

논문 2011-48SP-2-11

Histogram of Oriented Gradient를 이용한 실시간 소실점 검출

(Real-time Vanishing Point Detection Using Histogram of Oriented Gradient)

최 지원*, 김 창 익**

(Jiwon Choi and Changick Kim)

요 약

소실점이란 실제 공간의 평행한 선들이 영상 내에 투영되면서 한곳에 모이는 점이다. 본 논문에서는 이러한 소실점의 특성을 이용한 실시간 소실점 검출 알고리즘을 제안한다. 기존의 소실점 검출 방법은 1) 복잡한 계산이 요구되거나 2) 알고리즘에 따라 소실점을 검출할 수 있는 영상이 제한되어 있다. 제안하는 방법은 블록 기반의 HOG(Histogram of Oriented Gradient)를 구하여 영상의 구조적 특성을 이용하는 것으로 영상 내에 존재하는 소실점을 실시간으로 검출한다. 먼저 영상의 블록 단위로 HOG 기술자를 구한 뒤, 제안하는 동적 프로그래밍(dynamic programming)을 이용하여 소실점의 위치를 예측한다. 본 논문에서는 다양한 영상에 대한 실험을 통해 제안하는 알고리즘이 효율적인 소실점 검출 방법임을 보이고자 한다.

Abstract

Vanishing point can be defined as a point generated by converged perspective lines, which are parallel in the real world. In this paper, we propose a real-time vanishing point detection algorithm using this fundamental feature of vanishing point. The existing methods 1) require high computational cost or 2) are restricted to specific image contents. The proposed method detects the vanishing point in images based on the block-wise HOG (Histogram of Oriented Gradient) descriptor. First, we compute the HOG descriptor in a block-wise manner, then estimate the location of the vanishing point using the proposed dynamic programming. Experiments are performed on diverse images to confirm the efficiency of the proposed method.

Keywords : vanishing point, vanishing point detection, HOG descriptor, dynamic programming

I. 서 론

소실점(vanishing point)은 실제 공간의 평행한 직선들이 2차원 영상으로 투영되면서 원근감에 의해 한 점에서 만나는 것으로 소실점 검출을 활용한 예는 다음과 같다. 세 개의 서로 직교하는 방향의 소실점과 소실선을 구함으로써, 인공 구조물(architectural)을 해석하여 건물

의 재해석이 가능하다^[1]. 이러한 방법은 영상의 구조적 해석을 위해서 소실점이 사용된다. 소실점 검출을 이용하여 도로(road) 검출이 가능하다^[2~3]. 일반적으로 소실점과 그 소실점을 잇는 주요한 에지(edge)를 연결함으로써 도로를 분리할 수 있다. 2차원 영상의 3차원 변환에 대해서도 소실점이 사용되는 것을 알 수 있다. 특히, 인공 구조물이 포함된 영상의 2차원 영상의 3차원 변환에서는, 소실점을 검출함으로써 깊이정보를 생성할 수 있다. 이는 3차원 공간이 2차원 영상으로 변하면서 소실점이 위치하는 부분은 일반적으로 영상 내에서 가장 먼 곳에 해당하므로, 상대적인 깊이 추정이 가능하기 때문이다. Battiattoa 등은^[4] 한 장의 영상으로부터 소실점 검출을 통해서 깊이정보를 예측하였다. 두 장의 영상을 획

* 학생회원, ** 정회원, 한국과학기술원 전기 및 전자공학

(Dept. of Electrical Engineering, KAIST)

※ 본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2010- (C1090-1011-0003))

접수일자: 2010년7월16일, 수정완료일: 2010년11월10일



그림 1. 영상 내에 소실점이 있는 영상의 예
 Fig. 1. The examples of images which have a vanishing point.

특하여 스테레오 영상을 만들 때 카메라 교정이 필요한데 이때 소실점을 이용하기도 한다. 먼저 영상내의 소실점을 검출한 뒤, 두 장의 영상에서 서로 대응하는 소실점을 매칭 함으로써, 카메라 교정 파라미터를 구하는데 사용된다^[5]. 또, Horry등은^[6] 영상내의 소실점 정보를 이용해 영상 안을 여행하는 Tour in picture를 제안하였다.

