

도시 구조물 분류를 위한 3차원 점 군의 구형 특징 표현과 심층 신뢰 신경망 기반의 환경 형상 학습

Spherical Signature Description of 3D Point Cloud and Environmental Feature Learning based on Deep Belief Nets for Urban Structure Classification

이 세 진⁺, 김 동 현¹

Sejin Lee⁺, Donghyun Kim¹

Abstract This paper suggests the method of the spherical signature description of 3D point clouds taken from the laser range scanner on the ground vehicle. Based on the spherical signature description of each point, the extractor of significant environmental features is learned by the Deep Belief Nets for the urban structure classification. Arbitrary point among the 3D point cloud can represent its signature in its sky surface by using several neighborhood points. The unit spherical surface centered on that point can be considered to accumulate the evidence of each angular tessellation. According to a kind of point area such as wall, ground, tree, car, and so on, the results of spherical signature description look so different each other. These data can be applied into the Deep Belief Nets, which is one of the Deep Neural Networks, for learning the environmental feature extractor. With this learned feature extractor, 3D points can be classified due to its urban structures well. Experimental results prove that the proposed method based on the spherical signature description and the Deep Belief Nets is suitable for the mobile robots in terms of the classification accuracy.

Keywords 3D Point Cloud, Spherical Signature Descriptor, Classification, Deep Belief Nets, Feature Extractor

1. 서 론

기계 학습(Machine learning)은 여러 분야에서, 특히 영상 처리 분야에서 비약적인 발전을 이루어 왔다. GPU와 같은 장비의 발달로 인해 컴퓨터의 계산 속도가 빨라짐에 따라 기계 학습 모델의 학습 속도 또한 더불어 급속

도로 향상되고 있다^[1]. 결국 구글 딥마인드 알파고(Google DeepMind AlphaGo)^[2]의 인공 지능이 이세돌 9단과의 바둑 대국에서 5전 4승을 거두는 역사적인 순간을 우리는 맞이하게 되었다. 이러한 인공 지능의 발전에 의해 사회 및 문화적 패러다임의 변화가 지금부터 시작될 것이라고 사료된다. 이러한 기류에 발맞추어 본 논문에서는 무인 이동 로봇의 자율 주행을 위한 기계 학습 기술을 기반으로 하는 새로운 환경 인식 및 분류 방법을 제안하고자 한다.

무인 이동 로봇이 자율 주행을 성공적으로 수행하기 위해서는 주변 환경을 강인하게 인지 및 분류할 수 있어야 하고 이를 위해서는 근본적으로 물체에 대한 형상학

Received : Apr. 14. 2016; Revised : Jun. 7. 2016; Accepted : Jul. 21. 2016

※This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning(NRF-2015R1C1A1A01054724)

⁺Corresponding author: Division of Mechanical & Automotive Engineering, Kongju National University, Cheonan-daero, Seobuk-gu, Cheonan-si, Chungcheong-nam-do, Korea (sejiny3@kongju.ac.kr)

¹Department of Computer Science, Kennesaw State University (donghyun.kim@kennesaw.edu)

적 정보를 잘 표현할 수 있어야 한다^[3]. 이러한 기능이 잘 뒷받침될 때, 위치 추정, 경로 계획, 장애물 회피, 그리고 의미 지도 작성 등의 무인 자율 주행 기술의 기본적인 기능들이 원만하게 수행될 수 있다.

무인 이동 로봇이 주행 환경을 강인하게 인지하기 위해 지금까지 다양한 종류의 센서들이 사용되었다. 그 중에서도 라이다(Lidar)나 레이저 스캐너(Laser scanner)와 같은 3차원 거리 측정 센서가 주로 사용되고 있다. 2005년 다르파 그랜드 챌린지(DARPA Grand Challenge)에서 우승한 스탠포드(Stanford) 대학의 스텐리(Stanley)^[4]나 최근의 구글 자동차(Google Car)가 그 대표적인 예가 될 것이다. 이러한 센서에 의해 제공되는 데이터는 단순히 물체의 존재 여부, 즉 공간적인 점유에 대한 정보만을 제공하기 때문에 무인 이동 로봇은 거리 측정 데이터 자체만으로 어떤 종류의 물체에서 측정되었는지를 직접적으로 판단하기 매우 어렵다. 따라서 환경 지도 생성에서 단순히 주변 환경의 지형/지물을 표현하는 것뿐만 아니라 지형/지물의 종류를 인식하는 것이 무엇보다도 중요하다^[5]. 이러한 지도의 생성을 의미 지도 작성(Semantic mapping)이라고 하는데, 이를 통해 무인 이동 로봇의 주변 환경에 존재하는 물체의 종류를 확인할 수 있고 물체 사이의 공간적인 관계를 파악할 수 있을 뿐만 아니라 앞에서 언급한 위치 추정, 경로 계획 등의 응용에 더욱 풍부한 정보를 제공하게 된다.

본 논문에서는 센서 데이터의 형상 분류를 위해 크게 구형 특징 표현과 환경 형상 학습, 이렇게 두 부분으로 나누어 설명하고자 한다. 첫째, 구형 특징 표현에서는 레이저 스캐너로 측정된 3차원 점 군의 단일 점에 대한 형상학적 정보를 직관적으로 서술할 수 있는 방법을 제안한다. 3차원 점 군에 속해 있는 임의의 단일 점에 대한 형상학적 정보는 기본적으로 그 점이 중심점이 되는 단위 구(하늘)의 표면에서 형상화될 수 있다. 단위 구의 표면을 격자화 시킨 후 주변의 다른 점 들을 이용하여 베이저안 업데이트(Bayesian updating) 모델^[6] 기반의 점유 증거를 누적시키는 방식을 기반으로 한다. 둘째, 환경 형상 학습에서는 이렇게 형상화된 구형 특징 표현 결과 들을 입력 값으로 하는 기계 학습의 한 종류인 심층 신뢰 네트워크(Deep Belief Networks)^[7]를 사용하여 구형 특

징 표현에 대한 환경 형상을 학습시킨다. 본 논문에서는 실험 데이터 기반의 시뮬레이션 결과를 통해 제안하고자 하는 센서 데이터의 형상 분류법에 대한 성능 검증 및 타당성을 증명하고자 한다.

2장에서는 관련 연구들에 대해 소개하며, 3장에서는 구형 특징 표현 방법에 대해 설명한다. 4장에서는 환경 형상 학습 방법에 대해 설명하며 5장에서는 실험 데이터 기반의 시뮬레이션 결과를 분석해본다. 마지막으로, 결론에서는 본 연구의 결과가 갖는 의미와 추후 연구 주제에 대해 고찰하고자 한다.

