

# 깊이 카메라를 활용한 제스처 인식 기반의 자율주행 차량용 HMI(Human Machine Interface) 시스템

## (HMI System for Autonomous Vehicles based on Gesture Recognition Using Depth Camera)

김 현 보<sup>\*</sup>, 김 용 훈<sup>†</sup>, 서 효 덕<sup>†</sup>, 윤 장 규<sup>†</sup>  
<sup>†</sup>경북IT융합산업기술원

(Hyeon-Bo Kim, Young-Hoon Kim, Hyo-Duck Seo, Jang-Kyu Yun)  
 (Gyeongbuk Institute of IT Convergence Industry Technology)

**Abstract** : Today, automobiles have established themselves as a representative mode of transport for transporting drivers and passengers to their destinations. The interior of such a vehicle consists of a space for accommodating occupants and a structure that has a great influence on occupant safety. Among the spaces for accommodating occupants for the task of 'driving', the driver's space is changing with the rise of autonomous driving technology. This change will lead to diversification of the infotainment system along with the change of the Human Machine Interface (HMI) between the driver and the vehicle. In this paper, we propose a gesture recognition-based vehicle content and display control system, a stable control interface that can ensure passenger comfort and pedestrian safety for commercialization of autonomous vehicles. The proposed system recognizes the user's gesture input using depth and color image data, and defines eight types of recognition gesture types and corresponding control scenarios according to the shape and movement of the hand. The performance of the system defines the image from the start frame to the last frame in which the target action occurs among a series of data collected through the sensor as a sequence. In addition, the recognition rate for each sequence was measured by performing predefined gestures within/outside the effective recognition distance.

**Keywords** : Gesture recognition, Depth camera, Hand gesture, HMI(Human Machine Interface)

### 1. 서 론

오늘날 자동차는 운전자와 탑승자를 목적지까지 수송하는 대표적인 운송수단으로써 자리매김하였다 [1]. 이러한 자동차의 실내는 탑승자 수용 공간과 탑승자 안전도에 큰 영향을 미치는 구조물로 이루어진다. 이 중 탑승자 공간은 '운전'이라는 작업을 위한 운전자 공간과 '거주성'이 중심이 되는 탑승자 공간으로 세분된다. 운전자가 운전이라는 작업을 안

전하게 수행하기 위해서는 전방 주시를 위한 차량 진행 방향 지향성과 적정 운전 자세 유지가 필요하다. 운전자 자세 기준은 발뒤꿈치를 기준점 (Accelerator Heel Point, AHP)으로 차량의 종류와 기능 등에 따라 높이와 위치는 변화하지만, 전방 지향성을 반드시 유지해야 한다는 특징을 지닌다. 그러나 자율주행 기술의 차량 적용을 통해 운전자 자세나 지향성의 중요성은 현저히 감소시킬 수 있다. 이로 인해 자율주행 차량의 실내는 운전 개념 변화로 운전석의 위치·비중 변화, 운전자 동선·동작·행위 변화로 공간 구성 변화, 주행 방향 지향성 요구 감소 또는 소멸, 좌석 기능 변화로 회전하거나 평평해지는 시트, 유효 점유 공간이나 공간 소비 증가, 주요 조작 장치의 의미나 형태 변화, 공간 활용에서 감성적 요소 비중 증가, 사고 확률 감소로 실내 충돌규제 비중 감소 등에 맞춰 변화하고 진화하

\* Corresponding Author (hbkim@gitc.or.kr)

H. Kim, Y. Kim, H. Seo, J. Yun: Gyeongbuk Institute of IT Convergence Industry Technology

※ 본 연구는 산업통상자원부와 한국산업기술진흥원이 지원하는 스마트특성화 기반구축사업으로 수행된 연구결과입니다 (과제번호: P0013655).

