JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 27, No. 10, pp. 3075-3087, Oct. 2026

Check for updates

생성형 AI 도구 연계 제작 파이프라인에서 스톱모션 스타일의 일관성 형성 과정 연구

최 현 주[1] · 옥 창 엽[2*]
[1]중앙대학교 첨단영상대학원 영상학과 석사과정
[2]중앙대학교 예술공학부 조교수

Study on Formation of Stop-Motion Style Consistency in Production Pipeline Linking Generative AI Tools

Hyun-Joo Choi[1] · Chang-Yeob Ok[2*]

[1]Master's Course, Department of Entertainment Technology, GSAIM, Chung-Ang University, Seoul 06974, Korea
[2]Assistant Professor, College of Art & Technology, Chung-Ang University, Anseong 17546, Korea

[요 약]

본 연구는 복수의 생성형 AI 도구를 연계하여 제작한 스톱모션 스타일 애니메이션 <엉뚱한 엉두>(2026)를 사례로, 시각적·시간적 일관성의 형성·조정 과정과 감독의 조정 기능을 질적 단일 사례연구를 통해 분석하였다. 분석 결과, 시각적 일관성은 캐릭터의 식별 요소와 클레이·펠트 재질, 색채와 조명 등의 기준을 유지하면서 장면별 변화를 허용하는 방식으로 형성되었다. 시간적 일관성은 시작·끝 프레임으로 움직임의 방향과 범위를 제한하고, 비의도적 형태 변화는 생성 오류로 판단하여 선별·재생성하는 한편, 분절된 운동 리듬은 후반 편집에서 의도적으로 부여하는 방식으로 조정되었다. 감독은 스타일 기준을 설정하고 생성 결과를 선별·수정하며 서로 다른 도구의 산출물을 연결하는 통합 조정 역할을 수행하였으며, 이를 통해 스타일 일관성은 개별 생성 모델의 성능만으로 확보되는 것이 아니라, 제작 단계별 기준 설정과 판단·조정을 통해 형성됨을 확인하였다.

[Abstract]

This study investigates the formation and adjustment of visual and temporal consistency as well as the director's coordinating role through a qualitative case study of Eongttunghan Eongdu (2026), a stop-motion style animation produced with multiple generative AI tools. The results show that visual consistency was achieved by maintaining key criteria—such as character-identifying features, clay and felt material qualities, color, and lighting—while allowing scene-level variation. Temporal consistency was managed by constraining movement through start and end frames, treating unintended deformation as a generation error subject to selection and regeneration, and constructing a segmented motion rhythm in post-production. The director established style criteria, selected and modified generated outputs, and connected outputs across tools, thereby demonstrating that style consistency cannot be achieved solely through individual model performance, but also requires stage-specific criteria and repeated human judgment and adjustments.

색인어 : 생성형 AI, 스톱모션 애니메이션, 스타일 일관성, 제작 파이프라인, 인간-AI 협업
Keyword : Generative AI, Stop-Motion Animation, Style Consistency, Production Pipeline, Human-AI Collaboration

http://dx.doi.org/10.9728/dcs.2026.27.10.3075

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-CommercialLicense(http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 05 August 2026; **Revised** 27 August 2026
Accepted 31 August 2026

***Corresponding Author; Chang-Yeob Ok**

Tel: +82-31-670-3196
E-mail: cyok@cau.ac.kr

Ⅰ. 서 론

1-1 연구 배경 및 문제 제기

생성형 AI 기술의 발전으로 시나리오, 스토리보드, 콘셉트 아트, 애니메이팅, 사운드, 합성 등 애니메이션 제작의 각 단계에서 AI 도구가 활용되고 있으며, 이를 제작 공정에 적용하려는 연구도 다양하게 시도되고 있다. 오동일[1]과 한혜진[2]의 연구는 생성형 AI가 애니메이션의 제작 영역을 확장하고 제작 과정의 효율성과 활용 가능성을 높일 수 있음을 보여준다. 이러한 선행연구를 통해 생성형 AI의 애니메이션 제작 적용 가능성은 점차 구체화되고 있으나, 복수의 생성형 AI 도구가 연계되는 과정에서 작품의 스타일이 어떻게 지속되고 조정되는지에 관한 논의는 상대적으로 부족하다. 본 연구는 개별 생성 도구의 성능이나 활용 가능성에 초점을 두기보다, 서로 다른 도구의 산출물이 연속적으로 연결되는 제작 공정을 하나의 분석 단위로 보고, 그 과정에서 시각적·시간적 일관성이 어떻게 형성되는지를 분석한다. 이를 통해 생성형 AI 기반 애니메이션의 스타일 일관성을 기술적 성능의 문제에 한정하지 않고, 제작 단계마다 이루어지는 인간의 판단과 조정이 누적되어 형성되는 과정적 문제로 설명한다는 데 학문적 의의가 있다.

본 연구에서는 스타일 일관성을 시각적 일관성과 시간적 일관성을 포괄하는 개념으로 정의한다. 시각적 일관성은 캐릭터와 배경의 형태, 재질, 색채, 조명 및 공간감이 작품의 기준 안에서 유지되는 상태를, 시간적 일관성은 프레임 사이와 개별 생성된 컷들 사이에서 형태와 동작이 안정적으로 이어지는 상태를 의미한다. 특히 스톱모션은 실제 오브젝트와 재료의 물질성이 시각적 표현에 관여하며[3], 프레임별로 설정된 움직임의 간격과 타이밍에 따라 운동감이 구성되므로[4], 시각적·시간적 일관성을 함께 검토할 필요가 있다.

그러나 생성형 AI 기반 영상 제작에서는 정지 이미지에서 형성된 캐릭터와 배경의 특성이 합성 및 영상화 과정에서 자동으로 유지되는 것은 아니다. 프레임 사이에서 캐릭터의 외형이나 배경이 비의도적으로 변형될 수 있으며, 생성형 영상 연구에서도 프레임 간 외형과 움직임을 일관되게 유지하는 문제는 주요 기술적 과제로 다뤄진다[5].

한편 스톱모션 특유의 분절된 리듬은 이와 성격이 다르다. 프레임 간 불안정성은 제거되어야 할 생성 오류인 반면, 분절된 운동 리듬은 의도적으로 구성되어야 할 표현 요소이기 때문이다. 다만 컷 단위로 개별 생성되는 제작 방식에서 이러한 분절을 어느 단계에서 어떤 방식으로 부여할 것인지는 자명하지 않다. 컷마다 분절의 정도가 달라지면 오히려 시간적 일관성이 저해될 수 있기 때문이다.

따라서 스톱모션 스타일의 일관성은 특정 생성 모델의 결과만으로 판단하기보다, 이미지 생성·합성 단계의 시각적 일관성과 영상 생성·후반 편집 단계의 시간적 일관성이 함께 확보되는지를 전체 제작 공정 안에서 분석할 필요가 있다.

1-2 연구 목적

본 연구의 대상인 <엉뚱한 엉두>(2026)는 복수의 생성형 AI 도구를 연계하여 클레이 질감의 스톱모션 스타일을 구현한 단편 애니메이션이다. 작품 개요는 3-1에서, 구체적인 제작 공정은 3-2에서 기술한다.

본 연구는 <엉뚱한 엉두>의 제작 과정을 사례로 하여 스톱모션 스타일을 구성하는 시각적·시간적 요소가 도구 연계 제작 공정에서 어떻게 형성되고 유지·조정되는지, 그리고 그 과정에서 이루어진 조정 개입이 어떠한 기능을 수행하며 어떠한 기준에 근거하는지를 분석하는 데 목적이 있다. 본 연구에서는 스타일 기준을 설정하고 생성 결과의 채택 여부를 판단하며 단계 간 결과물을 연결하는 통합 조정 기능을 수행한 역할을 감독으로 지칭한다.

이에 본 연구는 다음의 연구문제를 중심으로 분석을 진행한다. 첫째, 이미지 생성 및 합성 단계에서 스톱모션 스타일의 시각적 일관성은 어떻게 형성·조정되며, 어떠한 변동이 나타나는가? 둘째, 영상 생성 및 후반 편집 단계에서 시간적 일관성과 스톱모션의 운동 리듬은 어떻게 형성·조정되며, 어떠한 한계가 나타나는가? 셋째, 위 두 층위의 일관성이 형성되는 과정에서 감독은 어떤 조정 기능을 수행하며, 그 판단은 어떠한 기준에 근거하는가?

