

논문 2008-45SP-6-5

피부색 모델 기반의 효과적인 얼굴 검출 연구

(Efficient Face Detection based on Skin Color Model)

백 영 현*

(Young Hyun Baek)

요 약

피부색 정보는 컬러영상에 포함된 얼굴영역을 검출하는 중요한 요소이다. 피부색 정보로부터 생성된 통계 피부색 모델을 이용하여 얼굴영역을 검출할 수 있다. 하지만 다른 피부색 부분이 포함되어 있는 컬러영상에서는 일반적인 통계 피부색 모델만으로 정확한 얼굴영역 검출을 할 수 없는 단점을 가진다. 본 논문에서는 다른 피부색 부분이 포함되어 있는 다양한 컬러 영상에서 얼굴영역만을 정확히 검출하기 위한 방법을 제안한다. 제안된 방법은 YCbCr 피부 컬러 모델기반의 피부색 가우시안 분포를 적용하여 얼굴 후보영역 설정 하였고, 영상내의 잡음 부분과 얼굴 영역이외의 부분을 제거하기 위해 수학적 형태학을 적용하였다. 그리고 Haar-like 특성을 이용하여 정확한 얼굴 검출을 수행하였다. 모의실험 결과 제안된 방법이 목이나 팔과 같이 유사한 피부색을 포함한 영상과 다양한 크기의 영상에서도 효과적인 얼굴영역 검출하는 우수함을 보였다.

Abstract

Skin color information is an important feature for face region detection in color images. This can detect face region using statistical skin color model who is created from skin color information. However, due to the including of different race of people's skin color points, this general statistical model is not accurate enough to detect each specific image as we expected. This paper proposes method to detect correctly face region in various color image that other complexion part is included. In this method set face candidate region applying complexion Gaussian distribution based on YCbCr skin color model and applied mathematical morphology to remove noise part and part except face region in color image. And achieved correct face region detection because using Haar-like feature. This approach is capable to distinguish face region from extremely similar skin colors, such as neck skin color or arm skin color. Experimental results show that our method can effectively improve face detection results.

Keywords: skin color information, color model, face detection, mathematical morphology, Haar-like feature

I. 서 론

얼굴영역검출 방법은 자동화된 얼굴인식 및 표정인식을 위한 전처리 기술로서 연구가 시작되었으나 문제의 어려움과 중요성으로 인해 최근에는 하나의 독립적인 분야로 연구가 진행되고 있다. 현재 연구되고 있는 얼굴영역 검출 방법으로는 지식 기반법, 템플릿 정합법, 형태 기반법 등이 있다^[1-3]. 대표적인 얼굴영역 검출 방법으로 DiVAN(Distributed audioVisual Archives Network)이 있다. 이 방법은 제어 시스템, 내용기반 색

인 비디오 검색 시스템 접근 등 많은 응용분야에 매우 중요한 작업이 되었다^[4-6]. 하지만 피부색과 비슷한 영역인 목과 팔 피부 영역과 같은 부분이 포함된 영상에서는 정확히 얼굴영역을 검출할 수 없고, 영상크기에 따라 조정이 필요하다는 단점을 가지고 있다.

본 논문에서는 이러한 단점을 해결하기 위해 첫 번째 단계에서는 YCbCr 피부색 정보 기반의 가우시안 분포 피부색 모델과 수학적 형태학을 이용한 최적화된 검출 영역을 설정 방법을 제시하고, 두 번째 단계에서는 눈, 코, 입과 같은 얼굴 특징을 중심으로 얼굴영역과 얼굴이 아닌 영역을 구분하여 효과적인 얼굴영역 검출 방법을 제안한다.

본 논문의 구성은 II장에서 피부색 모델과 수학적 형

* 학생회원, 원광대학교 전기전자 및 정보공학부
(Department of Electrical Electronic and
Information Eng., Wonkwang University)

접수일자: 2008년2월9일, 수정완료일: 2008년10월10일

태학을 이용한 클러스터링과 III장에서는 얼굴 특징과 고유얼굴 거리매칭에 대해서 고찰하고, 제안하고자하는 얼굴 검출 방법에 대해서 살펴본다. IV장에서는 모의실험 및 검증에 위해 다른 피부색 영역이 포함되어 있는 영상과 각각 크기가 다른 영상을 이용하여 제안한 방법과 기존방법을 비교하여 분석한다. 마지막으로, 본 논문의 결론 순으로 작성한다.

