

차세대 콘텐츠를 위한 AI 기술 활용 동향 및 전망

Trends and Prospects in the Application of AI Technology for Creative Contents

홍성진 (S.J. Hong, sjhong0117@etri.re.kr)

이승욱 (S.W. Lee, tajinet@etri.re.kr)

윤민성 (M.S. Yoon, msyoon@etri.re.kr)

박지영 (J.Y. Park, jiyp@etri.re.kr)

이수웅 (S.W. Lee, suwoong@etri.re.kr)

김아영 (A.Y. Kim, kimay@etri.re.kr)

정일권 (I.K. Jeong, jik@etri.re.kr)

지능형지식콘텐츠연구실 연구원

CG/Vision연구실 책임연구원

VR/AR콘텐츠연구실 책임연구원

감성상호작용연구실 책임연구원

콘텐츠인식연구실 선임연구원

콘텐츠실증연구실 선임연구원

차세대콘텐츠연구본부 책임연구원/본부장

ABSTRACT

With the development of artificial intelligence (AI) and 5G technology, an ecosystem of digital content is gradually becoming intelligent, immersive, and convergent. However, there is not enough ultra-realistic content for the ecosystem. For ultra-realistic content services, creative content technologies using AI are being developed. This paper introduces the trends in and prospects of creative content technologies such as 3D content creation, digital holography, image-based motion recognition, content analysis/understanding/searching, sport AI, and content distribution.

KEYWORDS creative contents, content intelligence, content AI trends

1. 서론

디지털 콘텐츠는 우리 생활과 밀접한 관련이 있다. TV를 시청하고, 음악을 들으며, 인터넷 쇼핑을 하고, 컴퓨터 게임을 하는 등 대다수 사람은 살면서 한 번쯤은 경험해 보았을 것이다. 이렇게 부호 · 문자 · 도형 · 색채 · 음성 · 음향 · 이미지 및

영상 등의 정보가 디지털 형태로 제작되고, 정보통신망을 통해 다양한 부가가치를 창출할 수 있는 서비스들을 디지털 콘텐츠라고 한다[1].

과거의 디지털 콘텐츠는 주로 영상의 형태로 단편적으로 제공되고, 소비자들의 선택 폭은 제한적이었다. 하지만 최근에는 디지털 교과서로 수업이 진행되고, 실내에서 스크린 스포츠를 즐기며, 개

* DOI: <https://doi.org/10.22648/ETRI.2020.J.350511>

* 논문작성에 도움을 주신 통신미디어연구소 차세대콘텐츠연구본부의 양성일, 임성재, 양용연, 이해룡, 박지현, 김명규, 박창준, 이기석, 김진서, 유원영, 지형근에게 감사의 말씀을 전합니다.

* 본 연구는 한국전자통신연구원 주요사업의 일환으로 수행되었음[20ZH1200, 초실감 입체공간 미디어 · 콘텐츠 원천기술 연구].



본 저작물은 공공누리 제4유형

출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.

©2020 한국전자통신연구원

인 맞춤형 정보를 자동으로 추천받는 등 디지털 콘텐츠는 다양해지고 있다. 인공지능 기술과 결합하여 생성/분석, 가시화, 인터랙션, 플랫폼 등 디지털 콘텐츠의 생태계는 점차 지능화되고 있으며, 초실감 서비스를 제공할 수 있는 차세대 콘텐츠로 발전되고 있다.

초실감 서비스를 제공하기 위한 차세대 콘텐츠 기술은 3D 콘텐츠 생성, 디지털 홀로그래피, 영상 기반 동작인식, 콘텐츠 분석/이해/검색, 스포츠 인공지능, 콘텐츠 유통 플랫폼 등이 있다. 본 논문에서는 차세대 콘텐츠 핵심기술들의 국내·외 동향과 환경변화 및 전망을 소개하고자 한다.

II. 3D 콘텐츠 생성

1. 국내·외 동향

3D 센서의 발전 및 관련 컴퓨터 비전기술의 발전으로 3D 데이터의 확보가 가능함에 따라 3D 딥러닝에 대한 기술개발이 가속화되고 있다. 3D 콘텐츠에 대한 인공지능의 연구분야는 응용서비스를 기준으로 3가지 분야로 나눌 수 있다. 먼저 입력 3D 모델을 분석하여 분류(Classification), 구성요소 분할(Segmentation) 및 모델 간의 상관관계(Correspondence)를 분석하는 분야이다. 이 중 모델 간의 상관관계를 분석하는 것은 3D 모델의 변형을 위해 많이 사용된다. 또 다른 응용 영역은 3D 합성(Synthesis) 분야이다. 이는 불완전한 입력으로 완전한 3D 모델을 만드는 것이다. 예를 들어 스캔된 영상을 입력하여 빈 부분을 추론하거나, 2D 영상을 입력하여 3D 콘텐츠를 생성하는 것으로 현재 활발히 연구되고 있다. 마지막으로 2D 영상을 분석하기 위해 내부적으로 3D 모델을 생성하는 분야가 있다.

본 장에서는 현재 활발히 연구되고 있는 2D 영

상에서 3D 콘텐츠를 생성하는 3D 합성을 주로 고찰한다.

