

MediaPipe를 활용한 제스처 인식 기반 자연 사용자 인터페이스(NUI) 시스템 제안

황석영¹⁾ · 이동운²⁾ · 조해인³⁾ · 박효상⁴⁾ · 박기철^{*5)}

홍익대학교 국제디자인대학원 스마트디자인엔지니어링¹⁾, 홍익대학교 일반대학원 기계공학과²⁾,
경기대학교 예술대학원 시각속디자인³⁾, 홍익대학교 국제디자인대학원 디자인학⁴⁾, 홍익대학교 공과대학 기계시스템디자인^{*5)}

Proposal for a Natural User Interface (NUI) System Based on Gesture Recognition Using MediaPipe

Seokyoung Hwang¹⁾, Dongwoon Lee²⁾, Haein Cho³⁾, Hyosang Pak⁴⁾, Kicheol Pak^{*5)}.

¹⁾Master, Smart Design Engineering, Hongik University, 94, Wausan-ro, Mapo-gu, Seoul, Korea

²⁾Master, Mechanical Engineering, Hongik University, 94, Wausan-ro, Mapo-gu, Seoul, Korea

³⁾Master, Visual Communication & Metal Design, Kyonggi University, 154-42, Gwanggyosan-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, Korea

⁴⁾Ph.D. Student, Design Studies, Hongik University, 94, Wausan-ro, Mapo-gu, Seoul, Korea

^{*5)}Prof., Mechanical System Design and Engineering, Hongik University, 94, Wausan-ro, Mapo-gu, Seoul, Korea

Abstract : 본 연구는 기존 GUI(Graphical User Interface) 기반 차량 인터페이스의 조작 복잡성과 이에 따른 인지적 부담을 개선하기 위해, 직관적인 제스처 기반 자연 사용자 인터페이스(Natural User Interface, NUI) 시스템의 설계를 목적으로 한다. 이를 위해 차량 내 주요 편의장치인 선루프와 공조 장치를 대상으로, MediaPipe 기반 손 좌표 인식, 제스처 입력 알고리즘, Dynamixel 모터 제어 방식이 통합된 피지컬 컴퓨팅 구조를 설계하였다. 실제 차량 환경을 가정하여 입력-처리-출력 흐름에 따른 시스템 시퀀스를 구성하고, 하드웨어와 소프트웨어 연동을 고려한 동작 체계를 정의하였다. 또한 기존 NUI 시스템의 인식 오류, 피드백 부족, 사용자 위치 반영 미흡 등의 한계를 보완하고자, 카메라 위치와 제스처 방식을 주요 실험 변수로 설정하였다. 해당 평가 지표는 향후 인식 정확도, 응답 시간, 사용성 만족도 등을 중심으로 한 정량적·정성적 평가 실험의 기반으로 활용될 예정이다. 본 연구는 다양한 사용자 환경을 고려한 인터랙션 설계와 몰입형 사용자 경험 강화, 다기능 차량 인터페이스 통합을 위한 후속 연구에 활용될 수 있다.

Key words : 그래픽 사용자 인터페이스(GUI), 자연 사용자 인터페이스(NUI), 머신러닝(Machine Learning), 차량 내 인터페이스(In-vehicle Interface), 프로토타입 시스템(Prototype System)

1. 서론

1.1 연구 배경

사용자 인터페이스(User Interface, UI)는 '직접 조

작(Direct Manipulation)' 개념을 기반으로, CLI(Command Line Interface)에서 GUI(Graphical User Interface)로 발전하며 사용자 인지 부담을 줄이고 접근성과 사용성을 크게 향상시켰다 (Shneiderman, 1983). 그러나 GUI는 2D 평면 기반의