II. 피부색 모델과 클러스터링

1. 피부색 모델

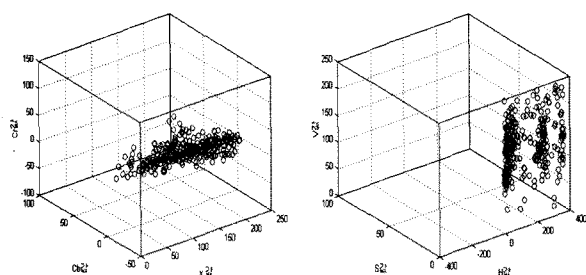
인간의 피부색은 컬러 영상에서 얼굴영역을 검출하는 중요한 요소이다. 피부색 정보를 이용한 피부색 영역 분할은 다양한 동작이나 이동, 복잡한 배경에서 얼굴영역만을 검색하기 위한 공간을 효과적으로 줄여줄 수 있다. YCbCr 컬러 모델은 MPEG과 JPEG 코딩에 관련된 모델에 주로 사용되고, HSV 컬러 모델은 컬러가 섞여있는 그림, 컴퓨터 그래픽 등에 사용된다. 본 논문에서는 MPEG과 JPEG 상에서 얻은 테스트 영상에 적용하기 때문에 HSV 컬러 모델을 생략하였다^[7~8].

그림 1과 같이 컬러 공간 3차원 분포도를 이용하여 YCbCr 컬러 모델의 컬러 분포보다 HSV 컬러 모델의 컬러 분포가 더 넓음을 알 수 있다. 즉 조밀도가 YCbCr 컬러 모델이 더 우수함을 알 수 있다.

본 논문에서는 RGB 영상으로부터 CbCr의 유사도에 따른 피부색 영역을 분할하기 위해 식 (1)과 같이 정의한다.

$$LSA = \exp[-0.5(x - C_m)^T C_c^{-1}(x - m)] \quad (1)$$

여기서, LSA (Likelihood Skin Area) 비슷한 피부색 영역을 나타내고, $x = (Cr, Cb)^T$ 이다. 그림 1의 피부색



(a) YCbCr color model

(b) HSV color model

그림 1. 컬러 공간 3차원 분포도

Fig. 1. Color space 3-dimension distribution chart.

정보를 이용하여 피부 색 영역과 피부색이 아닌 영역으로 나누기 위해서 Cb, Cr의 평균값 C_m 과 Cb와 Cr의 공분산의 결과 값 C_c 를 이용하여 가우시안 분포를 구한다. 피부색 가우시안 분포 적용한 피부색 필터는 식 (2)와 같이 나타낼 수 있다.

$$R_{Cb} = \{C_b \mid 77 \leq C_b \leq 127\},$$

$$R_{Cr} = \{C_r \mid 133 \leq C_r \leq 173\} \quad (2)$$

$$S(x,y) = \begin{cases} 1, & \text{if } [C_b(x,y) \in R_{Cb}] \cap [C_r(x,y) \in R_{Cr}] \\ 0, & \text{otherwise} \end{cases}$$

여기서, R_{Cb} 와 R_{Cr} 는 각각 77과 127 사이의 Cb값과 133과 173 사이의 Cr값을 의미하며, $S(x,y)$ 는 R_{Cb} 와 R_{Cr} 에 속하는 영역만을 얼굴후보영역으로 결정하는 피부색 필터이다.

2. 수학적 형태학을 이용한 클러스터링

수학적 형태학은 영상의 형태학적인 면을 조작하는 것으로 경계, 골격, 블록과 같은 영역 형태를 표현하거나 서술하는데 있어서 유용한 영상 요소들을 추출하기 위한 도구로 수학적 형태론의 개념을 이용하여 영상을 단순화하여 클러스터링 개념을 부가하고자 한다. 이 방법은 영상 전처리 작업이나 초기 객체 분류 또는 이러한 처리 뒤에 이어서 물체의 내재된 구조를 명확히 하는데 이용된다. 이는 물체의 외각선을 물체의 가장 현저한 1-픽셀 (Pixel) 두께의 외각선이나 골격 선을 간단하게 함으로써 구성된다. 이진 형태학의 기본 연산은 Dilation과 Erosion으로 구성된다. 그레이 스케일 형태학은 이진 형태학을 확장한 집합의 Top Surface 개념과 Surface의 umbra 개념을 도입한 것이다^[9~11]. 그레이 스케일 Erosion은 최소 연산과 감산의 집합연산으로 계산할 수 있다. 따라서 설계자가 적절한 구조적 원소와 연산을 선택하여 데이터의 필요 유·무를 공간 영역에서 적절하게 활용할 수 있어 의미 전달이 용이하다.

