
서베이런스에서 Adaptive Boosting을 이용한 실시간 헤드 트래킹

강성관*, 이정현**

Real-Time Head Tracking using Adaptive Boosting in Surveillance

Sung-Kwan Kang*, Jung-Hyun Lee**

요약 본 논문에서는 복잡한 배경에서의 사람의 머리 추적에 있어서 효과적인 Adaptive Boosting에 의한 방법을 제안한다. 하나의 특징 추출 방법은 사람의 머리를 모델링하기에는 부족하다. 따라서 본 연구에서는 여러 가지 특징 추출 방법을 병행하여 정확한 머리 검출을 시도하였다. 머리 영상의 특징 추출은 sub-region과 Haar 웨이블릿 변환(Haar wavelet transform)을 이용하였다. Sub-region은 머리의 지역적인 특징을 나타내고, Haar 웨이블릿 변환은 얼굴의 주파수 특성을 나타내기 때문에 이들을 이용하여 특징을 추출하면 효과적인 모델링이 가능해진다. 실시간으로 입력되는 영상에서 사람의 머리를 추적하기 위하여 제안하는 방법에서는 3가지 형태의 Harr-wavelet 특징을 AdaBoosting 알고리즘으로 학습한 후 결과를 이용하였다. 원래 AdaBoosting 알고리즘은 학습시간이 매우 길며 학습데이터가 변하면 다시 학습을 수행해야 하는 단점이 존재한다. 이 단점을 극복하기 위하여 제안하는 방법에서는 캐스케이드를 이용한 AdaBoosting의 효율적인 학습방법을 제안한다. 이 방법은 머리 영상에 대한 학습시간을 감소시키며, 학습데이터의 변화에도 효율적으로 대처할 수 있다. 이 방법은 학습과정을 레벨별로 분리한 후 중요도가 높은 학습데이터를 다음 단계에 반복적으로 적용시킨다. 제안하는 방법이 적은 학습 시간과 학습 데이터를 사용해서 우수한 성능을 가지는 분류기를 생성하였다. 또한, 이 방법은 다양한 머리데이터를 가진 실시간 영상데이터에 적용한 결과 다양한 머리를 정확하게 검출 및 추적하였다.

주제어 : 머리 추적, AdaBoosting, 얼굴 검출, 특징점 추출, Haar Wavelet

Abstract This paper proposes an effective method using Adaptive Boosting to track a person's head in complex background. By only one way to feature extraction methods are not sufficient for modeling a person's head. Therefore, the method proposed in this paper, several feature extraction methods for the accuracy of the detection head running at the same time. Feature Extraction for the imaging of the head was extracted using sub-region and Haar wavelet transform. Sub-region represents the local characteristics of the head, Haar wavelet transform can indicate the frequency characteristics of face. Therefore, if we use them to extract the features of face, effective modeling is possible. In the proposed method to track down the man's head from the input video in real time, we use the results after learning Harr-wavelet characteristics of the three types using AdaBoosting algorithm. Originally the AdaBoosting algorithm, there is a very long learning time, if learning data was changes, and then it is need to be performed learning again. In order to overcome this shortcoming, in this research propose efficient method using cascade AdaBoosting. This method reduces the learning time for the imaging of the head, and can respond effectively to changes in the learning data. The proposed method generated classifier with excellent performance using less learning time and learning data. In addition, this method accurately detect and track head of person from a variety of head data in real-time video images.

Key Words : Head Tracking, AdaBoosting, Face Detection, Feature Extraction, Haar Wavelet

이 논문은 인하대학교의 지원에 의하여 연구되었음.

*인하대학교 컴퓨터정보공학부 박사과정

**인하대학교 컴퓨터정보공학과 교수

논문접수: 2013년 1월 9일, 1차 수정을 거쳐, 심사완료: 2013년 1월 30일, 확정일: 2013년 2월 20일

1. 서론

최근까지 서베일런스 분야에서 카메라로부터 입력되는 실시간 영상에서 사람의 얼굴을 검출하고 머리를 추적하는 연구가 활발하게 진행되어져 왔다. 이 연구들은 얼굴 검출 시스템, 생체 인식 시스템, 화상회의, 이미지 및 비디오 데이터베이스 인덱싱, 인간과 컴퓨터와의 상호작용에 관련한 대화형게임기 등에 주로 적용되어지고 있다.

