

생성형 AI 영상의 미장센 제어 수준이 미학적 완성도와 시청자 몰입에 미치는 영향

공간 연속성과 피사체 동선 제어를 중심으로

The Effects of Mise-en-Scène Control in Generative AI Videos on Aesthetic Completeness and Viewer Immersion

Focusing on Spatial Continuity and Subject Trajectory Control

주 저 자 : 조한주 (Jho, Han Joo)

서원대학교 미디어콘텐츠학부 조교수
orkdman@gmail.com

<https://doi.org/10.46248/kidrs.2026.2.643>

접수일 2026. 04. 27. / 심사완료일 2026. 05. 26. / 게재확정일 2026. 06. 08. / 게재일 2026. 06. 30.

Abstract

This study analyzes how the level of mise-en-scène control, spatial continuity, and subject trajectory design in generative AI video production influences the aesthetic quality and viewer immersion of the final output. To achieve this, a comparative analysis was conducted between videos generated primarily through text prompts and videos in which creators intentionally controlled spatial structures and subject movements using multi-reference inputs and trajectory control techniques. Although recent generative AI technologies have significantly lowered the barriers to video production, they still reveal limitations in maintaining the continuity and structural stability of on-screen physical space according to the creator's intentions. Beyond producing visually impressive imagery, deep viewer immersion is difficult to achieve without deliberate spatial planning that carefully coordinates subject blocking and the geometric logic of cinematic space. The findings suggest that higher levels of spatial and trajectory control reduce visual inconsistencies occurring during camera movement and scene transitions, thereby lowering viewers' cognitive fatigue. In addition, intentionally designed spatial composition and movement trajectories were found to enhance narrative plausibility, viewer immersion, and the overall aesthetic quality of AI-generated videos. This study suggests that for generative AI videos to function as aesthetically complete works, creators' spatial planning and intentional control should play a central role rather than relying solely on the probabilistic generation process of AI systems.

Keyword

Generative AI Video(생성형 AI 영상), Spatial Continuity(공간의 연속성), Viewer Immersion(시청자 몰입), Mise-en-scène(미장센)

요약

본 연구는 생성형 AI 영상 제작 과정에서 미장센, 공간의 연속성, 그리고 피사체 동선에 대한 창작자의 제어 수준이 영상의 미학적 완성도와 시청자 몰입에 미치는 영향을 분석하고자 하였다. 이를 위해 텍스트 프롬프트 중심으로 생성된 영상과, 다중 레퍼런스 및 동선 제어 기술을 활용하여 창작자가 공간 구조와 피사체 움직임을 의도적으로 통제 한 영상을 비교 분석하였다. 최근 생성형 AI 기술은 영상 제작의 접근성을 크게 향상시켰으나, 화면 내부 공간의 구조적 안정성과 시공간적 연속성을 창작자의 의도에 따라 정교하게 유지하는 데에는 여전히 한계를 보이고 있다. 특히 단순히 시각적으로 화려한 이미지를 생성하는 것을 넘어, 피사체의 동선과 공간의 기하학적 구조가 유기적으로 연결될 때 시청자의 몰입 경험 또한 안정적으로 형성될 수 있다. 연구 결과, 공간 및 동선에 대한 인위적 제어 수준이 높을수록 카메라 이동 및 장면 전환 과정에서 발생하는 시각적 오류가 감소하는 경향이 나타났으며, 이에 따라 시청자의 인지적 피로감 또한 낮아지는 것으로 분석되었다. 또한 의도적으로 설계된 공간 구성과 피사체 동선은 서사적 개연성과 영상의 전반적인 몰입도 및 미학적 완성도 향상에 긍정적인 영향을 미치는 것으로 나타났다. 결론적으로 생성형 AI 영상이 단순한 기술적 결과물을 넘어 하나의 미학적 작품으로 기능하기 위해서는, AI의 확률적 생성 방식에만 의존하기보다 창작자의 공간 기획과 연출 통제가 중요한 요소로 작용할 필요가 있음을 확인하였다.

