

# 2025년 한국방송·미디어공학회 추계학술대회 후기

최해철 (조직위원장, 국립한밭대학교), 이진영 (프로그램위원장, 세종대학교)



## 조직위원회

- 대회장 : 윤경로 회장 (한국방송·미디어공학회/건국대학교)
- 조직위원장 : 최해철 교수 (국립한밭대학교)
- 프로그램위원장 : 이진영 교수 (세종대학교)
- 프로그램위원
  - 강정원 실장 (ETRI)
  - 김상욱 교수 (중앙대학교)
  - 류은석 교수 (성균관대학교)
  - 서영호 교수 (광운대학교)
  - 심동규 교수 (광운대학교)
  - 이영호 교수 (목포대학교)
  - 정진우 책임 (KETI)
  - 추현곤 실장 (ETRI)
  - 강제원 교수 (이화여자대학교)
  - 김휘용 교수 (경희대학교)
  - 서광덕 교수 (연세대학교)
  - 서정일 교수 (동아대학교)
  - 오병태 교수 (한국항공대학교)
  - 장한얼 교수 (국립한밭대학교)
  - 조병철 교수 (동아방송예술대학교)



## 추계학술대회 프로그램 구성 및 내용

최근 방송과 미디어 산업은 기술 융합과 사용자 중심의 경험 혁신을 중심으로 빠르게 재편되고 있다. 메타버스, 디지털 휴먼, 확장현실 기술은 방송 콘텐츠를 단순히 보는 미디어에서 경험하는 미디어로 확장시키고 있으며, AI 기반의 자동 편집, 음성 합성 등은 제작 효율성과 창의성을 동시에 추구하고 있다. 이러한 변화 속에서 한국방송·미디어공학회는 급변하는 기술 트렌드를 선도하고 학계와 산업계의 협력을 통해 미래 미디어 생태계의 주도적 역할을 수행하기 위하여 2025년도 추계학술대회를 개최하였다.

이번 추계학술대회는 총 9개의 특별 세션, 1개의 일반 세션, 그리고 5개의 포스터 세션으로 구성되었다. 특별 세션에서는 〈기계를 위한 부호화〉, 〈비디오 부호화〉, 〈3D 공간 미디어〉, 〈메타버스 산업을 위한 영상 미디어 압축〉, 〈3D 정보 추출, 표현 및 압축〉, 〈생성형 콘텐츠 및 보호〉, 〈디지털 휴먼〉, 〈Audio Coding for Machines〉 등 최신 기술 동향과 연구 결과가 심도 있게 논의되었다. 또한 일반 세션과 포스터 세션에서는 지능형 미디어 신호처리 및 산업 응용을 비롯하여 딥러닝을 활용한 미디어 처리 등 다양한 연구 결과가 발표되었으며, 이는 기존 미디어의 한계를 넘어서고 시청자들에게 더욱 몰입감 있는 경험을 제공하는 데 기여하는 기술들로 주목받았다.

이번 학술대회에서는 국내외 미디어 기술을 선도하는 전문가들의 특별 초청 강연을 통해 미래 기술의 방향성과 비전을 제시하는 계기를 마련하였다. KETI 임태범 본부장은 ‘미디어 AX 가속화 및 안착을 위한 고려사항’을 주제로, GS리테일 이찬종 팀장은 ‘GEN AI를 활용한 홈쇼핑 콘텐츠 개선 및 NDI system 구축’에 대해, NEC Corporation의 Jianquan Liu Director는 ‘An Industry Perspective: Video Analytics meets Generative AI’에 대해 강연하였다. 또한 일본 게이오 대학교의 Hideo Saito 교수는 ‘Journey of Leveraging Multiple Cameras in Computer Vision’을, 리콘랩스의 반성훈 대표는 ‘AI 영상 생성 파이프라인을 위한 에이전틱 시스템 구축 사례’를 주제로 심도 있는 내용을 전달하였다.

아울러 차세대 방송·미디어 산업의 주역이 될 인재들을 위한 〈대학생 논문 및 캡스톤디자인 경진대회〉도 함께 개최되었다. 학생들은 자신만의 창의적인 아이디어와 연구 성과를 발표하고, 분야별 전문가들로부터 직접 피드백을 받는 시간을 가짐으로써 실질적인 연구 경험을 쌓고 미래 미디어 기술에 대한 이해를 넓히는 소중한 기회를 얻을 수 있었다.

## 초청 강연

KETI 임태범 본부장은 ‘미디어 AX 가속화 및 안착을 위한 고려사항’을 주제로, 미디어 산업이 디지털 전환을 넘어 인공지능 전환으로 나아가는 과정에서 필수적으로 고려해야 할 요소들을 소개하였다. 성공적인 미디어 AX를 실현하기 위해서는 양질의 데이터를 지속적으로 확보하고 AI 기술의 신뢰성을 검증하는 것이 무엇보다 중요함을 강조하였으며, 미디어 콘텐츠의 생성과 이해 그리고 신뢰를 아우르는 기술 개발 방향과 이를 통해 산업 현장의 문제를 해결하고 AX를 안착시키기 위한 구체적인 전략을 제시하였다.

GS리테일 이찬중 팀장은 ‘GEN AI를 활용한 홈쇼핑 콘텐츠 개선 및 NDI system 구축’을 주제로, 생성형 AI를 활용하여 상품 연출 이미지나 방송 배경화면 등 방송 제작에 필요한 시각 결과물을 생성하고, 이를 실제 콘텐츠 제작에 적용하여 효율성을 높인 실증 사례를 공유하였다. 또한, 기존 하드웨어 중심의 방송 시스템을 네트워크 기반의 NDI 시스템으로 전환하여, 물리적 제약을 줄이고 유연한 방송 제작 환경을 구축한 성과를 소개하였다.

NEC Corporation의 Jianquan Liu Director는 ‘An Industry Perspective: Video Analytics meets Generative AI’를 주제로, 생체 인식 및 군중 분석 등 NEC의 핵심 비디오 분석 기술이 생성형 AI 및 대규모 비전 모델(LVM)과 결합된 최신 산업 동향을 소개하였다. 이를 통해 공공 안전 및 마케팅 분야에서의 활용성을 설명하고, 물리적 세계를 이해하고 추론하는 월드 모델로의 기술 발전 방향을 제시하였다.

