

영상 영역 특징 추가 및 유전 알고리즘 기반 최적화를 통한 스틱셀 분할 개선 방법

이 선 영¹⁾ · 서 재 규¹⁾ · 정 호 기^{*2)}

한양대학교 자동차전자제어연구소 · 한양대학교 미래자동차학과

Improvement of Stixel Segmentation Using Additive Image Domain Features and Genetic Algorithm-based Optimization

Sunyoung Lee¹⁾ · Jae Kyu Suhr¹⁾ · Ho Gi Jung^{*2)}

¹⁾Research Institute of Automotive Control and Electronics, Hanyang University, Seoul 04763, Korea

²⁾Department of Automotive Engineering, Hanyang University, Seoul 04763, Korea

(Received 28 October 2014 / Revised 4 May 2015 / Accepted 7 July 2015)

Abstract : Recently, a medium-level representation named “Stixel” has been extensively researched in stereo vision-based environmental perception. Obstacle detection using Stixel representation consists of three steps: static Stixel generation, dynamic Stixel generation, and Stixel segmentation. This paper focuses on the Stixel segmentation step and has two contributions. One is that it shows that Stixel segmentation performance can be enhanced by utilizing both image domain and real world domain features. The other is that it suggests that parameters used for Stixel segmentation can be effectively tuned based on genetic algorithm. The proposed method was quantitatively evaluated and the result showed that the proposed method increased Stixel segmentation accuracy compared with the previous method.

Key words : Stixel segmentation(스티셀 분할), Stixel matching(스티셀 정합), Stereo vision(스테레오비전), Driver assistant system(운전자보조시스템), Genetic algorithm(유전 알고리즘)

1. 서 론

스테레오 비전은 영상 정보와 거리 정보를 동시에 제공한다는 장점으로 인하여 최근 주행 환경 인식에 널리 사용되고 있다. 스테레오 비전 기반 주행 환경 인식을 수행하는 경우, 다양한 응용 시스템들(예를 들면, 보행자, 차량, 차선 인식기)이 모두 저수준(low-level)의 영상 정보와 거리 정보를 전달 받아 사용하게 되면, 불필요한 중복 연산이 증가하고 차량 내부 통신으로 전달해야 하는 정보의 양이 증가하게 된다. 이러한 문제를 해결하기 위해 주행 상황을 간략화하여 다양한 응용 시스템들이 공통적으로 사용할 수 있는 형태로 표현하는 방법에 대한 연구가

수행되고 있으며, 이를 중간수준표현 (medium-level representation) 방법이라고 부른다. 스테레오 비전 기반 중간수준표현 방법으로 사용될 수 있는 대표적인 방법으로는 점유 격자(occupancy grid),¹⁾ 디지털 고도 맵(digital elevation map),²⁾ 스틱셀(Stixel)³⁾ 등이 있다. 이들 중, 다임러(Daimler)에서 제안된 스틱셀 기반 방법이 가장 활발하게 연구되고 있으며,³⁾ 다임러의 자율주행 차량인 Bertha에 적용되어 그 효율성을 증명한 바가 있다.⁴⁾ 이 방법은 전방 상황을 위치와 높이 정보를 갖는 스틱셀(stick와 pixel의 합성어)이라 불리는 다수의 막대로 표현한다.

스티셀 기반 주행 환경 표현 방법은 크게 세 단계로 구성된다. 첫 번째 단계에서는 일차 장애물 경계

*Corresponding author, E-mail: hogijung@hanyang.ac.kr

와 이의 높이를 추정하여 스틱셀을 생성하고,³⁾ 두 번째 단계에서는 생성된 스틱셀의 영상 상에서의 움직임 추정하며,⁵⁾ 마지막 단계에서는 스틱셀의 위치 및 움직임 정보를 기반으로 실제 세계에서 동일한 물체를 구성하는 스틱셀들을 묶어주는 과정을 수행한다.⁶⁾ 일반적으로 움직임 정보가 포함된 스틱셀을 동적 스틱셀(dynamic Stixel)이라고 부르고, 스틱셀들을 묶어주는 과정을 스틱셀 분할(Stixel segmentation)이라고 부른다.

본 논문은 위와 같이 구성된 스틱셀 기반 주행 환경 표현 단계 중 스틱셀 분할 과정의 성능을 향상시키는 방법을 제안한다. 스틱셀 분할 방법은 자율학습(unsupervised) 방법⁶⁾과 지도학습(supervised) 방법⁷⁾으로 구분할 수 있다. 자율학습 방법은 유사한 특징을 갖는 스틱셀들끼리 묶어주는 군집화 방법을 기반으로 스틱셀 분할을 수행하고, 지도학습 방법은 수동으로 라벨이 입력된 학습용 스틱셀을 사용하여 학습된 스틱셀의 특징과 라벨의 대응 관계 기반으로 스틱셀의 라벨을 설정함으로써 스틱셀 분할을 수행한다. 이 중 지도학습 방법은 사전 정보 학습을 위해 매우 많은 수의 데이터를 필요로 하고 처리 속도가 느리다는 한계를 갖는다. 따라서 본 논문은 실시간 처리 측면에서 유리한 자율학습 기반 방법으로 스틱셀 분할을 수행하였다. 다임러는 스틱셀 분할을 위해 DBSCAN(Density Based Spatial Clustering of Applications with Noise)⁸⁾이라는 군집화 방법을 사용하였다.⁶⁾ 이 방법은 노이즈를 포함한 공간 데이터를 다루는데 적합하며, 미리 군집(cluster)의 개수를 지정해줄 필요가 없으며 다양한 모양과 크기의 군집을 구분할 수 있다는 장점을 가진다.

