

SoC 칩 구현을 위한 얼굴 인식

김대진, 전봉진 (포항공과대학교, 연세대학교 생체인식 연구센터)

1. 서론

기존의 전통적인 인식 방법을 대체할 수단으로써 생체 인식 기술에 대한 연구가 다양하게 진행되고 있다. 기존의 인식 방법은 패스워드처럼 사용자가 알고 있는 것이나, 스마트카드처럼 사용자가 갖고 있는 것을 이용하여 출입 통제 여부를 결정하였다. 하지만 이러한 방법은 사용자가 패스워드나 스마트카드를 잊어버릴 수도 있고 악의를 가진 다른 사용자가 패스워드를 예측하여 사용하거나 스마트카드를 훔쳐서 사용할 수 있다는 단점을 갖고 있었다. 한편, 생체 인식 기술은 홍채, 지문, 손바닥, 얼굴, 서명, 행동 양식, 및 목소리와 같이 사용자가 알고 있거나 갖고 있는 것이 아닌, 사용자 그 자신이 가진 물리적 혹은 행동적 특징을 인식에 사용한다. 이러한 특징들은 패스워드처럼 잊어버리거나 스마트카드처럼 잃어버리지도 않고, 다른 사용자가 훔쳐서 사용하기 어려우므로 기존의 방법들 보다 더욱 안전하고 확실한 인식 방법이라고 할 수가 있다. 생체 인식의 여러 특징들 가운데에서도 얼굴 특징은 특정한 장치를 이용하여 직접 입력을 받아야 하는 다른 특징들에 비해서, 사용자의 직접

입력 없이 카메라의 입력 영상만을 이용하여 좀 더 친숙하고 편안한 방법을 제공하므로, 그림 1과 같이 다양한 분야에서 활용하거나, 적용 중에 있다.

얼굴 인식 알고리즘은 조명, 포즈, 표정, 장신구, 나이 등 사용자의 외부환경 또는 얼굴변화에서 발생하는 다양한 변화에 신뢰성 있는 알고리즘을 개발하기 위하여 다양한 방법으로 개발되고 있다. 현재 개발되고 있는 얼굴 인식 방법은 얼굴 데이터 입력 방식에 따라 이차원 인식(일반적인 얼굴 인식 데이터)과 삼차원 인식(얼굴의 포즈 변화에 강인함, 학습/테스트 데이터의 부족)으로 구분할 수 있으며, 데이터의 특징에 따라 적외선 카메라를 이용한 얼굴의 열 패턴 영상(조명 변화에 강인함, 응용 분야가 한정됨)과 컬러(얼굴의 피부색으로 모델링, 처리 속도가 빠름, 조명, 인종 변화에 민감함) 및 흑백 데이터(일반적인 얼굴 인식 데이터) 등으로 나눌 수가 있다.

본고에서는 이차원 흑백 데이터를 입력으로 하는 얼굴 검출 및 인식 알고리즘의 실시간 처리를 위한 하드웨어 구조를 제시한다. 현재까지 개발된 얼굴 인식 시스템은 실시간 처리를 위해서

응용분야	활용 방법
비디오폰 분야	홈 네트워크(사이버/유비쿼터스) 아파트의 출입 개폐 시스템 및 방범 시스템, 가정내외 네트워크 인지정보 및 명령 교환 시스템
모바일 분야	카메라센서와 연동하여 각종 정보활용(금융, 통화 선택) 무선 인터넷 컨텐츠 성인물의 원천적인 차단
ATM 분야	금융 자동화기기에 현금인출기용 안면인식시스템의 활용
지능 로봇분야	로봇의 다양한 네트워크를 위한 명령자의 각종 지시인식정보(sign)활용 로봇 지능형 인식 시스템으로 활용
보안/감시 /통제분야	공항, 항망 등 신속 데이터 판독 후 검출 판별, 요주의 인물 관리
고령자/지체자 활용분야	안면인식 표정검출로 정보기기 및 기타 명령, 지시 등 의사소통 가능
기타	Smart Toy, Volp 화상전화 등

