

論文2000-37TE-3-11

ATM 보안 시스템을 위한 모델 인증 알고리즘

(Model Verification Algorithm for ATM Security System)

鄭 憲*, 林春煥**, 片 爽 範**

(Heon-Jeong, Chun-Hwan Lim, and Suk-Bum Pyeon)

요 약

본 연구에서는 ATM 보안 시스템을 위한 DCT와 신경망 기반 모델 인증 알고리즘을 제안한다. CCD 카메라를 이용하여 일정한 조도와 거리에서 30명의 얼굴영상을 획득한 후 데이터 베이스를 구성한다. 모델 인증 실험을 위해 동일인에 대해 학습영상 4장 그리고 실험 영상 4장을 각각 획득한다. 얼굴영상의 에지를 검출한 후 에지 분포에 의해 얼굴영상에서 사각형태로 특징영역을 검출한다. 특징영역에는 눈썹, 눈, 코, 입, 그리고 뺨이 포함된다. 특징영역에 대해 DCT를 수행한 후 대각방향의 계수 합을 구해 특징벡터를 추출한다. 특징벡터는 정규화되어 신경망의 입력 벡터가 된다. 패스워드를 고려하지 않는 경우, 데이터 베이스를 검색한 결과 학습된 얼굴영상에 대해서는 100%의 인증율을 나타내었고 학습되지 않는 얼굴영상의 경우에는 92%의 인증률을 나타내었다. 그러나 패스워드를 고려한 경우 모두 100%의 인증률을 보였다.

Abstract

In this study, we propose a model verification algorithm based on DCT and neural network for ATM security system. We construct database about facial images after capturing thirty persons facial images in the same lumination and distance. To simulate model verification, we capture four learning images and test images per a man. After detecting edge in facial images, we detect a characteristic area of square shape using edge distribution in facial images. Characteristic area contains eye bows, eyes, nose, mouth and cheek. We extract characteristic vectors to calculate diagonally coefficients sum after obtaining DCT coefficients about characteristic area. Characteristic vectors is normalized between +1 and -1, and then used for input vectors of neural networks. Not considering passwords, simulations results showed 100% verification rate when facial images were learned and 92% verification rate when facial images weren't learned. But considering passwords, the proposed algorithm showed 100% verification rate in case of two simulations.

I. 서 론

21세기 멀티미디어 정보화 사회에서는 개인에 대한

* 正會員, 草堂大學校 電子工學科
(Dept. of Electronic Engineering, Chodang University)

** 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3月6日, 수정완료일:2000年7月2日

정보관리와 보호가 가장 중요한 문제가 될 것이다. 컴퓨터를 이용한 개인의 신분확인을 위해 인체특징을 이용하는 바이오메트릭 기법이 많이 이용되고 있다. 그 중 얼굴인식, 지문인식, 안구인식, 정맥 패턴인식과 같이 컴퓨터 비전 기술을 이용하는 방법과 화자인식과 같이 음성정보를 이용하는 방법이 있다.

이들 중 얼굴인식 방법은 사용자의 특별한 동작이나 행위에 대한 요구가 없이 비접촉식으로 자연스럽게 사용자 확인을 할 수 있는 장점 때문에 가장 편리하고 경쟁력 있는 바이오메트릭 기술로 인정되고 있다.

그러므로 향후 21세기 정보화 및 지능화 사회의 국가적 실현을 위해서는 핵심기술인 얼굴인식 기술이 반드시 확보되어야 한다.

얼굴인식을 위한 연구는 여러 가지 방법으로 연구되고 있다. L.D. Harmon은 옆얼굴 사진에서 9개의 기점(fiducial marks)을 이용하여 기점간 거리, 기점간 기울기, 턱이나 이마 곡면의 반경, 코의 돌출정도, 얼굴의 면적등 모두 11개의 특징값을 이용하여 얼굴 인식을 시도하였고^[1], K.H. Wong은 정면 영상을 이용하여 얼굴 윤곽선(face outline) 및 특징점간 거리 등을 이용하여 거리값의 변화정도에 따라 가중치를 적용하는 방법으로 인식을 시도하였다.^[2] Sirovich와 Kirby는 수염이 없는 얼굴 영상에 대해 KL 변환을 적용하여 얼굴을 표현하는 방법을 소개하였다^[3]. 이 방법은 모든 영상이 KL 변환에 의한 고유벡터 값의 가중치 조합을 사용하여 재구성될 수 있으며, 영상의 재구성에 사용하는 고유벡터의 수를 증가시키면 입력 영상에 보다 가까운 영상을 구성할 수가 있다는 점에 바탕을 두고 있다.