일본 게이오 대학교의 Hideo Saito 교수는 ‘Journey of Leveraging Multiple Cameras in Computer Vision’을 주제로, 다중 카메라를 활용하여 수술실 등의 임의 시점 영상을 생성하는 자유 시점 비디오(Free-viewpoint Video) 기술과 장애물 뒤의 상황을 시각화하는 투시 비전(See-through Vision) 기술에 대해 소개하였다. 또한, 최근 주목받고 있는 NeRF 및 3D Gaussian Splatting(3DGS) 기술을 활용한 스포츠 영상 분석 및 3D 시각화 연구 성과를 공유하였다.

리온랩스 반성훈 대표는 ‘AI 영상 생성 파이프라인을 위한 에이전틱 시스템 구축 사례’를 주제로, 기존 생성형 AI 비디오의 제어력 문제를 해결하기 위해 3D Gaussian Splatting을 활용한 4D 생성 파이프라인을 소개하였다. 특히, 단순한 도구 활용을 넘어 대규모 언어 모델(LLM)이 기획부터 3D 에셋 생성, 모션 적용, 렌더링까지 전 과정을 스스로 제어하고 수행하는 에이전틱 워크플로우 기반의 시스템 구축 사례를 통해 상용 콘텐츠 제작의 새로운 가능성을 제시하였다.



## 특별 세션

특별 세션은 9개로 구성되었으며, 총 38편의 논문이 발표되었다.

### 1. 특별 세션 1: 기계를 위한 부호화

특별 세션 1에서는 ‘기계를 위한 부호화’를 주제로 4편의 논문이 발표되었으며, 한국항공대학교/인시그널(1편), 연세대학교/ETRI(1편), 국립한밭대학교/ETRI(2편)가 참여하였다. 논문 발표는 기계 인식 성능을 고려한 비디오 부호화 및 데이터 압축 최적화 방안을 중심으로 이루어졌다. VCM(Video Coding for Machines)은 기계 학습 및 비전 처리를 위한 비디오 부호화 표준이며, FCM(Feature Coding for Machine)은 비디오에서 추출된 특징을 효율적으로 압축해 기계 학습 성능을 최적화하는 기술이다.

‘FCM을 위한 NNC 내부 코덱 적용기법’은 FCM 과정에서 VVC를 통한 특징맵 양자화 과정을 NNC로 대체하는 방식을 제안하였다. 제안된 방식은 기존 대비 압축률 절감과 기계 성능 향상을 동시에 달성함으로써 효율적인 특징맵 압축 가능성을 확인하였다. ‘VCM의 공간적 리샘플링을 위한 하이브리드 스케일 계수 보정’은 VCM에서 공간적 리샘플링 단계의 효율을 향상시키기 위해 입력 영상의 해상도 특성에 따라 스케일 계수 보정 방식을 유연하게 전환하는 하이브리드 조정 기법을 제안하였다. 이로 인해 다양한 해상도 조건에서 압축 효율과 객체 탐지 정확도의 일관된 성능을 확보하는 결과를 보였다. ‘FCM을 위한 양자화 매개변수 적응적 특징 스케일링 방법’은 기존 VVC 양자화 단계에서 발생하는 양자화 왜곡을 줄이고자 양자화 매개변수 값에 따라 변화하는 왜곡 보정을 위한 적응적 스케일링 방법을 제안하였다. 제안 방법은 양자화 단계에서 발생한 왜곡을 완화하고 특징의 통계적 안정성을 확보하여 개선된 압축 효율을 달성했다. ‘FCM을 위한 통계적 특성 기반 비선형 변환 방법’은 VVC로 부호화하기 위한 기존 균등 양자화 방법을 입력 특징의 통계적 특성에 따라 비선형 강도를 조절하는 비선형 변환 기반 특징 압축 방법을 제안하였다. 제안 방식은 특징 분포 조정으로 양자화 대푯값의 발생 빈도를 높여 압축 효율을 높이는 결과를 보였다. 본 세션은 기계를 위한 부호화 관점에서 압축 효율과 기계 인식 성능을 동시에 개선하는 기계 중심 부호화 기술의 발전 방향을 논의하였다.

### 2. 특별 세션 2: 비디오 부호화

특별 세션 2에서는 ‘비디오 부호화’를 주제로 5편의 논문이 발표되었으며, 한국항공대학교/ETRI(1편), 경희대학교(1편), 국립한밭대학교/ETRI(3편)가 참여하였다. 주요 발표는 VVC 이후 차세대 비디오 부호화 표준 ECM(Enhanced Compression Model)을 목표로 ECM 기반 화면 내 예측 구조 개선과 예측 모드 선택 효율화 그리고 새로운 영상 콘텐츠 특성에 대응하는 부호화 기법을 중심으로 이루어졌다.

