

연 구 소
소 개

한국전자기술연구원 지능형영상처리연구센터 주요 연구 소개

김용환, 송혁, 장성준, 김성제, 김제우 / 한국전자기술연구원 지능형영상처리연구센터

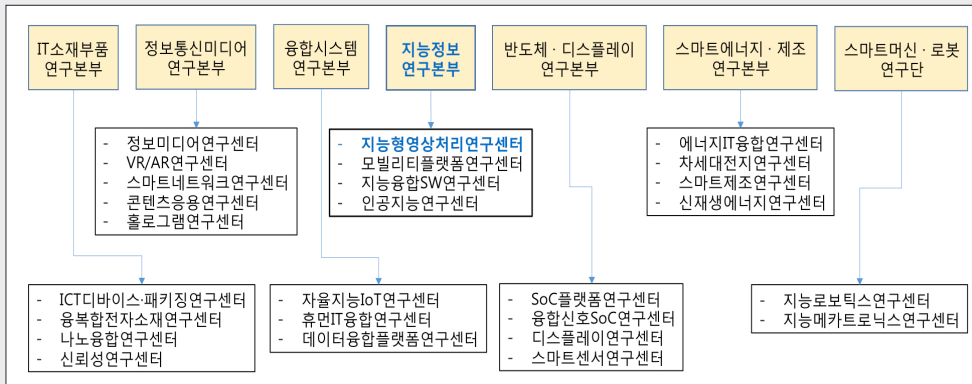


<그림 1> KETI 본원 전경

한국전자기술연구원(Korea Electronics Technology Institute; KETI)은 산업통상자원부 산하 전자 IT분야 전문생산연구기관으로써 우리나라 중소·중견기업이 필요로 하는 핵심기술 연구/개발 및 지원을 목적으로 1991년에 설립되었다. 현재 약 1000여 명의 연구/개발 인력을 보유하고 있으며, 본원은 경기도 성남시 분당구 야탑동에 위치하고, 수도권권을 포함하여 전국에 6개의 분원을 두고 있다.

KETI는 <그림 2>와 같이 7개의 연구본부 내 총 26개의 연구센터에서 전자 IT의 다양한 분야(융복합 전자소재, VR/AR, 홀로그래프, 인공지능, IoT¹⁾, SoC²⁾, 차세대전지, 지능메카트로닉스 등)에 대한 연구/개발을 수행하고 있다.

1) Internet of Things
2) System on Chip



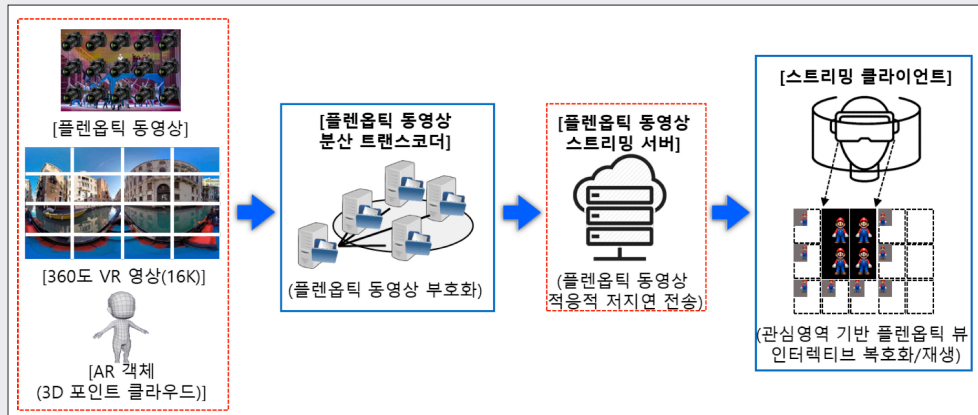
<그림 2> KETI 연구분부와 연구센터

본 소개에서는 지능정보연구본부 내에서 영상처리 관련 연구/개발을 주로 수행하는 지능형영상처리 연구센터의 다양한 연구/개발 주제 및 내용을 소개한다. 지능형영상처리연구센터는 다양한 영상센서로부터 획득된 영상 정보(오디오 포함)를 기반으로 차세대 융합·미디어 서비스를 위한 핵심 알고리즘을 연구하여 HW IP(Intellectual Property) 및 SW 모듈을 연구·개발하는 목적으로 2008년 12월에 ‘멀티미디어IP연구센터’로 출발하여, 2015년에 현재의 ‘지능형영상처리연구센터’로 센터명을 변경하였다. 현재 3개의 팀(멀티미디어팀, 시각지능IP팀, 컴퓨터비전팀)으로 구성된 약 40여 명의 연구원이 연구·개발을 수행하고 있으며, 주 연구 분야는 방송/콘텐츠, 영상 보안, 자율주행차, 로봇, 패션, 의료, 스마트 시티 등의 분야에 활용되는 딥러닝 기반 영상신호처리 기술을 연구하고 있다. 최근에 수행완료했거나 현재 수행하고 있는 주요 연구/개발 주제 및 내용은 다음과 같다.