1-3 연구 방법 및 범위

본 연구는 문헌연구와 질적 단일 사례연구를 병행한다[6]. 생성형 AI를 활용한 애니메이션 제작은 사용 도구의 구성과 제작 조건이 사례마다 달라 표준화된 비교가 어렵고, 제작 과정에서 이루어진 판단의 근거는 완성된 결과물만으로 확인되지 않기 때문이다. 선행연구 검토를 통해 분석 기준을 도출한 뒤 <엉뚱한 엉두>의 제작 과정에 이를 적용하며, 구체적인 분석 대상과 기준은 3-3에서 제시한다.

본 연구자는 <엉뚱한 엉두>의 제작자 겸 감독으로 직접 참여하였다. 이는 제작 과정의 내부 맥락에 접근할 수 있다는 이점과 분석 관점의 제약을 함께 가지며, 이에 관한 논의는 3-4에서 다룬다.

연구 범위는 <엉뚱한 엉두> 한 작품과 해당 작품의 제작 과정으로 한정한다. 개별 AI 도구의 성능 비교와 관객 반응 평가는 연구 범위에서 제외한다.

Ⅱ. 이론적 배경 및 선행연구 검토

본 장에서는 생성형 AI 기반 스톱모션 스타일 애니메이션의 일관성 형성 과정을 분석하기 위한 기준을 선행연구 검토를 통해 도출한다. 먼저 스톱모션 애니메이션의 미학적 특성을 검토하고, 생성형 AI 제작에서 제기되는 스톱모션 스타일

일관성의 문제와 제작 과정에서 인간의 개입이 수행하는 기능에 관한 논의를 살펴본다. 이를 바탕으로 선행연구의 한계를 정리하고 본 연구의 분석틀을 제시한다.

2-1 스톱모션 애니메이션의 미학적 특성

스톱모션 애니메이션은 모델, 퍼펫과 같은 물리적 대상을 프레임마다 조작하고 촬영하여 움직임을 구성하는 방식이다. 본 절에서는 이러한 제작 방식이 형성하는 미학적 특성을 재료성과 수공성, 프레임 단위 운동과 시간적 리듬으로 구분하여 살펴본다.

1) 재료성과 수공성

스톱모션의 재료성은 점토, 천, 목재 등 캐릭터와 오브젝트를 구성하는 실제 재료의 물리적 특성이 화면에 드러나는 데서 형성된다. 재료가 지닌 표면과 물리적 특성은 캐릭터와 배경의 외형뿐 아니라 작품의 정서와 표현에도 영향을 미친다. Paul Wells는 물질 기반 애니메이션에서 오브젝트의 형태와 재료가 조작 및 연출 방식과 결합하여 의미를 형성한다고 보았다[3].

재료성은 재료의 종류만으로 성립하지 않는다. 스톱모션에서는 애니메이터가 프레임마다 물리적 대상을 직접 조작하며, 이 과정에서 표면의 변화와 미세한 불균질성이 나타날 수 있다. Alice Gambrell은 스톱모션의 노동집약적 제작 과정에 주목하여 애니메이터의 손과 작업이 작품 안에서 드러나거나 감춰지는 방식을 분석한다[7]. 이러한 제작 흔적은 화면에 수작업적 인상을 형성하는 요소로 작용한다. 그림 1은 실제 스톱모션 제작에서 애니메이터가 클레이 퍼펫을 프레임 단위로 조작하는 과정을 보여준다.



그림 1. 클레이 퍼펫을 프레임 단위로 조작하는 스톱모션 제작 과정(출처: Aardman Animations, <Wallace & Gromit: Vengeance Most Fowl> Behind the Scenes, 2025)

Fig. 1. Frame-by-frame manipulation of a clay puppet in stop-motion production (Source: Aardman Animations, Behind the Scenes of Wallace & Gromit: Vengeance Most Fowl, 2025)

2) 프레임 단위 운동과 시간적 리듬

스톱모션의 움직임은 실제 대상의 운동을 실시간으로 연속 촬영한 것이 아니라, 정지된 오브젝트의 위치와 형태를 프레임마다 조정하여 촬영한 이미지의 연속으로 구성된다. 따라서 움직임의 성격은 애니메이터가 설정한 동작의 지속 시간과 프레임별 위치 변화에 따라 달라진다. Richard Williams는 동작에 할당된 시간을 의미하는 타이밍과 그 시간 안에서 대상의 위치를 배분하는 스페이싱을 구분하고, 이 두 요소가 움직임의 속도와 성격을 결정한다고 설명한다[4]. 스톱모션에서도 프레임 사이의 이동량과 자세 변화, 동일한 자세를 유지하는 프레임 수에 따라 움직임의 속도와 리듬이 달라진다.

2-2 생성형 AI 제작에서 스톱모션 스타일 일관성의 문제

1) 캐릭터 정체성과 시각적 일관성

생성형 AI로 동일한 캐릭터를 여러 이미지에 걸쳐 생성할 때에는 캐릭터의 정체성과 세부 특성을 일관되게 유지하는 문제가 제기된다[8],[9]. 동일한 프롬프트와 참조 이미지를 사용하더라도 캐릭터의 얼굴과 신체 비례, 의상과 세부 장식이 장면마다 달라질 수 있다.

물리적 스톱모션에서는 퍼펫과 세트가 실제 제작물로 존재하므로, 캐릭터의 형태와 재료적 특성을 비교할 수 있는 물리적 기준점이 유지된다. 다만 복수의 퍼펫과 교체 부품이 사용되거나, 장면별 조명 조건과 반복된 조작에 따른 표면 변화가 발생할 수 있으므로 장면 사이의 시각적 연속성은 별도로 관리되어야 한다.

반면 생성형 AI 기반 제작에서는 장면 사이에 지속되는 물리적 대상이 존재하지 않으므로, 캐릭터와 배경의 형태 및 재질적 특성을 참조 이미지, 프롬프트와 생성 결과의 선별을 통해 반복적으로 설정해야 한다. 이때 유지되어야 하는 것은 캐릭터의 형태와 색채만이 아니다. 본 연구가 다루는 스톱모션 스타일에서는 클레이·펠트의 재질적 외관, 미니어처형 공간의 조형과 조명, 캐릭터와 배경 사이의 공간적 정합성이 함께 유지되어야 한다. 또한 실제 조작 과정이 존재하지 않으므로, 표면의 미세한 요철이나 비대칭적 형태와 같이 수작업적 인상을 형성하는 특성 역시 프롬프트와 결과 선별을 통해 구현해야 한다.

따라서 본 연구의 시각적 일관성은 캐릭터의 형태와 색채뿐 아니라 이상에서 언급한 재질적 외관, 수작업적 표면 특성, 공간의 조형·조명과 정합성을 포함하는 개념으로 정의한다.

2) 프레임 간 형태와 시간적 일관성

기존의 스톱모션 제작 과정에서는 프레임마다 부여하는 대상의 이동량과 동일한 자세를 유지하는 프레임 수를 제작자가 직접 설정하므로 운동 리듬을 계획적으로 통제할 수 있다. 장면 사이의 리듬 일관성 역시 공통된 타이밍과 스페이싱 기준을 적용해야 확보되지만, 이 경우 설정한 기준은 촬영 과정에서 그대로 실현된다. 반면 생성형 AI로 제작할 경우에는 움

직임이 컷 단위로 생성되므로, 개별 컷 내부의 프레임 안정성과 독립적으로 생성된 컷들 사이의 속도 및 리듬 편차를 각각 확인해야 한다.

이 가운데 프레임 사이의 안정성은 생성 영상 평가 연구에서 주요 항목으로 다루어진다. VBench는 생성 영상의 평가 차원으로 피사체 일관성, 배경 일관성, 시간적 깜박임, 움직임의 부드러움과 동적 움직임의 정도 등을 구분하여 제시한다[5]. VBench의 모델 비교에서는 시간적 품질이 높은 모델이 낮은 동적 움직임을 보이는 경향이 관찰되어, 큰 움직임과 프레임 간 일관성을 함께 확보하는 일이 과제로 제시되었다[5]. 따라서 본 연구에서는 시간적 일관성을 움직임의 부드러움뿐 아니라, 동작이 진행되는 동안 캐릭터의 형태와 재질이 유지되는지를 함께 포괄하는 개념으로 본다. 다만 이러한 평가 기준은 하나의 생성 영상을 대상으로 하므로, 개별적으로 생성된 컷들 사이의 균일성은 별도의 검토를 필요로 한다.

