
공간 제약 특성과 WPA를 이용한 얼굴 영역 검출 및 검증 방법

송호근*

Face Region Detection and Verification using both WPA and Spatially Restricted Statistic

Ho-Keun Song*

요 약

본 논문에서는 컬러 정지 영상을 대상으로 상반신 인물 영상이 입력되었을 때, 얼굴 영역을 추출하고 검증하는 방법을 제안한다. 본 논문의 얼굴 추출과정은 1단계로 영상 내 피부색 영역을 추출한 다음, 후보 영역들에 대한 공간적 제한조건을 이용하여 1차 얼굴 후보 영역을 결정한다. 2단계에서는 얼굴 구성 요소 중 가장 두드러진 특징으로서 눈 영역을 탐색하고, 눈 영역을 기준으로 한국인의 얼굴에 대한 구조적 통계값을 적용한다. 이로서 얼굴 포함 최소 사각형 후보 영역을 결정한다. 마지막 3단계에서는 영상 내 색상 정보와 공간 정보 그리고 구조적 통계치로부터 결정된 얼굴 후보 영역에 대하여 얼굴 영역의 텍스처(texture)를 Wavelet Packet Analysis를 이용해 조사함으로써 얼굴 영역을 확정하게 된다.

ABSTRACT

In this paper, we propose a face region detection/verification method using wavelet packet analysis and structural statistic for frontal human color image. The method extracts skin color regions from input images, first. and then applies spatial restrictive conditions to the region, and determines whether the region is face candidate region or not. In second step, we find eye region in the face candidate region using structural statistic for standard korean faces. And in last step, the face region is verified via wavelet packet analysis if the face texture were satisfied to normal texture conditions.

키워드

Face Region Detection, Wavelet Packet Analysis, Color Facial Image, Structural Statistic

I. 서 론

멀티미디어의 급격한 발달로 영상 정보의 검색과 활용이 매우 중요해졌다. 특히 얼굴 영상 인식은 원거리 화상 회의에서의 얼굴 추적 시스템, 눈동자 추적을 통한 운전자의 시야 조절 시스템, 얼굴 인식을 통한 출입 통제 시스템, 헤드 제스처 인식을 이용한 컴퓨터 또는 로봇 제어 시스템 등 많은 분야에서 활용되고 있다.

한편 얼굴 영역 추출과 얼굴 특징 요소의 추출은 얼굴 영상 처리 분야에서 기본적이면서 매우 중요한 부분이다. 그러나 얼굴의 모양, 크기, 색상, 방향, 머리스타일, 복잡한 배경과 조명 등 다양한 요소들로 인하여 추출 방법이 쉽지 않다.[1][2][3]

얼굴 인식을 하기 위한 얼굴 영역 추출에 관한 연구는 90년초반부터 수행되었다. 이 중에서 대표적인 얼굴 영역 추출 방법은 동영상을 대상으로 움직임 부분을 추출하고 이를 바탕으로 얼굴의 경계를 추출하는 연구, 영상의 해상도를 줄여가면서 눈, 코, 입 부분의 명암 특징을 사용하여 얼굴 영역을 추출한 연구, 눈, 코, 입의 명암특징을 이용하여 얼굴 영역을 추출하는 시스템을 개발한 연구, 계층적 영상에서 후보 영역을 추출하고 삼각형 구조 모델을 이용한 얼굴 영역을 추출하는 연구 등이 있다.[4][5]

본 논문에서는 컬러 정지 영상을 대상으로 상반신 인물 영상이 입력되었을 때, 얼굴 영역을 추출하고 검증하는 방법을 제안한다. 본 논문의 얼굴 추출 과정은 1단계로 영상 내 피부색 영역을 추출하고, 후보 영역들에 대한 공간적 제한 조건을 이용하여 1차 얼굴 후보 영역을 결정한다. 2단계에서는 얼굴 구성 요소 중 가장 두드러진 특징인 눈 영역과 입 영역을 탐색하고, 눈과 입 영역을 기준으로 동양인 얼굴 크기에 관한 구조적 통계값을 적용한다. 이로서 얼굴 포함 최소 사각형 후보 영역을 결정한다. 마지막 3단계에서 영상 내 색상 정보와 공간 정보 그리고, 구조적 통계치로부터 결정된 얼굴 포함 최소 사각형에 대하여 얼굴 영역의 텍스처(texture)를 Wavelet Packet Analysis를 이용해 조사함으로써 얼굴 영역을 확정하게 된다. 일반적으로 2단계에서 대략적인 얼굴 영역이 결정되지만 3단계에서 얼굴 내 텍스처 정보를 활용하면 보다 적절한 얼굴 포함 사각형의 범위를 결정할 수 있다.(그림1)