일본의 NTT 연구소의 Takaaki Akimoto는 정면 얼굴과 측면얼굴을 두 대의 카메라를 이용하여 동시에 찍어서 측면영상의 굴곡 특성을 이용하여 정면영상에서 특징점을 찾는 방법을 사용했다.^[4] C. vonder Malsburg는 Gabor 함수를 이용하여 그래프 매칭방법에 의해 얼굴영상을 인식하였다.^[5]

Nixon은 얼굴 특징 추출을 위해 Hough변환을 사용하였다^[6]. Hough변환은 기울기 연산자에 의해 제공되는 방향 정보와 기울기의 크기 정보를 사용하여 원하는 형태를 찾아낸다. 이 방법에서 눈동자는 원으로 표현되었으며, 타원이 공막의 둘레를 표현하기 위해 사용되었다. 그러나 타원은 얼굴의 중심에서 멀리 떨어져 있는 눈꼬리 부분을 표현하는데 부적절하기 때문에, 각 얼굴의 중심에 맞도록 지수함수를 사용하여 수정되었다.

특징값을 추출한 후 인식 작업을 수행할 때는 일반적으로 두가지 방법을 사용한다. 첫째, 신경망을 이용하는 방법은 다양한 입력 환경에 적응하여 얼굴영상을 인식할 수 있는 장점이 있으나 입력 수가 너무 많으면 신경망의 입력 노드 수가 증가하게 되고 아울러 연결개수도 증가하기 때문에 시스템 구현에 어려움이 있다. 그러나 얼굴 영상에 대해 주파수 영역으로 변환한 후 소수의 특징 파라미터를 추출하면 신경망의 입력 노드 수와 연결선의 수를 줄일 수 있어 학습 시간 단축 및 시스템 구현이 용이하다.

둘째, 유클라디안 거리를 이용한 방법은 데이터가 많은 경우 인식률이 떨어지는 단점이 있는 반면 시스템 구현이 용이하다는 장점이 있다. 따라서 인식률을 높이면서 신경망의 입력노드를 줄일 수 있는 특징추출 방법이 필요하다.

본 논문에서는 일정한 조도 하에서 CCD 카메라로부터 취득된 얼굴영상을 에지 정보를 이용하여 얼굴영상에서 사각의 특징 영역을 분리하고 이를 DCT하여 특징파라미터를 추출하는 방법을 제안한다. 특징영역에는 눈썹, 눈, 코, 입이 포함된다. 다음으로 추출된 특징 파라미터를 정규화하여 다층신경망의 입력 벡터로 사용한다. 각 인물에 대해 학습된 네트워크를 이용하여 임의의 입력 얼굴영상에 대하여 출력층의 오차를 비교함으로써 인증하도록 설계하였다. 이를 위하여 제 II장에서는 모델 인증 알고리즘을 제안하고 III장에서는 시물레이션을 통해 결과를 분석하고 제 IV장에서 결론을 맺는다.

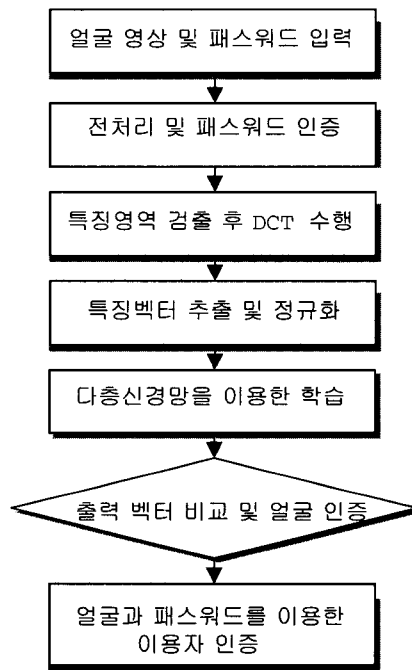


그림 1. 인증 알고리즘
Fig. 1. Verification Algorithm.