* 교신저자, E-mail: hide@hongik.ac.kr

상호작용 방식에 제한되어 있어, 복잡한 조작이나 정보 탐색 과정에서 여전히 높은 인지적 부담을 초래한다는 지적이 지속되어 왔다. 이러한 한계를 보완하고자 등장한 자연 사용자 인터페이스인 NUI(Natural User Interface)는 제스처, 음성, 시선 등 사용자의 직관적인 행위를 기반으로 상호작용을 구성하며, 별도의 사전 교육 없이도 사용자가 점진적으로 기능을 익히도록 유도하는 ‘스캐폴딩(Scaffolding)’ 방식의 설계를 통해 몰입감 있는 사용 경험을 제공하는 것이 특징이다(Wigdor & Wixon, 2011). NUI는 기존 GUI가 제공하지 못했던 3차원 공간에서의 자유로운 상호작용을 가능하게 하며, 증강현실(AR), 가상현실(VR) 등 차세대 인터페이스 기술의 핵심 기반으로 자리잡고 있다. 이에 따라, 맨손 제스처 기반 상호작용 시스템(서동우 외, 2012), 환경 인지형 UX 설계(전관중, 2016), 차량용 멀티모달 인터페이스(김명섭 외, 2019) 등 다양한 형태의 NUI 응용 연구가 진행되어 왔으며, NUI가 사용자 조작의 직관성을 높이고 인지적 부담을 완화하는 가능성을 보여준 바 있다. 본 연구에서 참고한 주요 선행연구는 다음과 같이 표로 나타내었다.

Table 1. Preliminary research on NUI

구분	연구 주요 내용	연구자명 (년도)
증강현실 기반 NUI 개발	증강현실 환경에서 맨손 기반 상호작용 기술을 개발하여 NUI의 가능성을 입증. Depth Camera와 Natural Marker 기반 알고리즘을 통해 직관적인 2D-3D 좌표 변환을 구현	서동우 외(2012)
환경 인지 기반 UX 설계	환경 복합 센서를 활용하여 사용자인지 기반 UX 설계를 제안. 날씨 정보와 사용자 행동을 통합하여 NUI 기반의 직관적 입력 방식을 통해 사용자 인지 시스템의 가능성을 개념적으로 제시	전관중 (2016)
차량용 멀티모달 인터페이스	대면적 윈드셸드 기반 헤드업 디스플레이와 터치 제스처, 음성 명령을 결합한 멀티모달 인터페이스 프로토타입을 개발. 주행 시물레이션 기반 실험을	김명섭 외(2019)

	통해 운전자 워크로드 감소 및 주행 안전성 향상 효과를 실증	
--	-----------------------------------	--

상기 선행 연구들은 NUI가 사용자의 자연스러운 행동을 반영하여 조작의 직관성을 높이고 인지적 부담을 감소시키는 데 기여할 수 있음을 시사한다. 그러나 기존 제스처 기반 NUI 시스템은 의도하지 않은 동작에 대한 잘못된 반응(False responses)이나 응답 없음(Feedback is lacking) 등으로 인해 인식의 명확성이 떨어지며, 그 결과에 대한 피드백이 부족하다는 한계가 존재한다(Norman, 2010).

1.2 연구 목적

본 연구의 목적은 기존 GUI 기반 차량 인터페이스가 지닌 조작의 복잡성, 물리적 위치 제약, 정보 약자에 대한 접근성 한계 등을 개선할 수 있는 직관적이고 사용자 중심의 제스처 기반 자연 사용자 인터페이스(Natural User Interface, NUI) 시스템을 설계하고 구현하는 데 있다. 이를 위해 실제 차량 환경을 고려한 상황에서 작동 가능한 NUI 프로토타입의 개발을 중심으로 연구를 수행하였으며, 제스처 인식 알고리즘 설계, MediaPipe 기반 손 좌표 추적, 모터 제어 알고리즘 구현, 그리고 선루프 및 공조 장치 제어를 위한 피지컬 컴퓨팅 구조 설계를 포함하였다. 구축된 시스템은 카메라 위치와 제스처 방식에 따른 입력 처리 구조를 고려하여, 추후 다양한 사용 시퀀스에서의 적용 가능성을 탐색할 수 있도록 구성되었다. GUI 시스템 대비 NUI 시스템의 특징은 다음과 같이 표로 나타내었다.