1. 몰입형 비디오 코딩(Immersive Video Coding) 기술

기존 2D 형식의 2K 및 4K 영상보다 몰입감이 높은 새로운 서비스를 원하는 산업계의 니즈에 맞춰, 몇 년 전부터 ISO/IEC MPEG³⁾-I 및 JPEG⁴⁾-Pleno에서는 몰입형 비디오/이미지 코딩 표준화를 진행 중이고, 최근에 일부 표준화를 완료하였다. 대표적인 몰입형 비디오/이미지는 6-DoF⁵⁾/3-DoF+를 지원하는 3D 포인트 클라우드, 라이트필드(플렌옵틱) 및 홀로그램과, 3-DoF를 지원하는 360도 VR⁶⁾ 등이 있다. 세계적인 연구/개발 동향에 발맞춰, 차세대 몰입형 비디오 서비스 시스템(<그림 3>)에서 몰입형 비디오 압축/전송/재생을 위한 요소 기술들과 시스템을 개발하고 있다. 특히, 초대용량 데이터 특성을 갖는 몰입형 비디오의 스트리밍 서비스를 위해서 1)고효율 부호화 기술, 2)관심영역 기반 적응적 전송 기술, 3)관심영역 기반 적응적 복호화 기술의 연구/개발을 수행 중이다.

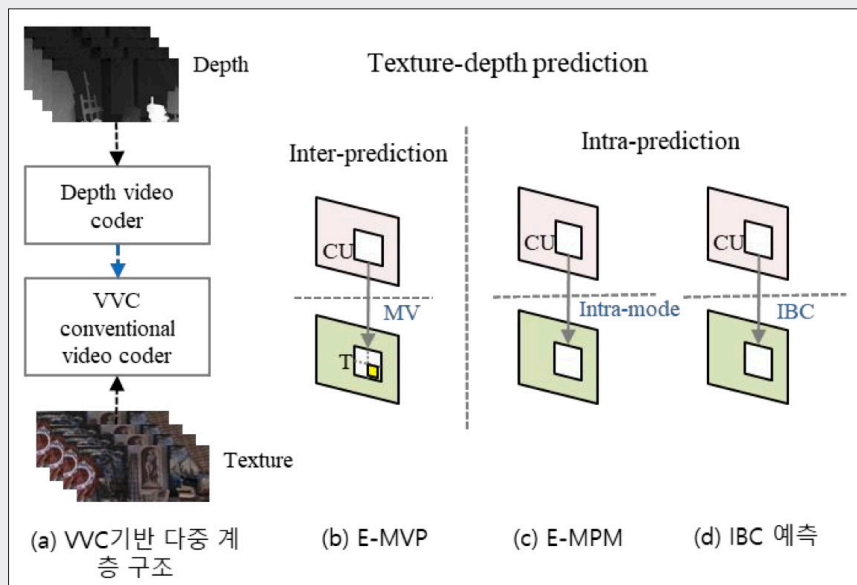
3) Moving Picture Experts Group
 4) Joint Photographic Experts Group
 5) Degree of Freedom
 6) Virtual Reality