3) 다중 도구 연계 과정의 스타일 변동

기존 생성형 이미지·영상 연구는 주로 개별 모델 내부에서 나타나는 캐릭터 및 프레임 일관성의 개선과 평가에 초점을 둔다. 반면 본 연구가 다루는 다중 도구 제작에서는 한 단계의 산출물이 다른 도구의 입력으로 사용된다. 이에 본 연구는 캐릭터 생성 단계에서 확보한 외형이 합성 과정에서 달라지거나, 정지 이미지의 재질과 조명이 영상 생성 과정에서 변형되는 현상을 다중 도구 연계 과정의 스타일 변동으로 설정한다.

따라서 다중 도구 제작의 스타일 일관성은 각 단계의 결과물 품질뿐 아니라, 이전 단계의 산출물을 참조 이미지나 시작·끝 프레임으로 전환하고, 변형된 결과를 선별·재생성·편집하는 단계 간 조정의 문제를 포함한다.

2-3 생성형 AI를 활용한 파이프라인에서 인간 개입의 기능

1) 스타일 기준 설정

인간-AI 공동창작 연구에서는 인간과 AI 사이의 주도권과 통제 수준이 창작 결과에 영향을 미치는 중요한 요소로 다뤄진다. Caterina Moruzzi와 Solange Margarido는 공동창작 과정에서 사용자가 AI의 역할과 개입 정도를 조정할 수 있어야 한다고 보았으며[10], Jeba Rezwana와 Mary Lou Maher는 인간과 AI의 기여 방식, 상호작용 순서 및 공유 결과물에 대한 조정이 공동창작 구조를 형성한다고 설명한다[11].

이러한 논의를 생성형 AI 애니메이션 제작에 적용하면, 인간 제작자가 작품의 재질, 색채, 조명, 캐릭터 비례와 움직임의 방향을 먼저 설정해야 생성 결과의 적합성을 판단할 수 있다. 따라서 스타일 기준은 단순한 프롬프트 문장이 아니라 생성 결과를 채택하거나 제외하는 판단의 기준으로 기능한다. 나아가 복수의 인원이 제작 단계를 분담하는 환경에서는, 스타일 기준이 인간과 AI 사이의 지시 기준을 넘어 서로 다른 단계를 담당하는 참여자들이 작업 방향을 공유하고 산출물을 평가·통합하기 위한 공통의 준거로도 기능한다.

2) 생성 결과의 선택·수정·통합

제작자는 생성형 AI와의 상호작용 과정에서 생성 결과를 수정·편집하며 결과를 구체화할 수 있다[12]. 국내에서도 생성형 AI 영상 제작 사례를 통해 창작자가 중심이 되어 결과물을 제어하는 활용 방안이 검토된 바 있다[13].

특히 다중 도구 제작에서는 각 단계의 산출물을 다음 단계의 입력 조건에 맞게 수정하고 연결하는 역할이 필요하다. 이때 인간 개입은 개별 결과의 오류를 수정하는 데 그치지 않고, 제작 단계 전반의 스타일을 통합하는 기능을 수행한다.

2-4 선행연구의 한계와 본 연구의 분석틀

앞서 검토한 선행연구는 스톱모션의 물질성과 수작업[3],[7], 생성형 이미지의 캐릭터 일관성과 생성 영상의 프레임 간 일관성[5],[8], 인간-AI 공동창작에서의 주도권과 상호작용[10],[11]을 각각 설명한다. 그러나 본 연구가 검토한 선행연구의 범위에서는 복수의 생성형 AI 도구와 팀원들의 분업이 결합된 애니메이션 제작에서 감독의 조정이 스톱모션 스타일의 시각적·시간적 일관성 형성에 어떻게 작용하는지를 전체 제작 공정 차원에서 분석한 논의는 제한적으로 나타난다.

이에 본 연구는 스톱모션 미학 자체를 포괄적으로 규명하기보다, 스톱모션의 시각적·운동적 특성이 생성형 AI 제작 공정에서 어떻게 구현되고 유지·조정되는지를 분석한다. 이를 위해 분석틀을 첫째, 캐릭터의 형태, 클레이·펠트 재질, 수작업적 인상을 형성하는 표면 특성, 색채, 조명 및 공간감에 관한 시각적 일관성, 둘째, 프레임 간 형태의 지속성, 동작의 연결, 컷 간 속도·리듬의 균일성과 본 작품이 지향한 분절된 운동감에 관한 시간적 일관성, 셋째, 스타일 기준 설정과 생성 결과의 선별·수정·통합에 관한 감독의 조정 개입으로 구성한다.

이러한 분석틀을 바탕으로 <엉뚱한 엉두>의 이미지 생성·합성, 영상 생성 및 후반 편집 과정에서 나타난 스타일 변동과 이를 조정하는 감독의 역할을 분석한다.

Ⅲ. 연구 설계

본 장에서는 사례 선정의 근거와 작품의 기본 정보를 제시하고, <엉뚱한 엉두>에 적용된 생성형 AI 제작 파이프라인을 단계별로 정리한다. 이어서 시각적 일관성, 시간적 일관성, 감독의 조정 개입을 중심으로 분석 대상과 기준을 구체화하고, 연구자의 제작 참여에 따른 자료 접근상의 이점과 분석상의 제약을 밝힌다.

3-1 사례 선정 및 작품 개요

본 연구는 생성형 AI 기반 스톱모션 스타일 애니메이션 <엉뚱한 엉두>를 질적 단일 사례로 선정하였다. 이 작품은 시

나리오 작성부터 후반 편집에 이르기까지 복수의 생성형 AI 도구와 후반 제작 도구를 연계하여 제작되었다. 한 단계의 산출물이 다음 단계의 입력으로 사용되는 구조이므로, 도구와 제작 단계가 전환되는 과정에서 스톱모션 스타일이 어떻게 구현되고 유지·조정되는지를 추적하기에 적합하다. 국내에서는 <갤럭시키즈>의 제작 과정을 중심으로 전통적 수작업과 3D 프린팅을 결합한 스톱모션 제작 사례가 연구된 바 있으며 [14], 본 사례는 생성형 AI 기반 제작 공정을 대상으로 한다는 점에서 구별된다.

또한 스톱모션 양식은 재료의 물질적 표면과 프레임 단위로 구성되는 리듬을 함께 요구하므로, 시각적 일관성과 시간적 일관성을 하나의 사례에서 동시에 검토할 수 있다.

본 작품은 복수의 참여자가 제작 단계를 분담하였으며, 연구자는 제작자 겸 감독으로 참여하여 중간 산출물과 조정 과정에 접근할 수 있다는 점을 사례 선정의 근거로 삼았다. 연구자의 구체적인 참여 범위와 그에 따른 제약은 3-4에서 기술한다.

<엉뚱한 엉두>는 4분 14초 분량의 패밀리 코미디이자 AI 스톱모션 애니메이션이다. 작품은 청개구리의 외형을 지닌 엉뚱한 초등학생 '엉두'의 일상을 다루며, 어린이 등장인물은 각자의 기질에 따라 동물의 모습으로 표현된다. 작품의 주요 시각적 방향은 따뜻한 클레이 질감과 생활 공간의 수공적 표현으로 설정되었으며, 이러한 기준은 캐릭터와 배경뿐 아니라 그림 2와 같이 타이틀을 포함한 화면 요소 전반에 적용되었다.



*The title is presented in Korean as it appears in the original work.
그림 2. <엉뚱한 엉두> 타이틀 화면
Fig. 2. Title Card of Eongttunghan Eongdu

3-2 생성형 AI 기반 제작 파이프라인

<엉뚱한 엉두>의 제작 파이프라인은 프리프로덕션, 메인프로덕션, 포스트프로덕션의 세 단계로 구성된다. 프리프로덕션 단계에서는 세계관과 캐릭터의 기본 설정을 수립하였다. 이어 Claude를 활용하여 시나리오와 컷 구성의 초안을 생성하고, 이를 이야기의 흐름과 캐릭터 행동을 기준으로 검토·수정하여 최종안을 확정하였다. 이후 Gemini를 활용해 콘티 이미지를 생성하고, 장면의 구도와 연출 방향을 사전에 검토하였다.

메인프로덕션 단계에서는 먼저 캐릭터 스케치를 통해 주요 캐릭터의 형태와 기본적인 시각적 특성을 설정하였다. 이어 ChatGPT를 활용하여 클레이·펠트 재질, 조명, 공간감에 관한 이미지 생성 프롬프트를 작성하였다. Nano Banana Pro를 통해 캐릭터, 배경 및 타이틀 이미지를 생성하였다. 본 연구에서 Nano Banana Pro는 정지 이미지의 생성과 이미지 기반 편집·합성에 활용되었으며, 개별 생성한 캐릭터와 배경을 하나의 장면으로 통합하는 용도로 사용하였다. 캐릭터는 장면별 표정과 동작에 따라, 배경은 공간별로 각각 생성하였으며, 이후 두 요소를 하나의 장면 이미지로 합성하였다. 생성과 합성 과정에서 이루어진 선별·재생성·검수의 기준은 4-1에서 다룬다.