‘ECM의 템플릿 기반 Intra Prediction Fusion 개선기법’은 ECM 환경에서 화면내 예측 성능을 향상시

키는 방법을 제안하였다. 기존 ECM 18.0에서 동일한 방향성 모드에 대해 서로 다른 참조라인으로 생성된 예측 블록을 융합하는 Intra Prediction Fusion 구조를 대상으로 융합 생략을 포함한 템플릿 비용 기반 가중치 결정 기법을 적용하여 불필요한 융합을 줄이고 예측 효율을 개선함으로써 휘도 및 색차 성분에서 비디오 부호화 성능 향상을 달성하였다. ‘Evaluation of Compression Artifacts and Visual Quality in AI-Generated Videos Using VVC/H.266’은 최근 확산되고 있는 AI 생성 영상을 대상으로 기존 비디오 코덱의 부호화 특성을 분석하였다. VVC로 AI 생성 영상을 압축 후 복원하여 객관적 지표 및 시각적 열화를 분석한 결과, 자연 영상과는 다른 플리커, 밴딩, 텍스처 붕괴와 같은 특성이 두드러져 AI 생성 영상에 특화된 인지적 비디오 부호화 프레임워크의 필요성을 제시하였다. ‘ECM 기반 화면내 예측 후보 리스트 구성을 위한 OBIC 후보 확장 기법’은 ECM의 화면내 예측 과정에서 사용되는 Most Probable Mode(MPM) 리스트의 후보 다양성을 확장하는 기법을 제안하였다. OBIC(Occurrence-Based Intra Coding) 모드를 MPM 후보로 추가하고, 비인접 후보 탐색과의 중복성을 고려하여 탐색 범위를 조정함으로써 불필요한 후보를 제거하여 예측 모드 선택 효율을 개선하고 부호화 성능 향상을 달성하였다. ‘ECM 기반 색차 화면내 예측을 위한 Directional Planar 적용 방법’은 ECM의 화면내 예측에서 기존 Planar 예측 모드의 한계를 분석하고, 휘도 성분에서 사용되던 Directional Planar 예측을 색차 성분으로 확장하는 방법을 제안하였다. 제안 기법은 색차 블록 내 방향성과 구조적 특성을 반영함으로써 경계 및 텍스처 영역에서 예측 정확도를 향상시키는 결과를 보였다. ‘ECM 화면내 예측 모드 부호화를 위한 경사도 분포 기반 정렬 기법’은 ECM 화면내 예측 과정에서 Non-MPM 리스트의 비효율적인 후보 정렬 문제를 개선하였다. 블록 내 기울기 분포를 Histogram of Gradients(HoG)로 분석하여 방향성 정보를 반영한 재정렬 기법을 제안하였으며, 이를 통해 발생 확률이 높은 예측 모드에 더 적은 비트를 할당하도록 유도하여 부호화 효율 개선 가능성을 확인하였다. 발표를 통해 차세대 비디오 부호화 표준을 위한 ECM 기반 화면내 예측 구조의 고도화와, 새로운 영상 콘텐츠 특성에 대응하는 부호화 기술 동향을 공유하고 실질적인 적용 가능성을 논의하였다.

### 3. 특별 세션 3: 3D 공간 미디어

특별 세션 3에서는 ‘3D 공간 미디어’를 주제로 5편의 논문이 발표되었으며, 경희대학교(1편), 이화여자대학교/ETRI(1편), 한국항공대학교/ETRI(1편), 성균관대학교/연세대학교(1편), 성균관대학교(1편)가 참여하였다. 주요 발표는 3DGS를 기반으로 한 3차원 장면 표현, 압축, 품질 평가 및 응용 기술을 중심으로 이루어졌다.

‘3차원 객체 모델링을 위한 3D Gaussian Splatting 기반 Segmentation Loss 및 Gaussian Pruning 방법’은 3DGS를 활용한 객체 중심 3D 모델링에서 발생하는 가우시안 아티팩트 문제를 개선하였다. 분할 이미지를 기반으로 마스크 손실을 설계하고 학습 이후 가우시안 프루닝을 적용함으로써, 배경과 분리된 객체를 보다 안정적으로 복원할 수 있음을 보였다. ‘가우시안 스플랫 코딩에서 구면조화함수의 채널 간 상관 제거를 위한 변환 기법’은 Gaussian Splat Coding(GSC)에서 색상 표현에 사용되는 구면조화(SH) 계수의 상관성을 제거하기 위한 SH Decorrelation Transform(SHDT)을 제안하였다. SH 계수의 통계적 특성을 반영한 데이터 기반 직교 변환을 통해 압축 효율을 향상시키고, 부호화 오버헤드와 성능 간의 절충 관계를



분석하였다. ‘구면조화(SH) 계수 예측 기반 효율적인 3DGS 압축’은 MPEG GSC 맥락에서 SH 계수 간 상관성을 활용한 예측 부호화 기법을 제안하여, 기존 비디오 기반 부호화 프레임워크의 중복성을 효과적으로 제거하였다. 이를 통해 압축 성능과 부호화·복호화 복잡도를 동시에 개선하였다. ‘LightMarkGS: MLP 임베딩을 통한 3D 가우시안 스플래팅 압축-강건 워터마킹’은 3DGS 환경에서 압축 후에도 워터마크를 안정적으로 유지할 수 있는 워터마킹 구조를 제안하였다. 워터마킹 경로와 압축 경로를 분리하고 MLP 임베딩을 활용함으로써, 고압축 환경에서도 렌더링 품질과 워터마크 복원 성능을 유지하였다. ‘3D Gaussian Splatting 기반 임의 시점 렌더링의 비참조 품질평가’는 기준 영상이 없는 상황에서도 시점 간 일관성과 기하적 안정성을 함께 고려할 수 있는 비참조 품질평가 프레임워크를 제안하였다. 깊이 단서와 공간 특성을 결합한 평가 방식을 통해 다양한 데이터셋에서 신뢰성 있는 품질 추정이 가능함을 보였다. 위와 같은 발표를 통해 3D Gaussian Splatting을 중심으로 한 3D 공간 미디어의 모델링, 압축, 보호 및 품질평가 기술 동향을 공유하고, 실감형 3D 콘텐츠의 실질적인 응용 가능성을 논의하였다.

#### 4. 특별 세션 4: 메타버스 산업을 위한 영상 미디어 압축

특별 세션 4에서는 ‘메타버스 산업을 위한 영상 미디어 압축’을 주제로 4편의 논문이 발표되었으며, 세종대학교(1편), 경희대학교(1편), 한양대학교(1편), (주)아티스(1편)가 참여하였다. 주요 발표는 메타버스 산업을 위한 주요 표준화 기술 공유 중심으로 이루어졌다.

‘Lenslet Video Coding(LVC) 표준화 동향 및 AI 기술 적용 소개’는 차세대 몰입형 미디어 서비스를 위한 MPEG Lenslet Video Coding(LVC) 표준화의 최근 동향과 기술 적용 사례를 소개하였다. ‘JVET 39차 Neural Network based Video Coding 표준 기고내용 논의’는 VVC 이후 차세대 비디오 압축 기술로 주목받고 있는 Neural Network based Video Coding(NNVC)의 표준화 진행 상황과 주요 실험 결과를 소개하였다. ‘MPEG 152차 MPEG-I 그래픽스 표준 주요 내용 분석’은 MPEG WG7에서 추진 중인 3D 그래픽스 및 햅틱스 부호화 표준화 동향을 소개하였다. ‘40차 JVET Enhanced Compression beyond VVC capability 표준 기고내용 논의’는 VVC 이후 차세대 비디오 압축 표준을 위한 CfE 및 EE2 활동 결과를 중심으로 표준화 논의 내용을 소개하였다. 위와 같은 발표를 통해 차세대 비디오 및 몰입형 미디어 표준화 동향과 AI 기반 부호화 기술의 발전 방향을 공유하며, 향후 실질적인 응용 가능성과 표준 확장 방향을 논의하였다.