<그림 3> 몰입형 비디오 적응적 스트리밍 시스템

▶ 플렌옵틱 동영상 부호화/복호화 기술

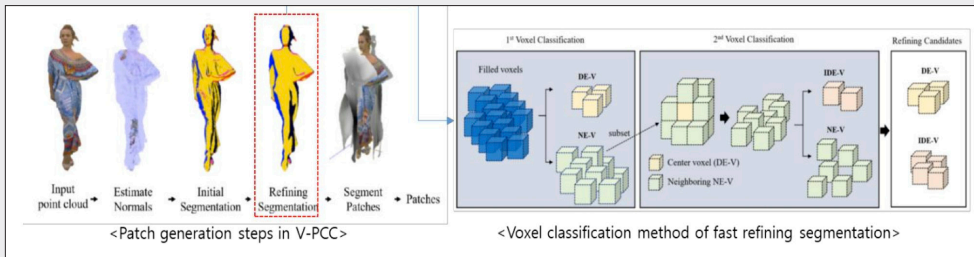
- 관련 표준 : ISO/IEC MPEG-I Part 12 MIV(MPEG Immersive Video)
ISO/IEC MPEG-I Part 3 VVC(Versatile Video Coding)
- 깊이 영상 부호화 효율 향상 알고리즘 개발 중(그림 4)
- Hadoop/Spark Streaming 기반 몰입형 비디오 분산 인코딩 시스템 개발 중
- 초대용량 플렌옵틱 동영상 분산 디코딩 시스템 개발 중



< 그림 4> 컬러 영상 기반 깊이 영상 예측 부호화 방법

▶ 3D 포인트 클라우드 동영상 부호화 기술

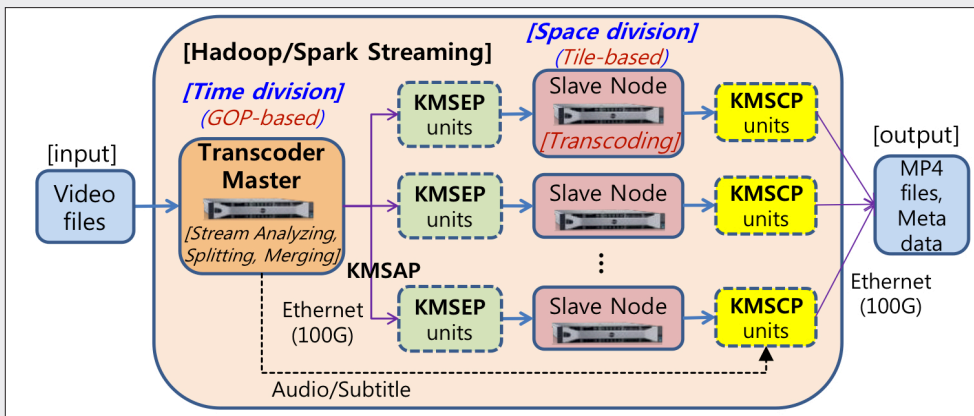
- 관련 표준: ISO/IEC MPEG-I V3C(Visual Volumetric Video-based Coding) and V-PCC (Video-based Point Cloud Compression)
- 고속 V-PCC 부호화 알고리즘 개발 중
- 고속 패치 생성 알고리즘 개발*(**그림 5**)
- 실시간 V-PCC 복호화기 개발 중



<그림 5> 복셀 분류를 통한 고속 refining segmentation 방법

▶ 360도 VR 영상 동영상 부호화 기술

- 관련 표준: ISO/IEC MPEG-H HEVC(High Efficiency Video Coding) AOMedia AV1
- 타일 스트리밍을 위한 움직임 제한(Motion constrained) 부호화 기술 개발
- Hadoop/Spark Streaming 기반 실시간 16K급 360도 VR 영상 분산 인코딩 시스템 개발(**그림 6**)
- 실시간 HEVC/AV1 복호화기 SW 개발
 - 프레임 기반 병렬 디코딩 구조 개발
 - 다수 모듈의 SIMD(Single Instruction Multiple Data) 구조 개발 및 구현