합성된 장면 이미지는 각 컷의 시작·끝 프레임으로 활용하여 움직임의 방향과 범위를 설계하였으며, Kling과 Veo를 활용하여 컷 단위의 영상을 생성하였다. Kling과 Veo는 정지 이미지를 영상으로 확장하는 단계에서 활용되었으며, 본 제작에서는 시작·끝 프레임과 텍스트 프롬프트를 입력 조건으로 하여 캐릭터의 동작과 카메라 움직임을 생성하는 방식으로 사용하였다. 영상 생성 시 모션의 범위와 정도는 별도의 고정 수치로 일괄 설정하기보다, 시작·끝 프레임 사이의 캐릭터 위치와 자세 변화, 그리고 프롬프트에 지정한 캐릭터 동작과 카메라 이동의 범위를 기준으로 장면별로 조정하였다. 캐릭터의 형태와 배경 구도를 안정적으로 유지해야 하는 장면에서는 카메라를 고정하고 고개 회전, 방향 전환, 이동 경로와 종료 위치 등을 구체적으로 지정하였다. 반면 상대적으로 작은 움직임이 필요한 경우에는 slow, gently, gradually 등의 표현을 통해 동작과 카메라 변화의 정도를 제한하였다. 또한 fast motion, jerky movement, choppy frames, camera shake, abrupt zoom 등의 요소를 네거티브 프롬프트로 설정하여 과도한 움직임과 비의도적인 카메라 변화를 억제하였다.

실제 영상 생성 프롬프트의 예로, 캐릭터가 방향을 전환하여 화면 밖으로 이동하는 장면에서는 “Static camera. The character tilts its head to the right once, then turns and walks off to the left, exiting the frame completely.”와 같이 카메라를 고정한 상태에서 동작의 순서, 이동 방향 및 종료 위치를 지정하였다. 한편 피사체를 강조하는 장면에서는 “Slow push-in toward the wardrobe lower shelf. The tiny knitted socks sit upright, gently wobbling … The depth of field gradually shifts …”와 같이 느린 카메라 이동, 피사체의 미세한 움직임 및 점진적인 초점 변화를 지정하였다. 이러한 방식으로 시작·끝 프레임과 프롬프트를 함께 사용하여 장면별 움직임의 범위를 조정하였다.

포스트프로덕션 단계에서는 일부 효과음을 Adobe Firefly를 활용하여 생성하고, 적합한 결과를 선별하여 배치하였다. After Effects에서는 생성된 영상의 컷을 연결하고 장면의 흐름과 길이를 조정하였다. 또한 프레임 처리와 재생 속도 조절을 통해 스톱모션 특유의 분절된 운동감을 적용하였다. 구체적인 처리 방식은 4-2에서 다룬다.

이 파이프라인에서 본 연구가 주목하는 것은 각 도구의 성능이 아니라, 단계별로 어떠한 스타일 기준이 설정되고 어떤 판단을 통해 관리되었는가이다.

표 1. <엉뚱한 엉두> 제작 단계별 생성형 AI 도구와 감독의 역할
Table 1. Generative AI tools and director's roles by production stage

Production Stage	Process	Tool	Director's Role
Pre-production	World · Character Setup	Manual	Planning, Character Setup
	Script · Storyboard	Claude	Review, Revision, Shot Finalization
	Storyboard Images	Gemini	Scene Composition Review
Main Production	Character Sketches	Manual	Initial Character Design
	Prompt Writing	ChatGPT	Set Texture, Lighting, and Depth
	Character, Background, and Title Generation	Nano Banana Pro	Output Selection, Regeneration, Consistency Check, Expression and Motion Direction
	Character –Background Compositing	Nano Banana Pro	Lighting, Scale, and Spatial Fit Review
	Start–End Frame Setup	Manual	Motion Range Design
	Video Generation	Kling, Veo	Shot Selection, Error Review, Regeneration
Post-production	Sound Effect Generation	Adobe Firefly	Sound Selection and Placement
	Shot Editing	After Effects	Shot Assembly, Pacing Adjustment
	Stop-Motion Effect	After Effects	Frame Processing Criteria

3-3 분석 대상 및 분석 기준

본 연구의 분석 대상은 최종 완성 영상과 제작 과정의 이미지·영상 자료 및 관련 기록이며, 분석 단위는 최종 영상의 개별 컷과 대응하는 중간 제작 결과물이다.

1) 시각적 일관성

시각적 일관성은 캐릭터와 배경의 시각적 특성이 서로 다른 장면과 제작 단계에서도 작품의 스타일 기준 안에서 유지되는 정도를 의미한다. 세부적으로는 캐릭터의 얼굴과 신체 비례, 의상과 색채, 클레이·펠트 표면 질감, 표면의 요철과 지문 등 수작업적 인상을 형성하는 특성, 배경의 조형과 공간감, 캐릭터와 배경 사이의 조명 및 색조 통합을 분석한다. 캐릭터 생성 이미지, 합성 이미지와 최종 영상 프레임을 비교하여 어떤 특성이 유지되었고 어떤 부분에서 변형이 발생했는지 검토한다.

2) 시간적 일관성

시간적 일관성은 영상이 진행되는 동안 캐릭터와 오브젝트의 형태 및 움직임이 안정적으로 이어지는 정도를 의미한다. 분석 항목은 프레임 간 얼굴과 신체 형태의 지속성, 오브젝트와 배경의 안정성, 동작 방향과 자세의 연결, 시작·끝 프레임 사이의 움직임 변화, 그리고 개별적으로 생성된 컷들 사이의 속도와 분절 정도의 균일성으로 구성한다. 또한 비의도적 형태 변형은 생성 오류와 조정의 대상으로, 스톱모션 스타일을 위해 의도적으로 적용된 분절된 운동 리듬은 표현 요소로 구분하여 분석한다.

3) 감독의 조정 개입

감독의 조정 개입은 생성 결과가 작품의 스타일 기준에 부합하도록 선택·수정하고, 서로 다른 제작 단계의 결과물을 연결한 행위를 의미한다. 분석 항목은 스타일 기준의 설정, 생성 결과의 채택과 수정, 단계 간 산출물의 연결로 구성한다. 이를 통해 감독의 개입이 개별 오류의 수정뿐 아니라 전체 제작 공정의 스타일을 통합하는 데 어떠한 기능을 수행했는지 살펴본다.

표 2. <엉뚱한 엉두> 중간 산출물 선별 및 검수 기준
Table 2. Criteria for selection and review of intermediate outputs in Eongttunghan Eongdu

Production Stage	Review Criterion	Acceptance Criteria	Corrective Action
Image Generation	Character Identity	Consistent body color, facial structure, and proportions	Regenerate / revise reference
	Material and Surface	Consistent clay/felt texture and handcrafted surface	Revise prompt / regenerate
Compositing	Color, Lighting, and Spatial Integration	Consistent lighting, tone, scale, perspective, and spatial fit	Recompose / adjust placement, scale, tone
Video Generation	Form Consistency and Background Stability	Stable facial/body form, material, and background	Reject / regenerate
	Motion Consistency	Intended direction, pose, range, and end position	Revise prompt and start/end frames / regenerate
Post-production	Rhythm and Color Consistency	Consistent speed, segmentation, and color tone	Adjust timing, frame rate, or color

3-4 연구자의 제작 참여 위치와 분석상의 제약

본 연구자는 <엉뚱한 엉두>의 제작자 겸 감독으로 참여하여 작품의 시각적 방향 설정, 생성 결과 선별, 이미지 합성, 영상 생성 결과의 검토 및 후반 편집을 조정하였다. 이를 통해 제작 의도와 결정 과정, 단계별 중간 산출물에 접근할 수 있으며, 최종 결과물만으로는 확인하기 어려운 변동과 조정 과정을 분석할 수 있다. 그러나 연구자가 제작자이자 분석자라

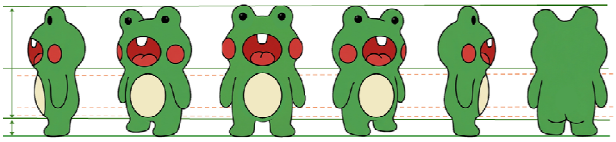


그림 3. 이미지 생성에 앞서 설정한 캐릭터 시트와 비례 기준
Fig. 3. Character reference sheet and proportion criteria established prior to image generation