사람의 머리 검출은 서베일런스 시스템에서 사람을 인식하기 전에 이루어져야 하는 중요한 단계이다. 이런 머리 검출은 여러 가지 방법으로 수행될 수 있다. 영상내의 윤곽선 정보, 색상 정보, 움직임 정보 등을 이용하는 방법과 템플릿 정합, 신경망, SVM (Support Vector Machine)과 방법 등이 있다[1][2].

본 논문에서 제안하는 방법에서는 머리 패턴을 sub-region과 Haar Wavelet 변환을 이용하여 모델링하고 AdaBoosting 분류기를 이용하여 머리의 검출을 시도한다. Sub-region은 얼굴 영상의 지역적인 특징을 나타낼 수 있다. 흑백 영상에서 얼굴의 영상은 각 점들의 밝기 값으로 표현이 된다. 이런 밝기 값들은 눈, 코, 볼 등과 같은 특정 부분에서 비슷한 값을 가지게 된다. 따라서 이런 영역들의 평균값을 이용해서 얼굴을 모델링 할 수 있다. Haar 웨이블릿 변환은 머리 영상의 주파수 특성을 나타낼 수 있다[3][4]. DCT(Discrete Cosine Transform), 푸리에 변환(Fourier transform), 웨이블릿 변환 등은 모두 주파수 특성을 분석하는데 이용할 수 있다. 주파수 영역으로 표현된 영상은 여러 주파수 성분 중에서 낮은 주파수 대역이 그 영상의 특성을 잘 나타내게 된다. 따라서 낮은 주파수 대역의 성분이 높은 주파수 대역의 성분보다 분별력이 크게 나타난다[5]. 이런 특성을 이용하여 영상을 나타내는 차원이 낮은 특징 벡터로 영상을 표현할 수 있다. 머리의 검출은 영상의 각 점의 값들을 그대로 이용하여 검출할 수도 있고, 또는 이 값들을 이용하여 특징을 추출하여 검출할 수도 있다.

실시간 비디오 영상에서 머리를 정확하게 검출하는 것은 매우 어려운 일이다. 왜냐하면, 다양한 조명환경이 발생하고 머리 주위에 복잡한 배경 및 사람의 얼굴 형태의 변화가 발생할 수 있기 때문이다. 사람의 머리를 검출하기 위하여 기존의 연구되었던 방법들이 머리 검출 성능을 매우 향상시켜 왔다. 그러나, 공통적으로 연산 시간이 오래 걸린다는 단점들이 대부분 존재한다. 그래서 서

베일런스에서 실시간으로 머리를 검출할 때 속도 문제나 머리 추적에 실패하는 경우가 자주 발생하였다.

이 논문에서는 효율성이 더 좋은 머리 검출 및 추적을 달성하기 위해 간단한 웨이블릿 특징을 기반으로 AdaBoosting 알고리즘을 이용하여 분류기를 학습한 후에 머리를 추적하도록 하였다. 기존의 AdaBoosting 알고리즘은 학습 시간이 너무 길며 학습 영상의 변화에 대처하는 데 어려움이 많았다. 본 논문에서는 카메라로부터 입력되는 영상으로부터 실시간으로 머리를 검출하는 시스템을 구현하고 실험하였다. 이 시스템은 학습 시간이 적게 걸리며, 학습데이터의 변화에도 일정한 머리 추적 성능을 보인다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 머리 추적 기술에 대해서 설명한다. 3장에서는 제안하는 머리 추적 알고리즘 및 시스템에 대하여 기술한다. 4장에서 제안하는 시스템의 실험 결과를 기술하고, 5장에서 결론 및 향후 연구에 대하여 기술한다.

