

論文2000-37TE-3-15

웨이브렛 변환과 신경망 기반 얼굴 인식

(Facial Image Recognition Based on Wavelet Transform and Neural Networks)

林春煥*, 李尙薰*, 片爽範*

(Chun Hwan Lim, Sang Hun Lee, Suk Bum Pyeon)

요약

본 연구에서는 웨이브렛 변환과 신경망 기반 얼굴 인식 알고리즘을 제안한다. 이 알고리즘은 일정한 조도 상태에서 두 개의 영상을 그레이 레벨로 취득하고 가우시안 필터를 이용하여 영상 내에 존재하는 잡음을 제거한 후 배경영상과 얼굴이 포함된 입력영상의 차를 구하여 차영상에 대해 축소와 팽창과정을 통한 전처리 과정을 거치게 된다. 그리고 팽창 영상으로부터 마스크를 생성하여 마스크를 얼굴이 존재하는 원 영상에 투영하여 배경과 얼굴을 분할하고 분할된 얼굴영상의 에지를 조사하여 눈, 코, 입, 눈썹 그리고 뺨이 포함된 사각 모양의 특징영역을 검출한다. 그리고 특징영역에 대해 이산 웨이브렛 변환을 수행하여 특징벡터를 추출하고 정규화한 후 신경망의 입력벡터로 하여 학습에 의한 인식을 수행한다. 시뮬레이션 결과 학습된 영상에 대해서는 100%의 인식률을 보였고 학습되지 않는 실험적 영상에 대해서도 92%의 인식률을 나타내었다.

Abstract

In this study, we propose facial image recognition based on wavelet transform and neural network. This algorithm is proposed by following processes. First, two gray level images is captured in constant illumination and, after removing input image noise using a gaussian filter, differential image is obtained between background and face input image, and this image has a process of erosion and dilation. Second, a mask is made from dilation image and background and facial image is divided by projecting the mask into face input image. Then, characteristic area of square shape that consists of eyes, a nose, a mouth, eyebrows and cheeks is detected by searching the edge of divided face image. Finally, after characteristic vectors are extracted from performing discrete wavelet transform(DWT) of this characteristic area and is normalized, normalized vectors become neural network input vectors. And recognition processing is performed based on neural network learning. Simulation results show recognition rate of 100 % about learned image and 92% about unlearned image.

I. 서론

뉴스 방송과 같이 배경이 잘 변하지 않는 경우 영상

을 연속적으로 전송할 때 사람의 얼굴과 같이 가변하는 객체를 분할하여 전송하면 훨씬 전송 효율을 높일 수 있다. 또한 영상 인식을 위해서도 사람의 모습을 분할하는 작업은 필수적이다.

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 최종

* 正會員, 東岡大學 電子情報科

(Dept of Electronics & Information Eng., Dongkang College)

接受日字:2000年8月14日, 수정완료일:2000年9月18日

적인 머리 영역을 추출하는 방법은 제시하였다^[1]. 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^[2].

Craw 등은 주어진 영상에서 계층적인 크기의 영상과 형판을 사용하여 머리 영역을 추출하는 방법을 제시하였다^[3]. Govindaraju는 복잡한 배경의 영상에서 얼굴의 위치를 찾기 위해 에지 영상으로 이루어진 머리 윤곽을 형판으로 사용하였다^[4].

Sirohey는 배경이 있는 영상으로부터 Canny의 에지 탐색기로 추출된 에지 영상과 명도 영상을 함께 사용하여 얼굴을 분할하였다^[5]. 이 방법은 제약을 가하지 않은 48개의 영상에 대해 약 80%의 정확도를 보여주었다. 그러나 이와 같은 얼굴 분할 방법은 추출 알고리즘이 복잡하고 또한 얼굴이 정확히 분할되지 않으면 얼굴 인식이 용이하지 않다.

분할된 영상을 인식할 경우 특징 추출 방법에는 영상의 윤곽이나 어느 특정부분의 특징을 추출하는 지역적 특징 추출법과 영상을 블록화한 후 처리하여 특징 파라미터를 추출해 내는 전체적 특징 추출 알고리즘이 있다. 특징 위치 검출방법에는 기하학적인 대칭에 의한 방법, 눈·코·입 등의 특징 템플리트와 영상의 상호관계를 이용하는 방법, snakelets를 이용하여 얼굴 후보 에지들을 찾아 얼굴을 검출하는 방법, 자기 조직화 특징 지도를 이용하여 얼굴 특징을 검출하는 방법, FFT, DCT와 DWT와 같은 주파수 영역에서 특징을 추출하는 방법 등이 있다.

템플리트를 이용하는 방법은 여러 방향의 얼굴 특징을 검출하는데 적합하나 연산량이 큰 것이 단점이나 주파수 영역에서의 특징 추출방법은 영상을 주파수 영역으로의 변환하여 효율적으로 특징을 검출할 수 있다.