2. 관련 연구

의미 지도 작성은 현재 로봇공학 분야에서 활발히 진행 중인 연구 분야이다. Jean-Francois Lalonde 연구팀은 레이저 거리 센서를 이용하여 주변 환경을 측정하였고, 3차원 점 군 데이터의 통계적 특성을 통해 점 군 데이터를 ‘Scatter’, ‘Linear’, ‘Surface’의 3가지로 분류하는 방법을 제안하였다^[8]. [8]에서는 레이저 거리 센서를 이용하여 획득한 3차원 점 군의 공간상의 분포 특성을 계산하여 특징 형상(Saliency feature)을 추출한 후, 획득한 특징 형상으로부터 GMM (Gaussian Mixture Model)을 이용하여 지형에 따라 점 군을 분류하였다. Aleksey Golovinskiy 연구팀은 물체의 검출과 군집화, 그리고 인식까지 전 영역에 걸쳐 3차원 점 군을 인식하는 방법을 소개하였다^[9]. 레이저 거리 센서를 이용하여 점 군을 획득하였고, 점 군의 밀도가 높은 곳을 물체의 위치로 추정하여 전경과 배경을 그래프 컷(Graph-cut)의 방법으로 구분한다. 군집화된 점들에 대해 점의 수, 영역의 부피, 평균 높이, 높이의 표준 편차 등을 특징치로 추출하여 물체를 분류하는 방법을 제안하였다. Xialong Zhu 연구팀은 2차원 레이저 거리 센서를 차량의 양쪽 측면에 설치하여 수직으로 스캔하도록 시스템을 구성하고, 차량의 이동에 따라 연속적으로 2차원 데이터를 수집하도록 하였다^[9,10]. 이러한 방식으로 점 군을 수집하면 3차원 점 군 데이터의 각 점을 픽셀로 하는 3차원 거리 이미지(Range image)로 표현할 수 있다. 거리 이미지에 이미지 기반 객체 분할(Segmentation) 알고리즘을 적용하여 물

체를 효과적으로 분할하고 분류하는 방법을 제안하였다. 특히 [10]에서는 법선 벡터 추정(Normal vector estimation) 기법을 물체의 분류에 이용하고, [11]에서는 각각의 물체마다 물체의 기하학적인 특징을 정의하여 물체를 분류한다. 그리고 [12]에서는 획득된 3D 점 군 데이터를 4가지의 특징을 표현하는 복셀(Voxel) 격자로 나누어 분석 후 구조물마다 갖게 되는 경향성을 찾아 하나의 구조물로 인식하는 의미지도 작성법을 개발하였다. [13]에서는 점 군 획득을 하며 생긴 오차들을 논문에서 정한 5가지의 특징을 이용하여 가까운 점 군끼리 교합시켜 실제 데이터(True)와 최대한 일치하도록 만든 3D ICP outdoor SLAM을 개발하였다. [14]에서는 획득한 점 군에 이웃되는 점 군 그룹을 만들어 그 그룹의 법선 벡터(Normal vector)를 생성하고 이 벡터의 방향을 원하는 방향으로 설정한 후 이 벡터를 이용하여 점 군의 형상학적 특성을 파악하는 연구를 진행하였다. [15]에서는 컨볼루션 신경 회로망(Convolutional neural networks)과 순환 신경 회로망(Recursive neural networks) 기법을 이용하여 RGB-D 센서로 얻은 데이터로부터 다양한 물체를 분류하는 연구를 수행하였다.

본 연구에서는 기존의 연구자가 직접 모델링하는 방식의 수작업 형상 추출법(Hand crafted feature extractor) 대신에 기계 학습의 한 종류인 심층 신뢰 네트워크(Deep Belief Nets)를 사용함으로써 데이터의 물체 분류 성공률을 최대화 시키고자 함에 그 의미를 두고 있다.

3. 구형 특징 표현(Spherical Signature Description)

3차원 점 군 데이터 내에서 임의의 단일 점에 대한 특징을 분석하고자 할 때 일반적으로는 주변 점 군에 따른 주요 특징(Principal component analysis)을 추출해냄으로써 단일 점에 대한 대표적인 형상학적 특징을 표현한다. 이러한 방법은 국부적인 점 군의 주된 특징은 잘 표현할 수 있는 반면 상대적으로 약한 정보들은 손실될 수밖에 없다. 따라서 본 연구에서는 기본적으로 단일 점이 중심점이 되는 단위 구의 표면을 격자화 시킨 후 주변의 다른 점들을 이용하여 베이저안 업데이트(Bayesian

updating) 모델 기반의 점유 증거를 누적시킴으로써 그 특징을 모두 표현하고자 한다.

3.1 점유 증거 갱신을 위한 준비

전체 점 군 데이터를 T , 그리고 구형 특징을 표현하고자 하는 단일 점을 τ 라고 할 때, 우선 전체 점 군은 구형 특징 표현을 하고자 하는 점을 기준으로 다음의 식과 같이 이동된다.

$$X(\tilde{T}) = X(T) - X(\tau), \tau \in T \quad (1)$$

여기서 \tilde{T} 는 단일 점 τ 를 상대적인 원점으로 고려하였을 때 이동되는 점 군을 나타내며 $X(\cdot)$ 표시는 입력 점에 대한 직교 좌표 위치 (x, y, z) 를 나타낸다. 또한 단일 점에 대한 자신의 특징을 표현할 수 있는 구형 표면은 구형 좌표계, (r, θ, φ) 상의 각도 간격 $(\delta\theta, \delta\varphi)$ 으로 세분화될 수 있다. 여기서 r 은 단일 점, 즉 원점에서 세분화된 격자 패치(Patch) 중심까지의 거리가 되고, Fig. 1에서 보는 바와 같이 θ 는 xy 천정 방향에서 측정된 위도, 그리고 φ 는 평면 상에 투영된 패치의 중심점에 대한 경도가 된다. 여기서 구의 표면은 단일 구의 형태로 고려되기 때문에 세분화된 격자 패치의 반경 r 은 무시되고 단지 위도와 경도 관점에서만 주변 점 군에 의해 특징이 표현된다.

3.2 점유 증거 갱신

임의의 단일 점 τ 에 의해 이동된 모든 점들 \tilde{T} 는 Fig. 1에

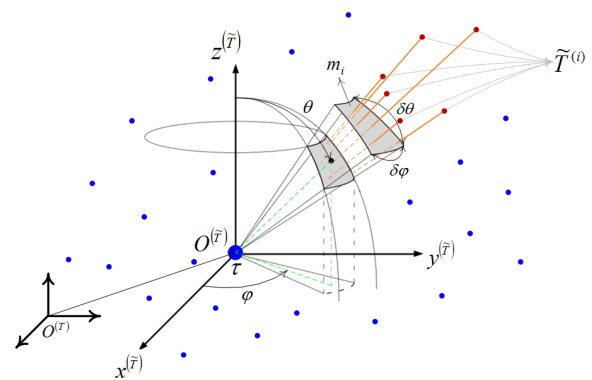


Fig. 1. Concept of the angular tessellation of the unit spherical surface with its grouped points

