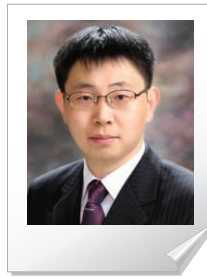


특집 2

자동차용 스테레오 비전 개발동향

Stereo Vision Systems for Automotive Applications



서재규 • 한양대학교
Jae Kyu Suhr • Hanyang University



정호기 • 한양대학교
Ho Gi Jung • Hanyang University

1. 서론

스테레오 비전은 카메라 두 대로 구성되어 있는 스테레오 카메라를 사용하여 주변 환경을 인식하는 방식을 의미한다. 스테레오 카메라는 삼각법을 기반으로 거리 정보를 획득할 수 있는 센서로서 영상과 거리 정보를 동시에 제공할 수 있다는 장점을 갖는다. 이 방식은 모노(단안) 카메라 사용 방식에 비해 하드웨어 가격 및 알고리즘 복잡도가 높다는 단점을 갖지만, 최근 하드웨어 가격의 하락, 프로세서 성능의 향상, 효율적인 알고리즘의 개발로 인해 장점이 부각되고 있는 상황이다.

최근, 스테레오 비전이 근중거리 주행 환경 인식 분야에서 부각되는 이유는 크게 두 가지로 설명될 수 있다. 첫 번째 이유는 스테레오 비전이 자율주행 측면에서 3D Scanning Lidar와 단안 카메라의 조합을 대체하는 용도로 주목 받고 있기 때문이다. Google과 대부분의 대학들은 자율주행 자동차의 주행 환경 인식 목적으로 3D Scanning Lidar와 단안 카메라의 조합을 선호하고 있다^①.

하지만 완성차 및 자동차 부품 업체에서는 센서의 높은 가격, 구동 장치로 인한 약한 내구성, 차량 외관을 심하게 변형시켜야 하는 센서 부착 방법 등의 이유로 3D Scanning Lidar의 사용을 기피하고 있는 상황이다. 이러한 이유로 차량용 인식 시스템 개발의 선두 업체인 독일 Daimler의 경우, 3D Scanning Lidar와 카메라 조합을 스테레오 비전으로 대체하여 103km 완전 자율주행을 수행하였다^②.

〈그림 1〉은 대표적인 자율주행 차량인 Google Car와 Daimler Bertha를 보여준다. 차량 디자인, 센서 가격, 내구성 측면에서 Daimler Bertha가 선호됨을 쉽게 알 수 있다.

스테레오 비전이 부각되는 두 번째 이유는 ADAS(Advance Driver Assistance System) 측면에서의 필요성 때문이다. 유럽 자동차 안전성 등급 판정기관인 Euro NCAP(New Car Assessment Programme)은 AEB(Autonomous Emergency Braking System)를 자동차 안전등급 판정항목으로 추가할 예정이다. 이를 제공하지 않을 경우 최고 등급을 받지 못할 것으로 보이며, 단계적인 의무 장착 규제가 진행될 가능성 또한 높다. 다양한 종류의 장애물을 매우 낮은 오인식률로 인지해야 하는 AEB의 특성으로 인해, 3차원 거리 정보와 물체의 형상 정보를 동시에 제공하는 스테레오 카메라가 이를 위한 효율적인 센서로 대두되고 있다. 또한 추후 높아지게 될 안전성 평가 기준을 만족하기 위해서 스테레오 비전 센서의 필요성이 높아지고 있다.