〈그림 1〉 얼굴 인식의 활용 분야

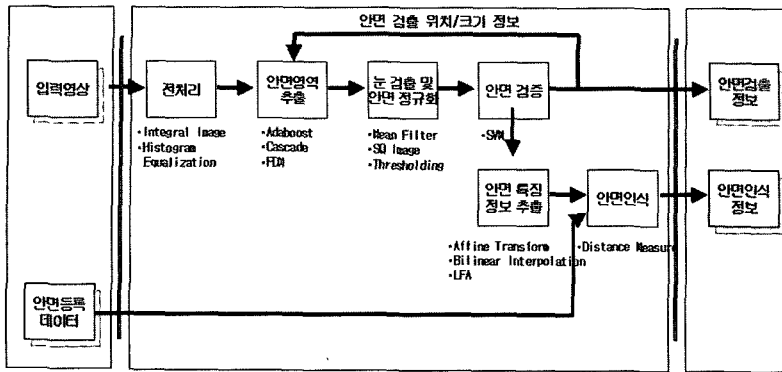
는 고사양의 시스템이 필요하다. 하지만, 가정용 출입 통제 시스템, 모바일 환경, 로봇과 같은 임베디드 시스템이 요구되는 곳에서는 적용이 힘들기 때문에, 적용이 매우 제한적이며 편리성, 가격 경쟁력이 떨어진다. 이를 위하여, 실시간으로 얼굴을 검출하고 인식하기 위한 메모리 구조 및 하드웨어 구조를 제시하고 이를 FPGA 및 SoC 칩으로 구현하여 설계한 하드웨어 구조가 임베디드 시스템의 핵심 코어로 적합함을 제시하려 한다.

논문의 구성은 다음과 같다. II장에서는 얼굴 인식 알고리즘에 대해 설명하며, III장에서는 얼굴 인식 알고리즘을 하드웨어로 구현하기 위한 방법을 제시하고 전체 하드웨어 구조를 설명한다. 그리고 IV장에서는 설계한 얼굴 검출 하드웨어를 검증하고 성능을 분석하며, 마지막으로 결론이 뒤따른다.

II. 얼굴 인식 알고리즘

얼굴 인식은 크게 얼굴 검출과 얼굴 인식의 두 단계로 나눌 수 있다.

얼굴 검출 과정에서는 입력된 카메라 영상에서 얼굴영역을 찾아 검출하며, 검출된 얼굴 영상을 전처리하여 얼굴 인식과정으로 넘겨주는 역할을 한다. 이때 입력된 영상에서 얼굴의 위치 및 크기를 미리 알 수 없기 때문에 입력 영상의 모든 영역에서 일부 영역의 위치와 크기를 변경해 가면서 얼굴인지 아닌지를 판별하는 방법으로 검출한다. 때문에 얼굴인지 아닌지를 판별하는 방법의 검출 성능 및 처리속도가 얼굴 검출의 성능에 가장 중요한 역할을 하게 되는데, 여기서는 미리 학습된 얼굴과 비-얼굴을 가장 잘 구분하는 간단한 필터들의 선형 조합으로 처리함으로써 매우 뛰어난 검출 성능과 실시간 처리가 가능하였다.