서와 같이 각도 간격 ($\delta\theta, \delta\varphi$)에 의해 세분화된 패치를 포함하는 각각의 사각뿔 영역에 의해 그룹 지어질 수 있다. 이 사각뿔 끝 꼭지점은 단일 점 τ 에 위치해 있으며 해당 패치를 포함하고 있는 형태로 밑면은 무한대의 거리에 의한 것이라고 가정할 수 있다. 기본적으로 각각의 패치에 대한 점유 증거는 베이시안 업데이트(Bayesian updating) 모델에 의해 갱신되고 누적된다. 세분화된 패치 세트의 점유 증거 확률은 아래와 같이 정의될 수 있다.

$$p(M|\tilde{T}) = \prod_{i=1}^N p(m_i|\tilde{T}^{(i)}) \quad (2)$$

여기서 M 은 각도 간격 ($\delta\theta, \delta\varphi$)에 의해 단위 구면을 구성하는 총 N 개로 세분화된 패치 세트를 나타내며, $\tilde{T}^{(i)}$ 는 Fig. 1에서 보는 바와 같이 i 번째 패치 M_i 에 해당되는 사각뿔 공간에 포함된 점 군을 나타낸다. m_i 의 점유 증거 확률은 베이즈 규칙(Bayes rule)에 의해 다음과 같이 정의된다.

$$p(m_i|\tilde{T}^{(i)}) = p(m_i|\tilde{T}_{1:k_1-D}^{(i)}) \quad (3)$$

$$\begin{aligned} p(m_i|\tilde{T}_{1:k}^{(i)}) &= p(m_i|\tilde{T}_{1:k-1}^{(i)}, \tilde{\tau}_k^{(i)}) \\ &= \frac{p(\tilde{\tau}_k^{(i)}|m_i)p(m_i|\tilde{T}_{1:k-1}^{(i)})}{p(\tilde{\tau}_k^{(i)}|\tilde{T}_{1:k-1}^{(i)})} \end{aligned} \quad (4)$$

여기서 k 는 해당 그룹에 속한 점들의 총 개수 D 까지 점진적으로 증가한다. 분자의 첫 번째 항은 역시나 베이즈 규칙(Bayes rule)을 통해 다음과 같은 갱신(Updating) 모델로 표현할 수 있다.

$$p(\tilde{\tau}_k^{(i)}|m_i) = \frac{p(m_i|\tilde{\tau}_k^{(i)})p(\tilde{\tau}_k^{(i)})}{p(m_i)} \quad (5)$$

따라서 m_i 의 점유 증거 확률은 다음과 같이 다시 쓰여질 수 있다.

$$p(m_i|\tilde{T}_{1:k}^{(i)}) = \frac{p(m_i|\tilde{\tau}_k^{(i)})p(\tilde{\tau}_k^{(i)})p(m_i|\tilde{T}_{1:k-1}^{(i)})}{p(m_i)p(\tilde{\tau}_k^{(i)}|\tilde{T}_{1:k-1}^{(i)})} \quad (6)$$

그리고 식 (6)의 반대 사건 \bar{m}_i 이 일어날 확률은 다음과 같다.

$$p(\bar{m}_i|\tilde{T}_{1:k}^{(i)}) = \frac{p(\bar{m}_i|\tilde{\tau}_k^{(i)})p(\tilde{\tau}_k^{(i)})p(\bar{m}_i|\tilde{T}_{1:k-1}^{(i)})}{p(\bar{m}_i)p(\tilde{\tau}_k^{(i)}|\tilde{T}_{1:k-1}^{(i)})} \quad (7)$$

식 (6)을 식 (7)로 나누면 여러 계산하기 어려운 확률 항들이 제거되며 결국 다음과 같이 표현된다.

$$\frac{p(m_i|\tilde{T}_{1:k}^{(i)})}{p(\bar{m}_i|\tilde{T}_{1:k}^{(i)})} = \frac{p(m_i|\tilde{\tau}_{1:k}^{(i)})p(m_i|\tilde{T}_{1:k-1}^{(i)})p(\bar{m}_i)}{p(\bar{m}_i|\tilde{\tau}_{1:k}^{(i)})p(\bar{m}_i|\tilde{T}_{1:k-1}^{(i)})p(m_i)} \quad (8)$$

식 (8)에 대한 Log odds ratio는 $l_k(m_i)$ 에 의해 다음과 같이 표현된다.

$$\begin{aligned} l_k(m_i) &= \text{observation model}(m_i, \tilde{\tau}_k^{(i)}) \\ &\quad + l_{k-1}(m_i) - l_0(m_1) \end{aligned} \quad (9)$$

여기서 $\text{observation model}(m_i, \tilde{\tau}_k^{(i)})$ 은 기준점 주변에 존재하는 임의의 이웃 점, $\tilde{\tau}_k^{(i)}$ 에 의해 정의되는 m_i 의 점유 증거 확률인 $p(m_i|\tilde{\tau}_k^{(i)})$ 에 대한 Log odds ratio값이고, sigmoid 함수를 변형하여 모델링 하였다. $p(m_i|\tilde{\tau}_k^{(i)})$ 은 다음과 같이 정의하였다.

$$p(m_i|\tilde{\tau}_k^{(i)}) = 1.0 - \frac{0.5}{1 + e^{-4(r_p - \mu)}} \quad (10)$$

여기서 r_p 는 기준점으로부터 이웃 점, $\tilde{\tau}_k^{(i)}$ 까지의 거리값이며, μ 는 m_i 의 점유 증거 갱신에 참여하는 이웃 점, $\tilde{\tau}_k^{(i)}$ 의 허용 구간을 조절할 수 있는 디자인 파라미터이다. 본 연구에서는 1.5 m로 설정하였으며 이에 대한 그래프 모양은 다음 Fig. 2와 같다. 이 그림에서 보듯이 대략

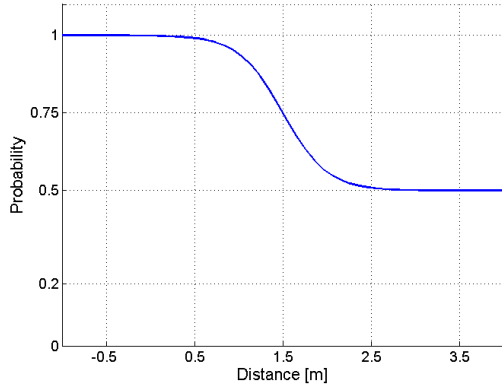


Fig. 2. Plot of the modified sigmoid function for the observation model in the Bayes updating model

1.0 m 거리에까지 존재하는 이웃 점들은 해당 격자의 점을 확률 갱신에 큰 영향을 미칠 수 있으며, 1.0 m부터 대략 2.0 m 거리에 존재하는 이웃 점들은 거리에 따른 영향이 줄어든다. 그리고 2.5 m 이상의 이웃 점들은 점유 확률 갱신에 영향을 거의 미치지 못하게 된다.

