

XR 글래스 기반 캐릭터 콘텐츠를 활용한 시각장애인 시기능 훈련 시스템

김갑열*, 정인철**, 조진수***

Visual Function Training System for the Visually Impaired using Character Content based on XR Glasses

Kapyol Kim*, Incheol Jeong**, and Jinsoo cho***

요 약

본 논문은 저시력 시각장애인을 위한 XR 글래스 기반 캐릭터 훈련 시스템을 제안한다. 제안 시스템은 기존의 오프라인 훈련 대비 장소 제약 없이 개인 시력 유형에 맞춘 맞춤형 콘텐츠 제공이 가능하다는 장점이 있다. 관리자 단말에서 등록된 캐릭터 이미지 기반으로, 객체 탐지(YOLOv8), 키워드 추출(Transformer), 설명문 생성(KoGPT-3), TTS 변환 등 AI 기반 캐릭터 상황 인지 절차를 수행하여 훈련용 콘텐츠를 자동 생성하며, 생성된 콘텐츠는 XR 글래스를 통해 실시간 제공된다. 성능 평가 결과, 평균 응답 속도 0.11초, 동시 접속 처리량 307.2건/초, 객체 탐지 정확도 96.96%, 키워드 추출 정확도 98.98%, 문장 생성 정확도 89.64%로, 실제 활용 환경에서도 높은 신뢰성과 처리 성능을 입증하였다.

Abstract

This study proposes a character-based training system utilizing XR glasses to enhance the visual function of individuals with visual impairments. The system provides personalized training content tailored to users' visual characteristics and enables flexible use beyond spatial limitations. It comprises user devices, an administrator interface, and a central platform server. Character images registered by administrators are automatically transformed into training materials through an AI-driven character situation recognition pipeline, which includes object detection (YOLOv8), keyword extraction (Transformer), sentence generation (KoGPT-3), and TTS conversion. The generated content is delivered to users in real-time via XR glasses, facilitating immersive visual training. Performance evaluations revealed an average response time of 0.11 seconds and a concurrent throughput of 307.2 requests per second. The system also achieved accuracies of 96.96% in object detection, 98.98% in keyword extraction, and 89.64% in sentence generation, demonstrating its reliability and effectiveness in real-world environments.

Keywords

visual impairment, visual function training, XR glasses, YOLOv8, transformer, KoGPT-3

* 가천대학교 컴퓨터공학과

- ORCID: <https://orcid.org/0009-0007-8754-2951>

** 가천대학교 컴퓨터공학과 대우교수

- ORCID: <https://orcid.org/0009-0001-7180-1759>

*** 가천대학교 컴퓨터학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0002-8119-8068>

· Received: Apr. 01, 2025, Revised: Apr. 21, 2025, Accepted: Apr. 24, 2025

· Corresponding Author: Jinsoo cho

Dept. of Computer Engineering, Gachon University, 1342, Seongnam-daero, Sujeong-gu, Seongnam-si, Gyeonggi-do, Korea

Tel.: +82-31-750-5820, Email: jscho@gachon.ac.kr

I. 서 론

시각장애인은 시각 감각의 제한으로 인해 교육과 일상생활에서 다양한 어려움을 경험한다[1]. 정안인이 시각을 통해 복합적인 학습이 가능한 반면, 시각장애인은 주요 시각정보가 제한된 환경에서 학습이 이루어지므로 시각적 이해가 필요한 내용 습득에 상당한 어려움을 겪는다. 이러한 제약은 고등 교육의 장벽으로 작용하여 결과적으로 취업 기회를 제한하는 요인이 되고 있다.

시각장애는 전맹, 저시력, 시야 협착, 중심 시야 손실, 색각 이상 등으로 구분되며, 이 중 저시력은 잔존 시력이 있으나 일상생활에 어려움을 겪는 상태를 의미한다. 특히 저시력 시각장애인의 경우, 잔존 시력을 효과적으로 활용할 수 있는 시기능 훈련이 중요하다. 시기능 훈련은 저시력자가 제한된 시력을 최대한 활용할 수 있도록 돕는 과정으로, 시각적 인지 능력을 향상시키고 자립성을 높이는데 기여한다. 그러나 현재 대부분의 시기능 훈련은 오프라인 환경에서 제한적으로 이루어지고 있으며, 다양한 시각장애 유형에 맞춤형된 훈련 콘텐츠가 부족한 실정이다.

XR(eXtended Reality, 확장현실) 기술은 가상현실(VR, Virtual Reality), 증강현실(AR, Augmented Reality), 혼합현실(MR, Mixed Reality)을 아우르며, 현실과 가상을 융합해 새로운 경험을 제공한다. 특히 XR 글래스는 안경 형태의 웨어러블 기기로, 사용자 시야에 정보를 직접 표시할 수 있어 시각장애인의 시기능 훈련 및 보조 도구로서 잠재력을 가진다[2]. 본 논문에서 제안하는 시스템은 저시력 사용자에게 초점을 맞추어, 개인의 시력 특성에 따라 콘텐츠를 최적화하여 제공하는 것을 목표로 한다. 저시력자는 이를 통해 영상 확대, 대비 조정, 주변 정보 인지 등 기능을 활용할 수 있다.

하지만 XR 글래스를 활용한 시각장애인 전용 콘텐츠 및 플랫폼은 아직 초기 단계로, 실질적인 적용 사례는 매우 드물다.

국내에서는 시각장애인을 위한 보행 안전 시스템 개발도 일부 이루어지고 있으나[3], 대부분의 지원 서비스는 맞춤형 학습이나 시기능 훈련과 같은 특화된 교육은 부족하다. 국외의 Seeing AI, Aira,

OrCam MyEye 등도 높은 비용이나 사용 제약으로 접근성이 떨어진다.