#### 5. 특별 세션 5: 3D 정보 추출, 표현 및 압축

특별 세션 5에서는 ‘3D 정보 추출, 표현 및 압축’을 주제로 6편의 논문이 발표되었으며, 광운대학교(3편), 서강대학교(2편), 건국대학교(1편)가 참여하였다. 주요 발표는 3차원 정보의 추출부터 압축까지의 처리 과정 중 핵심 기술을 중심으로 이루어졌다.

‘의료 안면 정합을 위한 표면 재구성 성능 평가’는 CT 영상이 제공하는 구조 정보와 이미지 기반 3D 안면 복원 데이터의 색상·텍스처 정보를 정합하여 시각적 사실성과 기하 정확성을 동시에 확보하

는 통합 안면 모델링 방법을 제안하였다. Scale ICP와 Nonrigid Registration을 적용하고 Chamfer Distance, Point-to-Point, Normal Consistency 지표를 통해 임상적 활용 가능성을 검증하였다. ‘Skeleton 기반 행동 인식: 최근 동향에 대한 종합적 고찰’은 Skeleton 기반 행동 인식 분야의 연구 흐름을 지도학습과 자기 지도학습 관점에서 정리하였다. 그래프 신경망 기반 모델링과 대조학습 및 복원 기반 사전학습 기법을 중심으로 표현 학습의 발전 양상을 분석하고, 보편적 Skeleton 백본의 필요성을 향후 연구 방향으로 제시하였다. ‘다시점 역투영 오차 보정을 통한 고정밀 SMPL 생성’은 다중 시점 영상에서 관절 신뢰도를 활용해 시점별 오차를 선택적으로 보정하는 SMPL 생성 방법을 제안하였다. 점진적 관절 갱신을 통해 기존 방식 대비 관절 위치 오차를 효과적으로 감소시켜 고정밀 인체 모델 복원이 가능함을 보였다. ‘효율적인 3차원 데이터 표현을 위한 Point Cloud와 3D Gaussian Splatting의 표현 및 압축 성능 분석’은 Point Cloud로부터 3DGS 데이터를 생성하고, G-PCC 압축을 적용해 두 표현 방식의 시각적 품질과 데이터 크기를 비교 분석하였다. 실험 결과, 동일한 비트스트림 조건에서 3DGS가 Point Cloud 대비 우수한 렌더링 품질을 제공하였다. ‘IF-MoDGS: Initial-Free Monocular Dynamic Gaussian Splatting’은 단일 시점 동적 장면을 복원할 수 있는 3DGS 기반 렌더링 기법을 제안하였다. 정적 배경 기반 포즈 추정과 동적 객체 변형 모델링을 결합해, 단일 시점 동적 장면에서 우수한 렌더링 성능을 달성하였다. ‘객체 신원 유지 기반 다중시점 특징 혼합을 통한 다중객체 3차원 메쉬 복원’은 다중 시점에서 추출된 특징을 객체 신원 유지 알고리즘으로 정합하고, 특징 혼합 모듈을 통해 가려진 환경에서도 강인한 3차원 메쉬 복원을 가능하게 하였다. 위와 같은 발표를 통해 3차원 재구성 및 렌더링 기술 전반의 최신 연구 동향을 공유하고, 다양한 실감형 미디어 분야에서의 실질적인 응용 가능성을 논의하였다.

## 6. 특별 세션 6: 생성형 콘텐츠 및 보호

특별 세션 6에서는 ‘생성형 콘텐츠 및 보호’를 주제로 총 4편의 논문이 발표되었다. 한국항공대학교(1편), 서울시립대학교(1편), UNIST(1편), 중앙대학교(1편)가 참여하였다. 발표는 생성형 AI의 발전으로 인해 발생하는 위협을 방어하기 위한 영상 및 음성 탐지 기술과 차세대 실감 미디어 구현을 위한 3D 렌더링 기술의 확장 연구를 중심으로 이루어졌다.

‘AI 생성 영상의 일반화된 검출 방식’은 학습 과정에서 보지 못한 새로운 생성 모델로 만들어진 영상까지 효과적으로 탐지하기 위한 기술을 제안하였다. 연속된 영상 프레임 간의 미세한 시간적 불일치와 저수준 디테일 차이를 증폭시켜 학습함으로써, 특정 생성 모델에 과적합되지 않고 다양한 AI 생성 영상에 대해 높은 일반화 성능을 보이는 이상치 탐지 기법을 소개하였다. ‘진짜 음성 쌍 학습을 통한 오디오 딥페이크 탐지 강화’는 날로 정교해지는 음성 합성 공격에 대응하기 위해, 진짜 음성 간의 유사성을 강화하는 진짜 음성 쌍 학습(Bonafide-Pair Learning, BPL) 방법론을 제시하였다. 동일 화자의 진짜 음성에서 파생된 증강 샘플들을 쌍으로 구성하여 임베딩 거리를 좁히도록 학습시킴으로써, 진짜 음성 클래스 내의 응집도를 높이고 위조 음성과의 경계를 명확히 하여 탐지 성능을 획기적으로 개선하였다. ‘Bio-signal Based Detection of AI-Generated Videos’는 생성형 AI가 인간의 외형은 모방할 수 있어도 미세한 생리적 신호

까지는 재현하지 못한다는 점에 착안하였다. 얼굴 영상에서 원격 광혈류측정(rPPG) 기술을 이용해 심박 등 생체 신호를 추출하고, 이를 분석하여 AI가 생성한 영상 특유의 생체 신호 부재나 불규칙성을 탐지해냄으로써, Sora나 Veo와 같은 최신 고품질 생성 모델에 대해서도 높은 정확도로 가짜 영상을 판별할 수 있음을 입증하였다. ‘굴절 효과를 반영한 Gaussian Splatting 기법의 초기 연구’는 기존의 3D Gaussian Splatting 기술이 반사 표현에는 강점이 있으나, 유리나 물과 같은 투명한 물체의 굴절 표현에는 한계가 있음을 지적하고 이를 개선하였다. 깊이 추정 방식을 조정하여 투명 굴절면을 인식하게 하고, 간단한 광선 추적 알고리즘을 결합하여 배경이 왜곡되는 굴절 현상을 시각적으로 구현함으로써, 더욱 사실적인 3D 렌더링을 위한 기술적 토대를 마련하였다. 본 세션은 AI 기술이 가져온 미디어 생성의 혁신과 보안 위협이라는 두 가지 핵심 이슈를 심도 있게 다루었으며, 신뢰할 수 있는 미디어 환경 조성을 위한 방어 기술과 몰입감을 극대화하는 표현 기술의 구체적인 발전 방향을 제시하였다.