(a) Google car^①



(b) Daimler Bertha^②

〈그림 1〉 대표적인 자율주행 차량인 Google Car와 Daimler Bertha

2. Stereo Matching

스테레오 카메라로부터 거리 정보를 획득하기 위해서는 두 카메라의 내/외부 변수를 추정하는 카메라 칼리브레이션 과정과 두 카메라에서 획득된 영상 상에서 서로 대응되는 위치를 찾는 스테레오 매칭 과정이 반드시 필요하다. 카메라 칼리브레이션의 경우 Offline으로 비교적 간단하게 수행될 수 있기 때문에 오랜 기간 구동하면서 발생하는 카메라의 작은 변화를 Online으로 보정하는 것이 주요 연구 대상이다. 스테레오 매칭은 오랜 기간 연구되어온 주제로 다양한 방법들이 존재하며 많은 연산량을 필요로 한다. 이러한 연산량의 문제로 인해 과거에는 세로 Edge만을 매칭하는 것과 같은 단순한 방식들이 주로 개발되었다. 하지만 최근 효율적인 알고리즘들의 개발 및 프로세서의 성능의 향상으로 인해 영상 전체를 매칭하여 사용하는 방법이 선호되고 있다. 이렇게 계산된 조밀한 스테레오 매칭 결과를 Dense Disparity Map이라고 부른다.

자동차용 스테레오 비전 연구자들이 다양한 스테레오 매칭 방법을 비교 분석하였으며, 그 결과 SGM(Semi-Global Matching)이 성능 및 연산량을 모두 고려하였을 때 가장 합리적인 방법으로 알려지게 되어 자동차 응용에서 널리 사용되고 있다. 이 방식은 Benz S-class의 스테레오 비전 시스템에 양산된 방식

으로 알려져 있다. SGM은 픽셀 단위의 매칭 Cost와 다수의 1D Constraint로 근사화된 2D Smoothness를 최소화함으로써 스테레오 매칭을 수행하는 방법이다. <그림 2>는 스테레오 카메라의 좌측 영상과 SGM을 통해 계산된 Dense Disparity Map의 예를 보여준다.



(a) 스테레오 카메라 좌측 영상⁹⁾



(b) SGM으로 생성된 Dense Disparity Map⁹⁾

<그림 2> 스테레오 카메라의 좌측 영상과 SGM을 통해 계산된 Dense Disparity Map의 예

3. General Obstacle Detection (GOD)

GOD란 주행 상황에 존재하는 일반적인 장애물을 검출하는 방식을 의미한다. 단안 카메라를 사용하는 모노 비전은 영상의 밝기 정보만을 제공하기 때문에 보행자, 차량과 같이 유형이 정해진 물체의 외형을 학습하는 방법으로만 물체를 검출할 수 있다. 이에 반해 스테레오 비전은 조밀한 거리 정보를 제공하기 때문에 물체의 유형과 관계없이 거리 정보를 기반으로 장애물을 인식할 수 있다. 이러한 GOD는 모노 비전에 비해 스테레오 비전이 갖는 가장 큰 장점 중 하나이다.

GOD는 중간수준표현(Medium-level Representation)으로 볼 수 있으며, 이에 대한 연구는 다음과 같은 중요성을 갖는다. 다양한 응용 시스템들(보행자, 차량, 차선 인식기 등)이 모두 스테레오 비전 센서로부터 저수준(Low-level)의 정보(Color Image, Dense Disparity Map)를 가져가서 사용하게 되면, 차량 내부 통신으로 전달해야 하는 정보양과 불필요한 중복 연산이 증가하게 된다. 따라서 주행 상황을 간략화하여 다양한 응용 시스템들이 공통적으로 사용할 수 있는 중간수준의 형태로 표현하여 공유하게 되면 위의 문제를 해결할 수 있게 된다.

먼저 GOD 혹은 중간수준표현 단계를 효과적으로 수행하기 위해서는 전방 도로면의 형태를 추정하는 과정이 필요하다⁹⁾. 도로면은 일반적으로 3차원 곡면이다. 따라서 평면, 2차 곡면, B-spline 곡면 등의 3차원 곡면 모델을 사용하여 추정할 수 있다. 하지만 일반적인 포장도로 상황에서는 좌/우 방향의 기울기는 무시 가능하기 때문에 대부분의 경우 2차원 곡선 모델을 사용하여 V-disparity 혹은 YZ-plane 도메인에서 추정한다. 이러한 근사화는 도로면 추정 연산량을 줄여주며, 스테레오 매칭 오류와 장애물에서 생성된 3차원 점들에 대한 강인성을 확보해줄 수 있다는 장점을 갖는다.