특징을 추출하여 인식하는 방법은 주로 유클리디안 거리를 이용한 방법과 신경회로망을 이용하는 방법을 많이 이용하고 있다. 전자는 시스템 구현이 용이하다는 장점이 있으나 데이터가 많을 경우 인식률이 떨어지는 단점이 있다. 후자를 이용하여 패턴 인식을 할 경우 입력 패턴의 잡음이나 왜곡, 크기의 다양성, 위치의 변화 등에 유연하게 적응하여 패턴을 인식할 수 있는 장점이 있으나 입력 수가 너무 많으면 그에 따른 많은 노드 수가 필요하며 연결 개수도 증가하기 때문에 패턴 인식 시스템 구현에 있어서 어려움이 있다. 그러나 입

력 대상에 따라 공간 영역 또는 주파수 영역으로 변환한 후 물체에 대한 특징 파라미터를 추출한 후 신경회로망을 이용하면 소수의 입력벡터를 사용하기 때문에 노드 수와 연결선의 수를 줄일 수 있어 처리 시간 단축 및 시스템 구현이 용이하다.^{[6][7][8][9]}

따라서 본 연구에서는 두 개의 입력영상 이용이 가능한 경우 차영상을 적용하여 마스크를 생성하고 이를 이용하여 얼굴을 분할하는 방법을 제안한다. 또한 얼굴의 특징영역을 정의하여 특징추출시 연산량을 줄인다. 그리고 웨이블릿 변환을 이용하여 정교한 특징벡터를 추출하고 동시에 특징 벡터의 수를 감소시켜 신경망 학습시 필요한 학습 데이터를 최대한 줄이고 잡음과 미세한 조도변화에 유연하게 대처하도록 한다.

이를 위하여 2장에서는 차영상기반 얼굴 분할 알고리즘을 제안하고 3장에서는 웨이블릿 기반 얼굴인식 알고리즘을 제안한다. 그리고 제 4장에서는 시뮬레이션을 수행하고 제 5장에서 결론을 맺는다.

II. 차영상 기반 얼굴 분할 알고리즘

이 장에서는 동일한 환경에서 자동적으로 입력영상에서 얼굴을 분할하고 분할된 얼굴에서 특징을 검출하는 알고리즘을 제안한다.

1. 차영상과 마스크를 이용한 얼굴 분할

일정한 조도 상태에서 취득된 배경영상과 얼굴이 포함된 두 입력영상에서 차영상을 적용하여 얼굴을 분할한다. 이를 위해 동일거리에 있는 정경을 256×256 크기의 그레이 스케일 256 레벨로 취득하여 가우시안 필터링에 의해 영상 내에 존재하는 잡음을 제거한다.

그리고 배경 영상과 얼굴이 포함된 입력 영상의 차를 구하는데 얼굴이 포함된 입력영상에서 얼굴 밖의 픽셀 값과 얼굴이 포함되지 않는 배경영상의 픽셀 값이 정확히 일치하지 않으므로 다음 식 (1)과 같이 임계값을 주어 동일 카메라로 얻어진 영상의 픽셀 값이 미소하게 변화해도 원하는 영상을 구할 수 있도록 한다.

$$\begin{aligned} & \text{if } |Image1(x, y) - Image2(x, y)| < threshold \\ & \text{then } Differ_Image(x, y) = 0 \\ & \text{else} \\ & Differ_Image(x, y) = |Image1(x, y) - Image2(x, y)| \end{aligned} \quad (1)$$

단, *Image1* : 배경 영상, *Image2* : 얼굴이 포함된 입력영상, *Differ_Image*: 차영상

배경의 픽셀 값과 얼굴의 픽셀 값이 비슷한 경우, 차영상을 적용하면 얼굴 영역의 일부가 배경 픽셀 값에 의해 제거되고 배경 영역에서 임계값보다 크게 변화되는 부분은 차영상 내에서 잡음으로 발생하게 되므로 이를 직접 마스크로 사용할 수가 없다. 그러나 약간의 조도 변화와 빛의 반사에 의하여 생기는 잡음은 발생 확률이 적어 영상 내에서 미소 면적을 차지하게 되므로 차영상을 이진화 하면서 영상을 축소시킨다. 영상의 축소는 물체의 경계면에서 픽셀 값을 바꾸는 방법으로 지정된 이웃 픽셀 값들을 이용하여 픽셀을 변화시킨다. 출력 이미지에서 지정된 픽셀의 상태(0 또는 1)는 입력 이미지에서 부합되는 이웃 픽셀에 대한 규칙성을 적용시켜 결정한다. 즉, 입력 픽셀의 이웃 픽셀에서 어떤 픽셀 값이 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.

축소된 이진이미지는 얼굴영상의 크기에 맞게 팽창이 되어야한다. 다음으로 픽셀 값을 조사하여 마스크를 생성한 후 마스크를 얼굴이 존재하는 원 영상에 투영하여 배경에서 얼굴을 분할한다. 그림 1은 얼굴을 분할하기 위한 순서도를 나타낸다.

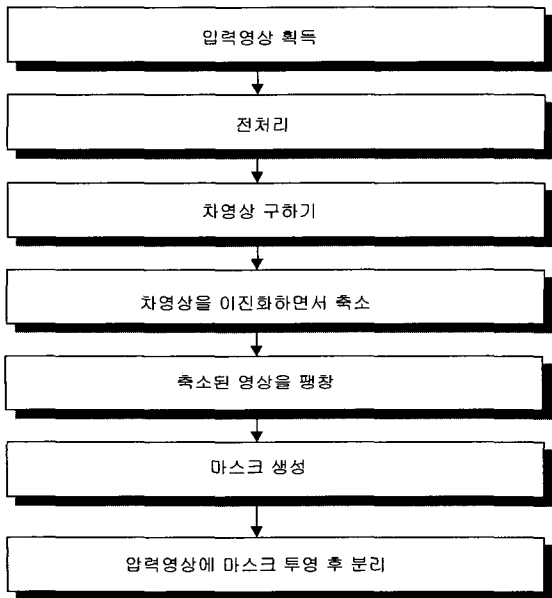


그림 1. 얼굴 분할을 위한 순서도
Fig. 1. Flowchart for facial image segmentation.