본 논문은 DBSCAN 기반 스틱셀 분할의 성능을 개선하는 방법을 제안하였다. 기존 논문에 비해 본 논문이 갖는 차별점은 다음과 같다.

- 1) 스틱셀의 위치 및 움직임 정보와 더불어 외양(appearance) 정보 및 높이 정보를 함께 사용하는 방법을 제안하여 스틱셀 분할 성능을 향상시켰다.
- 2) 스틱셀 분할에 필요한 다수의 파라미터(parameter)들을 유전 알고리즘(genetic algorithm)에 의해 최적화하는 방법을 제안하여 스틱셀 분할 성능을 향상시켰다.

다양한 상황에서 획득된 데이터를 사용하여 정량적으로 비교 평가한 결과 기존 방법⁶⁾에 비해 제안된 방법이 향상된 스틱셀 분할 성능을 보이는 것을 확인할 수 있었다.

2. 스틱셀 기반 주행 환경 인식

Fig. 1은 스틱셀 기반 주행 환경 인식의 일반적인 수행 단계를 보여준다. 본 논문에서는 적은 연산량으로 조밀한 거리 정보를 획득할 수 있는 ELAS(Efficient Large-Scale Stereo Matching) 방법⁹⁾을 통해 조밀 디스패리티 맵(dense disparity map)을 획득하여 스틱셀 기반 주행 환경 인식을 수행하였다.

스티셀 기반 주행 환경 인식의 첫 번째 단계는 도로면 추정 단계이다. 스틱셀은 도로면에 수직인 막대로 정의되므로 도로면 추정 결과는 스틱셀 생성 결과에 직접적인 영향을 미치게 된다. 도로면은 3차원 평면¹⁰⁾ 혹은 곡면²⁾으로 표현되기도 하지만, 좌우 방향 높이 차이가 작다고 가정하고 직선,¹¹⁾ 2차 함수,¹²⁾ B-spline 함수¹³⁾와 같은 2차원 선으로 표현되는 경우가 보편적이다. 본 논문에서는 도로면 추정을 위해 B-spline 함수 기반 방법을 개선하여 사용하였다.

두 번째 단계는 일차 장애물과 도로면이 만나는 위치인 장애물 경계(free space)를 검출하는 단계이다. 스틱셀 생성 관점에서 장애물 경계는 스틱셀의 하단 위치를 결정하는데 사용된다. 과거에는 연산량의 한계로 인하여 조밀 디스패리티 맵을 계산하지 않는 방법들이 연구되어 왔다.¹⁴⁾ 하지만 최근에는 실시간 조밀 디스패리티 맵 생성이 가능해짐에 따라 이를 사용하는 방법들이 연구되고 있으며, 이들 중 u-디스패리티(u-disparity)와 동적 계획법(dynamic programming)을 사용한 방법이 가장 널리 사용되고 있다.³⁾ 본 논문에서도 이 방법을 사용하여 장애물 경계를 추정하였다.

세 번째 단계는 장애물 경계에 위치하는 장애물의 높이를 알아내는 장애물 높이 추정 단계이다. 스틱셀 생성 관점에서 장애물 경계는 스틱셀의 상단 위치를 결정하는데 사용된다. 장애물의 높이추정에 사용되는 대표적인 방법으로는 디지털 고도 맵 기반 방법²⁾과 멤버십 투표(membership voting) 기반 방

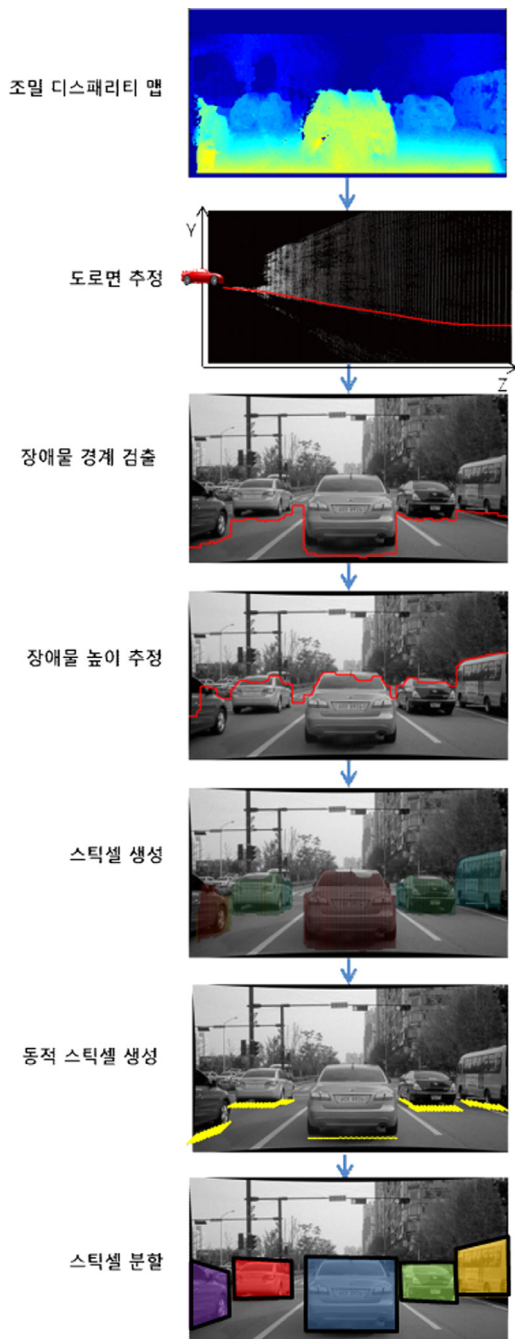


Fig. 1 Flowchart of Stixel-based driving environment recognition system

법³⁾이 있다. 디지털 고도 맵 기반 방법은 각 위치 별로 가장 높은 곳에 위치하는 3차원 점의 높이를 해당 위치의 장애물 높이로 설정하기 때문에 스테레

오 매칭 오류와 공중에 떠서 존재하는 장애물로 인해 오류를 발생시킨다. 따라서 본 논문에서는 멤버십 투표 기반 방법으로 장애물의 높이를 추정하였다.