그레이 스케일 영상을 c_j 라 하고, 영상의 처리 공간 E^N 의 함수를 k 라 정의하면, x 는 공간 E^N 에서의 한 점일 때, 두 함수의 Dilation은 두 함수의 umbra의 dilation에 의한 surface로 정의되며 k 에 의한 c_j 의 Dilation은 식 (3)과 같다.

$$c_j \oplus k = \mathcal{T} \{ U[c_j] \oplus U[k] \} \quad (3)$$

Dilation을 합집합에 의한 최대값을 얻을 수 있도록

변환하면 $c_j \oplus k$ 는 식 (4)와 같이 계산할 수 있다. 그리고 $x \in D$, $K = [-v, \dots, 0, v]$ 이다.

$$(c_j \oplus k)(x) = \max_{\substack{z \in K \\ x-z \in C}} \{c_j(x-z) + k(z)\} \quad (4)$$

$c_j(x)$ 가 샘플링 함수 $c_j(i)$, $i \in Z$ 이고, K 는 구조적 집합이면 Dilation 필터의 출력은 식 (5)가 된다.

$$y_i = [c_j \oplus K](i) = \max \{c_j(i-v), \dots, c_j(i), \dots, c_j(i+v)\} \quad (5)$$

구조적 원소 구조에 의해 정의된 이웃에서 c_j+k 의 최대값을 선택하는 결과이며, 그레이 스케일 영상에 대해서 Dilation을 수행하였을 때 일반적인 결과는 2가지 특성을 가지는데, 첫째로 구조적 원소의 값이 모두 양수이면 입력 영상보다 밝아지는 경향이 있으며, 둘째로 어두워지는 것은 구조적 원소의 값이 감소되거나 제거되는 것으로 설명할 수 있다. 따라서 그레이 스케일 영상의 값과 모양은 Dilation에 이용되는 구조적 원소에 관계된다. Erosion에 대한 정의는 Dilation의 정의와 같은 방법으로 정의되며, Erosion은 $c_j \ominus k$ 로 표시되고 식 (6)과 같이 정의된다.

$$c_j \ominus k = T[U[c_j] \ominus U[k]] \quad (6)$$

Erosion은 차집합의 최소값으로 얻을 수 있도록 변환하면 $c_j \ominus k$ 는 식 (7)과 같이 계산할 수 있다.

$$(c_j \ominus k)(x) = \min_{\substack{z \in K \\ x+z \in C}} \{c_j(x+z) - k(z)\} \quad (7)$$

$c_j(x)$ 가 샘플링 함수 $c_j(i)$, $i \in Z$ 이고, K 는 구조적 집합이면 Erosion 필터의 출력은 식 (8)과 같다.

$$y_i = [c_j \ominus K](i) = \min \{c_j(i-v), \dots, c_j(i), \dots, c_j(i+v)\} \quad (8)$$

Erosion은 구조적 원소 k 의 이동이라기보다는 c_j 의 이동으로 이해할 수 있으며, Erosion은 구조적 원소에 의해 정의되는 이웃에서 c_j-k 의 최소값을 선택한다. 그레이 스케일 영상에 대해 Erosion을 수행하였을 때 일반적인 결과는 구조적 원소의 모든 원소들이 양수이면 출력영상은 입력영상에 비해 어두워지며, 또한 영상의 밝은 부분은 형태소의 모양과 값의 크기에 따라 줄어들거나 사라지게 된다. 또한 Opening과 Closing은 Dilation과



(a) 원영상



(b) RCb ∩ RCr 결과영상



(c) 휘도 정규화



(d) 형태학 적용 결과영상

그림 2. 형태학 연산자를 적용한 결과 영상

Fig. 2. Result images applying morphology operators.

Erosion을 연속적으로 적용함으로써 정의되며, 만약 구조적 원소 g 에 의한 c_j 의 Opening c_j^g 는 식 (9)와 같다.

$$c_j^g(x) = [(c_j \ominus g^s) \oplus g](x) = [c_j(x) \ominus g(-x)] \oplus g(x) \quad (9)$$

또한 구조적 원소 g 에 의한 c_j 의 Closing c_j^g 는 식 (10)과 같다.

$$c_j^g(x) = [(c_j \oplus g^s) \ominus g](x) = [c_j(x) \oplus g(-x)] \ominus g(x) \quad (10)$$

본 논문에서는 그림 2와 같이 검출된 피부색 영역에서 해상도 감소를 이용하여 처리 데이터양을 줄이고, 휘도 정규화와 형태학 연산자를 이용하여 얼굴 검출영역을 설정 하였다.