4. 환경 형상 학습(Environmental Feature Learning)

심층 학습(Deep learning)은 여러 비선형 변환기법의 조합을 통해 다량의 데이터나 복잡한 자료들 속에서 핵심적인 내용 또는 기능을 요약하는 작업을 높은 수준으로 시도하는 기계 학습 알고리즘의 집합으로 정의되며, 큰 틀에서 사람의 사고방식을 컴퓨터에게 가르치는 기계학습의 한 분야라고 이야기할 수 있다.

어떠한 데이터가 있을 때 이를 컴퓨터가 인지할 수 있는 형태(예를 들어 이미지의 경우는 픽셀정보를 열벡터로 변환)로 표현하고 이를 학습에 적용하기 위해 적합한 표현기법과 학습 모델을 만들지에 대한 많은 연구가 진행되고 있으며, 이러한 노력의 결과로 심층 신경 회로망(Deep neural networks), 컨볼루션형 심층 신경 회로망(Convolutional deep neural networks)^[16], 심층 신뢰 회로망(Deep belief networks)과 같은 다양한 심층 학습(Deep learning) 기법들이 컴퓨터 비전(Vision), 음성인식, 자연어 처리, 음성/신호처리 등의 분야에 적용되어 괄목할만한 결과들을 보여주고 있다.

본 장에서는 3장에서 설명한 구형 특징 표현 데이터를 입력 값으로 하여 심층 신뢰 회로망(Deep belief networks)을 통해 환경 형상 정보를 학습시키고자 한다. 특히 도시 환경에서 측정된 거리 데이터를 사용함으로써 야외 환경에서 주된 형상이 될 수 있는 지면, 벽면, 차량, 나무 등에 대한 강인한 형상 추출자를 학습시키고자 한다.

4.1 제한된 볼츠만 머신(Restricted Boltzmann Machine)

볼츠만 머신은 1985년 Geoffrey Hinton과 Terry Sejnowski가 발명한 방법으로 확률적으로 순환하는 신경망 네트워크^[17]이며 볼츠만 머신에서, 층간 연결을 없앤 형태의 모델이 제한된 볼츠만 머신(RBM)이다. 층간 연결을 없애면, 머신은 가시 유닛(Unit) x 와 은닉 유닛 h 로 이루어진 무방향 이분 그래프 형태의 모양이 되며, 이러한 이점으로 Neural network는 깊어질 수 있다. 본 연구에서는 제한된 볼츠만 머신의 입력 값, v 는 기본적으로 제 3장에서 설명한 임의의 단일 점 관점에서 작성된 구형 특징 표현, $p(M|\tilde{T})$ 이 된다.

Fig. 3에서는 제한된 볼츠만 머신(RBM)의 개념도를 나타내며 [16, 17]에 자세한 내용이 기술되어 있으므로 본 논문에서는 중요한 사항만 간단히 아래와 같이 요약하고자 한다. 가시 유닛 x 와 은닉 유닛 h 에 대한 제한된 볼츠만 머신(RBM)의 에너지 함수 $E(x, h)$ 는 다음과 같이 정의된다.

$$E(v, h) = -b^T h - c^T x - h^T W x \quad (11)$$

여기서 W 은 은닉층과 가시층 사이를 연결하는 가중치를 나타내고, b 와 c 는 각각 은닉층과 가시층의 바이어스(bias)가 된다. 이 에너지 함수 기반의 확률 분포 함수 $p(x, h)$ 는 다음과 같이 정의할 수 있다.

$$p(x, h) = \frac{\exp(-E(x, h))}{Z} \quad (12)$$

여기서 분모 Z 는 노말라이저(Normalizer)가 된다. 또한 모든 은닉 유닛 h 에 대한 가시 유닛 x 의 확률 합은 다음

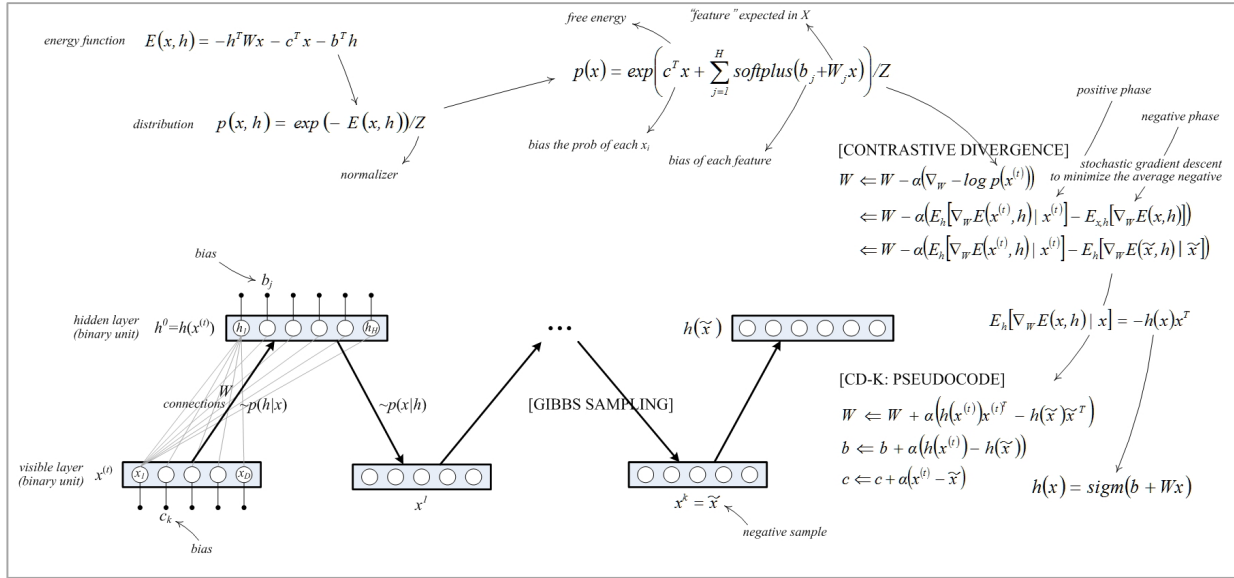


Fig. 3. Conceptual chart of the restricted Boltzmann machine^[16]

과 같다.

$$\begin{aligned}
 p(x) &= \sum_h p(x, y) \\
 &= \frac{1}{Z} \exp\left(c^T x + \sum_{j=1}^H \text{softplus}(b_j + W_j x)\right) \\
 &= \frac{\exp(-F(x))}{Z}
 \end{aligned} \tag{13}$$

여기서 $F(x)$ 는 식(13)에 있는 지수(exponential) 함수의 지수 부분이며, 자유 에너지 함수로 정의되고 *softplus* 함수는 다음과 같이 정의된다.

