

프롬프트텔링기반의 AI 영상 콘텐츠 제작 과정 연구

경기도 지역 랜드마크와 전통 캐릭터의 하이브리드 파이프라인을 중심으로

A Study on AI Video Content Production Process Based on Prompt-telling

Focusing on the Hybrid Pipeline of Gyeonggi-do Landmarks and Traditional Characters

주 저 자 : 권동은 (Kwon, Dong Eun) 한국폴리텍대학 디지털콘텐츠과 교수
goldkwon@kopo.ac.kr

<https://doi.org/10.46248/kidrs.2026.2.1114>

접수일 2026. 05. 19. / 심사완료일 2026. 06. 01. / 게재확정일 2026. 06. 08. / 게재일 2026. 06. 30.

Abstract

This study proposes a prompt-telling-based hybrid pipeline designed to control the probabilistic variability inherent in generative AI video production and to ensure visual consistency between real-world footage and virtual assets. A hierarchical prompt architecture consisting of Universe, Scene, and Direction levels was developed and empirically tested across ten major landmarks in Gyeonggi-do. The analysis demonstrates that the proposed methodology maintains a 92.2% design consistency rate by suppressing character design variance to within 7.8%, a significant improvement over single-prompt methods. Furthermore, by synchronizing real-world lighting data with AI parameters, the study achieved a 94.1% lighting synchronization rate, effectively eliminating the physical dissonance between elements. Expert evaluations further validated the model's commercial viability, with both visual and narrative congruence scoring above 4.5 on average. This research holds significant academic and industrial value by standardizing a creator's aesthetic judgment into digital assets. It establishes a standard production pipeline capable of generating uniform, high-quality content regardless of the practitioner's technical proficiency.

Keyword

Generative AI(생성형 AI), Prompt-telling(프롬프트텔링), Hybrid Video(하이브리드 영상)

요약

본 연구는 생성형 AI 영상 제작 과정에서 발생하는 확률론적 변동성과 실사 영상 및 가상 에셋 간의 시각적 불일치 문제를 개선하기 위해 프롬프트텔링 기반 하이브리드 제작 파이프라인을 제안하였다. 이를 위해 유니버스, 씬, 디렉션의 단계로 구성된 계층형 프롬프트 아키텍처를 설계하고, 경기도 주요 랜드마크 10개소와 전통 모티프 캐릭터를 결합한 실증 실험을 수행하였다. 분석 결과, 제안한 방법론은 단일 프롬프트 방식 대비 캐릭터 디자인 변동률을 7.8% 이내로 억제하며 92.2%의 디자인 항상성을 확보하였고, 실사 광원 데이터와 AI 파라미터의 동기화를 통해 94.1%의 조도 일치율을 달성하였다. 또한 전문가 평가에서 시각적 정합성과 서사적 적합성이 모두 평균 4.5점 이상으로 나타나 하이브리드 영상 제작 방식의 실무적 활용 가능성을 확인하였다. 해당 수치는 디자인 항상성, 조도 동기화율, 지형 접지 정밀도 및 전문가 평가를 기반으로 산출하였다. 본 연구는 프롬프트를 단순 명령어가 아닌 서사 및 연출 제어 체계로 확장하여, 생성형 AI 기반 영상 제작의 표준화된 제작 방법론과 프롬프트 라이브러리 기반 활용 가능성을 제시하였다는 점에서 학술적·산업적 의의를 가진다.

목차

1. 서론

- 1-1. 연구의 배경
- 1-2. 연구의 목적
- 1-3. 연구의 범위와 방법

2. 이론적 배경 및 선행연구

- 2-1. 생성형 AI 영상 툴 현황

- 2-2. 프롬프트텔링 개념 정의

- 2-3. 하이브리드 영상 제작 방식 고찰

3. 프롬프트텔링 기반 기획 구조 설계

- 3-1. 단계별 프롬프트 아키텍처 설계
- 3-2. 서사 기반 프롬프트 라이브러리 구축
- 3-3. 실험 환경 및 분석 지표 설정

4. 하이브리드 제작 파이프라인 구축과 구현

- 4-1. 실사 기반 하이브리드 워크플로우 설계
- 4-2. 경기도 랜드마크 영상 콘텐츠 구현
- 4-3. 대조군 · 실험군 비교 분석
- 4-4. 전문가 평가 결과 분석