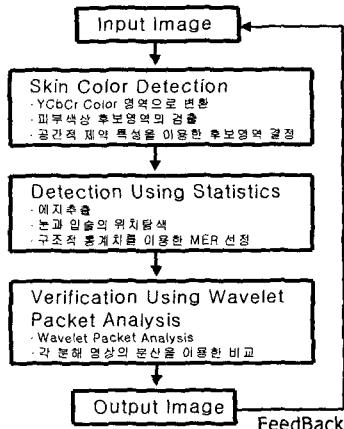


그림 1. 3단계 Face Detection
Fig. 1 Three step Face Detection

II. 피부색 영역 추출

2.1 YCbCr 색상계로 변환

본 논문에서는 YCbCr 좌표계를 이용하여 피부 색상 정보를 표현하였다. HSI 좌표계와 비교했을 때 YCbCr 좌표계의 경우 HSI 좌표계보다 피부색 분포가 한곳에 집중되어 나타나는 경향이 있다.[4] 이를 바탕으로 피부색이 집중된 영역을 선정, 얼굴 영역의 대략적 범위를 설정하였다.

2.2 피부색상 후보 영역의 검출

피부 색상을 추출하기 위해 입력 영상을 YCbCr 좌표계로 변환한 후 Cb와 Cr 성분만을 이용하여 피부 색상 후보 영역을 추출한다.

본 논문에서는 여러 입력 영상들에 대한 피부색을 조사하고, 아래 표와 같은 일반적인 피부색 분포 범위를 찾았습니다.

표 1. 피부색영역에서 Cr/Cb값의 변화

Table 1. Variation of Cr/Cb within skin color region

황백색	황갈색	적갈색	적색
0	16	32	48
진노랑	황갈색	연갈색	황백색
-64	-48	-32	-16

찾아낸 피부색상 영역 모두에서 피부 색상을 추출한다 면 피부색 이외의 많은 물체들이 추출될 것이다. 때문에 Cr값과 Cb값을 각각 다음과 같이 제한하였다.

$$10 < Cr < 50, \quad -40 < Cb < 0 \quad (1)$$

한편, 일반적인 실내조명(형광등)하에서의 피부의 명도값을 알아보았다. 명도는 0~255의 값으로 나타나었으며, 일반적인 실내조명 아래서 얼굴색의 명도 범위는 80~200 정도에 속했다. 그런데 빛에 의한 반사나 음영으로 밝거나 어두운 피부색도 고려해야 한다. 따라서 본 논문에서는 영상 내 피부색 영역의 평균 밝기를 기준으로 영상의 피부색상의 밝기에 따라 피부색상 범위값을 변화시켰다. 아래는 피부색 부분의 명도를 히스토그램화하여 나타낸 것이다.(그림 2)

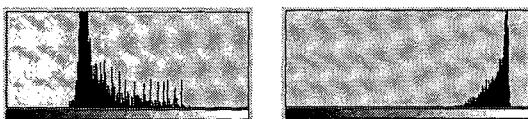


그림 2. 조명에 따른 피부색영역의 범위
Fig. 2 Range of the skin color according to the illumination

처음에 CbCr값을 이용하여 찾아낸 피부색상 영역 내에서 평균 명도값을 구한다. 이 값이 속하는 밝기범위의 피부색상 영역에 다른 영역보다 높은 가중치를 주었다. 때문에 똑같이 찾아진 피부색상 영역이라도 가중치가 높은 곳이 얼굴 피부색상일 확률이 높아진다.

2.3 공간제약특성을 이용한 얼굴후보 영역결정

피부 색상 후보 영상 내에는 얼굴 피부 색상과 유사하게 염색된 갈색 머리나 유사 색상의 배경 물체나 영역 등도 함께 검출된다.

따라서 피부 색상 후보 영역들 중 얼굴 후보 영역을 분리해 내기 위하여 다음과 같은 공간적 제약 특성을 사용한다.

먼저 후보 색상 영역의 면적을 조사한다. 본 논문에서 대상으로 하는 입력 영상은 증명 사진이나 상반신을 포함하는 인물 영상이다. 이때 얼굴 영역이 전체 영상에서 차지하는 최소 비율이 존재하므로 피부 색상 후보 영역들 중 일정 면적 비율(30%) 이하의 영역을 후보에서 탈락시키는 전략을 사용한다.