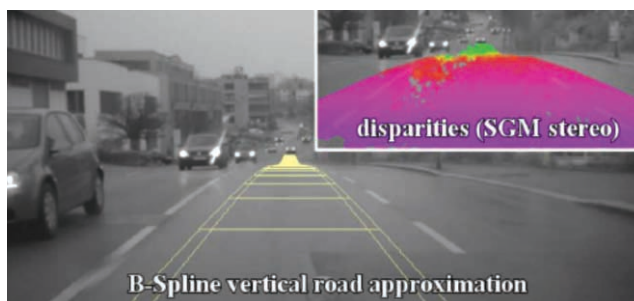
도로면 모델링에 사용되는 2차원 곡선 모델로는 1, 2차 다항식(Polynomial), 구분적 선형(Piecewise Linear) 함수가 있으며, 이를 일반화한 형태인 B-spline 함수도 있다. 이러한 모델들 중 Cubic B-spline 함수가 높은 자유도로 인해 다양한 형태의 도로면을 정확하게 표현할 수 있다고 알려져 있다. 하지만 이는 높은 자유도로 인해 스테레오 매칭 오류와 장애물에서 생성된 3차원 점들에 강인하게 추정되기 힘들다는 한계를 갖기 때문에 이에 대한 추가 연구가 필요한 상황이다.

〈그림 3〉은 도로면 추정 결과의 예를 보여준다. GOD 혹은 중간수준표현으로 사용되는 대표적인 방법은 세 가지가 있다.

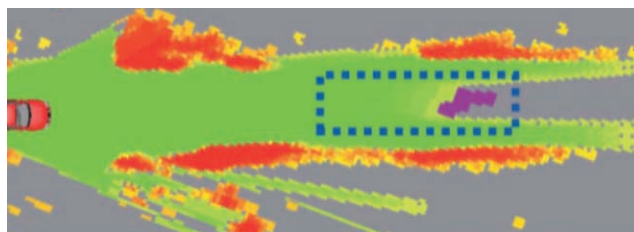
첫 번째는 Occupancy Grid 기반 방법이다^⑤. 이는 전방 주행 환경을 격자로 표현한 후 격자를 구성하는 각 칸이 점유되었을 확률을 스테레오 비전 센서를 기반으로 계산하여 전방 상황을 표현하는 방법으로 로봇틱스 분야에서 오랜 기간 사용되어온 방법이다.

두 번째는 Digital Elevation Map 기반 방법이다^⑥. 이는 주행 환경을 격자로 표현한 후, 격자를 구성하는 각 칸에 존재하는 장애물의 높이를 계산하여 이를 기반으로 전방 상황을 표현하는 방법으로 주로 루마니아의 S. Nedeveschi 교수에 의해 연구되고 있다.

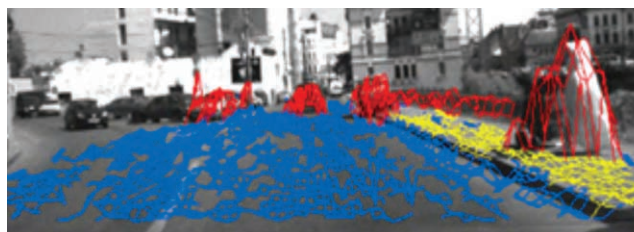
세 번째는 Stixel 기반 방법이다^⑦. 이는 장애물과 도로면의 경계와 장애물의 높이를 추정하여 이를 이어주는 다수의 막대들로 전방 상황을 표현하는 방법으로 Stixel은 Stick과 Pixel의 합성어이다. 이 방법은 주로 Daimler의 U. Franke 팀에 의해 연구되고 있다. 〈그림 4〉는 세 가지 대표적인 GOD