Kar는 학습 영상의 태깅정보와 객체의 분할정보를 활용하여 한 장의 입력 영상에서 3D 객체를 생성하는 연구를 수행했다[2]. Rock은 깊이정보를 가진 한 장의 영상으로 DB에 있는 가장 유사한 데이터를 획득했고, 깊이정보의 차이를 보정해 온전한 3D 객체를 생성했다[3]. Lun은 한 장(정면) 혹은 두 장(정면/측면)의 영상에서 360° 다시점 영상을 생성하고, 해당 영상에 멀티뷰 기반 영상 복원 알고리즘을 적용하는 기법을 개발했다[4]. 제안하는 방법은 총 14개의 시점 영상을 사용했고, 시점의 개수가 증가함에 따라 학습의 부하는 증가했지만 결과의 품질은 향상되었다. 또한, 입력 데이터를 격자형 구조로 만들 수 있어 기존의 합성곱 기반의 네트워크를 직접 적용할 수 있었다. Dibra는 한 장 혹은 두 장의 실루엣 영상이나 음영 정보를 가진 영상에서 3D 인체 형태정보를 추정하는 방법을 제안했고, CAESAR DB에서 한 장의 정면 실루엣 영상을 추출하고 합성곱 네트워크를 활용해 3D 전신 인체모델을 복원했다[5].

2. 환경변화 및 전망

인공지능을 이용한 3D 콘텐츠 생성 연구의 이슈는 학습데이터 포맷이다. 기존의 1D/2D 데이터 기반의 인공지능 연구는 학습데이터가 격자구조 모양으로 정의되어 있어 합성곱 기반의 네트워크가 잘 적용된다. 하지만 3D 모델의 경우 질감, 연결정보, 리깅정보 등 다양한 데이터가 있어서 기존의 합성곱 네트워크에 적용하기 어렵다. 이를 해결하기 위해 데이터 포맷에 대한 연구가 진행 중이다. 만약 3D 학습에 적합한 데이터 구조가 개발된다면, 다시점 영상이나 복셀 등으로 변환 없이 기

존의 다양한 네트워크를 활용하여 3D 콘텐츠를 생성할 수 있을 것으로 예상된다.

III. 디지털 홀로그래피

1. 국내·외 동향

디지털 홀로그래피에서 캡처된 또는 합성된 홀로그램으로부터 고품질의 이미지 복원을 위해 0차 회절, 복소공역항, 배경 노이즈 등의 억제와 같은 복잡하고 시간 소모적인 추가적 처리가 요구된다. 홀로그래피는 광학 필드의 위상과 진폭을 모두 사용함으로써 단일 이미지로부터 사용가능한 학습 훈련용 데이터양을 두 배 늘릴 수 있는 장점이 있다. 딥러닝이라는 기계학습 형태는 실시간 음성 인식, 자동화된 이미지 변환 처리 및 비디오 레이블링과 같은 응용 분야에서 최근 발전한 주요 기술 중 하나이다. 데이터 분석을 자동화하기 위해 다층 인공 신경 네트워크를 사용하는 이 접근 방식은 헬스케어 분야에서 중요한 가능성을 보여주고 있다.

예를 들어, 환자의 X-선, CT 스캔 및 기타 의료용 이미징에서 이상을 자동으로 식별하는 데 사용될 수 있다. 최근 UCLA의 Ozcan 교수 연구팀은 딥러닝의 새로운 용도를 발표했다. 홀로그램을 복원(Reconstruction)하여 현미경에서 물체의 미세한 이미지를 형성함으로써 광학 현미경의 성능을 개선했다[6,7]. 혈액 샘플 및 유방 조직 샘플의 홀로그램을 복원하기 위해 딥러닝 신경망을 이용했고, 원하지 않는 광 간섭 및 이미지 복원 과정의 다른 물리적 부산물로부터 물체의 실제 이미지 특징을 추출하고 분리했다. 해당 연구는 전통적인 물리기반 홀로그램 재구성 방법을 딥러닝기반 계산 방식으로 대체한 대표적인 사례이다. 이 새로운 홀로그래피 이미징 기술은 여러 홀로그램들을 사용하는 기존 방법보다 더 나은 이미지를 생성하고, 더 적은

횟수의 측정을 필요로 하며, 컴퓨터 계산도 더 빠르기 때문에 구현과정이 매우 용이하다.

다음으로 디스플레이용 홀로그램 콘텐츠 합성을 위한 대표적인 기술은 컬러 영상을 입력으로 하는 인코더-디코더 구조를 갖는 DenseDepth 모델을 통해 깊이맵 이미지를 추출하고, 획득된 깊이맵과 사전 준비된 컬러 이미지로부터 홀로그램을 계산하는 방법이 있다[8].

DenseDepth 모델은 기존에 저해상도 및 낮은 선명도 근사 방법의 개선된 솔루션이다. 인코더-디코더 구조는 더 정확하고 상세한 경계면(Boundary)을 나타내는 고품질의 깊이맵 결과를 도출하도록 마련된 증강 학습 전략(Augmentation and Training Strategy)과 더불어 사전에 훈련된 고성능 네트워크를 사용하여 특징을 추출할 수 있다. 2018년에 개발된 인코더-디코더 구조 기반 양질의 깊이맵 추정 방법은 단일 컬러 영상 외에도 한 쌍의 스테레오 이미지들을 사용하거나 다시점 영상들로 구성된 이미지 세트로부터 깊이정보를 추정할 수 있다. DenseDepth 아키텍처를 토대로 추출된 고품질의 깊이맵 이미지는 디지털 홀로그램 콘텐츠를 고속 획득하기 위한 FFT 알고리즘을 적용하여 CGH 합성 기술에 적용되고 있다[9].