본 논문에서는 XR 글래스를 활용한 시각장애인 시기능 훈련 시스템의 설계 및 구현에 대해 다룬다. 캐릭터 콘텐츠를 기반으로 잔존 시력을 효과적으로 훈련하고, 다양한 시각장애 유형에 맞춤형된 인터페이스를 제공하는 시스템을 개발하였다. 특히, AI(Artificial Intelligence) 기반 캐릭터 상황 인지, KoGPT-3 대체텍스트 생성[4], TTS(Text-to-Speech) 변환 기능[5]을 통해 정보 접근성을 향상시키는 데 중점을 둔다.

본 논문의 구성은 다음과 같다. 2장에서는 시각장애인의 인지특성과 시각장애 지원 기술, 그리고 XR 글래스 기술 동향을 분석하고, 3장에서는 제안하는 시스템의 전체 구조, 연구 대상 특성, 그리고 캐릭터 기반 시각장애인 시기능 훈련 서비스와 맞춤형 시각 보정 기능의 설계 및 구현에 대해 설명한다. 4장에서는 개발된 시스템의 성능 평가 및 AI 처리 성능 결과를 제시하고, 마지막으로 5장에서는 결론과 향후 연구 방향에 대해 논의한다.

II. 관련 연구

2.1 시각장애인 지원 기술 동향

시각장애인 지원 기술은 정보 접근성, 이동성, 일상생활 지원 등 다양한 영역에서 발전해왔다. 특히 교육 분야에서의 지원 기술은 시각장애인의 학습 경험을 개선하는 데 중요한 역할을 하고 있다[6].

현재 국내에서 운영 중인 시각장애인 교육 지원 서비스로는 EBS 장애인 서비스, Samsung anycom, 실로암 이러닝 센터 등이 있다. EBS 장애인 서비스는 화면해설 기반의 학습 영상을 제공하여 교육 콘텐츠의 접근성을 높이고 있으나, 개별 맞춤형 학습이나 상호작용 기능이 부족하다. Samsung anycom은 기초 센서리더를 활용한 정보화 기기 사용법 교육을 제공하지만, 주로 정보기기 사용법에 집중되어 있어 다양한 학습 콘텐츠 제공에 한계가 있다. 실로암 이러닝 센터는 여가 및 외국어 강좌, 모바일 앱 사용법 등 다양한 콘텐츠를 제공하지만, 시기능 훈련과 같은 특화된 교육 지원은 미흡한 실정이다.

국제적으로는 Microsoft Seeing AI, Aira, OrCam MyEye 등의 기술이 주목받고 있으며, 시각장애인을 위한 실내 내비게이션, 객체 인식, 경로 안내 시스템도 함께 개발되고 있다[7][8]. Microsoft Seeing AI는 AI를 활용하여 주변 환경이나 텍스트를 인식하고 음성으로 설명해주는 앱으로, 시각장애인의 정보 접근성을 높이는 데 기여하고 있다. Aira는 스마트 안경을 통해 전문 상담사가 실시간으로 시각장애인을 지원하는 서비스로, 독립적인 생활을 돕지만, 비용이 높고 인터넷 연결이 필수적이라는 제약이 있다. OrCam MyEye는 안경에 부착하는 웨어러블 장치로, 텍스트 읽기, 얼굴 인식, 제품 식별 등의 기능을 제공하지만, 역시 높은 비용과 제한된 기능으로 인해 다양한 사용자 요구를 충족시키기 어렵다. 또한, 가상현실을 활용하여 시각장애 아동의 이동 능력 훈련에 적용한 사례도 일부 연구되고 있다[9].

이러한 기존 기술들은 주로 일상생활 지원이나 기본적인 정보 접근성 향상에 중점을 두고 있으며, 시각장애인의 시기능 훈련과 같은 특화된 교육 지원은 상대적으로 부족한 상황이다. 특히, 시각장애 유형별 맞춤형 서비스와 상호작용적 학습 경험을 제공하는 기술은 더욱 제한적이다.

2.2 XR 글래스 기술 및 응용

XR 기술은 가상현실, 증강현실, 혼합현실을 포괄하며, 현실과 가상의 융합을 통해 몰입형 경험을 제공한다. 특히 XR 글래스는 안경 형태의 웨어러블 기기로, 사용자 시야에 정보를 직접 제공할 수 있어 교육, 의료, 제조 등 다양한 분야에서 활용되고 있다. 대표적인 XR 글래스로는 Microsoft HoloLens, Magic Leap, XREAL Light가 있으며, 각각 MR 기반의 산업용 기기부터 소비자용 AR 글래스까지 다양한 스펙트럼을 형성한다. 이들 장치는 영상 확대, 색상 보정, 객체 인식, 음성 안내 등 다양한 기능을 통해 시각정보 제공을 보조할 수 있으며, 시각장애인을 위한 정보 접근성 향상에도 활용 가능성이 제기되고 있다[10]-[12]. 표 1은 대표적인 XR 글래스 기기의 종류와 그 특징을 정리한 것이다.

XR 글래스를 활용한 시각장애인 지원은 영상 확대, 색상 보정, 객체 인식, 음성 안내 등의 기능을

통해 정보 접근성을 향상시키려는 시도로 이어지고 있다. 특히 저시력자의 시기능 훈련 도구로서 XR 글래스는 실시간 시각 자극 제공과 훈련 난이도 조절을 통해 맞춤형 훈련이 가능할 것으로 기대된다. 그러나 캐릭터 콘텐츠를 기반으로 한 시기능 훈련 시스템은 전례가 드물며, 본 연구는 이러한 기술적 공백을 보완하고 시각장애인 맞춤형 교육 플랫폼의 가능성을 제시한다.

표 1. XR 글래스 종류 및 특징

Table 1. Types and characteristics of XR glasses

XR glass	Key features	Primary application areas	Training focus
Microsoft HoloLens	Holographic display, spatial mapping MR headset	Industry, healthcare	Peripheral vision training, development of alternative visual zones
Magic Leap	Optical waveguide-based MR glasses	Education, entertainment	Field-of-view expansion training, efficient gaze movement strategy
XREAL light	Lightweight AR glasses, smartphone connectivity	Consumer market, education	Contrast enhancement training, color discrimination strategy

III. XR 글래스 기반 시각장애인 맞춤형 시기능 훈련 시스템 설계 및 구현

3.1 시스템 구조 및 플랫폼 구성

본 연구에서 개발한 XR 글래스 기반 시각장애인 맞춤형 시기능 훈련 시스템은 시각장애인의 시기능 특성을 고려하여 맞춤형 콘텐츠 제공과 시각 보정 기능을 통합한 플랫폼이다. 제안 시스템은 실시간 시기능 훈련 콘텐츠 제공과 시각 보정 기능 통합을 지원하도록 설계되었다.