영상 내에 소실점을 포함하고 있는 다양한 영상들을 그림 1에 나타내었다. 그림 1에서 알 수 있듯이, 영상 내의 소실점은 그 소실점을 기준으로 하여 여러 개의 투영된 선이 모인다. 본 논문에서는 이러한 영상의 특징을 이용하여 소실점을 검출하는 방법을 제안하고자 한다.

소실점 검출에 대한 연구는 다양한 응용에 대해 각각 적용할 수 있는 방법으로 연구되어 왔다. 영상 내의 정보를 이용하여 소실점을 검출하는 연구를 살펴보면 다음과 같다.

먼저, 영상에서 직선 또는 소실선을 검출한 뒤, 이를 이용해 소실점을 구하는 것을 살펴보겠다. Banard^[7]는 광학 중심(optical center)을 그 중심으로 하는 가우스 구(Gaussian sphere)에 영상 내의 직선들을 누적하여 가장 많이 누적되는 점을 소실점으로 검출하였다. Banard가 제안한 가우스 구 방법은 유한 및 무한 소실점을 표현할 수 있지만 카메라의 광학 중심을 알아야 하고, 평면 공간을 구 공간으로 투영하면서 에러가 생기기도 한다^[8]. Tuytelaars 등^[9]은 허프 변환(Hough Transform)을 통하여 각 레벨마다 소실선과 소실점을 검출하였다. 그러나 허프 공간으로 변환 되면서 실제 직선과 점간의 거리가 보존되지 못하는 단점이 있다.^[11]

Tardif^[10]는 인공 구조물 영상 내의 비반복적(non-iterative) 소실점 검출 방법을 제안하였다. 제안하는 방법은 영상 내에서 검출된 에지들을 최근에 제안된 J-Linkage 방법을 사용하여 예상되는 소실점 후보군을 추출한 뒤 소실점을 검출하는 것이다.

영상 내의 텍스처 방향(texture orientation)에 대한 누적 투표(voting)를 통해 소실점을 검출하는 방법은 다음과 같다. Rasmussen^[2]는 다중 스케일(multi-scale) 가버 웨이블렛 필터(Gabor wavelet filter)를 이용해 도로 위의 소실점 검출을 제안하였다. 각 픽셀에서 주요 텍스처 방향을 구한 뒤, 투표(voting)를 통해 최종 소실점을 검출한다. Hui 등은^[3] 먼저 가버 필터(Gabor filter)를 사용하여 각 픽셀(pixel)의 텍스처 방향을 계산한 후, 국부 영역(local region)의 모든 픽셀과 소실점 후보 간의 거리를 이용하여 도로 영상 내 소실점을 검출하였다. 이는 적응적(adaptive) 크기의 국부 영역 내에서 투표 과정을 통해 신뢰도가 낮은 텍스처 방향은 제거되고 누적된 점수(voting score)가 높은 소실점 후보가 소실점으로 선택되는 것이다.

그러나 앞서 제안한 소실점 검출 방법은 복잡한 계산과 더불어 소실점 검출에 상당한 시간이 소요된다. 이에 본 논문에서는 영상 내의 구조적 특징을 반영하는 HOG 기술자를 이용하여 실시간으로 소실점을 검출하는 알고리즘을 제안하고자 한다. 제안하는 방법은 특정 영상 군(인공적인 구조물 영상에 제한되거나 도로 영상에 제한되는)에 한정되지 않고 영상 내에 소실점이 있으면 검출할 수 있다.

논문의 구성은 다음과 같다. 먼저 II장에서는 HOG 기술자를 간략히 설명하고, 동적 프로그래밍 을 통해 소실점을 구하는 방법을 제안한다. III장에서는 다양한 영상에 대해 제안하는 알고리즘의 강건함과 효율성을 보이고 IV장에서 결론 및 앞으로 과제를 논의한다.