목차

1. 서론

- 1-1. 연구의 배경 및 필요성
- 1-2. 연구의 목적 및 가설

1.3 연구의 범위 및 방법

- 1-3-1 연구의 대상 및 범위
- 1-3-2 연구 및 평가 방법

2. 이론적 배경

- 2-1. AI 영상의 한계와 시각적 오류
- 2-2. 영상 매체에서의 공간 기획과 시청자 몰입

3. 연구 방법론

- 3-1. 생성 영상물 제작 및 비교 집단 설계
- 3-2. 연구 변수 및 분석 방법

4. 연구 결과 및 분석

- 4-1. 공간 통제와 인지적 피로도 감소
- 4-2. 동선 연출이 몰입에 미치는 영향

5. 결론 및 시사점

- 5-1. 연구결과 요약
- 5-2. 실무적 시사점
- 5-3. 연구의 한계점 및 향후 과제

참고문헌

1. 서론

1-1. 연구의 배경 및 목적

생성형 AI 기술의 발전은 영상 제작의 진입 장벽을 크게 낮추었다. 과거에는 대규모 자본과 전문 인력이 요구되었던 고품질 영상 제작이 이제는 텍스트 프롬프트 입력만으로도 단시간 내 구현 가능한 단계에 이르렀다. 특히 최근 비주얼 리얼리즘 측면에서 현재 AI 영상은 실사 수준에 준하는 레벨에 도달하였고, 미디어와 디자인 산업 전반에 혁신을 일으켰다. 마노비치는 디지털 미디어 환경에서 영상이 데이터와 알고리즘 기반으로 재구성되는 새로운 미디어 언어 체계로 변화하고 있다고 설명한다.¹⁾ 이러한 변화 속에서 생성형 AI 영상은 기존 영상 제작 방식과는 다른 새로운 시각적 생산 구조를 형성하고 있다. 그러나 시각 및 공간 디자인의 관점에서 볼 때, 현재의 생성형 AI 영상 기술은 여전히 구조적 한계를 지닌다. 영상은 단순한 2차원 픽셀의 집합이 아니라, 피사체와 카메라가 유기적으로 상호작용하는 3차원 공간 구조를 기반으로 구성된다. 따라서 실내의 공간의 원근감, 조명에 따른 심도 변화, 피사체의 동선 및 시선 방향 등 공간 연출 요소가 안정적으로 유지되지 않을 경우, 시청자의 몰입 또한 저해될 가능성이 높다. 최근 AI 영상 생성 분야에서도 이러한 공간적 비일관성은 주요 기술적 한계로 지적되고 있다. 현재의 텍스트 기반 비디오 생성 모델은 3차원 세계의 물리적 구조를 완전히 이해한 상태에서 영상을 생성하는 것이 아니라, 대규모 2차원 이미지 데이터의 확률적 패턴을 기반으로 프레임임을 예측하는 방식으로 작동한다. Tong Wu 외 6 (2025)는 이러한 기존 영

상 생성 모델이 지속적인 3차원 공간 이해의 부재로 인해 공간적 일관성을 유지하는 데 한계를 가진다고 설명한다.²⁾ 이러한 구조적 한계는 카메라 이동이나 장면 전환 과정에서 더욱 두드러지게 나타난다. 프레임 외부 공간의 형태가 갑작스럽게 변형되거나, 배경 구조와 피사체 위치가 비일관적으로 변화하는 현상은 시공간적 연속성을 약화시키며, 결과적으로 창작자의 연출 통제력을 제한하는 요인으로 작용한다. 특히 보드웰과 톰슨은 연속성 편집 이론에서 화면 내부의 공간 축과 시선 방향의 일관성이 유지될 때 관객이 장면의 공간 구조와 서사 흐름을 안정적으로 인지할 수 있다고 설명한다. 본 연구에서 관찰된 생성형 AI 영상의 공간 붕괴 현상은 이러한 연속성 편집 원리를 저해함으로써 시청자의 몰입과 공간 인지에 부정적인 영향을 미칠 가능성이 있다.³⁾ 또한 Yue Ma 외 21 (2025)는 기존 텍스트 기반 영상 생성 모델이 텍스트 프롬프트에 지나치게 의존함으로써 사용자의 세밀한 연출 의도를 충분히 반영하지 못한다고 지적한다. 이는 현재 생성형 AI 영상이 높은 시각적 품질에도 불구하고, 창작자의 정교한 공간 연출과 미장센 통제를 완전히 대체하지 못하는 이유로 해석될 수 있다.⁴⁾ 따라서 생성형 AI 영

1) 레프 마노비치, 뉴미디어의 언어, 커뮤니케이션북스, 2014, pp.65-75

2) Tong Wu, Shuai Yang, Ryan Po, Yinghao Xu, Ziwei Liu, Dahua Lin, Gordon Wetzstein, Video World Models with Long-term Spatial Memory, arXiv preprint, arXiv, 2025, p2

3) 데이비드 보드웰, 크리스틴 톰슨, Film Art: 영화예술, 지필미디어, 2011, pp.278-296

4) Yue Ma, Kunyu Feng, Zhongyuan Hu, Xinyu Wang, Yucheng Wang, Mingzhe Zheng, Bingyuan Wang, Qinghe Wang, Xuanhua He, Hongfa Wang, Chenyang Zhu, Hongyu Liu, Yingqing He, Zeyu

상이 단순한 기술적 시연을 넘어 하나의 미학적 작품으로 가능하기 위해서는 우연적 생성 결과에 의존하는 단계를 넘어설 필요가 있다. 본 연구는 창작자의 공간 기획력과 멀티뷰 및 카메라 궤적 제어 기반의 인위적 통제 방식이 영상의 공간적 안정성과 미학적 완성도에 어떠한 영향을 미치는지를 분석하고자 한다. 이를 통해 생성형 AI를 단순 자동화 도구가 아닌, 고도화된 공간 연출 매체로 확장하기 위한 이론적 기반을 제시하는데 목적이 있다.

따라서 생성형 AI 영상이 단순한 기술적 시연을 넘어 하나의 미학적 작품으로 가능하기 위해서는 우연적 생성 결과에 의존하는 단계를 넘어설 필요가 있다. 본 연구는 창작자의 공간 기획력과 멀티뷰 및 카메라 궤적 제어 기반의 인위적 통제 방식이 영상의 공간적 안정성과 미학적 완성도에 어떠한 영향을 미치는지를 분석하고자 한다. 이를 통해 생성형 AI를 단순 자동화 도구가 아닌, 고도화된 공간 연출 매체로 확장하기 위한 이론적 기반을 제시하는데 목적이 있다. 최근 국내 연구에서도 생성형 AI 기반 영상 제작 과정에서 미장센과 연출 구조에 대한 분석이 시도되고 있으며, AI 환경에서 창작자의 연출 개입 중요성이 지속적으로 논의되고 있다.⁵⁾

1-2. 연구의 목적 및 가설

본 연구의 목적은 생성형 AI 기반 영상 제작 과정에서 창작자의 미장센 제어 수준이 최종 결과물의 미학적 완성도와 시청자의 몰입 경험에 어떠한 영향을 미치는지를 분석하는 데 있다. 현재 대부분의 생성형 AI 영상 모델은 텍스트 프롬프트에 기반한 확률적 생성 방식으로 작동하기 때문에, 창작자가 화면 내부의 공간 구조와 피사체의 동선을 정교하게 통제하는 데 한계가 존재한다. 이러한 특성은 전통적인 영상 연출에서 중요하게 다루어지는 공간 연속성과 의도된 미장센 구성과 충돌할 가능성이 있으며, 결과적으로 영상의 서사적 설득력과 몰입도를 저하시킬 수 있다. 이에 본 연구는 단순 텍스트 프롬프트 중심의 영상 생성 방식을 넘어, 다중 레퍼런스 이미지 활용 및 카메라 궤적 제어와 같은

