

논문 2014-51-11-22

보행자 상반신 검출에서의 컬러 세그먼테이션 활용 (Exploiting Color Segmentation in Pedestrian Upper-body Detection)

박 래 정*

(Lae-Jeong Park[Ⓒ])

요 약

본 논문에서는 보행자 상반신 검출기의 성능을 향상하기 위한 세그먼테이션에 기반한 특징 추출 방법을 제안한다. 상반신의 부분별 색상 분포를 활용한 멀티 파트 컬러 세그먼테이션을 사용하여 국소 특징이 갖는 한계로 인해 발생하는 오검출의 감소에 효과적인 “전역적” 윤곽 특징을 추출한다. 컬러 공간과 히스토그램 분해도에 따른 성능을 분석하였으며, 자체 구축한 보행자 상반신 영상을 사용한 실험을 통해서 제안한 방법으로 추출한 특징이 국소 특징 기반 검출기의 오검출 감소에 효과적임을 확인하였다.

Abstract

The paper proposes a new method of segmentation-based feature extraction to improve performance in pedestrian upper-body detection. General pedestrian detectors that use local features are often plagued by false positives due to the locality. Color information of multi parts of the upper body is utilized in figure-ground segmentation scheme to extract an salient, “global” shape feature capable of reducing the false positives. The performance of the multi-part color segmentation-based feature is evaluated by changing color spaces and the parameters of color histogram. The experimental result from an upper-body dataset shows that the proposed feature is effective in reducing the false positives of local feature-based detectors.

Keywords : pedestrian detection, HOG, segmentation, color, false positive

I. 서 론

짙은 가림(occlusion)과 복잡한 배경 환경 등으로 인해서 쇼핑몰이나 지하철 역 내부처럼 붐비는 환경에서 보행자를 검출하는 것은 용이하지 않은 과제이다^[1]. 인파로 붐비는 상황에서의 보행자 검출을 위해서 상반신, 즉 머리 혹은 머리-어깨 패턴을 검출하는 시도가 있었다^[2-4]. 한편, 보행자 검출의 응용 특성 상, 실시간 동작

이 중요하므로 특징 기반(feature-based) 검출기의 대부분이 국소(local) 블록 내의 특징, 예를 들면, HOG (histogram of oriented gradients)^[5]를 사용한다. 하지만, 국소 블록 내의 특징이 갖는 한계로 인해서 그림 1 과 같이 검출 결과로는 의아해 보이는 영상이 검출되는 경우가 적지 않다. 이는 그림의 영상들이 공통적으로



그림 1. 국소 특징 기반 머리-어깨 검출기의 오검출 사례

Fig. 1. False positive samples of a local feature based head-shoulder detector.

* 정회원, 강릉원주대학교 전자공학과
(Department of Electronic Engineering,
Gangneung-Wonju National University)

Ⓒ Corresponding Author(E-mail: ljpark@gwnu.ac.kr)

접수일자: 2014년08월13일, 수정일자: 2014년10월04일
게재확정: 2014년11월01일

머리-어깨 패턴과 유사한 국소적 에지(edge), 특히 양 어깨 영역에서의 에지 패턴을 갖고 있기 때문이다.

이러한 국소 HOG에 기반한 검출기의 오검출 감소를 위해 비국소적인, 즉 전역적(global) “컬러” 정보를 활용한 특징 추출 방법들이 제안되었다. 보행자의 각 부위별 컬러 상관성을 반영한 국소 블록간의 컬러 유사도 특징^[6]과, 보행자 전신 혹은 상반신의 좌우 대칭 속성을 반영한 컬러 및 HOG 블록간의 대칭 유사도 특징^[7]등이 그 사례이다. 한편 전역적 컬러 정보를 보다 “직접적”으로 사용하기 위해서 세그멘테이션(segmentation)을 활용하는 방법이 제시되었다. Ramanan^[8]은 컬러 세그멘테이션 기법에 기반한 전역적 윤곽 특징 추출 방법을 제시하였으며 국소 특징 기반 검출기의 오검출을 효과적으로 줄일 수 있음을 보였다. [9]에서 제시한 방법을 머리-어깨 패턴 검출에 사용하여 보행자의 상반신 검출 문제에서의 적용 가능성을 확인하였다.

