

# 자율주행자동차에서의 제어권전환을 위한 영상 기반 운전자 모니터링 기술 동향

이동환\*,\*\*, 김경호\*,\*\*, 김도현\*\*  
\*과학기술연합대학원대학교 컴퓨터소프트웨어전공  
\*\*한국전자통신연구원 지능로보틱스연구본부  
ehdgl@setri.re.kr, kkh@etri.re.kr, dohyun@etri.re.kr

## Vision-Based Driver Monitoring Technology Trend for Takeover in Autonomous Vehicles

Dong-Hwan Lee\*,\*\*, Kyong-Ho Kim\*,\*\*, Do-Hyun Kim\*\*  
\*Major. of Computer Software, University of Science and Technology  
\*\*Intelligent Robotics Research Division, ETRI

### 요 약

운전자가 아닌 자율주행 시스템이 운전을 주도하기 위한 기술의 상용화를 위해 많은 기업이 노력 중이다. 특히 운전자의 안전을 보장하기 위한 운전자와 자율주행 시스템 간의 제어권전환이 중요하다. 운전자의 주행과 관련 없는 행동은 제어권전환 상황에서 운전자를 위험에 빠뜨릴 수 있으므로 제어권전환을 돕기 위한 운전자 모니터링 기술에 관한 많은 연구가 진행되고 있다. 운전자 모니터링 기술은 주로 생체 정보, 차량 정보, 영상을 사용하여 운전자의 상태와 부주의 행동 등을 감지하는 기술이다. 최근 머신 러닝, 딥 러닝을 사용한 영상처리 및 인식 기술 등의 발전으로 영상을 사용한 운전자 모니터링 기술이 활발하게 연구되고 있다. 따라서 본 논문에서는 영상기반 운전자 모니터링 기술 동향에 대해 상세히 기술하였다. 특히 운전자의 부주의 행동 중 졸음은 운전자가 주행 상황을 전혀 인지하지 못하게 할 수 있어 더욱 위험한 행동이다. 따라서 영상기반 운전자 모니터링 기술을 졸음 인식과 그 외의 행동 인식으로 분류하여 동향을 정리하였다.

### 1. 서론

2019 년에 도로교통공단에서 발표한 자료에 따르면 대한민국에서 2014~2018 년에 발생한 379,285 건의 중상해 교통사고의 주요 원인은 안전운전의무 불이행 (56.3%) 으로[1], 이는 도로교통법에 구체적 조항으로 규정되지 않은 행동들(동승자와 대화, 졸음, 음식물 섭취, 스마트폰 사용 등)로 인해 발생한 사고이다. 이와 같은 운전자의 주행과 관련 없는 부주의 행동들을 NDRA(Non-Driving Related Activities)라고 하며 운전자가 NDRA 를 하더라도 레이다, 라이다, GPS, 카메라 등의 센서와 고정밀지도를 사용하여 안전하게 목적지에 도달하기 위한 자율주행 시스템 개발이 활발히 진행 중이다.

자율주행 시스템은 SAE J3016 표준 규격을 따른다. 0 단계는 사람이 직접 운전하는 단계이며, 1~2 단계는 시스템이 운전자를 보조하는 단계이다. 3 단계는 시스템이 운전을 주도하지만 제어권전환을 위해 운전자의 개입이 필요하다. 4~5 단계로 갈수록 운전자의 개입을

최소화하며 시스템이 운전을 주도한다.

국내의 자동차 제조 기업들은 기술 개발에 많은 투자를 통해 3 단계 기술이 적용된 아우디 A8, 창안 UNI-T, 테슬라 모델 3 자동차를 양산하고 있다. 뿐만 아니라 4~5 단계 자율주행 기술이 적용된 현대 아이오닉(2017,4 단계), BMW7 시리즈(2019,4 단계), 아우디 AI:ME(2020,4 단계), GM 크루즈 오리진(2020,5 단계) 자율주행 자동차를 공개했다. 하지만 나라마다 3 단계 이상의 자율주행 시스템에 대한 안전기준이 미흡하고, 미국과 중국의 무역전쟁으로 인해 기업들의 투자가 감소하고, 도로 상황과 지형 정보를 제공하는 고정밀 지도 제작 시간이 길어지는 등의 이유로 대부분 기업의 상용화 된 자율주행 자동차들은 운전자를 보조하는 수준의 2 단계 기술이 적용되어 있다.

