논문 2024-4-28 http://dx.doi.org/10.29056/jsav.2024.12.28

애니메이션 제작과 생성형 인공지능의 실용적 활용에 관한 연구

박헌진*†

A Study on the Practical Application of Generative AI in Animation Production

Hunjin Park*†

요 약

생성형 인공지능은 멀티미디어 콘텐츠 제작 분야에서 급속한 발전을 이루며, 특히 동영상 생성 기술이 OpenAI의 Sora, Runway, Kling 등을 통해 획기적으로 진보하고 있다. 이러한 기술은 프롬프트와 참고 이미지만으로도 영화나 극장용 애니메이션에 준하는 고품질 영상을 생성할 수 있으며, 이는 게임 엔진이나 VFX 제작 환경의 미래 발전 방향에도 상당한 영향을 미칠 것으로 전망된다. 그러나 현재 생성형 인공지능이 보여주는 뛰어난 시각적 퀄리티에도 불구하고, 특히 캐릭터 애니메이션 분야에서는 스토리텔링을 지원하는 맥락 있는 제스처와 포즈 구현에 있어 분명한 한계를 보이고 있다. 본 연구는 애니메이션 제작에서 생성형 인공지능의 현재 활용 방식의 비효율성을 분석하고, 실질적인 활용 방안을 모색하여 효과적인 개발 방향을 제시하는 것을 목적으로 한다. 이를 위해 인공지능 생성 동영상과 실제 극장용 애니메이션의 제작 과정 및 원리를 비교 분석하고, 특히 캐릭터 움직임의 한계점을 중점적으로 고찰한다.

Abstract

This study examines the evolving role of generative AI in animation production, focusing on recent advancements in video generation technology through platforms like OpenAI's Sora, Runway, and Kling. While these technologies demonstrate impressive capabilities in producing high-quality visuals comparable to theatrical animations, they show significant limitations in character animation, particularly in creating contextually appropriate movements that support storytelling. Through comparative analysis of AI-generated videos and traditional animation production processes, this research identifies current inefficiencies in generative AI applications and proposes practical development directions for more effective integration in animation production workflows.

한글키워드 : 시각콘텐츠, 생성형 인공지능, 애니메이션, 게임아트, 일러스트레이션

keywords : Visual contents, Generative AI, Animation, Game art, Illustraion

1. 서 론

멀티미디어 콘텐츠 창작 분야에서 생성형 인공지능은 이제 인간의 창작을 보조를 할 수 있는

* 중부대학교 문화콘텐츠학부
† 교신저자: 박헌진(email: hunjinn@gmail.com)
접수일자: 2024.12.03. 심사완료: 2024.12.13.
게재확정: 2024.12.20.

역할로 점차 자리매김 해 나가고 있다. 최근에는 이미지 생성을 넘어 동영상 생성까지 급속히 기술이 발전하고 있으며 OpenAI의 Sora 발표 이후 많은 기업들이 경쟁적으로 동영상 제작을 위한 생성형 AI 기술을 발표하고 있다. 현재는 대표적으로 두각을 나타내는 Runway와 Kling을 필두로 많은 AI기술기반의 기업들이 영상 콘텐츠 제작 산업 지형을 변화 시키고 있다. 이러한 추세는 향후 멀티미디어 콘텐츠 제작 방식과 결과물에 혁명적인 변화를 가져올 것으로 전망된다.

동영상은 정지된 이미지와 근본적으로 차별화되는 매체로, 시간의 흐름에 따라 연속적으로 변화하는 시각적 이미지와 청각적 사운드를 유기적으로 통합하여 수용자에게 다차원적이고 총체적인 지각 경험을 제공한다. 이 매체는 단순한 시각적 정보 전달을 넘어 시간성, 움직임, 음향의 상호작용을 통해 감각적, 감정적, 인지적 차원에서 풍부하고 입체적인 의미를 구성하며, 관객으로 하여금 더욱 몰입적이고 지적인 커뮤니케이션 경험을 가능하게 한다.