네 번째 단계는 장애물 경계 검출 결과와 장애물 높이 추정 결과를 사용하여 스틱셀을 생성하는 단계이다. 본 논문에서는 스틱셀의 너비를 5픽셀로 설정하였으며, 스틱셀의 거리는 막대 내부의 거리 정보의 중간값(median)으로 설정하였다. 이와 같은 스틱셀 생성 과정을 거치면 주행 상황을 표현하는데 필요한 데이터의 양을 대폭 줄일 수 있게 된다. 예를 들어, 640 × 480픽셀의 영상의 경우 307,200개의 밝기 값과 거리 값을 단지 128개의 스틱셀로 표현할 수 있게 된다.

다섯 번째 단계는 연속 영상에서 스틱셀의 움직임을 추정하는 단계이다. 이와 같이 움직임 정보가 포함된 스틱셀은 동적 스틱셀이라고 불린다. 동적 스틱셀을 생성하는 방법은 조밀 광류(dense optical flow) 기반 방법¹⁵⁾과 스틱셀 매칭(Stixel matching)⁵⁾ 기반 방법이 있다. 조밀 광류 기반 방법은 성능 측면에서 우수하지만 조밀 광류 계산에 많은 연산량이 소요된다는 한계를 갖는다. 따라서 본 논문에서는 움직임 추정 정확도는 상대적으로 낮지만 연산량 측면에서 유리한 스틱셀 매칭 방법을 사용하여 동적 스틱셀을 생성하였다.

마지막 단계는 실세계에서 동일한 물체를 구성하는 스틱셀들끼리 묶어주는 스틱셀 분할 단계이다. 스틱셀 분할 방법은 크게 군집화(clustering) 기반 방법⁶⁾과 영상 분할(image segmentation) 기반 방법⁷⁾으로 구분할 수 있다. 이 중 영상 분할 기반 방법은 사전 정보 학습을 위해 매우 많은 수의 데이터를 필요로 하고 처리 속도가 느리다는 한계를 갖는다. 따라서 본 논문은 실시간 처리 측면에서 유리한 군집화 기반 방법으로 스틱셀 분할을 수행하였으며, 이의 성능을 개선하는 방법을 제안하였다.

3. 제안된 스틱셀 분할 방법

본 논문에서는 군집화 기반 스틱셀 분할 방법의 성능을 개선하는 방법을 제안한다. 기존에 제안된 군집화 기반 스틱셀 분할 방법⁶⁾은 실세계에서의 스틱셀 위치와 이동 방향을 특징으로 DBSCAN을 기

반으로 유사한 특징을 갖는 스틱셀들끼리 묶어줌으로써 스틱셀 분할을 수행하였다. DBSCAN 방법은 객체의 밀도에 기초한 군집화 방법으로 노이즈를 포함한 공간 데이터를 다루는데 적합한 군집화 방법이며, K-means 방법과 다르게 미리 군집(cluster)의 개수를 지정해줄 필요가 없으며 다양한 모양과 크기의 군집들을 구분할 수 있다는 장점을 가진다. DBSCAN을 사용한 기존 방법은 두 가지 한계를 갖는다. 첫째, 스틱셀이 포함하고 있는 다양한 특징들 중 실세계에서의 위치와 이동 방향만 사용하기 때문에 분별력의 한계를 갖는다. 둘째, 서로 다른 단위를 갖는 두 특징(거리와 방향)을 통합하여 유사도를 계산하는데 필요한 파라미터와 DBSCAN에 필요한 파라미터를 최적화하는 방법이 제시되어 있지 않다. 이러한 한계를 극복하기 위해서 본 논문은 스틱셀이 포함하는 다양한 특징들을 추가적으로 사용하여 스틱셀 사이 분별력을 높이고, 유사도 계산 및 DBSCAN에 사용되는 파라미터를 최적화하는 방법을 제안하였다.

3.1 군집화에 사용되는 스틱셀 특징 추가

기존에 제안된 군집화 기반 스틱셀 분할 방법⁶⁾은 스틱셀이 포함하고 있는 다양한 특징들 중 두 가지(실세계에서의 스틱셀 위치와 이동 방향)만을 사용하였다. 본 논문에서는 이에 다섯 가지 특징을 추가로 사용하여 총 7 가지 특징을 사용하여 스틱셀 분할을 수행하는 방법을 제안한다. 사용된 스틱셀의 특징은 다음과 같다.