2. 얼굴의 특징 검출

사람의 얼굴에서 대부분의 특징정보는 눈썹, 눈, 코, 입, 뺨에 존재한다. 따라서 이 영역만을 포함하는 국소 영역을 선정한 후 얼굴의 특징을 검출한다. 분할된 얼굴영상의 특징검출을 위해 sobel 연산자를 이용하여 에지를 검출한 후 에지성분들의 수평성분과 수직 방향 성분을 이용하여 눈의 영역과 얼굴의 중심을 찾는다. 그림 2는 얼굴의 특징을 검출하기 위한 순서도를 나타낸다.

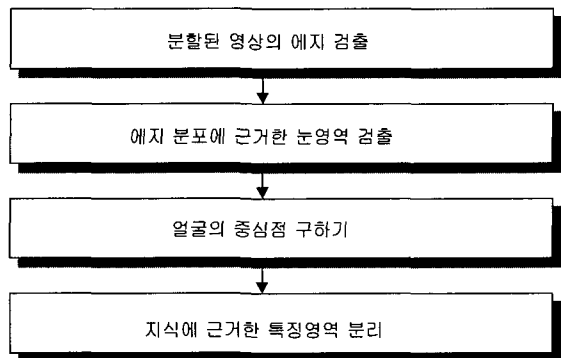


그림 2. 특징 검출을 위한 순서도
Fig. 2. Flowchart for characteristic detection.

일정 거리에 있는 사람의 얼굴의 크기는 비슷하므로 그림 3과 같이 눈썹, 눈, 코, 뺨 그리고 입이 포함되도록 사각영역을 정의하여 얼굴인식을 위한 특징영역으로 한다. 이렇게 추출된 특징영역에 대해 이산 웨이브렛 변환을 한 후 특징을 추출한다. 그림 3에서 a는 특징영역의 가로 방향 거리이고 b는 특징영역의 세로 방향 거리를 나타낸다.

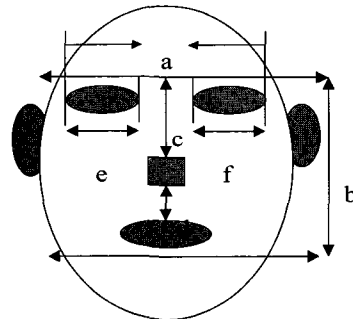


그림 3. 특징 검출을 위한 정규화된 거리
Fig. 3. The normalized distance for characteristic detection.

Ⅲ. 웨이브렛(DWT) 기반 얼굴 인식 알고리즘

이장에서는 웨이브렛 변환영역에서 얼굴의 특징을 추출한 후 신경망을 이용하여 얼굴영상을 인식하는 방법을 제안한다. 다음 그림 4는 얼굴인식 알고리즘 순서도를 나타낸다.

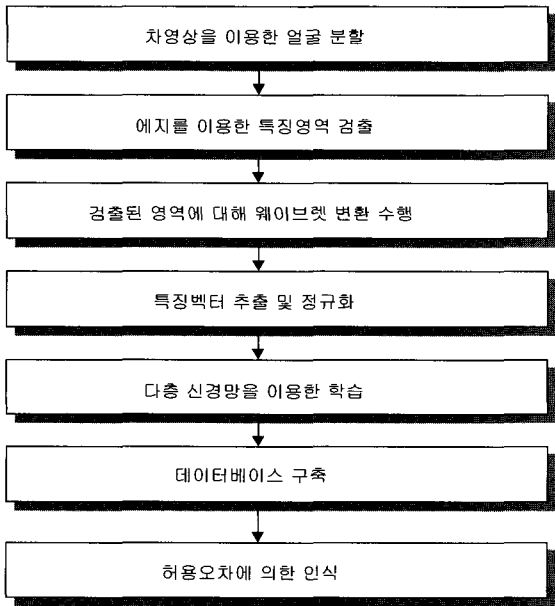


그림 4. 얼굴인식 알고리즘 순서도
Fig. 4. Flowchart for face recognition algorithm.

웨이브렛은 원형 웨이브렛인 $\psi(x)$ 를 신축/병진하여 구하며 식 2와 같다.

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (2)$$

여기서, a 는 스케일링 파라미터이고, b 는 병진 파라미터 그리고, \sqrt{a} 는 정규화 인자이다.

임의의 신호 $f(x)$ 의 웨이브렛 변환 $Wf(x)$ 는 식 3과 같이 원신호 $f(x)$ 와 웨이브렛 $\psi(x)$ 와의 상승 적분으로 정의된다.

$$Wf(x) = f * \psi(x) \quad (3)$$

또한, 2차원 함수 $f(x,y)$ 의 웨이브렛 변환에서 $\phi(x,y) = \phi(x)\phi(y)$ 를 2차원 스케일링 함수라 할 때, 1차원 스케일링 함수 $\phi(x)$ 를 갖는 1차원 웨이브렛을

$\phi(x)$ 라고 하면 2차원 웨이브렛들은 다음 식 4와 같이 분리 가능한 다해상도 근사화로 설명할 수 있다.

$$\begin{aligned} \phi^A(x,y) &= \phi(x)\phi(y) \\ \phi^H(x,y) &= \phi(x)\phi(y) \\ \phi^V(x,y) &= \phi(x)\phi(y) \\ \phi^D(x,y) &= \phi(x)\phi(y) \end{aligned} \quad (4)$$

식 (4)의 정규 직교 거리로 분해된 2차원 신호는 공간적으로 방향성을 지닌 주파수 성분으로 분해된다. [8] 웨이브렛 변환을 이용한 특징 파라미터의 추출 과정은 먼저 해상도가 $[256 \times 256 \times 2^8]$ 인 두 개의 입력 영상 신호를 차영상과 에지특성을 이용하여 크기가 91×91 인 화소를 갖는 특징영역으로 분리한다. 다음으로 DWT를 하여 DWT 계수 행렬을 구한다. 아래 그림 5는 4레벨 DWT에 대한 계수 행렬의 분포를 나타내는데 그림 5에서 cA4는 4레벨 저주파 계수 행렬을 의미하고, cH(i)는 (i)레벨의 수평 고주파 계수 행렬, cV(i)는 (i)레벨의 수직 고주파 계수 행렬 그리고, cD(i)는 (i)레벨의 대각 고주파 계수 행렬을 의미한다. 이와 같은 계수행렬을 이용하여 특징벡터를 추출한다.