홀로그래피 회절 광학 분야는 딥러닝기반 SW/HW 시스템에 대한 많은 잠재적인 응용을 갖고 있다. 대표적으로, 회절 광학 소자(DOE: Diffractive Optical Element)를 설계 및 제작하는 것은 산업적 응용 측면뿐만 아니라 광학 데이터 처리 분야에서 매우 중요한 과제이다[10]. DOE는 슬림한 프리즘·수렴 렌즈, 3D 장면 표시용 광부품, 패턴 인식용 광학 소재 등에서 많이 활용되고 있으며, 전통적으로 Gerchberg-Saxton 알고리즘이나 핀업(Fienup)과 같은 반복적 연산 방법이 DOE 패턴 설계를 위해 사용되고 있다. 그러나 딥러닝 알

고리즘을 사용하면, 이러한 계산 집약적인 반복 연산 방법을 단일 작업 내에 포함시킴으로써 계산 시간의 단축 및 복잡성 문제를 개선시킬 수 있다. 임의의 광 세기 이미지에서 역광 전파를 근사하도록 훈련된 딥러닝 모델은 임의의 이미지에서 위상 전용 DOE를 생성할 수 있다. 2020년 러시아 NRNU대학 Rymov 교수 연구팀은 깊은 잔여 신경망(Deep Residual Network)을 사용하여 임의의 데이터 세트에서 회절 광학 요소(DOE)를 생성하는 방법을 제안했다[11, 12]. 이미징의 역 문제(Inverse Problem)를 해결하기 위해 개발된 심층 CNN 구조와 유사한 이 네트워크는 역광 전파(Inverse Light Propagation) 과정을 근사하도록 훈련되었으며, 훈련된 모델이 사용되지 않은 테스트 이미지로부터 위상-DOE를 성공적으로 생성할 수 있음을 실증하였다. 유사한 방법을 토대로 네트워크 구조의 최적화를 통해 디지털 홀로그램의 품질을 향상시키고, 나아가 컴퓨터 홀로그램 생성의 계산 및 기타 회절 광학 소자의 성능을 더욱 향상시킬 수 있을 것으로 기대된다.

2. 환경변화 및 전망

최근 수년 동안에 심층 신경망은 홀로그래피에서의 이미지 재구성 및 위상 복구를 위한 잠재적인 응용에 관하여 상당한 검증 연구가 이루어졌다. 오늘날 급격히 발전 중인 바이오·의료 영상 획득을 위한 현미경용 딥러닝 기반 홀로그래픽 이미징 기술의 최신 결과는 모든 위상 복구 및 홀로그램 이미징 문제에 광범위하게 적용할 수 있다. 따라서 이 심층학습 기반 프레임워크는 가시광 대역 파장 및 X-선을 포함하여 전자기 스펙트럼의 여러 부분을 걸쳐 근본적으로 새로운 공간섭성 이미징 시스템 설계를 위한 연구로 발전될 것으로

전망한다.

딥러닝 기반 디스플레이용 홀로그래픽 콘텐츠 합성 기술은 향후 사용자와 콘텐츠 간 상호작용이 가능하도록 주어진 장면에 관하여 3차원 공간 위치에 정확히 3D 이미지가 복원되도록 디자인되어야 한다. 이를 위해 깊이맵 정보의 추정 속도 향상 기술뿐만 아니라 동시에 목표로 하는 깊이정보를 정확히 추정하는 기술이 확보되어야 한다. 기계학습 알고리즘을 활용한 DOE 패턴용 딥러닝 접근 방식은 전통적인 방법인 계산 집약적인 반복 알고리즘 대비 성능이 우수한 단일 작업으로 근사화하는 것을 목표로 개발이 진행되고 있다. 이를 위해 종래의 이미지 인식 및 컴퓨터 비전에서 축적된 경험들을 적용한다면, 고품질의 홀로그램 생성과 회절 광학 디자인뿐만 아니라 3D 이미지 복원, 컴퓨터 이미징 등에서 사용할 수 있는 일반화된 딥러닝 모델을 구현할 수 있을 것으로 예상된다.

딥러닝을 이용한 홀로그래픽 이미지 처리 관련 새로운 접근법들은 빛과 물질 간 상호작용의 모델링 또는 파동 방정식의 솔루션 없이도 구현할 수 있는 장점이 있다. 고해상도의 복잡한 홀로그래픽 콘텐츠일수록 개별 샘플과 빛의 형태에 대한 모델링과 계산이 요구되며, 이 경우 도전적인 컴퓨팅 파워와 장시간 소요되는 전통적인 물리 기반 접근 방식들을 대체할 수 있을 것으로 예상된다.

IV. 영상기반 동작인식

1. 국내외 동향

사람의 음성과 동작은 콘텐츠와 사용자가 소통하는 수단으로 사용자 친화적인 자연스러운 UI(User Interface)를 제공할 수 있는 신호이다. 사용자의 동작을 인식하는 기술은 깊이 카메라나 컬러 카메라를 활용하여 사용자를 관찰하고 동작의 의

미를 이해하는 기술이다.

동작인식 기술은 인체의 관절을 바탕으로 골격 구조를 인식하여 움직임 해석하는 접근이 주를 이룬다. 동작인식에 활용되는 대표적인 장치로는 전신 동작인식이 가능한 깊이 카메라 MS Kinect와 손동작 인식이 가능한 Leap Motion이 있으며, 웨어러블 장치로는 팔근육의 움직임을 인식하는 암밴드 형태의 Myo가 대표적이다.