1) 실세계에서의 스틱셀 위치, 이동 방향, 높이

2장에 설명된 방식으로 동적 스틱셀을 생성하게 되면, 각 스틱셀의 3차원 위치를 복원할 수 있게 된다. 이때, 실세계에서 스틱셀의 위치는 영상 상에서 스틱셀의 하단 위치를 3차원으로 복원한 위치가 되고, 스틱셀의 이동 방향은 영상 상에서의 스틱셀의 이동을 나타내는 움직임 벡터의 양 끝 점을 3차원으로 복원한 두 점을 잇는 방향이 되며, 스틱셀의 높이는 스틱셀의 하단과 상단의 위치를 3차원으로 복원한 두 점 간의 거리가 된다. 이와 같이 계산된 스틱셀의 위치를 $X=(X,Y,Z)$, 이동 방향을 나타내는 벡터

를 $V=(dx,dy,dz)$, 스틱셀의 높이를 H 라고 하면, 임의의 두 스틱셀(S_i, S_j)의 위치, 이동 방향, 높이에 대한 상이도는 각각 식 (1)~(3)과 같이 계산될 수 있다.

$$LD(S_i, S_j) = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2 + (Z_i - Z_j)^2} \quad (1)$$

$$AD(S_i, S_j) = \arccos\left(\frac{dx_i dx_j + dy_i dy_j + dz_i dz_j}{\sqrt{dx_i^2 + dy_i^2 + dz_i^2} \sqrt{dx_j^2 + dy_j^2 + dz_j^2}}\right) \quad (2)$$

$$HD(S_i, S_j) = abs(H_i - H_j) \quad (3)$$

위 수식에서 $LD(S_i, S_j)$ 는 두 스틱셀 간의 위치 상이도를 의미하고, $AD(S_i, S_j)$ 는 이동 방향 상이도를 의미하며, $HD(S_i, S_j)$ 는 높이 상이도를 의미한다. 위와 같은 특징들을 사용하여 스틱셀 분할을 수행하는 이유는 실세계에서 동일한 물체를 구성하는 스틱셀들은 3차원 공간에서 유사한 위치에 존재하고, 유사한 방향으로 움직이며, 유사한 높이를 갖기 때문이다.

2) 영상 상에서의 스틱셀 위치, 이동 방향, 높이

제안된 방법은 영상 상에서의 스틱셀의 위치, 이동 방향, 높이에 대한 정보를 스틱셀 분할을 위한 추가적인 정보로 사용한다. 이상적인 상황에서는 영상 상에서의 스틱셀 정보는 실세계에서의 스틱셀 정보와 동일한 정보를 가지기 때문에 중복되어 사용될 필요가 없다. 하지만 실제 상황에서는 3차원 복원 오차가 존재하기 때문에 영상 상에서의 스틱셀 정보가 추가적인 정보를 포함한다. 즉, 동일한 물체를 구성하는 스틱셀이 영상 상에서는 인접해서 존재하지만, 스테레오 매칭 오류로 인해 3차원 복원 후에는 두 스틱셀의 위치가 매우 달라질 수 있게 된다. 영상 상에서 위치를 $x=(u,v)$, 이동 방향을 나타내는 벡터를 $v=(du,dv)$, 스틱셀의 높이를 h 라고 하면, 임의의 두 스틱셀(S_i, S_j)의 위치, 이동 방향, 높이에 대한 상이도는 각각 식 (4)~(6)과 같이 계산될 수 있다.

$$ld(S_i, S_j) = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (4)$$

$$ad(S_i, S_j) = \arccos\left(\frac{du_i du_j + dv_i dv_j}{\sqrt{du_i^2 + dv_i^2} \sqrt{du_j^2 + dv_j^2}}\right) \quad (5)$$

$$hd(S_i, S_j) = abs(h_i - h_j) \quad (6)$$

위 수식에서 $ld(S_i, S_j)$ 는 두 스틱셀 간의 위치 상이도를 의미하고, $ad(S_i, S_j)$ 는 이동 방향 상이도를 의미하며, $hd(S_i, S_j)$ 는 높이 상이도를 의미한다. 위와 같은 특징들을 사용하여 스틱셀 분할을 수행하는 이유는 실제 세계에서 동일한 물체를 구성하는 스틱셀들은 영상에서 유사한 위치에 존재하고, 유사한 방향으로 움직이며, 유사한 높이를 갖기 때문이다.

3) 스틱셀의 외양(appearance) 정보

제안된 방법은 효과적인 스틱셀 분할을 위해서 스틱셀 내부의 외양 정보를 추가적인 정보로 사용한다. 외양 정보를 추출하기 위해서는 컴퓨터 비전 분야에서 특징 추출 용도로 광범위하게 사용되고 있는 LBP(Local Binary Pattern)¹⁶⁾ 방식을 사용하였다. LBP는 특정 픽셀과 인접 픽셀들 간의 밝기 차이를 사용하여 해당 픽셀을 표현하는 방식으로, 밝기 차이의 유형에 따라 해당 픽셀에 코드를 부여하는 형태로 구현된다. 예를 들어 특정 픽셀 주위의 3×3 픽셀의 밝기값이 Fig. 2의 좌측 같은 경우, 주변 픽셀의 밝기값들을 중앙 픽셀의 밝기값을 임계치로 사용하여 이진화시켜 중앙과 같은 결과를 획득한다. 그 후, 이진화 결과를 우측과 같이 일렬로 나열하여 해당 이진수에 대한 코드를 중앙 픽셀에 부여한다. 본 논문에서는 스틱셀 영역에서 LBP 코드를 생성한 후에 이들의 히스토그램(histogram)을 생성하여 해당 스틱셀의 특징으로 사용하였다. 이와 같은 방법으로 생성된 LBP 히스토그램을 \mathbf{b} 라고 하면, 임의의 두 스틱셀(S_i, S_j)의 외양 상이도는 식 (7)과 같이 맨하탄 거리(Manhattan distance)에 의해 계산될 수 있다.