<그림 6> 16K급 360도 VR 영상 분산 부호화 시스템

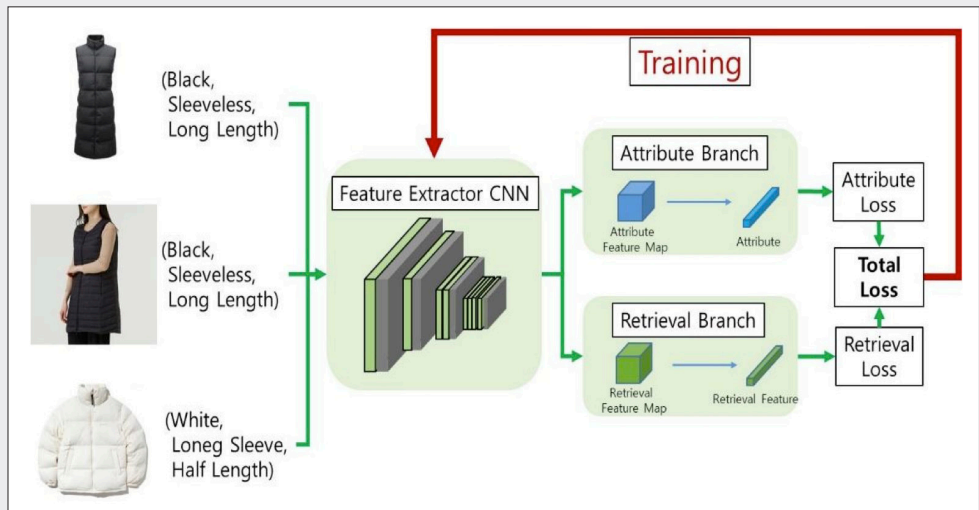
7) Jieon Kim and Y.-H. Kim, ※Fast grid-based refining segmentation method in video-based point cloud compression, IEEE Access, Vol. 9, June 2021
 Jieon Kim and Y.-H. Kim, ※ V-PCC Fast grid-based refining segmentation method in V-PCC, ISO/IEC JTC1/SC29/WG7 Doc. m56635, Online, April 2021

2. 패션 객체 속성 분석 및 검색 기술

온라인 커머스 시장에서 가장 큰 비중을 차지하는 것이 패션 분야이며, 최근에는 패션 분야에 인공지능 기술 적용이 시작되고 있다. 온라인 패션시장에서 가장 큰 문제인 30%에 달하는 재고 폐기 비용문제를 줄이는 방안으로 인공지능 기술을 활용한 유통 기술, 제조 기술, 구매추천 기술, 검색 기술들을 개발하고 있다.

▶ 인공지능 기반 스타일 검색 기술 개발(〈그림 7〉)

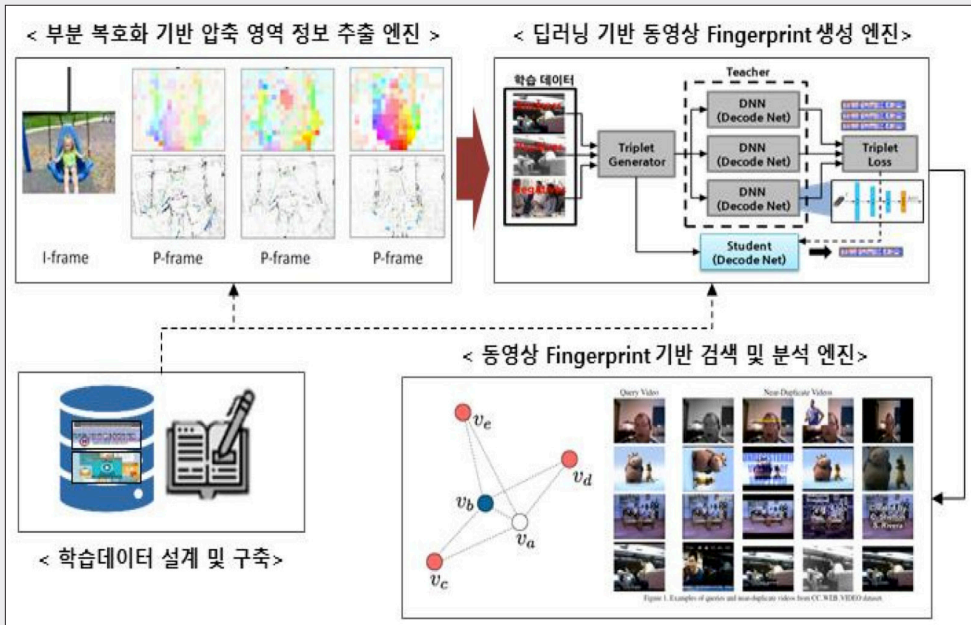
- 패션 스타일 추출 모델 개발
- TPO(Time, Place, Occasion) 추출 모델 개발
- 연령대 추출 모델 개발
- 고속 검색 기술 개발



<그림 7> 속성(Attribute) 기반 검색(Retrieval) 모델 구조

3. 콘텐츠 이해 기술

기존 콘텐츠 이해 기술은 객체의 특정 색상 또는 행동 검출 위주였으나 인공지능 기술의 발전과 함께 콘텐츠 내에서 이뤄지는 상황 및 환경 등을 이해하는 기술이 개발되고 있다. 미디어의 중심이 동영상이 되면서 동영상 내용 분석 및 검색 관련 기술 필요성이 증가하고 있다.