## 7. 특별 세션 7, 9: 디지털 휴먼

특별 세션 7, 9에서는 ‘디지털 휴먼’을 주제로 6편의 논문이 발표되었으며, 오모선(주)/광운대학교(1편), 엠앤엔에이치(1편), 서울과학기술대학교(1편), 광운대학교(2편), KBS(1편)가 참여하였다. 주요 발표는 디지털 휴먼의 사실적 생성, 평가 방식 그리고 디지털 휴먼 콘텐츠 제작 및 응용 기술을 중심으로 이루어졌다.

‘사실적 움직임 생성 및 재현 디지털휴먼 기술’은 해부학적 구조와 물리 기반 실사화 요소를 결합하여 사실적이고 자연스러운 디지털 휴먼 움직임을 생성, 편집, 재현하는 기술과 이를 지원하는 통합 제작 플랫폼을 제안하였다. AI 기반 데이터 정합과 물리 법칙 적용을 통해 환경과의 이질감을 최소화하고, 연령, 성별, 신체 특성에 따른 움직임 변형과 디에이징 기술을 구현하여 다양한 인체 표현 가능성을 제시하였다. ‘생성형 AI 기반 디지털휴먼을 활용한 IC-VFX AI 저작도구 개발’은 생성형 AI와 디지털 휴먼 기술을 결합한 IC-VFX 제작 파이프라인을 제안하였다. 극사실 디지털 휴먼, AI 기반 장면 생성, 실시간 렌더링 및 합성 기술을 통합하여 방송 콘텐츠 제작 환경에서의 실질적인 활용 가능성을 보였다. ‘근전도 신호를 통한 자세 분류’는 표면 근전도 신호만을 이용해 전신 자세를 분류하는 방법을 제안하였다. Transformer 기반 특징 추출과 다양한 신호처리 구조를 비교 분석하여 전신 동작 분류의 가능성과 높은 분류 정확도를 실험적으로 검증하였다. ‘디지털 휴먼 사실성 평가 척도의 개발 및 검증’은 디지털 휴먼의 시각적, 정서적, 사회적 특성을 반영한 다차원 사실성 평가 척도를 개발하고 검증하였다. 이를 통해 Eeriness, Likability, Warmth, Trust, Humanlikeness의 다섯 가지 핵심 요인을 도출함으로써 정량적 사실성 평가 기준을 제시하였다. ‘생체 신호를 활용한 디지털 휴먼의 정적 사실성 평가 방법에 대한 문헌 고찰’은 EEG, fMRI, EDA, HRV, EMG, 시선 추적 등 생체 신호 기반 사실성 평가 연구를 체계적으로 분석하였다. 생체 신호가 디지털 휴먼 외형에 대한 무의식적 인식과 정서 반응을 반영할 수 있음을 정리하며, 향후 객관적 평가 지표 개발의 이론적 기반을 제시하였다. ‘디지털 휴먼 캐릭터화 기획’은 디지털 휴먼을 콘텐츠 관점에서 캐릭터로 설계하고 활용하기 위한 기획 방향을 제시하였다. 서사, 정체성, 활용 시나리오를 고려한 캐릭터화 전략을 통해 콘텐츠 산업에서의 확장 가능성을 논의하였다. 위와 같은 발표를 통하여 디지털 휴먼의 생성 기술부터 콘텐츠 제작에 이르는 전반적인 기술 흐름을 공유하고, 방송, XR, 콘텐츠 산업에서의 실질

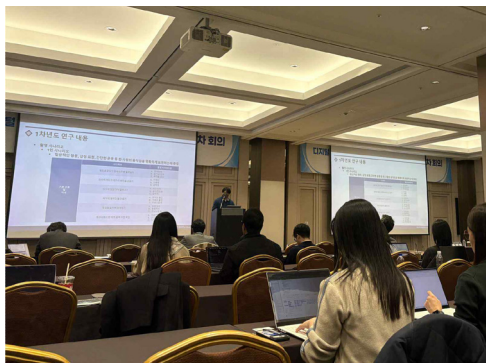


적인 응용 가능성을 논의하였다.

## 8. 특별 세션 8: Audio Coding for Machines

특별 세션 8에서는 ‘Audio Coding for Machines’를 주제로 총 4편의 논문이 발표되었다. 참여 기관으로는 동아대학교(3편), 중앙대학교(1편)가 참여하였다. 주요 발표는 기계를 위한 오디오 신호 부호화 표준화 동향과 더불어 오디오 처리 기술의 발전 방안을 중점적으로 다루었다.

‘스펙트로그램 기반 CNN 분류 모델을 이용한 FCTM 오디오 적용 방법론’은 선택적 학습 전략과 경량화된 신경망 구조를 활용하여 FCTM 인코더-디코더 구조 내에 CNN을 통합하였으며, ESC-50 및 Urban-Sound8K 데이터셋을 활용한 실험에서 DenseNet 모델이 91% 이상의 분류 정확도를 달성하여, FCTM 기반 특징 압축과 CNN 분류를 효과적으로 결합하는 방법을 입증하였다. ‘기계를 위한 오디오 신호 부호화 표준화 동향’에서는 멀티미디어 신호를 압축 및 표현하는 표준 기술을 개발하는 MPEG 오디오 부호화 그룹의 기계를 위한 오디오 압축 기술에 대한 표준화 동향을 다뤘다. ‘AuralSynth: Metadata-Free Sound Anomaly Synthesis for Industrial Monitoring’은 산업 시설 모니터링을 위한 기계용 음성 이상 탐지의 훈련 데이터 부족 문제를 해결하기 위해 확산 모델 기반 정상 음성 합성과 그라디언트 상승 기반 이상 음성 생성을 결합한 프레임워크를 제안하였다. 제안된 방식은 기존 대규모 모델 대비 파라미터를 99.7% 감소시킨 경량 분류기를 사용하면서 경쟁력 있는 성능을 달성하여, 메타데이터 없이 효과적인 이상음 탐지가 가능함을 입증하였다. ‘텍스트-음향 변환을 이용한 이상 음향 탐지 데이터 증강’은 AudioLDM 등 텍스트-오디오 생성 모델을 활용하여 도메인 시프트 환경에서의 이상음 탐지 성능 향상 방안을 제시하였다. ToyADMOS2 데이터셋의 ToyCar 클래스를 대상으로 한 실험에서 TTA 기반 합성 음성 데이터가 이상음 탐지 성능을 향상시킬 수 있음을 보여주었으며, LoRA(Low-Rank Adaptation)와 VAE를 활용한 효율적인 모델 적용 방안을 제시하였다. 본 세션은 MPEG의 ACOM(Audio Coding for Machines) 표준화 과정에서 제시된 기술 요구사항을 반영하여, 경량화된 신경망 모델, 메타데이터 없는 시스템 설계, 생성형 AI 기반 데이터 합성을 통해 산업용 음성 모니터링 시스템의 효율성과 실용성을 높이기 위한 다양한 접근 방안을 제시하였다.