둘째, 피부 색상 후보 영역 중 얼굴 영역은 일반적으로 가로와 세로의 비율(Aspect Ratio)이 일정 범위 내에 존재한다는 조건을 활용한다. 동양인의 얼굴은 보통 1:1.4의 비율을 가진다. 그러나 개인적인 차를 수용하기 위하여 최대 1:1.2에서 1:2까지의 범위 영역 내의 후보 영역을 얼굴 후보로 선정한다. [4][5]

셋째, 피부 색상 후보 영역들에 대하여 최소 외접 사각형을 구하고, 사각형 내부에 포함되는 피부 색상 화소의 개수를 조사한다. 그리고나서 최소 외접 사각형 면적에 대한 피부 색상 화소 영역의 면적 비율을 계산한다.

이것은 정상적인 얼굴 후보 영역인 경우 사각형 내부에서 피부 색상 화소가 차지하는 비율이 일정 비율 이상일 것이라는 지식을 사용하는 것이다. 본 논문에서는 85% 이상일 때 얼굴 후보 영역으로 선정하였다.

그림 3에서는 피부 색상 후보 영역에 대한 공간적 제약 특성을 이용하여 얼굴 후보 영역을 결정한 예이다. 얼굴 영역 이외의 후보 영역들이 제거된 모습을 볼 수 있다.



그림 3. 피부색 영역 추출
Fig. 3 Extraction of skin color region

III. 구조적 통계값 특성을 이용한 얼굴 포함 최소 사각형 후보 선정

3.1 수평, 수직투영을 이용한 최대 군집영역 선정

공간적 제약 특성을 이용한 1차 얼굴 후보 영역은 얼굴

구성 요소들과 함께 귀 영역이나 목 영역등이 함께 포함될 수 있다. 따라서 구조적 통계값을 적용하기 전에 1차 얼굴 후보 영역 중에서 최대 군집 영역을 선정하기 위하여 피부색 영역을 2진화하고 수평, 수직 투영한 뒤 평균 투영값 이상의 영역을 선택하여 최대 군집 영역을 선정한다.[5]

3.2 눈과 입술의 위치 탐색

얼굴 구성 요소를 찾는 가장 중요한 단계는 정확한 눈의 위치를 찾는 것이다. 대부분의 상반신 정면 영상에서 눈의 위치는 뚜렷하게 나타난다. 그러나 얼굴의 기울어짐에 따라 기울기 보정을 할 필요가 있다.[6][7]

본 논문에서는 위 단계에서 선정된 얼굴 후보 영역 내의 소벨 에지값과 얼굴의 구조적 통계치를 적용하여 눈과 입의 위치를 탐색한다.

우선 영역 내에서 수직 소벨 에지 영상을 구한뒤, 수평 투영값을 구하고, 최대값을 탐색하여 눈의 y 좌표를 구한다. 구해진 눈 위치에서 눈동자 유효성 검증을 통하여 양쪽 눈의 x 좌표를 구하였다. 탐색된 눈의 위치로부터 얼굴의 구조적 통계치를 기반으로 입 후보 영역을 탐색한다. (그림4-3) 그 후에 두 눈의 거리와 눈과 입까지의 수직거리를 기준으로 다시 구조적 통계치를 적용하여 얼굴 구성 요소 포함 영역을 선정한다.(그림4-4)



그림 4. 얼굴 구성 요소
Fig. Structural components of Face

3.3 얼굴 구성 요소 포함 최소 범위 사각형 후보 선정

두 눈의 위치를 알아낸 후, 이 구성 요소들을 포함하는 최소 범위 사각형 후보를 선정한다. 이는 동양인 얼굴 표준의 구성 요소간의 거리에 대한 통계치를 이용하여 결정하게 된다.

본 논문에서는 이 통계적 수치값보다 작은값을 적용하였는데 이는 눈과 입을 포함하는 사각형 후보영역은 얼굴 전체에 대한 사각형 크기보다 작기 때문이다.

