

연구소 소개

건국대학교 스마트미디어연구소

윤경로 교수 / 건국대학교 컴퓨터공학부

I. 스마트미디어연구소 현황

스마트미디어연구소는 미디어에 대한 이해를 기반으로 하는 다양한 응용 기술의 개발과 국제 표준화를 수행하고 있다. 핵심 연구 분야의 첫 번째는 인공지능 기술을 이용한 영상 이해에 기반한 기술 개발로 영상 내 객체 탐지, 포즈 추정, 영상 콘텐츠 이해 등 다양한 딥러닝 기반 연구를 진행 중에 있으며, 이러한 기술을 기반으로 기계를 위한 영상 특징 추출 및 압축 기술에 대한 과제를 진행 중에 있다. 두 번째는 메타버스 관련 기술 중 특히 현실 융합형 메타버스 기술의 개발로, 가상 세계와 물리 세계의 동기화를 위한 기술과 이에 기반한 디지털트윈 기술의 개발을 진행 중에 있다. 세 번째는 국제 표준화 활동으로 ISO/IEC JTC1 SC29 산하의 MPEG 표준화 활동과 IEEE SA Digital Content Technology Standards Committee 표준화 활동을 활발히 진행 중이다.



현재 연구실에서 진행하는 과제의 주제로는 기계를 위한 비디오 코딩 표준 기술 개발, 메타버스를 위한 미디어사물 인터넷 인공지능 및 NFT 인터페이스 표준 기술 개발, 인공지능을 이용한 태권도 판정 보조 시스템 개발, 인공지능을 이용한 수입신고의 위험관리 모델 개발, 가상 세계와 물리 세계의 인터페이스 기반 표준화 등이 있다.

II. 스마트미디어연구소 수행 연구

1. 기계를 위한 비디오 코딩 표준 기술 연구

압축되지 않은 원본 비디오 데이터는 막대한 저장 공간을 차지하며, 저장하거나 전송하는 데 막대한 비용이 소모된다. 또한, 대용량 데이터를 네트워크로 전송할 경우 매우 넓은 대역폭이 필요하며, 제한된 대역폭 환경에서는 전송 속도가 저하되고 실시간 전송 시 지연 및 끊김 현상이 발생할 수 있다. 이러한 문제들을 해결 또는 개선하기 위하여 비디오 코딩 기술이 개발되기 시작하였으며, 이는 인간의 인지 체계를 기반으로 체감 측면에서의 무손실을 지향하며 시간

<표 1> 기계를 위한 비디오 코딩의 특성

	사람을 위한 비디오 코딩	기계를 위한 비디오 코딩
개념도		
영상신호의 목적	사람에게 정보 전달	기계의 임무수행을 위한 정보 전달
영상부호화 목표	인지적 무손실/화질 유지하며 압축률 최대화	임무성능을 유지하며 압축률 최대화
부호화 복잡도 및 전력소비	상대적으로 중요치 않음	기계가 실시간 연동을 위하여 매우 중요
중앙집중처리	높은 복잡도로 인하여 부적합	저복잡도 구현을 고려하므로 적합

적, 공간적, 통계적 중복성을 줄이며 동시에 인간이 인지하지 못하는 정보를 제거하는 방식으로 압축하는 방향으로 발전해 왔다. 그러나 인공지능 기술의 발전으로 최근 동영상의 소비 주체가 변화하고 있다. 감시 시스템, 자율 주행 자동차, 교통 관리 시스템 등 다양한 산업 분야에서 딥러닝 기반 기술이 적용되고 있으며 막대한 양의 미디어 데이터가 인공지능 학습 및 지능형 미디어 분석을 위해 사용되고 있다. 이는 기존 인간 중심 멀티미디어 소비 환경에서 기계 중심 소비 환경으로의 전환을 시사한다.