는 점은 당시의 판단을 사후적으로 정당화할 가능성을 수반한다. 이에 본 연구에서는 최종 영상, 중간 생성 이미지, 프롬프트와 시작·끝 프레임 등 확인 가능한 자료를 상호 비교하고, 자료에서 직접 확인되는 변화와 연구자의 해석을 구분하여 기술하였다. 또한 중간 산출물의 선별과 검수에 적용된 질적 판단 기준을 시각적 일관성과 시간적 일관성의 세부 항목으로 구체화하였다. 시각적 검수에서는 캐릭터의 식별 요소와 신체 비례, 클레이·펠트 재질 및 표면 특성, 색채와 조명, 배경과의 크기·원근 관계를 확인하였으며, 영상 검수에서는 프레임 간 캐릭터 형태의 지속성, 배경 안정성, 동작 방향과 자세의 연결, 시작·끝 프레임과의 부합 여부, 움직임의 속도와 범위를 확인하였다. 각 항목에서 작품의 기준을 벗어난 결과는 그대로 채택하지 않고 재생성, 재합성 또는 후반 보정의 대상으로 분류하였다. 이러한 검수 기준을 표 2와 같이 명시함으로써 감독의 판단이 개인적 선호에만 의존하지 않고 작품 제작 이전 및 과정 중 설정된 스타일 기준에 근거했음을 확인할 수 있도록 하였다.

Ⅳ. 사례 분석: 〈엉뚱한 엉두〉

본 장에서는 <엉뚱한 엉두>의 제작 과정을 이미지 생성 및 합성, 영상 생성 및 후반 편집, 감독의 통합 조정이라는 세 측면에서 분석한다. 분석의 초점은 개별 생성형 AI 도구의 성능보다, 제작 단계가 전환되는 과정에서 스톱모션 스타일의 시각적·시간적 특성이 어떻게 유지되고 조정되었는가에 있다.

4-1 이미지 생성 및 합성 단계의 시각적 일관성 형성

1) 캐릭터 정체성 유지 과정

<엉뚱한 엉두>의 캐릭터 정체성은 이미지 생성 이전에 설정한 캐릭터의 형태와 색채를 기준으로 형성되었다. 주인공 엉두의 경우 녹색 몸체, 돌출된 눈, 넓은 입, 붉은 볼, 밝은색 복부를 주요 시각적 특징으로 설정하였다. 장면에 따라 의상, 자세, 표정과 소품은 달라지지만 이러한 핵심 특징은 동일한 캐릭터로 인식되기 위해 지속적으로 유지되어야 했다. 그림 3의 캐릭터 스케치와 그림 4의 장면별 생성 이미지를 비교하면 엉두의 기본적인 색채와 얼굴 구조가 공통 기준으로 사용된 것을 확인할 수 있다.

장면별 캐릭터 이미지는 기준 이미지를 참조하여 생성하였



그림 4. 캐릭터 시트를 기준으로 생성한 장면별 캐릭터 이미지
Fig. 4. Character images generated across scenes based on the reference sheet

으나, 생성 결과마다 얼굴의 비율, 눈과 입의 크기, 신체의 형태와 재질 표현에 차이가 나타났다. 이에 복수의 결과를 비교하여 기존 설정에 가까운 이미지를 선별하고, 필요한 경우 표정과 자세를 다시 지정하여 재생성하였다. 이때 선별의 기준은 개별 이미지의 완성도가 아니라 기준 이미지와의 부합 여부였다. 완성도가 높더라도 식별 요소가 기준에서 벗어난 결과는 채택하지 않았다. 캐릭터 정체성은 하나의 결과가 반복된 것이 아니라 이러한 비교와 조정을 통해 유지되었다.

캐릭터와 배경을 분리하여 생성한 뒤 합성하는 방식 또한 정체성 관리에 활용되었다. 배경과 캐릭터를 한 번에 생성할 경우 장면 구성에 따라 캐릭터 외형이 크게 달라지므로, 먼저 캐릭터 이미지를 확보하고 이를 장면별 배경에 배치하였다.

표 3. 캐릭터와 배경의 분리 생성 및 장면별 합성 결과
Table 3. Separately generated characters and backgrounds and their composited scenes

Character	Background	Composited Scene
		
		
		
		
		
		

합성은 확보된 요소를 배치하는 작업에 그치지 않았다. 캐릭터의 위치와 크기를 조정하는 과정에서 장면의 구도가 함께 결정되었으며, 스토리보드에서 설정한 구도가 생성된 배경과 맞지 않는 경우 배치나 프레이밍을 다시 조정하였다. 합성 이미지는 영상 생성 단계에서 시작 프레임으로 사용되므로, 이 단계에서 확정된 구도가 해당 컷의 구도가 되었다. 다만 합성 과정에서도 캐릭터의 크기와 형태가 변하거나 배경 색조에 따라 캐릭터 색상이 달라지는 현상이 나타나 결과를 다시 검토해야 했다.

2) 클레이·펠트 질감 및 공간 통합 과정

작품의 주요 시각적 방향은 실제 클레이 퍼펫을 촬영하는 대신 생성형 AI를 이용하여 클레이와 펠트로 제작된 듯한 외관을 구현하는 것이었다. 이를 위해 이미지 생성 프롬프트에는 재질뿐 아니라 표면의 요철, 부드러운 조명, 미니어처와 같은 공간감과 수작업적 형태를 함께 지정하였다. 특히 클레이로 제작된 캐릭터의 경우 표면에 손자국이 남은 듯한 질감이 드러나도록 조건을 부여하였다. 지문 자국과 눌린 흔적, 완전히 대칭적이지 않은 형태와 같이 물리적 조작 과정에서 자연스럽게 발생하는 특성을 프롬프트의 조건으로 명시한 것이다. 그림 5에서 확인되듯 생성된 캐릭터의 표면에는 이러한 흔적이 드러나며, 이는 2-1에서 검토한 수공성이 실제 조작의 결과로 남겨진 것이 아니라 프롬프트를 통해 의도적으로 설계된 것이라는 점에서 물리적 스톱모션과 구별된다.

재료의 지정은 캐릭터와 배경에 국한되지 않았다. 스톱모션에서는 연기 등 유동적인 현상을 솜과 같은 재료로 양식화하여 표현하는 방식이 사용되어 왔다[15]. 본 연구의 관점에서 이러한 재료적 대체는 기술적 제약을 우회하는 방식인 동시에, 유동적인 현상까지 물질적 재료로 표현함으로써 화면의 재료적 통일성을 형성하는 방식으로 볼 수 있다. 본 작품에서도 이러한 관습에 따라 그림 6과 같이 불꽃과 수증기를 섬유 재질의 형태로 지정하였으며, 국물과 채소 역시 반투명한 수지나 점토를 연상시키는 질감으로 구성하였다.



그림 5. 프롬프트로 지정하여 구현한 클레이 표면의 수작업 흔적(지문·눌린 자국)

Fig. 5. Handcraft traces on clay surfaces (fingerprints and press marks) specified via prompts

배경은 거실, 엉두의 방, 부엌, 욕실, 집 앞 골목 등 서로 다른 생활 공간으로 구성되었지만, 따뜻한 색조와 부드러운 표면, 단순화된 미니어처 구조를 공통적인 시각 기준으로 사용하였다. 이를 통해 공간의 기능과 구도는 달라지더라도 그림 7처럼 전체 장면이 동일한 작품 세계에 속한 것으로 인식되도록 하였다. 캐릭터와 배경의 합성에서는 재질의 유사성뿐 아니라 크기, 원근, 조명과 그림자의 관계가 중요하게 작용하였다. 캐릭터의 광원 방향과 배경의 명암이 일치하지 않거나 공간에 비해 캐릭터의 크기가 부적절한 경우 두 이미지가 분리된 요소처럼 보였다. 이에 캐릭터의 배치와 크기, 색조, 조명 방향을 조정하여 배경 안에 자연스럽게 포함되도록 하였다.

여기서 확인되는 것은 시각적 일관성이 모든 장면을 동일하게 만드는 방식이 아니라, 캐릭터의 식별 요소와 재질·조명 기준을 유지하면서 장면별 변화를 허용하는 방식으로 형성되었다는 점이다. 즉 일관성의 기준은 개별 이미지의 동일성이 아니라 작품 세계의 동일성에 있었다.