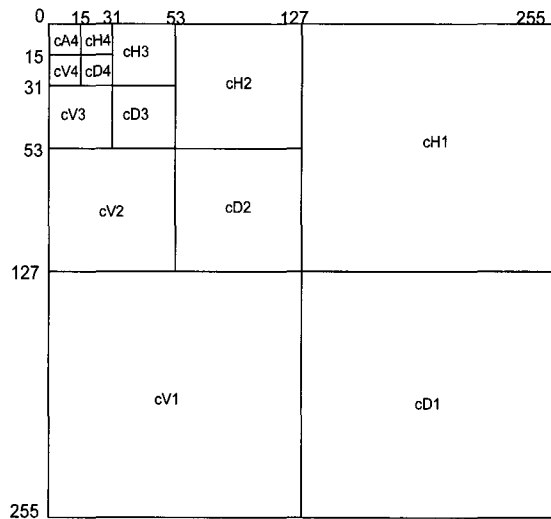


그림 5. 4-레벨 DWT 계수 행렬
Fig. 5. Four level DWT coefficient matrix.

분리된 얼굴의 특징영역에 대해 4레벨 DWT를 수행하면 6×6 의 크기를 가진 계수행렬 cA4, cH4, cV4 그리고 cD4를 구할 수 있다. 여기서 cA4는 4레벨 저주파

계수행렬, cH4는 수평 고주파 계수행렬, cV4는 수직 고주파 계수 행렬 그리고 cD4는 대각 고주파 계수행렬 인데 이들의 분포 특성을 분석한 후 특징벡터 추출을 한다.

먼저 동일인의 학습영상에 대한 계수행렬의 분포를 알아보기 위해 4가지 샘플의 학습 영상을 취해 각각 4 레벨 DWT를 수행한다. 다음으로 구해진 6×6의 크기를 갖는 4개의 계수 행렬에 대해 계수들의 절대값을 구한 다음 36개의 특징벡터를 추출하여 +1과 -1사이로 정규화를 한다. 마지막으로 4가지 샘플의 학습영상에서 각각 구해진 cA4, cH4, cV4 그리고 cD4에서 추출된 정규화 벡터의 평균자승 오차를 구한 후 오차의 크기에 근거하여 이를 신경망의 학습 벡터로 이용한다.

IV. 시뮬레이션

1. 얼굴 분할 및 특징 검출

본 연구에서는 일정한 조도상태에서 CCD 카메라로부터 동일거리에 있는 입력 영상을 256×256 크기의 그레이 스케일 256 레벨로 취득하여 PC(266MHz)상에서 시뮬레이션을 수행하였다.

이를 위해 그림 6, 7과 같은 입력 영상에 대해 가우시안 필터를 이용하여 잡음을 제거한 후 차영상을 구하여 배경영상에서 얼굴을 분할하였다. 30인을 대상으로 배경 및 얼굴이 포함된 영상을 각각 취득하였다. 이 때 얼굴이 포함된 영상에서 얼굴을 쉽게 분할하기 위하여 배경영상을 먼저 취득하고 다음으로 얼굴이 포함된 영상을 취득한다.



그림 6. 입력 영상(I)
Fig. 6. Input image(I).



그림 7. 입력영상(II)
Fig. 7. Input image(II).

다음으로 취득된 두 입력영상에 대해 차영상을 구했다. 그림 8은 구해진 차영상을 나타내고 그림 9는 이진 영상을 나타낸다.

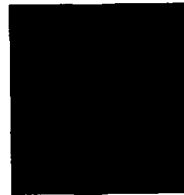


그림 8. 차영상
Fig. 8. Differential image



그림 9. 이진 영상
Fig. 9. Binary image

차영상을 바로 마스크로 이용할 수는 없으므로 그림 9와 같이 차영상을 이진화 한 후 경계면을 축소하였다. 그러나 경계면의 축소로 인해 얼굴 영역이 축소되는 결과를 초래하므로 축소된 영상을 그림 10과 같이 팽창시켰다. 그림 11은 팽창된 영상에서 픽셀값을 조사하여 생성한 마스크 영상이다.



그림 10. 팽창된 이진 영상
Fig. 10. Dilated binary image.

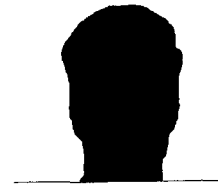


그림 11. 마스크 영상
Fig. 11. Mask image.

다음으로 마스크 영상을 얼굴이 존재하는 원영상에 투영하여 그림 12와 같이 배경에서 얼굴을 분할하였다.



그림 12. 분할영상
Fig. 12. Segmentation images



그림 13. 에지 영상
Fig. 13. Edge image.

또한 분할된 얼굴영상에서 sobel 연산자를 이용하여 에지를 검출한 후 에지성분의 가로방향 분포에 의해 눈과 눈썹의 위치를 검출한 후 눈썹 하단부를 분리하였다. 다음으로 에지의 수직방향 분포를 구해 얼굴의 중심선을 구했다.