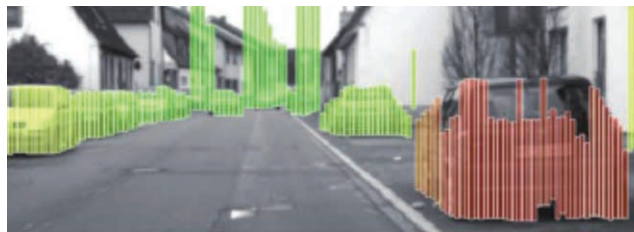
〈그림 3〉 도로면 추정 결과의 예^⑤



(a) Occupancy Grid^⑤

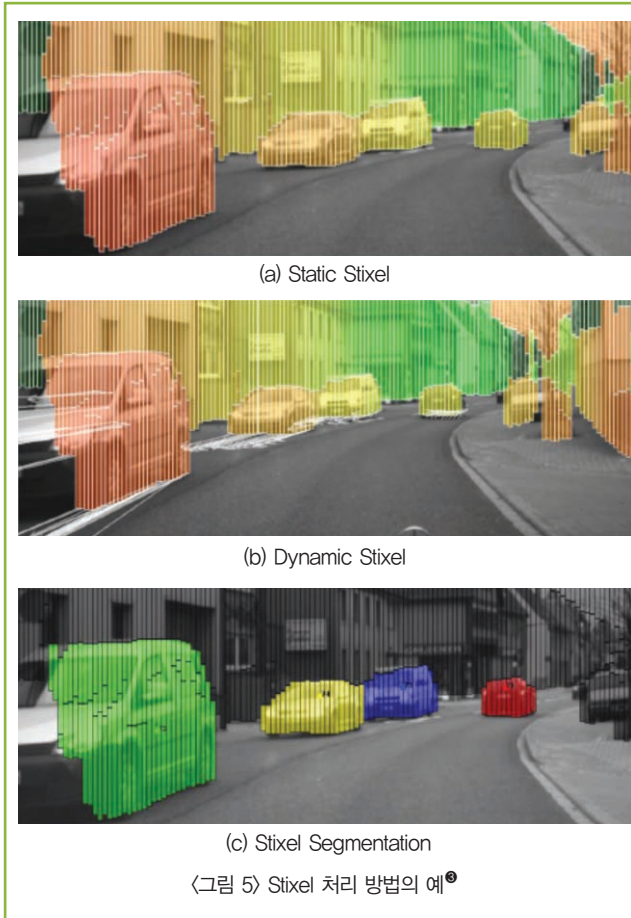


(b) Digital Elevation Map^⑥



(c) Stixel^⑦

〈그림 4〉 대표적인 GOD 방법의 예



방법의 예를 보여준다.

GOD 방법들 중 가장 활발히 연구되고 있는 방법은 독일 Daimler에서 개발 중인 Stixel 기반 방법이다. 이는 매우 많은 정보량을 갖는 Dense Disparity Map을 위치, 높이, 움직임 정보를 갖는 Stixel이라 불리는 소수의 막대로 표현하는 방법으로 정보 축약 측면에서 매우 효율적이라고 알려져 있다.

이 방법은 전방 주행 상황에 존재하는 장애물이 도로면에 수직하게 서 있다는 가정 하에 이를 고정 픽셀 너비의 Stixel로 모델링한다. 그 후 SGM을 통해 생성된 Dense Disparity Map을 사용하여 Stixel의 위치 및 높이를 추정한다. 검출된 Stixel의 움직임은 Dense Optical Flow와 Kalman Filter를 기반으로 추정되며, 이와 같은 움직임 정보를 포함하는 Stixel을 Dynamic Stixel이라고 부른다. 마지막으로 Dynamic Stixel들은 위치, 높이, 움직임의 유사성을 기반으로 구분되어 실세계에서 서로 같은 물체를 구성하는 Stixel들끼리 묶여지게 된다.