다음 그림 1은 제안하는 XR 글래스 기반 시각장애인 시기능 훈련 시스템의 구조를 도식화한 것이다.

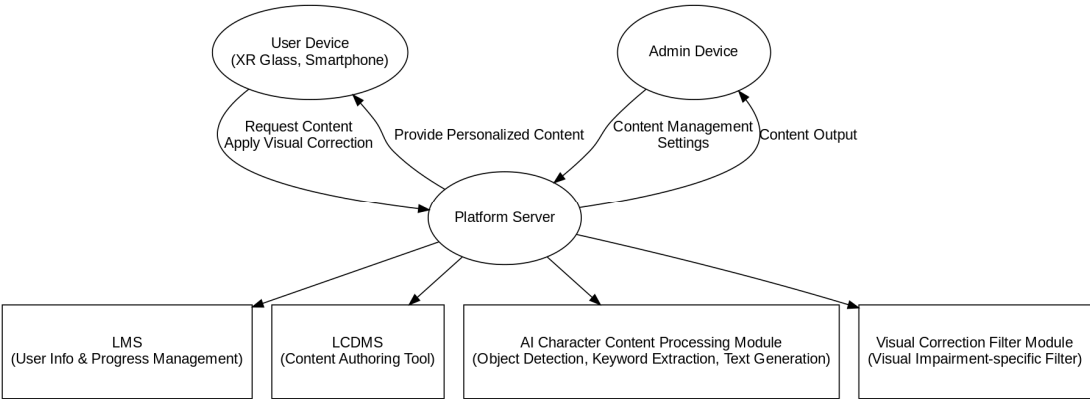


그림 1. XR 글래스 기반 시각장애인 시기능 훈련 시스템 구조
Fig. 1. XR glass-based visual function training system architecture

그림 1에 나타난 바와 같이, 제안 시스템은 사용자 단말과 관리자 단말이 플랫폼 서버와 상호 연결되어 있으며, 플랫폼 서버 내부에는 학습 관리 시스템(LMS, Learning Management System), 콘텐츠 저작 관리 시스템(LCDMS, Learning Content Development Management System), AI 기반 캐릭터 콘텐츠 처리 모듈, 시각 보정 필터 모듈이 포함된다. 사용자는 스마트폰과 XR 글래스를 통해 플랫폼에 접속하여 시각장애 유형 등록, 콘텐츠 이용, 훈련 이력 관리를 수행하고, 관리자는 관리자 단말을 통해 콘텐츠 제작 및 보정 설정을 수행한다. 각 주요 구성 요소의 기능은 다음 표 2와 같다.

표 2. 시스템의 구성 요소 및 기능
Table 2. Components and functions of the system

Component	Main functions
User device	Content usage, visual correction application, training history management
Administrator device	Content creation and management, visual correction configuration
Platform server	User data management, content delivery, visual correction processing
LMS	User registration, visual impairment type management, training history tracking
LCDMS	Authoring tools for character content, scenario and difficulty level configuration
AI character content processing module	Object detection, keyword extraction, description text generation, difficulty assessment
Visual correction filter module	Application of customized visual filters based on visual impairment types

본 시스템은 각 구성 요소가 유기적으로 연계되어, 사용자 맞춤형 시기능 훈련 환경을 제공한다. 플랫폼 서버 내 LMS와 LCDMS는 사용자 정보 및 콘텐츠를 효율적으로 관리하며, AI 캐릭터 콘텐츠 처리 모듈과 시각 보정 필터 모듈은 사용자 특성에 최적화된 콘텐츠 제공과 시각 보정 기능을 지원한다. 이를 통해 사용자는 자신의 시각장애 유형에 맞춘 시기능 훈련 콘텐츠를 실시간으로 제공받고, 지속적인 학습이 가능하도록 설계되어 있다.

3.2 연구 대상 특성 분석

시각장애는 세계보건기구(WHO)의 기준에 따라 전맹(Total blindness)과 저시력(Low vision)으로 구분된다. 본 연구에서는 잔존 시력이 있는 저시력 시각장애인을 위한 시기능 훈련에 초점을 맞추고 있으므로, 주요 저시력 유형에 대한 특성을 분석하였다.

황반원공은 황반 중심의 망막조직 결손으로 중심 시야에 암점(Scotoma)이 생기는 현상으로, 글자 읽기나 얼굴 인식 등 세부 시각 정보 처리에 어려움을 겪는다. 황반원공 환자들은 편심 시점(Eccentric viewing) 기법을 활용하여 주변 시야로 정보를 인지하는 훈련이 효과적이다[13][14].

터널 시야는 망막색소변성이나 녹내장 등의 질환으로 발생하며, 마치 터널을 통해 보는 것처럼 시야가 좁아지는 현상이다. 중심부는 인지할 수 있으나 주변 환경 파악에 어려움을 겪어 공간 인지와 이동

에 제약이 생기며, 효율적인 시선 이동 및 시야 확장 훈련이 도움이 된다.

색각이상은 특정 색상을 정상적으로 구분하지 못하는 것으로, 색상 기반 학습 자료나 인터페이스 사용에 제약이 있다. 색상 대비를 강화하거나 보조색을 활용하는 훈련이 효과적이다[15].

본 연구에서는 이러한 저시력 유형별 특성을 반영하여, 사용자의 시기능 특성에 맞춘 맞춤형 XR 클래스 기반 시기능 훈련 시스템을 설계하였다.

표 3은 저시력 유형에 따른 주요 특성을 정리한 것이다.

