색상정보와 AdaBoost 알고리즘을 이용한 얼굴검출

나종원* · 강대욱** · 배종성***

Face Detection using Color Information and AdaBoost Algorithm

Jong-Won Na* · Dae-wook Kang** Jong-sung Bae***

요 약

얼굴 검출은 대부분 얼굴의 움직임 정보를 이용한다. 기존에 얼굴 검출 방법은 프레임 간의 차를 이용하여 움직임을 검출하는 방법이 사용되어 왔으나 대부분이 실시간을 고려하지 않은 수학적 접근법을 사용하거나 알고리즘이 지나치게 복잡하여 실시간 구현에 용이하지 않았다.

본 논문에서는 실시간 얼굴검출을 위하여 감시카메라에서 입력된 RGB영상을 YCbCr 영상으로 변환한 후 연속된 두 영상의 차를 구하고 Glassfire 라벨링을 실시했다. 라벨링 결과 가장 넓은 구역의 면적과 Area 임계치 값을 비교하여 임계값 이상의 면적이면 동작변환으로 인식하고 영상을 추출하였다. 이렇게 추출된 동작변환 영상을 대상으로 얼굴 검출을 실시하였다. 얼굴 검출에 필요한 특징을 추출하기 위해 AdaBoost알고리즘을 사용하였다.

ABSTRACT

Most of face detection technique uses information from the face of the movement. The traditional face detection method is to use difference picture method are used to detect movement. However, most do not consider this mathematical approach using real-time or real-time implementation of the algorithm is complicated, not easy. This paper, the first to detect real-time facial image is converted YCbCr and RGB video input. Next, you convert the difference between video images of two adjacent to obtain and then to conduct Glassfire Labeling. Labeling value compared to the threshold behavior Area recognizes and converts video extracts. Actions to convert video to conduct face detection, and detection of facial characteristics required for the extraction and use of AdaBoost algorithm.

키워드

face detection, adaboost, color information

1. 서 론

일반적으로 얼굴 검출은 영상 내에서 배경, 비얼굴과 차별화되는 얼굴만이 가지고 있는 고유한 특징을 이용한다[1]. 얼굴의 고유 특징들은 독특한 영역의 피부색상, 타원형의 형태, 대칭적 얼굴 구조, 얼굴의 요소(눈, 코, 입) 등이다. 현재 연구는 이와 같은 고유 특징들을 중심으로 이루어지고 있다. 그러나 얼굴의 다양한 크기 및 형태, 위치, 기울기와 같은 문제점은 얼굴 영역을 분리하는데 많은 어려움을 발생시킨다. 또한, 얼굴 영상 자체뿐만 아니라 복잡한 배경, 영상의 질, 조명에 따른 영상의 색상 변화 등도 얼굴 검출을 어렵게 하는 원인이다. 이러한

* 전남대학교 대학원 소프트웨어공학협동과정
** 전남대학교 공과대학 전자컴퓨터공학부 교수
*** 전남대학교 자연과학대학 수학통계학부 교수

접수일자 2008. 03. 28

문제를 해결하기 위해서 지금까지 많은 연구가 진행되어 왔고 현재도 정확도와 효율성을 높이기 위한 새로운 얼굴 검출 방법들이 제안되고 있다[2][3].

얼굴 검출은 대부분 얼굴의 움직임 정보를 이용한다. 기존에 얼굴 검출 방법은 프레임 간의 차를 이용하여 움직임을 검출하는 방법이 사용되어 왔으나[5] 대부분이 실시간을 고려하지 않은 수학적 접근법을 사용하거나 알고리즘이 지나치게 복잡하여 실시간 구현에 용이하지 않았다.

본 논문에서는 실시간 얼굴검출을 위하여 감시카메라에서 입력된 영상을 YCbCr 영상으로 변환한 후 연속된 두 영상의 차를 구하고 Glassfire 라벨링을 이용하여 동작변환 영상을 추출한 후 AdaBoost알고리즘을 사용하여 얼굴 검출에 필요한 특징을 추출하도록 제안하였다.