그림 6. 불꽃과 수증기를 섬유 재질로 구성한 조리 장면

Fig. 6. Cooking scene with flame and steam rendered in fiber material



그림 7. 미니어처 세트의 질감을 공통 기준으로 구성한 장면별 배경

Fig. 7. Background scenes constructed with shared miniature set texture

4-2 영상 생성 및 후반 편집 단계의 시간적 일관성 형성

1) 시작·끝 프레임 기반 움직임 생성

이미지 생성과 합성을 통해 완성한 장면은 영상 생성 단계에서 시작 프레임과 끝 프레임으로 활용되었다. 시작 프레임은 캐릭터의 형태와 위치, 배경 구성 및 카메라 구도를 제시하고, 끝 프레임은 동작이 도달해야 하는 최종 자세와 위치를 제한한다. 이러한 방식은 텍스트 프롬프트만 사용할 때보다 이동 방향과 컷의 종료 상태를 구체적으로 통제할 수 있게 하였다. 또한 합성 단계에서 확정한 캐릭터의 외형과 구도가 컷의 기준으로 유지되므로, 이미지 단계에서 형성한 시각적 일관성을 영상으로 이어가는 연결 지점으로도 기능하였다. 이때 모션의 범위는 3-2에서 제시한 바와 같이, 캐릭터의 이동 방향과 자세 변화, 카메라의 고정 또는 제한적 이동을 프롬프트에 구체적으로 명시하는 방식으로 제어하였다.

다만 두 프레임이 고정되어 있더라도 중간 프레임에서 캐릭터의 얼굴, 신체 비율과 재질이 유지되지는 않았으므로, 생성된 영상은 동작의 자연스러움뿐 아니라 캐릭터 형태의 지속성, 배경의 안정성, 움직임 방향의 의도 부합 여부를 함께 검토하였다. 가장 빈번한 문제는 캐릭터가 의도한 경로와 다르게 이동하거나 앞뒤 컷과 이어지지 않는 동선의 불일치였으며, 배경 세부 요소의 변형과 모션의 비일관성도 확인되었다. 이러한 결과는 제외하거나 시작·끝 프레임과 프롬프트를 수정하여 재생성하였고, 컷 간 색감 불일치는 후반 단계의 톤 보정으로 처리하였다.

2) 스톱모션 리듬 구현과 생성 오류 조정

영상 생성 과정에서 나타나는 프레임 간 불안정성과 스톱모션 특유의 분절된 움직임은 서로 다른 성격의 현상으로 구분하였다. 캐릭터의 얼굴이나 신체가 프레임 사이에서 변형되거나 배경 구조가 흔들리는 현상은 시간적 일관성을 저해하는 생성 오류에 해당한다. 반면 형태를 유지한 채 자세의 갱신 간격을 넓혀 얻는 분절된 움직임은 스톱모션의 운동감을 형성하기 위해 의도적으로 부여하는 표현 요소이다. 따라서 본 작품에서는 스톱모션 특유의 분절된 움직임을 생성 단계에서 직접 요청하지 않고, 후반 편집에서 별도로 조정하였다. 이는 개별적으로 생성되는 다수의 컷에 동일한 정도의 분절을 적용하기 어렵고, 컷마다 리듬의 정도가 달라질 경우 작품 전체의 시간적 통일성이 약화될 수 있다고 판단했기 때문이다.

리듬의 형성은 후반 편집 단계에서 이루어졌다. 그림 8과 같이 After Effects의 포스터라이즈 시간(Posterize Time) 이펙트를 사용하여 프레임 레이트를 8~12fps 범위에서 설정하였으며, 대부분의 컷에는 12fps를 적용하였다. 이는 원본 프레임을 홀드하여 프레임 갱신 간격을 넓히는 처리로, 형태를 변형하지 않고 시간 해상도만을 조정하며, 생성 단계의 편차와 달리 분절의 정도를 수치로 지정할 수 있다.

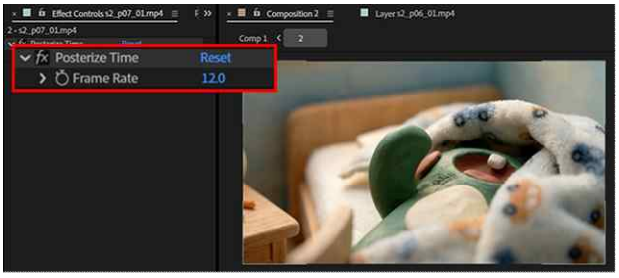


그림 8. After Effects 포스터라이즈 시간 이펙트를 활용한 프레임 간격 조정

Fig. 8. Frame interval adjustment using the posterize time effect in After Effects

또한 일부 컷에서는 캐릭터를 배경으로부터 레이어로 분리한 뒤, 캐릭터 레이어의 프레임 갱신 간격을 배경보다 크게 설정하였다. 인물의 분절감은 강조하되 공간은 안정적으로 유지하기 위한 조정이었으며, 적용 여부는 사전 기준이 아니라 장면을 확인하며 판단하였다. 이러한 조작은 장면 전체를 동일한 촬영 패스에서 기록하는 전통적 물리 스톱모션 방식에서는 직접 구현하기 어렵다.

이 지점은 2-1에서 검토한 스톱모션의 미학적 조건과 대비된다. 물리적 스톱모션에서 움직임의 불규칙성은 프레임마다 대상을 조작하는 과정에서 발생할 수 있으며, 그 비의도성이 수공적 제작의 흔적으로 기능한다. 반면 본 사례에서 분절된 리듬은 남겨진 흔적이 아니라 후반 단계에서 수치로 지정되어 부여된 요소이며, 생성 과정에서 실제로 발생한 불안정성은 오히려 제거의 대상이 되었다. 스톱모션 특유의 운동감은 유지되었으나, 그 효과를 발생시키는 조건은 조작의 흔적에서 편집상의 설계로 전환된 것이다.

4-3 감독의 통합 조정 역할

1) 스타일 기준 설정

감독은 작품의 기획과 세계관, 캐릭터와 전반적인 시각적 방향을 설정하고, 이를 이미지 생성, 영상 생성 및 컷 편집 각 단계에 적용되는 제작 기준으로 제시하였다.

스타일 기준은 특정 프롬프트 한 문장이나 단일 참조 이미지로 고정되지 않았다. 각 도구가 동일한 지시를 서로 다르게 해석하였기 때문에, 결과를 확인하고 작품의 기준에 맞는지를 다시 판단하는 과정이 매 단계에 요구되었다. 따라서 감독은 작품의 시각적 방향을 재질, 형태, 색채, 조명과 공간감 등의 구체적인 조건으로 변환하고, 이를 제작 단계별 판단 기준으로 사용하였다. 본 작품은 제작 단계를 분담하는 방식으로 진행되었으므로, 이렇게 구체화된 스타일 기준은 감독의 판단 준거인 동시에 각 단계를 담당한 참여자들이 제작 방향을 공유하기 위한 공통의 기준으로도 기능하였다.



그림 9. 서로 다른 장면에서 유지된 재질과 색조
Fig. 9. Material and color consistency maintained across scenes

2) 생성 결과 선별과 수정

생성형 AI는 동일한 입력에서도 복수의 결과를 산출하므로, 생성 자체보다 결과의 채택 여부를 판단하는 과정이 중요하게 작용하였다. 이미지 단계에서는 캐릭터의 비례와 재질, 구도와 조명을, 영상 단계에서는 형태의 지속성과 배경 안정성, 동작 방향과 속도를 기준으로 사용 가능한 결과를 선별하였다.

기준에 부합하지 않는 결과에는 프롬프트 수정, 참조 이미지 변경, 재생성, 재합성 및 편집이 적용되었다. 이때 어떤 층위에서 조정할 것인지에 대한 판단도 함께 이루어졌다. 앞서 4-2에서 확인한 바와 같이 동선과 모션 일관성의 문제는 생성 단계의 재생성으로, 컷 간 색감의 차이는 후반 단계의 보정으로 처리되었다. 이러한 판정과 조정은 오류의 사후 보완에 그치지 않고 작품에 사용될 이미지와 움직임을 결정하는 창작적 판단으로 기능하였다.

3) 단계 간 결과물 연결

다중 도구 기반 제작에서는 각 단계의 결과물을 다음 단계의 입력 조건에 맞게 변환하는 과정이 필요하였다.

각 도구는 이전 단계에서 설정된 제작 의도와 판단 기준을 공유하지 않는다. 따라서 단계가 전환될 때마다 유지해야 할 요소와 새롭게 조정해야 할 요소를 구분하는 판단이 요구되었다. 감독은 이 과정에서 서로 다른 형식의 결과물을 해석하고 연결하며, 파이프라인 전체의 스타일이 단절되지 않도록 조정하는 매개 역할을 수행하였다.