II. 얼굴과 패스워드가 결합된 모델 인증 알고리즘

제안된 모델 인증 알고리즘은 인증의 정확성을 위해

패스워드와 얼굴의 두 정보를 이용한다. 먼저 얼굴영상의 에지 분포를 이용하여 특징영역을 검출하고 특징영역에 대해 DCT를 수행한 후 특징벡터를 추출한다. 다음으로 추출된 특징벡터를 정규화하여 신경망의 입력 벡터로 한다. 학습된 얼굴영상의 특징은 신경망의 가중치 값으로 저장되고 이를 이용하여 인식을 한다. 그림 1은 인증을 위한 알고리즘을 나타낸다

1. 모델인증을 위한 특징 검출

모델인증을 위한 데이터 베이스를 구성하기 위해 일정한 조도와 거리에서 고정된 CCD 카메라로 단조로운 배경을 갖는 얼굴영상을 256×256 크기의 그레이 스케일 256 레벨로 취득한 후 가우시안 필터를 이용하여 영상 내에 존재하는 잡음을 제거하였다.



그림 2. 입력 영상
Fig. 2. Input image.



그림 3. 에지 영상
Fig. 3. Edge image.

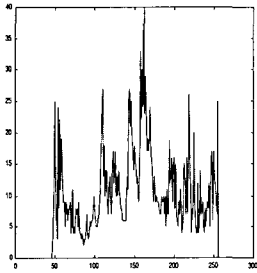


그림 4. 에지 합
Fig. 4. Edge sum.



그림 5. 분리된 얼굴 영상
Fig. 5. Separated face image.

특징영역 검출을 위해 분리된 얼굴 영상에 대해 sobel 연산자를 이용하여 영상의 에지를 검출한다. 그림 4는 검출된 에지 성분들의 수평 방향 합을 나타내고 있는데 그림에서 최대 값은 40이다. 최대값 근방이 눈의 영역이다. 따라서 가장 많은 에지를 갖는 위치에서 위쪽방향으로 픽셀값을 조사하여 눈썹의 상단부위를 분리한다. 또한 카메라로부터 일정한 거리에 있는

사람의 얼굴의 크기는 비슷하므로 얼굴 특징간 픽셀의 거리를 지정하여 눈썹, 눈, 코 그리고 입을 모두 포함하도록 얼굴의 중심을 기준으로 좌우 동일한 크기로 분리한다.

그림 5는 크기가 91×91 의 그레이 스케일을 갖는 검출된 얼굴의 특징영역을 나타내고 있으며 이 영역에 대해 DCT를 수행한 후 특징을 검출한다.

2. DCT를 이용한 특징 파라미터 추출

이산 코사인 변환은 많은 영상 압축 알고리즘의 기초가 된다. 이산 푸리에 변환에 비해 DCT의 장점은 복소수 계산을 하지 않아도 된다는 것이다[7]. 또한 DCT는 화상 데이터의 분포에 잘 맞기 때문에 자주 이용되고 있다. DCT와 역변환인 IDCT의 식은

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

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

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

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

로 표현된다. DCT를 화상에 이용하는 경우에는 m 화소 $\times n$ 화소를 1구획으로서 2차원 DCT를 행하고 있으며 8화소 $\times 8$ 화소 또는 16화소 $\times 16$ 화소의 구획이 자주 사용되고 있다. $M \times N$ 화소의 2차원 DCT와 IDCT는

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

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

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

$$1, \gamma > 0 \text{ 일 때} \quad (2)$$

DCT를 이용한 특징 파라미터의 추출 과정은 먼저 해상도가 $[256 \times 256 \times 2^8]$ 인 입력 영상 신호를 에지 분포를 이용하여 크기가 91×91 인 화소를 갖는 특징영역으로 분리한다. 특징영역에 대해 2차원 DCT를 수행한 후 DCT 계수들을 추출한다. DCT계수 중 영상에 대한 대부분의 정보를 포함하고 있는 저주파수 영역에서 특징을 추출한다. 그림 6은 특징 영역에 대한 DCT

영상을 나타내고 있는데 91×91 픽셀의 DCT 영상에서 30×30 픽셀 영역의 계수 행렬값 중에서 좌상삼각행렬만을 추출해 내어 같은 주파수대(대각방향)의 DCT계수합을 구하여 학습영상에 대해 30개의 특징 벡터를 추출한다.