2. 얼굴검출 시스템의 구성

본 논문에서 제안한 얼굴검출 시스템의 전체적인 구성은 그림 1과 같이 나타낼 수 있다. 실시간 얼굴검출을 위하여 전 단계로 감시카메라에서 입력된 RGB영상을 YCbCr 영상으로 변환한 후 연속된 두 영상의 차를 구하고 Glassfire 라벨링을 실시했다. 라벨링 결과 가장 넓은 구역의 면적과 Area 임계치 값을 비교하여 임계값 이상의 면적이면 동작변환으로 인식하고 영상을 추출하였다. 이렇게 추출된 동작변환 영상을 대상으로 얼굴 검출을 실시하였다. 얼굴 검출에 필요한 특징을 추출하기 위해 AdaBoost알고리즘을 사용하였다.

2.1 YCbCr 영상으로 변환

YCbCr색체모델은 영상 시스템에서 사용되는 색공간의 일종이다. Y는 휘도 성분이며 Cb와 Cr은 색차 성분이다. YCbCr 은 가끔 YCC 라고 줄여 부르기도 한다. YCbCr 은 절대 색공간이 아니며 RGB 정보를 인코딩하는 방식의 하나로, 실제로 보여지는 이미지의 색은 신호를 디스플레이 하기 위해 사용된 원본 RGB 정보에 의존한다[6]. 여기서 RGB에서 YCbCr로 변환은 식(1)과 같다.

$$Y = k_r R + (1 - k_b - k_r) G + k_b B$$

$$C_b = \frac{0.5}{1 - k_b}(B - Y)$$

$$C_r = \frac{0.5}{1 - k_r}(R - Y) \quad \text{----------} (1)$$

인간의 시각시스템은 색상 간의 변화보다는 밝기의 변화에 더욱 민감하므로 YCbCr색상모델을 사용하면 색체정보에서 밝기 값의 영향을 배제한 데이터를 이용할 수 있어 감시영역의 조명의 변화에 의한 얼굴검출의 에러율을 최소화할 수 있다.

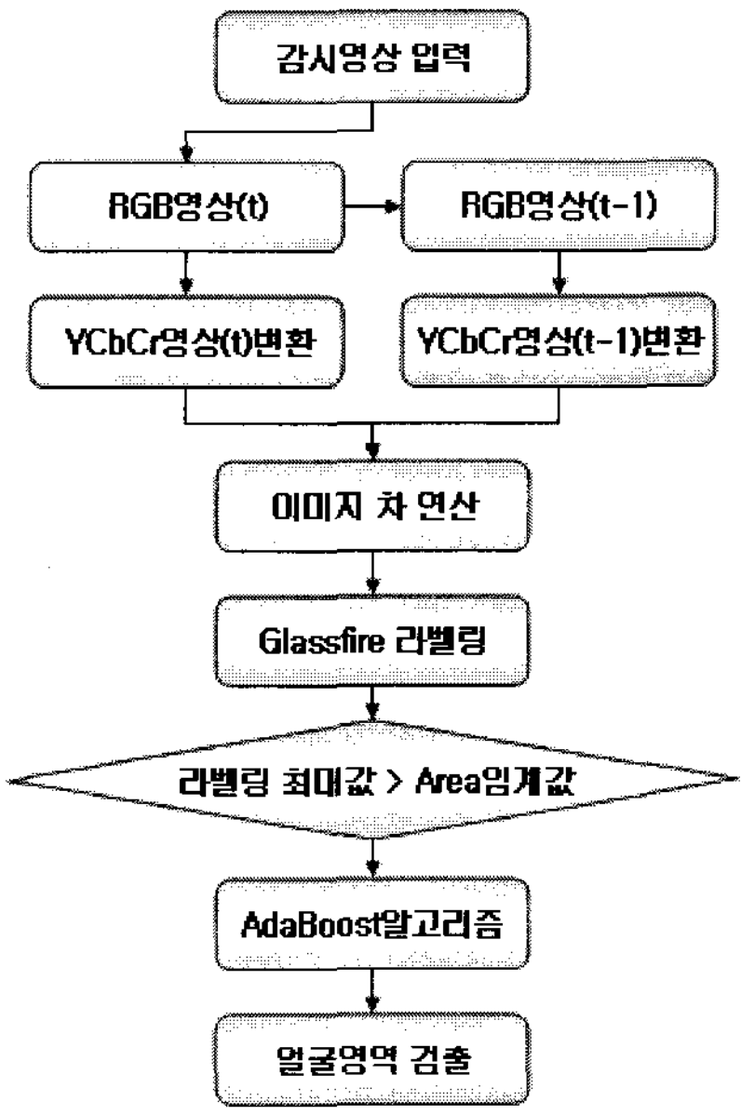

그림 4. 얼굴검출 시스템 흐름도
Fig. 1. Flow chart of face detection system

본 논문에서는 변환을 위하여 k_b=0.114와 k_r=0.299를 변환 가중치로 선택하였다. 또한 얼굴 색차의 정보의 YCbCr 표본형식은 사용자가 선택할 수 있도록 하였고 기본값은 4:4:4를 기본 표본 형식으로 채택하였고 사용자의 요구에 따라 조정할 수 있도록 구현하였다.