$$\text{softplus}(x) = \log(1 + \exp(x)) \tag{14}$$

결론적으로 확률 경사 하강법(Stochastic gradient descent) 기반으로 각 파라미터를 갱신하기 위한 규칙은 다음과 같이 나타낼 수 있다. 여기서 α 는 학습율로 디자인 파라미터가 된다. 본 연구에서는 0.1로 설정하였다.

$$\begin{aligned}
 W &\leftarrow W - \alpha (\nabla_w - \log p(x^t)) \\
 &\leftarrow W + \alpha (h(x^{(t)})(x^{(t)})^T - h(\tilde{x})(\tilde{x})^T)
 \end{aligned} \tag{15}$$

$$b \leftarrow b + \alpha (h(x^{(t)}) - h(\tilde{x})) \tag{16}$$

$$c \leftarrow c + \alpha (x^{(t)} - \tilde{x}) \tag{17}$$

대조 발산(Contrastive Divergence)을 위해 $x^{(t)}$ 로부터 시작되는 깁스 샘플링(Gibbs sampling)이 이루어지며 이를 통해 네가티브 샘플(negative sample) \tilde{x} 을 구할 수 있고 각각의 파라미터를 갱신할 수 있다.

4.2 심층 신뢰 신경망(Deep Belief Nets)

심층 신뢰 신경망(DBN)이란 기계학습에서 사용되는 그래프 생성 모형으로, 심층 학습(Deep learning)에서는 잠재변수의 다중계층으로 이루어진 심층 신경망을 의미한다. 계층 간에는 연결이 있지만 계층 내의 유닛(Unit) 간에는 연결이 없다는 특징이 있다. 심층 신뢰 신경망(DBN)에 관한 내용은 [16, 17]에 자세히 기술되어 있으므로 본 논문에서는 중요한 사항만 다음과 같이 요약하고자 한다.

심층 신뢰 신경망(DBN)은 생성 모형이라는 특성상 선행학습에 사용될 수 있고, 선행학습을 통해 초기 가중치를 학습한 후 역전파 알고리즘을 통해 가중치의 미조정을 할 수 있다. 이러한 특성은 훈련용 데이터가 적을

때 굉장히 유용한데, 이는 훈련용 데이터가 적을수록 가중치의 초기값이 결과적인 모델에 끼치는 영향이 커지기 때문이다. 선행 학습된 가중치 초기값은 임의로 설정된 가중치 초기값에 비해 최적의 가중치에 가깝게 되고 이는 미조정 단계의 성능과 속도향상을 가능케 한다.

심층 신뢰 신경망(DBN)은 비지도 방식으로 계층마다 학습을 진행하는데 이때 각각의 계층은 보통 제한된 볼츠만 머신(RBM)의 형태를 띠고 있다. 제한된 볼츠만 머신(RBM)은 에너지 기반의 생성 모형으로 가시 유닛과 은닉 유닛으로 이루어진 무방향 이분 그래프 형태이다. 가시 유닛들과 은닉 유닛들 사이에만 연결이 존재한다. 제한된 볼츠만 머신(RBM)이 훈련되고 나면 다른 제한된 볼츠만 머신(RBM)이 그 위에 쌓아 올려짐으로써 다중 계층 모형을 형성한다. 제한된 볼츠만 머신(RBM)이 쌓아 올려질 때마다, 이미 훈련된 제한된 볼츠만 머신(RBM)의 최상위 계층이 새로 쌓이는 제한된 볼츠만 머신(RBM)의 입력으로 쓰인다. 이 입력을 이용하여 새 제한된 볼츠만 머신(RBM)이 훈련되고 원하는 만큼의 계층이 쌓일 때까지 해당 과정이 반복된다.

5. 실험 결과

5.1 실험 시스템 및 데이터

제안된 방법론을 평가하기 위해 본 연구에서는 카이스트 캠퍼스를 측정한 데이터 세트를 사용하였다. 약 690 m와 680 m 영역에서의 카이스트 캠퍼스 내에서 측정된 위치와 거리 센서 데이터가 수집되었다. 거리 데이터는 전기 자동차에 장착된 3개의 SICK LMS 291 센서에 의해 수집되었다^[5].

Fig. 4(a)는 카이스트 내의 장소에서 측정된 점 군들을 보여주고 있다. 이 카이스트 데이터 세트는 수작업을 통해 건물, 도로, 차, 나무, 그리고 다른 것들에 대한 야외 구조물로 분류하였다. 그림에서 보이는 점 군들은 분류에 따라 색깔 별로 표시하였다. 도로는 파란색, 건물은 분홍색, 차는 하늘색, 나무는 연두색으로 표시하였다. 이 데이터 세트들은 추후 학습을 통한 분류 결과를 평가하기 위한 기준(Ground true) 값으로 사용된다. 수기 분류된(Hand labeled) 카이스트 데이터 세트의 통계들은 아

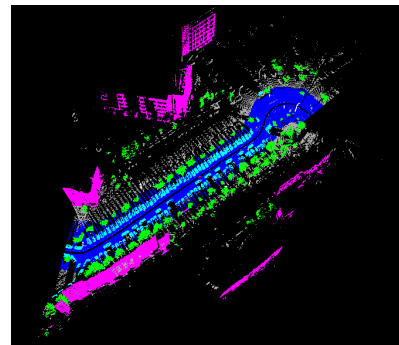
Table 1. Statistics of hand-labeled KAIST dataset^[12]

Num. of Pts.	Building (%)	Tree (%)	Road (%)	Car (%)	Others (%)
822,909	9.90	7.87	64.74	12.39	5.10

래 표에 정리되었다. 점들이 가장 크게 차지하는 분류 항목은 도로이다. 데이터 세트들은 대부분 야외 환경에서 쉽게 찾아지는 나무, 건물, 길, 그리고 차들로 구성된다. 그 외의 것들에 대한 점들은 단지 5% 정도로 차지된다.

5.2 구형 특징 표현 결과

먼저 효과적인 형상 학습을 위한 구형 특징 표현에서의 각도에 대한 해상도를 설정해야 한다. 이를 위해 본 연구에서는 θ 와 φ 의 각도 구간을 2° 에서 30° 의 간격 내에서 구분하여 계산하였다. 여기서 θ 와 φ 는 항상 같은 각도 인터벌(interval)로 설정하였다. 그리고 허용 가시 영역은 2 m, 이동 간격 μ 는 1.5 m로 설정하여 식 (10)의 Sigmoid 함수를 설정하였다. 이를 위해 Fig. 4에서와 같



(a)



(b)

Fig. 4. Hand-labeled KAIST dataset (a) point clouds and (b) satellite picture

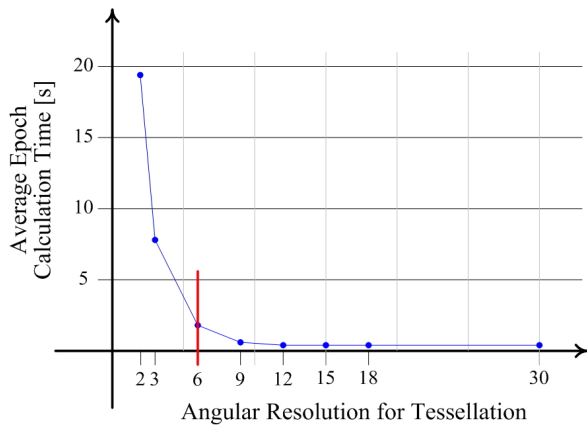


Fig. 5. Average epoch calculation time according to the angular resolution for tessellation