본 논문에서는 보행자 머리-어깨 검출에서의 추가 성능 개선을 위하여, 목표 패턴의 영역별 컬러 차이를 활용한 멀티 파트 컬러 세그멘테이션 기법을 제시하고, 이를 통해 추출한 특징의 분류 성능을 검증하고자 한다. 또한, 멀티 파트 컬러 세그멘테이션의 최적 성능을 위한 컬러 공간 및 컬러 히스토그램 변수 등에 대해 논의한다.

본 논문의 구성은 다음과 같다. II장절에서는 제안하는 멀티 파트 컬러 세그멘테이션을 활용한 전역적 윤곽 특징 추출 방법을 소개한다. III장에서는 보행자 상반신 데이터에 대해서 제시한 방법의 실험 결과를 기술하고 타당성을 검증하며, IV장에서 결론을 맺고 향후 연구 과제에 대해서 검토한다.

II. 멀티 파트 컬러 세그멘테이션 기반 특징 추출

2.1. 컬러 세그멘테이션을 통한 윤곽 추출

그림 2는 컬러 세그멘테이션을 활용한 전역적 윤곽 특징을 추출하는 과정^[8]을 나타낸다. 검출하고자 하는 목표 패턴, 즉 머리-어깨 패턴(foreground, FG) 영역과 배경(background, BG) 영역의 색상 정보를 히스토그램 형태로 추출하고, 그래프 컷(graph-cut) 방법^[10]으로 figure-ground 세그멘테이션을 수행하여 각 픽셀을 머리-어깨 패턴 영역과 배경 영역으로 양분한다. 그림에서 보듯이 보행자 상반신 영상(그림 2(b))은 BG 영역과

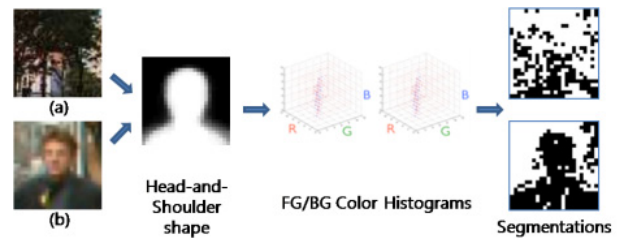


그림 2. 컬러 세그멘테이션을 이용한 전역적 윤곽 특징 추출 과정

Fig. 2. Extraction of global shape features through color segmentation.

의 색상 차이로 인해 머리-어깨 패턴의 윤곽이 생성되는데 반해서, 국소특징 기반 검출기의 오검출 영상(그림 2(a))은 BG 영역과 FG 영역에 유사한 색상이 존재하기 때문에 머리-어깨 패턴과는 다른 윤곽이 나타남을 볼 수 있다.

상세히는, 그래프 컷(graph-cut) 세그멘테이션은 식 (1) 함수의 최적화로 표현된다.

$$E(l_1, \dots, l_K) = \sum_k c(l_k) + \alpha \sum_{j, k \in N} I(l_j \neq l_k) \quad (1)$$

여기서, l_k 는 픽셀 k의 라벨(FG 혹은 BG)이고, $c(l_k)$ 는 픽셀 k가 FG나 BG로 할당될 경우의 비용함수이며, $I(\cdot)$ 는 인근(neighborhood) 픽셀 라벨과의 불일치에 대한 비용, 즉 평탄화(smoothness) 제약 함수이다. 실험에 사용하는 $c(l_k)$ 는 식 (2)와 같다.

$$\begin{aligned} c(l_k = FG) &\propto -\log[H_{FG}(im(k))], \\ c(l_k = BG) &\propto -\log[H_{BG}(im(k))] \end{aligned} \quad (2)$$

H_{FG} 과 H_{BG} 는 각각 FG 영역(머리-어깨 패턴 영역)과 BG 영역의 컬러 히스토그램을 나타내고, $im(k)$ 는 히스토그램에서 픽셀 k의 색상이 속하는 히스토그램의 bin의 인덱스를 나타낸다. 식 (1)-(2)에 의해서, 각 픽셀은 해당 픽셀의 색상과 동일한 색상의 빈도가 높은 영역으로 할당된다.