2.2 차영상 기법

본 논문에서 제안하는 얼굴검출 시스템은 기존의 시스템에서 연산양이 많아 실시간 얼굴검출이 어려운 부분을 개선하기 위해 영상처리 기법인 차영상(Difference Picture) 기법[4]을 이용하여 먼저 얼굴검출의 대상이 되

는 동작검출 영상을 추출하였다.

차 영상을 직접 사용하여 이동 물체의 영역을 구하는 방법은 잡음이나 카메라의 움직임, 조명 변화의 영향을 심하게 받는다는 단점을 갖는다. 또한 이동 물체의 밝기가 배경과 비슷하거나 느린 속도의 움직임일 경우에는 차 영상에서 이동 물체가 일으킨 변화가 잡음에 의해 발생한 변화보다 크지 않을 경우에 움직임이 전혀 나타나지 않을 수도 있다.

본 논문에서는 이러한 문제를 다루기 위해서 먼저, 입력된 RGB영상을 YCbCr 영상으로 변환하였다. YCbCr 색상모델을 사용하면 색체정보에서 밝기 값의 영향을 배제한 데이터를 이용할 수 있어 감시영역의 조명의 변화에 의한 얼굴검출의 에러율을 최소화할 수 있다. 이렇게 YCbCr 영상으로 변환한 후 연속된 두 영상의 차를 구하게 된다.

2.3 Glassfire 라벨링

감시영상에서 움직임을 감지하기 위하여 현재 프레임의 영상 정보와 이전 프레임의 영상 정보와의 차영상 기법을 이용하였다. 차영상 기법을 사용하기 위해서 그레이화시킨 값을 이용하고 차연산 결과가 실험에서 구한 임계값을 넘는다면 움직임이 발생했다고 판단하는데 이때 임계값 비교시 속도를 줄이기 위해서 픽셀 기반 프로세싱 (Pixel-based processing)이 아닌 블록 기반 프로세싱 (Block-based processing)을 한다. 픽셀 기반 프로세싱은 1 대 1 매칭을 수행하기 때문에 초당 많은 프레임을 처리해야 하는 실시간 환경에서는 다소 무리가 있다. 따라서 블록 기반 프로세싱을 하여 정확도는 약간 떨어지지만 속도를 높이는데 중점을 뒀다. 블록 기반 프로세싱이란 픽셀 기반 프로세싱과는 달리 여러 픽셀들을 한 블록으로 정하고 블록 대표 값을 구하여 프로세싱하는 라벨링(labeling)방법을 말한다[5].

라벨링 단계는 이진화된 영상을 탐색하다가 밝기가 255인 화소값을 만나면 라벨링을 수행하고 본 논문에서는 이 라벨링을 점을 8-근방 중심으로 이동 후 다시 인접 화소의 미방문 255화소값을 라벨링하는 방식으로 반복한다. 그림 2의 삼각형 꼭지부분을 라벨링하는 그림 2의 단계로 라벨링이 수행된다.

