 유도해낼 수 있다.

역전파 학습 알고리즘의 기본 원리는 다음과 같다. 입력층의 각 유니트에 입력패턴을 주면, 이 신호는 각 유니트에서 변환되어 중간층에 전달되고 최후에 출력층에서 신호를 출력하게 된다. 이 출력값과 기대값을 비교하여 차이를 줄여나가는 방향으로 연결강도를 조절하고, 상위층에서 역전파하여 하위층에서는 이를 근거로 다시 자기층의 연결강도를 조정해나간다. [1-3]

2.2 DCT (Discrete Cosine Transform)

DCT는 많은 영상 압축 알고리즘의 기초가 되며 화상 데이터의 분포에 잘 맞기 때문에 자주 이용되는데, 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,

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

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

$$C(r) = \begin{cases} \frac{1}{\sqrt{2}} & r=0 \text{ 일때} \\ 1 & r > 0 \text{ 일때} \end{cases} \quad (1)$$

x 및 u가 가로방향울, y 및 v가 세로방향울 나타내고 있다. 식 (1)의 코사인과 같은 2차원 기본함수를 화상을 이용하여 표현하는 경우를 특히 기본 화상이라고 한다. u = 0, v = 0 의 기본 화상은 직류성분을 나타내고, u가 클수록 높은 수평 주파수를, v가 클수록 높은 수직 주파수를 나타낸다. 즉, 기본 화상에서는 주파수별로 화상을 볼 수 있다. [4-7]

3. 얼굴검색 알고리즘과 특징 검출

3.1.1. 얼굴인식 알고리즘

본 연구에서는 동일장소와 일정한 환경하에서 취득된 두 입력영상을 차영상법을 이용하여 배경영상에서 인물영역을 분리하고 얼굴을 검색한다. 먼저 카메라로부터 동일배경과 동일거리에 있는 영상을 256×256 크기의 그레이 스케일(Gray Scale)로 취득하고, 영상 내의 잡음 제거를 위해 가우시안(Gaussian) 필터를 사용한다. 취득된 두 입력영상은 배경영상과 배경영상 내에 인물이 포함된 영상이다. 이 두 영상에 차영상법을 적용하게되는데, 실험환경이 아닌 실제 환경에서는 같은 영역에 대해 일정한 픽셀 값을 얻기가 어렵고, 조도도 일정하지않으므로 차영상과정에서 임계값을 주어 동일 카메라로 얻어진 영상의 픽셀 값이 미소하게 변화해도 원하는 영상을 구할 수 있도록 한다.

$$\begin{aligned} \text{if } |Image1(x, y) - Image2(x, y)| < \text{threshold} \\ \text{then } Differ_Image(x, y) = 0 \quad \text{else} \\ Differ_Image(x, y) = \end{aligned} \quad (2)$$

|Image1(x, y) - Image2(x, y)|
 단, Image1 : 배경 영상,
 Image2 : 얼굴이 포함된 입력영상,
 Differ_Image : 차영상

식 (2)에서 배경의 픽셀 값과 얼굴의 픽셀 값이 비슷한 경우, 차영상법을 적용하면 얼굴 영역의 일부가 배경 픽셀 값에 의해 제거되게되고, 배경 영역에서도 임계값보다 큰 부분은 잡음으로 나타나게 되므로 마스크로 사용하기엔 부적절하다. 그러나 이러한 잡음

은 미소한 확률이나 크기를 가지므로 이진화하여 영상을 팽창(dilation)시켜 제거할수 있다. 영상의 팽창은 물체의 경계면에서 0의 픽셀 값을 1로 바꾸는 방법으로 지정된 이웃 픽셀 값들을 이용하여 픽셀을 변화시킨다. 그러나 마스크로 사용하기위한 차영상 영역이 지나치게 팽창되면 마스크가 분리하고자하는 인물영역보다 커질 우려가 있다. 따라서 얼굴영상의 크기에 맞게 영상을 축소(erosion)한 후 좌, 우에서 픽셀 값을 조사하여 마스크를 생성한다.[8]

그리고 마스크를 인물이 포함된 입력영상에 투영하여 배경영역이 분리된 인물영역을 추출하게 된다. 아래 그림 1은 얼굴영역을 분리하기 위한 순서도를 나타낸다.