생성형 인공지능이 동영상 제작 영역에 가져올 혁신적 잠재력은 기존 미디어 제작 패러다임의 근본적인 전환을 예고하고 있다. 이 기술은 사용자의 창의적 의도를 정교하게 번역할 수 있는 능력을 통해, 스타일라이즈드(Stylized)된 비주얼부터 사실적 재현까지 영화나 극장용 애니메이션 수준의 시각적 품질을 구현할 수 있는 가능성을 열어젖혔다. 특히 프롬프트와 참고 이미지를 기반으로 한 동영상 생성 기술은 단순한 기술적 진보를 넘어, Unreal, Unity와 같은 게임 엔진과 Renderman, Arnold 같은 첨단 렌더링 기술의 미래 발전 궤적을 재정의할 잠재력을 지니고 있다. 이는 미디어 제작의 민주화와 창작 과정의 근본적인 재구조화를 의미하며, 기존 시각 미디어 산업의 제작 워크플로우와 창작 방식에 패러다임적 전환을 촉발할 것으로 전망된다.

현재의 생성형 인공지능 기술은 비주얼 영역에서 놀라운 잠재력을 보여주고 있음에도 불구하고, 애니메이션 제작의 본질적이고 핵심적인 요소에서 중대한 한계를 드러내고 있다. 특히 캐릭터 애니메이션 분야에서 인공지능 기술의 실질적 활용을 위해서는 단순한 시각적 구현을 넘어서 내러티브의 심층적 맥락을 반영할 수 있는 일관된 제스처(Gesture)와 포즈(Pose)의 구현이 필수적인데, 현재의 기술은 이러한 내러티브적 깊이와 맥락적 일관성을 확보하는 데 상당한 어려움을 겪고 있다.

이 연구의 핵심 목적은 애니메이션 제작 과정에서 생성형 인공지능을 활용하는 현재의 접근 방식이 내포하고 있는 구조적 비효율성을 분석하고, 이를 토대로 향후 인공지능 기술이 애니메이션 제작에 실질적으로 기여할 수 있는 구체적인 개발 방향을 제시하는 데 있다.

본 연구는 기존 문헌 분석과 함께, 현재까지 발표된 생성형 인공지능 영상들을 분석하여 인공지능 기반 애니메이션 영상 생성 기술이 실제 애니메이션 제작 과정에 어떻게 실질적으로 기여할 수 있을지 그 가능성을 도출하고자 한다. 생성형 인공지능이 만들어내는 동영상은 관점에 따라 특수효과, 실사 합성, 3D 모델링 생성 등 다양한 측면에서 접근할 수 있으며, 비주얼적인 측면에서도 극사실적인 표현부터 실험적이거나 과장된 스타일까지 폭넓은 스펙트럼을 보여준다. 그러나 본 연구에서는 캐릭터 애니메이션의 핵심 원칙과 제작 과정에서 필수적인 절차들을 중심으로 연구 범위를 한정하여 고찰할 것이다.

2. 키프레임 애니메이션 제작 방식과 생성형 인공 지능

2.1 애니메이션 제작방식의 차이

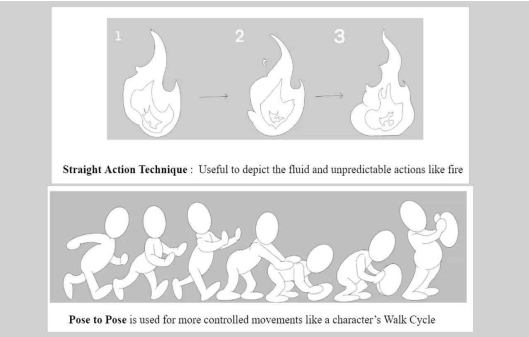

그림 1. 포즈 투 포즈와 스트레이트 어헤드
Fig. 1. Pose to pose and Straight ahead

애니메이션 기술의 초기 단계에는 체계적인 제작 방법론이 부재했기에, 제작자들은 수많은 시행착오를 거치며 효율적인 애니메이션 제작을 위한 지침들을 정립해 나갔다. 그 결과 탄생한 것이 바로 '애니메이션 12원칙'[1]과 같은 일련의 가이드라인이다. 이러한 지침들은 다수의 제작진이 참여하는 협업 환경에서 일관성 있게 현실적이고 자연스러운 애니메이션을 구현하기 위한 필수적인 기반이 되었다.