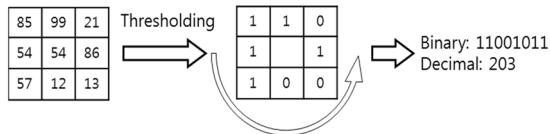


Fig. 2 Local binary pattern code assignment

$$MD(S_i, S_j) = \sum_{n=1}^N |b_i(n) - b_j(n)| \quad (7)$$

위 수식에서 $MD(S_i, S_j)$ 는 두 스틱셀 간의 외양 상

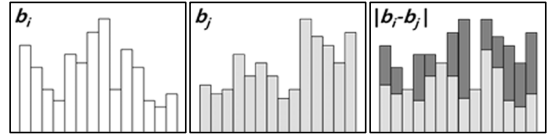


Fig. 3 Local binary pattern code assignment

이도를 의미하며, $b_i(n)$ 는 i 번째 스틱셀 영역에서 추출된 LBP 히스토그램의 n 번째 칸의 누적값을 의미하고, N 은 LBP 히스토그램의 차원을 의미한다. 본 논문에서는 8개의 코드를 갖는 LBP를 사용하였으므로 N 을 59로 설정하였다. 식 (7)의 의미는 Fig. 3으로 설명될 수 있다. 좌측과 중앙은 각각 서로 다른 스틱셀에서 추출된 LBP 히스토그램을 의미한다. 우측은 좌측과 중앙의 LBP 히스토그램을 겹쳐 놓은 것으로 이 그림에서 진하게 표시된 부분이 두 히스토그램의 차이를 의미한다. 맨하탄 거리는 이 차이의 합으로 정의된다. 따라서 맨하탄 거리가 클수록 두 스틱셀의 상이도가 커지게 된다.

앞서 설명된 7개의 상이도 값들은 가질 수 있는 값의 범위가 상이하기 때문에 min-max 정규화(normalization) 방법에 의해서 모두 0에서 1 사이의 값을 갖도록 정규화된 후 사용되었다.

제안된 방법에서 사용한 min-max 정규화 방식은 각 특징들이 가질 수 있는 값의 최소값 및 최대값을 이용하여 각 특징의 상이도 값들이 0에서 1 사이의 값을 갖도록 만든다. 한 특징의 상이도 x 가 가질 수 있는 최소값을 Min_x , 최대값을 Max_x 라 하면, 정규화한 후의 값 $N(x)$ 은 식 (8)과 같다.

$$N(x) = \frac{(x - Min_x)}{(Max_x - Min_x)} \quad (8)$$

3.2 유전 알고리즘 기반 파라미터 최적화

본 논문에서는 기존 방법과 동일하게 DBSCAN 기반 군집화를 사용하여 스틱셀 분할을 수행한다. DBSCAN 기반 군집화는 노이즈를 포함한 공간 데이터를 다루는데 적합한 방법이며, 미리 군집의 개수를 지정해줄 필요가 없고, 다양한 모양과 크기의 군집을 구분할 수 있다는 장점을 가진다. DBSCAN 기반 군집화는 특정 샘플(sample)의 반경(r) 안에 최소 개수(m) 이상의 샘플이 존재하는가를 확인하는 방식으로 동작한다. 따라서 이 방식을 위해서는 두

개의 파라미터(r, m)를 설정할 필요가 있다.

또한, 특정 샘플이 다른 샘플의 반경 안에 들어오는가를 판단하기 위해서는 스칼라 값을 갖는 거리를 정의할 필요가 있다. 본 논문에서는 두 스틱셀 간의 거리를 계산하기 위하여 스틱셀의 다양한 특징들로부터 유도된 상이도를 사용한다. 상이도의 단위가 픽셀, 미터, 누적값 등으로 매우 상이하기 때문에 식 (9)와 같이 가중합 방식으로 통합되며, 이를 위해 추가적으로 7개의 가중치가 파라미터로 필요하게 된다.

$$D(S_i, S_j) = w_1 \cdot LD(S_i, S_j) + w_2 \cdot AD(S_i, S_j) + w_3 \cdot HD(S_i, S_j) + w_4 \cdot ld(S_i, S_j) + w_5 \cdot ad(S_i, S_j) + w_6 \cdot hd(S_i, S_j) + w_7 \cdot MD(S_i, S_j) \quad (9)$$

위에서 제시한 내용을 종합하면, DBSCAN 기반 군집화를 사용하여 스틱셀 분할을 수행하기 위해서는 총 9개의 파라미터(r, m, w_1-w_7)를 설정하는 것이 필요하다. 이들 파라미터는 스틱셀 분할 성능과 밀접한 관계를 가지기 때문에 최적의 값을 설정해 주어야 한다. 이를 위하여 본 논문에서는 유전 알고리즘을 사용하여 최적의 파라미터 값을 설정해주는 방법을 제안하였다. 유전 알고리즘을 사용한 파라미터 최적화는 오프라인(offline)에서 미리 수행되며, 온라인(online)으로 스틱셀 분할을 수행하는 동안에는 최적화 수행을 다시 수행하지 않고 오프라인 파라미터 최적화 과정에서 얻어진 파라미터를 사용한다.

유전 알고리즘은 자연 세계의 진화 과정을 기반으로 하는 전역 최적화(global optimization) 방법으로, 유전자들의 재생산(reproduction), 교차(crossover) 및 변이(mutation) 과정을 통해 주어진 조건에 최적인 유전자를 찾아내는 방법이다. 본 연구에서 대상으로 하는 스틱셀 분할 방법에 유전 알고리즘을 적용하기 위해서는 유전자를 설정하는 것과 유전자의 적합도를 정의할 필요가 있다. 본 논문에서 최적화하기 원하는 값은 DBSCAN의 파라미터 r, m 과 거리 계산 시 사용되는 가중치 파라미터 w_1-w_7 이다. 따라서 이들 9개의 값을 일렬로 늘어놓은 벡터를 유전자로 설정하였다.