인위적 개입 방식이 생성형 AI 영상의 시공간적 안정성과 미학적 완성도에 미치는 영향을 분석하고자 한다. 이를 위해 다음과 같은 세부 연구 목적을 설정하였다. 첫째, 생성형 AI가 확률적으로 구성한 공간과 창작자가 의도적으로 제어한 공간 사이에서 나타나는 시각적 오류와 공간적 일관성의 차이를 비교 분석한다. 둘째, 공간 연속성과 피사체 동선의 안정성이 시청자의 몰입 경험과 인지적 피로감에 어떠한 영향을 미치는지 검토한다. 셋째, 생성형 AI 영상이 하나의 미학적 작품으로 가능하기 위해 요구되는 공간 기획 및 연출 통제 요소를 도출하고, 실무적 활용 가능성을 탐색한다. 본 연구는 위와 같은 연구 목적을 바탕으로 다음과 같은 연구 가설을 설정하였다.

[표 3] AI 영상 생성 모델의 연출적 제어 수준에 따른 시각적·서사적 효과 가설

구분	연구가설	변인구조	핵심개념 및 이론적 근거
가설 1	창작자의 인위적 제어 수준이 높을수록 영상의 시각적 안정성이 향상될 것이다.	[독립변인] 인위적 제어 개입 수준 [종속변인] 시각적 안정성	시공간적 일관성 유지 인지적 부하 최소화
가설 2	의도된 공간 구성과 피사체 동선 설계는 시청자의 몰입 경험에 긍정적인 영향을 미칠 것이다.	[독립변인] 공간 및 동선 설계 수준 [종속변인] 몰입 경험 및 서사적 설득력	연속성 편집 공간 리얼리티 형성

본 연구는 이러한 가설 검토를 통해 생성형 AI 기반 영상 제작 환경에서 기술적 자동화와 창작자의 연출 통제 사이의 관계를 분석하고, 향후 AI 영상 제작의 연출적 활용 가능성을 탐색하고자 한다.

1-3 연구의 범위 및 방법

1-3-1 연구의 대상 및 범위

본 연구는 현재 상용화되어 실무 및 학계에서 널리 활용되고 있는 생성형 비디오 모델(예: Runway Gen-3, Kling, Luma 등)을 연구 대상으로 설정한다. 특히 단일 프레임의 시각적 해상도나 정지 이미지 품질 평가보다는, 영상 매체의 본질적 특성인 시간의 흐름에 따른 공간적 연속성과 피사체 동선의 물리적 타당성에 초점을 맞춘다. 최근 생성형 AI 영상 평가에 관한 연구들은 단순한 프레임 단위의 선명도를 넘어, 영

Wang, Zhifeng Li, Xiu Li, Sirui Han, Yike Guo, Wei Liu, Dan Xu, Linfeng Zhang, Qifeng Chen, Controllable Video Generation: A Survey, arXiv preprint, arXiv, 2025, p1

5) 박민서, '생성형 AI를 이용한 영상제작과 미장센 유형분석 연구', 한국엔터테인먼트산업학회, vol.19, no.3, 2025, pp91-104

상 내부에서 발생하는 공간 붕괴 현상과 피사체의 비일관적 변형 등 세계 일관성 문제의 중요성을 강조하고 있다. Akshat Rakheja 외 3 (2025)는 최신 영상 생성 모델이 개별 프레임 수준에서는 높은 사실성을 구현하더라도, 객체의 임의적 생성 및 소멸, 공간 관계의 비논리적 변화, 물리적 인과성 붕괴 등의 문제로 인해 전체 영상 차원에서는 비일관성과 비현실성이 발생할 수 있다고 설명한다.⁶⁾ 따라서 본 연구에서는 생성형 AI 영상 제작 과정에서 창작자가 개입하는 미장센 요소의 인위적 제어 수준을 주요 독립변인으로 설정한다. 이를 위해 텍스트 프롬프트 기반 생성에 주로 의존하는 통제 집단과, 다중 뷰 레퍼런스 및 카메라 궤적 제어 기술을 활용하여 피사체 동선과 공간 구조를 의도적으로 통제한 실험 집단으로 연구 범위를 구분하였다. 또한 종속변인은 영상 시청 과정에서 나타나는 공간적 안정성 인식, 인지적 피로감, 그리고 전반적인 몰입 경험으로 설정하였다. 본 연구는 이러한 변인 간 관계를 통해 생성형 AI 영상에서 미장센 제어 수준이 결과물의 미학적 완성도에 미치는 영향을 분석하고자 한다.

1-3-2 연구 및 평가 방법

본 연구는 가설 검토를 위해 질적 평가와 양적 평가를 결합한 혼합 연구 방법론을 채택하였으며, 다음과 같은 절차로 진행하였다. 첫째, 자극물 제작 단계에서는 동일한 서사 시나리오를 기반으로 제어 수준에 차이를 둔 두 종류의 생성형 AI 영상을 제작하였다. 실험 장면은 인물이 복도를 이동하며 180도 방향 전환을 수행하는 장면으로 설정하였으며, 통제 집단은 모델의 확률적 추론에 의존하여 영상을 생성하였다. 반면 실험 집단은 3D 공간 앵커링 및 ControlNet 기반 제어 기법을 활용하여 카메라 동선과 피사체 위치를 보다 안정적으로 유지하도록 설계하였다. 둘째, 전문가 집단 심층 면접(Focus Group Interview, FGI)을 진행하였다. 시각디자인 및 영상 연출 분야의 실무 경험을 보유한 전문가 그룹을 대상으로 생성 영상의 시각적 플리커링 현상, 180도 가상선 붕괴, 뎁스 왜곡 및 공간 위계 변화 등을 중심으로 질적 분석을 수행하였다. 셋째, 일반 시청자를 대상으로 정량적 설문 평가를 실시하였다. 참여자들은 두 유형의 영상을 교차 시청한 후 리커