3 단계 자율주행에서는 운전자의 안전을 위한 제어권전환이 무엇보다 중요하다. 하지만 2014~2017 년에 캘리포니아에서 발생한 자율주행 자동차 실험 중 사고 통계에 따르면 사고 상황에서 시스템과 운전자 간

의 제어권전환을 통해 사고를 대비한 경우는 총 26건의 사고 중 3건에 불과하였다.[2]

제어권전환을 돕기 위해 주로 생체 정보, 차량 정보, 영상을 사용하여 운전자의 상태나 NDRA 를 감지하는 운전자 모니터링 기술에 관한 연구가 진행되고 있다.

본 논문에서는 생체 정보 및 차량 정보기반 운전자 모니터링 기술 동향을 간략히 기술하고, 최근 머신러닝, 딥 러닝 기술의 발전에 힘입어 활발히 연구되고 있는 영상기반 운전자 모니터링 기술에 대해서 보다 상세한 동향을 정리하였다.

## 2. 운전자 모니터링 기술

운전자 모니터링 기술은 주행 중인 운전자의 상태나 NDRA 등을 감지하는 기술이며 생체 정보를 사용하는 생체기반 기술, 차량 정보를 사용하는 차량 정보기반 기술, 영상을 사용하는 영상기반 기술로 구분할 수 있다.[3]

### 2.1 생체기반 기술

생체기반 기술은 운전자의 신체에 생체 센서를 부착하여 얻을 수 있는 뇌전도, 심전도, 안구 전도 등의 생체 정보를 수집하고, 분석하는 방법을 통해 운전자의 스트레스, 심리 상태, NDRA 등을 감지하는 기술이다.

생체기반 기술은 운전자의 신체에서 나온 생체 신호를 통해 운전자의 상태나 NDRA 등을 빠르게 감지할 수 있다. 하지만 운전자 신체에 부착된 생체 센서가 운전자의 불편을 유발하여 주행 집중을 방해할 수 있음을 고려해야 한다.

### 2.2 차량 정보기반 기술

차량 정보기반 기술은 차량으로부터 얻을 수 있는 조향각, 차량 속도, 스티어링 포지션 등의 차량 정보를 수집하고, 분석하는 방법을 통해 운전자의 주행 상태나 NDRA 등을 감지하는 기술이다.

차량 정보는 차량의 고장 상태를 나타내는 DTC 와 운전자의 차량 주행을 통해 수집한 조향각, 차량 속도, 스티어링 포지션 등이 있다. DTC는 차량 결함을 알려주는 정보이며, 이는 운전자의 상태나 NDRA 등을 감지하는데 주로 사용되는 정보는 아니다. 운전자의 차량 주행을 통해 수집한 조향각, 차량 속도, 스티어링 포지션 등이 운전자의 상태나 NDRA 를 감지하기 위해 사용되는 정보이다.

차량 정보기반 기술은 자동차에 탑재된 OBD-II 커넥터에 다양한 OBD-II 스캐너를 연결하면 쉽게 차량 정보를 수집할 수 있다. 하지만 운전자의 나이, 성별, 운전 숙련도 등의 요인에 의하여 차량 정보가 운전자마다 큰 차이가 날 수 있음을 고려해야 한다.

## 2.3 영상기반 기술

영상기반 기술은 차량 내부에 카메라를 설치하여 촬영한 운전자의 영상을 수집하고, 분석하는 방법을 통해 운전자의 감정 상태나 NDRA 등을 감지하는 기술이다.

최근 머신러닝, 딥 러닝 기술을 사용한 영상처리 및 인식 기술 등의 발전으로 운전자의 상태나 NDRA 등을 감지하는 속도와 정확도 등을 향상시키는 연구가 활발하게 진행되고 있다. 하지만 주, 야간 운전자 촬영 시 변하는 빛의 밝기, 운전자별 최적의 촬영 영역 등을 고려해야 한다.

## 3. 영상기반 운전자 모니터링 기술

앞서 설명한 머신러닝, 딥러닝 기술을 사용한 영상처리 및 인식 기술 등의 발전으로 영상기반 운전자 모니터링 기술에 관한 연구가 활발하게 진행되고 있다.

운전자의 NDRA 는 운전자의 집중력을 분산시키며, 이는 고속으로 주행하는 상황에서 사고를 유발할 수 있는 위험한 행동이다. 특히 NDRA 중 졸음은 운전자가 주행 상황을 전혀 인지하지 못하게 하며, 운전자가 사고를 대비하지 못하여 더 큰 사고를 유발할 수 있는 행동이다. 따라서 영상기반 운전자 모니터링 기술을 졸음 인식과 그 외의 NDRA 인식으로 구분하여 정리하였다.