III. 얼굴특징을 이용한 얼굴영역 검출 방법

인간 얼굴특징을 잘 나타내는 매우 효율적인 방법으로는 1차원 Harr-like특징을 이용하는 방법이 있다. 1차원 Haar-like특징은 인접한 직사각형 영역들의 화소값들의 합의 차로 본 얼굴 검출 시스템에서 적용하고 있는 Haar-like특징들은 네 개로서 직사각형 두 개로 이루어진 특징과 세 개로 이루어진 특징들 모두 같은 모양, 같은 면적을 가진 수평, 수직으로 인접한 직사각형으로 구성되어있다^[12~13]. Haar-like특징을 이용하여 그레이 레벨의 영상 $A_{x,y}$ 에서 수평과 수직 성분 x, y

를 이동하며 수평과 수직성분의 영상을 추출은 식 (11) 과 같다.

$$G_x = abs(dx * A) = abs([1, -1] * A)$$

$$G_y = abs(dy * A) = abs\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix} * A\right) \quad (11)$$

$$G_{max}(x', y') = \max(G_x(x', y'), G_y(x', y'))$$

G_x 는 수직 성분인 코 부분 검출과 G_y 는 눈과 입술 부분인 수평 성분을 나타내고, $G_{max}(x', y')$ 은 G_x 와 G_y 의 최대값의 합성 영상이다. 식 (11)에 의해 추출된 영상은 그림 3과 같이 PCA를 이용한 고유 얼굴 영상과

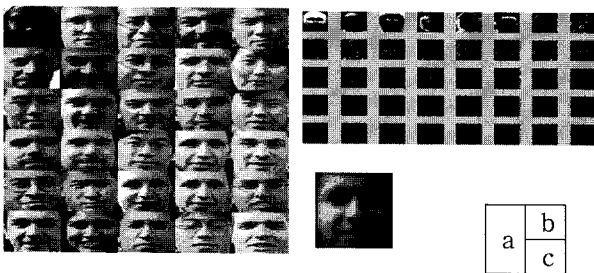


그림 3. PCA를 이용한 40개의 고유얼굴 영상 : (a) 학습 영상, (b) 40개의 고유 얼굴 집합, (c) 고유얼굴
Fig. 3. Result of a 40 units EigenFace image using PCA.

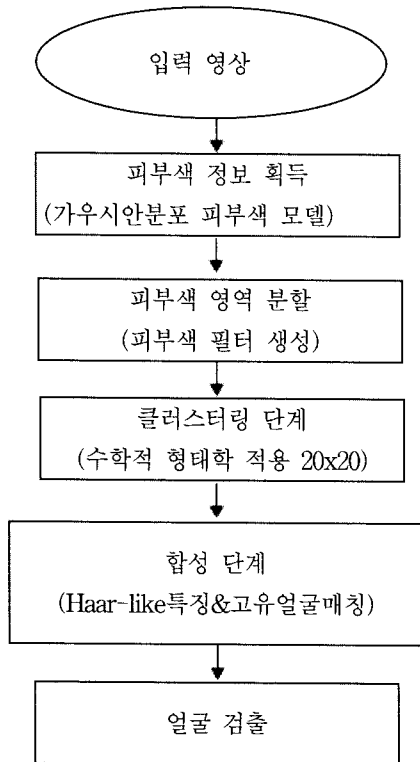


그림 4. 제안된 얼굴 검출 방법
Fig. 4. The proposed face detection method.

의 수평, 수직 거리매칭을 수행한다.

식 (12)와 같이 맨하탄 거리 척도 개념을 적용해 특징벡터 x_i 와 입력 영상의 특징벡터 y_i 사이의 거리를 비교하여 얼굴 영역 검출을 수행한다.

$$MD: D = \sum_{i=1}^n |x_i - y_i| \quad (12)$$

본 논문에서는 효과적인 얼굴영역 검출 방법을 구현하기 위해 기존의 단일 구조 피부색 모델이 아닌 피부색 가우시안 분포를 이용한 방법을 제안하였다. 또한 수학적 형태학 연산을 적용하여 잡음이나 불필요한 영역을 제거하고, Haar-like특징을 이용하여 유사 피부색 영역 중에서 얼굴특징을 포함한 영역만을 검출하는 방법을 제안하였다. 제안하는 얼굴영역 검출 방법의 전체 구조 4단계인 피부색 분포에 따른 가우시안 모델, 컬러 분할, 클러스터링, 합성 그리고 얼굴 검출 순으로 구성되었다. 얼굴 검출 시스템의 구조는 그림 4와 같다.