그림 1과 같이 키프레임(Keyframe) 애니메이션 제작 방식은 크게 두 가지 흐름[2]으로 나뉜다. 첫째는 명확한 계획 아래 주요 포즈를 설정하고 그 사이를 채워나가는 'Pose to Pose' 방식이며, 둘째는 즉흥적으로 다음 포즈를 결정하며 흐름을 만들어가는 'Straight Ahead' 방식이다. 이 두 방식은 모두 인간의 근원적인 창작 방식에서 비롯된 자연스러운 창작 방법론이라 할 수 있다. 'Key Pose'는 이러한 애니메이션 제작 과정에서 캐릭터의 움직임을 본질적으로 표현하는 핵심적인 요소로서, 애니메이션의 뼈대를 구성하는 역할을 수행한다. 전통 애니메이션의 관점에서 보면, 현대적인 3D 애니메이션 역시 동일한 원리를 기반으로 발전해 왔다. 그림 2의 'Squash and Stretch', 'Overlapping Action and Follow Through'[3]와 같은 애니메이션 원칙들은 자연스럽고 부드러운 캐릭터 동작을 표현하기 위해 대상을 면밀히 관찰하고, 이를 디자인 원리로 재구성하는 데 필수적인 요소라고 할 수 있다. 이러한 원칙들은 애니메이션 제작자들이 캐릭터의 움직임을 보다 생동감 넘치고 현실감 있게 구현할 수 있도록 안내하는 핵심적인 지침이 된다.

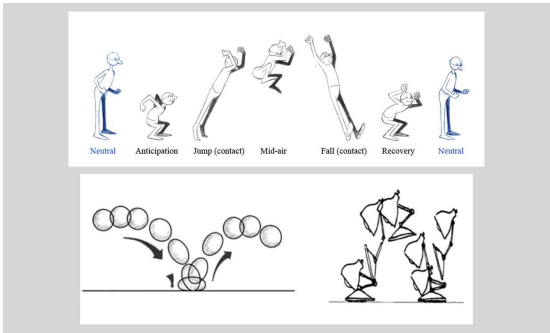

그림 2. 스쿼시와 스트레치 원리
Fig. 2. Squash and stretch

인공지능으로 만들어지는 영상은 일반적으로 두 가지 인풋[4] 방식을 사용한다. 하나는 텍스트 프롬프트를 기반으로 하는 Text-to-Video 방식이고, 다른 하나는 이미지와 프롬프트를 결합한 Image-to-Video 방식이다. 이러한 창작방식은 영상 제작의 패러다임을 바꾸고 있지만, 여전히 몇 가지 한계점을 지니고 있다. 먼저, 원하는 결과물을 얻기까지 여러 차례의 시행착오가 필요하다. 그림 3과 같이 특히 영상의 길이가 길어질수록 의도하지 않은 기괴한 요소[5]가 등장하거나 원래의 의도에서 벗어나는 경향이 있다. 또한, 각기 다른 학습 모델을 사용할 때마다 상이한 결과물이 나오기 때문에 콘텐츠의 일관성[6]을 유지하는 데 어려움을 가지고 있다. 이러한 이유로, 현재 생성형 인공지능이 만들어내는 애니메이션 영상은 장면 간 연결성이 크게 중요하지 않은 프로젝트에 활용될 때 가장 효과적이다.

반면, 전통적인 키프레임 애니메이션 제작 방

식[7]은 Pose to pose 접근법을 사용한다. 이 방식의 가장 큰 장점은 제작 과정에서 결과물을 예측할 수 있다는 점이다. 이는 협업 환경에서 매우 중요한 요소로 작용한다. 애니메이션 아티스트와 감독이 중간 단계의 결과물을 검토하고 서로 피드백을 주고받으며 애니메이션의 품질을 지속적으로 관리할 수 있기 때문이다.