### 3.1 졸음 인식

졸음 인식은 운전자를 촬영한 영상의 특징 영역을 추출하고, 특징 영역과 관련된 파라미터를 분석하여 운전자의 안전운전과 졸음 행동을 감지하며, 경고 알람을 통해 운전자가 졸음을 극복하도록 하는 목적이 있다.

졸음 인식에 주로 사용되는 특징 영역 및 파라미터를 졸음 인식 관련 동향분석을 통하여 정리하였고, 이를 사용한 일반적인 졸음 인식 프로세스를 제시하였다.

#### 3.1.1 특징 영역 및 파라미터

졸음 인식은 주로 안전 운전자와 졸음 운전자를 촬영한 영상의 변화 차이가 명확하게 나타나는 눈, 입, 머리 특징 영역을 추출하고, 특징 영역과 관련된 파라미터를 분석하여 졸음과 안전운전을 감지한다. 보통 하나의 특징 영역을 사용하며, 최근 하나 이상의 특징 영역을 사용하여 졸음 인식 정확도 등을 향상시키기 위한 연구가 진행되고 있다.

눈은 졸음 시 나타나는 깜빡임 횟수 감소, 눈 감긴 시간 증가 등을 통해 안전운전과 졸음을 감지할 수 있으며, 추출하기 쉬워 졸음 인식에 주로 사용되는 영역이다. 졸음 인식을 위한 눈과 관련된 파라미터는 PERCLOS(Percentage of Eye Closure), EAR(Eye aspect ratio), 깜빡임 횟수 등이 있다.

PERCLOS 는 특정 시간 동안 눈이 감긴 영상 비율을 계산한다. 눈 감김을 판단하는 비율은 70% 이상, 80% 이상 등이 있으며, 졸음과 상관관계가 더 높은 80% 이상이 주로 사용된다.[4] 시간은 30 초, 1 분, 3 분, 20 분 등을 사용하며, PERCLOS 계산 정확도와 실행 시간 사이의 트레이드 오프를 고려해야 한다.

EAR 은 눈의 특징점을 사용하여 가로와 세로 비율을 계산한다. EAR 을 통해 눈이 감겼는지, 열렸는지를 알 수 있고, 사람마다 변할 수 있는 눈 위치로 인한 계산오차 등을 감소시킬 수 있다.[5]

깜빡임 횟수는 눈이 열린 다음 눈이 감긴 횟수를 계산한다. 일반적으로 1 분당 눈 깜빡임 횟수는 성인 남성은 20 회, 여성은 15 회 정도로 알려져 있으며, 졸음 시 눈 깜빡임 횟수가 감소한다.

입은 지루함, 피로 등의 원인으로 발생하는 운전자의 하품을 통해 안전운전과 졸음을 감지할 수 있다. 입과 관련된 파라미터는 MAR(Mouth Aspect Ratio) 등이 있다.

MAR 은 입의 특징점을 사용하여 가로와 세로 비율을 계산한다. MAR 을 통해 입이 닫혔는지, 열렸는지를 알 수 있다. 하지만 단순히 입이 닫혔는지, 열렸는지만으로는 졸음을 감지하기 어렵다. 운전자가 대화, 음료 섭취, 기침 등의 행동 시 마찬가지로 입이 열리고, 닫히기 때문이다. 따라서 입이 열린 정도와 시간 등을 고려해야 한다. 일반적으로 하품 시 입이 더 넓게 열리고, 약 6 초간 지속된다.

머리는 졸음 시 머리 무게를 지탱하는 근육들이 이완되어 제 기능을 수행하지 못해 앞으로 기울어짐을 통해 안전운전과 졸음을 감지할 수 있다. 머리와 관련된 파라미터는 머리 자세 등이 있다.

머리 자세는 2 차원, 3 차원 영상을 통해 Yaw, Pitch, Roll 각도 변화를 계산하며, 3 차원 영상을 사용하면 계산 정확도를 향상시킬 수 있다.[6]

### 3.1.2 프로세스

졸음 인식 프로세스는 일반적으로 운전자의 안전운전 및 졸음 영상을 수집하는 영상 수집단계, 앞서 설명한 특징 영역을 추출하고, 파라미터를 계산하는 특징 영역 추출 단계, 파라미터를 분석하여 안전운전과 졸음 행동을 감지하는 졸음 인식 단계 순으로 진행된다.