1. 서론

1-1. 연구의 배경

최근 생성형 AI 기술은 텍스트-투-비디오와 비디오-투-비디오 모델의 발전을 통해 영상 제작 방식에 큰 변화를 가져오고 있다. 특히 실사 영상에 AI 기반 가상 에셋을 결합하는 하이브리드 영상 제작은 지역 랜드마크, 문화유산, 관광 콘텐츠 등에서 새로운 시각적 표현 가능성을 제시한다. 그러나 생성형 AI는 확률론적 생성 방식을 기반이기 때문에 동일한 프롬프트를 사용하더라도 장면마다 캐릭터의 외형, 색채, 조명, 질감 등이 달라지는 문제가 발생한다. 이러한 시각적 불일치는 실사 배경과 가상 에셋을 결합하는 영상에서 물리적 이질감을 유발하며, 연출자의 의도를 안정적으로 반영하는 데 한계로 작용한다.

본 연구는 생성형 AI를 단순한 자동 생성 도구가 아니라 연출 의도를 체계적으로 제어할 수 있는 제작 시스템으로 활용하기 위해, 프롬프트를 서사와 연출 구조를 포함하는 기획 언어로 확장한 프롬프트텔링 개념을 제안하고자 한다.¹⁾

1-2. 연구의 목적

본 연구는 생성형 AI 영상 제작 과정에서 발생하는 확률론적 변동성과 장면 간 시각적 불일치를 완화하고, 실사 영상과 AI 기반 가상 에셋 간의 정합성을 확보하기 위한 프롬프트텔링 기반 하이브리드 제작 방법론을 제안하는 데 목적이 있다. 이를 위해 첫째, 프롬프트텔링의 개념을 정립하고, 둘째, 유니버스(Universe), 씬(Scene), 디렉션(Direction) 단계로 구성된 계층형 프롬프트 아키텍처를 설계한다. 셋째, 경기도 주요 랜드마크와 전통 모티프 캐릭터를 결합

1) 김은정, 김정윤, 생성형 AI 게임 캐릭터 시스템 설계 및 구현 연구(2025), 차세대컨버전스정보서비스기술논문지, pp.2-3.

5. 결론

참고문헌

한 실증 실험을 통해 단일 프롬프트 방식과 제안 방식의 차이를 비교 분석한다. 이를 통해 생성형 AI 기반 영상 제작에서 시각적 일관성과 연출 제어력을 높일 수 있는 제작 가이드라인을 제시하고자 한다.²⁾

1-3. 연구의 범위와 방법

본 연구는 경기도 주요 랜드마크 10개소와 한국 전통 모티프 캐릭터를 결합한 하이브리드 영상 제작 과정을 연구 범위로 설정한다. 연구 방법으로는 생성형 AI 영상 제작 및 프롬프트 관련 선행연구를 고찰하고, 이를 바탕으로 유니버스, 씬, 디렉션 단계의 계층형 프롬프트 아키텍처를 설계하였다. 또한 실사 영상과 AI 기반 가상 에셋을 결합한 실증 실험을 수행하고, 디자인 항상성, 조도 동기화율, 서사적 적합성 등을 중심으로 단일 프롬프트 방식과 비교 분석하였다. 최종적으로 전문가 집단의 리커트 척도 평가를 통해 제안한 방법론의 타당성과 활용 가능성을 검토하였다

2. 이론적 배경 및 선행연구

2-1. 생성형 AI 영상 톨 현황

생성형 AI 영상 제작 기술은 텍스트 입력을 기반으로 영상을 생성하는 텍스트-투-비디오와 기존 이미지 또는 영상을 바탕으로 새로운 움직임과 장면을 생성하는 이미지-비디오 기반 생성 방식으로 발전하고 있다. OpenAI의 Sora, Runway Gen-3, Luma Dream Machine, Kling AI 등은 텍스트 프롬프트, 이미지 참조, 카메라 제어, 모션 생성 기능을 통해 기존 영상 제작 과정의 일부를 단축하고, 창작자의 시각적

2) 이연우, AI 기반 이미지 생성 프롬프트 활용을 통한 영상스토리텔링 교육 모델 연구(2025), 멀티미디어학회논문지 pp.2014-2015.