유전자의 적합도는 스틱셀 분할 결과가 얼마나

정확한가를 평가하였다. 스틱셀 분할 결과의 정확도는 수동으로 입력한 스틱셀 분할과 설정된 유전자를 파라미터로 사용해 얻은 스틱셀 분할 결과를 비교하여 계산하였다. 즉, 설정된 파라미터를 통해 얻어진 스틱셀 분할 결과가 수동으로 입력된 스틱셀 분할 정답(ground truth)과 유사할수록 더욱 적합한 파라미터라고 평가된다. 이를 위해 75개의 상황에서 실제계에서 동일한 물체를 구성하는 스틱셀들끼리 묶어주는 정답을 수동으로 입력하였다. 본 논문에서는 스틱셀 분할 결과의 정확도를 식 (10)과 같이 $F-measure^{17}$ 를 사용하여 평가하였다. $F-measure$ 는 알고리즘이 우수할수록 1에 근접한 값을 가지고 반대의 경우 0에 근접한 값을 가지므로, 유전 알고리즘은 $F-measure$ 를 최대화하는 유전자를 산출하도록 수행된다.

$$F-measure = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (10)$$

이 수식의 $Recall$ 과 $Precision$ 은 각각 식 (11)과 (12)와 같이 계산된다.

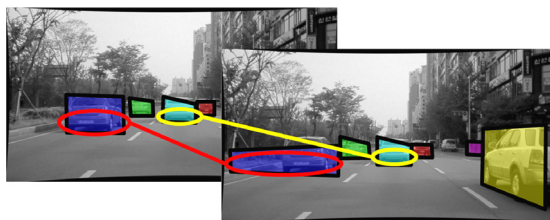
$$Recall = \frac{tp}{tp + fn} = \frac{tp}{GT(p)} \quad (11)$$

$$Precision = \frac{tp}{tp + fp} = \frac{tp}{Output(p)} \quad (12)$$

위 수식에서 tp, fn, fp 는 각각 true positive, false negative, false positive를 의미하며, 각각은 다음과 같이 정의된다.

- True positive(tp): 스틱셀 분할 결과에 존재하는 세그먼트(segment)가 정답에도 존재하는 경우
- False negative(fn): 스틱셀 분할 결과에 존재하지 않는 세그먼트가 정답에는 존재하는 경우
- False positive(fp): 스틱셀 분할 결과에 존재하는 세그먼트가 정답에는 존재하지 않는 경우

스티크셀 분할 결과에 존재하는 세그먼트와 정답에 포함된 세그먼트가 동일한가 여부는 자카드 계수(Jaccard coefficient)를 기반으로 판단하였다. 두 세그먼트 A 와 B 의 자카드 계수는 식 (13)과 같이 교집합 영역과 합집합 영역의 비율로 계산되며, 본 논문에서는 J 가 0.67 이상이면 두 세그먼트가 동일하다고 판단하였다.



(a) Ground truth (b) Segmentation result

Fig. 4 True positive is determined by calculating Jaccard coefficient between ground truth and segmentation result

$$J = \frac{(A \cap B)}{(A \cup B)} \quad (13)$$

이때, Recall의 분모($tp+fn$)는 정답에 존재하는 모든 세그먼트의 개수($GT(p)$)를 나타내며, Precision의 분모($tp+fp$)는 스틱셀 분할 결과에 존재하는 모든 세그먼트의 개수($Output(p)$)를 나타낸다. 그러므로 tp , $GT(p)$, 및 $Output(p)$ 를 통해 Recall 및 Precision, F -measure 값을 계산할 수 있다.

Fig. 4는 스틱셀 분할 결과와 정답에 포함된 세그먼트를 비교하여 tp 를 판단하는 과정을 보여준다. Fig. 4(a)는 정답에 포함된 세그먼트, Fig. 4(b)는 스틱셀 분할 결과에 포함된 세그먼트를 보인다. tp 인 지 판단하기 위해 스틱셀 분할 결과와 정답에 포함된 두 세그먼트의 이미지 상의 x 위치를 기반으로 자카드 계수(J)를 구하여, 그 값이 임계값(0.67)이상이면 동일 세그먼트(tp)로 간주한다. Fig. 4(a) 및 Fig. 4(b) 내 붉은색 원 안의 두 세그먼트는 해당 자카드 계수가 임계값을 넘지 못하여 다른 세그먼트로 판단되고, 노란색 원 안의 두 세그먼트의 자카드 계수는 임계값 이상이어서 동일 세그먼트(tp)로 판단된다.

Matlab의 Global Optimization Toolbox¹⁸⁾에 구현되어 있는 유전 알고리즘을 사용하였으며, 유전 알고리즘의 입력으로 사용되는 유전자들의 개체 수 (population)와 세대 수(generation)로는 각각 300과 70을 사용하였다.

4. 실험 결과

제안된 방법의 성능을 평가하기 위해서 75개의 전방 주행 상황에서 취득된 스테레오 영상을 사용하였으며, 조밀 디스패리티 맵은 ELAS 방법⁹⁾으로

Table 1 Quantitative performance evaluation result

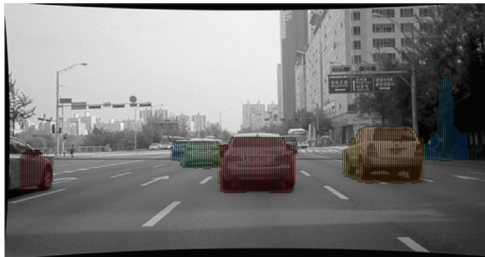
		Recall	Precision	F-measure
기존 방법 (특징 추가 전)	최적화 전	0.36	0.15	0.20
	최적화 후	0.51	0.68	0.56
제안 방법 (특징 추가 후)	최적화 전	0.29	0.18	0.21
	최적화 후	0.68	0.65	0.64

획득하였다. 사용된 영상의 크기는 1280×670 픽셀 이고, 실험은 Intel Core i7-2600 CPU 3.4GHz의 사양을 갖는 PC에서 진행되었다.