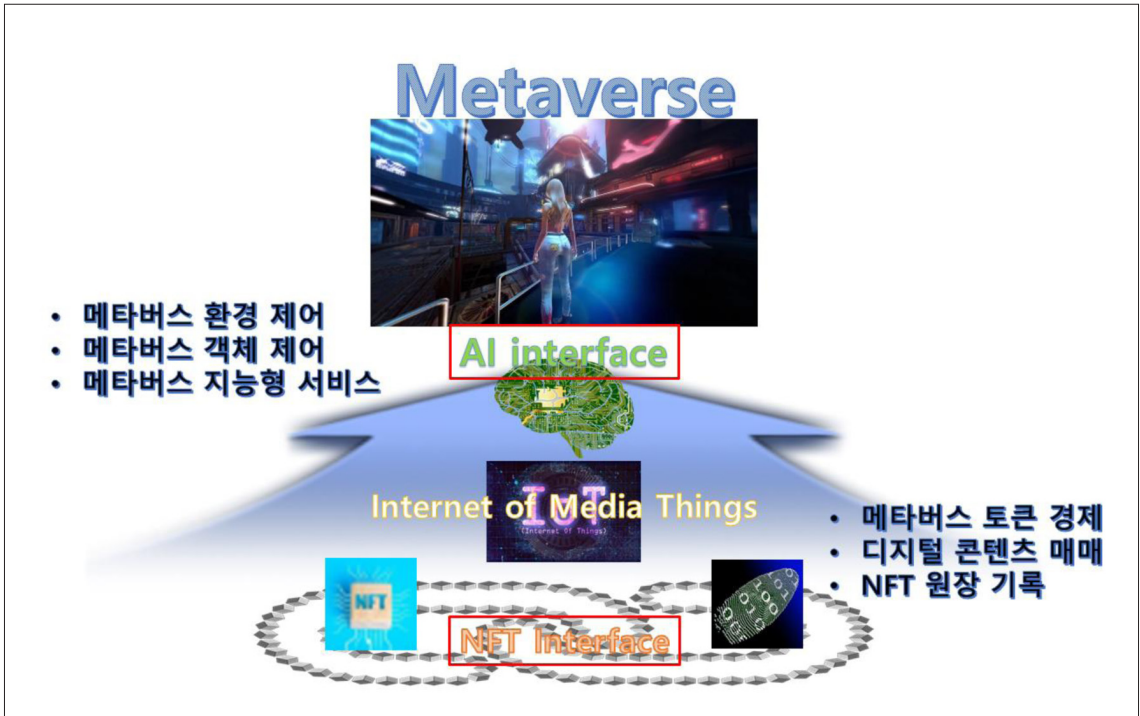
기존의 비디오 코딩 표준은 인간의 시각적 인지 특성에 맞추어 개발되었기 때문에, 압축률이 높아지는 경우 기계가 동영상 분석하는 데는 적합하지 않을 수 있으며, 머신 비전 작업을 위한 코딩 기술을 표준화하고, 데이터를 더 작은 크기로 압축하면서도 높은 분석 성능을 유지하는 것으로 하는 기계를 위한 비디오 코딩 기술의 표준화가 진행되고 있다.

<표 1>은 사람을 위한 비디오 코딩과 기계를 위한 비디오 코딩 요구사항의 특성적 차이점을 정리하여 보여 주고 있다. 코딩의 대상이 되는 영상신호를 저장/전송하는 목적 자체가 사람이 소비 주체가 되는 경우와 기계가 소비 주체가 되는 경우가 다르며, 사람이 영상을 소비하는 경우 바라보거나 인지하는 비디오 신호의 특징 요소와 기계가 컴퓨터 비전과 같은 임무 수행을 위하여 영상을 소비하는 경우 사용하는 비디오 신호의 특징 요소가 다를 수 있다는 것을 전제로 하고 있다. 또한 사람이 비디오를 소비하는 경우의 용례와 기계가 비디오를 소비하는 경우의 용례가 다르므로 그에 따른 요구사항도 달라지는 것을 보여 주고 있다.

본 연구실은 이렇게 기계가 비디오를 소비하는 경우 달라지는 용례에 따라서 머신 비전 임무의 성능이 떨어지지 않거나 적게 떨어지면서도 압축률은 올리는 효율적인 비디오 코딩 기술을 연구하며 이를 ISO/IEC JTC 1/SC 29 산하의 MPEG 표준화 그룹을 통한 표준에 반영시키기 위하여 노력하고 있다.