Table 2. Characteristics of NUI compared to GUI

구분 (유형)	GUI	NUI
주요 특징	그래픽 요소(아이콘, 버튼, 창 등)를 이용하여 조작	인간의 자연스러운 동작, 제스처, 음성, 시선 등을 활용
입력 방식	마우스, 키보드, 터치스크린	손짓, 시선, 음성, 터치, 몸동작
출력 방식	화면에 그래픽 요소 표시	화면, 햅틱 피드백, 소리, 기능 수행
구성 요소	아이콘, 메뉴, 버튼, 윈도우	터치 제스처, 음성 명령, 얼굴 인식,

	스크롤바, 체크박스, 라디오 버튼 외	동작 인식 외
연구자명 (년도)	Wigdor & Wixon (2011)	

2. 시스템 설계

시스템 설계는 퍼지컬 컴퓨팅을 위해 크게 하드웨어와 소프트웨어로 이루어져 있다. 하드웨어의 경우 제어를 위한 PC, 웹캠, Dynamixel 모터, 차량편의 장치(선루프, 공조 장치)의 프로토타입, 통신장치(U2D2)로 구성되어 있으며, 소프트웨어는 제스처를 인식하는 실행파일과 인식된 제스처를 바탕으로 모터를 제어하는 실행파일 두 가지로 구성되어 있다. 이를 그림으로 나타내면 다음과 같다.

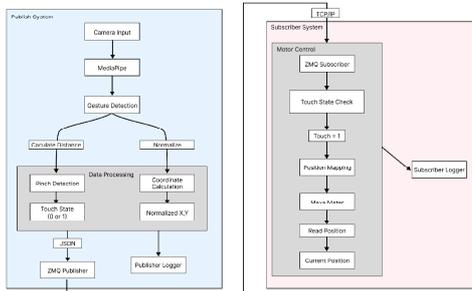


Fig 1. Algorithm architecture

2.1 소프트웨어

상기 그림<Fig 1>로 언급한 바와 같이 두 가지 소프트웨어 중에 Publisher 파일은 카메라를 통해 Input을 받고 제스처 상태 인식과 손의 위치를 x, y 축에 대하여 데이터를 변환하고, Subscriber 파일은 변환된 값을 받아 제스처 상태에 따라 모터를 제어한다. Publisher 파일의 경우 Fig 2에서와 같이 다시 “Pinch and drag” 방식과 “Pinch to drag” 방식으로 나뉘는데, “Pinch and drag” 방식의 경우 검지와 엄지가 맞닿아있는 상태를 Touch=1, 떨어져 있는 상태를 Touch=0으로 엄지와 검지가 맞닿은 상태에서 만 손이 움직일 때 모터를 구동하게 되는 방식이다. “Pinch to drag” 방식은 초기 실행 시 Touch=0이며, 검지와 엄지가 닿았다가 떨어질 때마다 On/Off(Touch=1, Touch=0) 값이 변환되며

On(Touch=1)일 때만 손이 움직일 때 모터가 구동된다. 이를 그림으로 나타내면 다음과 같다.

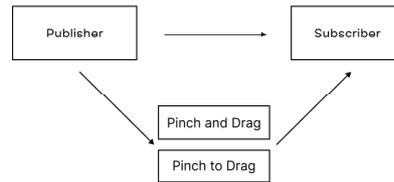


Fig 2. Drag type



Fig 3. Pinch and drag

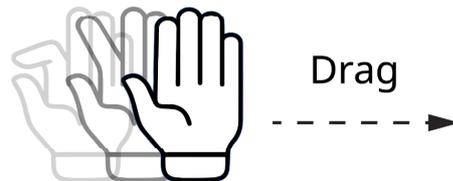


Fig 4. Pinch to drag

이때 손의 위치는 검지 끝의 좌표점을 인식한다. 카메라 프레임을 X, Y축에 따라 0~255로 나눈 뒤 해당 값을 변환하여 모터 구동 범위(ex) 0°~180°에 매핑하는 방식으로 데이터를 처리한다. 다수의 모터를 활용할 시에는 공조 장치 제어 등의 시퀀스와 같이 필요시 2-DOF의 구동이 가능하다. 손 인식은 MediaPipe7)를 통해 인식하며 각 손가락의 끝과 관절 점 등 21개의 포인트를 인식할 수 있으며 해당 포인트 간의 거리를 계산하는 방식으로 작성되었다. 예시 이미지는 다음과 같다.