〈그림 2〉 얼굴 인식 알고리즘의 구조 및 흐름

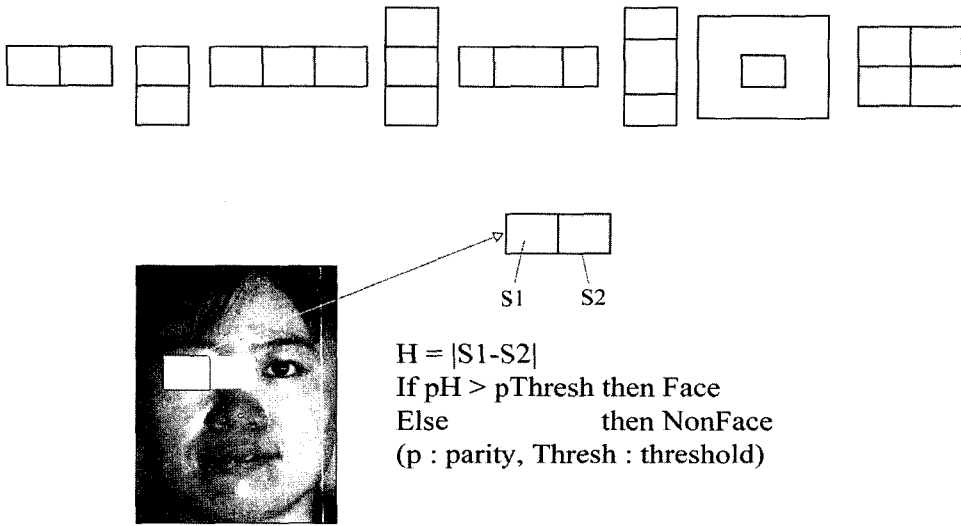
얼굴 인식은 입력 영상으로 들어온 얼굴이 데이터베이스에 저장된 얼굴 자료중의 누구와 가장 유사한지를 밝혀내는 과정이다. 얼굴 인식은 얼굴 내에서 어떤 정보를 추출하여 사용하느냐에 따라서 크게 기하학적인 특징 정합 방법과 템플릿 패턴 정합 방법으로 나눌 수 있다. 기하학적인 특징 정합 방법은 눈, 코, 입과 같은 얼굴을 구성하는 요소들의 특징점을 찾아서 각 점들 사이의 위치, 모양 등을 측정함으로써 얼굴영상들 사이의 유사도를 비교한다. 이 방법은 얼굴영상의 크기가 작거나 해상도가 낮을 때 얼굴의 특징점을 찾아내는데 어려움이 있고, 얼굴의 포즈가 바뀌었을 때 정면 얼굴이었을 때와 비교하여 특징점 사이의 거리 변화가 커서 인식이 어렵다는 단점이 있다. 한편, 템플릿 패턴 정합 방법은 얼굴내의 특징점을 찾는 것이 아니라, 얼굴 영상의 명암도를 이용하여 인식을 시도한다. 얼굴 인식에 사용되고 있는 대표적인 템플릿 패턴 정합 방법으로는 주성분 분석(PCA), PCA와 유사하지만 공분산 행렬대신 클래스 각각의 고유치 분석이 이루어진다는 차이점을 갖는 선형 판별 분석(LDA) 등이 있다. 본 연구에서는 통계적으로 계산된 국부적인 특징과 위치로 객체를 표현하는

국부적 특징 분석(LFA) 방법을 이용함으로써 조명 변화에 강한 얼굴 인식 시스템을 얻을 수 있었다.

1. 얼굴 검출 알고리즘의 수행 과정

그림 2는 본 연구에서 구현한 얼굴 검출 시스템으로 입력 영상에서 조명 성분을 제거하는 전처리, 정면 얼굴 영상 추출, 추출한 정면 얼굴에서의 눈 검출 과정과 정규화 과정, 마지막으로 얼굴인지 판단하는 안면 검증의 과정으로 구성되어 있다.

기존에는 얼굴과 비얼굴을 구별하기 위하여 복잡한 패턴 인식, Neural Network과 같은 방법을 이용함으로써, 검출 성능은 높지만, 계산하는데 필요로 하는 연산이 기하급수적으로 커져서 실시간 처리가 힘들 뿐만 아니라, 연산에 필요로 하는 데이터의 크기도 커서 칩으로 구현하기가 매우 힘들었다. 또한 실시간 처리를 위해서 칼라 영상에서 얼굴의 피부색을 모델링함으로써 입력 영상의 칼라 분포에서 피부색 분포를 추출하는 방법으로 얼굴을 검출하는 방법이 개발되기도 하였지만, 다양한 인종, 조명 변화에 민감하



〈그림 3〉 얼굴 검출에 사용된 사각형 필터 및 적용 방법

계 반응함으로써 원하는 수준의 검출 성능을 얻기가 힘들었다. 하지만, 본 연구에서는 얼굴과 비얼굴을 잘 구별하는 간단한 사각형 필터들의 선형 조합을 얻어낼 수 있는 Adaboost 알고리즘¹¹⁾을 이용하였다. 이때 얻어진 간단한 사각형 필터들을 칩 내부 메모리에 적재하고, 필터링 연산만을 이용하여 검출이 가능함으로써, 쉽게 칩으로 변환이 가능하도록 한다.

본 연구에서 구성한 사각형 필터는 그림 3과 같이 입력 영상에서 검은 영역과 흰 영역의 픽셀 값의 절대차에 의하여 구분하는 것으로, 영상의 전체적인 조명변화에도 이러한 절대차의 변화는 단순히 픽셀 값 자체를 이용하는 것 보다 훨씬 높은 검출 성능을 보일 수 있었다. 이때 가장 많이 차지하는 연산은 사각형 영역내부의 픽셀 합을 구하는 연산인데, 이는 전처리 단계에서 얻은 누적 영상(Integral Image)을 이용함으로써 4번 이하의 덧셈, 뺄셈 연산을 통하여 신속하게 얻음으로써 실시간 처리를 할 수 있었다. 또한 몇 개의 사각형 필터의 선형조합을 단계적으로

배치하는 캐스케이드(Cascade) 방법을 적용함으로써 간단한 배경 영상과 같은 것은 초기의 단계에서 배경영역으로 걸러낼 수 있었다.