마지막으로 인간얼굴에 대한 지식에 근거하여 그림

14와 같이 특징영역을 검출하였다.

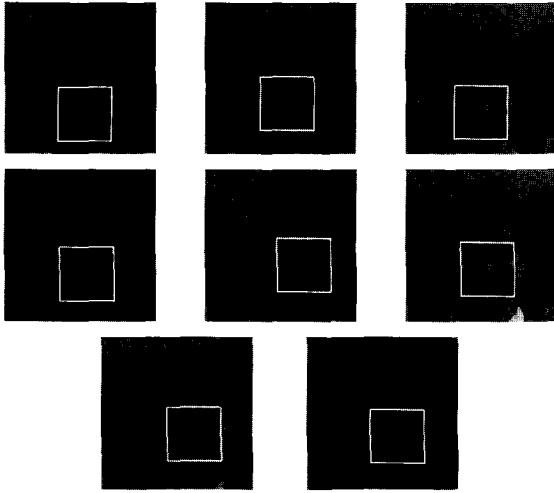


그림 14. 얼굴의 특징 영역
Fig. 14. Facial characteristic area.

2. 얼굴 인식 실험

그림 15(a)~15(d)는 동일인물에 대한 4개의 실험영상에서 분리한 특징영역이며 크기가 각각 91×91이다. 이들 특징영역에 대해 각각 4레벨 웨이브렛 변환을 한 후 동일 계수행렬을 정규화하여 데이터 분포를 분석한 후 신경망 입력벡터를 추출하였다.

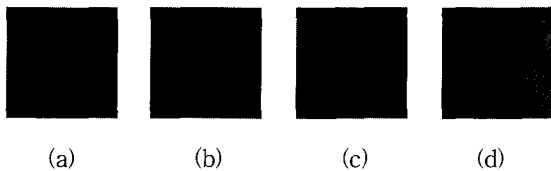


그림 15. 분리된 특징영상
Fig. 15. Example of Separated characteristic image.

그림 16은 그림15(a)에 대한 IDWT 영상이고 그림 17은 4레벨 웨이브렛 변환된 계수 행렬 cA4, cH4, cV4 그리고 cD4를 영상으로 나타낸 그림으로 이들 4개의 계수 행렬에 원영상의 정보가 집중되어 있다. 여기서 이들 계수 행렬의 크기는 각각 6×6이다.



그림 16. IDWT 영상
Fig. 16. IDWT image.

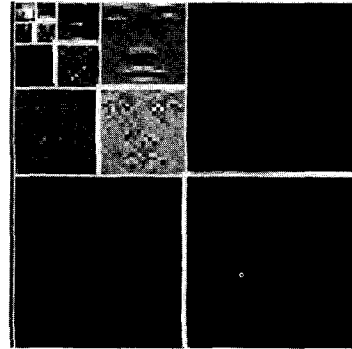


그림 17. 4 단계 분해 영상
Fig. 17. Four step decomposition image.

특징 추출은 그림 15(a)~15(d)와 같은 영상에 대해 4개의 학습 샘플을 선정한 후 이를 4레벨 Wavelet 변환하여 cA4, cH4, cV4 그리고 cD4를 각각 구했다. 다음으로 동일 계수행렬의 정규화 데이터 분포를 분석한 후 특징벡터를 추출하였고 이를 +1에서 -1 사이로 정규화 하여 신경망의 입력벡터로 하였다. 그림 18, 19, 20 그리고 21은 4개의 샘플영상에서 추출한 특징벡터 분포를 나타내고 있다.

여기서 가로축은 특징벡터의 수로 36개이고 세로 축은 특징벡터의 값을 나타내고 있다.

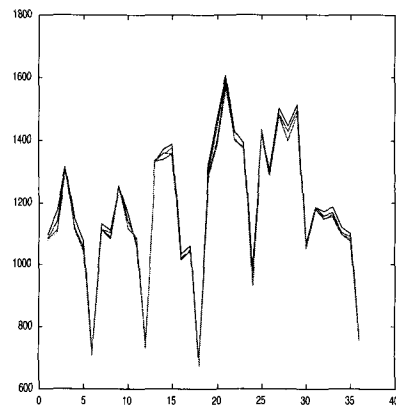


그림 18. 특징벡터의 분포 곡선(cA4)
Fig. 18. Distribution curve of characteristic vector(cA4).

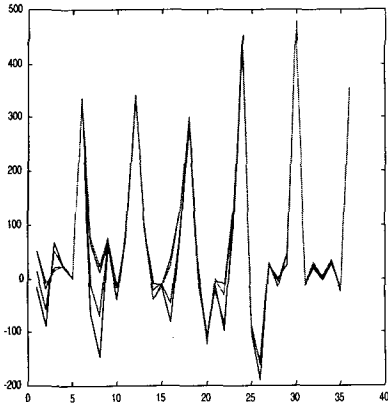


그림 19. 특징벡터의 분포 곡선(cH4)
 Fig. 19. Distribution curve of characteristic vector(cH4).

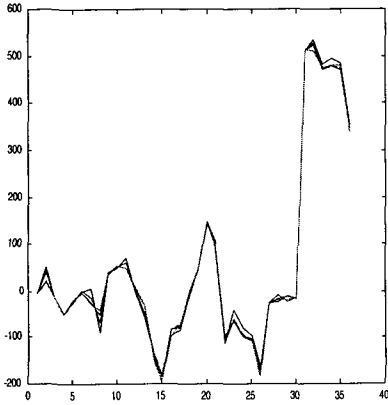


그림 20. 특징벡터의 분포 곡선(cV4)
 Fig. 20. Distribution curve of characteristic vector(cV4).