## 일반 세션

일반 세션에서는 ‘지능형 미디어 신호처리 및 산업 응용’이라는 주제에 대하여 2편의 논문이 발표되었다.

### 1. 일반 세션 1: 지능형 미디어 신호처리 및 산업 응용

‘지능형 미디어 신호처리 및 산업 응용’ 세션에서는 총 2편의 논문이 구두 발표되었는데, 아주대학교(1편), 경북대학교(1편)에서 발표하였다. 발표는 Vision Transformer(ViT)의 학습 최적화를 위한 신호처리 방법론과 산업 현장의 안전 제고를 위한 지능형 예측 시스템에 대한 연구들로 구성되었다.

‘의미론적 앵커 토큰: 메타데이터 가이드를 통한 Vision Transformer의 숏컷 학습 억제’는 최신 ViT가 이미지와 함께 제공되는 메타데이터를 활용하지 못하고 허위 상관관계에 의존하는 숏컷 학습(Shortcut Learning) 문제에 취약하다는 점에 주목하였다. 이를 해결하기 위해 메타데이터 임베딩을 추가적인 클래스 토큰으로 주입하는 의미론적 앵커 토큰(SAT) 방법론을 제안하여, 모델이 객체의 전체적인 특징을 학습하도록 유도하였다. ‘플랜트 건설현장에서의 공구 부피 추정 및 작업자의 근골격계 부하 예측 모델 타당성 평가’는 플랜트 건설 현장의 고위험 환경에서 근골격계 안전 관리 문제를 해결하고 안전 투자의 재무적 타당성을 입증하고자 하였다. RGB-D 이미지 처리를 이용한 공구 부피 추정과 YOLOv11 Pose 모델 기반의 키포인트 검출을 통해 작업자의 관절 각도와 근골격계 부하를 정량적으로 예측하는 시스템을 개발하여 안전 관리가 단순 비용이 아닌 기업의 핵심 전략임을 입증하였다. 위와 같은 발표를 통하여 미디어 데이터의 효율적인 처리 기술과 이를 산업 현장에 적용하여 실질적인 가치를 창출하는 융합 연구의 가능성을 논의하였다.



## 포스터 세션

포스터 세션은 총 5개로 구성되었으며, 이 중 3개는 일반 포스터 세션, 1개는 대학생 논문 경진대회, 1개는 대학생 논문 및 캡스톤디자인 경진대회로 이루어졌다. 각 세션에서는 다양한 주제와 연구 결과를 발표하며 학술 교류의 장을 마련하였다.

### 1. 포스터 세션 1

포스터 세션 1에서는 총 11편의 논문이 발표되었다. 참여 기관으로는 KETI(1편), 동아대학교(2편), 광주대학교(1편), 한양대학교(1편), 세종대학교(4편), 한화시스템(1편), 동아대학교/ETRI(1편)가 포함되었다.

‘Android 테더드 AR에서의 Mediapipe 랜드마크 기반 TFLite 손동작 분류 및 실행 연구’는 UVC 카메라 영상을 Unity에서 처리하여 Mediapipe Hands 그래프에 실시간으로 입력하고, 추출된 21개의 랜드마크로 학습한 경량 TFLite 분류기를 통해 사용자 정의 손동작을 인식하고 실행하는 파이프라인을 설계 및 구현하였다. 이 연구는 커스텀 ImageSource(UVCSource) 구현과 GestureAPI를 통해 UVC-Unity-Mediapipe 간 호환성 문제를 해결하였으며, 코드 수정 없이 Inspector 설정만으로 다양한 상호작용을 구성할 수 있는 사용자 중심의 직관적이고 몰입감 있는 AR 경험 설계 프레임워크를 제공하였다. ‘EdgeSplat: 엣지 디바이스를 위한 경량 3D Gaussian Splatting 최적화 프레임워크’는 고성능 GPU 환경의 3DGS 기반 장면 재구성 기법을 경량 엣지 디바이스에서 효율적으로 동작하도록 최적화하는 모듈형 프레임워크를 제안하였다. 연구는 연산량과 메모리 접근을 줄이고 품질 손실을 완화하며, 경량 번들조정과 타일링 기반 렌더링 파이프라인으로 최적화 복잡도를 감소시켰다. ‘잔차 연결 채널 트랜스포머 기반의 초해상화’는 영상 해상도 저하 문제를 해결하기 위해 잔차 연결 구조를 기반으로 깊이별 분리 합성곱과 채널 트랜스포머를 활용한 RCCT 기반의 초해상화 네트워크를 제안하였다. 연구에서는 깊이별 분리 합성곱으로 연산량을 감소시키고 잔차 연결 구조로 배치하여 효율성을 유지하면서 특징 정보를 효과적으로 활용하는 방법을 소개하였다. 제시된 논문들 외에도 3D 얼굴 복원, 초해상화 기반 객체 탐지, 3D 씬 생성 및 경량 배포, 양자화 및 모델 압축 등 다양한 첨단 기술 연구가 다뤄졌다. 발표된 논문들은 차세대 미디어 기술의 핵심 영역에서 이루어진 선도적 연구들을 보여주었으며, 향후 스마트 미디어 생태계와 차세대 방송·통신 기술의 발전에 핵심적인 역할을 할 것으로 기대된다.