그림 4의 오른쪽과 같이 입력 영상으로 수정된 Self Quotient 영상을 얻어내어, 안면 영상 추출 단계에서 얻어진 안면 영역 내에서, 눈 위치를



〈그림 4〉 조명 성분 제거 및 눈 검출을 위해 사용한 수정된 Self Quotient Image



〈그림 5〉 얼굴 검출 결과

찾기 위해서 반복적 임계치 설정(Iterative Threshold Selection) 방법^[2]를 통하여 눈 영역을 검출하였다.

마지막으로 얼굴 검증을 수행하여 찾은 후보 얼굴 영역이 정확한지를 판별하기 위해 두 눈의 위치를 기준으로 한 정규화 변환을 수행하여 25x20 픽셀 이미지를 만든 후 SVM (Support Vector Machine)을 이용하여 학습한 500x977 배열 파라미터와 25x20 이미지와 내적의 합 연산을 수행하여 결과가 일정 임계치보다 높을 경우 추정된 얼굴 후보가 실제 얼굴임을 결정하였다.

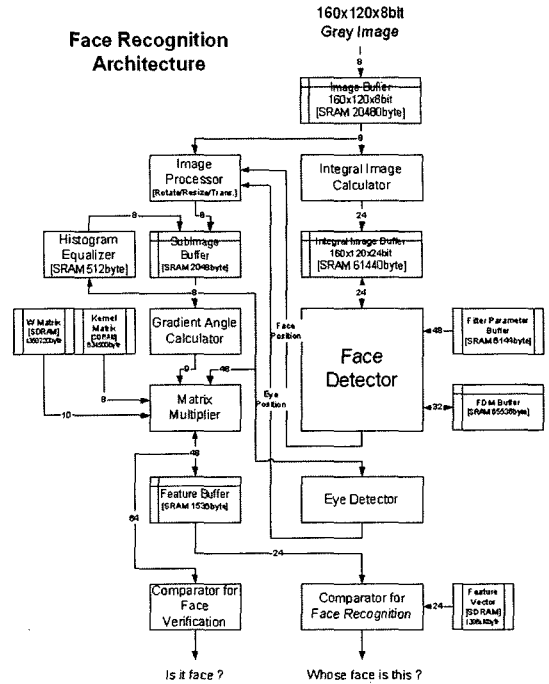
2. 얼굴 인식 수행 과정 및 알고리즘

얼굴 인식은 조명성분을 제거하기 위하여 그림 4와 같이 수정된 Self Quotient 영상에서 얼굴 검출 단계에서 얻어진 눈 위치를 이용하여 40x48픽셀 이미지로 정규화한 이미지를 사용한다. 이때, 이진 마스크로 필터링하여 얼굴 특징을 추출하기 위한 영역만을 선택한다. 얼굴 특징 추출은 LFA(Local Feature Analysis)^[3] 이용하여 정규화 영상을 40개 특징 벡터를 구한다. 마지막으로 추출한 특징 벡터는 등록된 특징 벡터와의 거리 연산을 이용하여 누구의 얼굴인지 인식할 수 있다.

III. 실시간 처리를 위한 얼굴 검출 및 인식의 하드웨어 구조 및 검증

1. 실시간 처리를 위한 얼굴 검출 및 인식의 하드웨어 구조

얼굴을 추정하고 인식하는 주요한 과정은 학습된 파라미터나 픽셀 데이터를 저장하고 불러오는 수행과 이 데이터를 연산하거나 주소를 계

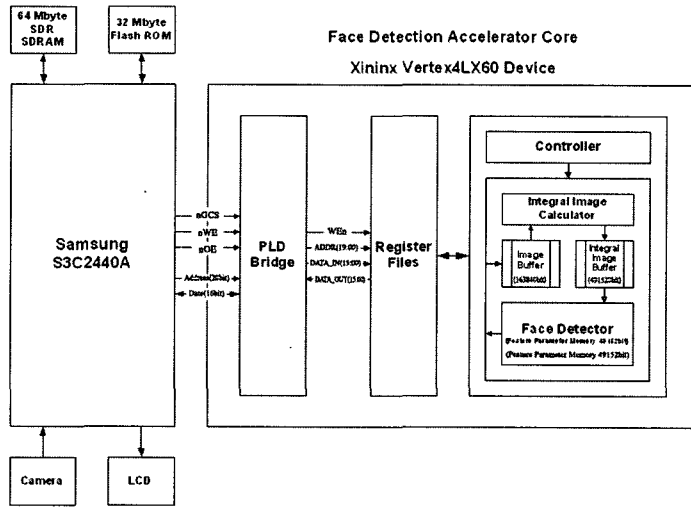


〈그림 6〉 얼굴 검출 및 인식기의 하드웨어 구조

산하는 수행이다. 실시간 얼굴 인식을 위한 하드웨어 구현을 위해서는 효과적인 메모리 구조와 병렬 처리 연산 모듈 설계가 필수이다.