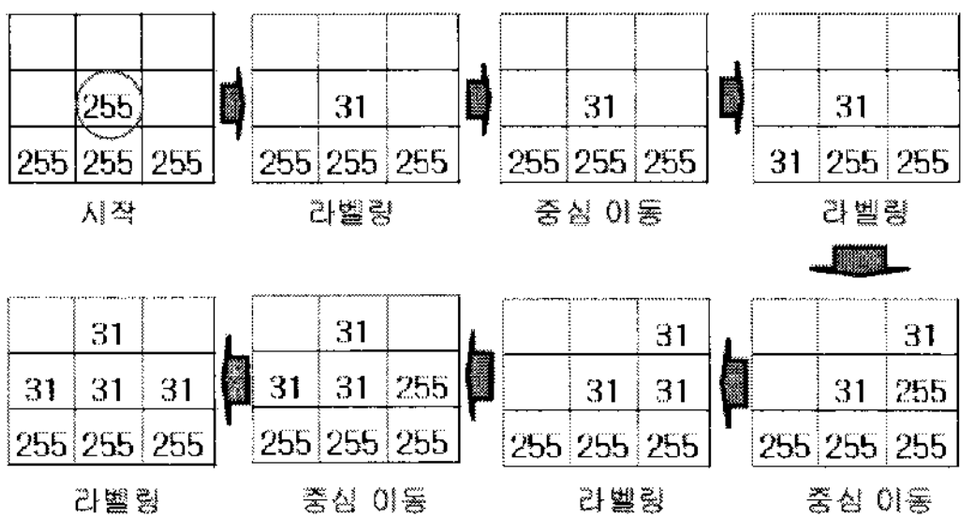

그림 5. 라벨링 단계
Fig. 2. Step of labeling

이진화된 영상으로부터 라벨링 알고리즘 식 2을 이용하여 해당 영상을 라벨링한다.

$$P\left(R_i^{(k)} \cup x\right) = \begin{cases} TRUE & \text{if } x > T \\ FALSE & otherwise \end{cases}, \text{for } i = 1,2,...N, \quad \text{-------- (2)}$$

P는 (R,x,T)의 논리적 판단 자, (k)는 각 단계, x는 영상의 픽셀 밝기 값, T는 임계치 값을 나타낸다. 각 영역 R은 region growing 알고리즘에 의하여 라벨링 된다.

Glassfire 알고리즘은 자기호출(recursive call)을 이용하여 모든 인접요소가 라벨링될 때까지 현재 관심화소의 주변 인접화소를 차례로 검사하면서 라벨링을 한다. 또한 자기홀출에 있어서 스택 자료구조를 사용하기 때문에 고속으로 아주 큰 영역을 라벨링하는 것도 가능하다.

2.4 AdaBoost알고리즘을 이용한 얼굴검출

본 논문에서는 얼굴 검출에 필요한 특징을 추출하기 위해 AdaBoost알고리즘을 사용하였다. AdaBoost알고리즘의 기본 개념은 약한 분류기(Weak classifier)를 선형적으로 결합하여 최종적으로 높은 검출 성능을 가진 강한 분류기(Strong classifier)를 생성하는 것이다. AdaBoost 알고리즘에 의해 생성된 강한 분류기는 계층적인 체계를 이룬다. 이것은 기존의 다른 방법들이 하나의 복잡한 마스크 형태의 분류기를 이용한 것과 달리 간단한 마스크를 여러 개의 층으로 형성한 것이다. 기존의 하나의 복잡한 분류기들은 실제 영상에서 얼굴을 검출하는데 복잡한 마스크와의 계산이 이루어지기 때문에 계산량이 많고 시간이 많이 소요되었다. 그러나 AdaBoost알고리

즘을 이용하여 생성된 계층적 분류기는 앞쪽 부분에 간단하면서 도 얼굴을 가장 잘 검출하는 것을 배치하고 뒤쪽 부분에 잘못 검출된 제거하는 형식으로 구성되어 있다. 이는 실시간 검출에 있어서 기존의 방법에 비하여 뛰어난 성능을 보인다. 이런 점을 고려하여 실시간으로 얼굴을 검출하고자 본 논문에서는 AdaBoost알고리즘을 사용하였다.

본 논문에서 사용한 AdaBoost알고리즘은 그림 3과 같다. 그림 3의 1에서 입력 값 중 x는 훈련영상을 y는 얼굴과 비얼굴을 나타낸다. 훈련 영상의 앞부분은 얼굴 영상이며 뒷부분은 비얼굴 영상이다. 그림 3의 2에서는 가중치(w)를 초기화한다. 가중치(w)는 각각의 훈련 영상의 중요도를 나타내고 후에 학습과정에서 어느 영상을 더 중요하게 학습할 것인지를 결정하는 역할을 한다. 그림 3의 3은 약한 분류기를 생성하는 단계이다.