2. 기존의 머리 추적 방법

머리 영역 검출에 관한 기존 연구들은 크게 형판 정합 모델(template matching model), 특징 기반의 상향식 모델(feature-based bottom-up model), 지식 기반의 하향식 모델(knowledge-based top-down model), 외형 기반 모델(appearance-based model)의 종류로 구분해서 볼 수 있다[6][7].

형판 정합 모델은 표준적인 머리의 패턴을 미리 구성하고, 입력 영상과의 상관관계를 비교하여 머리를 검출하는 방법이다. 그러나 머리의 자세나 모양, 크기 등이 다양하기 때문에 하나의 형판으로는 검출이 불가능 하다. 이의 대안으로 Yuille는 가변 형판 모델(deformable template model)을 제안하였다. Yuille는 기준적인 형판을 만들되, 매개변수에 의해 그 크기와 모양이 변형되도록 하였다[8].

특징 기반의 상향식 모델은 머리 영상들의 공통적인 특징들을 추출해 내고, 이를 기반으로 머리 검출을 수행한다. T.K. Leung, M.C. Burl와 P. Perona는 머리의 특징을 추출하고, 이를 바탕으로 머리 각 점의 특징들을 구성한 다음, 이를 정합시키는 알고리즘으로 머리 검출에 성공하였다[9].

지식 기반의 하향식 모델은 머리에 대한 일반적인 선행 지식을 바탕으로 일정한 모델 혹은 규칙을 정한 후, 이를 바탕으로 머리 검출을 수행한다. Yang 과 Huang은 3단계의 모자이크와 에지 검출을 바탕으로 머리 검출을 수행하였으며, Lanitis, Taylor 와 Cootes는 모양과 흑백 단계의 매개변수를 조절하여 머리 모양을 추출하는데 성공하였다[10].

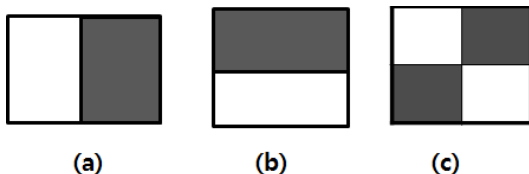
외형 기반 모델은 통계적이고 수학적인 방법으로 머리 패턴을 추출하고, 이를 바탕으로 머리 검출을 수행한다. 신경망(Neural Network), 고유 얼굴(Eigenface), SVM(Support Vector Machine), HMM(Hidden Markov Model)등을 이용한 머리 검출이 이에 해당된다[11]. 이들은 공통적으로 머리의 특정한 패턴을 학습 시키고, 새로 입력된 영상의 패턴과 학습된 통계적 특성을 비교하여 머리를 검출한다.

본 논문에서 제안하는 방법은 외형 기반 모델로써, 머리 패턴을 sub-region과 Haar Wavelet 변환을 이용하여 모델링하고 AdaBoosting 분류기를 이용하여 머리의 검출을 시도한다.

3. 제안하는 시스템의 설계 및 구현

제안하는 시스템은 온라인 부분과 오프라인 부분의 두 개의 모듈로 구성된다. 오프라인 부분은 입력영상의 전처리 단계와 Haar Wavelet 특징을 생성하는 단계와 특징을 이용한 학습단계로 구성된다.

전처리 단계에서는, 훈련 영상집합을 인테그럴 이미지로 변환하며 연산량을 줄이는 역할을 한다. 얼굴 영역을 추출하기 위해 피부색으로 얼굴 후보 영역을 추출한 후 특징 집합을 만들기 위해 Haar Wavelet특징을 사용하였다. 얼굴 특징은 Haar기저 함수를 사용해서 [그림 1]과 같이 사각형 특징으로 생성하였다. 전처리 단계를 거친 훈련 영상집합 중 얼굴 영상들의 특징에 대한 계수들을 구한다.