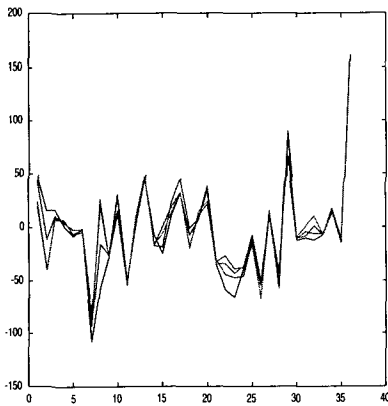


그림 21. 특징벡터의 분포 곡선(cD4)
 Fig. 21. Distribution curve of characteristic vector(cD4).

그리고 4개의 샘플영상에서 각각 구해진 6×6 계수

행렬에 대해 절대값을 취한 후 정규화를 수행하였다. 그림 22, 23, 24 그리고 25는 정규화된 특징 벡터의 분포 상태를 나타내고 있다.

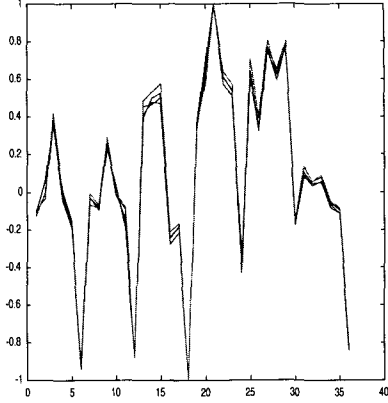


그림 22. 정규화된 특징벡터(cA4)
 Fig. 22. Normalized characteristic vectors(cA4).

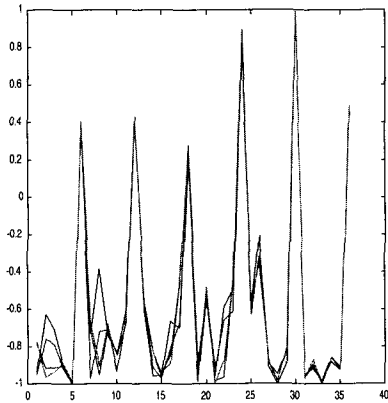


그림 23. 정규화된 특징벡터(cH4)
 Fig. 23. Normalized characteristic vectors (cH4).

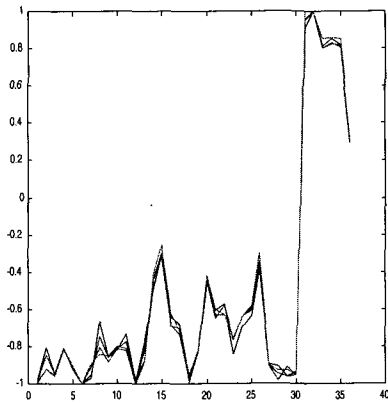


그림 24. 정규화된 특징벡터(cV4)
 Fig. 24. Normalized characteristic vectors(cV4).

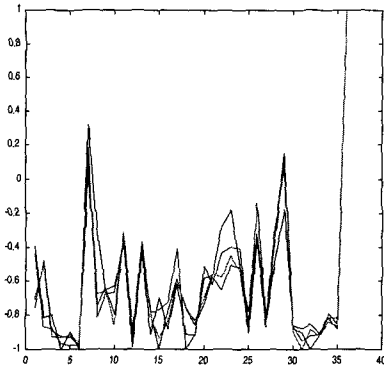


그림 25. 정규화된 특징벡터(cD4)
Fig. 25. Normalized characteristic vectors(cD4).

여기서 가로방향은 정규화 특징벡터의 수로 36개이고 세로방향은 정규화의 범위로 +1에서 -1의 값을 갖는다.

그림 15(a) ~15(d)와 같은 샘플 영상 4개에서 각각 구한 계수행렬에서 신경망의 입력벡터를 선정하기 위해 동일 계수 행렬에 대한 정규화 벡터간의 오차를 구했다. 먼저 cA4에서 구해진 정규화 특징벡터의 RMSE를 구했다. 다음으로 cH4, cV4 그리고 cD4에 대해서도 동일한 과정을 통해 RMSE를 구했다. 다음 표 1은 이들 각각의 계수행렬의 RMSE를 나타낸다.

표 1. 정규화 벡터간의 RMSE
Table 1. RMSE of normalized feature vectors.

계수종류	cA4	cH4	cV4	cD4
RMSE	0.003575	0.004867	0.006381	0.006580

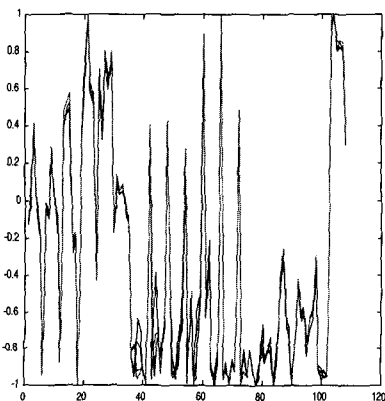


그림 26. 추출된 입력벡터
Fig. 26. Extracted input vectors.

특징벡터 추출의 당위성을 확인하기 위해 3인의 인물에 대해 각각 4개의 영상을 취득하여 웨이블릿 변환을 한 후 계수행렬을 정규화한 결과를 보인다.

이를 위해 그림 27(a), 27(b) 그리고 27(c)과 같은 실험영상에 대해 각각 4개의 샘플 영상을 취득하고 이들에 대해 4 레벨 웨이블릿 변환을 수행한 후 구해진 6×6 계수행렬을 정규화 하여 36개의 특징벡터를 추출하였다. 그리고 동일 계수간의 RMSE를 구했다. 표 2, 표 3 그리고 표 4는 정규화 벡터간의 RMSE이다 마찬가지로 cD4를 정규화한 벡터에서 가장 큰 오차를 보였다.