2.2. 멀티 파트 컬러 세그멘테이션

FG 영역과 BG 영역의 컬러가 상이한 경우(그림 2(b))에는, 히스토그램의 각 구간별 $H_{FG}(\cdot)$ 과 $H_{BG}(\cdot)$ 의 값이 큰 차이를 갖게 되어, 식 (1)-(2)에 의해서 목표 패턴의 윤곽을 갖는 세그멘테이션이 생성된다. FG 영역과 BG 영역이 동일한* 색상의 픽셀을 포함하는

경우(그림 2(a))에는, 해당 구간에서의 $H_{FG}()$ 과 $H_{BG}()$ 이 비슷한 값을 가지게 되고, 목표한 것처럼, 목표 패턴(머리-어깨 패턴)의 윤곽과는 다른 세그멘테이션이 생성된다. 하지만, 목표 패턴이 존재하지 않은 영상에서도 FG와 BG 영역에 일정 비율 정도로 동일한 색상이 분포된 경우가 존재하는데, 이 경우에도 목표 패턴의 세그멘테이션이 생성되기도 한다. 이는 동일 색상을 갖는 FG/BG 영역의 픽셀 수의 비율, 즉 $H_{FG}()$ 과 $H_{BG}()$ 의 값의 비율에 의해서 세그멘테이션 결과가 도출되기 때문이다. 동일한 컬러의 픽셀이 BG 영역보다 FG 영역에 많은 경우에, 실제 목표 패턴이 없음에도 불구하고 FG 영역과 유사한 윤곽을 갖는 세그멘테이션이 생성된다(그림 6 참조).

이러한 문제를 완화하기 위해서, 목표 패턴 영역 즉 머리-어깨 영역을 멀티 파트, 여기서는 머리 영역과 어깨 영역으로 구분하고, 각 파트별로 독립적으로 컬러 세그멘테이션을 수행하는 방식을 제시한다. 즉, 머리 영역을 기준으로 FG (FG_H)와 BG (BG_H) 영역을 구분하고 컬러 세그멘테이션을 수행하며, 추가적으로 어깨 영역을 기준으로 FG (FG_S)와 BG (BG_S) 영역을 구분해서 컬러 세그멘테이션을 수행한다. 각 파트별로 컬러 세그멘테이션을 수행함으로써, 머리-어깨 패턴을 사용할 때의 동일한 컬러(해당 컬러의 구간이 n 이라 가정)의 히스토그램 $H_{FG}(n)$ 과 $H_{BG}(n)$ 값을, $H_{FG_H}(n)$ 와 $H_{BG_H}(n)$, 그리고 $H_{FG_S}(n)$ 와 $H_{BG_S}(n)$ 의 값으로 구분할 수 있게 된다. 따라서 기존의 머리-어깨의 통합된

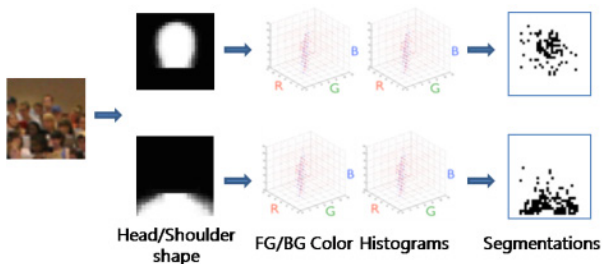


그림 3. 멀티 파트 세그멘테이션을 이용한 머리와 어깨 윤곽 특징 추출

Fig. 3. Extraction of two separate head and shoulder shape features through multi-part color segmentation.