실험 가능성을 확대하고 있다. 이들 도구는 고해상도 영상 생성, 카메라 움직임 제어, 이미지 기반 영상화, 장면 확장 등에서 강점을 보이지만, 생성형 AI의 확률론적 특성으로 인해 동일하거나 유사한 입력 조건에서도 결과가 달라질 수 있다. 특히 장면이 연속되는 영상 콘텐츠에서는 캐릭터의 형태, 색채, 조명, 질감, 공간감이 일관되게 유지되지 않는 문제가 발생한다. 이러한 한계는 실사 배경과 가상 에셋을 결합하는 하이브리드 영상 제작에서 시각적 이질감을 유발하므로, 장면 간 일관성과 정합성을 제어하기 위한 별도의 제작 체계가 요구된다. 주요 생성형 AI 영상 툴의 특징과 기술적 지표는 아래 [표 1]과 같다.³⁾

[표 1] AI 영상 툴별 특성 및 제어 범위 비교

구분	주요 강점	제어 방식	한계점
OpenAI Sora	고품질 영상 생성	텍스트 기반 제어	장면 일관성 유지 한계
Runway Gen-3	정교한 카메라 및 모션 제어	브러시·프롬프트 제어	프롬프트 해석 가변성
Luma Dream	자연스러운 움직임 생성	비디오 기반 생성	복잡한 구조 유지 한계
Kling AI	높은 동작 표현력	텍스트·비디오 기반 생성	세부 연출 제어 한계

2-2. 프롬프트텔링 개념 정의

프롬프트텔링은 프롬프트와 스토리텔링의 합성 개념으로, 생성형 AI에 입력되는 기술적 명령어를 서사와 연출 의도를 포함하는 기획 언어로 확장하는 방식이다. 기존의 프롬프트 엔지니어링이 결과물의 정확도와 효율 향상에 초점을 두었다면, 프롬프트텔링은 장면 간 관계, 캐릭터 설정, 공간 맥락, 미장센 등의 연출 요소를 구조화하여 영상의 시각적 일관성과 서사적 흐름을 유지하는 데 목적이 있다.⁴⁾

생성형 AI 영상 제작에서는 동일한 입력 조건에서도 결과가 달라지는 확률론적 특성으로 인해 캐릭터 형태, 색채, 조명, 공간감 등이 장면마다 변형되는 문제가 발생한다. 프롬프트텔링은 이러한 변동성을 줄이기 위해 세계관, 장면, 연출 요소를 계층적으로 구조화하여 AI가 반복적으로 참조할 수 있는 서사적 상수를 설정하는 방

식으로 활용된다. 이러한 관점에서 프롬프트텔링은 다음과 같은 특성을 가진다. 첫째, 프롬프트를 계층화하여 장면 간 시각적 일관성을 유지하는 통제된 창의성을 구현한다. 둘째, 실사 배경과 AI 기반 가상 에셋 사이의 광원, 색채, 질감 정보를 조정하여 시각적 정합성을 높인다. 셋째, 생성형 AI를 자동 생성 도구가 아닌 연출 보조 시스템으로 활용함으로써, 제작자가 조작보다 기획, 연출에 집중할 수 있는 환경을 제공한다.⁵⁾

[표 2] 프롬프트 엔지니어링과 프롬프트텔링의 비교

구분	프롬프트 엔지니어링	프롬프트텔링
지향점	결과 생성 및 최적화	서사 및 연출 구조화
핵심 요소	명령어 조합 및 문법 제어	장면·캐릭터·연출의 계층적 구성
제작 방식	단일 결과 중심 생성	장면 간 일관성 중심 생성
시각적 특징	개별 이미지·영상 생성	연속적 시각 흐름 유지
활용 목적	생성 결과 효율 향상	하이브리드 영상 제작 정합성 강화
연구 의의	AI 활용 방식 확장	AI 기반 영상 제작 방법론 제안

2-3. 하이브리드 영상 제작 방식 고찰

하이브리드 영상 제작 방식은 실사 촬영을 통해 확보한 공간의 현실감과 생성형 AI로 제작한 가상 에셋을 결합하여 하나의 영상 장면으로 구성하는 제작 방식을 의미한다. 기존의 하이브리드 영상은 주로 3D 모델링, VFX, 매치무빙, 합성 기술을 기반으로 제작되었으나, 최근 생성형 AI를 활용하여 이미지, 캐릭터, 배경, 움직임 등을 빠르게 생성하고 실사 영상과 결합하는 방식으로 확장되고 있다. 이러한 방식은 실제 장소가 지닌 현장성과 AI가 생성하는 이미지를 동시에 활용할 수 있다는 점에서 지역 홍보 영상, 문화콘텐츠, 관광 브랜딩 분야에서 활용 가능성이 높다. 특히 경기도 랜드마크와 같은 실사 기반 공간에 전통 모티프 캐릭터를 결합할 경우, 장소의 현실성과 캐릭터의 서사성을 함께 전달할 수 있다. 그러나 하이브리드 영상 제작에서는 실사 배경과 가상 에셋 사이의 시각적 정합성이 중요한 문제로 나타난다. 가상 캐릭터의 조명, 색채, 질감, 원근감이 실사 배경과 어

3) https://frarang.com/frarang_Lifestyle_Review/?bmode=view&idx=167692023

4) <https://wikidocs.net/318793>

5) 주종우, 2035년까지의 글로벌 사진 산업 전망과 생성형 AI(2025), 현대사진영상학회논문집, pp.167-169.