4-4 종합 분석

<엉뚱한 엉두>의 스톱모션 스타일 일관성은 특정 생성형 AI 도구의 결과만으로 형성되지 않았으며, 제작 단계별 선택과 조정이 누적되면서 형성되었다. 이미지 생성 및 합성 단계에서는 캐릭터의 식별 요소, 클레이·펠트 재질, 색채, 조명과 공간감이 주요 관리 대상이었다. 영상 생성 단계에서는 시작·끝 프레임을 통해 움직임의 방향과 범위를 제한하고, 중간 프레임에서 나타나는 형태 변화와 동선의 정합을 검토하였다. 후반 편집에서는 선택된 컷에 프레임 처리를 적용하여 스톱모션의 분절된 운동 리듬을 구성하였으며, 일부 컷에는 레이어별 차등 간격을 적용하였다. 이러한 조정을 거쳐 그림 9처럼 서로 다른 시간과 공간을 배경으로 하는 장면들이 공통된 재질과 색조를 유지하게 되었다.

시각적 일관성과 시간적 일관성은 서로 독립적으로 형성되지 않았다. 이미지 단계에서 안정적으로 설정된 캐릭터와 배경의 형태는 영상화 과정에서 형태 변화를 판단하는 기준점으로 기능하였으며, 개별 영상이 안정적으로 생성되더라도 컷마다 움직임의 속도와 리듬이 다르면 작품 전체의 통일성이 약해졌다. 따라서 두 층위의 일관성은 이미지 생성, 합성, 영상화와 후반 편집이 연결되는 과정에서 함께 관리되었다.

본 사례에서는 영상 생성 과정에서 나타나는 비의도적 형태 변화를 생성 오류로 판단하여 선별과 재생성을 통해 조정하였다. 반면 스톱모션 특유의 분절된 운동 리듬은 생성 단계에서 직접 구현하지 않고, 후반 편집에서 프레임 처리를 통해 별도로 구성하였다. 이는 생성 과정에서 발생하는 형태의 불안정성을 억제하는 한편, 작품 전체의 운동 리듬은 후반 단계에서 일관되게 조정하기 위한 방식이었다. 물리적 스톱모션에서는 실제 대상을 프레임마다 조작하는 과정 자체에서 운동 리듬이 형성되는 반면, 본 사례에서는 이러한 리듬의 형성을 후반 편집 단계로 분리하였다.

이를 종합하면 <엉뚱한 엉두>의 스타일 일관성은 스타일 기준 설정, 결과 생성, 비교와 선별, 수정·재생성, 단계 간 변환 및 후반 통합이 반복되는 조정 과정을 통해 형성되었다. 감독은 이 과정에서 생성 결과를 선택하는 역할뿐 아니라, 각 도구의 결과가 다음 단계에서도 동일한 시각적·시간적 방향을 유지하도록 연결하는 통합 기능을 수행하였다.

V. 논의 및 결론

앞선 분석 결과를 바탕으로, 본 장에서는 생성형 AI 기반 스톱모션 스타일의 일관성 형성 과정이 인간-AI 협업 구조와

감독의 조정 역할에 갖는 의미를 논의하고, 본 연구의 기여와 한계 및 후속 연구 방향을 제시한다.

5-1 인간-AI 협업 구조의 가능성과 한계

<엉뚱한 엉두>의 제작에서 AI는 각 단계에 필요한 초안과 복수의 결과물을 산출하였고, 감독은 이를 작품의 기준에 따라 선별·수정하여 다음 단계로 연결하였다.

이러한 협업 구조에서 확인되는 것은 이전 단계의 산출물이 다음 단계의 제작 자원으로 재사용된다는 점이다. 캐릭터와 배경을 분리하여 생성한 뒤 합성하고, 합성 이미지를 다시 영상 생성의 시작·끝 프레임으로 활용하는 방식이 그 예이다. 각 단계의 결과물은 그 자체로 완결되기보다 다음 도구의 입력 조건으로 전환되면서 제작을 이어가는 매개가 되었다.

그러나 도구의 연계가 제작 과정의 자동화나 스타일 일관성을 보장하는 것은 아니었다. 각 생성형 AI 도구는 이전 단계에서 설정된 캐릭터의 형태, 재질, 조명과 움직임의 의도를 완전하게 공유하지 않으므로, 도구가 전환될 때마다 결과가 새롭게 해석되거나 변형될 가능성이 있었다. 이미지 단계에서 확보된 캐릭터의 외형과 클레이 질감이 합성이나 영상화 과정에서 달라졌으며, 영상마다 움직임의 속도와 형태 안정성에도 차이가 나타났다.

이러한 변형은 매 단계에서 결과를 확인하고 판정하는 작업을 필연적으로 요구한다. 다수의 결과를 비교하고 오류를 판정하며, 필요한 경우 프롬프트와 참조 이미지, 시작·끝 프레임을 수정하여 재생성하는 과정이 반복되기 때문이다. 이는 생성형 AI 기반 제작의 효율성을 개별 결과물의 생성 시간만으로 평가하는 데 한계가 있음을 시사한다. 생성 이후에도 검수와 재생성 과정이 반복되기 때문이다.

본 사례에서 감독과 AI의 관계는 양측이 대등하게 기여하는 구조가 아니었다. 인간과 AI의 공동창작을 다루는 논의가 기여 방식과 상호작용의 순서를 구조의 구성 요소로 제시하듯[11], 본 사례에서도 AI는 복수의 결과를 산출하는 역할에, 감독은 그중 무엇을 채택하고 어떻게 수정할지를 판단하는 역할에 각각 배치되었다. 산출과 판정이 분리된 이러한 배치는 AI가 제작자를 대체하는 구조와 구별된다.

5-2 감독 조정 판단의 성격

본 사례에서 감독은 참여자들과 실무를 분담하였으나, 스타일 기준의 설정과 산출물의 판정, 단계 간 연결은 감독에게 집중되었다. 이때 판정 기준이 모두 명문화된 것은 아니었다. 앞서 확인한 레이어 분리 처리의 적용 여부가 사전에 정해진 기준이 아니라 장면을 확인하는 과정에서 결정된 것처럼, 일부 판단은 감독의 감각과 경험에 의존하였다.

또한 감독의 판단에는 부적합한 결과를 어느 단계에서 조정할 것인지에 대한 선택이 포함되었다. 동선과 모션의 문제는 생성 단계의 재생성으로, 컷 간 색감의 차이는 후반 단계의 보정으로 처리한 것이 그 예이다. 즉 감독의 조정은 결과의 채택 여부를 정하는 데 그치지 않고, 문제를 해결할 지점을 파이프라인 안에서 배치하는 판단을 포함하였다.

5-3 본 연구의 기여와 의의

본 연구의 첫 번째 기여는 스톱모션의 미학적 특성과 생성형 AI의 스타일 일관성 문제를 실제 제작 파이프라인 안에서 연결하여 분석했다는 점이다. 기존 스톱모션 연구에서 다뤄진 재료성과 수공성, 프레임 단위 운동의 특성을 생성형 AI 기반 이미지 생성, 합성, 영상화 및 후반 편집 과정과 연결함으로써, 스톱모션 스타일이 디지털 생성 환경에서 어떻게 구현되고 유지되는지를 구체적으로 살펴보았다. 특히 물리적 스톱모션에서 불꽃이나 수증기를 섬유 재질로 대체하여 표현해 온 관습을 AI 기반 제작에서도 재현한 사례는, 재료성의 재현이 캐릭터와 배경을 넘어 화면 요소 전반에 걸쳐 이루어졌음을 보여준다.

두 번째 기여는 생성 영상의 비의도적인 프레임 불안정성과 스톱모션의 의도적인 분절성을 구분했다는 점이다. 캐릭터의 얼굴이나 신체, 배경이 프레임 사이에서 변형되는 현상은 시간적 일관성을 저해하는 생성 오류에 해당한다. 반면 프레임 홀드와 단계적인 자세 갱신으로 나타나는 분절된 운동감은 스톱모션 스타일을 구성하기 위해 의도적으로 부여되는 표현 요소이다. 본 사례에서 전자는 생성과 선별 과정에서 조정되었고, 후자는 후반 편집에서 부여되었다. 본 연구는 이 두 요소가 서로 다른 단계에서 상반된 방향으로 처리되었음을 확인하였다.