최근에는 깊이 카메라나 착용형 장치를 사용하지 않고, 한 대의 컬러 카메라만을 활용하는 딥러닝기반의 사용자 동작인식 연구가 매우 활발히 진행되고 있으며, 정확도 측면에서도 상용 목적으로 활용 가능한 수준의 결과물들이 발표되고 있다. 컬러 영상을 활용하는 방식은 깊이 영상과 동일한 접근으로 연속적인 포즈 분석기반의 동작인식을 위해 인체의 관절 위치를 정확히 추정하기 위한 연구가 다수이며, 이를 위해 영상의 특징점 추출 네트워크를 다양화하거나[13-15], 표준 인체모델을 입력 영상에 정합하는 방법[16] 등이 있다.

딥러닝기반의 동작인식 기술개발을 위해서는 관절위치 기반의 자세 학습을 위한 대용량 데이터 확보가 반드시 선행되어야 한다. 가장 많이 활용되는 공개 데이터셋으로는 MS의 COCO 데이터셋[17]과 max plank의 MPII 데이터셋[18]이 있다. 양질의 데이터 확보를 위해서는 그림 1과 같이 컴퓨



그림 1 인체모델 및 모션 변형을 통한 대용량 학습데이터 생성 과정

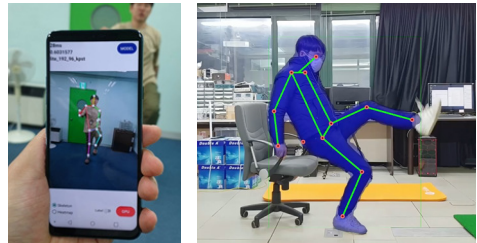


그림 2 컬러영상기반 2D 자세인식 기술

터그래픽스 기술을 적용하여 3D 휴먼모델의 변형과 다양한 모션 데이터 리깅, 배경 및 조명 변환 등의 과정을 수행함으로써 대규모 학습데이터를 자동으로 생성하는 방법을 사용할 수 있다.

대규모 데이터의 기계학습을 바탕으로 구축한 딥러닝 네트워크를 활용하면 그림 2와 같이 단일 RGB 영상에서 인체의 관절 위치를 파악할 수 있다. 해당 네트워크는 단안의 컬러 카메라를 활용하여 실시간 구동이 가능하므로 스마트폰 등의 모바일기기에서도 활용될 수 있다.

그림 3과 같이 단일 RGB 영상에서 추출한 2D 관절 위치를 대규모 3D 데이터 학습을 통해 3D로 변환하는 연구도 활발히 진행 중이다[19]. 하지만 이 경우 정확도는 학습데이터에 종속적인 한계가 있어 아직 상용화 단계에는 이르지 못하였다.

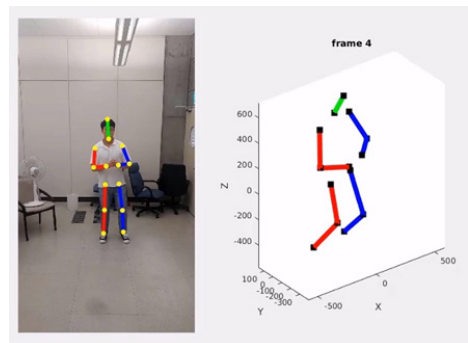


그림 3 컬러기반 3D 자세인식 기술

2. 환경변화 및 전망

최근에 인공지능 스피커가 다양하게 출시되며, 일상적으로 활용하는 키보드나 리모컨과 같은 기계식 인터페이스에서 벗어나, 보다 사용자 친화적인 인터페이스가 보급되는 계기가 되었다. 동작인식은 사람의 인지 과정에 70% 이상의 영향을 미치는 시각을 활용하는 과정으로, 기술적으로는 음성 인식과 비교하여 아직 정확도가 떨어진다. 하지만 동작인식 기술은 인터페이스의 측면에서 활용도가 매우 높은 수단으로, 향후 포스트 코로나 시대에 요구되는 비대면 환경의 다양한 산업, 교육, 엔터테인먼트 등의 분야에서 활용 가능한 핵심기술로써 자리 잡을 것으로 전망한다.

V. 콘텐츠 분석/이해/검색

1. 국내·외 동향

콘텐츠 분석/이해/검색은 이미지와 동영상 콘텐츠로부터 의미 있는 2차 콘텐츠를 생성하기 위한 기술을 개발하는 것이다. 예를 들어 영상에서 사람, 동물, 건물 등 의미 있는 정보들을 객체단위 또는 화소단위로 이해하며, 분별력 있는 특징을 추출하여 검색에 활용한다.

최근 콘텐츠 분석/이해/검색에서는 옛지 디바이스에서 영상 인식이 가능한 소프트웨어/하드웨어 플랫폼들이 발표되고 있다. xnor.ai에서는 Intel과 협업하여 OpenVINO플랫폼에서 1bit모델을 시연했고, Ai2go라는 플랫폼을 통해 모바일, 임베디드보드 등 다양한 하드웨어에서 구동되는 다양한 시나리오의 SDK를 배포하고 있다. Xilinx는 FINN[20]과 BISMO[21]라는 FPGA기반의 영상 콘텐츠 분석 하드웨어를 개발했으며, 2018년 7월 중국의 딥러닝 스타트업인 DeePhi Tech를 합병하며



그림 4 ETRI DeepMobileAR 어플리케이션

딩치를 키우고 있다.