II. 제안하는 알고리즘

제안하는 알고리즘은 먼저 영상을 블록 단위로 HOG 기술자를 구하는 단계와 블록 단위로 구한 HOG 기술자 중, 주요한 방향(dominant orientation) 성분을 취득한 뒤 동적 프로그래밍으로 소실점의 위치를 예측하는 단계로 나눌 수 있다. 영상을 블록 단위로 HOG 기술자를 구함으로써 블록 내의 에지 방향을 특성을 알 수 있다. 이때, 각 블록의 HOG 기술자의 주요한 방향 성분

표 1. 제안하는 알고리즘
Table 1. The proposed algorithm.

Main Algorithm
1. 영상을 $m \times m$ 크기의 블록으로 나눈다.
2. 각 블록에 대해 HOG 기술자를 구한다.
3. 각 블록의 HOG 기술자에 대해 n 개의 주요 방향을 선택한다.
4. 외곽 블록으로부터
4.a 주요 방향이 유사한 블록을 모두 찾는다.
4.b 찾은 블록 중, 외곽 블록과 가장 먼 거리에 있는 블록을 선택한다.
5. 선택된 블록들로부터 중간 위치에 있는 블록이 소실점을 포함하고 있는 블록이다.

은 그 블록을 대표하는 에지의 방향을 가리킨다. 제안하는 동적 프로그래밍으로 유사한 에지의 방향을 가지는 블록을 찾는 과정에서 소실점을 구성하는 소실선을 찾을 수 있다. 표 1에서 제안하는 알고리즘을 간략하게 나타내었다.

1. HOG 기술자

HOG(Histogram of Oriented Gradient) 기술자는 객체 검출을 목적으로 고안된 특징정보 기술자로 Dalal 와 Triggs^[11]가 처음 제안하였다. Dalal 와 Triggs는 [11]번 논문에서 HOG 기술자를 이용한 정지 영상에서의 사람 검출 알고리즘을 제안하였다.

HOG 기술자는 국부(local) 영역 내의 화소값 변화(intensity gradient) 또는 에지의 방향을 나타내는 것으로 다음과 같이 구한다.

우선 영상 I 는 식 (1)과 같이 $[-1, 0, 1]$ 기울기 필터 커널(gradient filter kernel)을 거친다.

$$G = I * [-1, 0, 1] + I * [-1, 0, 1]^T \quad (1)$$

기울기(gradient) 영상 G 를 $m \times m$ 픽셀 크기의 블록으로 나눈 후, 각 블록에서 에지의 방향에 대한 히스토그램을 구한다. 본 논문에서는 계산된 에지의 방향(orientation)을 18등분으로 분할하였으며, 각각의 방향에 대해 방향의 크기를 비교하여 주요한 n 개의 방향만 취득하였다. 이를 그림 2에 나타내었다. 그림 2에서는 18개 방향 중에서 5개의 주요 방향만 선택하였는데, 표시된 HOG 기술자가 영상의 에지 특성을 반영하고 있는 것을 알 수 있다. 이를 활용하여 소실점의 위치를 구

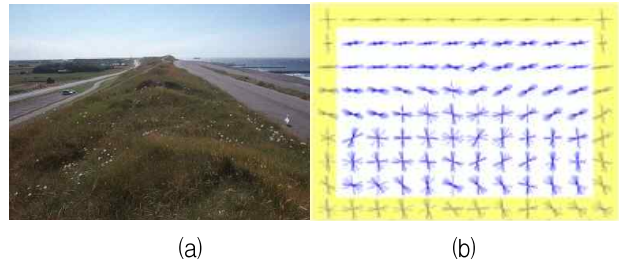


그림 2. HOG 기술자 예 (a) 원영상, (b) 5개의 주요 방향만 표시한 HOG 기술자, 노란색으로 표시된 영역이 외곽블록이다

Fig. 2. Example of HOG descriptor (a) original image, (b) HOG descriptor which has 5 dominant orientation, the outer blocks are marked yellow.

하는 방법을 다음 절에서 설명한다.