<그림 8> 동영상 부분 복호화 및 Fingerprint 생성기반의 디지털 성범죄 검색 기술

- ▶ 동영상의 Fingerprint 생성을 통해 디지털 성범죄 피해 영상 조기 확산 방지를 위한 고속 검색 기술 개발(<그림 8>)
 - 동영상 부분 복호화 기반 압축 영역 정보 추출
 - 주요 장면 및 압축 정보 기반 동영상 Fingerprint 생성
 - 동영상 Fingerprint 기반 검색 및 동영상 관련 정보 출력
 - 피해자 제공 정보 기반 디지털 성범죄 동영상 검색

4. 영상 기반 보안 기술

영상 분석 기술에 기반하여 이상행위를 탐지하고 정보를 전송하는 기술은 인공지능 기술 이전부터 꾸준히 연구되어왔으나 다양한 상황에서의 신뢰성 향상을 위하여 인공지능 기술의 적용 등으로 지속적으로 기술개발이 이뤄지고 있다. 카메라의 성능이 향상되고 있어 카메라 자체에서 인공지능 기술을 적용한 분석이 가능해지고 있으며, 서버와의 연계를 통하여 성능 향상 및 상황에 따른 효율적인 분산처리가 가능하다.

- ▶ 다양한 영상보안환경에 대응 가능한 엣지컴퓨팅 방식의 적응형 AI 영상보안장비 및 하이 프로파일 분석서버 개발(<그림 9>)



<그림 9> 적응형 AI 엣지 디바이스와 분석서버

- NPU⁸⁾ 기반 AI 카메라 개발 및 객체 검출 딥러닝 모델 개발
- NPU 기반 AI NVR⁹⁾ 개발 및 행동인식 딥러닝 모델 개발
- AI FPGA¹⁰⁾ 구현 및 적응형 AI NVR 구현
- 증류(Distillation) 기법을 통한 적응형 AI NVR 고도화

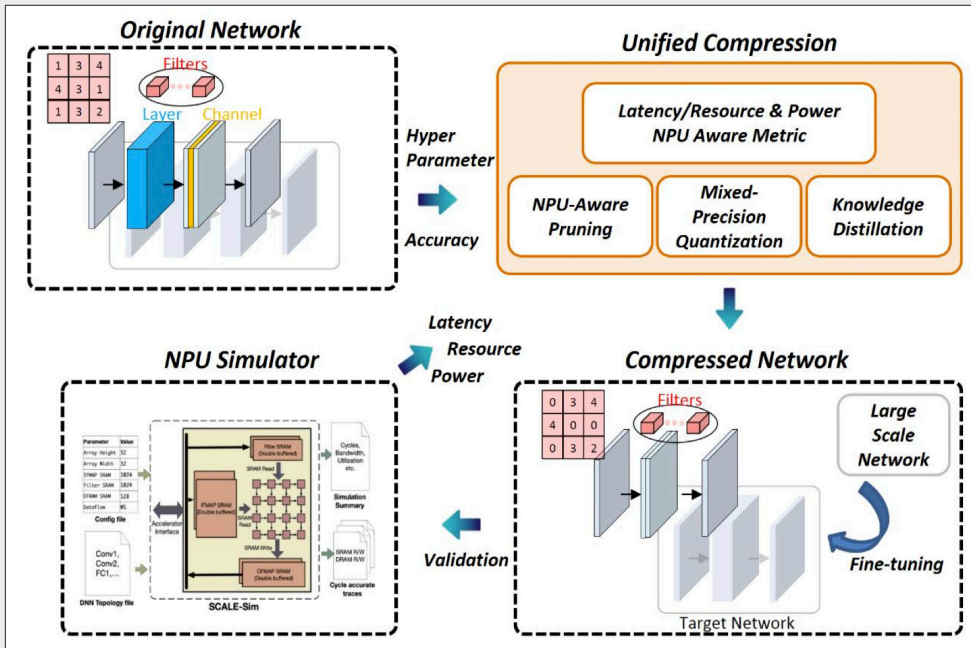
5. 시각지능 알고리즘 경량화 및 가속화 기술

스마트폰, 자율주행차, 드론 등의 지능형 모바일 디바이스에서 다양한 시각지능(객체 검출/인식/추적 및 분할, 모션/거리 추정, 자기위치추정 및 경로계획 등)의 탑재 요구가 커짐에 따라, 이를 디바이스에서 실시간/저전력으로 구동하기 위한 경량화/가속화 기술의 연구개발이 전세계적으로 활발히 진행 중이다. 시각지능IP팀에서는 다양한 컴퓨터비전/딥러닝 알고리즘의 경량화 연구와 이를 FPGA 또는 ASIC¹¹⁾에서 가속 구동을 위한 병렬 구조의 H/W IP를 개발하고 있다.