[그림 1] Haar 기저함수를 이용한 사각형 특징

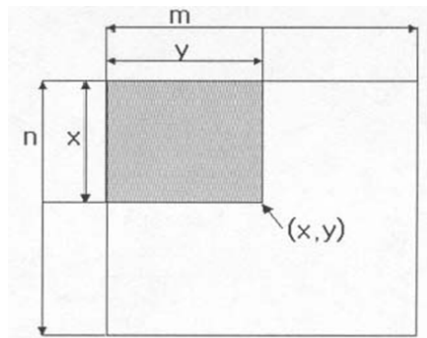
학습단계에서는 훈련영상집합(training data set)을 가지고 AdaBoosting 알고리즘을 이용해서 최종검출기를 생성한다. 훈련영상집합은 얼굴과 배경으로 구별되어진 입력영상집합이다. 본 논문에서는 AdaBoosting 학습 알고리즘을 효율적으로 적용하는 캐스케이드 구조에 기반한 학습시스템을 제안한다. 최종검출기는 설정한 스테이지 수만큼의 강한 분류기로 구성되며 캐스케이드 구조로 이루어진다.

온라인 부분은 입력 영상의 전처리 단계와 오프라인 부분에서 생성한 최종검출기를 이용해서 얼굴을 검출하는 과정이다. 최종검출기에서는 사각형 특징들을 스케일링(scaling), 쉬프팅(shifting)하며 입력 영상을 스캔하여 얼굴 영역을 검출한다.

훈련 영상집합을 인테그럴 이미지로 변환하게 되는데 픽셀마다 수행하는 방대한 연산량을 줄이기 위함이다. [그림 2]와 같이 영상의 좌표(x,y) 픽셀에 세로가 x, 가로가 y 인 사각형 내부의 모든 픽셀값을 더한 값이 저장된다.

$$I(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y') \quad (1)$$

여기서, $I(x,y)$ 는 인테그럴 이미지이며, $i(x',y')$ 는 입력영상이다. 이러한 인테그럴 이미지는 Haar Wavelet 특징의 계수를 계산할 때 연산량을 줄이는 역할을 한다.



[그림 2] 입력영상(m,n)에서 인테그럴 이미지(x,y) 픽셀의 값은 회색 사각형내부의 모든 픽셀의 그레이 값의 합

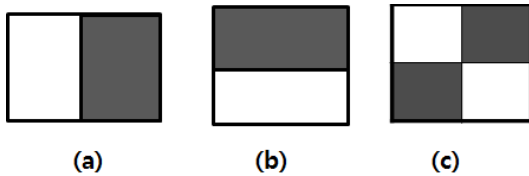
카메라로부터 입력되는 영상에서 얼굴은 유동적인 물체이다. 얼굴의 크기, 색깔, 모양 등의 변화가 심하며 조명의 영향을 많이 받는다. 제안하는 시스템에서는 얼굴의 특성에 강인한 검출을 위해서 얼굴 영상을 Haar Wavelet 특징으로 정의한다. Haar Wavelet 계수를 이용

한 Haar Wavelet 특징은 간단하므로 계산량을 감소시키며 여러 잡음이 있는 환경에서도 얼굴 검출에 있어서 강한 면을 보인다.

$$C(\text{scale, position}) = \int_{-\infty}^{\infty} f(t)\Psi(\text{scale, position, } t)$$

where, C : wavelet coefficient
 $\Psi(t)$:Harr wavelet basis function

입력 영상내에서 얼굴은 크기가 유동적이며 잡음이 존재한다. 본 논문에서는 아래와 같이 정의된 사각형 특징의 계수를 아래의 식과 같이 정의하여 값을 구한다.