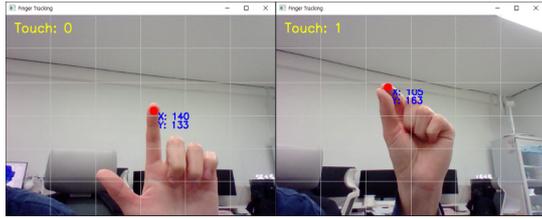


Fig 5. Hand Recognition via Webcam

모터 제어 함수 외에도 성능평가를 위한 데이터 저장 함수가 포함되어 있는데, Publisher 파일과 Subscriber 파일에 동일하게 데이터를 로깅하고 저장하는 함수이다. 로그에는 시간, 목표 위치, 실제 위치, 동작 종료 시각 등이 있으며 CSV 형식으로 저장되며 해당 로그 파일을 통해 인식 방식, 카메라 위치, 사용자 위치 등에 따른 인식 정확도와 레이턴시 등을 비교 및 분석을 가능하게끔 한다.

3. 프로토타입 설계

3.1 NUI 적용 사례 분석

실험 설계 이전, 차량 내 NUI 시스템의 초기 시퀀스 설계를 위해 NUI가 사용된 실제 제품 및 차량 적용 사례를 조사하였으며 이를 표로 정리하였다.

Table 3. Products and applications using NUI

이미지	사용 단계 요약	제품/회사명, 연도)
	차량 시동 시스템 자동 활성화 → 센터페시아 앞 제스처 수행 → 기능 실행 → 피드백 제공	BMW Gesture Control (2015)
	착용 및 초기화 → 시선 추적 활성화 → 손 제스처 인식 → 가상 객체 조작 → 작업 완료	Microsoft HoloLens 2 (2019)
	시스템 활성화 → 제스처 인식 및 명령 수행 → 피드백 제공	기아자동차 V 터치 (2019)

	시스템 활성화 → 제스처 수행 → 명령 인식 및 수행 → 피드백 제공	현대모비스 가상공간 터치 (2020)
	착용 및 초기화 → 시선 추적 및 손 제스처 활성화 → 공간 내 객체 조작 및 앱 실행 → 시각적 피드백 제공	Apple Vision Pro(2024)

상기 제품 및 적용 서비스는 제스처 및 시선 인식 기반의 NUI 시스템이 도입되어 오디오 제어, 내비게이션 조작, 인포테인먼트 기능 활성화 등 다양한 편의 기능을 지원하고 있다. 그러나 기존의 NUI 시스템은 한계점을 가지며 이를 다음과 같이 표로 나타내었다.

Table 4. Limitations of traditional NUI systems

유형	한계	보고 사례
제스처 인식 정확도 저하	외부 조명, 사용자의 위치, 손의 각도 변화에 따라 제스처 인식 오류가 발생하며, 명령이 인식되지 않거나 잘못 수행되는 문제가 보고됨	Frison et al., 2017
다중 사용자 권한 관리 미흡	운전자와 조수석 등 다수 사용자가 동일한 물리적 공간에서 제스처를 사용할 경우, 시스템이 사용자 간 명령을 구분하지 못해 입력 충돌이나 기능 오작동이 발생할 수 있음	김명섭 외. 2019
피드백 시스템의 부족	제스처 수행 결과에 대한 시각적 또는 청각적 피드백이 부족하여, 사용자 입장에서는 시스템의 현재 상태나 동작 성공 여부를 파악하기 어려움	Wigdor & Wixon, 2011
인지적 부담 및 접근성 저하	고령자나 아동 등 정보 약자는 복잡한 제스처 조작이나 인식 오류에 취약하며, 점진적인 학습 단계를 고려하지 않은 NUI 설계로 사용성 저하 가능성 존재	
기능 확장성 한계	대부분의 제스처 기반 NUI 시스템은 오디오 제어나 간단한 명령 실행 등에 국한되어 있으며, 복합적인 기능 흐름이나 차량 내 다기능 인터페이스와의 통합에는 한계 존재	