또한, 실내·외 영상정보를 기반으로 영상 검색을 수행하여 식당 정보, 위치 정보 등 사용자에게 필요한 검색 정보를 제공하는 서비스가 출시되었다. Google은 구글 렌즈를 통해 구글 지도와 인공지능기반 이미지 검색을 결합해 AR 내비게이션 앱을 2018년 출시했다. 네이버는 시각적 위치추정(Visual Localization)기술을 적용한 실내 AR 내비게이션 시제품을 2019년에 전시했고, 2022년까지 서울 지역 모든 실내·외 공간의 고정밀 지도화 작업을 수행 중이다.

ETRI에서는 2019년 경량 객체 검출 모델과 장면 분할 모델을 개발했다. 해당 모델을 활용하여 증강현실 서비스가 가능한 DeepMobileAR이라는 어플리케이션을 개발했고, 그림 4와 같이 모바일 기기에서 테이블 영역과 테이블 위의 객체(종이컵)를 인식해 증강현실 기반의 인터랙션 기능을 제공한다[22].

진화하는 불법 콘텐츠 생성 기술에 대응하여 딥러닝 기술로 정교하게 영상을 식별할 수 있는 솔루션들이 발표되고 있다. Google은 YouTube에서 저작권 침해 관리를 위해 사용하고 있는 저작물 식별 시스템인 콘텐츠 아이디(content ID)에 인공지능 기

술을 적용하여 영상 왜곡을 감지할 수 있도록 기능을 향상시켰다[23]. Irdeto는 인공지능 기술 활용해 영상 내 자동 로고 감지, 텍스트 인식, 얼굴 인식이 가능하도록 불법 콘텐츠 차단 시스템을 확장함으로써 정교하게 편집된 불법 영상을 검출할 수 있도록 했다[24].

음악검색 분야에서 모바일용 소규모 딥러닝 모델기반 음악검색 서비스와 표절검사 서비스가 등장했다. Google은 모바일기기에서 저전력 stand-alone형 음악식별을 위해 딥러닝기반의 핑거프린트 기술을 개발했고 음악검색 서비스인 Now Playing에 적용했다[25]. Fraunhofer 연구소는 신호분석 기술과 딥러닝 기술을 활용해 음악 샘플링 표절과 멜로디 표절을 탐지할 수 있는 툴킷을 제공하고 있다[26].

2. 환경변화 및 전망

대부분의 딥러닝 모델은 깊은 신경망 구조로 모바일기기에서의 활용은 한계가 있다. 이를 해결하기 위해 가지치기(Pruning), 양자화(Quantization), 깊이별 컨볼루션(Depthwise Convolution) 등 딥러닝 모델의 경량화를 위한 연구와 하드웨어 최적화 연구가 진행되고 있다. 하지만 아직 실시간성을 보장하기 어렵고, 인식 결과만을 활용하는 등 단순 증강 콘텐츠로 제한되고 있다. 앞으로도 모바일 환경에서 실시간 서비스를 제공하기 위한 관련 연구들이 지속될 것으로 예상되며, 모바일기기에서 직접 객체를 검출 및 인식하고, 컨텍스트 정보를 분석하며 실제와 가상이 연동된 증강 콘텐츠를 제공할 수 있을 것으로 전망된다.

또한, 불법 콘텐츠의 진화로 인공지능을 활용하여 기존 저작권 보호 기술을 회피하는 등 저작권 침해 기술이 고도화되고 있다. 이에 대응하여 향상

된 저작권 침해판단 기술이 요구되고 있다. 전체 콘텐츠의 동일성뿐만 아니라 부분적 특징정보를 활용한 저작권 판단 기술이 연구된다면 사용자의 편집을 통해 유포되는 부분적 불법 콘텐츠의 확산을 방지할 수 있을 것으로 전망한다.

VI. 스포츠 인공지능

1. 국내·외 동향

스포츠 분야는 선수, 경기에 대한 기록들이 정량화되어 인공지능 기술을 적용하기 적합하다. 풋볼, 야구 등 다양한 스포츠에서는 팀의 전략 수립 및 운영을 위해 인공지능기술을 활용하고 있다. NFL의 Next Gen Stats는 경기장에 배치된 센서를 이용하여 선수들의 속도, 방향, 거리, 투구 속도 등의 정보를 측정했고, 아마존 웹서비스(AWS: Amazon Web Services)의 머신러닝 기술을 활용하여 선수들의 캐치 확률, 키커의 골을 놓칠 확률 등을 계산했다[27]. Stats Perform은 축구의 세트피스 상황을 분석했고, 어떤 종류의 전술이 팀에 위협적이거나 효과적으로 활용될 수 있는지 분석했다[28]. 또한, 농구경기의 상황정보와 선수들의 모션정보를 기반으로 각 선수들의 공격과 수비의 움직임 예측했다[29].