$$C = \frac{\sum_m^b}{m} - \frac{\sum_n^a}{n}$$

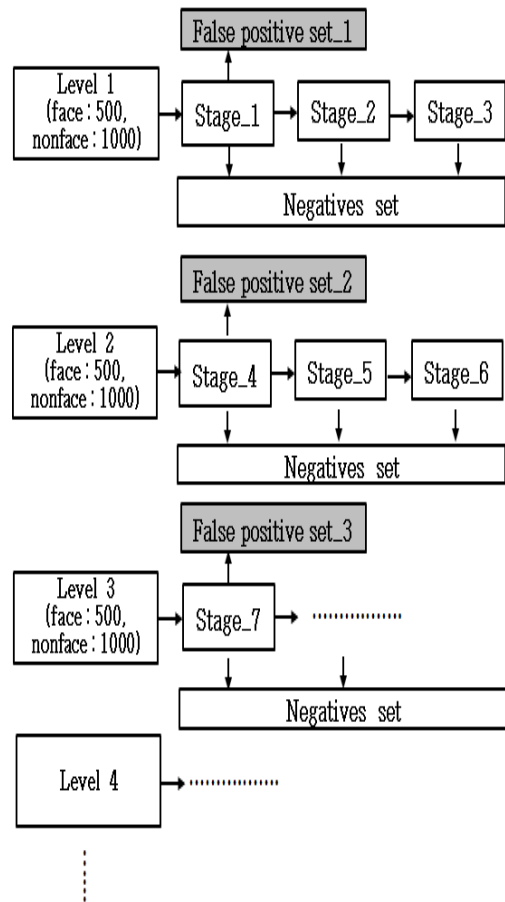
where, b : grey rectangle's pixel
 c : white rectangle's pixel
 m : the number of grey rectangle's pixels
 n : the number of white rectangle's pixels

제안하는 방법에서는 얼굴 영상의 특징을 Harr-wavelet을 이용하여 구한 후 그 특징을 AdaBoosting 알고리즘으로 학습한 후 결과를 이용하였다. 시스템이 얼굴 영역에서 하나의 특징만으로는 높은 검출율과 낮은 에러율을 가지는 분류기를 만들기 어렵다. 제안하는 방법은 AdaBoosting 알고리즘을 적용하여 여러 개의 특징을 선형적으로 결합해서 효율적인 강한 분류기를 만들 수 있다. AdaBoosting 알고리즘은 특징을 선택할 때 매우 많은 특징들로부터 적은 수의 특징을 과감하게 제거하여 찾아낸다.

기존의 AdaBoosting 알고리즘은 학습시간이 오래 걸리며 학습데이터가 변하게 되면 결과에 변화가 오기 때

문에 다시 처음부터 학습을 다시해야 하는 단점이 있다. 학습 얼굴 영상 중에서 영상의 질이 떨어지는 데이터가 있을 경우 이를 고려하지 않기 위해서는 재학습을 해야만 한다. 영상의 질이 떨어지는 데이터란 눈과 같은 얼굴의 중요한 특징이 손실된 데이터를 의미한다.

제안하는 방법에서는 [그림 3]과 같이 캐스케이드를 이용한 AdaBoosting의 효율적인 학습방법을 이용한다. 이 방법에서는 학습단계를 i 개의 레벨로 분리한다. 각 레벨은 캐스케이드 구조로 되어있기 때문에 많은 수의 학습데이터를 제거할 수 있다. 또한 다음 단계를 학습할 때 연산량을 대폭 감소시키는 효과가 있다. 각 단계의 학습과정에서 오검출된 영상들은 얼굴과 비슷한 패턴을 갖는 영상들이다. 따라서, 오검출된 영상을 반복적으로 고려해서 좀 더 강한 검출 성능을 가지는 특징들을 선택할 수 있으며 훈련 영상 집합의 크기도 작아진다.



[그림 3] 학습시스템 흐름도

각 단계의 학습과정은 순차적으로 진행이 되지만 오검출 영상집합을 두어서 재학습을 할 경우 독립적으로 학습시킨다. i 번째 레벨의 학습 영상을 변화할 경우 i 번째 레벨의 학습만을 다시하게 하면 된다. i 번째 레벨의 입력은 $i - 1$ 번째의 오검출된 데이터와 i 번째 배경 영상 집합과 i 번째 얼굴 영상 집합이다.

AdaBoosting 알고리즘을 이용해서 첫 번째 단계는 [그림 4]와 같이 세 개의 특징을 가지는 강한 분류기를 제작한다. 첫 번째 특징은 얼굴의 콧대와 왼쪽 눈 사이의 밝기 차이이고 두 번째 특징은 콧대와 오른쪽 눈 사이의 밝기 차이이다. 세 번째 특징은 오른쪽 눈과 눈썹 이마와 관련된 특징이다.