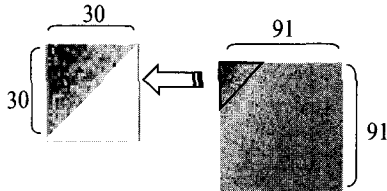


그림 6. 계수 행렬에서 특징추출
Fig. 6. Characteristic extraction from coefficient matrix.

그림 7은 DCT를 이용한 정규화 입력 벡터의 추출과정을 나타낸다. 추출된 특징 벡터들을 신경망에 입력시키기 위해서는 특징벡터 중 최소값을 -1, 최대값을 1로 정규화 시켜야 한다.

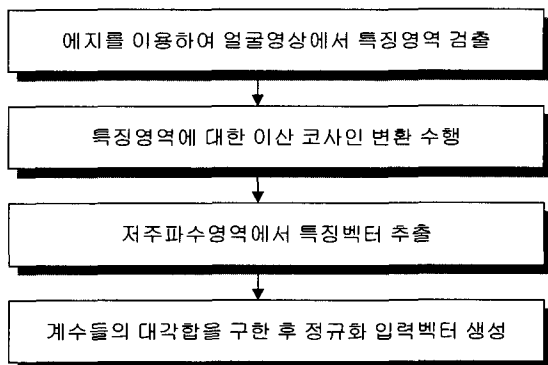


그림 7. 정규화 입력 벡터 추출
Fig. 7. Normalized extraction of input vectors.

이와같이 정규화된 신경망 입력 벡터를 각각의 인물에 대하여 동일한 구성을 가지는 다층 신경망에 입력하여 학습시키고, 각 학습된 신경망의 가중치 벡터와 출력벡터를 이용하여 입력된 패스워드를 고려하여 얼굴 인증을 수행한다.

III. 시뮬레이션

일정한 조도에서 고정된 CCD 카메라와 F360 DSP board를 이용하여 동일 거리에 있는 얼굴 영상에 대해

표정과 자세를 약간 변화시켜 학습영상 4장 그리고 테스트 영상 4장을 각각 취득하였다. 30인에 대해 연속적으로 영상을 취득하여 총 240장으로 데이터베이스를 구성한 후 Pentium PC상에서 시뮬레이션을 수행하였다.

특징 영역 검출을 위해 먼저 Sobel 연산자를 이용하여 에지를 검출하였다. 검출된 에지 성분들의 가로 방향 에지합을 구해보면 가장 많은 에지를 포함하는 부분이 눈 영역 근처가 되므로 여기에서 위쪽 방향으로 픽셀값을 조사하여 눈썹의 상한 위치를 찾아내었다.

그리고 카메라로부터 일정한 거리에 있는 사람의 얼굴의 크기는 일정하고 또한 에지가 눈썹과 입주위에서 강하게 나타나므로 이를 근거로 눈썹, 눈, 코, 입을 모두 포함하도록 얼굴의 중심선을 기준으로 좌우로 눈썹의 상한 위치로부터 동일한 크기로 잘라내었다

에지분포에 근거하여 입력영상은 아래 그림 8과 같이 특징 영역이 91×91인 사각 형태로 분리된다. 이렇게 하여 검출된 특징 영역에 대한 IDCT 영상은 아래 그림 9와 같다.



그림 8. 분리된 얼굴 영역
Fig. 8. Seferated face region.



그림 9. IDCT 영상
Fig. 9. IDCT image.

그림 6의 DCT 영상을 보면 좌상삼각행렬에 영상의 특징이 집중되므로 본 연구에서는 좌상삼각행렬만을 인증에 이용하였다. 그림 8과 같은 동일 인물에 대해 그림 12와 같이 4개의 학습 샘플영상을 사용하여 DCT를 수행한 후 특징벡터를 각각 추출하였다.

그림 10은 학습영상에 대해 추출된 특징벡터의 분포를 나타내고 있다. 그림에서 최대값은 직류성분이며 동일인에 대한 학습 영상의 특징벡터들이 거의 일정하게 분포함을 알 수 있다.