트 척도 기반 질문에 응답하였으며, 공간적 불안정성으로 인해 발생하는 인지적 피로감과 의도된 동선 설계가 몰입 경험 및 서사적 설득력에 미치는 영향을 평가하였다. 본 연구는 일반 시청자 30명을 대상으로 진행되었으며, 실험 영상은 모두 동일한 길이(20초), 해상도(1920×1080), 프레임레이트(24fps) 조건에서 제작하였다. 수집된 설문 데이터는 IBM SPSS Statistics 29.0을 활용하여 독립표본 t-test 방식으로 분석하였다. 이와 같은 다각적 평가 방식을 통해 본 연구는 생성형 AI 영상 제작에서 창작자의 공간적 통제 개입이 영상의 시공간적 안정성과 미학적 완성도에 어떠한 영향을 미치는지를 분석하고자 한다.

2. 이론적 배경

2-1. AI 영상의 한계와 시각적 오류

현재 생성형 AI 영상 기술은 단일 프레임 수준에서는 높은 시각적 완성도를 구현하고 있으나, 시간의 흐름 속에서 물리적 일관성을 유지하는 데에는 여전히 구조적 한계를 드러낸다. 이는 AI가 3차원 공간의 물리 법칙을 구조적으로 이해한 상태에서 영상을 생성하는 것이 아니라, 대규모 데이터의 확률적 패턴을 기반으로 2차원 이미지 프레임을 예측하는 방식으로 작동하기 때문이다. 그 결과 카메라 시점이 이동하거나 피사체가 움직이는 과정에서 시공간적 연속성이 불안정하게 나타난다. 프레임 간 미세한 이미지 흔들림, 피사체 형태 왜곡, 배경 구조 변화와 같은 할루시네이션 현상이 발생하며, 이는 시청자의 공간 인지 및 몰입 경험을 저해하는 요인으로 작용한다. Zhiyu Yin 외(2025)는 영상 생성에서 시간적 일관성(temporal consistency)이 유지되지 않을 경우 이미지 점프(image jump), 플리커링(flickering), 부자연스러운 움직임 등이 발생할 수 있다고 설명한다. 이는 개별 프레임의 품질이 높더라도 전체 영상 차원에서의 연속성이 확보되지 않으면 시청자가 영상의 흐름을 비현실적이거나 불안정하게 인식할 수 있음을 의미한다.⁷⁾ 결국 공간의 기하학적 논리가 결여된 생성형 AI 영상은 단순한 시각적 노이즈를 넘어, 시청자에게 인지적 부조화를 유발할 가능성이 있다. 이러한 문제는 영상 매체의 핵심 요소인 서사적

6) Akshat Rakheja, Aarsh Ashdhir, Aryan Bhattacharjee, Vanshika Sharma, World Consistency Score: A Unified Metric for Video Generation Quality, arXiv preprint, arXiv, 2025, p1

7) Zhiyu Yin, Kehai Chen, Xuefeng Bai, Ruili Jiang, Juntao Li, Hongdong Li, Jin Liu, Yang Xiang, Jun Yu, Min Zhang, A Survey: Spatiotemporal Consistency in Video Generation, arXiv preprint, arXiv, 2025, p2

몰입을 방해하는 주요 기술적 한계로 작용한다.

2-2. 영상 매체에서의 공간 기획과 시청자 몰입

전통적인 영상 연출론에서 미장센은 화면 내부의 시공간적 요소를 창작자의 의도에 따라 배치하고 통제하는 연출 방식으로 이해된다. 특히 피사체의 동선과 카메라의 움직임은 인물의 심리 상태와 서사 전개를 시각적으로 전달하는 핵심 요소로 기능한다. 로벨로는 영화 연출에서 배우의 동선과 카메라 움직임이 장면의 감정 구조와 시각적 리듬 형성에 중요한 역할을 수행한다고 설명한다.⁸⁾ 따라서 공간 구성과 동선 설계는 단순한 시각적 배치를 넘어, 관객이 장면의 흐름과 상황을 자연스럽게 인지하도록 만드는 중요한 장치라고 볼 수 있다. 보드웰은 관객이 서사를 안정적으로 인지하기 위해서는 화면 내부의 공간 축과 시선 방향의 연속성이 유지되어야 한다고 설명한다. 이러한 연속성 편입은 장면 간 공간 관계를 안정적으로 유지함으로써 관객의 몰입을 유도하는 영화 문법의 핵심 원리 중 하나이다.⁹⁾ 그러나 생성형 AI 영상에서 확률적으로 생성된 피사체 동선과 카메라 워킹은 이러한 공간적 연속성을 불안정하게 만들 가능성이 존재한다. 피사체의 움직임과 카메라 시점 사이의 인과관계가 자연스럽게 연결되지 않을 경우, 시청자는 장면의 공간 구조와 서사 흐름을 직관적으로 이해하기 어려워진다. 이는 결과적으로 영상의 개연성과 몰입 경험을 저해하는 요인으로 작용할 수 있다. 생성형 AI 영상에서 확률적으로 생성된 피사체 동선과 카메라 워킹은 이러한 공간적 연속성을 불안정하게 만들 가능성이 존재한다. 피사체의 움직임과 카메라 시점 사이의 인과관계가 자연스럽게 연결되지 않을 경우, 시청자는 장면의 공간 구조와 서사 흐름을 직관적으로 이해하기 어려워진다. 이는 결과적으로 영상의 개연성과 몰입 경험을 저해하는 요인으로 작용할 수 있다. 보드웰은 관객이 영화의 시공간 정보를 인지하고 이를 바탕으로 서사를 구성하는 과정 자체가 영화 몰입 형성의 핵심 메커니즘이라고 설명한다.¹⁰⁾ 따라서 생성형 AI 영상에서 공간 구조와 피사체 동선의 연속성이 불안정하게 나타날 경우, 시청자는 장면의 흐름을 자연스럽게 인지하기 어려워질 가능성이