2. 소실점 검출

가. 동적 프로그래밍

앞 절에서 구한 HOG 기술자를 이용하여 소실점의 위치를 예측하기 위한 동적 프로그래밍을 다음과 같이 제안하였다.

우선, 영상을 $m \times m$ 크기의 블록으로 나눈 뒤, 각 블록에 대해 HOG 기술자를 구한 다음, 이 중에서 n 개의 주요한 방향을 구한다. 블록 b 의 HOG 기술자를 H_b^i 라 표시하고 이때 $i = 1, \dots, n$ 이다.

소실점이 영상 내에 있다고 가정하여 외곽(border) 블록으로부터 주요한 방향이 유사한 블록을 모두 찾은 뒤, 그중에서 가장 멀리 떨어져 있는 블록을 선택한다.

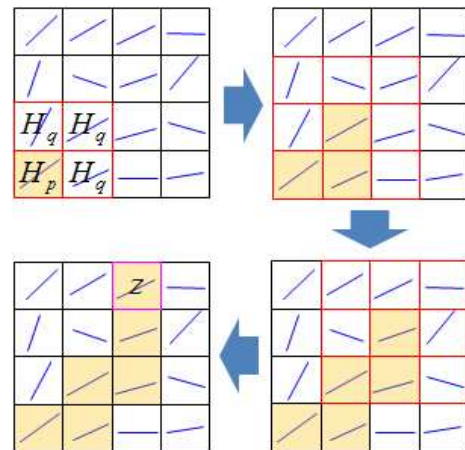


그림 3. 유사 블록 찾는 과정과 가장 멀리 떨어진 최종 블록의 선택

Fig. 3. Finding similar block and selecting a block which has the maximum distance.

먼저 외곽 블록을 p 라하고, 블록 p 의 주변 블록 q 의 집합을 N 이라고 한다. 외곽 블록 p 와 유사한 모든 블록의 집합을 Q 라하고 이를 다음과 같이 나타낸다.

$$Q = \{q | s(H_p, H_q, n) < T, q \in N\}$$

$$Q = \{q_1, \dots, q_{n_q}\} \tag{2}$$

외곽 블록 p 와 주변 블록 q 의 유사도를 측정하여 그 값이 특정 임계값 T 보다 작으면, 블록 q 는 p 와 유사하다고 판단한다. 이렇게 구한 외곽 블록 p 와 유사한 블록의 총 수는 n_q 이다. 외곽 블록 p 와 유사한 정도를 측정하기 위한 식은 다음과 같다.

$$s(H_p, H_q, n) = \sum_{i=1}^n w_i (H_p^i - H_q^i) \tag{3}$$

블록 p 와 블록 q 간의 유사한 정도는 앞서 구한 HOG 기술자에서 n 개의 주요한 방향을 선택하여, 그 주요한 방향의 값이 큰 순서대로 방향의 차이를 구한 후, 가중치를 곱한 값의 합으로 표시할 수 있다. 블록 p 와 블록 q 가 유사할수록 그 값이 작아진다.

나. 소실점 블록 계산

소실점은 실제 3차원 공간이 2차원 영상 형태로 이루어지면서 영상 내에서 가장 멀리 있는 곳에 있으며, 원근감에 의해 영상을 구성하는 주요 에지 성분이 한곳으로 모이는 특성이 있다. 이 특성을 이용하여 소실점의 위치를 다음과 같이 예측할 수 있다.

외곽 블록 p 의 모든 유사한 블록 Q 에 대해서, 블록 p 와 가장 멀리 떨어져 있는 블록 q_i 을 z 라하고, 이를 구하는 식은 다음과 같다.

$$Z = \{z | z = q_i, \text{argmax}_{q_i \in Q} d(p, q)\},$$

$$Z = \{z_1, \dots, z_{n_z}\} \tag{4}$$

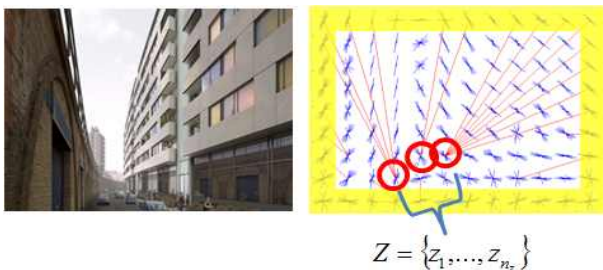


그림 4. 외곽 블록 P 로부터 구한 블록 Z . 블록 Z 의 평균 위치가 소실점이 있는 위치이다

Fig. 4. Block Z from outer block P . Vanishing point locates at the average placement of block Z .