IV. 실험 및 고찰

제안된 알고리즘의 성능 평가를 위해 MIT-CMU face database, Caltech face database, Essex UK database, Georgia Tech face database를 사용하였다. 고유 얼굴 템플릿을 만들기 위해 Bio-ID 데이터베이스 1,200개의 정면 얼굴의 샘플들을 검출하였고, 다양한 MPEG 비디오와 일반 증명사진, 인터넷 사진 등을 테스트 영상으로 사용하였다. 모의실험은 Matlab 7.0을 이용하였다. 그림 5는 다양한 크기의 영상을 통하여 기존의 DiVAN 시스템과 제안하는 얼굴 검출 시스템을 비교한 결과 영상들이다.

그림 5는 다양한 크기의 유사 피부색 영역이 포함된 영상에 DiVAN 검출방법과 제안한 검출방법을 적용한 결과 영상이다. 실험 결과 제안된 검출 방법이 다양한 크기와 피부색 영역 검출 시에도 얼굴의 구조적 특징인 눈, 코, 입을 포함한 부분을 더 효과적이고 정확하게 검출함을 확인하였다.

그리고 100개 테스트 영상 결과는 표 1과 같으며, 실험 결과 제안된 얼굴 검출 방법을 적용한 검출율은 94%이고, DiVAN 방법은 86%의 결과 나타내었다. 따라서 제안된 얼굴 검출 방식이 DiVAN시스템보다 8% 정도 우수함을 확인하였다.

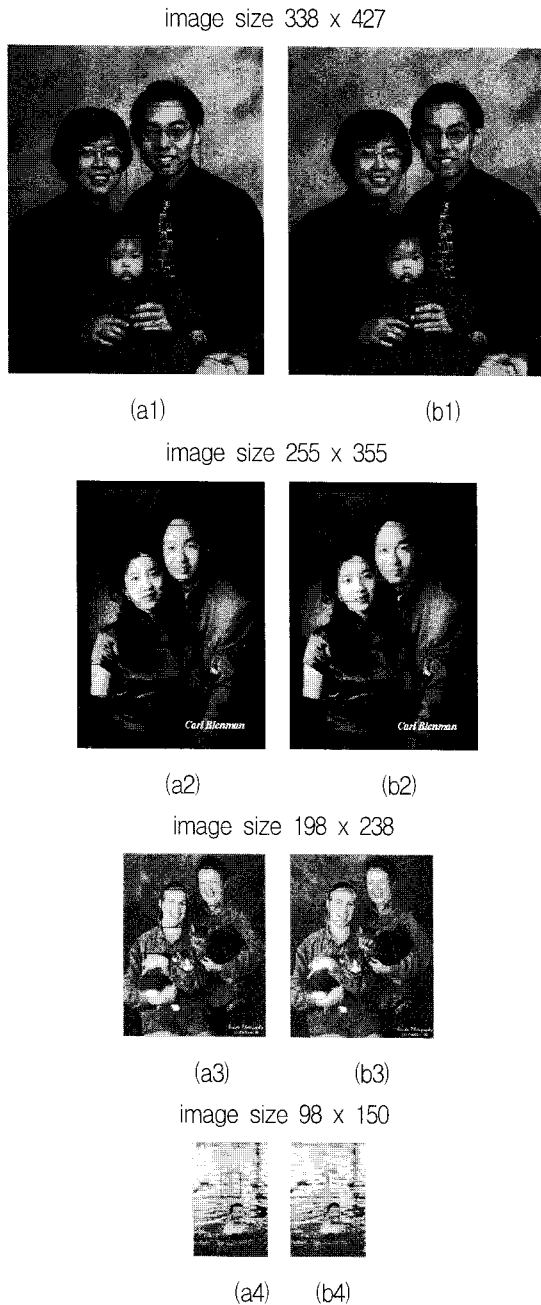


그림 5. 테스트 영상에 기존 알고리즘과 제안된 알고리즘을 적용한 결과 영상: (a1-4) Divan 방법, (b1-4) 제안된 방법

Fig. 5. Result image applying the conventional algorithm and the proposal algorithm in test image.