### 2. 포스터 세션 2

포스터 세션 2에서는 총 9편의 논문이 발표되었다. 참여 기관으로는 동아대학교/ETRI(1편), 한양대학교(2편), 시그마케이(주)/서울과학기술대학교(1편), 서울과학기술대학교(1편), 서울대학교(1편), 홍익대

교/㈜아이닉스(1편), KETI(2편)가 포함되었다.

‘항공 영상 복원을 위한 비지도 기반 그림자 제거 및 초해상화 통합 모델’은 고해상도 항공 영상의 그림자와 저해상도 문제를 동시에 해결하기 위해 비지도 학습 기반의 통합 프레임워크를 제안하였다. Mirror Padding과 Overlapping 기법을 적용하여 패치 처리 시 경계 부근의 아티팩트를 효과적으로 완화하였으며, 순차 처리 방식 대비 처리 속도를 12.8% 단축하면서도 복원 품질을 유지하여 실용적인 항공 영상 분석 솔루션을 제시하였다. ‘초고해상도 영상복원 및 복잡도 분할 모델을 결합한 영상 워터마킹 시스템’은 생성형 AI 이미지 진위 여부 판별을 위해 초고해상도 복원, 엔트로피 기반 영역 복잡도 분할, 블렌딩 기반 워터마킹을 통합한 시스템을 제안하였다. 연구에서는 복잡한 영역에는 상대적으로 강하게, 단순한 영역에서는 약하게 워터마크를 블렌딩하여 시각적 비가시성과 높은 검출률을 동시에 달성하였다. ‘단어 경계 기반 가변 블록을 통한 스트리밍 FastSpeech2 개선’은 대규모 언어 모델과 결합되는 실시간 음성 합성 시스템을 위해 단어 경계를 고려한 가변 블록 구조를 제안하였다. 제안된 방법은 기존 스트리밍 합성 방식 대비 품질과 지연 측면에서 모두 개선된 성능을 보였으며, CPU 환경에서도 1초 이하의 낮은 첫 블록 지연을 달성하여 실제 TTS 시스템 구현에 효과적임을 입증하였다. 음성 합성, 영상 복원 외에도 초해상화 통합 모델, 비전-언어 모델 성능 평가, FCTM 추론 효율 최적화 등과 같이 실시간 처리, 도메인 적응성, 경량화 및 효율성에 대한 발표가 이루어졌다.

### 3. 포스터 세션 3

포스터 세션 3에서는 총 9편의 논문이 발표되었다. 참여 기관으로는 국립창원대학교(3편), 광운대학교/ETRI(1편), 서울과학기술대학교/시그마케이(㈜)(1편), 서울과학기술대학교(2편), 덕성여자대학교/원광대학교/경희대학교(1편), 중앙대학교(1편)가 포함되었다.

‘생성형 AI 단편영화 제작에서의 인물, 장면, 음향 일관성 및 한국어 표현 기법에 관한 실험적 연구’는 생성형 AI 도구를 이용한 영상 제작 시 발생하는 기술적 한계를 분석하고 해결 방안을 제시하였다. 동일 장면에서의 인물 복장, 조명, 구도, 샷 사이즈 불일치, 한국어 표현의 부자연성 등을 실험적으로 분석하였으며, Adobe Photoshop, Illustrator, After Effects 등 전통적인 편집 도구와 참조 기능, 프롬프트 엔지니어링을 병행하여 해결하였다. ‘생성형 AI를 활용한 한국적 브랜딩 아이덴티티 연구: 로고 디자인에 미치는 키워드의 영향’은 생성형 AI 기술이 서구적 미학에 편중되는 한계를 극복하기 위해 프롬프트에 한국적, 동양적 키워드를 전략적으로 활용하는 방법을 제안하였다. DALL·E를 활용하여 가상의 차 브랜드 ‘선차’의 로고를 디자인하였으며, 기본 프롬프트와 한국적 키워드가 추가된 확장 프롬프트의 효과를 설문 조사를 통해 비교하였다. 연구에서는 문화 키워드가 생성된 이미지에 한국적 정체성을 효과적으로 부여함을 확인하였으며, 단순 키워드 추가가 아닌 기본 스타일과의 상호작용을 예측한 정교한 프롬프트 설계의 중요성을 제시하였다. ‘부분 조작 기반 설명 가능한 얼굴 합성 탐지 데이터셋’은 생성형 AI 시대에 딥페이크 탐지의 정확도와 해석 가능성을 동시에 향상시키기 위해 XFace 데이터셋을 제안하였다. CelebA 데이터셋을 기반으로 주요 안면 부위의 부분적 조작에 초점을 맞춰 총 144,000개의 조작 이미지를 구축

하였으며, 각 이미지에 자연어 설명을 함께 제공하여 시각적 탐지와 언어적 해석을 결합한 멀티모달 학습을 가능하게 하였다. 이 외에도 군중 영상 분석, 로봇 제어 시스템, 센싱 및 통신 기술, 네트워크 보안, 가상현실 학습 환경 등 폭넓은 주제의 연구들이 논의되었다.

#### 4. 포스터 세션 4: 대학생 논문 경진대회

포스터 세션 4는 ‘대학생 논문 경진대회’로서 총 16편의 논문이 출품되었다. 참여 기관으로는 명지대학교(1편), 동아대학교(1편), 홍익대학교(2편), 서울과학기술대학교(4편), 국립창원대학교(1편), 연세대학교(1편), 서울시립대학교(1편), 성신여자대학교(3편), 동아방송예술대학교/마젠타컴퍼니(1편), 광운대학교(1편)가 포함되었다.

학부생들의 연구 활동을 고취하고, 향후 석사 및 박사 과정으로의 심화 연구로 자연스럽게 유도하기 위해, 학부생 연구원으로서 소속 대학의 연구실에서 수행하여 도출한 연구 성과들을 발표할 수 있는 세션이었다.