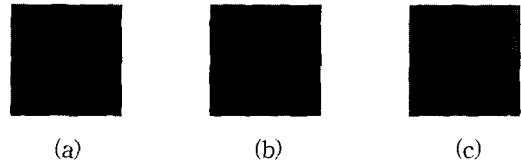


그림 27. 실험영상의 예
Fig. 27. Example of experimental images.

표 2. 정규화 벡터간의 RMSE(그림 27(a))
Table 2. RMSE of normalized feature vectors.

계수종류	cA4	cH4	cV4	cD4
RMSE	0.003076	0.005565	0.003042	0.006726

표 3. 정규화 벡터간의 RMSE(그림 27(b))
Table 3. RMSE of normalized feature vectors.

계수종류	cA4	cH4	cV4	cD4
RMSE	0.001788	0.003830	0.002196	0.006224

표 4. 정규화 벡터간의 RMSE(그림 27(c))
Table 4. RMSE of normalized feature vectors.

계수종류	cA4	cH4	cV4	cD4
RMSE	0.000848	0.002928	0.001082	0.003476

그림 28(a), 28(b), 그리고 28(c)는 각각 그림 27(a), 27(b) 그리고 27(c)에서 추출한 신경망 입력벡터를 나타낸다.

그림 28과 같이 추출된 108개의 정규화 입력 벡터를 다층신경망의 입력으로 하였다. 신경망의 학습알고리즘은 오류 역전파 학습 알고리즘을 이용하였다. 출력층의 오차는 0.005 그리고 학습률은 0.7로 하였다. 학습 후

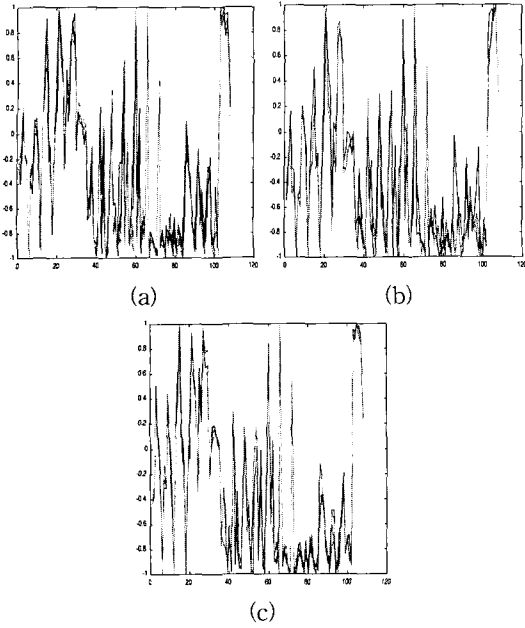


그림 28. 추출된 입력벡터
Fig. 28. Extracted input vectors.

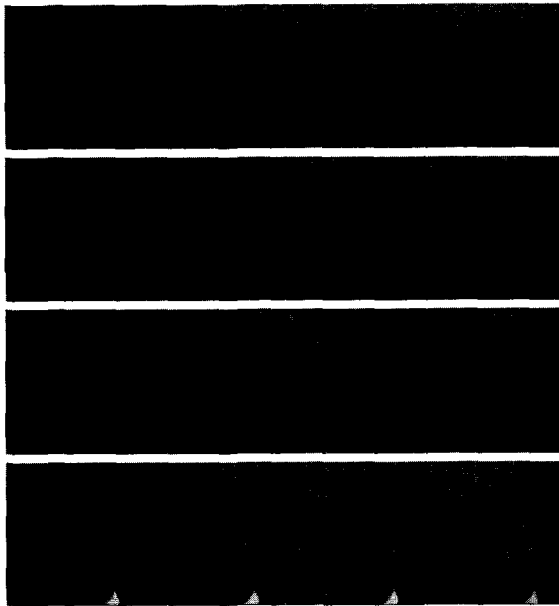


그림 29. 실험영상(I)
Fig. 29. Test image(I).

생성된 네트워크의 가중치와 입력된 비교영상의 특징 벡터를 연산하여 출력층의 오차를 비교하여 인식을 수행하였다. 이 때 출력층의 오차가 0.005보다 작으면 동일인물로 판별하였다.

그림 29의 테스트 영상(I)은 인식 영상의 유형을 나

타내고 있고 그림 30의 테스트 영상(II)은 오인식 영상의 예를 나타내고 있는데 오인식 영상을 보면 얼굴이 기울어져 있음을 알 수 있다.

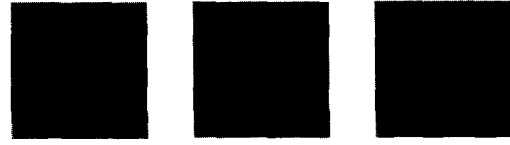


그림 30. 실험영상(II)
Fig. 30. Test image(II).

이와 같이 30명의 영상에 대해 1인당 학습 영상 4장, 비학습 영상 4장을 가지고 실험한 결과 학습 영상 120장에 대해서는 100%의 인식률을 보였고 비학습 영상 120장에 대해서는 92%의 인식률을 나타내었다.

3. 비교 및 검토

제안된 알고리즘들의 시뮬레이션 결과는 다음과 같다.