먼저 2장에서 얻어진 각각의 알고리즘은 부동소수점 모델링(Float-point modeling)으로 구성되었다. 이를 칩으로 구성하기 위하여 각각의 모듈을 고정 소수점 모델링(Fixed-point modeling)으로 변환하였고, 이렇게 변환된 알고리즘을 이용하여 칩으로 변환하고, 테스트 벡터를 생성하였다. 이때 고정소수점 모델링을 위하여 얼굴 정규화의 과정과 같은 경우 얼굴의 확대, 축소, 회전 연산의 경우 sin, cos 연산뿐만 아니라 여러 부동 소수점 연산이 요구된다. 이를 Affine Transform과 Bilinear Interpolation을 이용함으로써 sin, cos 연산에 필요한 데이터 테이블이 필요가 없이 변환할 수 있었다.

얼굴 검출 및 인식은 학습한 파라미터를 저장



〈그림 7〉 얼굴 검출, 인식 검증 플랫폼

하기 위한 메모리와 이를 연산하기 위한 곱셈기가 요구된다. 파라미터 값을 외부 DDR SDRAM 메모리에 저장하도록 하였고, 필요시에 그 값을 불러들여 수행하도록 구성하여 내부 메모리 요구량을 최소화 하도록 하고, 외부 메모리 버스 크기만큼 처리할 수 있도록 곱셈기를 병렬로 두도록 하였다.

그림 6은 얼굴 검출 및 인식기의 하드웨어 블록 다이어그램을 보여주고 있다. 얼굴 추정 및 인식의 각 과정에 일어나는 연산 중, 유사 변환은 눈 위치 추정, 얼굴 검증 및 인식에서 수행되므로 로직 및 배열 내적 연산을 공유하였다.

2. 하드웨어 검증

얼굴 검출, 인식 하드웨어 구조를 검증하기 위하여 동일한 결과가 나오도록 각각의 처리 모듈에서 부동소수점으로 되어 있는 것을 최대 32비트가 되는 정수형의 고정소수점으로 변환하였다. 그리하여, Pentium 기반, ARM9(OMAP5912), FPGA에서 동일한 정수형의 고정소수점으로 변

환된 알고리즘으로 적용하여 동작시켜 표1의 결과를 얻을 수 있었다.

설계한 얼굴 검출, 인식 하드웨어 구조를 검증하기 위하여 그림 7과 같은 검증 플랫폼을 구성하였다. 하드웨어를 제어할 수 있도록 카메라 컨트롤러와 LCD 컨트롤러가 내장된 삼성사의 S3C2440A 프로세서를 사용하여 얼굴 이미지 입력과 결과 출력을 추가적인 로직 없이 수행할 수 있도록 구성하였다. 그리고 하드웨어 로직을 에뮬레이션하기 위하여 Xilinx사의 Vertex4LX60 FPGA 칩을 사용하고 하드웨어 내부에 브리지를 추가하여 프로세서와 시스템버스로 통신할 수 있도록 연결하였다. 리눅스 디바이스 드라이버 및 어플리케이션 프로그램을 작성하여 하드웨어를 초기화하기 위한 데이터를 플래시 메모리에서 불러들여 온 칩 SRAM에 저장하도록 하도록 하였으며, 수행 과정을 플래그 레지스터를 통해 읽을 수 있도록 하였고 수행이 끝날 시점에 폴링 방식으로 결과 값을 읽도록 구성하였다.