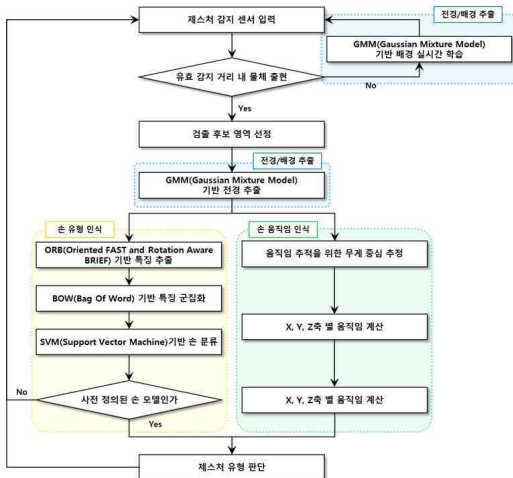


그림 1. 제스처 인식 시스템 흐름도

Fig. 1 Flow chart of hand gesture recognition system

게 될 것이다.

이러한 변화는 운전자와 차량 간 HMI(Human Machine Interface) 변화와 더불어 인포테인먼트 시스템의 다양화를 초래할 것이며 [2, 3], 자율주행 차량의 상용화에 승객의 안락함과 보행자의 안전을 보장할 수 있는 안정적인 제어 인터페이스인 제스처 인식 기반의 차량의 콘텐츠 및 디스플레이 제어 시스템을 제안한다.

## II. 본 론

### 1. 제스처 인식 시스템의 구성

본 논문에서 제안하는 자율주행 차량에 적용 가능한 제스처 인식 시스템은 깊이(Depth)와 색상(RGB color) 데이터 수집이 가능한 센서를 사용하며, 깊이 영상과 색상 영상을 모두 활용하여 사용자의 입력을 인식한다. 깊이 영상을 바탕으로 사용자의 입력인 제스처를 인식하기 위한 유효 인식 거리를 설정하고, 인식 거리 내 객체 출현 여부를 감지한다.

센서를 통해 수집되는 깊이 영상은 정확한 손 형태를 판단하기에 정밀도 측면에서 색상 영상보다 상대적으로 인식을 위한 특징이 부족하다. 이에 따라, GMM(Gaussian Mixture Model)을 활용하여 배경과 그림자, 전경의 분리를 수행하며 분리된 전경과 깊이 영상 내 출현 객체 정보를 활용하여 색상 영상 내 객체의 분리를 수행한다. 또한, 실시간으로 배경, 그림자 및 전경의 모델링 및 학습을 수



그림 2. 자율주행 차량 시스템 제어를 위한 제스처 유형 및 제어 시나리오

Fig. 2 Gesture types and control scenarios for autonomous vehicle system control

행함으로써 차량 실내 환경의 지속적인 변화에 적응이 가능하도록 모듈을 설계하였다.

출현 객체의 분류를 위해 사전 학습된 특징 추출기와 분류 모델을 사용하며, 색상 영상 내 추출 가능한 다양한 특징들을 바탕으로 사용자의 입력 유형 중 하나인 손 형태를 결정한다. 제스처 인식 시스템의 실시간성을 고려하여, 지역적 특징 추출을 위해 빠른 처리속도를 가지는 ORB(Oriented FAST and Rotation Aware BRIEF) 알고리즘을 적용하였다. 또한, BOW(Bag Of Word) 방식에 기반하여 지역적 특징을 군집화 및 손 형태별 핵심 특징 모델링을 수행한 후, SVM(Support Vector Machine)을 통해 출현 객체의 유형을 인식한다. 이때, 사용되는 분류기인 SVM 또한 실시간성 보장을 위해 멀티 클래스 모델이 아닌 바이너리 클래스 모델을 중첩하여 활용하였으며, SVM을 통한 객체 분류 시 회귀(Regression) 문제 판단하여 유형별 신뢰도 평가를 수행하여 제스처 인식 시스템의 인식 성능 확보하였다.