- 1) 차영상 기반 얼굴 분할 알고리즘을 이용하여 제안된 실험환경에서 얼굴이미지를 100% 분할 할 수 있었고 분할 알고리즘을 보완하면 영상 전송시 효율적인 코딩을 할 수 있다.
- 2) DCT를 이용하여 특징추출을 하는 방법보다 더 정교한 특징 벡터를 구할 수 있었고 인식률도 4% 증가했다. 그림 31을 보면 동일인물에 대해 DCT를 수행한 후 추출한 정규화 벡터의 분포 특성보다 Wavelet 변환을 한 후 추출한 특징벡터의 정규화 특성이 고르게 분포함을 알 수 있다.

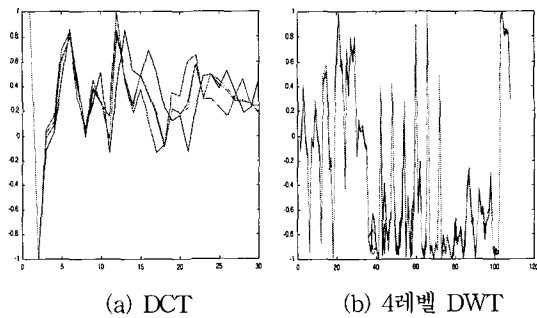


그림 31. 특징벡터의 분포
Fig. 31. Distribution of characteristic vectors.

- 3) 공간영역의 데이터를 주파수 영역으로 변환하여 특징을 추출함으로써 데이터량을 줄일 수 있었다. 또한 DWT 수행시 얼굴 전체를 처리하지 않고 특징

영역만 처리함으로써 연산을 위한 데이터량이 감소되었다.

- 4) 얼굴이 기운영상에 대해서는 오인식을 하는 경우가 발생하였으므로 기술기 보정이 필요하다. 그리고 완벽한 보완 시스템에 적용하려면 패스워드 입력이나 지문인식과 같은 알고리즘을 병행할 필요가 있다.

V. 결 론

본 연구에서는 복잡한 배경하에 사람이 존재할 때 얼굴을 분할하는 방법을 제안하였고 분할된 영상에서 특징을 추출하여 웨이블릿 변환을 적용하는 효율적인 인식시스템을 제안하였다.

먼저 일정한 조도 상태에서 취득된 두 입력영상에 대해 차영상을 구하여 배경영상에서 얼굴을 분할하였다. 이를 위해 CCD 카메라로부터 동일거리에 있는 영상을 256×256 크기의 그레이 스케일 256 레벨로 취득하여 가우시안 필터를 이용하여 영상 내에 존재하는 잡음을 제거하였다. 그리고 배경 영상과 얼굴이 포함된 입력 영상의 차를 구하고 차영상을 팽창하고 축소하여 얼굴 내의 잡음을 제거한 후 픽셀 값을 조사하여 마스크를 생성하였다. 다음으로 마스크를 입력영상에 투영하여 얼굴을 분할한 후 얼굴영상에서 특징영역을 검출하고 웨이블릿 변환을 적용하여 구해진 계수행렬로부터 특징을 추출하였다. 마지막으로 특징 벡터를 +1과 -1사이로 정규화한 후 다층 신경망에 입력하여 학습하였다.

시뮬레이션 결과 입력영상에 대해 얼굴이 전부 분할됨을 확인하였고 DCT 알고리즘을 이용한 경우의 인식률은 120장의 학습영상에 대해 100%, 120장의 실험영상에 대해 88%의 인식률을 보였으나 Wavelet 변환을 이용하여 인식을 수행했을 경우 120장의 학습영상에 대해 100%, 120장의 실험영상에 대해 92%의 높은 인식률을 입증하였다.

따라서 제안된 분할 알고리즘은 뉴스와 스포츠 중계를 하는 아나운서의 얼굴을 분할하여 부호화하는데 효

과적으로 응용될 수 있고 인식률 개선과 알고리즘 보완을 통해 멀티미디어 환경에서 얼굴인식을 효과적으로 할 수 있다.

참 고 문 헌

- [1] T. Sakai, M. Nagao, and Fukibayashi : Line extraction and pattern recognition in a photograph : *Pattern Recognition*, Vol. 1, pp. 233-248, 1969.
- [2] M.D. Kelly : Visual identification of people by computer : *Tech. Rep. AI-130*, Stanford AI Proj., Stanford, CA, 1970.
- [3] I. Craw, H. Ellis, and J. Lishman : Automatic extraction of face feature : *Pattern Recognition Lett.*, Vol. 5, pp.183-187, 1987.
- [4] 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.
- [5] S. A. Sirohey : Human face segmentation and identification : *Tech. R. CAR-TR-695* : Center for Automation Research, Univ. of Maryland, 1993.
- [6] 이명길, 이주신, "DCT와 신경회로망을 이용한 패턴인식에 관한 연구", *한국통신학회논문지*, Vol. 22 No. 3, pp.481-492, 1997
- [7] 이명길, 이주신, "웨이블릿 변환과 신경회로망을 이용한 SMD IC 패턴인식", *전자공학회 논문지*, 제34권 S편 제 7호, pp.768-777. 1997년 7월
- [8] M.D. Kelly : Visual identification of people by computer : *Tech. Rep. AI-130*, Stanford AI Proj., Stanford, CA, 1970.
- [9] I. Craw, H. Ellis, and J. Lishman : Automatic extraction of face feature : *Pattern Recognition Lett.*, Vol. 5, pp.183-187, 1987.

저 자 소 개

林 春 煥(正會員) 第 32卷 T編 3號 參照
현재 : 동강대학 전자정보과 겸임교수

片 爽 範(正會員) 第 32卷 T編 3號 參照
현재 : 동강대학 전자정보과 교수

李 尚 薰(正會員) 第 36卷 T編 1號 參照
현재 : 동강대학 전자정보과 교수