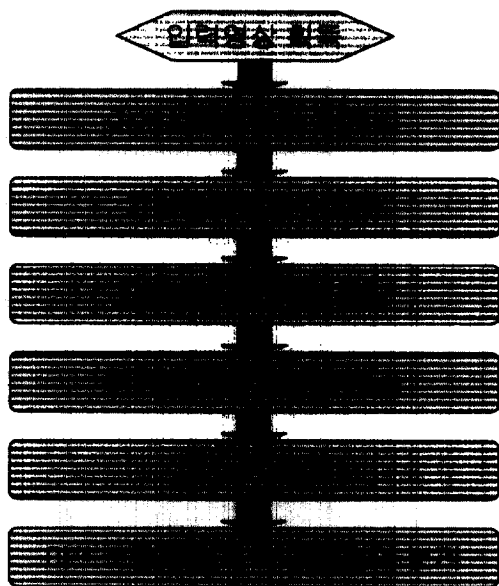


그림 1. 얼굴영역 분리를 위한 순서도

3.1.2 얼굴의 특징 검출

특징을 가지는 대부분의 정보는 사람의 얼굴에서 눈썹, 눈, 코, 입, 뺨에 존재한다.[9]

특징검출을 용이하게 하기위해 검출된 얼굴영상에서 소벨 연산자를 이용하여 에지를 검출한다. 검출된 에지 성분들의 수평성분과 수직 방향 성분을 이용하여 코의 영역을 찾는다. 그림 2는 얼굴의 특징을 검출하기 위한 순서도를 나타낸다.

일정한 거리에서 추출한 사람의 얼굴의 크기는 비슷하므로 그림 3과 같이 얼굴영역에 대해 픽셀의 거리를 지정하여 눈썹, 눈, 코 그리고 입을 각각 분할

하고 DCT 변환을 이용하여 특징을 추출하기 위해 눈썹, 눈, 코, 뺨 그리고 입이 포함되도록 T형영역을 정의하여 얼굴인식을 위한 특징영역으로 한다.

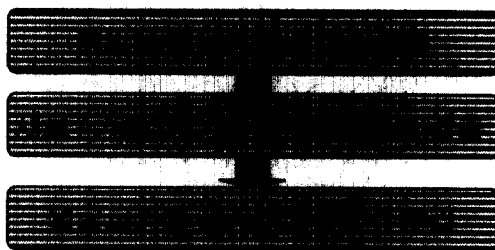


그림 2. 특징 검출을 위한 순서도.

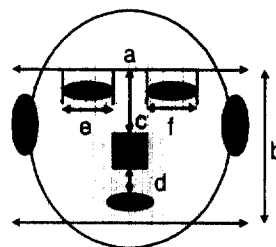


그림 3. 특징 추출을 위한 정규화된 거리와 T영역

3.2. DCT를 이용한 특징 추출 알고리즘

DCT를 이용한 특징 파라미터의 추출 과정은 먼저 $[256 \times 256 \times 2^8]$ 의 해상도의 두 입력 영상 신호를 차영상과 에지특성을 이용하여 91×91 픽셀크기의 특징영역으로 분리하여 2차원 DCT를 수행한 후 DCT 계수들을 추출하여 영상에서 다른 인물과 구분되어지는 대부분의 정보를 포함하는 저주파수 영역에서 특징을추출한다.[10]

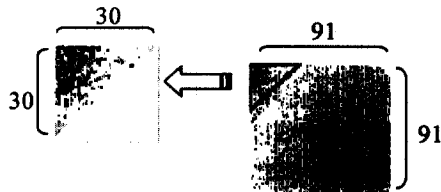


그림 4. 계수 행렬에서 특징추출(대각방향)

그림 4는 91×91 픽셀의 DCT 영상에서 30×30 픽셀 영역의 계수 행렬 중에서 좌상 삼각행렬만을 추출해 같은 주파수대(대각방향)의 DCT계수 합을 통해서 30개의 특징 벡터를 추출한다. 그림 5는 DCT를 이용한 정규화 입력 벡터의 추출과정을 나타낸다.