표 2. 각 알고리즘의 수행 시간

Table 2. Processing time.

	제안하는 방법	Rasmussen[6]	Tardif [3]
수행시간	0.061 s	2.187 s	0.81 s

외곽 블록 p 로부터 유사한 블록을 모두 찾아 그중에서 가장 멀리 떨어진 블록 z 을 구하는 과정을 그림 3에 나타내었다.

이렇게 모든 외곽 블록 P 로부터 블록 Z 를 구한 다음, 소실점이 있는 위치 VP 를 식 (5)와 같이 계산하고 이를 그림 4에 나타내었다.

$$VP = \frac{1}{n_z} \sum_{i=1}^{n_z} z_i \tag{5}$$

III. 실험

제안하는 방법을 이용하여 다양한 영상에 대해 소실점 검출 실험을 하였다. 제안하는 방법은 특정 영상군에서만 적용되는 것이 아니라 소실점이 영상 내에 존재

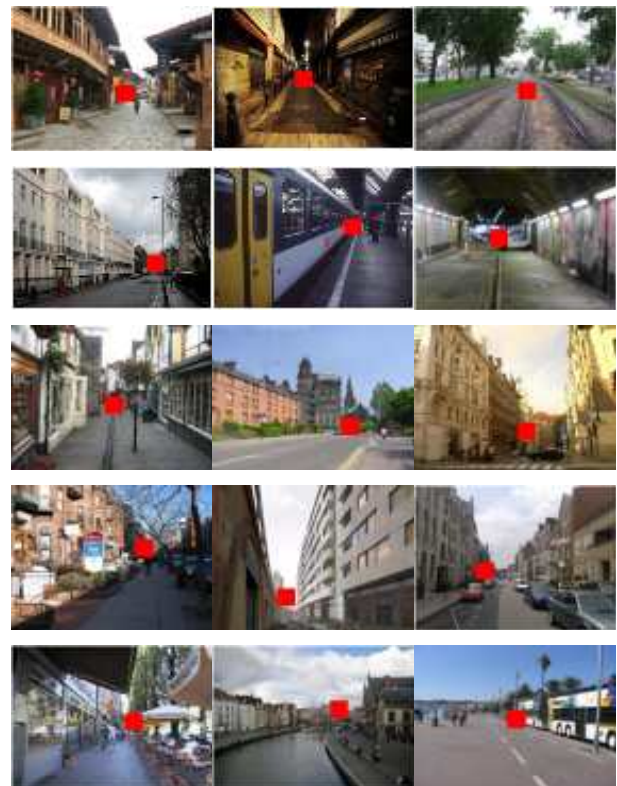


그림 5. 실험 결과. 붉은색 상자로 표시된 블록이 소실점을 포함하는 블록이다

Fig. 5. Experiment result of proposed method. The red box contains vanishing point.