* 여기서, “동일한” 컬러의 의미는 컬러 히스토그램의 bin 인덱스가 동일하다는 의미이다.

패턴의 유무가 아니라, 머리 패턴의 윤곽 여부와 어깨 패턴의 윤곽 여부를 각각의 색상 분포를 활용하여 독립적으로 추출하게 된다. 이를 통해서 머리와 어깨 영역에 동일한 컬러가 존재하는 경우에 파트별로 한 단계 더 윤곽 여부를 반영하는 특징을 추출하여 검증하는 장점을 갖는다. 그림 3에 제시한 멀티 파트세그멘테이션에 의한 파트별 윤곽 특징 추출 과정을 도시하였다.

2.3. 특징 처리 및 분류기

학습을 통해서, 멀티 파트 세그멘테이션으로 추출한 윤곽 특징을 근거로 상반신 여부를 판단하는 분류기를 구축한다. 그림 3의 두 세그멘테이션이 계산되면 분류 특성을 향상하기 위해서, 0과 1의 이진 데이터(세그멘테이션 결과)를 슬라이딩 윈도우(예를 들면, 3x3)를 사용하여 각각 평탄화(smoothing) 처리한다. 실험을 통해서 이진 형태에 비해 일반화 성능의 향상에 효과적임을 확인하였다. 이후, 평탄화된 두 세그멘테이션 결과를 연결(concatenation)하여 특징을 만들고, 이를 SVM (support vector machine) 분류기의 입력으로 사용하여 최종적으로 상반신 패턴 여부를 검증한다. 실제 동작 소요 시간을 고려해서 계산량이 적은 선형 SVM 분류기를 사용하였다.

III. 실험 결과 및 토의

3.1 실험 데이터

제시한 방법의 성능을 비교 검증하기 위해, INRIA 보행자 데이터^[5]로부터 후면과 측면의 보행자 영상을 제외한, 680개의 근정면(near-frontal) 보행자 영상 데이터를 생성하였다. INRIA 네거티브* 영상집합에 지하철역, 쇼핑몰 내부 등의 영상을 추가하고, 이 중 국소 HOG 특징을 사용한 Viola-Jones(VJ) cascade 분류 방식을 채택한 머리-어깨 패턴 상반신 검출기^[7]을 통과한 1,460개의 오검출 영상을 사용하였다. 사용한 상반신 검출기는 10^{-4} fppw(false positive per window)에서 검출율 75%의 성능을 갖는다. 가용한 데이터(상반신 680개, 오검출 1,460개)에서 학습에 75%, 테스트에 25%를 무작위로 배분하여 총 10회의 실험을 수행하였다. 그림 4는 보행자 상반신 영상과 네거티브 영상의 샘플을 나

* 보행자 상반신이 포함되지 않은 영상

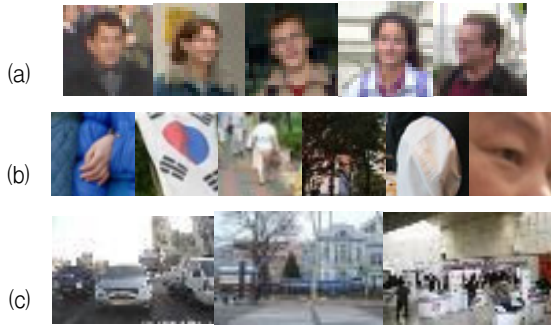


그림 4. 실험에 사용한 데이터. (a) 상반신 영상. (b) 오검출 영상. (c) 오검출 샘플 생성용 영상
 Fig. 4. Data set. (a) upper-body images. (b) false positive images. (c) negative images for generation of false positives.

타낸다.