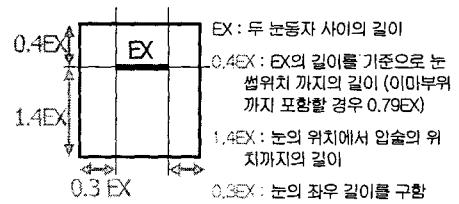


그림 5. 얼굴 후보영역과 관계 통계치들
Fig. 5 Candidate of rectangular face region and its statics

위 그림 5에서 보면 두 눈의 위치가 정해졌다면 두 눈 사이의 거리를 EX라 정하고, 이를 기준으로 통계값을 이용해 1.4EX의 거리 내에 입술이 포함되어 있다는 전제하에 얼굴의 주요 구성 요소를 포함하는 최소 범위 사각형 영역을 선정할 수 있다. 이 과정을 통하여 피부색상 영역으로 넓게 선정하였던 얼굴 후보 영역을 구성 요소를 포함한 정확한 최적 사각형 영역으로 확정할 수 있었다. 이는 통계적인 값을 이용함으로써 사람에 따라 개인차가 있는 얼굴 크기를 고려해 얼굴의 최소 범위를 결정할 수 있다.

IV. Wavelet Packet Analysis를 이용한 얼굴 영역 검증

탐색된 얼굴 후보 영역에 대하여 얼굴의 텍스처 정보를 조사하고 표준 얼굴 텍스처 정보와의 매칭 과정을 통하여 얼굴 후보 영역에 대한 검증을 수행하였다.

4.1 Wavelet Packet Analysis(W.P.A.)

웨이블릿 패킷 분석 방법(WPA)은 1단계는 일반 웨이블릿 변환과 동일하나, 2단계 이후에서 근사 영상뿐 아니라 상세 영상을 각각에 대해서도 동일한 변환을 수행하게 된다. 이로서 보다 풍부한 신호 분석을 제공하게 된다. 따라서 본 논문에서는 WPA를 사용하여 얼굴의 텍스처 정보를 표현하고, 이를 얼굴 영역 검증에 활용하였다. [3][8][9][10]

4.2 WPA의 각 분해영상의 분산을 이용한 비교

추출된 얼굴 영역을 2단계 WPA로서 1개의 근사 영상과 15개의 상세 영상들로 분해한 뒤, 식(2)를 이용하여 각 영상들에 대한 분산을 구한다.

$$\sum_{i=0}^n \sum_{j=0}^m \frac{|W(i,j) - \sqrt{t}|^2}{|255 - \sqrt{t}|^2} \quad (2)$$

여기서 $W(i,j)$ 는 WPA분해 영상의 평균값에 대해 화소 위치(i,j)에서의 화소값의 차이를 의미한다. 분산을 구한 후, 얼굴 영상 데이터베이스에서 미리 추출한 얼굴 영역에 대하여 평균 분산값들을 저장한 뒤 입력 영상으로부터 추출된 분산값들과의 차이를 조사한다. 이들 값들이 한계 값 이내일 때 얼굴 영역을 확정한다.

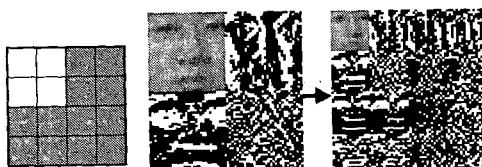


그림 6. 2단계 Wavelet 변환 예
Fig. 6. Two step wavelet transform

얼굴을 정확히 찾았을 경우와 잘못 찾았을 경우 그리고, 유사 얼굴의 경우의 분산값들을 비교해 보았다. 유사 얼굴이란 피부색 바탕에 눈과 입 부위만을 수작업으로 그린 그림의 경우를 말한다. 이는 실제 얼굴이 아닌 그림이나 만화영화 등의 경우에도 추출이 되는지 확인하기 위함이었다. 그림7의 분산값 중에서 색상 쌍을 이룬 네모칸이 쳐져 있는 것이다. 이 값들은 분산값을 비교하기 위하여 임의로 선택한 값으로 가장 크거나 작은 값은 아니다. 색상 쌍을 이룬 네모칸의 분산값을 서로 비교해 보면 얼

굴 영역이 아닌 경우가 얼굴 영역인 경우보다 높은 값들이 나옴을 알 수 있다. 일부 영역의 값은 얼굴 영역과 비 얼굴 영역의 값이 비슷한 경우도 있었다. 그러나, 비 얼굴 영역일 경우 얼굴 영역보다 분산값이 대체적으로 높게 나타나는 것을 확인할 수 있었다. 유사 얼굴의 경우 근사 영역의 값은 얼굴 영역과 유사하고, 나머지 영역은 조명이나 개인차를 반영하지 않기 때문에 비 얼굴 영역 분산값에 근접하였다. 이는 분산값을 이용하여 얼굴 영역을 잘 못 찾았는지 아닌지를 판가름하여, 재검사를 하는 기준으로 사용될 수 있음을 알 수 있었다.