이 카이스트 데이터 세트의 일부를 사용하였으며 구형 특징 표현 데이터 샘플링(Sampling)을 위한 점의 샘플링 간격은 10으로 설정하였다. 적절한 각도 해상도를 결정하기 위해 구형 특징 표현 샘플링 데이터를 입력값으로 하는 제한된 볼츠만 머신(RBM)의 평균 에픽(epoch) 시간을 이용하였다. 이때 제한된 볼츠만 머신(RBM)을 수행하기 위해 에픽(epoch)을 15회, 미니 배치(mini batch) 개수를 10개, 은닉층의 노드 개수는 100개로 임의 설정하였다. 매틀랩(Matlab)으로 구현된 코드를 사용하였으며, 컴퓨터의 CPU 사양은 Intel(R) Core(TM) i7-6700K @ 4.00 GHz이고 RAM은 32.0 GB이다. 운영체제는 Windows 10 pro이다. Fig. 5에서 보듯이 각도 해상도는 6°에서

9° 정도 사이에서부터 계산량이 일정 수준 이하로 수렴되고 있음을 확인할 수 있다. 본 연구에서는 계산량과 각도 사이의 상관 관계에 따라 각도 해상도를 6°로 설정하였다. 단일 구형 특징 표현 데이터는 60×30 개($360^\circ/6^\circ \times 180^\circ/6^\circ$)의 픽셀로 구성되며 학습을 위한 1×1800 크기의 벡터로 변환되어 사용된다.

Fig. 6은 앞서 설명한 조건 하에서 수행한 지면, 자동차, 벽면, 나무 등에 속한 임의의 점들의 구형 특징 표현 결과를 각각 3개씩 나타내고 있다.

5.3 환경 형상 학습 결과

5.3.1 학습 파라미터 설정

심층 신뢰 회로망(DBN)을 실행하기 위해서는 몇 가지 파라미터를 설정해야 할 필요가 있다. 그 중에서 대표적인 것이 에픽(epoch)의 횟수, 배치(batch) 개수, 은닉층의 개수, 은닉층의 노드(Node) 개수 등이 있다. 이 중에서 우선 은닉층 노드의 개수에 따른 분류 성공율의 변화를 살펴본다. 이를 위해 은닉층은 2개의 층으로 하였으며 첫 번째 층과 두 번째 층의 노드 개수는 동일하게 설정하였다. 그리고 에픽(epoch) 횟수는 20, 배치(batch) 개수는 100으로 설정하였다. 카이스트 데이터 세트를 사용하였으며, 구형 형상 표현 데이터 세트를 위해 전체 822,909 개의 포인트 중 샘플링하여 10%인 82,292개의 포인트에 대한 구형 형상 표현 데이터 세트를 구성하였다. 이 중

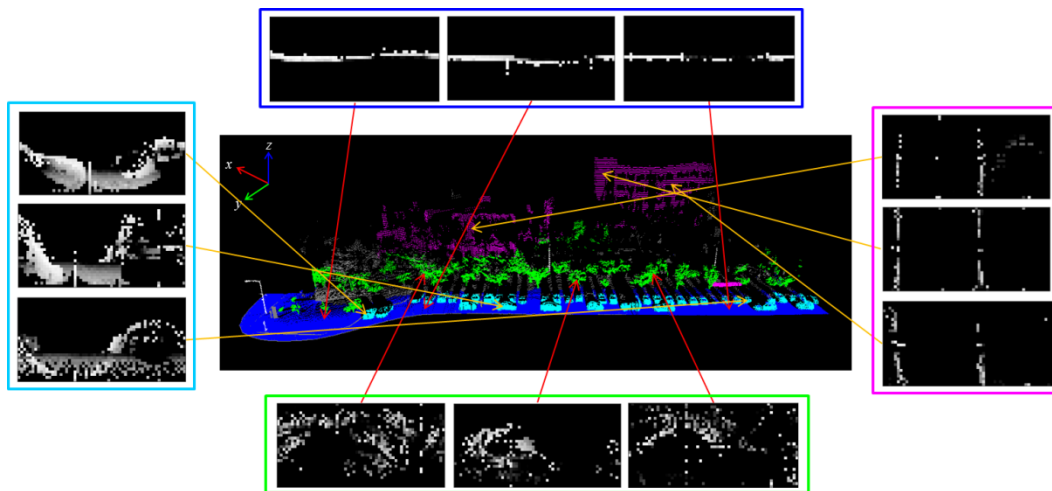


Fig. 6. Representative results of the spherical signature description

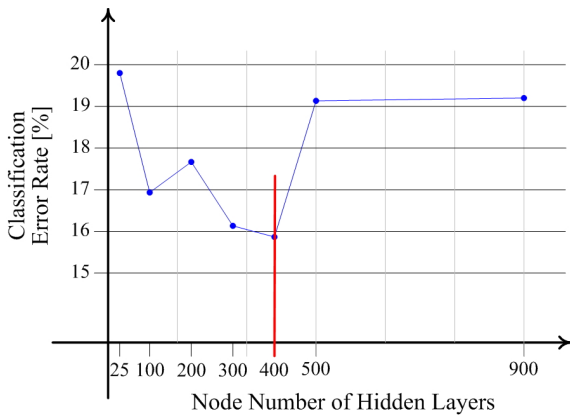


Fig. 7. Error rate of the classification according to the node number of hidden layers

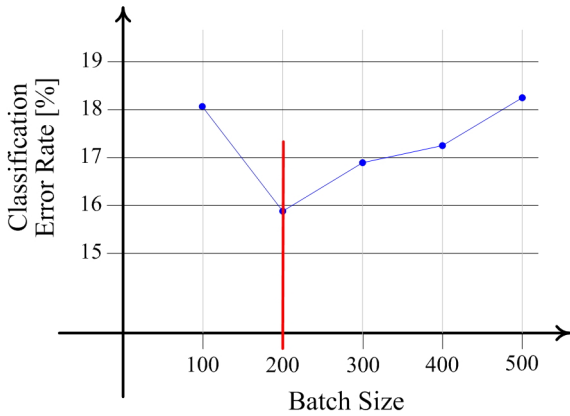


Fig. 8. Error rate of the classification according to the number of batch size

60,000개는 비지도 학습(Unsupervised learning)을 위한 훈련(Training) 데이터로 사용하였고, 10,000개는 미세 조정(fine-tuning)을 위한 훈련 데이터로 사용하였다. 나머지 12,292개의 데이터는 테스트(test) 데이터로 사용하였다. 은닉층의 노드 개수는 25, 100, 200, 300, 400, 500, 900개로 각각 나누어 분별 성공률을 측정하였다. 레이블(Label)은 도로, 건물, 차, 나무, 기타로 총 5개로 나누어 진행하였다. Fig. 7에서 보듯이 노드가 400개일 때 오차율이 가장 낮았으므로 본 연구에서는 400으로 설정하였다. 또한 배치(batch) 개수를 정하기 위해 배치(batch) 개수를 100부터 500개까지 100개씩 증가하여 분별 성공률을 측정하였다. Fig. 8에서 보듯이 배치(batch) 개수가 200일 때 오차율이 가장 낮았으므로 본 연구에서는 배치(batch) 개수를 200으로 설정하였다.