▶ 딥러닝 모델(객체 검출/인식) 경량화 기술(<그림 10>)

- Knowledge Distillation 학습 기술 및 S/W Framework 개발 중
- 다양한 수준의 Pruning 및 Data(Kernel Weight/Activation) Quantization를 위한 학습 기술 및 S/W Framework 개발 중
- NPU H/W Architecture Aware Pruning/Quantization 기술 개발 중

8) Neural Processing Unit
 9) Network Video Recorder
 10) Field-Programmable Gate Array
 11) Application-Specific Integrated Circuit



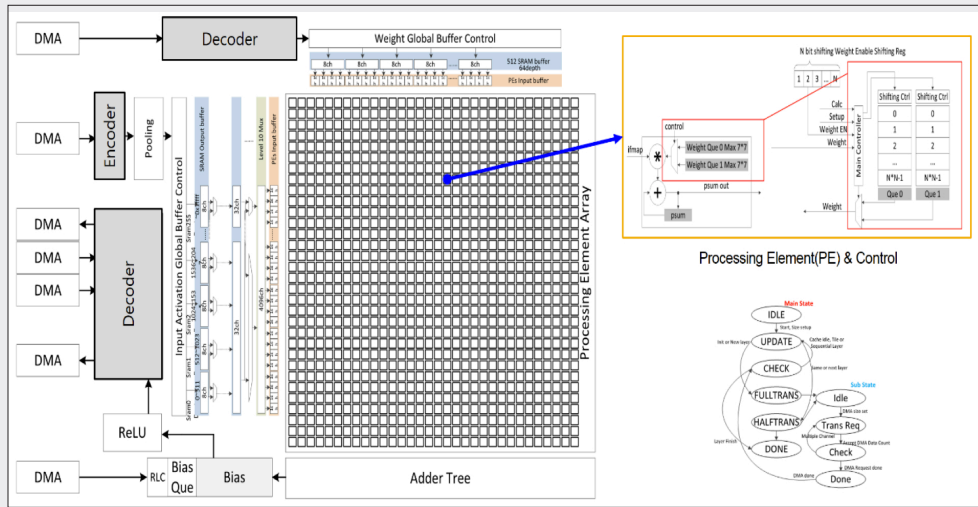
< 그림 10> NPU-Aware/Independent 경량화 Framework

- 딥러닝 가속 프로세서(NPU) IP 설계 기술(〈그림 11〉)
- 초병렬 연산 유닛 구조 기반 초고속 딥러닝 연산 모듈 및 연산 유닛의 Utilization 극대화를 위한 제어 방법/모듈 개발
- 외부메모리 접근 최소화 및 데이터(Kernel Weight/Activation) 재사용률 극대화를 위한 최적 메모리 구성 및 제어 방법 개발
- 다양한 딥러닝 모델 구동을 위한 Reconfigurable(S/W Controllable) 아키텍처 개발
- Kernel Weight/Activation 데이터의 Layer-Wise 압축/복원 방법 및 모듈(코덱) 개발
- 프로세서 IP 및 통합 SoC 개발(삼성 28nm 공정 기반 반도체 개발 중)
- NPU IP 기반 ADAS용(보행자/차량/Cyclist/차선 검출, 주행도로 Segmentation, 교통표지판 인식 등) IP, 로봇 자동물류용 객체 검출/분류 IP의 확장 개발

▶ 컴퓨터비전 알고리즘 가속 IP 설계 기술(〈그림 12〉)

- 머신 러닝(SURF¹², ORB¹³, SVM¹⁴, Adaboost 등) 기반 객체/장면인식/얼굴검출 가속 IP 개발 및 소형 로봇 통합
- 실시간 Optical Flow(Motion Map) 및 Stereo-Depth 추출 H/W IP 개발
- 초고속 객체 추적 IP 개발 및 드론 통합

12) Speeded Up Robust Features
 13) Oriented FAST and Rotated BRIEF
 14) Support Vector Machine



<그림 11> 딥러닝 가속 프로세서 IP 아키텍처



<그림 12> 컴퓨터비전 알고리즘 가속 IP 개발 및 디바이스 통합

6. 3D 휴먼 복원 및 변형 기술

컴퓨터비전분야에서는 360도 전방위에서 시청이 가능한, 몰입감 있는 3차원 모델을 복원하는 연구를 진행하고 있으며, 특히 3D 휴먼을 단일 또는 다중 카메라를 이용하여 복원하고, 복원된 3차원 모델에 움직임 을 더하는 모델 변형 기술을 개발하고 있다.