존재한다. 최근 생성형 AI 영상 연구에서도 단순한 배경 유지 수준을 넘어, 카메라 움직임과 서사 전개가 유기적으로 연결되는 공간적 개연성의 중요성이 강조되고 있다. 최근 생성형 AI 영상 연구에서도 단순한 배경 유지 수준을 넘어, 카메라 움직임과 서사 전개가 유기적으로 연결되는 공간적 개연성의 중요성이 강조되고 있다. 즉, 영상 생성 과정에서 창작자의 의도된 공간 설계와 동선 통제가 확보될수록 시청자는 보다 안정적으로 화면 내부의 상황을 인지하고 몰입할 가능성이 높아진다. 바쟁 또한 영화의 현실감은 단순한 시각 효과가 아니라, 공간과 시간의 연속성을 통해 형성된다고 설명하였다.¹¹⁾ 이러한 관점에서 볼 때 생성형 AI 영상에서도 공간 구조의 안정성과 시공간적 개연성은 몰입 형성의 핵심 요소로 작용할 가능성이 존재한다. 따라서 창작자가 카메라 동선과 피사체 위치를 적극적으로 제어할 경우, 생성형 AI 영상은 단순한 기술적 시연을 넘어 미학적 가치를 지닌 영상 콘텐츠로 확장될 수 있다. 의도적으로 설계된 공간 기획과 동선은 시공간적 오류를 완화할 뿐만 아니라, 시청자에게 서사적 타당성과 공간적 안정감을 제공하는 역할을 수행한다. 본 연구에서 '미학적 완성도'란 영상 내부 공간의 연속성, 카메라 동선의 자연성, 피사체 움직임의 개연성, 그리고 시청자의 몰입 지속 정도가 통합적으로 유지되는 상태를 의미한다.

3. 연구 방법론

3-1. 자극물 제작 및 비교 집단 설계

본 연구는 생성형 AI 영상 제작 과정에서 미장센 제어 수준에 따른 시각적·서사적 차이를 분석하기 위해 두 가지 비교 집단의 자극물을 제작하였다. 실험의 변별력을 확보하기 위해, 인물이 복잡한 기하학적 구조를 가진 실내 공간을 이동하며 180도 방향 전환을 수행하는 시공간적 연속성 중심의 시퀀스를 실험 장면으로 설정하였다.

[표 4] 자극물 제작 및 비교 집단 설계 구조

구분	집단A 프롬프트 의존형	집단B 창작자 통제형
핵심	AI의 확률적 생성	창작자의 인위적 통제

11) 안드레 바쟁, 영화란 무엇인가, 바다출판사, 2005, pp.33-49

8) 빈센트 로벨로, 영화 연출론, 비즈앤비즈, 2018, pp.112-128

9) 데이비드 보드웰, 크리스틴 톰슨, Film Art: 영화예술, 지필미디어, 2011, pp.278-296

10) 데이비드 보드웰, 픽션 영화의 내레이션, 커뮤니케이션북스, 2017, pp.45-63

개념	방식 및 유연성 중심 생성	와 의도적 연출 기반 생성
제어 도구	텍스트 프롬프트 단독 입력	다중 레퍼런스 및 카메라 궤적 제어 기술 활용
공간 구성	AI 알고리즘 기반의 확률적 공간 매핑	다중 뷰 레퍼런스를 통한 공간 구조 고정
피사체 동선	2D 데이터 패턴에 의존한 비예측적 이동	사전 설계된 이동 경로 및 카메라 워킹 적용
시각적 일관성	프레임 간 플리커링 및 할루시네이션 발생 가능성	시공간적 연속성 유지 및 시각적 오류 최소화

본 연구에서 설정한 비교 집단은 생성형 AI의 확률적 생성 방식과 창작자의 의도적 공간 제어 방식 사이의 차이를 비교하기 위해 구성되었다. 특히 공간 구조 유지, 피사체 동선의 안정성, 그리고 카메라 움직임의 연속성을 중심으로 두 집단 간 시각적 차이를 분석하고자 하였다. 이를 통해 생성형 AI 영상 제작에서 창작자의 공간적 개입 수준이 영상의 시공간적 안정성과 몰입 경험에 어떠한 영향을 미치는지 검토하고자 한다.

3-2. 연구 변수 및 분석 방법

본 연구는 생성형 AI 영상 제작 과정에서 창작자의 미장센 제어 수준이 영상의 시각적 안정성과 시청자의 몰입 경험에 어떠한 영향을 미치는지를 분석하기 위해 다음과 같은 독립변수와 종속변수를 설정하였다. 연구 변수의 구성과 측정 항목, 그리고 분석 방법은 [표 3]과 같다. 특히 본 연구는 창작자의 공간 기획 및 동선 제어 수준에 따라 발생하는 시각적 차이를 비교하고, 이러한 차이가 시청자의 인지적 반응과 몰입 경험에 미치는 영향을 정량적·정성적 방법을 통해 분석하고자 한다. 이를 통해 생성형 AI 영상에서 인위적 공간 통제가 시공간적 안정성과 미학적 완성도 형성에 어떠한 역할을 수행하는지를 검토한다.