표 1. 얼굴 검출율
Table 1. Face Detection rate.

	DiVAN 방법	제안된 방법
얼굴 검출율	86개(86%)	94개(94%)
오 검출율	8개(8%)	4개(4%)
미 검출율	6개(6%)	2개(2%)

V. 결 론

본 논문에서는 피부색 모델과 1차원 Haar-like 특성을 이용한 효과적인 얼굴 영역 검출방법을 제안하였다. 제안된 검출 방법은 피부색 분포를 이용한 가우시안 컬러 모델 생성과 유사 피부색 영역을 분할하기 위한 피부색 필터를 적용하였다. 그 결과 얼굴 검출 영역 크기를 감소하는 장점을 보였으며, 수학적 형태학 연산을 적용함에 따라 불필요한 영역 제거 및 피부색 영역을 단순화하였다. 마지막으로 검출된 피부색 영역에서 얼굴 구성요소인 눈, 코, 입을 포함한 얼굴영역을 정확히 검출하기 위해 1차원 Haar-like 특징과 고유얼굴 특징벡터간의 거리척도 개념을 적용한 얼굴영역 검출 방법을 제안하였다. 모의실험 결과 유사 피부색이 포함된 다양한 크기의 영상에서 제안된 검출 방법이 얼굴 구성요소를 중심으로 기존의 DiVAN 시스템보다 목(目)시적인 정확도와 검출율면에서 8%정도 더 우수한 성능을 보임을 알 수 있었다.

참 고 문 헌

- [1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," Proc. IEEE, Vol. 83, No. 5, pp705-740, 1995.
- [2] A. J. Colmenarez and T. S. Huang, "Face Detection with Information-Based Maximum Discrimination," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.782-787, 1997.
- [3] A. Rajagopalan, K. Kumar, J. Karlekar, R. Manivasakan, M. Patil, U. Desai, P. Poonacha, and S. Chaudhuri, "Finding Faces in Photographs," Proc. Sixth IEEE Int'l Conf. Computer Vision, pp.640-645, 1998.
- [4] C. Garcia, G. Zikos, and G. Tziritas, "Face Detection in Color Images using Wavelet Packet Analysis," Proc. IEEE Intern. Conf. Multimedia Computing and Systems, Florence, vol. 5, pp. 703-708, June 1999.
- [5] C. Garcia and G. Tziritas, "Face Detection Using Quantized Skin Color Regions Merging and Wavelet packet Analysis," IEEE Trans. on Multimedia, vol. 1, no. 3, pp. 264-277, September 1999.
- [6] DiVAN : Distributed audioVisual Archives Network(European Esprit Project EP 24956). <http://divan.intranet.gr/info>, 1997.

- [7] 백덕수, 변오성, 백영현, “공간컬러모델을 이용한 얼굴객체 검출 시스템연구,” 대한전자공학회논문지, 제43권 IE편, 제2호, 30-38쪽, 2006년 6월
- [8] Randy Crane, A simplified approach to Image processing, Prentice Hall, 1997.
- [9] S. R. Moon, “Design of Hybrid Median Filter Using Gray Scale Morphology,” Chonbuk University, Ph D., 1993.
- [10] Y. K. Yoon, “DTCNN Hardware Implementation and Application Using Morphology,” Wonkwang University, Master, 1998.
- [11] E. D. Dougherty, An Introduction to Morphological Image Processing, volume TT9.SPIE Press, 1992.
- [12] Viola P, Jones M, “Robust Real-time Object Detection”, International Journal of Computer Vision, vol.57(2), pp.137-154, 2004.
- [13] Viola P, Jones M, “Rapid Object Detection using a Boosted Cascade of Simple Feature”, IEEE CS Conf. on Computer Vision and Pattern Recognition, vol.1, pp.511-518, 2001.

— 저 자 소 개 —



백 영 현(학생회원)

2002년 원광대학교 전자공학과
학사 졸업.

2004년 원광대학교 전자공학과
석사 졸업.

2007년 원광대학교 전자공학과
박사 졸업.

2007년~현재 원광대학교 전기전자 및
정보공학부 전임강사

<주관심분야 : 영상신호처리, 얼굴인식, 영상 검
색, 해상도 복원, 객체 검출 및 인식 >