3.2 실험 결과 및 토의

제시한 특징의 결과를, HOG 국소 블록 특징과 싱글 (머리-어깨 패턴) 세그멘테이션 특징의 결과와 함께, 그림 5의 ROC (receiver of operating characteristic) 그래프*로 도시하였다. 컬러 히스토그램은 RGB 컬러 공간에서 계산하였으며, 한 축당 bin 개수는 18로 설정하였다. 싱글 세그멘테이션 특징과 제시한 멀티 파트 세그멘테이션 특징 모두가 HOG 블록 특징 (6블록×6블록, 총 36블록)에 비해 전 구간에서 높은 검출율 성능을 보였다. 제시한 방법에 의한 특징이 싱글 세그멘테이션 특징에 비해, [0.0~0.4]의 오검출율 구간에서 약 10% 이상의 검출율 개선과, [0.7~0.9]의 검출율 구간에서 약 5% 정도의 오검출 개선 성능을 보였다. 90%의 검출율 동작 점에서 오검출 영상의 감소 비율은 HOG 블록 특징 대비 38%, 싱글 세그멘테이션 특징 대비 22%이다. 실제로 멀티 파트 세그멘테이션 특징은 기존 세그멘테이션 특징보다 차원이 2배 높다. 그림 5에서 볼 수 있듯이, 싱글 세그멘테이션 특징을 단순 2배 증가한 경우보다 높은 성능 보였으며, 이는 성능 개선의 원인이 단순 차원 증가에 의한 것이 아니라 특징의 유효성에 있음을 의미한다.

그림 6에 국소 특징 기반 검출기의 오검출 영상에 대해서, 멀티 파트를 사용한 경우와 싱글 파트를 사용한 경우의 세그멘테이션 결과를 도시하였다. 싱글 파트 세

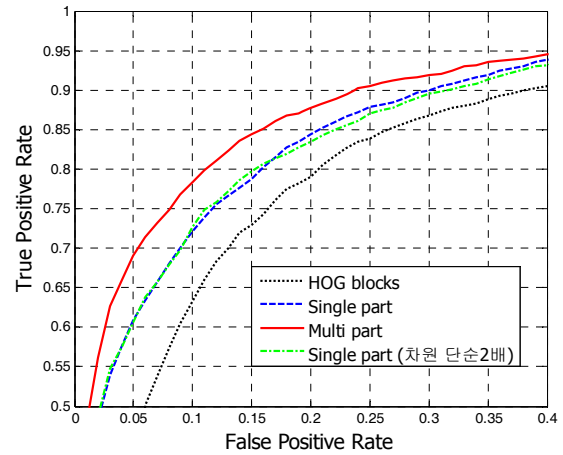


그림 5. 멀티 파트 세그멘테이션 특징의 ROC 그래프 (RGB 컬러 공간)
 Fig. 5. ROC curves of features produced by multi-part segmentation (RGB color space).

그멘테이션의 경우, 머리-어깨 패턴과 유사한 윤곽이 나타나는 것을 확인할 수 있다. 의아한 결과이지만, 육안으로 유사해 보이는 FG 영역과 BG 영역의 컬러가 실제로 컬러 공간에서 “동일한” 색상이 아니기 때문이다. 더 큰 원인은, “동일한” 색상을 갖는 픽셀 수가 FG 영역에 상대적으로 더 많이 분포하기 때문이다. 이에 반해, 파트별로 독립적으로 컬러 세그멘테이션을 적용한 경우에는, 상대적으로 더 넓어진 배경 영역 탓으로 인해, 즉, 식 (2)의 $H_{BG_H}(n)$ 과 $H_{BG_S}(n)$ 가 $H_{FG_H}(n)$ 와 $H_{FG_S}(n)$ 를 압도하여, 머리와 어깨 내의 컬러가 배경과 유사한 경우에 배경 컬러에 의해 억제되어 머리와 어깨의 윤곽이 사라지게 된다. 그림 6 (b)-(c) 오검출 영상에 대해서 멀티 파트 세그멘테이션에 의해 특히 어깨 영역의 윤곽을 불분명하게 하는 효과가 있음을 확인할 수 있다.