그림 3. Sora에 의해 생성된 체조 선수 영상
Fig. 3. Sora generated video of gymnastics

애니메이션 품질을 관리하기 위해 중간 과정을 확인할 수 있다는 점은 애니메이션 영상 제작에 있어 핵심적인 요소이다. 현재까지 개발된 인공지능 기술은 이와는 대조적으로, 중간 과정 없이 마치 사진을 찍듯이 최종 결과물을 한 번에 생성해낸다. 이로 인해 창작 과정에서 주로 경험하게 되는 두 가지 모습이 있다. 하나는 만족스러운 결과물이 나올 때까지 프롬프트와 이미지를 반복적으로 조정하는 것이고, 다른 하나는 어느 정도 이야기적 맥락이나 컨셉이 구현되면 그 상태에서 작업을 마무리하는 것이다.

2.2 절차적 제작과정과 스토리텔링

애니메이션 영상 제작은 체계적이고 절차적인 과정[8]을 통해 이루어진다. 이러한 접근 방식은 단순히 효율성을 높이는 것을 넘어서, 여러 애니메이션 아티스트들이 개인의 취향에 매몰되지 않고 공동의 비전을 향해 창의성을 발휘할 수 있도록 하는 데 그 목적이 있다. 이 과정은 크게 두 단계로 구분할 수 있다. 첫번째는 Blocking 단계로써, 그림 4와 같이 이 단계에서는 기본적인 키포즈와 타이밍을 설정한다. 이를 통해 전체적인 움직임의 흐름과 리듬을 잡고, 최종 결과물을 어느 정도 예측할 수 있게 된다.

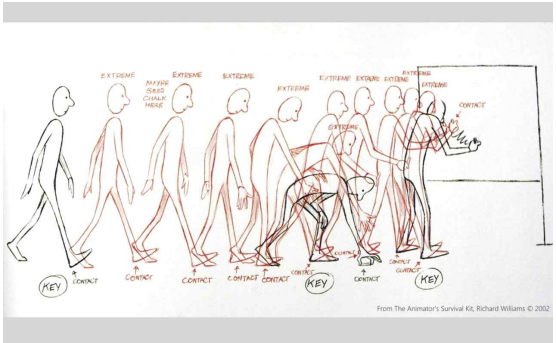

그림 4. 블로킹 단계의 키포즈
Fig. 4. Key poses in the blocking stage

이는 마치 건축에서의 기초 공사와 같이, 전체 작품의 뼈대를 세우는 중요한 과정이다. 두 번째로, Polishing 단계는 Blocking을 통해 구성원들이 공동의 비전을 충분히 공유하고 나면, 이 단계로 넘어간다. 여기서는 앞서 잡아놓은 키포즈 사이에 인비트윈 디테일을 더해 애니메이션을 고품질로 완성시킨다. 이 과정에서 캐릭터의 표정, 미세한 동작, 특수 효과 등이 추가되어 작품에 생동감을 불어넣는다. 물론 제작사마다 이러한 단계를 세분화하거나 변형하여 사용하기도 한다. 그러나 대부분의 경우, 이와 같은 절차적 접근을 통해 애니메이션 영상을 완성해 나간다. 이러한 방식은 대규모 프로젝트에서 특히 중요한데, 여러 팀원들의 작업을 조율하고 일관된 퀄리티를 유지하는 데 도움이 되기 때문이다.

이러한 단계별 접근법은 단순히 작업의 효율성만을 높이는 것이 아니다. 각 단계마다 품질을

점검하고 필요한 수정을 가할 수 있어, 최종 결과물의 완성도를 크게 높일 수 있다. 또한 이 과정을 통해 아티스트들은 개인의 창의성을 발휘하면서도 전체적인 작품의 방향성을 잃지 않을 수 있다. 결과적으로 이러한 제작 과정은 높은 품질의 애니메이션을 일관되게 제작할 수 있게 하는 핵심적인 요소라고 할 수 있다.