1. Given N examples $(x_1, y_1), \cdots, (x_N, y_N)$
   with $x \in R^k, y_i \in \{1, -1\}$
2. Initialize weighs $w_1(i) = \frac{1}{m}, \frac{1}{n}$ $i = 1, \cdots, N$
   m: number of positive, n: number of negative
3. Repeat for t=1,···, T
(a) Train weak learner using weight w_i
(b) Repeat for j=1,..., number of intial feature
   Get weak classifier $h_i = X \rightarrow \{1, -1\}$
   with error $\epsilon_{j1} = \sum_i^m w_t(i)|h_j(x_i) - y_i|$ (positive)
   $\epsilon_{j2} = \sum_i^n w_t(i)|h_j(x_i) - y_i|$ (negative)
   choose the classifier using error (ϵ_t, h_t)
(c) weight update
   $\alpha_t = \frac{1}{2}\ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$
   $w_{t+1}(i) = w_t(i) \cdot \begin{cases} e^{-\alpha_t} & \text{if } h_t(x_i) = y_i \\ e^{\alpha_t} & \text{if } h_t(x_i) \neq y_i \end{cases}$
4. Output the final hypothesis(strong classifier)
   $H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$

그림 6 AdaBoost 알고리즘
Fig. 3. Algorithm of AdaBoost

약한 분류기는 최소의 에러를 갖는 하나의 특징만을 선택하나 본 논문에서는 효율성을 고려하여 에러의 임계치를 주어 한 단계에서 여러 개의 특징을 선택하도록 하였다. 이는 후에 강한 분류기가 계층적으로 분류되는 역할을 한다. 그리고 T는 단계를 나타내는 역할을 한다. (b)에서 j는 특징의 수를 나타내며 j만큼 반복하면서 임계치보다 높은 에러값을 찾는다. 에러값은 positive일 경우와 negative일 경우를 나누어서 계산하였다. 이는 학습 초기에는 positive의 에러율을 줄이는데 집중하고 후반부에는 negative의 에러율을 줄이기 위함이다. (c)에서는 가중치를 업데이트하는 부분이다. 이때 잘못 분류된 훈련 영상은 가중치 w(i)를 증가시키고, 옳게 분류된 훈련 영상은 가중치 w(i)를 감소시킨다. 이것은 초기에 선택된 특징들은 쉽게 얼굴 영상과 비얼굴 영상을 구별할 수 있는 역할을 담당하지만 후반부에 선택된 특징들은 얼굴 영상과 비얼굴 영상의 구별이 어려운 것을 구별하는 역할을 담당하기 위함이다.

3. 실험 및 분석

본 논문에서 제안한 얼굴검출 시스템은 실시간 얼굴검출을 위하여 전 단계로 감시카메라에서 입력된 RGB 영상을 YCbCr 영상으로 변환한 후 연속된 두 영상의 차를 구하고 Glassfire 라벨링을 실시했다. 라벨링 결과 가장 넓은 구역의 면적과 Area 임계치 값을 비교하여 임계값 이상의 면적이면 동작변환으로 인식하고 영상을 추출하였다. 이렇게 추출된 동작변환 영상을 대상으로 얼굴 검출을 실시하였다. 얼굴 검출에 필요한 특징을 추출하기 위해 AdaBoost알고리즘을 사용하였다.

그림 4 다양한 환경에서 얼굴 검출
Fig. 4. Face detection from variable environment

이렇게 제안된 시스템의 성능을 검증하기 위해서 일반 퍼스널 컴퓨터와 웹캠(webcam)을 이용하여 실험을 실시하였다. 본 실험에서는 1 초당 30 프레임을 입력으로 받았다.

본 실험을 위해 방문자를 대상으로 얼굴을 검출하였고 검출된 얼굴 영상은 캡쳐를 하였다. 실험 결과는 표 1과 같다.

<표 1> 사람 수에 따른 얼굴 검출 결과
Table 1. Result of Men face detect

구 분		한명 경우	두명 경우	세명 경우	네명 경우
총 횟수		123	120	116	94
검출성공	제안기법	112	108	102	62
	기존기법	111	105	96	51
검출실패	제안기법	11	12	14	32
	기존기법	12	15	20	43
검출확률	제안기법	91.0%	90.0%	87.9%	65.9%
	기존기법	90.2%	87.5%	82.8%	54.2%

표 1 에서는 사람 수에 따른 얼굴 검출 결과를 보여 주고 있다. 한 명일 경우 검출되는 확률이 91%, 두 명일 경우 검출되는 확률이 90%, 세 명일 경우 검출되는 확률이 87.9%이다. 네명 이상일 경우는 65.9%로 낮아지고 있는 것을 볼 수 있다. 또한 기존의 기법에 비해 제안된 기법으로 얼굴을 검출할 경우 검출하고자하는 얼굴의 수가 많아질수록 검출확률이 높아지는 것을 알 수 있다.