표 3. 저시력 유형별 특성 분석

Table 3. Analysis of characteristics by type of low vision

Type of visual impairment	Visual characteristics	Cognitive features	Training approach
Macular hole	Central visual field loss due to macular retinal tissue defect	Central scotoma, difficulty recognizing text and faces	Peripheral vision utilization, development of alternate visual areas
Tunnel vision	Loss of peripheral vision, only narrow central vision remains	Clear central perception, impaired spatial awareness and navigation	Field expansion training, efficient gaze shift strategy
Color vision deficiency	Inability to distinguish specific colors (e.g., red-green, blue-yellow)	Difficulty processing color-based information, limited content usability	Contrast enhancement training, color differentiation strategies

3.3 캐릭터 기반 시각장애인 시기능 훈련 서비스 설계

본 연구에서 제안하는 캐릭터 기반 시각장애인 시기능 훈련 서비스는 시각장애인의 시기능 특성에 최적화된 맞춤형 훈련 콘텐츠를 제공하기 위해 설계되었다. 제안 서비스는 시각장애 유형별로 개별화된 시각 자극 제공과 실시간 콘텐츠 생성, 그리고 난이도 조정 기능을 통합하고 있다. 특히, 관리자 단말을 통해 제작된 캐릭터 콘텐츠를 플랫폼

서버의 AI 캐릭터 콘텐츠 처리 모듈에서 자동으로 분석하고 가공하여 사용자에게 실시간으로 제공하는 점이 특징이다. 캐릭터 기반 시기능 훈련 서비스는 크게 캐릭터 상황 인지 기반 콘텐츠 제공 과정과 콘텐츠 난이도 판별 및 시퀀스 제공 과정으로 구성된다.

플랫폼 서버 내 AI 캐릭터 콘텐츠 처리 모듈은 캐릭터 이미지 데이터를 입력으로 받아, YOLOv8 기반 객체 탐지 모델을 통해 캐릭터 이미지 내 객체들을 탐지하고 객체의 위치 및 속성을 파악한다 [16][17]. 이어서 Transformer 기반 키워드 추출 모델을 통해 객체의 주요 특징을 바탕으로 핵심 키워드 집합을 도출한다[18]. 이후, KoGPT-3 기반 텍스트 생성 모델을 통해 키워드와 객체 정보를 바탕으로 설명 텍스트를 생성하고, 해당 설명은 TTS 변환을 통해 음성으로 사용자에게 제공된다. 이때 텍스트 생성 과정에서는 처리 속도와 간결한 문장 생성을 위해 경량화된 KoGPT-3 기반 모델을 사용하였다. 또한, 캐릭터 콘텐츠의 난이도는 객체 수, 색상 다양성, 선분 길이의 정량적 요소를 기반으로 계산된다. 난이도 점수는 다음과 식 (1)과 산출된다.

$$D = \alpha \times N_{obj} + \beta \times C_{div} + \gamma \times L_{seg} \quad (1)$$

여기서 N_{obj} 는 객체 수, C_{div} 는 색상 다양성, L_{seg} 는 선분 총 길이이며, 각 가중치는 실험적으로 조정된다. 판별된 난이도에 따라 콘텐츠는 초급, 중급, 고급으로 분류되고, 콘텐츠 제공 순서는 코사인 유사도 기반 시퀀스 생성 알고리즘을 통해 유사도가 낮은 콘텐츠가 연속 제공되도록 배열된다.

각 난이도 구간은 설명 텍스트의 문장 구조와 어휘 수준, 정보 밀도에 따라 차별화된다. 예를 들어, 초급 수준에서는 “파란 모자를 쓴 곰이 있다”와 같은 단문 중심의 설명이 생성되며, 고급 수준에서는 “파란 모자를 눌러쓴 곰이 오른손에 가방을 들고 길을 건너고 있다”와 같이 복합문과 동작 중심의 상세 설명이 생성된다. 이러한 텍스트 난이도 차이는 사용자 인지 부하에 맞게 조절된 학습 콘텐츠를 제공함으로써 훈련 집중도 및 이해도를 높이는 데 기여한다.

본 시스템에서는 일반적인 코사인 유사도 수식
이 아닌, 캐릭터 콘텐츠의 주요 시각적 특성을 고려
하여 콘텐츠 간 유사도를 계산한다. 콘텐츠 i 와 j 의
유사도 스코어 $S_{i,j}$ 는 다음 식 (2)와 같이 정의된다.

$$S_{i,j} = \frac{\sum_{k=1}^n w_k \times f_k(V_i, V_j)}{\|V_i\| \times \|V_j\|} \quad (2)$$

여기서 V_i, V_j 는 각각 캐릭터 콘텐츠의 특징 벡터이
며, 각 벡터는 객체 수, 색상 다양성, 선분 길이 등
주요 시각적 특성을 포함한다. $f_k(V_i, V_j)$ 는 k -번째
특성에 대한 비교 함수이고, w_k 는 시각장애 유형별
중요도에 따라 가중치가 부여된다.

유사도 스코어가 낮을수록 서로 다른 특성을 가
진 콘텐츠로 간주되어 연속 제공된다. 이를 통해 사
용자는 반복 학습 시 다양한 시각 자극을 경험할
수 있다.

본 서비스는 캐릭터 상황 인지 기반 콘텐츠 제
공, 콘텐츠 난이도 판별, 시각장애 유형별 콘텐츠

조정, 그리고 특화된 유사도 기반 콘텐츠 시퀀스 제
공 기능을 통합하여 시각장애인 맞춤형 시기능 훈
련 환경을 제공한다. 그림 2는 본 시스템의 콘텐츠
처리 절차를 시각적으로 정리한 것이다.