캐릭터 애니메이션[9]과 스토리텔링이 중심이 되는 애니메이션 영상은 본질적으로 심층적인 내러티브 구조를 내재하고 있다. 이러한 매체적 특성으로 인해 연속된 샷들 간의 유기적 연결성과 내적 일관성은 극히 중요한 요소로 작용한다. 특히 자연스러운 카메라 앵글과 장면 간 맥락적 전환은 관객의 정서적 몰입을 결정짓는 핵심 메커니즘이 된다. 애니메이션 제작 과정에서 상당한 시간과 노력이 투입되는 영역은 단순히 캐릭터의 물리적 액션이나 감정 표현을 넘어, 장면들 사이의 유연하고 설득력 있는 연출에 있다. 이는 내러티브의 흐름을 자연스럽게 유지하고 관객의 감정적 지속성을 확보하기 위한 필수적인 창작 과정이다. 현재의 생성형 인공지능 기술은 이러한 복합적이고 섬세한 내러티브 연결성을 수정하고 대응하는 데 있어 근본적인 한계를 보이고 있다. 따라서 현 단계의 생성형 인공지능이 생산하는 영상물은 주로 단발적이고 고립된 장면의 시각적 구현에 치중하는 양상을 보인다.

3. 애니메이션 제작에서 생성형 인공지능의 활용과 전망

3.1 사용자 편의 중심의 기능 강화

최근 생성형 인공지능의 애니메이션 생성 기술은 전통적인 컴퓨터 그래픽스 소프트웨어의 워크플로우와 인터페이스 디자인[10] 철학을 적극적으로 수용하며 실무적 활용도를 제고하고 있다. 특히 Stable diffusion의 CompyUI가 도입한 노드 기반 시스템은 그림 5에서 예시된 바와 같이 비주얼 스크립팅 방식을 채택함으로써, 애니메이션 아티스트들의 직관적 접근과 섬세한 제어를 가능하게 하는 혁신적인 인터페이스를 제공하고 있다. 이러한 노드 기반의 비주얼 스크립팅 인터페이스는 두 가지 중요한 의의를 지닌다. 첫째, 기술적 진입 장벽을 낮춤으로써 창작자들의 접근성을 획기적으로 향상시켰으며, 둘째, 이미지 생성 과정에 대한 정교한 제어를 요구하는 전문가들의 실무적 욕구를 효과적으로 충족시키고 있다. 이러한 발전 양상은 생성형 인공지능 기술이 독자적인 발전 경로를 고수하기보다, 기존 산업 표준 소프트웨어의 사용자 경험과 인터페이스 디자인을 전략적으로 참조하며 실용적 가치를 극대화하는 방향으로 진화하고 있음을 시사한다. 이는 기술 혁신과 실무적 적용 사이의 간극을 좁히는 의미 있는 진전으로 평가될 수 있다.

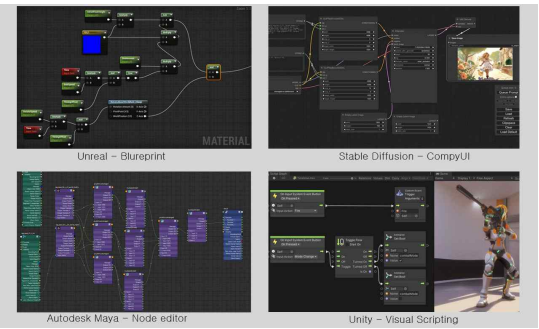

그림 5. 다양한 응용 프로그램에서의 시각적 노드 기반 프로그래밍
Fig. 5. Visual Node-Based Programming Across Different Applications

Runway가 최근 프로토타입으로 공개한 혁신적 기능은 전통적인 키프레임 애니메이션의 핵심 원리를 생성형 인공지능과 융합한 주목할 만한 시도를 보여준다. 그림 6에서 예시된 바와 같이, 이 시스템은 아티스트가 지정한 키포즈 이미지들

사이의 중간 프레임을 인공지능이 자동으로 생성함으로써, 현재 생성형 인공지능이 가지고 있는 문제점을 극복할 수 있도록 도와준다. 이러한 접근 방식은 캐릭터 애니메이션 제작 패러다임의 근본적인 변화를 촉발할 수 있는 혁신적 기술 발전 방향으로 평가된다.