실제로는 그림 6 (b)-(c) 영상에 대해서 컬러 히스토그램의 bin 개수를 줄이면 싱글 파트 세그멘테이션의 경우에도 머리-어깨 윤곽이 나타나지 않는다. 하지만 bin 개수가 적정 값보다 작은 경우에는 색상 구분도 (distinguishability)가 낮아져서 상반신이 포함된 영상의 검출도 함께 낮아진다. 따라서 bin 개수에 따른 저분해도(low resolution)와 고분해도(high resolution) 사이의 적절한 타협점이 필요하다. 대략 16~24 사이의 값을 갖는 경우에 좋은 성능을 보였다.

컬러 히스토그램 차원에서 각 영상의 구분도는, bin

* 10회 실험 결과를 평균한 성능임.



그림 6. 오검출 영상의 멀티 파트 세그멘테이션 예시
Fig. 6. Examples of multi-part segmentation for false positive samples.

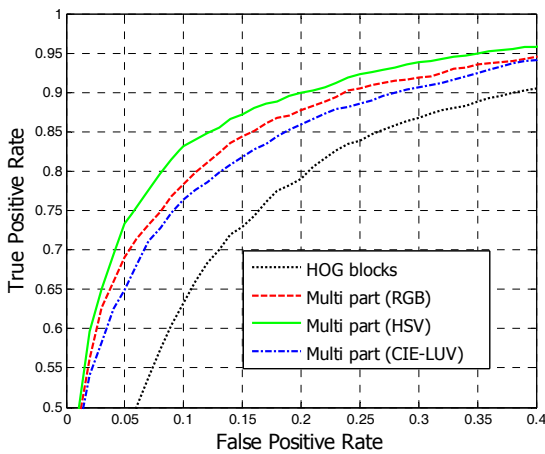


그림 7. 컬러 공간에 따른 멀티 파트 세그멘테이션 특징의 ROC 그래프
Fig. 7. ROC curves of features produced by multi-part segmentation in three different color spaces.

개수 이외에도 사용하는 컬러 공간^[11]에 따라 달라질 수 있으므로, 컬러 공간에 따른 성능에 대해 실험 분석하였다. 테스트해 본 컬러 공간은 HSV 공간과 CIE-LUV 공간이다. 컬러 공간에서 각 컬러 축마다 그 기여도가 다르기 때문에 3 축마다 동일하지 않게 분할(non-equal quantization)하는 것이 효과적일 수 있다. 특히, HSV 공간에서는 채도(saturation)과 value(명암) 정보보다 색상(hue) 정보가 색상 구분 정도에 중요하므로, 색상 축을 보다 세분해서 분할하였다*.

* RGB 공간과 CIE-LUV 공간은 3축에 대해 18/18/18로 분할하였으며, HSV 공간은 18/12/12로 분할하였다.

그림 7은 컬러 히스토그램을 각각 RGB 공간, CIE-LUV 공간, HSV 공간에서 계산하고 이를 사용하여 멀티 파트 세그멘테이션 한 결과의 성능 그래프이다. CIE-LUV공간을 사용한 경우는 RGB 공간의 경우보다 유사하거나 낮은 성능을 보였으나, HSV 공간을 사용한 경우는 RGB 공간의 경우보다 일관되게 개선된 성능을 보였다. 특히 0.9 이상의 검출율 구간에서는 약 8% 정도의 오검출 개선 성능을 나타냈다. RGB 공간과 CIE-LUV 공간에 비해, HSV 공간이 색상 구분 측면에서 채도와 명암보다 더 유의미한 “색상” 정보를 표현할 수 있는 공간이기 때문으로 추정된다. 멀티 파트 세그멘테이션 과정에서, 머리와 어깨 영역별로 다른 bin 개수를 사용할 수 있는 옵션이 가능한데, 대략적으로 머리 영역의 구간 개수를 어깨 영역 (다양한 의상)의 경우에 비해 낮게 설정하는 편이 나은 성능을 보였다.