[그림 4] AdaBoosting 알고리즘으로 선택된 특징

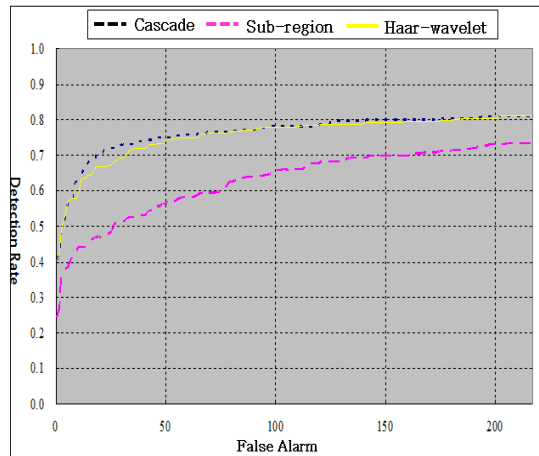
초기의 단계에서는 많은 수의 배경 영상을 제거하고 그 이후의 단계에서는 이전의 단계에서 제거되지 않은 배경 영상을 제거하여 고려할 영상의 개수를 낮춘다. 초기에 배경 영상들을 많이 제거함으로써 연산량을 대폭 감소시킨다.

4. 실험 결과

훈련 영상 집합은 국제 표준 FERET [11]과 카메기 멜론 대학(C.M.U)의 얼굴 데이터 집합과 인터넷을 이용하여 수집한 얼굴 데이터를 이용하였다. 사용한 영상의 크기는 $20 * 20$ 픽셀이며 얼굴 영상은 정면 영상만을 고려하였다. 얼굴 영상 4000장과 배경 영상 약 10000장을 사용하여 학습하였다. 특징 집합은 23,328 개의 사각형 특징으로 구성하였다. 각 특징들은 $20 * 20$ 영상 내에 정방향 사각형 특징부터 비대칭 형태로도 존재하여 다양한 위치에 존재한다. 각 특징들은 훈련 영상내의 계수값을 구하게 되며 계수값들은 AdaBoosting 알고리즘의 입력이 된다. 레이블링 된 모든 훈련 데이터에 대해서 에러값을 구하게 되며 결과는 여러 특징들의 조합이 되며 이는 강한 분류기를 의미한다.

본 논문에서 제안하는 학습시스템으로 각 단계를 생

성하게 되면 각 단계는 하나의 강한 분류기로 구성된다. 최종 검출기는 200여개의 특징과 30개의 단계로 구성된다. 현재의 dual core 3.30 GHz CPU 컴퓨터로 모든 레벨을 학습하는데 약 90분 정도가 소요된다. 카메라로부터 입력되는 영상이 $320 * 240$ 일 경우에 얼굴을 검출할 경우 0.5 ~ 0.7초가 걸린다. 이와 같은 결과를 바탕으로 제안하는 머리 추적 시스템은 실시간 시스템에 적용이 가능하다. 입력 영상내의 얼굴 영상의 크기는 매우 유동적이므로 스캔하는 과정에서 스케일링 작업이 필요하다. 본 실험에서는 187개의 특징으로 구성된 최종 검출기를 스케일링하며 얼굴을 검출하고 추적하게 된다. 초기의 스케일은 1.0 이며 1.2, 1.4로 스케일을 증가시키며 구현하였다. [그림 5]에서 볼 수 있는 것과 같이 각각의 분류기를 사용하는 것보다 두 개의 분류기인 Haar Wavelet과 캐스케이드를 이용한 AdaBoosting 알고리즘을 적용한 분류기를 조합하여 이용하였을 때 더 좋은 결과를 나타내었다. [그림 6]은 카메라에서 실시간으로 입력되는 영상에서 머리 영역을 검출하고 추적하는 장면을 보여주고 있다.