2. 메타버스를 위한 미디어사물인터넷 인공지능 표준 기술 연구

미디어사물인터넷은 현실 세계에서 수집된 미디어 데이터를 서로 교환할 수 있는 인터페이스를 제공하는 것을 목표



<그림 1> 메타버스를 위한 미디어사물인터넷 인공지능 및 NFT 인터페이스 개요도

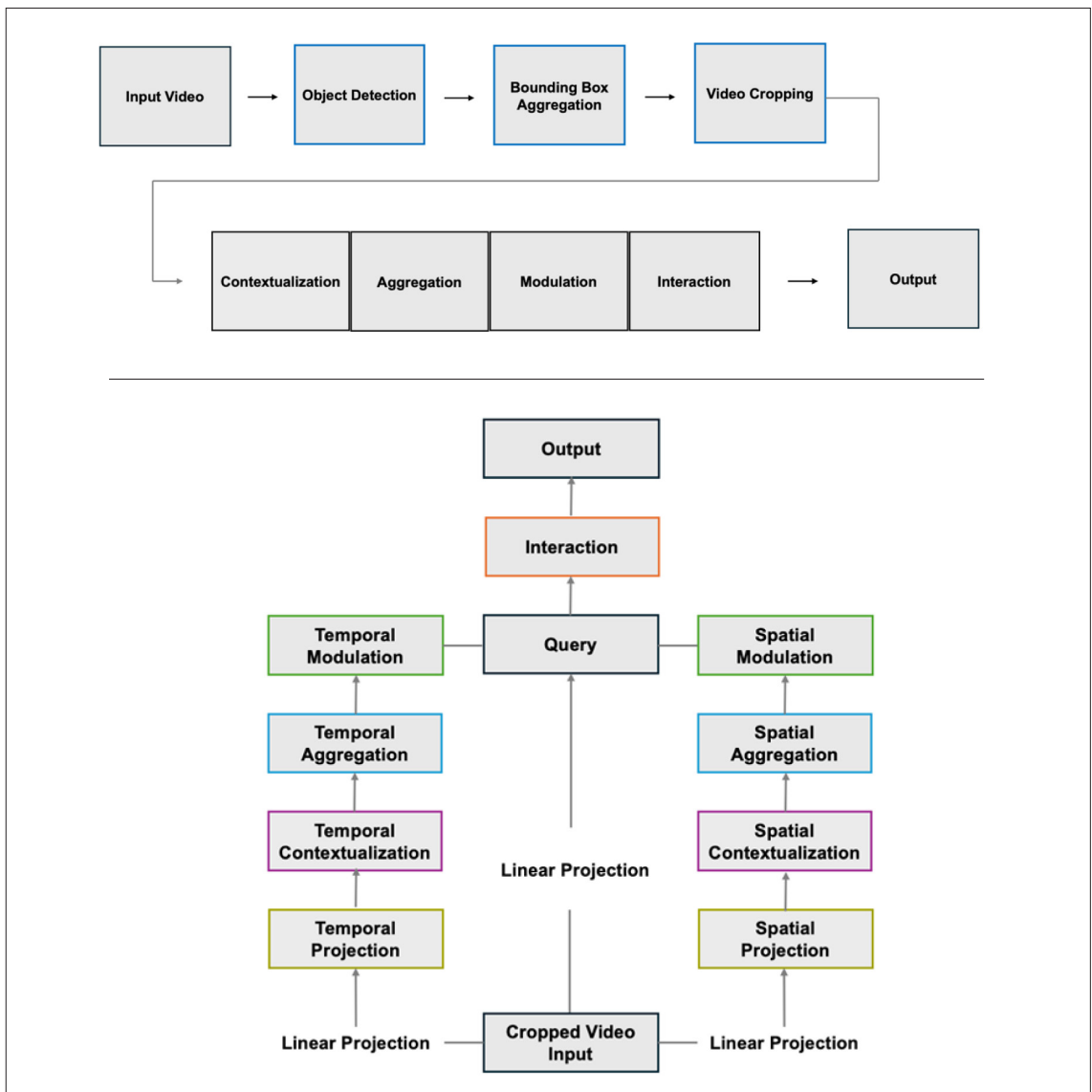
로 한다. 미디어 센서를 통해 획득한 데이터(예: 사진, 동영상, 소리)는 인공지능 기반의 미디어 분석기를 통해 분석되고, 이를 통해 연결된 다른 미디어사물을 제어할 수 있으며, 이러한 데이터의 수집과 분석을 통해 다양한 지능형 미디어 서비스가 가능하게 된다. 또한, 현실 세계의 미디어 데이터를 인공지능으로 분석한 결과는 메타버스 공간에서 환경을 제어하기 위한 현실 세계의 미디어 데이터 분석, 메타버스 공간 내 객체를 조작하기 위한 미디어 데이터 분석, 지능형 서비스의 컨텍스트 및 시나리오 제어를 위한 데이터 분석 등에 활용될 수 있으며, 이러한 인공지능 분석 결과를 효과적으로 표현하기 위해선 표준화된 데이터 포맷과 메타버스 공간으로 전달할 수 있는 API를 개발한다. 미디어사물(센서)로부터 얻어진 콘텐츠는 NFT로 생성되어 거래될 수 있으며, 현실 세계의 미디어사물에서 생성된 콘텐츠를 메타버스에서 블록체인과 연결된 NFT로 만드는 표준 또한 필요하다. 블록체인을 통해 NFT의 생성, 거래, 소각과 거래 기록 및 소유권을 확인할 수 있는 표준 인터페이스가 요구되며, 이를 위해 미디어사물로부터 생성된 콘텐츠에 대한 메타데이터 표준화와 블록체인 연계 API를 개발한다. <그림 1>은 메타버스를 위한 미디어사물인터넷 인공지능 및 NFT 인터페이스 개요도로 미디어사물인터넷으로 수집된 콘텐츠의 인공지능 분석을 통하여 메타버스의 환경제어, 객체제어 및 지능형 서비스를 제공할 수 있으며, 미디어사물인터넷을 통하여 생성된 콘텐츠를 NFT로 등록하여 NFT 토큰 기반의 경제 생태계를 구축할 수 있음을 보이고 있다.