3.2 대상 선정 및 시퀀스 도출

현재 차량 내 편의장치 조작은 센터페시아 및 카펫 중심으로 제어 인터페이스가 집중되어 있어, 특히 2열 탑승자의 조작 접근성이 낮고, 운전자의 도움 없이 사용이 제한되는 경우가 많다. 제스처 기반 NUI는 이러한 공간적 제약을 극복하고 사용자 중심의 접근을 가능하게 할 수 있는 잠재력을 지닌다. 그러나 <Table 4>에서 언급된 바와 같이, 기존의 NUI 시스템은 조작 가능 범위가 협소하며, 사용자의 신체 위치나 상호작용 맥락을 충분히 반영하지 못하는 한계를 지닌다.

이에 본 연구는 상기 언급된 한계를 보완하고 차량 내 편의장치 조작의 사용성을 개선하기 위해, 평가 대상 편의장치로 선루프와 공조 장치를 선정하였다. 이 두 기능은 조작 위치가 1열에 한정되어 있으며, 실사용 빈도가 높은 대표적 편의장치로 간주된다. 따라서 해당 기능을 중심으로 평가 시퀀스를 설계하였으며, 각 단계는 논리성과 기능적 일관성을 기반으로 구분되도록 구성하였다. 또한 기존 NUI 적용 사례들의 인터페이스 흐름을 참고하여 다음과 같이 표로 정리하였다.

Table 5. Designed evaluation sequence

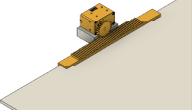
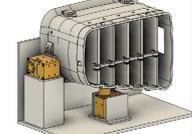
단계	설명
시스템 활성화 (System Initialization)	전원 인가 시 제스처 인식 시스템 자동 초기화
명령 입력 (Command Input)	사용자는 제스처(On/Off, Swipe, Pinch)를 입력하며, MediaPipe 기반 실시간 손 위치에 대한 좌표 분석
명령 수행 (Execution)	인식된 제스처에 따라 선루프 개폐, 공조장치 방향 조절의 기능을 Dynamixel 모터로 제어
피드백 제공 및 종료 (Feedback & Termination)	기능 수행 후 종료 단계를 알리는 LED, 비프음 등을 제공. 일정 시간을 초과한 입력 없음 시 시스템 자동 비활성화

3.3 프로토타입

기존 NUI 시스템의 단계를 기반으로 도출한 평가 시퀀스를 바탕으로 선정된 차량 편의장치(선루프, 공조장치)의 피지컬 컴퓨팅 구현을 위해 구조

설계를 진행하고, 3D Print를 활용하여 프로토타입을 제작하였다. 각각의 프로토타입의 구조 설계, DOF 및 구동 방식 등은 아래의 표와 같다. 아래 표의 구동 방식은 앞서 도출한 평가 시퀀스에서 명령 수행(Execution)의 단계이다.

Table 6. Prototype

구분	이미지	DOF	구동 방식
선루프		1	모터 회전 → 피니언 기어(Pinion Gear) 회전 → 랙(Rack)의 직선 운동 (Surge) → 선루프 개폐
공조 장치		2	제스처 입력에 맞는 모터 회전 → 블레이드 각도 조절 → 송풍

설계된 프로토타입은 앞서 언급된 2.1 소프트웨어 시스템과 연동되어 동작한다. 특히 공조 장치는 상하, 좌우 제어 모터 2개를 사용하여 2-DOF를 가능하게 하였다. 사용자의 제스처 입력은 웹캠을 통해 감지되며, 이 입력을 기반으로 제어 시스템이 모터의 동작을 결정한다. 이러한 구조 설계를 통해 하드웨어와 소프트웨어를 유기적으로 통합하여 차량 내 NUI 시스템이 적용된 편의 장치의 사용성을 향상하도록 하였으며, 이를 바탕으로 추후 실험을 위한 설계를 완료하였다.

