



보고서

시각장애인을 위한 지능형 장애물 탐지 및 간판 인식 시스템

과목명 : 캡스톤디자인

팀명 : 하얀조끼

팀장 : 20221052 서민경

팀원 : 20221044 김내경, 20221047 김윤희

지도교수 : 현장훈 교수님

1. 연구 배경 및 목적

시각장애인의 이동 보조 도구로 가장 널리 사용되는 '흰 지팡이'는 물리적인 장애물을 감지하는 데 유용하지만, 복잡한 도심 환경에서는 한계를 갖는다. 특히 변화가나 상점가와 같은 밀집된 공간에서는 고정된 장애물과 움직이는 장애물이 혼재되어 있어 경로를 정확히 파악하기 어렵다. 또한, 흰 지팡이는 주로 발 앞의 장애물 감지에 초점이 맞춰져 있어, 상체 높이에 있는 간판이나 돌출된 구조물과 같은 장애물은 효과적으로 탐지하기 어렵다. 이와 더불어, 시각장애인은 특정 목적지를 찾거나 새로운 경로를 탐색할 때 어려움을 겪는다. 예를 들어, 처음 방문하는 지역에서는 주변 환경을 직관적으로 인식하기 어렵고, 보행 중 가게나 건물의 위치를 파악하는 것이 쉽지 않다. 특히 간판과 같은 시각적 정보를 활용할 수 없는 환경에서는 특정 장소를 식별하는 것이 더욱 어렵다. 이는 시각장애인의 이동 효율성과 안전성을 저하시켜 사회적 활동의 제한을 초래할 수 있다. 최근 객체 탐지(Object Detection) 및 텍스트 인식(OCR, Optical Character Recognition) 기술이 발전함에 따라, 장애물을 인식하고 주변의 상점 정보를 제공하는 시스템을 개발할 수 있는 가능성이 열리고 있다. 이러한 기술을 활용하면 기존 흰 지팡이의 한계를 보완하고, 시각장애인의 이동 편의성을 대폭 향상시킬 수 있을 것으로 기대된다.

2. 시스템 설계 및 구현

2.1 상대 속도 추정 기반 동적 정적 장애물 분류 알고리즘

2.1.1 사용자 이동 거리 추정 (IMU 기반)

사용자의 움직임을 정확히 추정하기 위해 IMU 센서(가속도계)를 활용한다.

IMU로부터 얻은 가속도 데이터를 시간 간격(Δt) 기준으로 1차 적분하여 속도(v)를, 다시 2차 적분하여 위치(p)를 계산한다.

이를 통해 두 프레임 간의 물리적 위치 변화 Δs 를 추정할 수 있다:

- 속도:

$$v(t_n) = v(t_{n-1}) + \int_{t_{n-1}}^{t_n} a(\tau) d\tau$$

- 위치:

$$p(t_n) = p(t_{n-1}) + \int_{t_{n-1}}^{t_n} v(\tau) d\tau$$

- 이동 거리:

$$\Delta s_n = |p(t_n) - p(t_{n-1})|_2$$

2.1.2 객체 검출 및 트래킹 (YOLO + ByteTrack)

정확한 객체 추적을 위해 세그멘테이션 기반 검출과 바운딩 박스 트래킹을 결합한다.

- 객체 검출: YOLOv11-seg 모델을 사용하여 객체의 바운딩 박스 및 픽셀 단위 마스크를 추출한다.
- 중복 제거: Non-Maximum Suppression(NMS) 기법을 적용하여 중복 탐지를 제거 한다.
- 객체 추적: ByteTrack 알고리즘을 사용하여 프레임 간 객체를 연결하고 고유 ID를 유지한다.
(칼만 필터 기반의 예측 및 최장 경로 매칭을 통해 일시적 가림에도 안정적인 추적을 보장한다.)

2.1.3 객체별 거리 산출 (깊이 센서 기반)

깊이 정보를 통해 객체와 사용자의 거리를 추정한다.