[표 5] 연구 변수 정의 및 분석 방법 구조

변수 구분	측정 항목	조작적 정의 및 측정 내용	분석 및 검증 방법
독립 변수	인위적 통제 수준	-텍스트 프롬프트 의존 무작위 생성 -다중 레퍼런스 및 궤적 제어 적용	집단별 자극물 분류 및 교차 비교 실험
종속 변수	시각적 오류 빈도	-프레임 간 이미지 플리커링 발생률	영상 프레임 데이터 정량

1		-공간 붕괴 및 형태 변이 횟수	분석 및 전문가 FGI 비평
종속 변수 2	인지적 피로도	-시각적 모순 인지 시 발생하는 심리적 저항감 -화면 전환 시 시청자가 느끼는 부하 정도	일반 시청자 대상 리커트 척도 설문
종속 변수 3	전반적 몰입도 및 미학적 가치	-설계된 미장센의 서사적 타당성 부여 수준 -단순 기술 시연을 넘어선 작품성 평가	전문가 집단 인터뷰 및 일반 시청자 설문 통계 분석

4. 연구 결과 및 분석

4-1. 인위적 통제 따른 인지 피로 감소 효과

본 연구는 가설 1인 창작자의 인위적 제어 수준이 높을수록 시각적 오류와 시청자의 인지적 피로도가 감소할 것이라는 검토하기 위해 집단별 자극물에 대한 비교 분석과 설문 평가를 실시하였다. 분석 결과, 공간 및 동선 제어 수준은 영상의 시각적 안정성과 시청자의 인지적 피로감 형성에 유의미한 영향을 미치는 것으로 나타났다. 우선 영상 내 시각적 오류 발생 양상을 비교한 결과, 텍스트 프롬프트 중심으로 생성된 [집단 A]에서는 카메라 앵글이 회전하는 장면 구간에서 프레임 간 플리커링, 공간 구조 왜곡, 피사체 형태 변형 등의 현상이 반복적으로 관찰되었다. 특히 3초~5초 구간에서 배경 구조와 객체 위치가 불안정하게 변화하는 양상이 두드러지게 나타났다. 반면 다중 레퍼런스 이미지와 카메라 궤적 제어 기술을 적용한 [집단 B]에서는 공간 구조의 안정성과 피사체 동선의 연속성이 상대적으로 안정적으로 유지되는 경향을 보였다. 이는 창작자의 공간 제어 개입이 장면 전환 과정에서 발생하는 시각적 불안정성을 완화하는 데 일정 부분 효과를 가질 수 있음을 시사한다. 또한 일반 시청자 30명을 대상으로 두 자극물에 대한 인지적 피로도를 리커트 5점 척도로 조사한 후 독립표본 t-test를 실시한 결과는 다음과 같다.

[표 6] 집단별 인지적 피로도 및 시각적 오류 체감 차이 테스트 ($p < .01$, $p < .001$)

평가 항목	[집단 A] 프롬프트 의존형 (평균)	[집단 B] 창작자 통제형 (평균)	t값

화면 전환 시 피로감	4.37	1.83	6.42
공간 구조의 불협화음	4.51	1.64	7.15
시각적 일관성 저해 통증	4.12	1.90	5.89

[표 4]에서 확인할 수 있듯이, 두 집단 간 인지적 피로도 평균 점수는 통계적으로 유의미한 차이를 보였다($p < .001$). [집단 A]를 시청한 참여자들은 인터뷰 과정에서 “카메라가 회전할 때 배경 구조가 계속 변형되어 시각적으로 불안정하게 느껴졌다”, “인물 형태가 미세하게 변형될 때마다 장면에 집중하기 어려웠다”와 같은 반응을 보였다. 이러한 결과는 생성형 AI 영상에서 발생하는 시공간적 불안정성이 시청자의 몰입 경험과 인지적 부담감 형성에 영향을 미칠 가능성을 보여준다. 실제로 Zhiyu Yin 외 9 (2025)는 영상 생성 과정에서 시간적 일관성이 유지되지 않을 경우 이미지 점프, 플리커링, 부자연스러운 움직임 등이 발생할 수 있다고 설명한다.¹²⁾ 따라서 생성형 AI 영상에서 공간 구조와 피사체 동선에 대한 창작자의 제어 수준이 높아질수록 시청자는 보다 안정적으로 장면의 흐름을 인지할 수 있으며, 결과적으로 인지적 피로감 또한 감소하는 경향을 보이는 것으로 해석할 수 있다. 이는 미장센 기반의 공간 제어가 단순한 시각적 보정을 넘어, 시청자의 몰입 경험을 향상시키는 중요한 연출 요소로 기능할 가능성을 시사한다.

4-2. 동선 연출이 몰입에 미치는 영향

창작자의 공간 및 동선 제어는 시청자의 인지적 피로감을 감소시키는 기술적 보완 요소일 뿐만 아니라, 영상의 미학적 완성도와 서사적 개연성을 형성하는 중요한 연출 장치로 기능한다. 본 연구는 가설 2인 “의도적 공간 및 동선 설계는 서사적 타당성과 몰입 경험 향상에 긍정적인 영향을 미칠 것이다”를 검토하기 위해 영상 연출 및 시각디자인 분야 전문가 10인을 대상으로 심층 면접(Focus Group Interview, FGI)을 실시하고, 집단별 자극물의 미학적 특성을 질적으로 분석하였다. 전문가 패널 구성과 주요 평가 내용은 [표 5]와 같다.