	Wavelet 영역의 분산값			
0001	0.053714	0.002774	0.000888	0.003999
	0.006513	0.002438	0.002403	0.003280
	0.000873	0.003356	0.002564	0.002138
	0.001467	0.002036	0.008956	0.002348
0002	0.044766	0.002260	0.000801	0.003742
	0.005217	0.001985	0.002428	0.003215
	0.000821	0.002771	0.000258	0.002164
	0.001216	0.001833	0.000957	0.002224
0003	0.212059	0.003532	0.001072	0.004383
	0.006055	0.003355	0.002715	0.003913
	0.000889	0.003715	0.000362	0.003200
	0.003046	0.004275	0.001618	0.004075
0004	0.161173	0.002237	0.000714	0.003746
	0.005612	0.001981	0.002370	0.003180
	0.000744	0.002765	0.000259	0.002165
	0.001019	0.001744	0.000938	0.002170
0005	1.572076	0.005389	0.001442	0.006963
	0.006789	0.003473	0.002495	0.004336
	0.001123	0.005389	0.000671	0.004998
	0.002167	0.004043	0.003438	0.006007
0006	0.293614	0.007445	0.001808	0.008395
	0.013295	0.006762	0.003693	0.006367
	0.001675	0.006355	0.000837	0.005984
	0.005525	0.008319	0.003718	0.008060
0007	0.163392	0.005241	0.001428	0.005848
	0.006878	0.002347	0.002600	0.003295
	0.001054	0.004632	0.000925	0.002482
	0.000975	0.001549	0.000945	0.002481

그림 7. 얼굴 영역에 대한 WPA 결과

Fig. 7 WPA result of the face region

V. 실험결과

그림 8은 제안한 알고리즘을 사용하여 추출된 얼굴 영역들이다. 얼굴 요소를 포함하는 최소 사각형 영역이 만족할 만하게 추출되었다.



그림 8. 정확하게 추출된 얼굴 영역들.
Fig. 8 Correct extraction of the face region



그림 9. 잘못 추출된 얼굴 영역들
Fig. 9 Incorrect extraction of the face region

반면에 얼굴이 너무 기울어지거나, 전체적인 명도값이 너무 밝거나 어두운 경우 그리고 배경 잡음 등의 영향으로 그림9에서 보듯이, 얼굴 영역 추출이 실패하는 경우도 있었다. 하지만 위와 같이 일부 특수한 경우의 영상을 제외하면 대부분의 실험 영상에서 얼굴 영역 추출에 성공하였다. 그러나 여러 이미지에서의 얼굴 영역 추출에 관한 정확도에 대해 성능 평가를 해 보았다.

5.1 성능평가

여러 가지 밝기의 피부색이 다양한 입력영상을 이용하여 얼굴영역 추출을 실험하였다. 기존방법의 피부색만을 이용한 추출방법과 피부색과 에지를 이용한 추출 방법 그리고, 본 논문에서 제시한 방법을 비교하여 정확도를 측정하였다. 정확도를 계산하는 방법으로 첫 번째, 추출된 얼굴 영역내에 눈과 입이 위치하는지를 알아보고, 두 번째, 미리 수작업으로 영상에 표시한 최적 사각형 영역과의 오차를 비교하였다. 오차는 적을수록 높은 값을 부여하였다. 정확도는 이 두가지 결과값을 합산한 후 백분율

을 구해 누적한 후, 사진 장수에 대한 평균값으로 산출하였다.

실험에 사용된 데이터는 조명에 따른 피부색상 밝기별로 분류했으며, 황인종에 한정하였다. 그러나 얼굴의 회전각도에 대한 제한은 두지 않았다.

표 2. 이전의 방법들과 성능비교

Table 2. Comparision with the previous methods

이미지	방법1	방법2	방법3	장수
밝음	57.0%	58.0%	77.9%	50장
보통	72.6%	80.8%	91.6%	80장
어두움	53.0%	55.2%	76.0%	40장

실험결과 일반적인 피부색만을 이용한 경우보다 피부색과 에지를 이용한 경우 얼굴영역 추출의 정확도가 높았다. 그러나 본 논문에서 제시한 방법으로 얼굴 영역을 추출한 경우의 정확도가 가장 높았음을 알 수 있다. 본 논문의 경우는 피부색 검출을 통한 얼굴 구성 요소포함에서 비슷한 성능을 보였으나 구조적 정보를 이용한 기울기 보정과 WPA를 이용한 검증을 통하여 최적 사각형 영역의 정확도를 높일 수 있었다.