〈표 1〉 시스템 별 분석 결과

시스템	Pentium	ARM9	FPGA	SoC 칩(예상)
환경	CPU : Pentium2.8GHz RAM: 1.0GBytes	CPU: 192MHz RAM: 32MBytes	종류: Virtex4 Clock: 33MHz	Clock: 100MHz 게이트 : 40만 게이트 크기: 7*7mm
처리속도	0.2초	4~6초	0.5~0.7초	0.1초 이내
성능	Float-point modeling과 Fixed-point modeling은 허용범위내의 오차가 발생하도록 구성하였기 때문에 검출, 인식 성능은 동일, 때문에 모든 시스템에서 성능은 동일함. 얼굴 검출 성능 : FAR(0.01%), FRR(1%) 얼굴 인식 성능(등록100명 이내) : FAR(0.01%), FRR(3%)			
장단점	장점 : 범용적으로 활용 단점 : 고사양 컴퓨터	장점 : 임베디드 시스템 단점 : 속도가 느림	SoC칩을 위한 중간 개발 단계	고속 처리 다양한 응용가능 제품 단가 저렴
응용	출입통제시스템 ATM기 위변조 판별	Video 폰 출입통제시스템	SoC칩을 위한 중간 개발 단계	로봇, Mobile, PDA 출입통제시스템 Video 폰 비자/여권 생체인식

동작 주파수		100MHz
게이트 수		약 40만 게이트
메모리 크기	온 칩 메모리(SRAM)	737,280 비트
	오프칩 메모리(SDRAM)	약 2.8M 바이트
도입 IP		없음
Analog RF 회로		PLL
Test Vector Cycle 수		약 837,500,000클럭
Package 핀 수	Input 핀	51
	Output 핀	2
	Bi-direction 핀	32
	Power Vdd 핀 (Ext/Int)	14 (7/7)
	Power Vss 핀 (Ext/Int)	14 (7/7)
	Un-used 핀	31
	총 핀 수	144

〈그림 8〉 SoC 공정 계획

IV. 결 론

본고에서는 얼굴 인식 알고리즘의 실시간 처리를 위한 하드웨어 구조를 제안하고 이를 FPGA에 에뮬레이션 하였다. 먼저 얼굴 검출 및 인식 알고리즘을 설명하고, 알고리즘을 하드웨어로 구현하기 위한 효과적인 구조를 제시하고 이를 통합한 전체 구조를 설명하였다. 설계한 구조를 검증하기 위하여 FERET 이미지를 사용하여 부동 소수점 및 정수 소프트웨어의 FRR과 FAR를 하드웨어의 성능과 비교하였다. 그리고 이를 임베디드 프로세서와 FPGA 디바이스를 연결한 후 카메라에서 영상을 받아 얼굴을 검출하고 인식하는 과정을 수행하였다.

현재는 설계한 얼굴 인식의 하드웨어 구조를 매그너칩스 0.25m ASIC 공정을 거치기 위한 과정에 있다. 그리고 보다 높은 인식률 및 높은 성능을 위하여 여러 가지 추정 및 인식 알고리즘을 개선 중에 있으며, 내부 메모리를 최적화고 성능을 향상시키기 위한 소프트웨어 하드웨어 통합 설계 방법을 진행할 계획이다.

<Acknowledgement>

"This work was supported by the Korea Science and Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center(BERC) at Yonsei University."

참고문헌

[1] P. Viola and M. Jones, "Robust Real-time Object Detection", IEEE ICCV Workshop on Statistical and Computational Theories of Vision, Vancouver, 2001.
 [2] 정조남, 이필규, "얼굴 인식을 위한 효과적인 인

눈 위치 추출", 정보처리학회 논문지 제 12-B 권 2호 2005년 4월.

[3] P. Penev and J. Atick, "Local feature analysis: a general statistical theory for object representation", IOP Publishing, Network : Computation in Neural System Vol. 7, Aug. 1996.
 [4] Samsung Electronics, "S3C2440A - 32-BIT CMOS MICROCONTROLLER USER'S MANUAL", Samsung Electronics Co. LTD., 2004.

저자소개



김 대 진

1981년 연세대학교 전자공학과 학사졸업
 1984년 한국과학기술원 전기전자공학과 석사졸업
 1991년 Syracuse 대학 컴퓨터공학과 박사졸업
 1984년-1987년 KBS 기술연구소 연구원
 1992년-1999년 동아대학교 컴퓨터공학과 재직
 1999년-현 재 포항공과대학교 컴퓨터공학과 재직
 주관심분야 생체인식, 인간 컴퓨터 상호작용, 유비쿼터스 컴퓨팅



전 봉 진

2000년 부산대학교
 2002년 포항공과대학교
 2002년-2003년 삼성SDS 연구소 연구원
 주관심분야 생체인식, 인공지능, 영상처리, HCI