[그림 5] 각 레벨에 따른 머리 영역 검출률



[그림 6] 머리 영역에 대한 추적

5. 결론 및 향후 연구 방향

얼굴 검출 및 추적하는 시스템은 생체인식시스템, 대화형 게임기, 화상회의 등의 핵심기술이 되었다. 본 논문에서 제안하는 학습시스템으로 학습을 한 결과, 많은 학습시간을 줄일 수 있었으며, 얼굴검출 및 추적 성능 또한 우수한 결과를 보여주었다. 제안하는 시스템에서는 정면 얼굴만을 검출하므로 앞으로 개선의 여지가 많다. 실제적으로 다양한 조명환경과 복잡한 배경 속에서 얼굴의 특징을 찾아내는 것은 매우 어려운 일이다. 왜냐하면, 임의의 영상 내에서 얼굴영상은 다양한 자세와 크기를 가지는 매우 유동적인 물체이기 때문이다.

본 실험에서는 많은 훈련데이터를 이용해서 학습을 하였으며 AdaBoosting 알고리즘은 정확한 얼굴검출 및 추적을 가능하게 하였다.

참 고 문 헌

[1] M. Skurichina and R. P. W. Duin, "Bagging, boosting and the random subspace method for linear classifiers," *Pattern Anal. Appl.*, vol. 5, no.2, pp. 121 - 135, Jun. 2002..

[2] Michael J. Jones and Paul Viola. "Face Recognition Using Boosted Local Features". MERL Technical Reports. TR 2003-25, April 2003.

[3] C. W. Ng and S. Ranganath, "Real-time Gesture Recognition System and Application," *Image and Vision Computing*, Vol. 20, Issues 13-14, pp. 993-1007, 2002.

[4] H. Schneiderman and T. Kanade, "Object Detection Using the Statistics of Parts," *Int'l J. Computer Vision*, Vol. 56, No. 3, pp. 151-177, 2004.

[5] L. I. Kuncheva, J. C. Bezdek, and R. P. W. Duin, Decision templates for multiple classifier fusion, An experimental comparison, *Pattern Recognition.*, Vol. 34, No. 2, (2001) 299~314

[6] P. J. Philips, P. J. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M. Bone, *Face Recognition Vendor Test 2002: Evaluation report*, Technical Report, NISTIR 6965, National Institute of Standards and Technology, 2003, <http://www.frvt.org>.

[7] R. Gottumukkal and K. V. Asari, "A Novel Face Recognition Algorithm Based on PCA," *CVPRIP'03, Proc. JCIS 2003*, Cary, NC, USA, pp. 744-747, September 26 - 30, 2003.

[8] R. Duda, P. Hart, and D. Stork, "Pattern Classification," Second Edition, John Wiley & Sons Publications, New York, 2001.

[9] M. - H. Yang, N. Ahuja, and D. Kriegman, "Mixtures of linear subspaces for face detection," in *Proceedings of the Fourth International Conference Automatic Face and Gesture Recognition*, pp. 70 - 76, 2000.

[10] P. Phillips, "The FERET Database and Evolution Procedure for Object Recognition Algorithms," *Image and Vision Computing*, Vol. 16, No. 5, pp. 295-306, 1999.

[11] S.Z.Li, ZhenQiu Zhang. "FloatBoost Learning and Statistical Face Detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Accepted, 2004.

강 성 관



- 2001년 2월 : 인하대학교 컴퓨터공학부(학사)
- 2005년 8월 : 인하대학교 정보통신공학과(석사)
- 2007년 9월~현재 : 인하대학교 컴퓨터정보공학부(박사과정)
- 관심분야 : 컴퓨터 비전, HCI
- E-Mail : kskk1111@empas.com

이 정 현



- 1977년 2월 : 인하대학교 전자공학과(학사)
- 1980년 2월 : 인하대학교 전자공학과(석사)
- 1988년 2월 : 인하대학교 전자공학과(박사)
- 1979년 3월~1981년 12월 : 한국전자기술연구소 시스템 연구원
- 1984년 3월~1989년 2월 : 경기대학교 전자계산학과 교수
- 1989년 3월~현재 : 인하대학교 컴퓨터공학부 교수
- 관심분야 : IT융합기술, 서버일련스, HCI, USN
- E-Mail : jhlee@inha.ac.kr