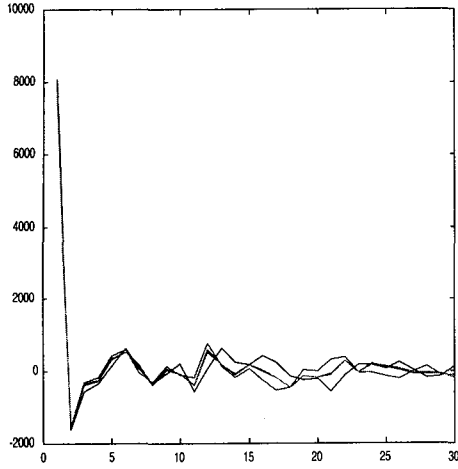


그림 10. 특징 벡터의 분포 특성
Fig. 10. Distribution curve characteristic vectors.

다음으로 그림 10의 특징벡터를 정규화하면 표 1과 같고 이때 정규화된 특징벡터의 RMSE는 0.0044이었다.

다음 그림 11은 표 1의 정규화 입력 벡터의 데이터 분포 곡선을 나타내고 있는데 특징벡터의 정규화로 인해 정규화 벡터간에 약간의 오차가 발생했으나 각각의 입력 영상에 대한 정규화 벡터는 일정하게 분포하였다.

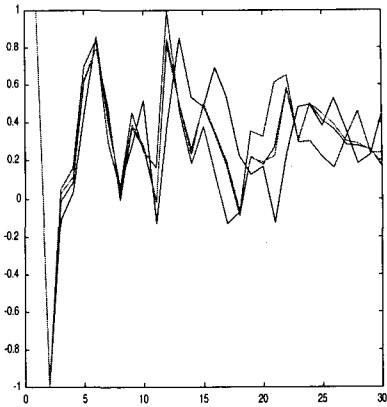


그림 11. 정규화된 입력 벡터
Fig. 11. Normalized input vectors.

이와 같이 추출된 정규화 입력벡터를 다층 신경망에 입력하여 학습을 수행하게 되는데 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.

30명의 영상에 대해 1인당 학습 영상 4개, 비학습 영

표 1. 정규화 입력 벡터
Table 1. Normalized input vectors.

특징벡터의 정규화 입력벡터	
그림(a)	+1.000000000 -1.000000000 -0.113297186 +0.033987866 +0.447664036
	+0.857262555 +0.406803297 +0.059167089 +0.283387330 +0.521533844
	-0.132028413 +0.403325810 +0.851591587 +0.538250210 +0.482208369
	+0.696179558 +0.538710888 +0.223221744 +0.128483673 +0.167121938
	-0.127178304 +0.244409439 +0.481599586 +0.497596791 +0.384386599
	+0.530736184 +0.363461415 +0.186711835 +0.231990147 +0.449302506
그림(b)	+1.000000000 -1.000000000 +0.047866113 +0.163013534 +0.698236295
	+0.839383835 +0.296545799 +0.047884696 +0.456574194 +0.251847497
	+0.163679222 +0.992887020 +0.454864442 +0.184838687 +0.377120645
	+0.122709037 -0.134220781 -0.073634109 +0.354292910 +0.323626446
	+0.613181385 +0.653181133 +0.294032925 +0.301843467 +0.219778608
	+0.161577522 +0.323626446 +0.461760569 +0.240820607 +0.243083906
그림(c)	+1.000000000 -1.000000000 -0.015128106 +0.083634478 +0.618736166
	+0.795024633 +0.482001261 -0.011365998 +0.371436593 +0.268976757
	-0.018554520 +0.846562716 +0.491554209 +0.261649431 +0.489609102
	+0.335307648 +0.169059658 -0.088189898 +0.217898736 +0.176302945
	+0.270598473 +0.583802674 +0.306396826 +0.502474714 +0.415725852
	+0.368570745 +0.282879116 +0.280527847 +0.260953572 +0.165103202
그림(d)	+1.000000000 -1.000000000 +0.024029092 +0.115860493 +0.633284751
	+0.796983993 +0.456791603 +0.029157007 +0.387815054 +0.279329418
	-0.015994021 +0.812114724 +0.490391139 +0.228942547 +0.503034984
	+0.350806159 +0.191888203 -0.062900856 +0.212763496 +0.192475215
	+0.228271250 +0.578749465 +0.303151940 +0.502379877 +0.448947076
	+0.389817119 +0.303689765 +0.292664983 +0.257297858 +0.181826853

상 4개를 가지고 실험하였다. 패스워드 없이 실험한 경우 전체학습 영상에 대해서는 100%의 얼굴 인증률을 보였고 비학습 영상에 대해서는 92%의 얼굴 인증률을 나타내었다. 그러나 패스워드를 이용하여 실험한 결과 두가지 경우 다 100%의 인증률을 보였다. 그림 12는 학습 영상의 예를 나타내고 있다.