5.3.2 형상 학습 결과

본 연구에서 사용된 카이스트 데이터 세트는 총 822,909개의 점으로 이루어져 있다. 이 중 구형 특징 표현 데이터를 구성하기 위해 중심이 된 점은 82,290개이다. 이는 전체 데이터의 약 10%이며 점 군이 표현할 수 있는 전체 공간 영역에 대해 균일하게 수집되었다. 이렇게 수집된 점에 대한 구형 특징 표현 데이터 세트를 구성하였으며 (training_x1, training_y1), (training_x2, training_y2), (test_x, test_y)와 같이 세 종류의 세트로 다시 구분하였다. 여기서 (training_x1, training_x2, test_x)는 구형 특징 표현 데이터 세트이며, (training_y1, training_y2, test_y)는 구형 특징 표현에 사용된 기준 점에 대한 레이블 세트가 된다. 이 데이터 세트 또한 전체 공간 영역에 대해 균일하도록 구분되었다. (training_x1, training_x2, test_x)는 각각 27,430개로 동일한 개수의 기준 점으로부터 생성되었다.

(training_x1, training_y1)는 심층 신뢰 신경망에 비지도 학습(Unsupervised learning)을 위해 사용되었으며, 레이블 데이터인 training_y1은 실제로 사용되지 않았다. (training_x2, training_y2)는 비지도 학습 후 미세 조정 학습을 위해 사용되었고, (test_x, test_y)는 미세 조정 학습 후, 실제로 학습된 신경망의 분류 성능을 검증하기 위해 사용되었다.

심층 신뢰 신경망을 통해 학습된 첫 번째 가중치 행렬(weight matrix)(입력층과 첫 번째 은닉층 사이)을 가시화해보면 Fig. 9와 같다. 도로, 건물, 차, 나무 등과 같은 물체를 구성하는 점들의 구형 형상 표현을 뚜렷이 나타내는 노드의 행렬(matrix)을 다수 발견할 수 있다. 그러나 어떠한 형상을 잘 표현하지 못하는 노드의 행렬(matrix) 또한 다수 포진되어 있음을 알 수 있다. 하지만 이렇게 형상을 잘 표현하지 못하는 노드의 행렬(matrix)가 다수 포진되어 있다 할지라도 Fig. 8에서 확인하였듯이 이러한 현상으로 인해 분류 성능에 악영향을 미친다고 단정지을 수는 없다.

Table 2는 각 레이블에 대한 기존 연구 방법과 본 연구에서 제안한 방법의 분류 성능을 각각 나타내고 있다. 도로에 대한 분류 성능은 기존 87.45%에서 95.1%로 7.65%만큼 소폭 증가하였다. 이는 전체 도로 레이블 점

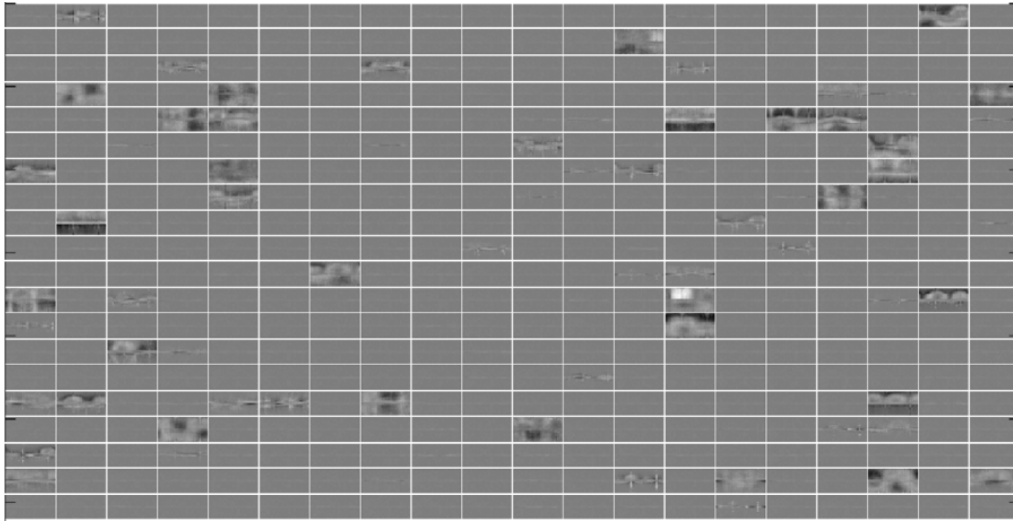


Fig. 9. Visualization result of weight parameters learned in the first layer

Table 2. Precision measure of the previous and proposed method

	Road	Building	Car	Tree	Others	Avg.
[12]	87.45	93.53	51.41	82.49	56.06	74.18
DBN	95.1	91.7	90.1	83.6	66.9	85.48
	7.65 ↑	1.83 ↓	38.69 ↑	1.11 ↑	10.84 ↑	11.3 ↑

데이터 35,465개 중 1,761개가 오분류된 수준이다. 건물에 대한 분류 성능은 기존 93.53%에서 91.7%로 1.83%만큼 소폭 감소하였다. 이는 전체 건물 레이블 점 데이터 5,435개 중 452개가 오분류된 수준이다. 차에 대한 분류 성능은 기존 51.41%에서 90.1%로 38.69%만큼 대폭 증가하였다. 이는 전체 차 레이블 점 데이터 6,821개 중 681개가 오분류된 수준이다. 나무에 대한 분류 성능은 기존 82.49%에서 83.6%로 1.11%만큼 소폭 증가하였다. 이는 전체 나무 레이블 점 데이터 4,343개 중 714개가 오분류된 수준이다. 결과에서 보듯이 차에 대한 점의 분류 성능이 월등히 증가했음을 알 수 있고 다른 레이블에서도 전반적으로 분류 성능이 증가했음을 확인할 수 있다.