▶ 딥러닝 기반 휴먼/객체 복원 기술(<그림 13>)

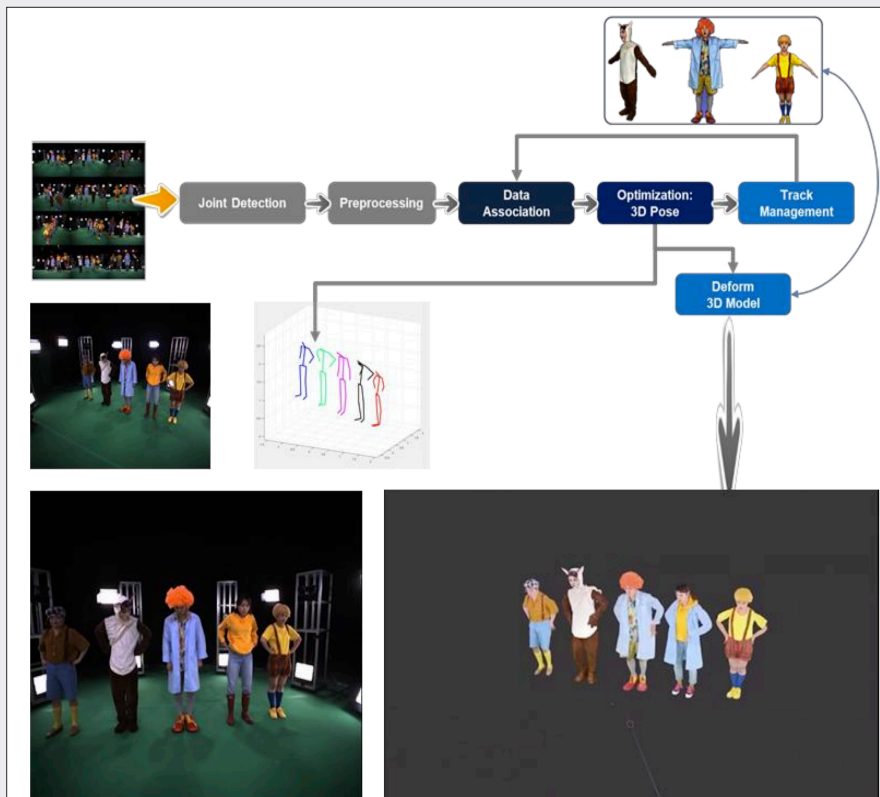
- 단일 영상 기반의 고훈해상도 3차원 복원 기술
- 다중 시점 기반의 다중 휴먼 복원 기술(공연, 스포츠)
- 단일 영상 기반 템플릿 모델 정합 기술



<그림 13> 단일 시점 영상 기반 고휘상도 3차원 모델 복원 기술

▶ 비마커 모션 캡처 및 모델 변형 기술(<그림 14>)

- 단일/다중 시점 카메라를 이용한 비마커 모션 캡처 기술
- 인체 사전 정보 기반 3차원 자세 최적화
- 실제 크기를 반영한 자세 추정 및 복원 모델 변형 기술