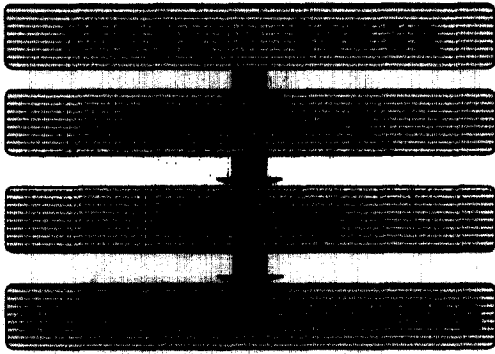


그림 5. 정규화 입력벡터 추출

4. 시뮬레이션 및 결과 분석

4.1. 얼굴 영상의 특징 검출에 대한 결과

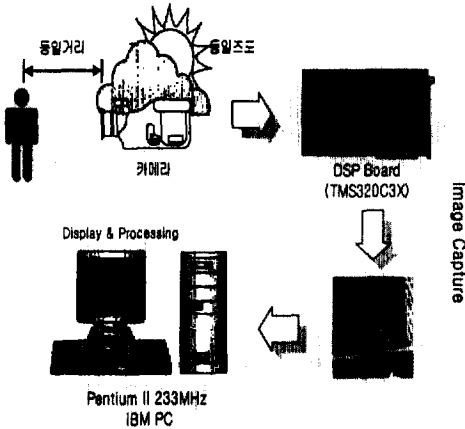


그림 6. 얼굴 인식 시스템

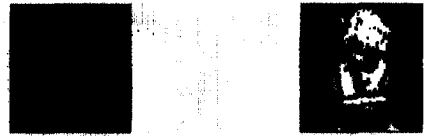
본 연구에서는 일정한 조도하에서 그림 6과 같이 고정된 카메라로부터 동일한 거리의 영상을 펜티엄 II PC(233MHz)에 장착된 TMS320C3X 보드를 이용하여 그레이 스케일(Gray Scale)의 256×256 크기의 입력영상을 취득하여 시뮬레이션을 수행하였다.

이를 위해 그림 7과 같은 배경영상과 인물이 포함된 배경영상의 두 개의 입력 영상을 가우시안(Gaussian) 필터링으로 영상 잡음을 제거하고 전처리하여 차영상을 통해 배경영상에서 얼굴을 검출하였다.

8인을 대상으로 비디오로부터 입력영상으로 배경 및 얼굴이 포함된 영상을 취득하였는데, 얼굴을 쉽게 검출하기 위하여 배경영상을 먼저 취득하고 다음으로 얼굴이 포함된 영상을 취득한다.



그림 7. 입력영상 (배경영상과 인물포함영상)



(a) 차영상

(b) 이진영상

그림 8. 차영상과 이진영상

취득된 두 입력영상에 대해 그림 8(a)와 같이 차영상을 구했다. 이를 마스크로 사용하기 위하여 그림 8(b)와 같이 이진화한 후에 영상의 경계면에 대해 팽창(dilation)과 축소과정을 거쳐서 마스크를 생성하였다.

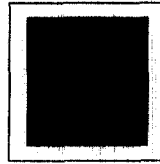


그림 9. 에지 영상

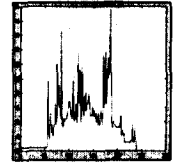


그림 10. 에지 합

다음으로 마스크에 의해 분리된 영상의 잡음을 제거하기 위해 메디안 필터를 수행한 후 소벨 연산자를 이용하여 그림 9와 같이 에지를 검출하였다.

그리고 그림 10에서 알 수 있듯이 에지의 분포 특성은 얼굴 영상의 측면의 수직 방향에서 제일 큰 값을 가졌다. 따라서 이 지점에서 출발하여 중심 방향으로 에지 값을 검색하여 코의 위치를 검출하고 수평 방향에서 눈을 검출한 후 인간 얼굴에 대한 지식에 근거하여 눈썹과 눈 그리고 입의 위치를 검출해서 T형으로 특징을 검출하였다.