영상 수집단계는 운전자의 모습을 촬영하여 영상을 수집하는 단계부터 시작한다. 운전자를 직접 촬영한 영상을 수집하기 어려운 경우 NTHU, UTA 등에서 만든 운전자 졸음 영상을 사용하면 영상을 직접 촬영하지 않고 영상을 수집할 수 있다.

특징 영역 추출단계는 영상에서 앞서 설명한 눈, 입, 머리 특징 영역을 추출하고, 파라미터를 계산하는 단계이다. 특징 영역 추출 속도 및 정확도를 향상시키기 위해서 운전자의 얼굴을 먼저 탐색한 다음 특징 영역을 추출하는 방법 등이 연구되고 있다.

졸음 인식단계는 파라미터 계산 결과를 분석하여 안전운전과 졸음을 감지하는 단계이다. 머신 러닝은 안전운전과 졸음 시 나타나는 파라미터 계산 결과의

특징을 기계가 스스로 분석하여 안전운전과 졸음을 감지한다. 딥 러닝은 머신 러닝의 한 분야로 인간의 뉴런 구조와 비슷한 인공신경망을 사용하며, 하드웨어, 빅 데이터 등의 발전으로 활발히 연구되고 있다. 최근 딥 러닝을 통해 파라미터를 계산하지 않고, 특징 영역만 추출하여 안전운전과 졸음 감지 정확도 등을 향상시키기 위한 연구가 진행되고 있다.[7] 졸음을 감지하면 경고 알람을 통해 운전자가 위험한 상태에서 벗어날 수 있도록 한다.

졸음 인식에 주로 사용되는 특징 영역 및 파라미터를 사용한 일반적인 졸음 인식 프로세스를 (그림 1)에 제시하였다.



(그림 1) 졸음 인식 프로세스

## 3.2 NDRA 인식

NDRA 인식은 운전자를 촬영한 영상의 특징 영역을 추출하고, 분석하여 운전자의 안전운전과 통화, 문자 전송, 음료 섭취 등의 NDRA 를 감지하며, NDRA 를 감지한 경우 알람 등을 통해 운전자가 안전운전에 집중하도록 하는 목적이 있다.

NDRA 인식에 주로 사용되는 특징 영역을 NDRA 인식 관련 동향분석을 통하여 정리하였고, 이를 사용한 일반적인 NDRA 인식 프로세스를 제시하였다.

### 3.2.1 특징영역

NDRA 인식은 주로 신체 모습, 손, 얼굴 특징 영역을 영상에서 추출하여 사용한다. 특히 신체 모습은 영상에서 추출하기 쉬워 다양한 행동들을 구별하기 위해 많이 사용되는 특징 영역이다.

얼굴 및 손 또한 NDRA 인식을 위해 사용되는 특징 영역이며, 영상에서 얼굴만 추출하거나, 손만 추출한다. 얼굴 또는 손 영역을 추출하지 못하면 안전운전과 NDRA 인식 정확도에 영향을 줄 수 있으며, 영역 추출을 하지 못한 경우 다른 영상으로 대체하는 방법 등을 적용해야 한다.[8]

최근 안전운전과 NDRA 인식 정확도 등을 향상시키기 위하여 영상에서 피부색을 통해 얼굴과 손 등의 특징 영역을 같이 추출하여 안전운전과 NDRA 를 인식하는 방법에 관한 연구가 진행되고 있다.

### 3.2.2 프로세스

NDRA 인식 프로세스는 일반적으로 운전자의 안전운전 및 NDRA 영상을 수집하는 영상 수집단계, 앞서

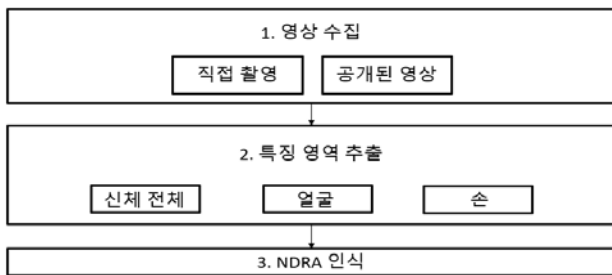
설명한 특징 영역을 추출하는 특징 영역 추출 단계, 특징을 분석하여 안전운전 행동, NDRA 를 감지하는 NDRA 인식 단계 순으로 진행된다.