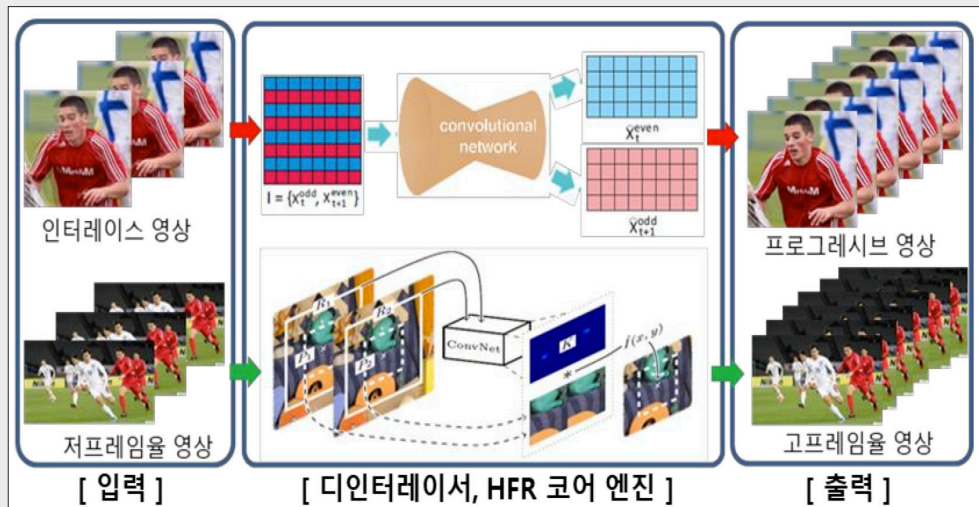
<그림 14> 다중시점 비마커 모션캡처/아바타변형 기술

7. 지능적 미디어 처리 및 변환 기술

SD/HD급으로 촬영된 저화질 구작(舊作) 미디어를 공간/시간 해상도 향상, 색역 변환, 화면비 확대 등을 통해 고품질 시청환경에 적합한 AI 기반 미디어 화질 개선 기술을 개발 중에 있으며, 이런 미디어 처리 및 변환 기술을 최적/경량화하여 모바일 단말에서 수행하는 기술을 개발 중이다.

▶ 딥러닝 기반 고프레임을 변환 기술(〈그림 15〉)

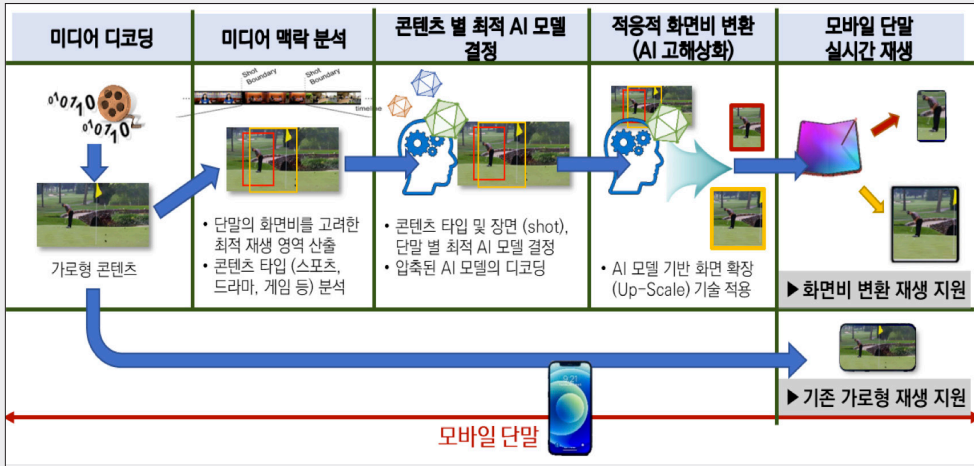
- 딥러닝 기반 실시간 디인터레이싱 기술(interlaced → progressive, 60i → 30p)
- 영상특성 적응형 딥러닝 기반 2배~8배 프레임율 향상 기술
- 시간/공간 해상도 융합 적용 네트워크 개발



〈그림 15〉 딥러닝 기반 고해상도 프레임율 변환 네트워크

▶ 모바일향 미디어 처리 및 변환 기술(〈그림 16〉)

- 딥러닝 기반 미디어 맥락 분석 기술
- 딥러닝 기반 화면 확장(Super-resolution) 기술
- 모바일향 딥러닝 모델 최적/경량화 기술 개발



<그림 16> 모바일 단말 기반 미디어 화면비 변환 재생 기술

영상처리 분야에 있어, 향후 10년은 몰입형 비디오, 5G, AI, 메타버스 등의 키워드를 중심으로 기술 개발이 이루어질 전망이다. 본 연구센터는 ‘카메라의 영상을 통해서 획득되는 정보로 사람이 눈을 통해서 인식하고 판단하는 세상을 동일한 수준으로 이해하고 판단할 수 있는 기술’의 개발을 목표로, 메타버스 환경 구축의 기반이 되는 실사영상 기반으로 가상 공간을 실현하기 위한 컴퓨터 비전 기술, 다양한 객체와 상황에 대한 세밀한 이해/판단/예측 기술 및 몰입형 비디오 서비스를 위한 요소 기술 개발을 추진할 계획이다.