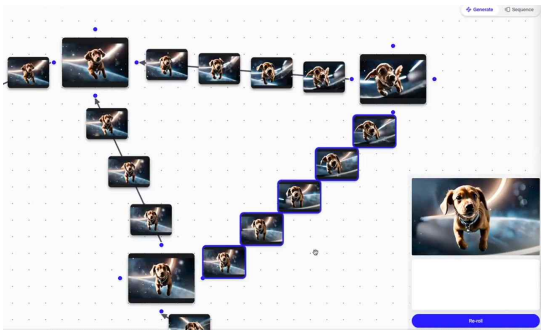

그림 6. 런웨이 AI의 비디오 키프레이밍 프로토타입
Fig. 6. Runway AI video keyframing prototype

또 다른 주목할 만한 사례로, 모션브러쉬 기술을 들 수 있다. 그림 7에서 확인할 수 있듯이, 이 기술은 이미지 내 객체의 이동 경로를 직관적으로 스케치하면 해당 궤적을 따라 자연스러운 애니메이션을 자동 생성하는 기능을 제공한다. 이러한 혁신적 도구들은 특히 VFX 분야에서 실무

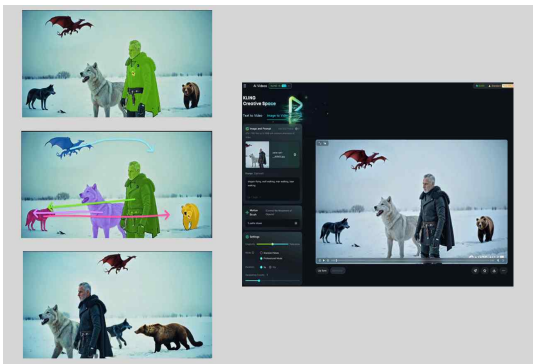

그림 7. 클링AI의 모션브러쉬
Fig. 7. Kling AI motion brush

적 활용도가 높을 것으로 전망되며, 향후 시각효과 제작 워크플로우의 효율성을 획기적으로 향상시킬 잠재력을 지니고 있다.

3.2 애니메이션 영상 분야의 활용 전망

심미적 가치 판단과 예술적 창의성은 본질적으로 인간 고유의 영역으로 남아있다. 특히 스토리텔링 기반의 캐릭터 애니메이션에서는 클리셰를 전략적으로 활용하면서도, 까다로운 관객의 기대를 충족시킬 수 있는 참신하고 혁신적인 창작적 발상이 요구된다. 현재의 생성형 인공지능은 내러티브의 맥락에 기반한 유의미하고 자연스러운 결과물을 도출하기보다는, 학습된 데이터의 기계적 재생산에 머물러 있는 실정이다. 이는 논리적이고 풍부한 언어 구사력을 보여주는 현세대 생성형 인공지능의 언어적 능력과 현저한 대조를 이루는 지점이다.

전통적인 컴퓨터 그래픽스 소프트웨어 기반의 제작 환경에 익숙한 창작자들에게 있어, 생성형 인공지능을 통한 애니메이션 제작 방식은 상당한 괴리감을 야기한다. 특히 고품질 키프레임 애니메이션이 추구하는 섬세한 심미적 요소들을 충실히 구현하지 못하는 한계를 보인다. 예를 들어, 숙련된 애니메이션 아티스트는 캐릭터의 무게감을 더욱 극적이고 설득력 있게 표현할 수 있는 반면, 현재의 인공지능은 이러한 미학적 섬세함을 구현하는 데 한계를 보인다. 따라서 생성형 인공지능의 실질적인 애니메이션 제작 활용을 위해서는, 완성된 애니메이션을 일괄 생성하는 방식보다는 기존 애니메이션 제작 프로세스 내에서 보조적 도구로서의 역할에 충실할 필요가 있다. 특히 반복적이고 시간 소모적인 작업들을 효율적으로 처리하는 방향으로 발전해야 할 것이다.

이러한 맥락에서, 생성형 인공지능이 애니메이션 제작의 각 단계에서 수행할 수 있는 이상적인 보조적 역할들을 제시할 수 있다. 구체적으로,

Timing과 Pose, Arc, Anticipation, Overlapping, Spacing, Overshoot과 Settle 등 기본적인 애니메이션 원리들을 적용하는 과정에서 생산성 향상을 위한 지원 도구로서 기능하는 것이 바람직하다. 이러한 접근은 인공지능 기술의 한계를 인정하면서도, 그 실용적 활용 가능성을 극대화하는 균형 잡힌 방향성이라 할 수 있다.