〈그림 5〉는 Daimler에서 진행하고 있는 Stixel 처리 방법의 예를 보여준다.

4. Classifier-based Object Detection (COD)

COD란 물체의 외형을 학습시켜 생성한 Classifier를 사용하여 영상에서 차량, 보행자 등을 검출하는 방식을 의미한다. COD 관점에서 스테레오 비전을 사용함으로써 얻을 수 있는 이점은 크게 세 가지로 설명될 수 있다.

첫째, GOD 및 도로면 추정 결과를 사용하여 COD 과정을 효율화할 수 있다^⑦. 도로면 굴곡 정보 혹은 GOD에서 검출된 장애물의 위치 및 높이 정보를 사용하면 영상의 각 위치 별로 예상되는 보행자 혹은 차량의 크기를 예상할 수 있기 때문에 COD를 적용할 관심 영역을 제한할 수 있다. 보다 적극적으로는 GOD 결과를 기반으로 장애물 뒤 혹은 장애물 사이와 같은 위험 지역을 검출하고, 그 위치에만 COD를 적용할 수도

있다. 이러한 GOD 기반 관심 영역 설정 방법은 COD 과정의 연산량을 줄여줄 수 있을 뿐만 아니라 오검출을 낮춰주는 효과 또한 가져올 수 있다.

둘째, Dense Disparity Map과 영상의 밝기 정보를 융합하여 COD에 사용되는 Classifier의 성능을 향상시킬 수 있다. 일반적으로 모노 비전을 사용하는 경우에는 영상의 밝기 정보를 사용하여 COD에 사용되는 Classifier를 학습한다. 스테레오 비전의 경우에는 영상의 밝기 정보뿐만 아니라 거리에 대한 정보를 담고 있는 Dense Disparity Map을 함께 제공하기 때문에 이 둘을 융합하여 Classifier의 성능을 향상시킬 수 있다. 보행자 검출의 경우, 영상의 밝기 정보와 Dense Disparity Map에 각각 Classifier를 적용하여 그 결과를 융합하는 방법이 인식률을 향상시킨다는 결과가 보고되었다⁹⁾.

셋째, 원거리 차량의 거리 추정 정확도를 높일 수 있다. 모노 비전의 경우에는 Classifier로 인식된 차량의 거리를 추정하기 위해 차량의 너비 혹은 하단 위치 정보를 사용한다. 하지만 이는 원거리에서 심한 거리 오차를 발생시킨다는 한계를 갖는다.

GOD의 경우에도 Disparity Map의 분별력이 원거리로 갈수록 급격히 저하되기 때문에 원거리 물체 검출에 실패할 가능성이 높으며, 원거리 물체가 검출된다 하더라도 Disparity Map의 Resolution 한계로 인하여 올바른 거리 정보를 추출하기 힘들게 된다. 이러한 한계는 COD 방법으로 먼저 물체 위치를 인식하고 해당 물체 영역에만 정밀한 스테레오 매칭 방법을 적용하여 거리 정보를 획득함으로써 보완 가능하다⁹⁾.

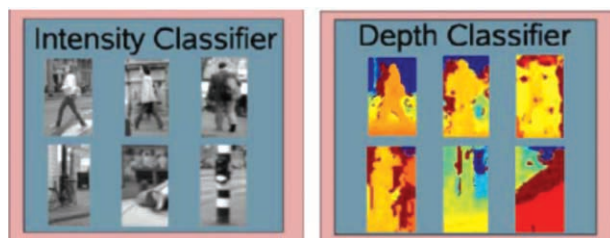
〈그림 6〉은 스테레오 비전을 기반으로 COD를 수행하는 방법의 예를 보여준다.