4.2 DCT를 이용한 얼굴 인식 성능 실험

DCT를 이용하여 특징 추출을 하기 위해 인간얼굴에 대한 일반적인 지식에 근거하여 눈썹, 눈, 코, 입, 이 포함되도록 얼굴의 중심선을 기준으로 눈과 코 부위를 에지 성분의 수직 수평 방향 분포에 의해 분

리하였다. 하단부도 마찬가지로 에지 분포와 얼굴크기를 근거로 T형 영역을 검출하여 얼굴의 특징영역으로 하였다. 그림 11은 T형으로 검출된 얼굴의 특징영역이다. 제안된 방법을 이용하면 전체의 입력영상에서 영상 내의 얼굴 위치에 무관하게 특징영역을 100% 검출할 수 있다.

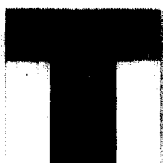


그림 11. 검출된 특징영상

아래 그림 12는 검출된 얼굴영역에 대한 IDCT 영상이고 그림 13은 DCT 계수를 영상으로 나타낸 그림이다.



그림 12. IDCT 영상

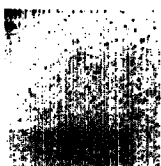


그림 13. DCT 영상

위의 그림 13의 DCT 영상에서 좌상삼각행렬에 영상의 특징이 집중되므로 본 연구에서는 좌상삼각행렬만을 인식에 이용한다. 동일 인물에 대해 조금씩 변화된 준 그림 11과 같은 4개의 샘플 영상을 사용하여 DCT를 수행한 후 아래와 같이 특징벡터를 각각 추출하였다.

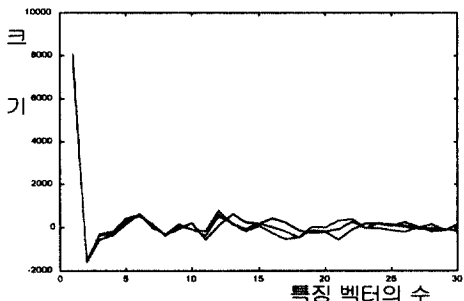


그림 14. 특징 벡터의 분포 특성

그림 14는 4개의 샘플영상에 대해 추출된 특징벡터의 분포를 나타낸다. 그림을 보면 최대값은 DCT의 직류성분이며 동일인에 대한 샘플 영상의 특징벡터들

은 거의 일정하게 분포함을 알 수 있다.

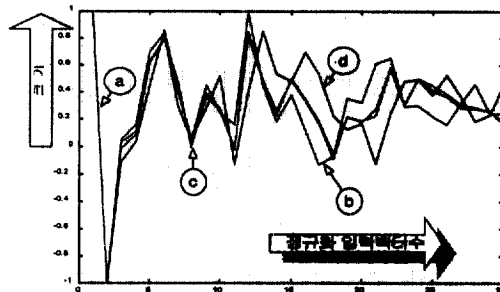


그림 15. 정규화된 입력 벡터

위 그림 15는 특징벡터들에 대한 정규화 입력 벡터의 분포 곡선을 나타내고 있는데 특징벡터의 정규화로 인해 정규화 벡터간에 약간의 오차가 발생했으나 각각의 입력 영상에 대한 정규화 벡터는 그래도 일정하게 분포함을 알 수 있다.

이와 같이 추출된 정규화 입력벡터를 그림 16에서처럼 다층 신경망에 학습 시켰는데 입력층, 은닉층 및 출력층 뉴런수를 각각 30, 15, 5개로 하였다. 또한 학습 계수로써 학습률은 0.7, 시그모이드 함수는 0.3으로 하였고 출력벡터와 목표벡터의 오차가 모두 0.005 보다 적을 때까지 학습하였다.