영상 수집단계는 운전자를 촬영한 영상을 수집하는 단계이다. 운전자의 안전운전과 다양한 NDRA 영상을 수집하기 어려운 경우 StateFarm, AUC 등에서 만든 운전자 NDRA 영상을 사용하면 영상을 직접 촬영하지 않고, 안전운전, 문자 전송, 통화, 음료 섭취, 동승자와 대화, 화장, 라디오 작동 등의 10 개 행동으로 구분된 영상을 수집할 수 있다.[9]

특징 영역 추출 단계는 앞에서 설명한 신체 전체, 얼굴, 손 특징 영역을 영상에서 추출하는 단계이다.

NDRA 인식단계는 특징을 분석하여 안전운전과 NDRA 등을 감지하는 단계이다. 머신 러닝은 추출한 특징 영역의 특징점을 비교하는 방법 등을 통해 안전운전과 다양한 NDRA 를 감지한다. 최근 딥 러닝을 사용하여 특징점을 비교하지 않고 NDRA 인식 정확도와 속도 등을 향상시킬 수 있는 방법이 연구되고 있다.[10]

NDRA 인식에 주로 사용되는 특징 영역을 사용한 일반적인 NDRA 인식 프로세스를 (그림 2)에 제시하였다.



(그림 2) NDRA 인식 프로세스

#### 4. 결론

현재 자율주행 시스템 기술은 운전자를 보조하는 단계 수준이며, 다음 단계로 넘어가기 위해선 무엇보다 원활한 제어권전환을 통해 운전자의 안전을 보장할 수 있어야 한다. 제어권전환을 돕기 위한 생체 정보, 차량 정보, 영상을 사용한 운전자 모니터링 기술에 관한 많은 연구가 진행되고 있다.

본 논문에서는 운전자 모니터링에 대해 간단히 정리한 뒤 특히 머신 러닝, 딥 러닝 기술 등의 발전으로 많은 연구가 진행 중인 영상기반 운전자 모니터링에 대해 상세히 정리하였다. NDRA 중 졸음은 운전자가 빠르게 주변 상황을 의식하지 못하게 하여 더 큰 피해를 유발할 수 있는 행동이다. 따라서 NDRA 를 졸음 인식과 그 외의 NDRA 인식으로 나누어 동향 분석을 수행하였다. 동향 분석을 통해 졸음과 NDRA 인식에 주로 사용하는 특징 영역 등을 파악할 수 있었고, 이를 사용한 일반적인 졸음과 NDRA 인식 프로세스를 제시하였다.

#### Acknowledgement

본 연구는 한국전자통신연구원 연구운영비지원사업의 일환으로 수행되었음. [20ZS1200, 인간 중심의 자율지능시스템 원천기술 연구]

#### 참고문헌

- [1] 도로교통공단, “최근 5 년간 중상해 교통사고 특성 분석”, 2019, [online] Available: [http://taas.koroad.or.kr/web/bdm/srs/selectStaticReportsDetail.do/중상해 교통사고 특성분석\(게시용\).pdf](http://taas.koroad.or.kr/web/bdm/srs/selectStaticReportsDetail.do/중상해 교통사고 특성분석(게시용).pdf)
- [2] Favaro FM, et al. “Examining accident reports involving autonomous vehicles in California”, PLoS ONE, 12, 9, p. e0184952, 2017.
- [3] 오영달, 류동운, 박선홍, “자율주행자동차 운전제어권 전환을 위한 운전자 모니터링 기술 동향”, 한국통신학회지(정보와통신), 35, 12, 29-35, 2018.
- [4] D. F. Dinges, et al. “Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management”, NHTSA, 1998.
- [5] T. Soukupova and J. Cech, “Real-time eye blink detection using facial landmarks”, 21st Computer Vision Winter Workshop(CVWW), Slovenia, 2016.
- [6] V.K. Diddi and S. Jamge, “Head pose and eye state monitoring (HEM) for driver drowsiness detection: overview”, International Journal of Innovative Science, Vol. 1, Issue 9, 504–508, 2014.
- [7] M. F. Shakeel, et al. “Detecting Driver Drowsiness in Real Time Through Deep Learning Based Object Detection”, International Work-Conference on Artificial Neural Networks(IWANN), Spain, pp. 283-296, 2019.
- [8] Hesham M. Eraqi, et al. “Driver Distraction Identification with an Ensemble of Convolutional Neural Networks”, Journal of Advanced Transportation, vol. 2019, 2019, pp.1-12.
- [9] State Farm, State Farm Distracted Driver Detection, [online] Available: <https://www.kaggle.com/c/state-farm-distracted-driver-detection/data>.
- [10] Bhakti Baheti, Suhas Gajre, Sanjay Talbar, “Detection of Distracted Driver using Convolutional Neural Network”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR) Workshops, United States, 2018, pp. 1032-1038