이 세션에서는 대학생 연구자들이 창의적인 아이디어를 바탕으로 다양한 연구를 발표하였다. ‘분산식별자를 활용한 티켓 취소 및 검증 시스템’은 공연 티켓의 위조 방지 및 신뢰성 강화를 위해 분산식별자 기술을 활용했다. 검증 가능한 자격증명 형태로 티켓을 발급하여 사용자의 지갑에 안전하게 저장하고, 검증 시 상태 리스트 기반 취소 모델을 통해 환불과 재발급 등 운영 시나리오에 효율적으로 대응했다. ‘생성형 AI 모델들을 사용한 전과정 영상제작 워크플로우 설계 및 평가’는 ChatGPT, Stable Diffusion XL, WAN2.1 VACE 모델을 결합하여 시나리오 생성부터 영상 생성까지 전 단계를 자동화했다. 오픈소스 모델의 한계를 DALL·E 3와의 비교를 통해 검증하였으며, 자동 생성 프롬프트 기반 영상이 기본적인 시각적 일관성은 확보하되 동작의 자연스러움과 클립 연결성 개선이 필요함을 보여주었다. 이 외에도 멀티모달 AI를 활용한 원격감사 영상 분석, 픽셀아트 생성을 위한 확산모델 개선, 음성 인식 모델의 속도 변형 기반 데이터 증강, HRTF 기반 시각장애인용 청각화 시스템 등 접근성 및 포용성을 강화한 다양한 연구가 포함되었다. 참가 학부생 연구자들이 창의적인 아이디어를 바탕으로 AI, 컴퓨터 비전 등 첨단 기술을 활용하여 실제 문제 해결을 지향하는 연구들을 수행함으로써, 향후 학술적 성과를 통해 관련 분야에 기여할 수 있는 높은 잠재력을 보여주었다.

#### 5. 포스터 세션 5: 대학생 논문 및 캡스톤디자인 경진대회

포스터 세션 5에서는 ‘대학생 논문 경진대회’ 논문 6편, ‘캡스톤디자인 경진대회’ 논문 7편, 총 13편의 논문이 출품되었다. 참여 기관으로는 서울과학기술대학교(7편), 홍익대학교(1편), 경북대학교/KETI(2편), 국립한밭대학교(3편)가 포함되었다.

‘데이터 크기와 GCC 컴파일러 최적화 옵션에 따른 Quick Sort 성능 비교’에서는 GCC 컴파일러 최적화 옵션이 Quick Sort 성능에 미치는 영향을 분석하였다. 실험 결과, -O3 옵션의 평균 실행 시간이 default



대비 최대 72.2% 감소하였으며, 데이터 크기가 증가함에 따라 감소율도 증가하는 경향을 보였다. 이는 컴파일 최적화 옵션 선택이 동일 알고리즘의 성능을 크게 좌우할 수 있음을 입증하였다. ‘키워드 기반 음악 추천 시스템: 의미 및 맥락 인식 유사도 접근법의 사용자 중심 평가’는 사용자의 현재 상태를 반영하는 키워드 기반 추천 시스템을 제안하였다. 세부 단어 단위 유사도와 문맥 의미 유사도를 60%, 40%로 가중 결합한 결과, 개별 곡 평가에서는 기존 시스템 대비 최대 66.3%, 리스트 평가에서는 64.4%의 높은 만족도를 기록하였다.

이 외에도 AI 생성 텍스트 판별, 3D 비전 및 압축 기술, 개인 정보 보호, AIoT 스마트 케어 시스템 등과 같은 다양한 분야의 실질적인 문제를 해결하기 위한 창의적인 접근과 기술적 도전을 통해 독창적이고 실용적인 연구 성과를 보여주었다.





## 맺음말

이번 한국방송·미디어공학회 2025년 추계학술대회는 방송·미디어 산업의 기술적 패러다임이 보는 미디어에서 경험하는 미디어로 확장되고 있음을 확인한 자리였으며, 성공적인 개최를 통해 몇 가지 중요한 의미를 남겼다. 이번 학술대회는 미디어 AX, 생성형 AI, 디지털 휴먼, 3D 공간 미디어 등 혁신적인 기술들이 실제 산업 현장과 연구 개발 단계에서 어떻게 구체화되고 있는지를 심도 있게 다루었다. 특히 단순한 콘텐츠 제작을 넘어 AI 기반의 자동화 파이프라인과 기계를 위한 부호화 기술 등 미래 미디어 생태계의 효율성을 극대화하는 논의가 이루어졌다는 점에서, 이번 대회는 관련 연구 성과를 공유하고 미래 비전을 제시하는 핵심 플랫폼으로서의 역할을 충실히 수행하였다.

이를 통해 학계와 산업계는 미디어 데이터의 분석과 생성, 그리고 전송에 이르는 전 과정에서의 협력 방안을 모색할 수 있었다. 특히 이번 학술대회에서는 대학생 논문 및 캡스톤디자인 경진대회가 함께 열려, 차세대 미디어 주역들이 창의적인 연구 성과를 발표하고 분야별 전문가들과 직접 소통하며 성장할 수 있는 기회를 제공하였다. 이는 미래 미디어 기술을 선도할 젊은 인재들이 실질적인 연구 경험을 쌓고, 학문적 호기심을 산업적 통찰로 발전시키는 중요한 발판이 되었다.

마지막으로 이번 대회에서 발표된 기술들은 방송과 미디어 산업의 미래를 관통하는 핵심 키워드들이다. 미디어 AX 가속화, 홈쇼핑 및 영상 제작에서의 생성형 AI 활용, 그리고 디지털 휴먼의 사실성 구현 기술 등은 기존 미디어 환경의 한계를 뛰어넘는 새로운 가능성을 보여주었다. 또한 ‘미디어 테크 오픈랩’과 같은 네트워킹 프로그램을 통해 산·학·연이 머리를 맞대고 기술 교류를 활성화함으로써, 앞으로 방송·미디어 기술이 나아갈 방향을 구체적으로 정립하고 새로운 비즈니스 모델 창출을 위한 전략적 토대를 마련하였다. 이러한 성과들은 향후 지능형 미디어 서비스와 같은 몰입형 콘텐츠 시대를 이끌어갈 강력한 원동력이 될 것이다.



단체 사진. 앞줄 좌측부터 김용한(서울시립대), 심동규(광운대), 이수인(ETRI), 김재곤(한국항공대), 윤경로(건국대), 김규현(경희대), 박재홍(한양대), 박상일(서울과기대), 신재섭(취픽스트리), 서정일(동아대)