3.4 맞춤형 시각 보정 기능 설계

본 연구에서는 시각장애인의 시각적 특성에 최
적화된 맞춤형 시각 보정 기능을 설계하였다. 시각
보정 기능은 사용자의 시각장애 유형에 따라 실시
간으로 적용되며, XR 글래스를 통해 제공되는 훈
련 콘텐츠에 직접 반영된다. 사용자는 플랫폼에 최
초 접속 시 LMS에 시각장애 유형을 등록하고,
LMS에서 사용자 유형에 따라 사전 정의된 보정
필터 파라미터가 자동으로 설정된다. 이후 플랫
폼 서버의 시각 보정 필터 모듈은 콘텐츠 제공 시 실
시간으로 필터를 적용하여 사용자 맞춤형 훈련 환
경을 제공한다. 시각 보정 필터의 적용은 다음 식
(3)으로 표현된다.

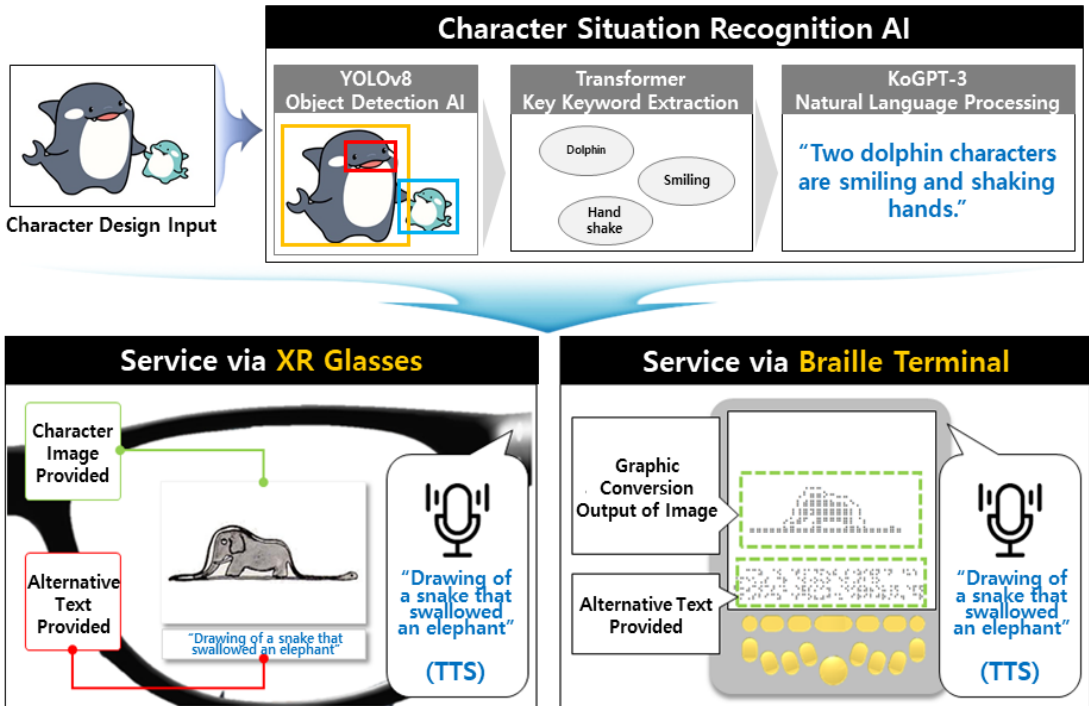


그림 2. 캐릭터 상황 인지 처리 절차
Fig. 2. Character situation recognition process

$$I_{adj}(x,y) = \lambda_{type} \times I(x,y) + \delta_{type} \quad (3)$$

여기서 $I(x,y)$ 는 원본 콘텐츠의 픽셀 값, $I_{adj}(x,y)$ 는 보정 필터가 적용된 픽셀 값, λ_{type} 은 시각장애 유형별 대비 조정 계수, δ_{type} 은 색상 보정 오프셋을 의미한다. 본 수식은 터널 시야 및 색각이상 사용자의 시각 보정 필터 적용에 활용된다.

한편, 황반원공 사용자의 경우 암점 영역의 영상을 주변 시야에 표시하고, 해당 위치 및 크기를 조정하는 별도의 영상 위치 조정 기능이 적용된다. 이에 따라, 중심부 정보 손실을 보완하고 주변 시야 활용을 극대화하도록 설계하였다.

시각장애 유형별 파라미터는 황반원공, 터널 시야, 색각이상 각각에 대해 설정되어 있으며, 터널 시야 사용자에게는 중심부 선명도를 강화하는 보정이 적용되며, 색각이상 사용자에게는 색상 대비가 조정되어 색상 차이 인식이 용이하도록 지원된다. 다음 그림 3은 황반원공 보정 실행 화면이다.

그림 3은 황반원공 사용자를 위한 시각 보정 기능이 적용된 훈련 콘텐츠 실행 화면으로, 중앙 시야의 암점(검은 원) 영역을 회피하여 주변 시야를 활용하도록 유도하는 구조로 구성되어 있다. 화면 중앙의 수직 및 수평 빨간 선은 전체 화면을 좌상, 좌하, 우상, 우하의 4분면으로 구분하는 기준선으로, 저시력자의 주변시 탐색 능력을 훈련하기 위한 시야 분할 도구로 사용된다. 사용자는 콘텐츠가 어느 분면에 위치하는지를 반복적으로 인식하고 시선이 나 고개 움직임을 통해 해당 위치를 추적함으로써, 손상된 중심 시야를 보완하고 주변시 인식 능력을 강화할 수 있다.

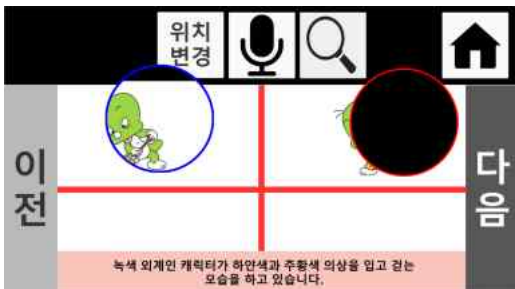


그림 3. 황반원공 보정 실행 화면

Fig. 3. Visual correction example for macular hole users

또한, 플랫폼 서버는 학습자의 학습 이력과 콘텐츠 난이도를 기반으로 보정 필터 매개변수를 지속적으로 최적화한다. 훈련 난이도가 증가함에 따라 보정 강도 역시 동적으로 조절되어 사용자 피로도를 최소화하면서도 효과적인 시기능 훈련이 가능하도록 설계하였다.