센서를 통해 수집되는 데이터를 바탕으로 x, y, z축에 대해 각각 독립적으로 추적하고, 축별 움직임을 종합함으로써 객체의 움직임을 추정함으로써 사용자의 입력 유형 중 움직임을 결정한다.

표 1. 인식 제스처 유형의 정의

Table 1. Definition of gesture type of recognition

구분	제스처 유형
1	(형태) 엄지와 검지를 사용하여 펜을 잡는 형태
	(움직임) 조수석에서 운전석으로 횡방향 이동
2	(형태) 엄지와 검지를 사용하여 펜을 잡는 형태
	(움직임) 운전석에서 조수석으로 횡방향 이동
3	(형태) 엄지와 검지를 사용하여 펜을 잡는 형태
	(움직임) 카메라와 가까워지는 수직방향 이동
4	(형태) 엄지와 검지를 사용하여 펜을 잡는 형태
	(움직임) 카메라와 멀어지는 수직방향 이동
5	(형태) 모든 손가락을 펼친 형태
	(움직임) 조수석에서 운전석으로 횡방향 이동
6	(형태) 모든 손가락을 펼친 형태
	(움직임) 운전석에서 조수석으로 횡방향 이동
7	(형태) 모든 손가락을 펼친 형태에서 모든 손가락을 접어 주먹을 권 형태
	(움직임) 없음
8	(형태) 모든 손가락을 접어 주먹을 권 형태
	(움직임) 카메라를 기준으로 수직방향 이동 반복

상기의 과정을 통해 추출된 손 형태 정보 및 움직임 정보를 바탕으로 제스처 유형을 판단하고, 사전 정의된 시나리오에 따라 차량의 콘텐츠 및 디스플레이를 제어한다. 그림 1은 본 논문에서 제안하는 제스처 인식 시스템의 흐름도이다.

## 2. 인식 제스처 유형 및 제어 시나리오

자율주행 차량의 콘텐츠 및 디스플레이 제어를 목적으로 제스처 인식 시스템을 적용할 경우, 제어를 위한 행동과 일상적으로 사용되는 행동이 구분되어야 한다. 반면, 시스템의 제어를 목적으로 하므로 제스처를 통한 제어 명령은 직관적으로 설계될 필요성이 있다. 이러한 점을 고려하여, 제안하는 시스템은 사용자의 손 형태와 움직임에 따른 8가지 인식 제스처 유형과 이에 상응하는 제어 시나리오를 표 1과 그림 2와 같이 정의한다.

## 3. 실험 및 결과

본 논문에서 제안하는 제스처 인식 시스템은 사용자 입력을 인식하기 위해 깊이와 색상 정보를 동시에 수집할 수 있는 Intel 社의 RealSense D435i [4]를 사용하였다. 센서는 차량 내부 룸미러와 앞

표 2. 제스처 인식 시스템 실험 결과

Table 2. Test result of gesture recognition system

in \ out	none	type1	type2	type3	type4	type5	type6	type7	type8
none	50							4	
type1		47				3			
type2			48				3		
type3				47					2
type4					46				2
type5		3				47			
type6			2				47		
type7								46	
type8				3	4				46
rate (%)	100	94	96	94	92	94	94	92	92

좌석의 조명 사이에 설치하는 것을 가정하여, 센서가 바닥을 향하도록 하여 실험을 수행하였다.

시스템의 성능은 센서를 통해 수집되는 일련의 데이터 중 인식 대상 행동이 발생하는 시작 프레임부터 마지막 프레임까지의 영상을 시퀀스로 정의하고, 유효 인식 거리 내/외에서 사전 정의된 제스처를 수행하여 시퀀스별 인식률을 측정하였다. 또한, 사용자의 입력에 따른 시스템의 결과를 혼동 행렬 (Confusion matrix)에 따라 분류하여 제스처 유형별 인식률을 평가하였다. 제안하는 시스템의 평가는 제스처 유형별로 50회씩 반복하여 수행하여, 실시간으로 인식되는 결과를 활용하여 측정하였다. 또한, 제스처 인식률은 식 (1)을 통해 도출하였으며, Odroid XU4 기반의 임베디드 시스템에서 평가를 진행하였다. 표 2는 제안하는 시스템의 인식률 평가 결과이며, 그림 3은 시스템의 인식 수행 결과이다.