그림 12. 학습 영상의 예
Fig. 12. Example of learning images.

패스워드를 이용하지 않는 경우에 오인식 이유를 분석하면 표정이 많이 변화되거나 상하좌우 기울어진 얼굴 영상에 대해 오인식이 일어났다.

IV. 결 론

본 연구에서는 ATM 보안 시스템을 위한 DCT와 신경망 기반 모델 인증 알고리즘을 제안하였다. 모델 인증 실험을 위해 30명의 인물에 대해 학습영상 4장 그리고 실험 영상 4장을 각각 획득하여 총 240장의 영상과 30개의 패스워드로 데이터 베이스를 구성하였다. 취득한 인물의 얼굴영상의 에지를 검출한 후 에지 분포에 의해 사각형태로 특징영역을 검출하였다. 그리고 특징영역에 대해 DCT를 수행한 후 대각방향의 계수 합을 구해 특징벡터를 추출하였다. 추출된 특징벡터는 정규화된 후 30개의 입력노드를 갖는 신경망에 입력되어 학습을 수행하였다.

30명의 영상에 대해 1인당 학습 영상 4개, 비학습 영상 4개를 가지고 실험하였다. 패스워드 없이 실험한 경우 전체학습 영상에 대해서는 100%의 얼굴 인증률을 보였고 비학습 영상에 대해서는 92%의 얼굴 인증률을 나타내었다. 그러나 패스워드를 이용하여 실험한 결과 두가지 경우 다 100%의 인증률을 보였다.

아울러 DCT를 이용하여 특징추출을 함으로써 신경망의 입력벡터의 수를 줄일 수 있었고 특징추출을 쉽게 할 수 있었다.

앞으로의 연구과제는 보다 많은 사람의 얼굴영상에 대해 데이터 베이스를 구성한 후 인식 성능을 비교하고 실제 ATM 보안 시스템에 적용하는 문제이다.

참 고 문 헌

- [1] L. D. Harmon, "Identification of human face profiles by computer," Pattern Recognition, Vol. 10, pp.301-312, 1978.
- [2] K. Wong, H. Law and P. Tsang, "A system for recognition human faces," Proc. ICASSP, pp.1638-1642, May, 1989
- [3] L. Sirovich and M. Kirby : Low-dimensional procedure for the characterization of human face : J. Opt. Society Amer., Vol. 4, pp.519~524, 1987.
- [4] T. Akimoto, R. Wallace and Y. Suenaga, "Feature extraction from front and side views of faces for 3d facial model creation," MVA '90, NTT Human Interface Lab., pp.291-294, 1990.
- [5] J. Buhmann, J. Lange and C. vonder Malsburg, "Distortion invariant object recognition by matching hierarchically labeled graphs," IJCNN, Vol. 1, pp.155-159, June, 1989.
- [6] M. Nixon : Eye spacing measurement for facial recognition : SPIE proc., Vol. 575, pp.279~285, 1985.
- [7] 임춘환 외 2명 "차영상과 DCT 기반 특징추출과 다층신경망을 이용한 얼굴인식에 관한 연구", 한국통신학회논문지, Vol. 24 No. 8B, pp.1549-1577, 1999

저 자 소 개



林 春 煥(正會員)

1993年 朝鮮大學校 電子工學科 卒業
(工學碩士). 2000年 朝鮮大學校 大學
院 卒業(工學博士). 2000年 現在 東
岡大學 電子情報科 兼任教授. 主 관심
분야는 멀티미디어, 영상통신, 네트

워크, 신경망, 자율이동로봇등임



片 爽 範(正會員)

1988年 朝鮮大學校 電子工學科 卒業
(工學碩士). 1994年 朝鮮大學校 大學
院 卒業(工學博士). 2000年 現在 東
岡大學 電子情報科 教授. 主 관심 분
야는 멀티미디어, 영상통신, 네트워

크, 신경망, 자율이동로봇 등

鄭 憲(正會員)

1999年 朝鮮大學校 大學院 制御計測工學科 卒業(工學博
士). 2000年 現在 草堂大學校 電子工學科 教授. 主 관심
분야는 인공지능, 신경망, 마이크로프로세서