이는 난이도가 높아질수록 시각적 정보의 밀도와 복잡도가 증가하게 되므로, 보정 강도를 그대로 유지할 경우 시각 과부하가 발생할 수 있다는 점을 고려한 것이다. 예를 들어, 황반원공 사용자의 경우 고난이도 콘텐츠에서는 중심 보정 영역의 크기를 소폭 축소하거나, 색상 대비 수준을 낮춤으로써 과도한 시각 자극을 완화한다. 이러한 방식은 장시간 훈련 중 시각적 피로 누적을 방지하고 학습 지속 가능성을 높이는 데 효과적이다.

본 시스템의 맞춤형 시각 보정 기능은 기존의 일괄 적용 방식과 달리, 시각장애 유형별 특성과 개인 학습 이력에 따라 정밀하게 조정된다는 점에서 차별성을 가진다.

IV. 시스템 성능 평가 및 사용자 연구

4.1 시스템 성능 측정

본 연구에서는 개발된 XR 글래스 기반 시각장애 맞춤형 시기능 훈련 시스템의 테스트 환경은 다음 표 4와 같다.

표 4. 테스트 환경

Table 4. Test environment specifications

Device	Specification item	Detailed specification
Personal PC	Processor	Intel Core i7 12700K
	Memory	32GB RAM
	Storage	1TB SSD
	Operating system	Windows 11 Pro
XR Glass (XREAL Light)	Display resolution	1920×1080 pixels
	Field of View (FOV)	52°
	Refresh rate	120Hz
	Connection type	USB-C Wired Connection
Smartphone (Samsung Galaxy S23)	Processor	Qualcomm Snapdragon8 Gen2
	Memory	8GB RAM
	Storage	256GB UFS 4.0
	Operating system	Android 13

시기능 훈련 시스템의 성능을 검증하기 위해 평균 응답 속도, 네트워크 동시 접속 처리량, 데이터 전송 정확도, 기능 신뢰성 항목을 평가하였다. 평균 응답 속도와 동시 접속 처리량은 일반 개인용 PC 환경에서 JMeter를 활용하여 100회 반복 측정하였으며, 데이터 전송 정확도와 기능 신뢰성은 스마트폰과 연동된 XR 글래스를 통해 기능 수행 여부를 기준으로 평가하였다. 다음 그림 4는 테스트 환경에서 훈련 콘텐츠를 실행하는 장면을 보여준다.

그림 4에서 화면 상단은 실제 사용자가 XR 글래스를 착용한 상태에서 훈련 서비스를 테스트하는 장면이며, 화면 하단은 캐릭터 이미지 및 설명, 제스처 기반 조작 기능이 제공되는 훈련 콘텐츠 실행 화면이다. 해당 콘텐츠는 XR 글래스를 통해 증강된 UI 형태로 사용자의 시야에 실시간으로 투영된다. 사용자는 캐릭터 이미지, 설명 음성, 난이도에 따라 조정된 텍스트 정보를 직관적으로 인지하며, 스마트폰 컨트롤러를 이용해 이전, 다음, 위치 변경, TTS 듣기 등의 기능을 직접 제어할 수 있다. 시스템 성능 측정 결과는 다음 표 5와 같다.

표 5와 같이, 평균 응답 속도는 0.11초로 측정되었으며 표준편차는 0.02초로 나타났다. 이는 실시간 콘텐츠 제공 서비스의 요구 조건을 충분히 만족하는 수치로, 시각장애인 대상 훈련 서비스에서 중요한 즉각성을 보장하는 수준임을 확인할 수 있다.

동시 접속 처리량은 100회 반복 측정 결과 평균

307.21건/초, 표준편차 10.3건/초로 나타났다. 이는 기존 플랫폼의 요구 기준인 100건/초 이상을 크게 상회하는 수치로, 다수의 사용자가 동시에 접속하는 환경에서도 안정적인 서비스를 제공할 수 있는 성능을 입증하였다.

데이터 전송 정확도와 기능 신뢰성은 XR 글래스와 스마트폰 환경에서 기능 수행 여부를 기준으로 평가하여 각각 99%로 측정되었다. XR 글래스와 스마트폰은 주로 클라이언트 단말로 사용되며, 플랫폼 서버와 통신하여 콘텐츠를 수신하고 출력하는 역할을 수행하기 때문에, 응답 속도나 처리량과 같은 서버 기반의 수치적 성능 측정은 어렵다. 이에 따라 해당 항목은 플랫폼 기능 수행 여부를 True/False로 판단하는 방식으로 평가하였다. 100회 반복 테스트에서 모든 요청이 정상적으로 처리되어 오류 없이 안정적으로 기능 수행이 이루어졌음을 확인하였고, 이러한 특성으로 인해 표준편차는 제시하지 않았다.

표 5. 시스템 성능 평가 결과

Table 5. System performance evaluation results

valuation item	Average value	Standard deviation
Average response time (s)	0.11	0.02
Concurrent throughput (req/s)	307.21	10.3
Data transmission accuracy (%)	99	N/A
Functional reliability (%)	99	N/A



그림 4. 캐릭터 기반 시기능 훈련 콘텐츠 테스트 화면
Fig. 4. Test screen of character-based visual training content

제안한 시스템은 평균 응답 속도, 동시 접속 처리량, 데이터 전송 정확도, 기능 신뢰성 모든 항목에서 우수한 성능을 보였으며, 실시간 시기능 훈련 서비스 제공에 필요한 안정성과 신뢰성을 충족함을 확인할 수 있다.

4.2 캐릭터 상황 인지 AI 모델 처리 성능 평가

본 연구에서는 제안한 XR 글래스 기반 시각장애인 시기능 훈련 시스템의 핵심 기능 중 하나인 캐릭터 상황 인지 AI 모델의 처리 성능을 정량적으로 평가하였다. 캐릭터 상황 인지 프로세스는 객체 탐지, 키워드 추출, 텍스트 생성, 음성 변환의 네 단계로 구성되며, 각 단계의 처리 속도 및 정확도를 독립적으로 측정하였다. 성능 평가는 Intel Core i7-13700K CPU, 128GB DDR5 RAM, NVIDIA RTX 4090 (24GB GPU), Ubuntu 22.04 LTS (64bit) 환경의 고성능 AI 전용 서버에서 수행되었다.