12) Zhiyu Yin, Kehai Chen, Xuefeng Bai, Ruili Jiang, Juntao Li, Hongdong Li, Jin Liu, Yang Xiang, Jun Yu, Min Zhang, A Survey: Spatiotemporal Consistency in Video Generation, arXiv preprint, arXiv, 2025, p2

[표 7] FGI 참여 전문가 패널 구성 및 주요 비평 요약

구분	전문 분야 및 실무 경력	주요 비평 코멘트 내용
패널 1~3	영상 제작 및 연출 (20년차 이상)	생성형 AI 영상이 실무 환경에서 활용되기 위해서는 카메라 동선과 공간 구성을 안정적으로 제어할 수 있어야 한다고 평가하였다. 또한 프롬프트 중심 생성 방식에서는 180도 가상선(axis) 유지가 불안정하게 나타난다는 의견이 제시되었다
패널 4~7	시각 및 공간 디자인 기획 (10년차 이상)	공간 구조에 대한 통제가 부족할 경우 뎁스 왜곡(depth distortion)과 공간 위계 혼란이 발생하며, 이는 장면 내부의 시각적 조화를 저해할 수 있다고 분석하였다.
패널 8~10	상업 영상 촬영 및 CG 포스트 프로덕션	생성형 AI 영상이 상업 영상 수준의 현실감을 확보하기 위해서는 피사체 위치와 카메라 움직임 사이의 공간적 정합성이 중요하다고 평가하였다.

분석 결과, 전문가 집단은 텍스트 프롬프트 중심으로 생성된 [집단 A]에 대해 상대적으로 낮은 수준의 서사적 개연성과 공간 안정성을 평가하였다. 특히 AI의 확률적 생성 방식에 의해 형성된 피사체 이동 경로와 카메라 움직임이 장면 내부의 서사 흐름과 유기적으로 연결되지 못한다는 의견이 다수 제시되었다. FGI에 참여한 한 전문가는 “인물의 이동 경로와 카메라 움직임 사이의 연출 의도가 명확하게 연결되지 않을 경우, 장면의 흐름이 단절된 것처럼 느껴질 수 있다”고 언급하며, 이러한 공간적 비일관성이 몰입 경험을 저해할 가능성을 지적하였다. 또한 전문가들은 [집단 A]에서 카메라 시점이 변화하는 과정에서 공간 축의 연속성이 불안정하게 유지되는 현상에 주목하였다. 보드웰이 설명한 연속성 편집 원리에 따르면, 화면 내부의 공간 축과 시선 방향이 안정적으로 유지될 때 관객은 장면의 흐름을 자연스럽게 인지할 수 있다. 그러나 생성형 AI 영상에서 이러한 공간 연속성이 무너질 경우, 시청자는 방향 감각의 혼란과 함께 서사 흐름의 단절을 경험할 가능성이 존재한다.¹³⁾ 반면 다중 레퍼런스 이미지와 카메라 궤적 제어 기술을 적용한 [집단 B]는 공간 구조와 피사체 동선의 연속성이 상대적으로 안정적으로 유

13) 데이비드 보드웰, 크리스틴 톰슨, Film Art: 영화예술, 지필미디어, 2011, pp.278-296

지되는 것으로 평가되었다. 특히 카메라 이동과 피사체 위치 변화가 보다 자연스럽게 연결되면서, 시청자가 장면 내부의 공간 관계를 안정적으로 인지할 수 있다는 의견이 제시되었다. 전문가들은 [집단 B]의 생성물이 단순한 기술 시연을 넘어, 창작자의 연출 의도가 반영된 영상 콘텐츠로 인식될 가능성이 높다고 평가하였다. 특히 카메라 앵글 변화 과정에서도 공간 축의 일관성이 비교적 안정적으로 유지되었으며, 피사체 이동과 조명 변화 또한 장면 흐름과 유기적으로 연결된다는 점이 긍정적으로 분석되었다. 종합적으로 볼 때, 생성형 AI 영상 제작에서 창작자의 공간 기획과 동선 제어는 단순한 기술적 보정 수준을 넘어, 영상의 서사적 설득력과 몰입 경험 형성에 중요한 역할을 수행하는 것으로 해석할 수 있다. 이는 생성형 AI 기반 영상 제작 환경에서 미장센 제어 기술이 향후 영상 연출의 핵심 요소로 기능할 가능성을 시사한다.

5. 결론 및 시사점

5-1. 연구결과 요약

본 연구는 생성형 AI 기반 영상 제작 환경에서 창작자의 미장센 제어 수준이 시청자의 인지적 피로감과 영상의 미학적 완성도에 어떠한 영향을 미치는지를 분석하고자 하였다. 이를 위해 텍스트 프롬프트 중심의 확률적 생성 방식인 [집단 A]와 다중 레퍼런스 및 카메라 궤적 제어를 적용한 창작자 통제형 방식인 [집단 B]를 비교 대상으로 설정하여 실험을 진행하였다. 연구를 통해 도출된 주요 결과는 다음과 같다. 첫째, 창작자의 공간 및 동선 제어 개입 수준이 높아질수록 영상 내 시각적 오류와 시청자의 인지적 피로감이 감소하는 경향이 나타났다. 특히 프롬프트 중심 생성 방식에서는 프레임 간 플리커링, 공간 구조 왜곡, 피사체 형태 변형과 같은 시공간적 불안정성이 반복적으로 관찰되었으며, 이는 시청자의 몰입 경험을 저해하는 요인으로 작용하였다. 반면 다중 레퍼런스와 카메라 궤적 제어를 적용한 영상에서는 공간 연속성과 피사체 동선의 안정성이 상대적으로 높게 유지되는 경향이 나타났다. 둘째, 의도적으로 설계된 공간 구성과 피사체 동선은 영상의 서사적 개연성과 몰입 경험 형성에 긍정적인 영향을 미치는 것으로 분석되었다. 전문가 FGI 결과에서도 카메라 움직임과 피사체 동선 사이의 공간적 연속성이 안정적으로 유지될수록 장면의 흐름과 연출 의도가 보다 자연스럽게 전달된다는 의견이 제시되었다. 특히 공간 축의 안정성과 카메라 위킹의 연속성

이 확보된 장면은 시청자가 화면 내부의 공간 관계를 직관적으로 이해하고 서사에 몰입하는 데 중요한 역할을 수행하는 것으로 나타났다. 종합적으로 볼 때, 생성형 AI 영상 제작 과정에서 창작자의 공간 기획 및 미장센 제어는 단순한 기술적 보정 수준을 넘어, 영상의 시공간적 안정성과 미학적 완성도를 형성하는 핵심 요소로 기능할 가능성이 있음을 확인할 수 있었다.