4. 결론

본 연구는 기존 차량 인터페이스가 지닌 조작의 복잡성, 공간적 제약, 정보 약자에 대한 접근성 한계를 개선하고자, 직관적인 제스처 기반 자연 사용자 인터페이스(Natural User Interface, NUI) 시스템의 설계 및 구현을 목적으로 수행되었다. 이를 위해 실제 차량 환경을 모사한 상황에서 작동 가능한 프로토타입을 구축하고, MediaPipe 기반 손 좌표 인식, 제스처 입력 알고리즘, Dynamixel 모터 제어 시스템 등을 통합하여 선루프 및 공조장치 제어가 가능한 피지컬 컴퓨팅 구조를 구현하였다.

또한 기존 NUI 시스템의 주요 한계로 지적되어 온 제스처 인식 오류, 피드백 부족, 사용자 위치 고려 미흡 등의 문제를 보완하기 위해, 보다 실사용 맥락에 기반한 평가 시퀀스를 도출하였다. 특히 사용자 위치나 상호작용 환경에 따른 인식 성능과 사용성 차이를 고려하여, 카메라 위치(위, 정면, 옆) 및 제스처 방식(Pinch and Drag, Pinch to Drag)을 주요 변수로 설정한 평가 기준을 수립하였으며, 이를 다음과 같이 표로 정리하였다.

Table 7. Directions for Follow-up research

평가 항목	실험 변수 설정
카메라 위치	위
	정면
	옆
제스처 방식	Pinch and Drag
	Pinch to Drag

향후 연구에서는 제안된 시스템의 효과성을 정량적·정성적 지표(예: 인식 정확도, 응답 시간, 사용성 만족도 등)를 통해 실험적으로 검증할 계획이다. 이를 위해 Low-fidelity Prototype 및 Working Mock-up을 기반으로 다양한 사용자 시나리오에 따른 사용성 평가 실험을 설계하고, 특히 고령자나 아동 등 정보 약자를 고려한 스캐폴딩(Scaffolding) 기반 인터랙션 설계의 타당성과 효과성 검토를 병행할 예정이다.

본 연구는 차량 내 제스처 기반 NUI 시스템의 초기 설계 및 구현 가능성을 실증적으로 제시하였으며, 향후 사용자 중심 인터페이스 설계, 차량 내 다기능 시스템 통합, 몰입형 사용자 경험 강화 등의 연구에 기초자료로 활용될 수 있을 것이다.

References

1) B. Shneiderman, "Direct manipulation: A step beyond programming languages," *Computer*, vol. 16, no. 8, pp. 57 - 69, 1983.

2) D. A. Norman, "Natural user interfaces are not natural," *Interactions*, vol. 17, no. 3, pp. 6 - 10, 2010.

3) G. Jeon, "A study on user cognition design and UX development through NUI (Natural User Interface): Focusing on the development of weather information content and design using environmental composite sensors," *Journal of the Korean Society of Industrial Design*, vol. 10, no. 2, pp. 177 - 186, 2016.

4) D. Wigdor and D. Wixon, *Brave NUI world: Designing natural user interfaces for touch and gesture*, Morgan Kaufmann Publishers, 2011.

5) D. Seo, J. Lee, J. Kim, and S. Lee, "Bare-hand-based augmented reality interaction for providing a Natural User Interface," in *Proceedings of the 2012 Spring Joint Conference of the Korean Institute of Industrial Engineers*, vol. 2, pp. 2658 - 2659, 2012.

6) M. Kim, D. Yeo, W. Kim, J. Lee, K. Cho, and S. Kim, "Large-area windshield head-up display and in-vehicle multimodal NUI interface: Touch gestures and voice commands," in *Proceedings of HCI Korea 2019*, pp. 501 - 506, 2019.

7) C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, et al., "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172*, 2019.

8) A.-K. Frison, P. Wintersberger, and A. Riener, "First evaluation of a multimodal driver-vehicle interface in a driving simulator," in *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 174 - 181, ACM, 2017.

이 연구는 2025년도 산업통상자원부 및 한국산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임.
(202401620001)