평가는 총 1,000장의 캐릭터 이미지를 대상으로 수행되었으며, 각 이미지에 대해 캐릭터 상황 인지 전체 절차를 1회씩 수행하여 성능을 측정하였다. 각 처리 단계의 평균 소요 시간은 그림 5와 같다.

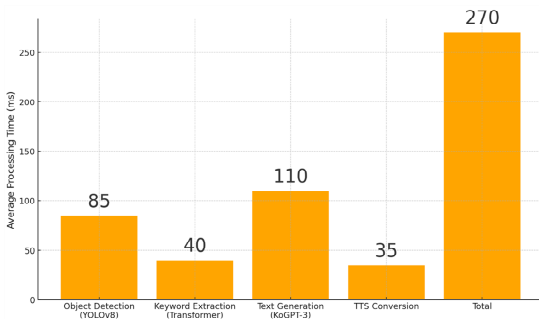


그림 5. 캐릭터 상황 인지 단계별 평균 처리 시간

Fig. 5. Average processing time per character recognition stage

객체 탐지 단계가 평균 85ms, 키워드 추출 40ms, 텍스트 생성 110ms, TTS 변환 35ms로 측정되었으며, 전체 처리 시간은 평균 270ms로 확인되었다. 이는 실시간 콘텐츠 제공에 적합한 수준으로, 실 사용 환경에서도 지연 없이 서비스가 가능함을 보여준다. 정확도 측정은 다음 표 6과 같이 세 가지 항목을 기준으로 수행되었다.

표 6. 캐릭터 상황 인지 정확도 평가 항목 정의
Table 6. Definitions of evaluation criteria for character situation recognition

Evaluation item	Definition
Object detection accuracy	Accuracy of detected bounding boxes compared to ground truth annotations
Keyword extraction accuracy	Semantic inclusion and correctness of keywords extracted from detected objects
Sentence generation accuracy	Appropriateness of generated sentence, rated by human evaluators (≥ 4 on 5-point scale)

위 기준에 따라 총 1,000장의 이미지를 평가한 정확도 측정 결과는 다음 표 7과 같다.

표 7. 캐릭터 상황 인지 정확도 평가 결과
Table 7. Accuracy evaluation results for character situation recognition

Evaluation item	Accuracy (%)	Description
Object detection accuracy	96.96%	Based on ground truth bounding boxes
Keyword extraction accuracy	98.98%	Based on semantic match with all detected objects
Sentence generation accuracy	89.64%	Rated ≥ 4 by at least 2 of 3 human evaluators on 5-point scale

Object Detection Accuracy는 정답 박스 기준 96.96%로 높은 수준의 객체 탐지 성능을 보였으며, Keyword Extraction Accuracy는 98.98%로 추출된 키워드의 의미적 일치도 또한 매우 우수한 수준이었다. Sentence Generation Accuracy는 전문가 평가 기준 89.64%로, 시각장애인이 이해하기에 충분한 문장 생성 품질을 확보하였다.

제안한 캐릭터 상황 인지 AI 모델은 실시간 처리 속도와 높은 정확도를 동시에 만족하며, 시각장애인 맞춤형 훈련 콘텐츠 자동 생성 시스템의 핵심 AI 모듈로서의 활용 가능성을 입증하였다.

V. 결론 및 향후 과제

본 연구에서는 시각장애인을 위한 XR 글래스 기

반 시기능 훈련 시스템을 설계하고 구현하였다. 제안된 시스템은 시각장애 유형별 특성을 반영하여 맞춤형 콘텐츠 제공과 실시간 시각 보정을 지원하며, LMS, LCDMS, AI 콘텐츠 처리 모듈, 보정 필터 모듈 등을 통합한 플랫폼 구조를 갖는다. 특히 캐릭터 기반 훈련 콘텐츠와 AI 모델을 결합함으로써, 기존 단순 정보 제공 방식에서 벗어나 상황 기반 실시간 학습 환경을 구현하였다.

캐릭터 상황 인지 기능은 객체 탐지, 키워드 추출, 텍스트 생성, TTS로 이어지는 자동화 과정을 통해 실시간 훈련 콘텐츠를 제공하며, 사용자 유형에 따라 시각 보정 필터가 동적으로 적용된다. 시스템 성능 평가 결과, 평균 응답 속도 0.11초, 동시 접속 처리량 307건/초, 데이터 전송 정확도와 기능 신뢰성 99%를 기록하였다. AI 처리 성능 평가에서는 전체 인지 과정이 평균 270ms 이내로 처리되었으며, 객체 탐지 정확도 96.96%, 키워드 추출 정확도 98.98%, 문장 생성 정확도 89.64%를 나타내어 실시간성과 정확성 모두에서 우수한 성능을 입증하였다.

본 시스템은 단순 보조 기능을 넘어, 시각장애인의 훈련 목적에 특화된 자율형 학습 플랫폼으로서 실질적인 활용 가능성을 보인다. 향후 연구에서는 다양한 시각장애 유형과 연령대에 맞춘 콘텐츠의 다양화, 사용자 피드백을 기반으로 한 보정 필터의 자동 조정 기능, 훈련 효과에 대한 장기적인 추적 분석, 다국어 지원을 위한 설명 생성 모델 고도화 등을 추진할 계획이다. 또한, 실제 교육기관 및 재활 훈련 현장에서의 적용 가능성에 대한 임상 기반 검증 연구도 필요할 것이다. 아울러, 본 연구에서는 시스템 성능과 AI 처리 정확성에 중점을 두어 평가를 수행하였으며, 실제 사용자 기반의 시기능 향상 효과 검증은 추후 임상 환경 기반의 장기적 연구로 확장해 나갈 예정이다. 특히 다양한 연령과 장애 정도를 아우르는 사용자 군을 대상으로 한 실사용 평가 및 훈련 효과의 객관적 지표화를 통해, 본 시스템의 임상적 활용 가능성을 심층적으로 분석하는 후속 연구가 필요할 것이다.