3. 인공지능 기반 태권도 판정 보조 시스템

태권도 겨루기 종목은 2000년 시드니 올림픽부터 2028년 LA 올림픽까지 8회 연속으로 올림픽 정식 종목으로 채

택되었으며, 이는 태권도가 세계적인 스포츠로 자리 잡았다는 것을 의미한다. 또한 2023년 12월 기준으로 태권도는 아시아, 유럽, 팬아메리카, 아프리카, 오세아니아 등 5개 대륙 211개 회원국에서 약 1억 명이 수련하고 있다.

하계 올림픽의 핵심 종목인 태권도 겨루기에서는 비태권도 동작(변칙 발차기)으로 인한 득점을 방지하기 위하여 득점 후 취소의 상황이 벌어지고 있으며 때로는 득점이 발생할 가능성도 있다. 이러한 비태권도 동작(변칙 발차기)으로 인한 득점이나 득점 후 취소의 상황이 줄어들게 된다면 경기의 운영이 더욱 공정해지고, 판정 신청으로 인한 지연이 최소화되어 경기가 원활하게 진행되며 관중의 흥미도 높아질 것으로 기대된다. 이러한 변칙 발차기 동작으로 인하여 발생하는 경기 판정 이슈를 줄이고 공정성과 객관성을 높이기 위해 인공지능(AI) 기반의 판정 보조 시스템을 개발할



<그림 2> 태권도 겨루기 동작 판정을 위한 인공지능 모델 구조도. YOLO 기반 객체 (선수) 탐지 모델(위) VideoFocalNet 기반 동작 인식 모델(아래)

수 있다.

2023년에는 인공지능 기반 판정 지원 시스템 프로토타입 개발의 결과를 통해 이러한 가능성을 확인하였고, 개발 중 얻은 주요 태권도 겨루기 동작 데이터를 활용하여 AI 기반 동작 판정을 위한 학습용 데이터를 확보하고 판정 엔진의 성능을 향상시키고 있다. 인공지능 모델의 성능 향상을 통해 실용화를 목표로 하고 개발하고 있으며, 동시에 경기 운영 시스템과 통합 운영이 가능하도록 개발하고 있다.

<그림 2>는 태권도 겨루기 동작 중 비태권도 동작의 판정을 위한 인공지능 모델의 구조도를 보여 주고 있다. 연구 중인 모델은 두 단계로 나누어져 있다. 먼저 1단계에서는 YOLOv8를 기반으로 하는 모델을 이용하여 카메라 입력 영상에서 사람들을 인식하고 상대적인 박스의 크기, 위치 등을 고려하여 선수라 추정되는 박스들만을 남기게 되고, 2단계에서는 VideoFocalNet을 기반으로 선수의 동작을 식별하게 된다. 이러한 과정에서 정확도의 향상을 위하여 적절한 loss function, optimizer, pooling method 등의 비교 분석 및 최적화를 진행하고 있다.

III. 스마트미디어연구실 최근 수상 현황

- 2024년 : 한국방송·미디어공학회 추계학술대회 최우수논문상 (이예지 외 3인)
- 2021년 : 한국방송·미디어공학회 하계학술대회 우수 학생논문상 (이예지)
- 2020년 : 한국방송·미디어공학회 추계학술대회 최우수 학생논문상 (이예지)

저 자 소개



윤 경 로

- 1987년 : 연세대학교 공과대학 전자전산공학과 학사
- 1989년 : 미시건대학교(앤아버) 전기컴퓨터공학과 석사
- 1999년 : 시라큐즈대학교 전산학 박사
- 1999년 ~ 2003년 : LG전자 전자기술원 책임연구원/그룹장
- 2003년 ~ 현재 : 건국대학교 컴퓨터공학부/스마트ICT융합공학과 교수

- 2017년 ~ 현재 : 과학기술정보통신부 멀티미디어부호화 전문위원회(SC29K) 대표전문위원
- 2019년 ~ 현재 : IEEE SA 2888 Working Group 의장
디지털가상화포럼 의장
- 2022년 ~ 현재 : 메타버스 얼라이언스 기술표준분과위원장
메타버스미래포럼 의장
- 2024년 ~ 현재 : IEEE SA 디지털콘텐츠기술 표준화위원회 (C/DCTSC) 의장
- 2025년 ~ 현재 : 한국방송·미디어공학회 회장
- 주관심분야 : 스마트미디어시스템, 영상처리, 컴퓨터비전, 메타버스, 디지털트윈