- RGB-D 센서에서 얻은 깊이 영상을 객체의 마스크에 맞춰 필터링한다.
- 노이즈, 그림자 등으로 인한 이상치를 제거한 후, 중앙값(Median)을 객체 거리로 정의한다:

$$d_i = \text{median}\{ \text{depth}(x, y) \mid (x, y) \in \text{mask}_i \}$$

2.1.4 정적 및 동적 장애물 분류

객체의 이동 여부를 사용자의 이동과 비교하여 분류한다.

- 이전 프레임과 현재 프레임 사이의 거리 변화량을 계산한다.

$$\Delta d_i = d_{i,\text{prev}} - d_{i,\text{cur}}$$

- 분류 기준:

- 동적 장애물:

$$\Delta d_i > \Delta s_n \quad \text{and} \quad d_{i,\text{cur}} < 7 \text{ m}$$

- 정적 장애물:

$$\Delta d_i \leq \Delta s_n \quad \text{and} \quad d_{i,\text{cur}} < 5 \text{ m}$$

이 방식은 사용자의 실제 이동 거리를 기준으로 상대적 거리 변화를 분석함으로써, 고정된 장애물과 이동 중인 장애물을 효과적으로 구분할 수 있도록 한다.

2.2 영상 기반 충돌 회피 보행 보조 시스템

2.2.1 시스템 설계 및 구현 (System Design & Implementation)

본 시스템은 시각장애인이 실내 환경에서 보행 중 충돌 위험을 사전에 인지하고 회피할 수 있도록 설계되었다. 스테레오 카메라로부터 수집한 RGB-D 영상을 입력으로 하여, 객

체 인식, 추적, 거리 계산, 동적/정적 분류, 충돌 판단까지 총 5단계에 걸쳐 처리된다.

2.2.2 객체 인식 (Object Detection)

보행자의 주변 환경을 정확히 인식하기 위해 YOLOv11 Segmentation 모델을 사용하였다.

- Instance Segmentation 방식으로 동작하여, 동일한 클래스(예: 사람)가 겹쳐있더라도 각각의 객체를 구분 가능하다.
- 입력은 RGB 영상과 깊이(Depth) 영상이며, 이를 통해 객체의 바운딩 박스와 픽셀 단위 마스크를 함께 추출한다.
- Semantic Segmentation과는 달리, 픽셀 수준의 객체 구분과 개별 ID 할당이 가능하여 이후 추적 및 거리 분석에 용이하다.

2.2.3 동적 객체 필터링 (Dynamic Object Filtering)

실내 환경에서는 정적 장애물(책상, 의자 등)이 다수 존재하며, 이들은 대부분 충돌 위험성이 낮다. 반면, 사람은 독립적으로 움직이며 돌발 상황을 유발할 수 있으므로, 본 시스템에서는 'People' 클래스에 속한 객체만을 위험 요소로 간주한다.

- 정적 객체는 분석 대상에서 제외됨으로써 불필요한 경고 발생을 줄인다.
- 동적 객체 필터링은 이후 추적과 위험 판단 단계의 효율성과 정확도를 향상시킨다.

2.2.4 객체 추적 (Object Tracking)

프레임 간 동일 객체를 식별하기 위해 SORT(Simple Online and Realtime Tracking) 알고리즘을 사용하였다.

- YOLOv11의 출력 결과를 직접 입력으로 사용하여 객체에 고유 ID를 할당한다.
- SORT는 칼만 필터 기반의 예측 및 Hungarian 알고리즘 기반의 프레임 간 매칭을 통해 빠르고 효율적인 실시간 추적이 가능하다.
- 추적된 객체는 ID 기반으로 시간에 따른 위치 변화와 속성 변화(면적, 거리 등)를 추적할 수 있다.

2.2.5 객체 검출-추적 결과 통합 (Matching Detection & Tracking Results)

객체 인식 결과(Class 정보)와 추적 결과(ID 정보)를 통합하기 위해 IOU(Intersection over Union) 기준의 매칭 기법을 적용하였다.

- 각 YOLO 바운딩 박스와 SORT 바운딩 박스 간의 IOU를 계산하여 가장 유사한 쌍을 찾는다.

- 매칭 기준:

$$IOU = \frac{\text{Area of Overlap}}{\text{Area of Union}} \geq 0.3$$

매칭 성공 시, 해당 객체는 Class, ID, 거리, 면적 정보가 통합된 단일 엔티티로 관리된다.