제안된 방법은 앞서 설명한 75개의 전방 주행 상황에서 F -measure를 사용하여 정량 평가되었으며, 기존에 제안된 군집화 기반 스틱셀 분할 방법⁶⁾과 비교 평가 되었다. Table 1은 정량적 평가 결과를 보여 준다. 기존 방법은 스틱셀 분할을 위해 실세계의 위치와 이동 방향 정보만을 사용한 방법을 의미하고, 제안 방법은 본 논문에 제시한 7가지의 정보를 모두 사용한 방법을 의미한다. 또한, 최적화 전은 DBSCAN 파라미터는 경험적으로 설정하고 가중치(w_1-w_7)들은 모두 동일한 값으로 설정한 상황을 의미하고, 최적화 후는 유전 알고리즘에 의해 최적화된 파라미터를 사용한 상황을 의미한다.

Table 1을 통해 두 가지 사실을 확인할 수 있다. 첫째, 파라미터 최적화 전에 비해 제안된 유전 알고리즘 기반 방법으로 파라미터 최적화를 수행한 후의 알고리즘 성능이 월등히 높은 것을 확인할 수 있다. 이는 제안된 최적화 방법이 스틱셀 분할을 위한 적절한 파라미터를 찾는 데 유용하다는 것을 보여준다. Fig. 5는 제안된 파라미터 최적화 방법 사용 전과 후의 스틱셀 분할 결과를 보여준다. 최적화 방법 사용 전의 스틱셀 분할 결과에서 하나의 세그먼트가 다수개의 세그먼트로 잘못 분할된 결과를 볼 수 있다. 최적화 방법 사용 후의 실험 결과는 최적화 전의 문제가 해결되었음을 보여준다.

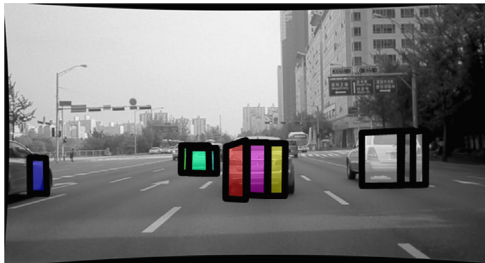
둘째, 최적화 후의 경우 기존 방법을 사용하는 것에 비해 제안된 방법을 사용하는 경우 F -measure 값이 0.56에서 0.64로 약 14% 향상된 것을 확인할 수 있다. 이는 제안된 방법이 추가적으로 사용한 특징들이 스틱셀 분할 성능 향상에 도움이 된다는 것을 보여준다. Fig. 6은 최적화 후의 기존 방법과 제안된 방법의 스틱셀 분할 결과의 예를 보여준다. Fig. 6(b)



(a) Stixel generation result



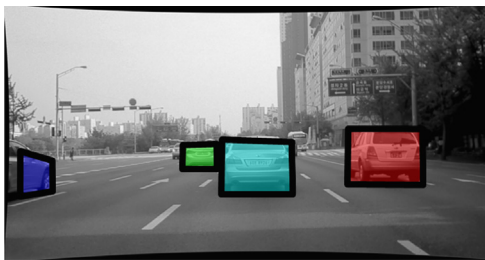
(a) Stixel generation result



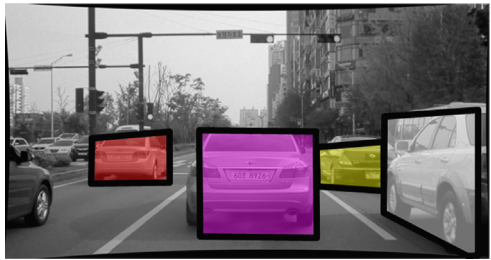
(b) Segmentation result of the conventional method



(b) Stixel segmentation result of the previous method



(c) Segmentation result of the proposed method



(c) Stixel segmentation result of the proposed method

Fig. 5 Segmentation result comparison between before and after parameter optimization

Fig. 6 Stixel segmentation result comparison

의 최적화 후의 기존 방법은 이미지 중앙의 자동차와 우측의 자동차를 하나의 세그먼트로 판단하고 있으나, Fig. 6(c)의 제안 방법은 두 자동차를 별개의 세그먼트로 분할하였음을 볼 수 있다.

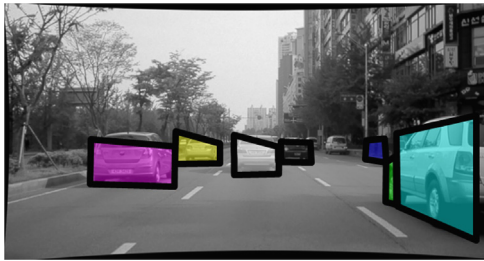
Table 2는 유전알고리즘에 의해 최적화된 가중치 값들을 보여준다. 각 특징들이 min-max 방법에 의해 정규화되었기 때문에 이 표에 제시된 가중치 값이 크다는 것이 해당 특징이 스틱셀 분할에 중요하게 사용된다는 것을 의미한다고 볼 수 있다.