4. 결론

생성형 인공지능의 실무적 활용 가능성에 대한 지속적인 연구 개발을 통해, 그 실질적 응용 방안이 점차 구체화되고 있다. 특히 어도비사가 최근 공개한 스프라이트 이미지 생성 기술은 이러한 진전을 상징적으로 보여주는 획기적 사례로 평가된다. 이 기술은 기초 디자인을 토대로 캐릭터의 다양한 각도와 포즈를 인공지능이 자동 생성하는 기능을 제공하며, 게임 개발과 애니메이션 제작 현장에서 즉시 활용 가능한 수준의 실용성을 확보했다는 점에서 주목할 만하다.

이러한 발전은 그동안 생성형 인공지능의 주요 난제로 지적되어온 이미지의 일관성 유지 문제에 대한 의미 있는 해결책을 제시한다. 특히 기존 컴퓨터 그래픽스 소프트웨어의 인터페이스와 워크플로우를 전략적으로 접목함으로써, 창작자들이 보다 정교하게 결과물을 제어할 수 있는 환경을 구축해나가고 있다.

이는 기존의 일방향적이고 제한적이었던 생성형 인공지능의 운용 방식에서 탈피하여, 창작자의 의도를 섬세하게 반영할 수 있는 양방향적 제작 환경으로의 진화를 의미한다. 결과적으로 전통적인 그래픽스 소프트웨어의 강점과 생성형 인공지능의 혁신적 기능이 유기적으로 결합되면서, 이미지의 세부적 요소에 대한 정밀한 제어가 가능한 새로운 창작 패러다임이 형성되고 있다고 볼 수 있다.

감사의 글
이 논문은 2024년도 중부대학교 학술연구비 지원에 의하여 이루어진 것임

참 고 문 헌

[1] Ollie Johnston, and Frank Thomas, “The illusion of life: Disney animation”, New York: Hyperion, 1995. ISBN: 9780786860708
[2] Eric Goldberg, “Character Animation Crash Course”, Silman-James Press, 2008. ISBN: 9781879505971
[3] Harold Whitaker, “Timing for Animation”, Focal Press, 2022. ISBN: 9780240517148
[4] S. Vinothkumar, S. Varadhaganapathy, R. Shanthakumari, S. Dhanushya, S. Guhan and P. Krisvanth. (2024) Utilizing Generative AI for Text-to-Image Generation, 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 1-6, DOI: 10.1109/ICCCNT61001.2024.10725454
[5] Anthropllogy news, Digital Symptoms, https://www.anthropology-news.org/articles/digit-al-symptoms/ Jun, 2024
[6] Forbes, The Problem With AI-Generated Art, Explained, https://www.forbes.com/sites/danidiplacido/2023/12/30/ai-generated-art-was-a-mistake-and-heres-why/ Dec, 2023
[7] George Maestri, Digital character animation, New Riders, 2006. ISBN: 9781562055592
[8] Andy Beane, 3D animation essentials. John Wiley & Sons, 2012. ISBN: 9781118147481
[9] N Sultana, P Lim Yan, M Nico. (2013 Sep). 2013 International Conference on Informatics and Creative Multimedia, Kuala Lumpur, Malaysia, 321-324. DOI: 10.1109/ICICM.2013.69

[10] H. Nishino, S Sueyasu, T Kagawa and K Utsumiya. (2011). A CG Learning Tool with Guided Programming Environment, 14th International Conference on Network-Based Information Systems, Tirana, Albania, 462-467, DOI: 10.1109/NBiS.2011.77

저 자 소 개

박헌진(Hunjin Park)

2002.2 홍익대학교 미술대학 졸업
2007.2 Academy of Art University 석사
2016.6-현재 : 중부대학교 만화애니메이션 전공 교수
<주관심분야> 3D 애니메이션, 애니메이션 제작, 문화콘텐츠