5-2. 실무적 시사점

본 연구는 생성형 AI 기반 영상 제작 환경에서 창작자의 공간 기획 능력과 연출 통제의 중요성을 제시한다는 점에서 실무적 의미를 가진다. 최근 생성형 AI 영상 기술은 높은 수준의 시각적 리얼리즘을 구현하고 있으나, 단순한 텍스트 프롬프트 중심 생성 방식만으로는 공간 연속성과 서사적 개연성을 안정적으로 유지하는 데 한계가 존재한다. 따라서 향후 생성형 AI 영상 제작 과정에서는 단순 프롬프트 입력 중심의 제작 방식에서 나아가, 다중 레퍼런스 이미지 활용, 카메라 궤적 제어, 공간 구조 고정 등 창작자의 의도적 개입이 가능한 제작 방식이 더욱 중요해질 것으로 예상된다. 특히 영상 연출 분야에서는 피사체 동선과 카메라 움직임 사이의 공간적 관계를 안정적으로 유지하는 것이 시청자의 몰입 경험 형성에 핵심적인 역할을 수행할 가능성이 있다. 또한 생성형 AI 기술이 상업 영상 및 콘텐츠 제작 환경으로 확장될수록, 창작자의 공간 설계 능력과 미장센 제어 역량은 AI 영상의 품질과 차별성을 결정하는 주요 요소로 작용할 수 있다. 이는 생성형 AI가 단순 자동화 도구를 넘어, 창작자의 연출 의도를 구현하기 위한 새로운 영상 제작 도구로 발전할 가능성을 시사한다.

5-3. 연구의 한계점 및 향후 과제

본 연구는 특정 생성형 AI 영상 모델과 제한된 실내 공간 시퀀스를 중심으로 실험을 진행하였다는 점에서 연구 범위의 한계를 가진다. 따라서 향후 연구에서는 더욱 다양한 생성형 AI 플랫폼과 복합적인 공간 환경을 대상으로 실험 범위를 확장할 필요가 있다. 특히 실내 공간뿐만 아니라 대규모 야외 환경이나 복잡한 군중 장면에서의 공간 연속성 유지 가능성에 대한 추가 검토가 요구된다. 또한 본 연구에서는 시청자의 인지적 피로감과 몰입 경험을 리커트 척도 기반 설문과 전문가 인터뷰를 중심으로 분석하였다. 향후 연구에서는 아이 트래킹, 시선 이동 분석, 뇌파 기반 반응 측정 등 생체 신호 기반 분석 방법을 병행함으로써 생성형 AI

영상이 시청자의 인지 과정에 미치는 영향을 보다 정확하게 분석할 필요가 있다. 아울러 생성형 AI 영상 기술은 빠르게 발전하고 있기 때문에, 향후에는 World Model 기반 영상 생성 기술이나 3차원 공간 이해 모델의 발전이 미장센 제어 및 공간 연속성 문제를 어느 수준까지 개선할 수 있는지에 대한 지속적인 연구가 필요할 것으로 판단된다.

참고문헌

- 데이비드 보드웰, 크리스틴 톰슨, 『Film Art: 영화예술』, 지필미디어, 2011.
- 데이비드 보드웰, 『픽션 영화의 내레이션』, 커뮤니케이션북스, 2017.
- 레프 마노비치, 『뉴미디어의 언어』, 커뮤니케이션북스, 2014.
- 빈센트 로벨로, 『영화 연출론』, 비즈앤비즈, 2018
- 안드레 바쟁, 『영화란 무엇인가』, 바다출판사, 2005
- 박민서, '생성형 AI를 이용한 영상제작과 미장센 유형분석 연구', 한국엔터테인먼트산업학회, 2025
- Tong Wu, Shuai Yang, Ryan Po, Yinghao Xu, Ziwei Liu, Dahua Lin, Gordon Wetzstein, 'Video World Models with Long-term Spatial Memory', arXiv preprint arXiv:2506.05284, 2025, 06
- Yue Ma, Kunyu Feng, Zhongyuan Hu, Xinyu Wang, Yucheng Wang, Mingzhe Zheng, Bingyuan Wang, Qinghe Wang, Xuanhua He, Hongfa Wang, Chenyang Zhu, Hongyu Liu, Yingqing He, Zeyu Wang, Zhifeng Li, Xiu Li, Sirui Han, Yike Guo, Wei Liu, Dan Xu, Linfeng Zhang, Qifeng Chen, 'Controllable Video Generation: A Survey', arXiv preprint arXiv:2507.16869, 2025, 07.
- Akshat Rakheja, Aarsh Ashdhir, Aryan Bhattacharjee, Vanshika Sharma, 'World Consistency Score: A Unified Metric for Video Generation Quality', arXiv preprint arXiv:2508.00144v1, 2025, 08.
- Zhiyu Yin, Kehai Chen, Xuefeng Bai, Ruili Jiang, Juntao Li, Hongdong Li, Jin Liu, Yang Xiang, Jun Yu, Min Zhang, 'A Survey: Spatiotemporal Consistency in Video Generation', arXiv preprint, arXiv:2503.06053 2025. 03.
- arxiv.org
- runwayml.com
- klimgai.com
- lumalabs.ai