5. 적용 사례

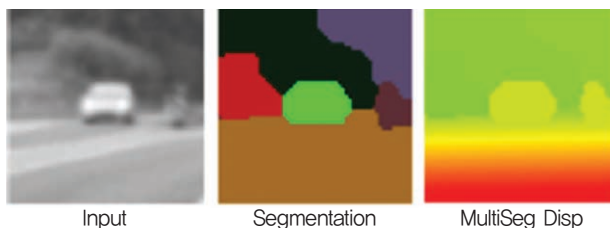
본 고에서는 Daimler에서 Benz S-class와 자율주행 차량 Bertha에 스테레오 비전을 적용한 사례를 소



(a) 관심영역 설정을 통한 효율화⁹⁾



(b) 영상과 거리 정보 융합을 통한 Classifier 성능 향상⁹⁾



(c) 원거리 차량의 거리 추정 정확도 향상⁹⁾

〈그림 6〉 스테레오 비전 기반 COD의 예

개한다. Benz S-class에 양산된 스테레오 비전 센서는 다음과 같은 다양한 기능을 수행한다⁹⁾.

첫 번째는 보행자, 전방 차량, 교행 차량에 대한 제동 보조 기능이다. 이 기능에서 스테레오 카메라는 레이더와 융합되어 보행자 및 전방 차량을 검출하는 용도와 교차로에서 교행하는 차량과의 위험 상황을 예측하는 용도로 사용된다.

두 번째는 ACC(Adaptive Cruise Control)와 Stop & Go 기능이다. 이 기능에서는 차선과 전방 차량이 스테레오 카메라로 인식되어 종/횡방향 제어를 보조하는데 사용된다. 마지막은 Active Suspension 기능이다. 이는 스테레오 카메라에서 획득한 조밀한 거리 정보로부터 근거리 도로면의 굴곡을 추정하여 현가시스템의 제어를 최적화한다.

Daimler의 자율주행 차량인 Bertha는 2013년 8월 103km의 완전 자율주행을 수행하였다⁹⁾. 이 차량은 주행 환경 인식을 위해 스테레오 및 모노 카메라, 레이더, 정밀 지도를 사용하며, 스테레오 카메라는 주행가능 영역 검출과 차량 및 보행자 인식에 이용된다. 주행 가능 영역은 Stixel 검출 및 추적을 실시한 결과를 기반으로 장애물이 존재하지 않는 영역을 계산하여 얻어지며, 차량 및 보행자는 영상의 밝기 정보와 Dense Disparity Map 정보의 융합을 통해 학습된 Classifier를 사용하여 검출된다. 이 때 근중거리 차량 인식의 관심 영역은 Stixel 생성 결과를 기반으로 설정되고, 원거리 차량 인식의 관심 영역은 영상 밝기 기반 차량 검출기를 기반으로 설정되며, 보행자 인식의 관심 영역은 도로면 추정 결과를 기반으로 설정된다.

6. 개발 방향

장기적으로 하드웨어의 저가화와 실시간 스테레오 매칭의 범용화가 진행되고 신차 안전성 평가 기준과 ADAS 의무 장착 요구가 높아지게 되면, 기존의 모노 비전 기반 방법들이 스테레오 비전 기반 방법으로 대체될 가능성이 높아지게 될 것으로 예상된다. 이에 따라 국내에서도 스테레오 비전에 대한 연구가 활발히 수행될 것으로 기대된다. 앞서 설명한 내용을 기반으로 현재 진행 중이고 추후에 진행될 스테레오 비전 기반 주행 환경 인식에 대한 연구 개발 방향을 정리하면 다음과 같다.