세 번째 기여는 스타일 일관성을 개별 생성 모델의 성능 특성만으로 설명하지 않고, 제작 단계 간 조정이 누적되어 형성되는 과정적 특성으로 설명했다는 점이다. <엉뚱한 엉두>의 사례에서는 기준 설정, 결과 선별, 재생성, 합성, 시작·끝 프레임 구성과 후반 편집이 반복되며 이러한 일관성이 형성되었다. 특히 캐릭터와 배경에 서로 다른 프레임 간격을 부여한 레이어 분리 처리는, 스톱모션의 운동 리듬이 물리적 제작 방식의 모사에 그치지 않고 생성형 AI 기반 제작의 조건에 맞추어 재구성될 수 있음을 보여준다. 이에 본 연구는 감독을 생성 결과의 채택 여부를 판단하는 역할뿐 아니라, 서로 다른 도구의 산출물을 하나의 작품으로 연결하는 통합 조정 역할로 제시하였다.

이러한 결과는 생성형 AI 기반 애니메이션 제작에서 중요한 것이 특정 도구의 활용 여부만이 아니라, 제작 단계별로 관리해야 할 기준과 인간의 판단 지점을 명확히 설정하는 것임을 보여준다. 특히 복수의 도구를 연계하는 제작 환경에서 활용할 수 있는 파이프라인 구성과 스타일 관리의 실제 사례를 제시했다는 점에서 실천적 의의를 갖는다.

5-4 연구의 한계 및 후속 연구 제언

본 연구는 <엉뚱한 엉두> 한 작품을 대상으로 한 질적 단일 사례연구이므로, 분석 결과를 모든 생성형 AI 애니메이션 제작에 일반화하기 어렵다. 작품이 클레이·펠트 기반의 스톱모션 스타일을 지향했다는 점에서도 다른 시각 스타일이나 장르에 동일한 분석 결과를 그대로 적용하기에는 한계가 있다. 또한 연구자가 제작자 겸 감독이자 분석자로 직접 참여했다는 점에서 제작 당시의 판단을 사후적으로 정당화할 가능성이 존재한다. 이를 보완하기 위해 본 연구에서는 3-4절에서 제시한 바와 같이 캐릭터 정체성, 재질과 표면 특성, 색채·조명·공간 통합, 프레임 간 형태 지속성, 배경 안정성, 동작 정합성 및 컷 간 리듬 등을 중간 산출물의 선별·검수 기준으로 구체화하고, 최종 영상과 중간 생성 이미지, 프롬프트, 시작·끝 프레임을 상호 비교하였다. 다만 이러한 기준 역시 감독이 설정한 작품의 스타일 기준에 근거한 것으로, 외부 평가자에 의한 독립적 검증이나 평가자 간 일치도 검증을 수행한 것은 아니다. 또한 생성·수정·채택된 결과가 전 제작 단계에 대해 일관된 형식으로 기록되지 않아 전체 채택률이나 불합격률을 정량적으로 산출하기 어렵다는 한계가 있다.

후속 연구에서는 서로 다른 스타일과 제작 조건을 적용한 복수의 생성형 AI 애니메이션 사례를 비교할 필요가 있다. 이를 통해 본 연구에서 제시한 시각적 일관성, 시간적 일관성, 감독의 조정 개입이라는 분석 기준이 다른 장르와 제작 환경에서도 적용될 수 있는지를 검토할 수 있을 것이다.

또한 생성 횟수, 수정 및 재생성 횟수, 결과 채택·제외 횟수, 제작 시간과 제외 사유 등을 제작 단계별로 기록하여 채택률과 수정 빈도를 정량화할 필요가 있다. 이를 본 연구에서 제시한 질적 검수 기준과 함께 분석한다면, 각 제작 단계에서 요구되는 인간 개입의 정도와 효율성을 보다 구체적으로 비교할 수 있을 것이다. 나아가 복수의 평가자가 동일한 검수 기준을 적용하여 판단 결과를 비교한다면, 캐릭터 형태의 지속성, 배경 안정성, 동작 정합성과 같은 질적 판단의 신뢰도를 보다 체계적으로 검토할 수 있다.

마지막으로 관객이 생성형 AI로 구현된 클레이·펠트 질감과 분절된 움직임을 실제로 스톱모션 스타일로 인식하는지에 관한 수용자 연구가 요구된다. 이를 통해 제작자가 의도한 스타일 일관성과 관객이 지각한 일관성 사이의 관계를 살펴본다면, 생성형 AI 기반 스톱모션 애니메이션의 표현 가능성과 평가 기준을 더욱 확장할 수 있을 것이다.

감사의 글

이 논문은 2026년도 중앙대학교 연구 장학기금 지원에 의한 것임.

참고문헌

[1] D. I. Oh, “Expansion of the Scope of Animation Production with AI Focusing on Generative Animation,” *The Korean Journal of Animation*, Vol. 20, No. 4, pp. 157-171, December 2024. https://doi.org/10.51467/ASKO.2024.12.20.4.157

[2] H. Han, “Generative AI Technology for the Production of Animation,” *A Treatise on The Plastic Media*, Vol. 27, No. 1, pp. 12-21, February 2024. https://doi.org/10.35280/KOTPM.2024.27.1.2

[3] P. Wells, “Chairy Tales: Object and Materiality in Animation,” *Alphaville: Journal of Film and Screen Media*, No. 8, pp. 6-23, 2014. https://doi.org/10.33178/alpha.8.01

[4] R. Williams, *The Animator's Survival Kit*, Expanded ed. London, UK: Faber and Faber, 2009.

[5] Z. Huang, Y. He, J. Yu, F. Zhang, C. Si, Y. Jiang, ... and Z. Liu, “VBench: Comprehensive Benchmark Suite for Video Generative Models,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle: WA, pp. 21807-21818, June 2024. https://doi.org/10.1109/CVPR52733.2024.02060

[6] R. K. Yin, *Case Study Research and Applications: Design and Methods*, 6th ed. Thousand Oaks, CA: Sage, 2018.

[7] A. Gambrell, “In Visible Hands: The Work of Stop Motion,” *Animation Practice, Process & Production*, Vol. 1, No. 1, pp. 107-129, March 2011. https://doi.org/10.1386/ap3.1.1.107_1

[8] Y. Zhou, D. Zhou, M.-M. Cheng, J. Feng, and Q. Hou, “StoryDiffusion: Consistent Self-Attention for Long-Range Image and Video Generation,” in *Advances in Neural Information Processing Systems*, Vol. 37, pp. 110315-110340, December 2024. https://doi.org/10.52202/079017-3501

[9] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, “DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, Canada, pp. 22500-22510, June 2023. https://doi.org/10.1109/CVPR52729.2023.02155

[10] C. Moruzzi and S. Margarido, “A User-Centered Framework for Human-AI Co-Creativity,” in *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, Honolulu: HI, pp. 1-9, May 2024. https://doi.org/10.1145/3613905.3650929

[11] J. Rezwana and M. L. Maher, “Designing Creative AI

Partners with COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems," *ACM Transactions on Computer-Human Interaction*, Vol. 30, No. 5, 67, pp. 1-28, September 2023. https://doi.org/10.1145/3519026

[12] V. N. Antony and C.-M. Huang, "ID.8: Co-Creating Visual Stories with Generative AI," *ACM Transactions on Interactive Intelligent Systems*, Vol. 14, No. 3, 20, pp. 1-29, August 2024. https://doi.org/10.1145/3672277

[13] Y.-C. Zhang, "Research on Video Content Production Process Using Generative AI," *Journal of Communication Design*, Vol. 88, pp. 315-326, July 2024. https://doi.org/10.25111/jcd.2024.88.22

[14] T. H. Kim and J. W. Park, "Case Study on Production for Stop-Motion Animation 'Galaxay Kids'," *The Journal of the Korea Contents Association*, Vol. 17, No. 1, pp. 444-454, January 2017. https://doi.org/10.5392/JKCA.2017.17.01.44

[15] K. A. Priebe, *The Advanced Art of Stop-Motion Animation*, Boston, MA: Course Technology, 2011.



최현주(Hyun-Joo Choi)

2026년 : 중앙대학교 예술공학부 (공학사)
2026년 : 중앙대학교 첨단영상대학원 영상학과 (석사과정)
※관심분야 : 미디어아트(Media Art), 애니메이션(Animation), 실감콘텐츠(Immersive Content), 생성형 AI(Generative AI), 에듀테크(Educational Technology) 등



옥창엽(Chang-Yeob Ok)

2009년~2011년: 이화여자대학교 우주망원경 연구실 연구원
2021년~2021년: 광주과학기술원 융합기술학제학부 연구원
2024년~2025년: 청주대학교 디지털미디어디자인학과 조교수
2025년~현 재: 중앙대학교 예술공학부 조교수
※관심분야 : 다학제적 예술(Interdisciplinary Art), 미디어아트(Media Art), 실감콘텐츠(Immersive Content), 피지컬 컴퓨팅(Physical Computing) 등