4. 결 론

얼굴 검출은 대부분 얼굴의 움직임 정보를 이용한다. 기존에 얼굴 검출 방법은 프레임 간의 차를 이용하여 움직임을 검출하는 방법이 사용되어 왔으나 대부분이 실시간을 고려하지 않은 수학적 접근법을 사용하거나 알고리즘이 지나치게 복잡하여 실시간 구현에 용이하지 않았다.

본 논문에서는 실시간 얼굴검출을 위하여 전 단계로 감시카메라에서 입력된 RGB영상을 YCbCr 영상으로 변환한 후 연속된 두 영상의 차를 구하고 Glassfire 라벨링을 실시했다. 라벨링 결과 가장 넓은 구역의 면적과 Area 임계치 값을 비교하여 임계값 이상의 면적이면 동작변환으로 인식하고 영상을 추출하였다. 이렇게 추출된 동작변환 영상을 대상으로 얼굴 검출을 실시하였다. 얼굴 검출에 필요한 특징을 추출하기 위해 AdaBoost알고리즘을 사용하였다.

제안된 시스템의 성능을 검증하기 위해서 일반 퍼스널 컴퓨터와 웹캠(webcam)을 이용하여 실험을 실시하였다. 본 실험에서는 1 초당 30 프레임을 입력으로 받았다. 사람 수에 따른 캡쳐 결과는 한 명일 경우 캡쳐되는 확률이 91%, 두 명일 경우 캡쳐되는 확률이 90%, 세 명일 경우 캡쳐되는 확률이 87.9%, 네 명 이상일 경우 캡쳐되는 확률은 65.9% 이다. 또한 기존의 기법에 비해 제안된 기법으로 얼굴을 검출할 경우 검출하고자하는 얼굴의 수가 많아질수록 검출확률이 높아지는 것을 알 수 있다.

향후에는 본 시스템을 응용하여 측면얼굴의 검출율과 다수 얼굴의 검출율을 향상시킬 수 있도록 하여야 할 것이며 또는 걸어 다니는 사람, 자동차 등 다양한 다른 물체의 검출에도 적용이 가능할 수 있을 것이다.

참고문헌

[1] H. Rowley, S. Baluja, T. Kanade, "Neural network-based face detection," In IEEE Patt. Anal. Mach. Intell, 20, pp 22-38, 1998.

[2] E. Osuna, R. Freund, and F. Girosi, "Traning support vector machines: An application to face detections," In IEEE Proc. of Int. Conf. on Computer Vision and Pattern Recognition, 6, 1997.

[3] A. L. Yuille, P. W. Hallinan, and D. S. Cohen, "Feature extraction from faces using deformable templates," Int. J. Comput. Vision, 8, pp 99-111, 1992.

[4] J. L. Crowley and F. Berard, "multi-model tracking of faces for video communications," In IEEE Proc. of Int, Conf. on Computer Vision and Pattern Recognition, Puerto Rico, June. 1997.

[5] H. Graf, E. Cosatto, and T. Ezzat, "Face analysis for the synthesis of photo-realistic talking heads," In Proceedings Fourth IEEE International Conf on Automatic Face and Gesture Recognition, 2000.

[6] R. Herpers, M. Michaelis, "Edge and keypoin detection in fatial regions," In IEEE Proc. of 2nd Int. Conf. on Automatic Face and Gesture Recognition, Oct. 1996, pp 212-217.

[7] Paul Viola, M. Jones, "Robust real-time object detection," International Conference on Computer Vision, 2001.

저자소개

나 종 원(Jong-won Na)

2003년 전남대학교 소프트웨어공학 협동과정 (공학석사)
2007년 전남대학교 소프트웨어공학 협동과정 (박사수료)
2005년~2007년 정보통신부 IT국내교수요원
동강대학 정보통신과 초빙전임강사
※관심분야 : Digital Image Processing, 사이버범죄, 정보시스템 감리

강 대 욱(Dae-wook Kang)

1985년 전남대학교 계산통계학과 졸업 (이학석사)
2000년 University of Newcastle, U.K. 수료 (박사수료)
1984년~ 전남대학교 전자컴퓨터공학부 교수
※관심분야 : Mobile Computing, Distributed Computing, Networks

배 종 성(Jong-sung Bae)

1975년 고려대학교 통계학과 졸업 (이학사)
1977년 고려대학교 대학원 통계학과 (이학석사)
1987년 고려대학교 대학원 통계학과 (이학박사)
1980년~ 전남대학교 수학통계학부 교수
※관심분야 : PBIB design