Fig. 7은 제안된 스틱셀 분할 방법의 오류 상황을

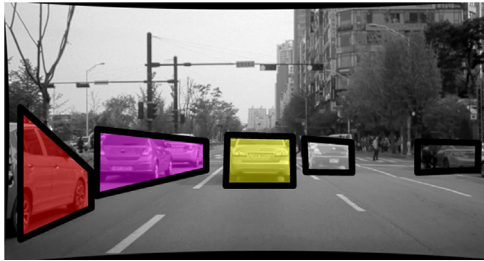
보여준다. Fig. 7(a)는 하나의 장애물이 다수의 세그먼트로 과분화된 경우를 보여준다. 이미지 내 우측의 과분화 된 자동차의 경우, 자동차의 높이가 급격하게 변함으로 인해 별개의 세그먼트로 오인하는 문제를 발생시켰다. Fig. 7(b)는 다수의 장애물이 하나의 세그먼트로 미분화 된 경우를 보여준다. 분홍색 세그먼트 내 두 대의 자동차가 실시계 및 영상 상에서의 스틱셀의 위치가 근접해 있고, 이동 방향이 유사하며, 두 장애물의 높이 또한 비슷하여 두 장애물을 하나의 세그먼트로 오인하는 문제를 발생시켰다. 추후에는 스틱셀 분할 결과를 연속 영상에서 지속적으로 추적하여 이와 같은 과분화 및 미분화 문제를 해결하는 접근법에 대한 연구를 수행할 예정이다.

Table 2 Weight values optimized by genetic algorithm

w_1	w_2	w_3	w_4	w_5	w_6	w_7
1.46	0.03	0.54	0.94	0.06	0.28	0.28



(a) Over-segmentation



(b) Under-segmentation

Fig. 7 Failure examples of the proposed method

5. 결론

본 논문은 스틱셀 기반 주행 환경 표현 단계 중 스틱셀 분할 과정의 성능을 향상시키는 방법을 제안하였다. 제안된 방법은 다임러의 DBSCAN 기반 스틱셀 분할의 성능을 개선하는데 그 목적을 두고 있다. 다임러의 기존 방법에 비해 본 논문이 갖는 차별점은 스틱셀의 다양한 특징(위치, 움직임, 높이 정보 및 외양 정보)를 사용하였고, 스틱셀 분할에 필요한 다수의 파라미터들을 유전 알고리즘을 사용하여 최적화 하였다는 점이다. 다양한 상황에서 획득된 데이터를 사용하여 정량적으로 비교 평가한 결과, 기존 방법에 비해 제안된 방법이 향상된 스틱셀 분할 성능을 보이는 것을 확인할 수 있었다.

후 기

본 연구는 현대모비스의 지원을 받아 수행되었습니다.

References

1) T.-N. Nguyen, B. Michaelis, A. Al-Hamadi, M. Tornow and M.-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, 2012.

2) 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.

3) D. Pfeiffer and U. Franke, "Modeling Dynamic 3D Environments by Means of the Stixel World," IEEE Intelligent Transportation Systems Magazine, Vol.3, No.3, pp.24-36, 2011.

4) 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. Haueis, M. Trepte, C. Brenk, A. Tamke, M. Ghanaat, M. Braun, Ar. 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.

5) B. Günzel, R. Benenson, R. Timofte and L. V. Gool, "Stixels Motion Estimation without Optical Flow Computation," European Conference on Computer Vision (ECCV), pp.528-539, 2012.

6) M. Muffert, T. Milbich, D. Pfeiffer and U. Franke, "May I Enter the Roundabout? A Time-to-contact Computation Based on Stereovision," IEEE Intelligent Vehicle Symposium (IV), pp.565-570, 2012.

7) F. Erbs, B. Schwarz and U. Franke, "Stixmentation - Probabilistic Stixel Based Traffic Scene Labeling," British Machine Vision Conference, pp.71.1-71.12, 2012.

8) M. Ester, H.-P. Kriegel, J. Sander and X. Xu, "A Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," International Conference on Knowledge Discovery and Data Mining (KDD), pp.226-231, 1996.

9) A. Geiger, M. Roser and R. Urtasun, "Efficient Large-scale Stereo Matching," ACCV2010,

- LNCS, Vol.6492, pp.25-38, 2011.
- 10) A. D. Sappa, F. Dornaika, D. Ponsa, D. Gerónimo and A. López, "An Efficient Approach to Onboard Stereo Vision System Pose Estimation," IEEE Transactions on Intelligent Transportation Systems, Vol.9, No.3, pp.476-490, 2008.
 - 11) R. Labayrade, D. Aubert and J.-P. Tarel, "Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through V-disparity Representation," IEEE Intelligent Vehicle Symposium (IV), pp.17-21, 2002.
 - 12) S. Nedeveschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, T. Graf and R. Schmidt, "High Accuracy Stereovision Approach for Obstacle Detection on Non-planar Roads," IEEE Intelligent Engineering Systems (INES), pp.211-216, 2004.
 - 13) 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.
 - 14) P. Cerri and P. Grisleri, "Free Space Detection on Highways Using Time Correlation between Stabilized Sub-pixel Precision IPM Images," International Conference on Robotics and Automation (ICRA), pp.2223-2228, 2005.
 - 15) D. Pfeiffer and U. Franke, "Efficient Representation of Traffic Scenes by Means of Dynamic Stixels," IEEE Intelligent Vehicle Symposium (IV), pp.217-224, 2010.
 - 16) T. Ahonen, A. Hadid and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.28, No.12, pp.2037-2041, 2006.
 - 17) Y. J. Huang, R. Powers and G. T. Montelione, "Protein NMR Recall, Precision, and F-measure Scores (RPF Scores): Structure Quality Assessment Measures Based on Information Retrieval Statistics," Journal of the American Chemical Society, Vol.127, No.6, pp.1665-1674, 2005.
 - 18) Global Optimization Toolbox, <http://www.mathworks.co.kr/products/global-optimization>, 2014.