스포츠 인공지능 기술은 사용자의 개인동작을 분석하여 퍼포먼스를 향상시킬 수 있도록 지원한다. Fathom AI은 3개의 모션센서를 사용하여 달리기 자세를 분석했다. 인체의 움직임 연구결과와 신체 회복 사례를 활용하여 사용자의 움직임을 분석했고, 자세 교정 피드백을 제공했다. Stats Perform은 인체 포즈 데이터를 활용하여 NBA 경기에서 나타나는 다양한 스타일의 슛 동작을 인식했고, 슛의 성공과 난이도에 영향을 미치는 속성을 분석했다[30]. ETRI에서는 인체 포즈 정보를 활용하여 음악

줄넘기 급수평가 동작을 인식했고[31], 사용자의 줄넘기 수준을 자동으로 평가했다. 또한, 축구의 킥/드루인 동작에 대한 딥러닝기반 구분동작 분석 기술[32]을 개발해 각 동작들을 단계별로 구분하여 코칭에 활용할 수 있게 했다.

스포츠 부상은 선수 생명과도 밀접한 연관이 있다. 선수와 구단 모두 관심을 갖는 부분으로 부상 방지 기술에도 인공지능 기술이 활용되고 있다. NFL에서는 선수들이 머리에 가해지는 잦은 충격으로 인해 뇌진탕이 발생하는 경우가 많다. 이를 예방하기 위하여 훈련이나 게임 중 머리에 가해진 충격량을 측정했고, 선수가 안정을 찾기까지의 필요한 휴식시간을 계산했다[27]. Sparta Science는 밸런스, 플랭크, 점프와 같은 간단한 동작을 포스 플레이트로 측정했고, 개인별 동작 패턴을 분석해 선수의 몸 상태와 부상의 위험도를 측정했다. QMIT는 기계학습을 활용한 코칭앱으로 선수들의 컨디션과 통증을 관리하고 선수들의 부상을 예측하고 예방하고 있다.

방송분야에서는 시청자들의 재미 요소를 위한 하이라이트 생성에 인공지능을 활용하고 있으며, IBM은 워블던 테니스 대회에서의 매치 데이터, 선수의 표현, 관객의 소리 정보 등을 활용해 하이라이트 영상을 자동으로 생성했다[33].

2. 환경변화 및 전망

프로 스포츠를 중심으로 전문 인력의 경험과 직관으로 판단되는 분석 기술은 인공지능과 결합하여 데이터로 정량화되고 있다. 데이터가 점차 쌓일수록 기존 분석 방법을 뛰어넘는 인공지능 기술이 개발되고, 또한 새로운 데이터를 통해 점차 분석 영역을 확대해갈 것이다.

현재 개발 중인 인공지능 스포츠 플랫폼들은 프

로 선수들을 대상으로 하지만 사회인/동호인 등 개인이 활용할 수 있는 플랫폼으로 확장될 것이다. 고가의 장비를 요구하는 대신 휴대폰, 웹캠 등 간단한 장비를 활용한 서비스가 가능할 것으로 전망된다. 또한, 종목에 특화되어 단일 종목 중심으로 개발되는 스포츠 인공지능 기술은 유사한 종목들 간의 결합으로 다종목을 지원하며 제한된 공간에서도 여러 스포츠 체험이 가능한 멀티 스포츠 플랫폼으로 확장될 것으로 전망된다.

Ⅶ. 콘텐츠 유통 플랫폼

1. 국내·외 동향

콘텐츠 유통 플랫폼은 생산된 콘텐츠를 소비자에게 제공하는 과정을 돕는 인프라를 말한다[34]. 콘텐츠 유통 분야에서는 인공지능 기술을 활용하여 사용자를 분석하고 맞춤형 서비스를 제공하기 위한 기술들이 개발되고 있다. 트루 슬리퍼는 POS기에 남은 데이터 이외에도 인공지능기반 영상분석 기술을 활용하여 내점률, 내점 고객의 구매 비율 등을 분석하고 고객의 동향을 파악해 마케팅 전략을 수립하고 있다[35]. Shop Direct는 SAS CI 솔루션을 활용하여 사용자, 재고, 판매 관련 데이터를 기반으로 개인 맞춤형 추천 서비스를 제공하였고, 고객의 행동 변화시점을 파악하는 모델을 통해 고객의 이탈을 방지했다[36]. Appier는 대용량 디지털 프로파일 기반 사용자의 성향과 행동방식을 학습하였고, 사용자 세분화 분석 기술을 통해 고객 예측, 개인 맞춤, 시장 창출 등 마케팅의 전반적인 영역을 지원했다[37].

또한, 사용자 세분화 분석 및 추천 서비스를 제공하기 위해 수집하는 개인정보를 보호하기 위한 표준화와 윤리지침 수립이 진행되고 있다[38]. 국제협약(Convention 108) 위원회에서는 인공지능과

데이터 보호에 대한 가이드라인을 발표하여 데이터 처리에 대한 정부 주체의 유의미한 통제를 허용하고 있고, 인공지능 기술이 데이터를 활용하는 경우 개인정보보호의 필요성을 강조했다[38]. Intel에서는 인공지능세계의 개인정보 및 데이터 보호에 대한 보고서를 발표했고, 인공지능 알고리즘의 투명성에 중점을 두는 등 데이터 활용에 대한 책임성을 강조했다[38].

2. 환경변화 및 전망

콘텐츠 유통은 온라인과 오프라인의 경계가 사라지고 있다. 사용자의 취향도 매우 다양해지고 있어 성별, 연령, 구매이력과 같은 단편적인 데이터만을 이용한 사용자 분석에는 한계가 존재한다. 사람의 행동에 영향을 미치는 감성 정보를 포함한 빅데이터는 더 나은 의사결정을 가능하게 할 것이다. 이와 관련하여 디지털 트윈 기술이 연구될 것으로 예상하며, 향후 대규모의 가상 사용자를 생성하고 다양한 상황을 시뮬레이션함으로써 기존보다 더 나은 맞춤형 콘텐츠를 사용자에게 제공해 줄 수 있을 것으로 예상된다.