방법1, 2에서는 색상의 밝기에 대한 범위값이 고정되어 있어 피부색의 밝기에 따른 가중치의 변화를 주지 않았다. 이로인해 평상시의 경우와 밝거나 어두운 경우의 추출율에서 큰 차이를 보였다. 그러나 세가지 방법 모두 평상시의 경우에서 가장 높은 추출율을 보인 것으로 보아 밝기변화가 클수록 추출의 정확도가 낮아짐을 알 수 있었다.

실험 결과를 통하여 정면얼굴을 포함하는 상반신 영상에 대하여 기존 방법보다 개선된 검출 결과를 확인할 수 있었다.

VI. 결 론

제안된 알고리즘에서는 먼저 피부 색상 영역에 대해 얼굴 포함 최대 사각 영역을 탐색하였다. 그리고 눈과 입 영역을 기준으로 동양인의 얼굴 통계치를 적용하여 얼굴 포함 최소 사각 영역을 결정하였다. 얼굴의 피부색상 범위와 최소 포함 사각형 영역의 조건은 되도록 얼굴 영역이 손상되지 않는 범위로 설정하고자 하였다. 마지막 단계에서 WPA를 사용한 얼굴 텍스춰 템플릿을 얼굴 후보 영역과 정합하여 얼굴 영역의 확인과 보정을 시도하였다.

그런데 영상 내 얼굴이 30° 이상 기울어져 있고, 조명이 밝거나 어두운 경우 추출에 실패한 경우도 있었다. 이는 1단계의 색상, 2단계에서 얼굴 후보 영역에 대한 기울어짐에 대한 보정이 좀더 강화될 필요가 있다고 생각된다. 앞으로 이러한 점이 개선된다면 WPA를 이용한 얼굴 영역 추출의 전체적인 정확도가 더욱 향상될 것으로 예상된다. 추후 보다 다양한 영상에서 얼굴 추출을 적용할 수 있는 방법에 관한 연구가 요구된다.

참고 문헌

- [1] 조영원, 조경은, 조형제, “위치 변화가 다양한 얼굴에 서의 특징 요소 추출”, 정보처리학회지, 제6권, 제2호, pp. 148-152, 1999년
- [2] 반가운, 유기형, 박정호, 최재호, 곽훈성, “웨이 브렛 변환을 이용한 내용기반 검색 시스템”, 전자공학회지, 제 21권, 제1호, pp. 733-763, 1998년 6월
- [3] 채영심, 위성두, 강현철, 김정규, “내용기반 검색을 위한 웨이브렛 변환을 이용한 텍스트 특징 추출”, 정보 과학회지, 제28권, 2001년 4월
- [4] Christophe Garcia, Georgios Tziritas, “Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis”, IEEE Trans. Multimedia, Vol. 1, No. 3, pp. 264-277, Sep. 1999
- [5] 김영일, 김정훈, 이응주, “HSI 정보와 얼굴 특징자들의 기하학적 특징각을 이용한 얼굴 인식 알고리즘”, 정보 처리학회지, 제8권, 1호, pp. 859-862, 2001년
- [6] 윤호섭, 왕민, 민병우 “눈영역 추출에 의한 얼굴 기울기 교정”, 전자공학회 논문지, 제33권, 제12호, 1996년 12월
- [7] 김영길, 한재혁, 안재형 “컬러 정지 영상에서 색상과 모양 정보를 이용한 얼굴영역 검출”, 멀티미디어 학회 논문지, 제4권, 제1호, 2001년 2월
- [8] C.Carcia, G.Zikos, G.Tziritas “Wavelet packet analysis for Face recognition”, Image and Vision Computing, 18 권, pp. 289-297, 2000
- [9] 강영미, 정성환, “Wavelet과 신경망을 이용한 내용기반 검색 시스템”, 컴퓨터전산교육학회 논문지, Vol.2, No. 3, March
- [10] Constantine Papageorgiou, Tomaso Poggio, “A Trainable System for Object Detection”, International Journal of Computer Vision 38(1), pp. 15-53, 2000

저자소개



송 호 근 (Ho-Keun Song)

1993년 중앙대,전자공학과,석사
1997년 중앙대,전자공학과,박사
1996년-현재 한서대학교 컴퓨터 정보학과 부교수

※ 관심분야 : 내용기반 영상/비디오 검색, 3차원 영상 처리/인식, 영상 통신/회의 등