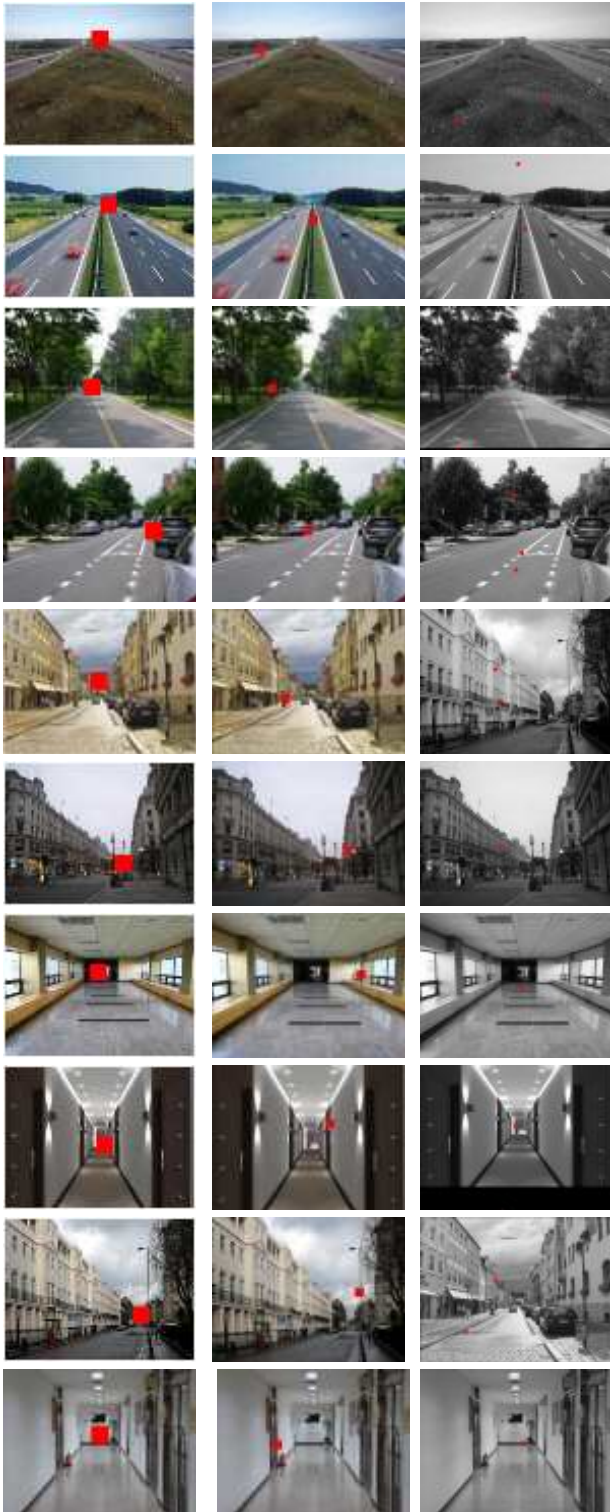


그림 6 첫 째열: 제안하는 방법의 실험결과, 둘째열: [6]의 방법 실험 결과, 셋 째열: [3]의 결과, 소실점의 위치를 붉은색으로 표시하였다.

Fig. 6. First column: Experiment results of the proposed method. Second column: the method in [6], third column: the method in [3]. The red box indicates vanishing point.

하면 실시간으로 소실점의 위치를 예측할 수 있다는 장점이 있다.

소실점 예측은 100개의 일반 영상들에 대해 실험하였다. 이 영상은 실내, 실외, 건축물, 도로 영상 등으로 구성되었다. 영상 대부분은 구글(Google) 이미지에서 다운로드 하였다. 실험에 사용된 영상의 크기는 가로가 400, 세로가 300으로 정규화 하였다.

본 실험에서는 HOG 기술자를 구하기 위해 영상을 32x32 크기의 블록으로 나누었다. HOG 기술자를 구하는 과정에서 방향을 0° ~ 180° 사이를 18등분 하였다. 블록 내의 주요한 방향은 2개를 선택하였다. 여기서 각 방향에 대한 가중치는 그 크기를 비교하여 큰 값을 0.7, 작은 값을 0.3으로 사용 하였다. 블록 간 유사도를 측정하기 위한 임계값 T 는 0.1이다.

본 논문에서 제안하는 방법으로 영상 내의 소실점을 검출한 결과를 그림 5에 나타내었다. 그림 5는 실내 영상, 실외 영상, 도로 영상 등으로 구성되어 있다. 소실점의 위치를 붉은색으로 표시하였다.