$$\begin{aligned} \text{제스처인식률}(\%) &= \frac{(\text{정답으로 인식된 시퀀스의 수})}{(\text{전체 입력된 시퀀스의 수})} \times 100 \quad (1) \end{aligned}$$

## III. 결론

본 논문에서는 자율주행 차량의 콘텐츠 및 디스플레이 제어를 위한 제스처 인식 시스템을 제안하였다. 제안된 시스템은 운전자의 제스처를 인식하여, 사전 정의된 시나리오에 따라 제어 신호를 송출한

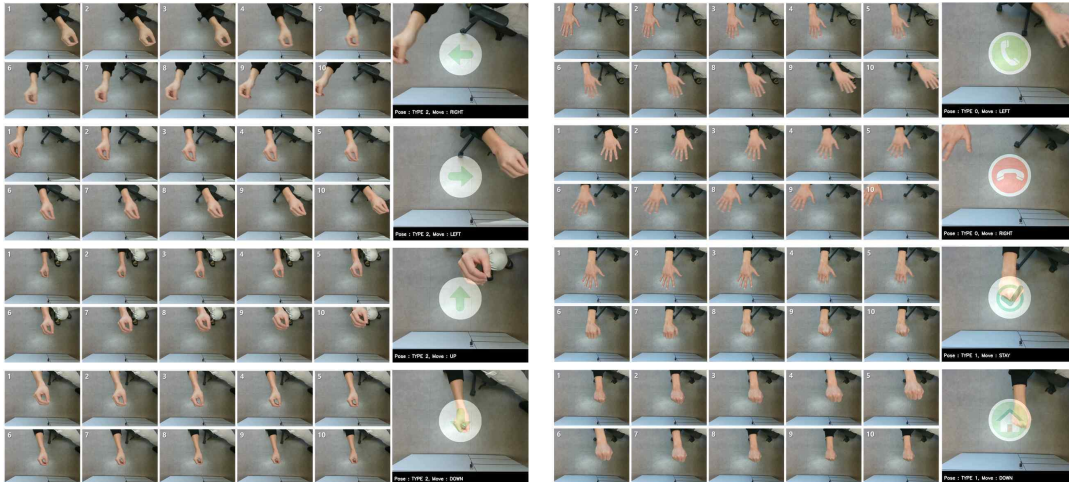


그림 3. 제스처 인식 시스템의 테스트 결과 예시  
 Fig. 3 Example of gesture recognition system test result

다. 개발 시스템은 평균 94.2%의 제스처 인식을 가지며, Odroid XU4 기반의 임베디드 시스템에서 10~15프레임의 처리 속도를 보였다. 향후, 최적화 및 인식 성능을 개선하여 시스템의 신뢰성을 확보할 예정이다.

### References

- [1] Zengeler, Nico, Kopinski, Thomas, Handmann, Uwe, "Hand Gesture Recognition in Automotive Human - Machine Interaction Using Depth Cameras," Journal of Sensors, Vol.19, No.59, 2018.
- [2] Pickering, Carl A., Keith J. Burnham, and Michael J. Richardson. "A research study of hand gesture recognition technologies and applications for human vehicle interaction." 3rd Conf. on Automotive Electronics. 2007.
- [3] Ohn-Bar, Eshed, and Mohan Manubhai Trivedi. "Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations." IEEE Transactions on Intelligent Transportation Systems 15.6 (2014): 2368-2377.
- [4] Intel RealSense D435, available on : <https://www.intelrealsense.com/depth-camera-d435/>