6. 결 론

본 연구에서는 센서 데이터의 형상 분류를 위해 크게 구형 특징 표현과 환경 형상 학습, 이렇게 두 부분으로

나누어 수행하였다. 구형 특징 표현에서는 레이저 스캐너로 측정된 3차원 점 군의 단일 점에 대한 형상학적 정보를 직관적으로 서술할 수 있는 방법을 제안하였다. 3차원 점 군에 속해 있는 임의의 단일 점에 대한 형상학적 정보는 기본적으로 그 점이 중심점이 되는 단위 구의 표면에서 형상화한 후, 단위 구의 표면을 격자화 시켜 주변의 다른 점 들을 이용하여 베이지안 업데이트(Bayesian updating) 모델 기반의 점유 증거를 누적시키는 방식을 기반으로 하였다. 환경 형상 학습에서는 이렇게 형상화된 구형 특징 표현 결과들을 입력 값으로 하는 기계 학습의 한 종류인 심층 신뢰 회로망(Deep Belief Networks)을 사용하여 구형 특징 표현에 대한 환경 형상을 학습시켰다. 결과에서 확인하였듯 차에 대한 점의 분류 성능이 기존 방법에 비해 월등히 증가했음을 알 수 있고 도로, 나무 등 다른 레이블에서도 전반적으로 분류 성능이 증가했음을 확인할 수 있었다. 특히 차에 대한 인식률이 높아진 것은 추후 무인 자동차의 자율 주행 성능에 큰 영향을 미칠 수 있을 것으로 사료된다.

본 논문에서 사용된 검증을 위한 테스트 데이터 세트는 학습을 위한 데이터 세트와 구분하여 사용하였다고는 하나 동일한 장소에서 측정된 전체 점 군 데이터 내에서 생성하였기 때문에 다른 장소에서 측정한 데이터를 테스트 데이터 세트로 사용하여 본 논문에서 제안한 방법에 대한 신뢰도를 추후 재확인 해 보고자 한다.

후 기

이 논문은 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2015R1C1A1A01054724). 연구를 위해 실험 데이터를 제공해 주신 카이스트 Robotics Research 연구실에 감사드립니다.

References

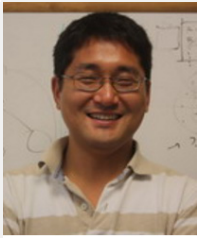
- [1] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NIPS, 2012.
- [2] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of Go with deep neural networks and tree search," *Nature* 529, pp. 484-489, January, 2016.
- [3] Sejin Lee, Dong-Woo Cho, and Jae-Bok Song, "Novel Sonar Salient Feature Structure for Extended Kalman Filter-Based Simultaneous Localization and Mapping of Mobile Robots," *Advanced Robotics*, vol. 26, no. 8-9, pp. 1055-1074, July 2012.
- [4] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, K. Lau, C. Oakley, M. Palatucci, V. Pratt, P. Stang, S. Strohband, C. Dupont, L. Jendrossek, C. Koelen, C. Markey, C. Rummel, J. van Niekerk, E. Jensen, P. Alessandrini, G. Bradski, B. Davies, S. Ettinger, A. Kaehler, A. Nefian, and P. Mahoney, "Stanley: The robot that won the DARPA Grand Challenge," *Journal of Field Robotics*, vol. 23, no. 9, pp. 661-692, September 2006.
- [5] Seunguk Ahn, Yungeun Choe, and Myung Jin Chung, "Fast Scene Understanding in Urban Environments for an Autonomous Vehicle equipped with 2D Laser Scanners," *Journal of Korea Robotics Society*, vol. 7, no. 2, pp. 92-100, June 2012.
- [6] Sejin Lee, Byung-Jae Park, Jong-Hwan Lim, Wan-Kyun Chung, and Dong-Woo Cho, "Grid Map Building through Neighborhood Recognition Factor of Sonar Data," *Journal of Korea Robotics Society*, vol. 2, no. 3, pp. 227-233, September 2007.
- [7] Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh, "A Fast Learning Algorithm for Deep Belief Nets," *Neural Computation*, vol. 18, no. 7, pp. 1527-1554, May 2006.
- [8] J.F. Lalonde, N. Vandapel, D.F. Huber, and M. Hebert, "Natural Terrain Classification using Three-Dimensional Ladar Data for Ground Robot Mobility," *Journal of Field Robotics*, vol. 23, no. 10, pp. 839-861, October 2006.
- [9] A. Golovinskiy, V.G. Kim, and T. Funkhouser, "Shape-based Recognition of 3D Point Clouds in Urban Environments," in *IEEE 12th International Conference on Computer Vision*, Tyoko, September 2009, pp. 2154-2161.
- [10] X. Zhu, H. Zhao, Y. Liu, Y. Zhao, and H. Zha, "Segmentation and Classification of Range Image from an Intelligent Vehicle in Urban Environment," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, October 2010, pp. 1457-1462.
- [11] H. Zhao, Y. Liu, X. Zhu, Y. Zhao, and H. Zha, "Scene Understanding in a Large Dynamic Environment through a Laser-based Sensing," in *IEEE International Conference on Robotics and Automation*, Anchorage, May 2010, pp. 128-133.
- [12] Yungeun Choe, Inwook Shim, and Myung Jin Chung, "Urban structure classification using the 3D normal distribution transform for practical robot applications," *Advanced Robotics*, vol. 27, no. 5, pp. 351-371, February 2013.
- [13] Yong-Ju Lee and Jae-Bok Song, "Three-dimensional iterative closest point-based outdoor SLAM using terrain classification," *Intelligent Service Robotics*, vol. 4, no. 2, pp. 147-158, 2011.
- [14] Martin Magnusson, "The Three-Dimensional Normal-Distributions Transform - an Efficient Representation for Registration, Surface Analysis, and Loop Detection," *Doctoral Dissertation*, Orebro University, 2009.
- [15] R. Socher, B. Huval, B.P. Bath, C.D. Manning, A.Y. Ng, "Convolutional-recursive deep learning for 3D object classification," *NIPS*, 2012.
- [16] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy Layer-Wise Training of Deep Networks," *Advances in neural information processing systems*, vol. 19, pp. 153-160, January 2007.
- [17] D.H. Ackley, G.E. Hinton, and T.J. Sejnowski, "A learning algorithm for Boltzmann machines," *Cognitive Science*, vol. 9, no. 1, pp. 147-169, 1985.



이 세 진

2003 한양대학교 기계공학과(공학사)
2005 포항공과대학교 기계공학과(공학석사)
2009 포항공과대학교 기계공학과(공학박사)
2013 ~ 현재 공주대학교 기계자동차공학부
조교수

관심분야: 이동로봇, 지도작성, 위치추정, 심화학습



김 동 현

2003 한양대학교 전자컴퓨터공학과(공학사)
2005 한양대학교 컴퓨터공학과(공학석사)
2010 University of Texas at Dallas 컴퓨터
공학과(공학박사)
2010 ~ 2016 North Carolina Central University
조교수

2016 ~ 현재 Kennesaw State University 조교수

관심분야: 보안, 사물인터넷, 빅데이터, 심화학습