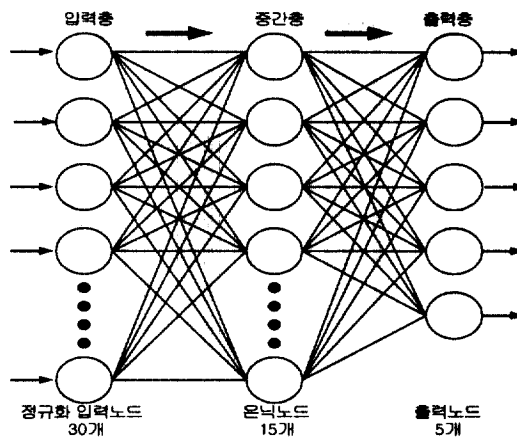


그림 16. 얼굴인식을 위한 신경망 구조



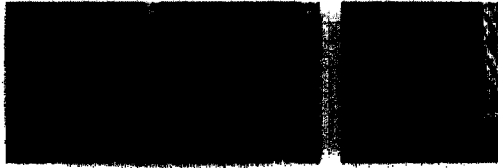


그림 17. 실험 영상의 예

위의 그림 17은 실험영상의 예이다. 실험에서는 8명의 영상에 대해 1인당 학습 영상 4개, 비학습 영상 4개를 가지고 실험한 결과 전체학습 영상에 대해서는 100%의 인식률을 보였고 비학습 영상에 대해서는 88%의 인식률을 나타내었다. 따라서 전체적으로 94%의 인식률을 보였다.

5. 결 론

본 연구에서는 일정한 조도 상태에서 취득된 두 입력영상에 대해 차영상을 적용하여 배경영상에서 얼굴을 분할한후 차영상을 팽창과 수축을 하여 얼굴영역 내의 잡음을 완전히 제거하고 좌, 우에서 픽셀값을 조사하여 생성한 마스크로 얼굴을 분리하였다.

메디안 필터를 이용하여 얼굴영역의 잡음을 제거한 후 에지를 검출하고 수직 방향 분포를 조사하여 T형의 특징영역을 검출하였다.

시뮬레이션 결과 DCT 알고리즘을 이용한 경우, 인식률은 학습영상에 대해 100%, 실험영상에 대해 88%, 전체적으로 94%의 인식률을 보였다. 그리고 T형만을 처리함으로써 얼굴전체를 처리하는 것보다 데이터량을 줄일 수 있었다.

이러한 연구결과는 현금자동지급기, 출입관리 시스템, 도난방지 시스템 등 보안 및 고속의 인증이 필요한 분야에 얼굴인식 시스템의 이용이 가능함을 보여주고 있다.

향후 연구과제는 조도의 변화가 심한 현장에서 제약을 최대한 줄이면서 양호한 특징 검출 및 인식 성능을 개선할 수 있는 알고리즘과 검출 속도가 향상된 인식방법에 대해 연구할 예정이다.

참 고 문 헌

- [1] 이병용, "자동차 번호 인식을 위한 역전파 학습 알고리즘 개선", 석사 학위논문, pp. 5-23, 1998.
- [2] Maureen Caudill and Charles Butler, "Understanding Neural Networks", Computer Explorations, MIT Press, 1992.
- [3] William, R.J. and Zipser, D., "A learning

algorithm for continually running fully recurrent neural networks", Neural Computation, v. 6(4), pp509-520, 1989.

- [4] J.S. Lim, "Two-Dimensional Signal and Image Processing", Prentice-Hall, 1990.
- [5] Chan, S.C., and Ho, K.L. "A new Two - Dimensional Fast Cosine Transform Algorithm", IEEE Trans. Sig. Proc., 39(2), pp. 481-485, 1986.[8]
- [6] 정연식, 이입건, 송홍엽, 박규태, "고속 이산 여현 변환(DCT) 알고리즘", 한국정보과학회논문지, pp. 65-68, 1998.
- [7] 이명길, 이주신, "DCT와 신경회로망을 이용한 패턴인식에 관한 연구", 한국통신학회논문지, Vol. 22 No. 3, pp. 481-492, 1997.
- [8] 유정현, 김형범, 김희년, 임춘환, 양원일, 박종안, "비디오 열에서 차영상 기반 얼굴 분할 및 특징 검출 알고리즘", 한국통신학회 학술발표지, pp. 84-90, 1999.
- [9] 최정일, 김수환, 이필규, "얼굴의 특징 추출을 위한 얼굴의 구성요소 분할" 한국정보과학회논문지, pp. 215-226, 1998.
- [10] 이문호, 염재훈 "C언어를 이용한 영상신호처리", 문영사. pp. 31-250, 1998.

* 본 연구는 정보통신부 대학 S/W 연구센터 지원 사업 연구비에 의하여 연구되었음.