Ⅷ. 결론

본 고에서는 차세대 콘텐츠 핵심기술의 국내·외 동향과 환경변화 및 전망을 소개했다. 콘텐츠 생성, 가시화, 분석, 유통 등 콘텐츠 생태계 전반에 걸쳐 인공지능 기술이 활용되고 있으며, 초실감 콘텐츠를 개발하기 위한 기술들이 연구되고 있었다. 3D 콘텐츠 생성에서는 2D 영상에서 3D 정보를 획득하기 위한 연구가 활발히 진행 중이고, 복잡한 3D 데이터를 처리하기 위한 포맷에 대한 연구가 요구되었다. 디지털 홀로그램에서는 홀로

그램 복원, 콘텐츠 합성, 회절 광학분야에서의 대표적인 사례들을 살펴보고, 위상 복구 및 홀로그램 이미징 문제, 홀로그래픽 이미지 처리 방법 등 다양한 분야에서 연구될 것으로 전망했다. 영상기반 동작인식에서는 대용량의 DB를 확보하고 사용자의 관절을 추정하는 연구가 진행되고 있었다. 하지만 2D에서 3D 자세를 추정하는 기술은 낮은 정확도로 인해 실환경에서 활용할 수 없어 상용화를 위해 많은 연구가 요구되었다. 콘텐츠 분석/이해/검색에서는 엣지 디바이스에서 활용 가능한 경량화된 인공지능과 하드웨어가 연구되고 있었다. 또한, 불법 저작행위를 검출하기 위한 기술들이 연구되고 있었다. 스포츠 인공지능에서는 전문 선수들의 퍼포먼스를 향상시키고 부상을 방지하기 위한 연구들이 진행되었다. 하지만 개인 및 가정용 등 보편화되기 위해 웹캠과 같은 간단한 기기활용이 요구되었다. 마지막으로 콘텐츠 유통에서는 이용자에게 맞춤형 콘텐츠를 제공하기 위한 추천기술들이 다양하게 연구되었고, 다양한 상황을 예측하기 위해 감성을 포함하는 빅데이터 구축과 디지털 트윈 기술 연구의 필요성을 확인할 수 있었다.

현재 우리나라는 세계 최초 5G 상용화로 선도적인 실증환경이 마련됨에 따라 실감콘텐츠 산업의 글로벌 주도권을 확보할 수 있는 환경이 조성되었다. 하지만 양질의 실감 콘텐츠 부재로 신시장 창출을 위한 콘텐츠 실증 프로젝트를 추진하고 있다[39]. 홀로그램 분야에서는 핵심기술을 활용한 문화유산·팩토리·상용차 3대 분야 실증 추진하고, 가상·증강·혼합현실 분야에서는 실감형 광화문 프로젝트, 증강도시, 문화 실감콘텐츠 분야에서 지역 문화 유희시설을 활용한 콘텐츠 실증을 지원하고 있다[39]. 이처럼 신시장 창출을 위한 초실감 콘텐츠에 대한 수요가 점차 증가할 것으로 예상되

며, 차세대 콘텐츠의 핵심기술들이 여러 방면에 적용되어 활용될 것으로 전망한다.

용어해설

DOE 투과된 레이저 광의 위상을 변경 또는 제어하는 미세패턴 구조를 갖는 회절 광학 부품

CGH 홀로그래픽 간섭무늬를 디지털로 생성하기 위한 계산 방법 또는 그 계산 결과로서 얻어지는 복소수 광파동 장(field)

엣지 디바이스 실시간으로 입력되는 데이터를 처리하는 장치

핑거프린트 원본 데이터에 삽입해 편집이 되더라도 본인이 작성했다는 것을 증명할 수 있는 것으로 유통되는 디지털 콘텐츠의 불법 복제를 방지하기 위한 식별 기술

디지털 트윈 가상환경에 현실과 같은 쌍둥이를 만들고, 현실에서 발생할 수 있는 기술을 가상으로 시뮬레이션함으로써 결과를 예측하는 기술

약어 정리

AI	Artificial Intelligence
AR	Augmented Reality
CGH	Computer-generated Hologram
CNN	Convolutional Neural Network
DB	DataBase
DOE	Diffraction Optical Element
FFT	Fast Fourier Transform
UI	User Interface

참고문헌

[1] <http://blog.daum.net/standards/314>

[2] A. Kar et al., "Category-Specific Object Reconstruction from a Single Image," IEEE Conf. comput. Vis. Pattern, Recogn. (CVPR), Boston, MA, USA, June 7-12, 2015, pp. 1966-1974.

[3] J. Rock et al., "Completing 3D Object Shape from one Depth Image," IEEE Conf. comput. Vis. Pattern, Recogn. (CVPR), Boston, MA, USA, June 7-12, 2015, pp. 2484-2493.

[4] Z. Lun et al., "3D Shape Reconstruction from Sketches via Multi-view Convolutional Networks," arXiv: 1707.06375, Sept. 2017.

[5] E. Dibra et al., "HS-Nets: Estimating Human Body Shape from Silhouettes with Convolutional Neural Networks," Int. Conf. 3D

Vision (3DV), Stanford, CA, USA, Oct. 25-28, 2016, pp. 108-117.

[6] Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," Light Sci Appl 7, 17141, 2018.