IV. 결 론

본 논문에서는 보행자 상반신 검출기의 성능을 향상하기 위한 컬러 세그멘테이션에 기반한 특징 추출 방법을 제안하였다. 국소 특징이 갖는 한계로 인한 오검출을 줄이기 위해서, 기존의 컬러 세그멘테이션을 확장한 부분별 색상 분포 차이를 활용한 “멀티 파트” 컬러 세그멘테이션을 통하여 “전역적” 특징을 추출하였으며, 자체 구축한 보행자 상반신 데이터를 통해서 제안한 방법의 성능 개선을 검증하였다. 또한, 컬러 공간 및 컬러 히스토그램 과정에 따른 성능 분석을 통해 HSV 컬러 공간이 효과적임을 확인하였다. 최종적으로 제시한 특징이 기존의 컬러 세그멘테이션 기반 특징에 비해 검출율 90% 동작점 기준으로 약 10% 이상의 오검출을 개선을 보였다. 향후에 컬러축의 균일한 분할 방식대신 주변 컬러와의 구분도를 높이기 위한 적응적 컬러 히스토그램 기법의 적용 가능성을 타진하고자 한다. 또한, 다양한 야외 시험 영상을 사용하여 영상 취득 환경 요인에 따른 성능을 검증할 예정이다.

REFERENCES

[1] D. Geronimo, A. M. Lopez, A. D. Sappa, and T. Graf, “Survey on pedestrian detection for advanced driver assistance systems,” *IEEE*

Trans. on PAMI, vol. 32, pp. 1239-1258, 2010.

- [2] Q. Zhu, M.-C. Yeh, K.-T. Cheng, and S. Avidan, "Fast human detection using a cascade of histograms of oriented gradients," *Proc. IEEE Conf. on CVPR*, pp. 1491-1498, 2006.
- [3] M. Li, Z. Zhang, K. Huang, and T. Tan, "Estimating the number of people in crowded scenes by MID based foreground segmentation and head-shoulder detection," *Proc. Int. Conf. on Pattern Recognition*, pp. 1-4, 2008.
- [4] X. Ding, H. Xu, P. Cui, L. Sun, and S. Yang, "A cascade SVM approach for head-shoulder detection using histograms of oriented gradients," *Proc. IEEE Int. Symposium on Circuits and Systems*, pp. 1791-1794, 2009.
- [5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *Proc. IEEE Conf. on CVPR*, pp. 886-893, 2005.
- [6] S. Walk, N. Majer, K. Schindler, and B. Schiele, "New features and insights for pedestrian detection," *Proc. IEEE Conf. on CVPR*, pp. 1030-1037, 2010.
- [7] L.-J. Park and J-H Moon, "Exploiting global self similarity for head-shoulder detection," *World Academy of Science, Engineering and Technology*, vol. 76, pp. 144-148, 2013.
- [8] D. Ramanan, "Using segmentation to verify object hypothesis," *Proc. IEEE Conf. on CVPR*, pp. 1-8, 2007.
- [9] L.-J. Park, "Reduction of false positives in pedestrian detection by color segmentation," *Proc. of KIIS Spring Conference*, pp. 157-158, 2013.
- [10] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *Proc. IEEE Conf. on Computer Vision*, pp. 377-384, 1999.
- [11] T. Gevers, A. Gijzenij, J. van de Weijer, and J-M. Geusebroek, *Color in Computer Vision: Fundamentals and Applications*, Wiley, 2012.

— 저 자 소 개 —



박 래 정 (정회원)

1991년 서울대학교 전기공학과
학사 졸업

1993년 한국과학기술원 전기및
전자공학과 석사 졸업

1997년 한국과학기술원 전기및
전자공학과 박사 졸업

1997년 ~ 1999년 LG종합기술원 선임연구원

2000년 ~ 현재 강릉원주대학교 (전 강릉대학교)
전자공학과 교수

<주관심분야 : 패턴 인식, 기계 학습, 지능 시스템>