2.2.6 충돌 위험 판단 (Collision Risk Detection)

충돌 위험 판단은 다음의 두 조건 중 하나 이상을 만족하는 경우 수행된다. 이 조건은 실내 보행 중 예상 가능한 위험 상황을 반영하여 설계되었다.

조건 1: 점진적 접근 판단

- 사람 객체의 바운딩 박스가 시간이 지남에 따라 커지며, 사용자를 향해 접근 중인 상황
- 다음의 두 조건이 동시에 만족할 경우 충돌 위험으로 판단:
 - 바운딩 박스 면적 증가율 $\geq 18\%$
 - 거리 $\leq 1.8m$

조건 2: 근접 정지 객체 판단

- 사람이 정지 상태이지만 사용자의 바로 앞에 위치한 상황
- 다음의 조건을 만족할 경우 충돌 위험으로 판단:
 - 바운딩 박스 면적 $\geq 250,000$ 픽셀

2.2.7 위험 알림 및 시각화

위험이 감지되면 다음과 같은 방식으로 사용자에게 경고를 제공한다.

- 해당 객체의 바운딩 박스 색상: 초록색 \rightarrow 빨간색으로 변경
- 경고 메시지 출력 예시:
- 충돌 위험 !! 1.5m 이내 사람이 있습니다.

2.3 시야 정보를 활용한 시각 안내 시스템

2.3.1 정지 상태 감지 (User Stillness Detection)

IMU 센서 데이터를 분석하여 사용자가 일정 시간 이상 움직이지 않을 경우 "정지 상태"로 간주한다. 이 시점부터 주변의 안내 텍스트를 인식하여 정보 제공을 시작한다.

2.3.2 시야 영역 필터링 및 전처리

정지 상태가 되면 카메라 시야 내 객체 중 일정 거리 이내에 있는 부분만 필터링하며, OCR 성능 향상을 위해 밝기 정규화, 왜곡 보정 등의 전처리를 수행한다.

2.3.3 OCR 후처리 및 클러스터링

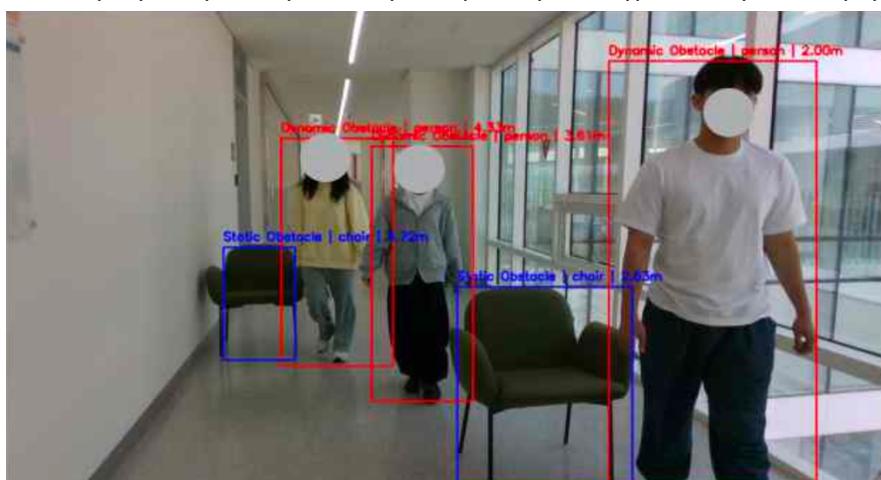
OCR 결과 중 중복되거나 깨진 텍스트는 제거하고, 위치 기반 클러스터링을 통해 중요 단어(예: 시설명, 안내문)를 추출한다. 필요 시 Depth 데이터를 이용해 거리 정보 기반 정렬을 수행한다.