[7] Y. Rivenson and A. Ozcan, "Deep learning microscopy," Optica, 2017, Vol. 4, Issue 11, pp. 1437-1443.

[8] I. Alhashim and P. Wonka, "High Quality Monocular Depth Estimation via Transfer Learning," arXiv preprint arXiv:1812.11941, 2018.

[9] M.S. Yoon et al., "Depth map Extraction from Multi-view Images for Holographic 3D Display," 2020 International Display Workshops (IDW), 2020.

[10] 윤민성, 회절광학 및 홀로그래피, BOOKK, 2019.

[11] K. H. Jin et al., "Deep Convolutional Neural Network for Inverse Problems in Imaging." IEEE Transactions on Image Processing 26 (9): pp. 4509-4522, 2017.

[12] P. Cheremkhin et al., "Machine learning methods for digital holography and diffractive optics," Procedia Computer Science 169, 2020, pp. 440-444.

[13] K. Sun et al., "Deep High-Resolution Representation Learning for Human Pose Estimation," IEEE Conf. comput. Vis. Pattern, Recogn. (CVPR), 2019.

[14] S. Kreiss et al., "PifPaf: Composite Fields for Human Pose Estimation," IEEE Conf. comput. Vis. Pattern, Recogn. (CVPR), 2019.

[15] B. Cheng et al., "HigherHRNet: Scale-Aware Representation Learning for Bottom-Up Human Pose Estimation," IEEE Conf. comput. Vis. Pattern, Recogn. (CVPR), 2020.

[16] H. Joo et al., "Total Capture: A 3D Deformation Model for Tracking Faces, Hands, and Bodies," IEEE Conf. comput. Vis. Pattern, Recogn. (CVPR), 2018.

[17] <https://cocodataset.org/#home>

[18] <http://human-pose.mpi-inf.mpg.de/>

[19] J. Y. Chang, "DR-Net: denoising and reconstruction network for 3D human pose estimation from monocular RGB videos," Electronics Letters, Volume: 54, Issue: 2, 2018, pp. 70-72.

[20] Y. Umuroglu et al., "Finn: A framework for fast, scalable binarized neural network inference," Proceedings of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays. 2017.

[21] Y. Umuroglu et al., "Bismo: A scalable bit-serial matrix multiplication overlay for reconfigurable computing," 2018 28th International Conference on Field Programmable Logic and Applications (FPL). IEEE, 2018.

[22] S. Lee et al., "DeepMobileAR: A Mobile Augmented Reality Application with integrated Visual SLAM and Object Detection," SIGGRAPH Asia 2019 Posters, November 2019.

- [23] John Paul Titlow, "YouTube is using AI to police copyright—to the tune of \$2 billion in payouts," Fastcompany, 2016, <https://www.fastcompany.com/4013603/youtube-is-using-ai-to-police-copyright-to-the-tune-of-2-billion-in-payouts>
- [24] John Moulding, "Anti-piracy solution harnesses machine learning for automatic logo and face recognition," Videonet, 2018, <https://www.v-net.tv/2018/01/11/anti-piracy-solution-harnesses-machine-learning-for-automatic-logo-and-face-recognition/>
- [25] IVAN MEHTA, "Google improves its song recognition service by using AI from the Pixel 2," The Next Web, 2018, <https://thenextweb.com/artificial-intelligence/2018/09/18/pixel-2s-song-recognition-tech-in-google-search/>
- [26] Fraunhofer IDMT, "Music Plagiarism Detection," <https://www.idmt.fraunhofer.de/en/institute/projects-products/music-plagiarism-detection.html>
- [27] Matthew Durborow, "How is Artificial Intelligence Improving the NFL?" 2019, <https://www.rebellionresearch.com/blog/how-is-artificial-intelligence-improving-the-nfl>
- [28] P. Power et al., "Mythbusting Set-Pieces in Soccer," MIT Sloan Sports Analytics Conference, 2018.
- [29] P. Felsen et al., "Where will they go? Predicting fine-grained adversarial multi-agent motion using conditional variational autoencoders," In Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 732-747.
- [30] P. Felsen and P. Lucey, "Body Shots: Analyzing Shooting Styles in the NBA using Body Pose," MIT Sloan Sports Analytics Conference, 2017.
- [31] S. Hong and M. Kim, "Time of Flight 카메라 기반 음악출넘기 동작분석 방법," 한국통신학회, 2019.
- [32] W. Kim and M. Kim, "On-Line Detection and Segmentation of Sports Motions Using a Wearable Sensor," Sensors, March, 2018.
- [33] https://www.wimbledon.com/en_GB/aboutwimbledon/wimbledons_cognitive_highlights.html
- [34] <http://blog.daum.net/kcc1335/5022>
- [35] 장연우, "성장하는 일본 AI 시장...소매업 현장에서 AI 활용으로 수익 창출," 뉴스비전, 2017, <http://www.nvp.co.kr/news/articleView.html?idxno=121941>
- [36] <https://blogsaskorea.com/134>
- [37] Junde Yu, "잠재고객 예측 세분화: 타깃 고객층을 알아내는 지름길," Mobsinside, 2018, <https://brunch.co.kr/@mobsinside/1001>
- [38] 한국인터넷진흥원, 해외 개인정보보호 동향 보고서, 2019. 05
- [39] 정보통신전략위원회, 5G 시대 선도를 위한 실감콘텐츠산업 활성화 전략(19~23), 2019. 10.