References

- [1] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: A survey", *IEEE Trans. Syst., Man, Cybern. C*, Vol. 40, No. 1, pp. 25-35, Jan. 2010.
- [2] J. Kasowski, B. A. Johnson, R. Neydavood, A. Akkaraju, and M. Beyeler, "A Systematic Review of Extended Reality (XR) for Understanding and Augmenting Vision Loss", *Journal of Vision*, Vol. 23, No. 5, pp. 5, 2023. <https://doi.org/10.1167/jov.23.5.5>.
- [3] S. Jeong, Y.-H. Lee, and E. Jung, "Semi-supervised learning YOLO framework based on few-shot learning and Grounding DINO for blind pedestrian safety", *Journal of the Korea Institute of Information Technology (JKIIT)*, Vol. 22, No. 12, pp. 39-46, Dec. 2024. <https://doi.org/10.14801/jkiit.2024.22.12.39>.
- [4] KakaoBrain, "KoGPT API: Korean Generative Pre-trained Transformer Service", Kakao DevTalk, <https://devtalk.kakao.com/t/kogpt-api-notice-kakao-brain-kogpt-api-is-released/125385>. [accessed: Dec. 05, 2024]
- [5] J. Cho, H. Kang, and J. Lee, "An end-to-end TTS system using Tacotron and WaveGlow", *IEEE Access*, Vol. 8, pp. 226941-226949, 2020.
- [6] M. A. Hersh and M. A. Johnson, "Assistive Technology for Visually Impaired and Blind People", Springer, 2008. <https://doi.org/10.1007/978-1-84628-867-8>.
- [7] A. Hub, J. Diepstraten, and T. Ertl, "Design and development of an indoor navigation and object identification system for the blind", *Proc. 6th Int. ACM SIGACCESS Conf. Comput. Accessibility (ASSETS)*, Atlanta, GA, USA, pp. 147-152, Sep. 2003. <https://doi.org/10.1145/1029014.1028657>.
- [8] R. Manduchi and J. Coughlan, "The last meter: Blind visual guidance to a target", *Proc. 32nd Annu. ACM Conf. Human Factors Comput. Syst. (CHI)*, Toronto, ON, Canada, pp. 3113-3122, Apr. 2014. <https://doi.org/10.1145/2556288.2557328>.
- [9] C. P. Janssen, J. Schoemans, and D. H. Brouwer,

"Using virtual reality to train children with visual impairments in mobility skills", Proc. ACM SIGACCESS Conf. Comput. Accessibility (ASSETS), Pittsburgh, PA, USA, pp. 249-250, Oct. 2009.

- [10] Microsoft, "Microsoft HoloLens official website", <https://www.microsoft.com/en-us/hololens>. [accessed: Dec. 05, 2024].
- [11] Magic Leap, "Magic Leap 2 product page", <https://www.magicleap.com/magic-leap-2>; XREAL, "XREAL official site", <https://www.xreal.com/>. [accessed: Dec. 05, 2024].
- [12] Startup Selfie, "Magic Leap 2 brings augmented reality to the operating room", <https://www.startupselfie.net/2024/10/05/magic-leap-2-augmented-reality/>. [accessed: Dec. 05, 2024].
- [13] D. C. Fletcher and R. A. Schuchard, "Preferred retinal loci relationship to macular scotomas in a low-vision population", Ophthalmology, Vol. 104, No. 4, pp. 632-638, Apr. 1997. [https://doi.org/10.1016/s0161-6420\(97\)30260-7](https://doi.org/10.1016/s0161-6420(97)30260-7).
- [14] T. Kuyk and J. L. Elliott, "Visual factors and mobility in persons with age-related macular degeneration", J. Rehabil. Res. Dev., Vol. 36, No. 4, pp. 303-312, Jul. 1999.
- [15] J. Birch and R. U. Kolle, "Efficiency of the Ishihara test for identifying red-green colour deficiency", Ophthalmic Physiol. Opt., Vol. 14, No. 2, pp. 143-148, Apr. 1994.
- [16] J. Terven and D. Cordova-Esparza, "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS", arXiv preprint, arXiv:2304.00501, Apr. 2023. <https://doi.org/10.48550/arXiv.2304.00501>.
- [17] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection", Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Venice, Italy, pp. 2980-2988, Oct. 2017. <https://doi.org/10.1109/ICCV.2017.324>.
- [18] A. Vaswani, et al., "Attention is all you need", Adv. Neural Inf. Process. Syst., Vol. 30, pp.

5998-6008, Jun. 2017. <https://doi.org/10.48550/arXiv.1706.03762>.

저자소개

김 갑 열 (Kapyol Kim)



2007년 2월 : 경원대학교
전자계산학과(공학석사)
2025년 2월 : 가천대학교
컴퓨터공학과(공학박사수료)
2014년 6월 ~ 현재 :
(주)씨티소프트 연구원
관심분야 : 객체인식, 보조공학,
인공지능, 엣지컴퓨팅, IoT, 네트워크

정 인 철 (Incheol Jeong)



2020년 2월 : 가천대학교
IT융합공학과(공학석사)
2025년 2월 : 가천대학교
IT융합공학과(공학박사)
2025년 3월 ~ 현재 : 가천대학교
컴퓨터공학과 대우교수
관심분야 : 객체인식, 보조공학,
인공지능, 엣지컴퓨팅, IoT

조 진 수 (Jinsoo Cho)



1998년 2월 : 콜롬비아대학교
전기공학과(공학석사)
2003년 2월 : 조지아공과대학교
전기공학과(공학박사)
2006년 4월 ~ 현재 : 가천대학교
컴퓨터공학과 교수
관심분야 : 보조공학, 스마트
헬스케어, IoT, 인공지능, 영상처리