2.3.4 TTS(Text-to-Speech) 변환 및 출력

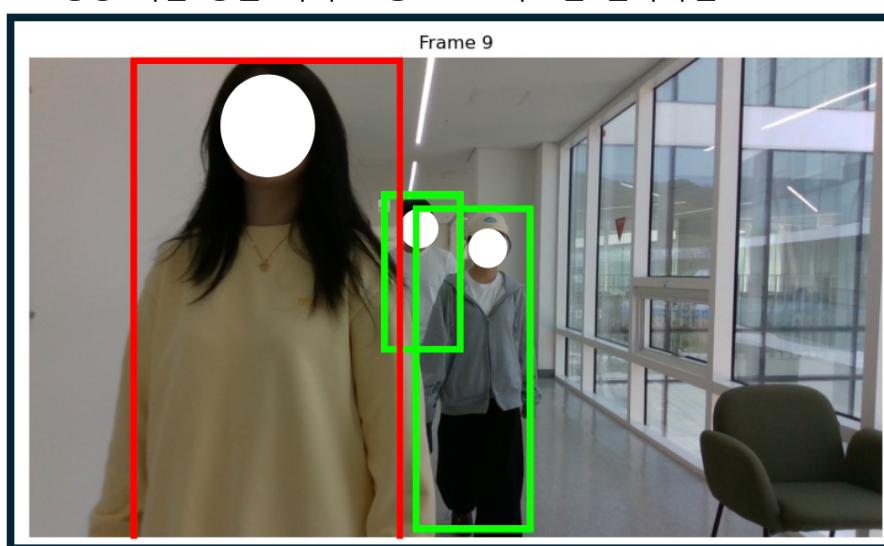
클러스터링된 텍스트 중 가장 중요한 내용을 TTS 엔진을 통해 음성으로 출력한다.

3. 결과

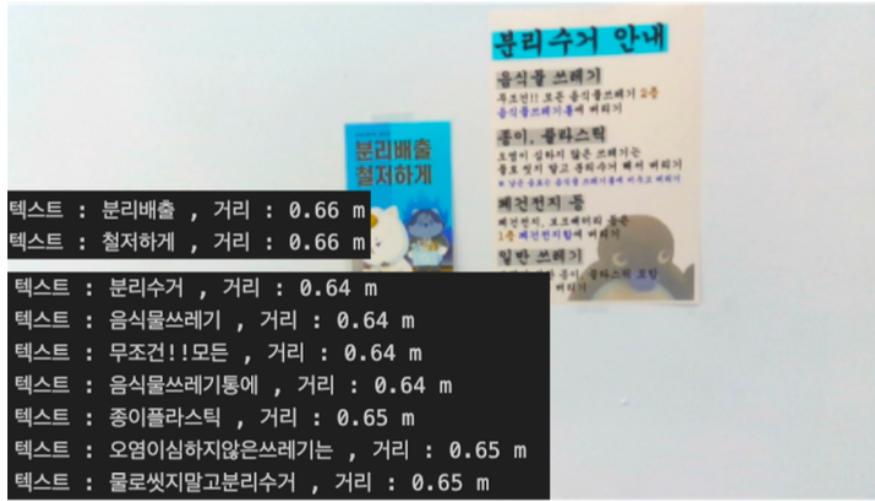
2.1 상대 속도 추정 기반 동적 장애물 분류 알고리즘 결과사진



2.2 영상 기반 충돌 회피 보행 보조 시스템 결과사진



2.3 시야 정보를 활용한 시각 안내 시스템 결과사진



본 시스템은 시각장애인의 자립적인 이동을 돋는 데 중요한 역할을 한다. 장애물 탐지 및 경로 안내 시스템은 사용자가 장애물과 충돌하지 않고 안전하게 이동할 수 있도록 도와주며, 간판 읽기 및 정보 제공 시스템은 주변 환경에 대한 중요한 정보를 실시간으로 제공함으로써 시각장애인의 생활을 크게 향상시킨다. 이 시스템은 특히 도시 환경에서의 이동을 지원하는 데 필수적인 도구로 자리 잡을 수 있다.

또한 비상 상황에서 제공되는 텍스트 인식 정보는 시각장애인의 중요한 안내판을 읽지 못해 겪을 수 있는 위험을 방지하는 데 도움이 된다. 이 시스템은 공공안전 분야에서 중요한 역할을 할 수 있으며, 비상 상황에서 정확하고 신속한 정보 전달을 가능하게 해준다.