제안하는 방법을 가버 필터를 이용하여 도로영상 내의 소실점을 검출하는 방법^[2]과 J-Linkage와 EM 알고리즘을 이용하여 인공(man-made) 영상의 소실점을 검출하는 방법^[3]과 비교하여 그림 6과 같이 나타내고, 수행 시간을 Visual C 2008로 측정한 결과를 표 2로 작성하였다. 제안한 알고리즘의 평균 수행 시간은 영상 한 장당 0.061초가 소요되는 것을 알 수 있다. 그림 6의 결과에서 소실점의 위치는 붉은색으로 표시하였고, 비교 결과 제안하는 방법의 정확도가 높은 것을 알 수 있다.

IV. 결론

본 논문에서는 HOG 기술자를 이용하여 영상 내의 소실점의 위치를 예측하는 방법을 제안하였다. 영상 내의 소실점은 영상의 주요한 에지성분이 한곳에 모이는 특성이 있다. 이 정보를 응용하여 영상의 블록 단위로 HOG 기술자를 구한 뒤, 그 중 주요한 방향 성분만을 가지고 제안하는 동적 프로그래밍 을 이용하여 소실점의 위치를 예측한다. 제안하는 알고리즘은 HOG 기술자를 이용하기 때문에 실시간으로 소실점의 위치를 예측할 수 있다.

참 고 문 헌

- [1] C. Rother, "A new approach to vanishing point detection in architectural environments," Image and Vision Computing, vol. 2, pp.647-655, 2002.
- [2] C. Rasmussen, "Grouping dominant orientations for ill-structured road following," in Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp.I-470-I-477, 2004.
- [3] K. Hui, J. Y. Audibert, and J. Ponce, "Vanishing point detection for road detection," in IEEE Conference on Computer Vision and Pattern Recognition, pp.96-103, 2009.
- [4] S. Battiato, A. Caprab, S. Curtib, M. La Casciac, "3D Stereoscopic Image Pairs by Depth-Map Generation," Proceedings of the 3D Data Processing, Visualization, and Transmission, 2nd International Symposium, pp.124-131, 2004.
- [5] B. Caprile and V. Torre, "Using vanishing point for camera calibration," International Journal of Computer Vision, vol. 4, pp.127-140, 1990.
- [6] Y. Horry, K. Anjyo, K. Arai, "Tour into the picture: using a spidery mesh interface to make animation from a single image," International Conference on Computer Graphics and Interactive Techniques, pp.225-232, 1997.
- [7] S. T. Banard, "Interpreting perspective images," Artificial intelligence, vol. 21, pp. 435-462, 1983.
- [8] V. Cantoni, L. Lombardi, M. Porta, and N. Sicard, "Vanishing point detection: representation analysis and new approaches," International Conference on Image Analysis and Processing, pp.90-94, Sep, 2001.
- [9] T. Tuytelaars, L. V. Gool, M. Proesmans, and T. Moons, "The cascaded hough transform as an aid in aerial image interpretation," IEEE International Conference on Computer Vision, pp67-72, 1998.
- [10] J.-P. Tardif, "Non-iterative approach for fast and accurate vanishing point detection," in IEEE International Conference on Computer Vision, pp.1250-1257, 2009.
- [11] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp.886-893, 2005.

저 자 소 개



최 지 원(학생회원)
 2004년 경북대학교 전자전기컴퓨터 학부 학사 졸업.
 2008년~현재 한국과학기술원 전기및전자공학과 석박사통합과정.

<주관심분야 : 영상처리, 객체추출, 3D 영상처리>



김 창 익(정회원)-교신저자
 1989년 연세대학교 전기공학과 학사 졸업.
 1991년 포항공과대학교 전자전기공학과 석사 졸업.
 1991년~1997년 SKC Ltd. R&D 센터 선임연구원

2000년 워싱턴주립대학교 전기공학과 박사 졸업.
 2000년~2005년 Senior member of Technical Staff, Epson Palo Alto LABORATORY, Epson R&D Inc.

2005년~2009년 한국정보통신대학교 부교수.

2009년~현재 한국과학기술원 부교수

<주관심분야 : 3D 영상처리, 패턴인식, 영상이해>