단기적으로는 스테레오 매칭과 Optical Flow 등의 기반 알고리즘에 대한 연구가 필요하고, 중기적으로는 주행 상황을 중간수준으로 표현하는 GOD 방법과 영상의 밝기와 거리 정보를 융합하여 개별 물체를 인식하는 COD 방법에 대한 연구가 필요하다. 장기적으로 모노 비전이 수행하던 일들을 스테레오 비전이 대체하여 응용 분야가 늘어나게 되면, GOD와 COD를 효과적으로 통합하여 전체 시스템을 효율화하는 연구가 필요하다. 또한 스테레오 매칭, Optical Flow, GOD, COD 기술들을 하드웨어 기반으로 구현하여 실시간성을 확보할 필요가 있다.

Daimler의 경우 스테레오 매칭과 GOD를 FPGA로, Optical Flow를 GPU로 구현하여 Benz S-class와 자율주행 차량인 Bertha에 적용한 것으로 보고된 바 있다⁹⁾.

〈서재규 교수 : jksuhr@hanyang.ac.kr〉

〈참고문헌〉

- ❶ E. Guizzo, How Google's Self-Driving Car Works, IEEE Spectrum, 2011, <http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works>, 2014.
- ❷ J. Ziegler, P. Bender, M. Schreiber, H. Lategahn, T. Strauss, C. Stiller, T. Dang, U. Franke, N. Appenrodt, C. G. Keller, E. Kaus, R. G. Herrtwich, C. Rabe, D. Pfeiffer, F. Lindner, F. Stein, F. Erbs, M. Enzweiler, C. Knöppel, J. Hipp, M. Hauels, M. Trepte, C. Brenk, A. Tamke, M. Ghanaat, M. Braun, A. Joos, H. Fritz, H. Mock, M. Hein and E. Zeeb, Making Bertha Drive? An Autonomous Journey on a Historic Route, IEEE Intelligent Transportation Systems Magazine, Vol.6, No.2, pp.8-20, 2014.
- ❸ R. Erbs, B. Schwarz and U. Franke, From Stixels to Objects – A Conditional Random Field based Approach, IEEE Intelligent Vehicles Symposium, pp.586-591, 2013.
- ❹ A. Wedel, H. Badino, C. Rabe, H. Loose, U. Franke and D. Cremers, B-Spline Modeling of Road Surfaces With an Application to Free-Space Estimation, IEEE Transactions on Intelligent Transportation Systems, Vol.10, No.4, pp.572-583, 2009.
- ❺ T.-N. Nguyen, B. Michaelis, A. Al-Hamadi, M. Tornow and M. Meinecke, Stereo-Camera-Based Urban Environment Perception Using Occupancy Grid and Object Tracking, IEEE Transactions on Intelligent Transportation Systems, Vol.13, No.1, pp.154-165, 2011.
- ❻ F. Oniga, and S. Nedeveschi, Processing Dense Stereo Data Using Elevation Maps: Road Surface, Traffic Isle, and Obstacle Detection, IEEE Transactions on Vehicular Technology, Vol.59, No.3, pp.1172-1182, 2010.
- ❼ M. Enzweiler, M. Hummel, D. Pfeiffer and U. Franke, Efficient Stixel-Based Object Recognition, IEEE Intelligent Vehicles Symposium, pp.1066-1071, 2012.
- ❽ C. G. Keller, M. Enzweiler, M. Rohrbach, D. F. Llorca, C. Schnorr and D. M. Gavrilu, The Benefits of Dense Stereo for Pedestrian Detection, IEEE Transactions on Intelligent Transportation Systems, Vol.12, No.4, pp.1096-1106, 2011.
- ❾ P. Pinggera, U. Franke and R. Mester, Highly Accurate Depth Estimation for Objects at Large Distances, Lecture Notes in Computer Science, Vol.8142, pp.21-30, 2013.
- ❿ U. Franke and S. Gehrig, How Cars learned to See, The 54th Photogrammetric Week, pp.3-10, 2013.

한국자동차공학회
회원가입 및 논문집 구독문의

www.KSAE.org

- Tel : 02-564-3971
- Fax : 02-564-3973
- e-mail : account@ksae.org

자동차공학은 한국의 힘!