곳날 경우 영상 전체의 몰입도가 저하될 수 있기 때문이다. 따라서 하이브리드 영상 제작에서는 단순히 가상 요소를 생성하는 것뿐만 아니라, 실사 공간의 광원, 색온도, 심도, 카메라 움직임 등을 고려하여 가상 에셋을 조정하는 과정이 필요하다. 이러한 하이브리드 영상 제작의 특성을 바탕으로, 프롬프트텔링을 통해 실사 배경과 AI 생성 에셋 간의 시각적 정합성을 제어하고자 한다. 즉, 프롬프트를 장면별 명령어가 아니라 장소성, 서사, 조명, 캐릭터 설정을 통합하는 제작 구조로 활용함으로써 장면 간 일관성과 영상 완성도를 높이는 방법을 고찰한다.⁶⁾

[표 3] 하이브리드 영상 제작 방식의 핵심 특성 분석

구분	주요 내용	기대 효과
리얼리티	실사 공간과 AI 에셋 결합	몰입감 향상
서사성	장소와 캐릭터의 이야기 구성	콘텐츠 확장성 강화
효율성	생성형 AI 기반 제작	시간비용 절감
시각적 정합성	조명·색채·공간감 조정	이질감 최소화
산업적 확장성	지역 기반 콘텐츠 활용	브랜딩 및 관광 활용

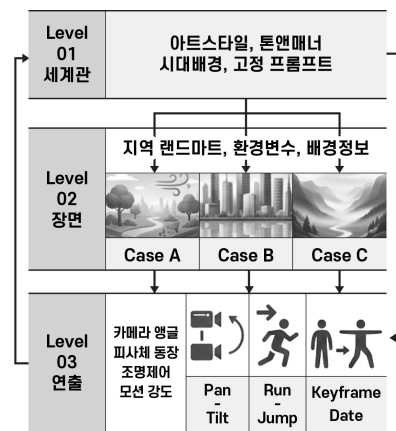
3. 프롬프트텔링 기반 기획 구조 설계

3-1. 단계별 프롬프트 아키텍처 설계

본 연구는 생성형 AI 영상 제작 과정에서 발생하는 장면 간 시각적 불일치 문제를 완화하고, 실사 영상과 가상 에셋 간의 정합성을 유지하기 위해 계층형 프롬프트 아키텍처를 설계하였다. 생성형 AI는 동일하거나 유사한 프롬프트를 사용하더라도 결과물이 달라지는 확률론적 특성을 가지므로, 영상 제작 과정에서는 캐릭터 형태, 색채, 조명, 질감 등의 시각 요소가 장면마다 변형되는 문제가 발생한다. 특히 하이브리드 영상에서는 이러한 변동성이 실사 배경과 가상 에셋 사이의 시각적 이질감을 증가시키는 요인으로 작용한다. 이에 프롬프트를 단일 명령어 방식이 아닌 계층적 구조로 구성하여 장면 간 시각적 일관성을 유지하고자 하였다. 이를 위해 단일 프롬프트 기반 생성 방식을 대조군으로 설정하고, 유니버스, 씬, 디렉션의 3단계

구조로 구성된 계층형 프롬프트 아키텍처를 실험군으로 설계하였다.

[그림 1]은 계층형 프롬프트 아키텍처의 구조를 나타낸 것이다. 유니버스 레벨은 영상 전체의 화풍, 색채, 조명 톤, 캐릭터 스타일과 같은 공통 시각 요소를 정의하는 단계이며, 씬 레벨은 장소별 조도, 색온도, 공간 분위기 등 환경 정보를 반영하여 실사 배경과 가상 에셋 간의 정합성을 조정하는 단계이다. 디렉션 레벨은 카메라 움직임, 모션 강도, 장면 전환 등의 연출 요소를 설정하여 개별 장면의 흐름을 제어한다.⁷⁾



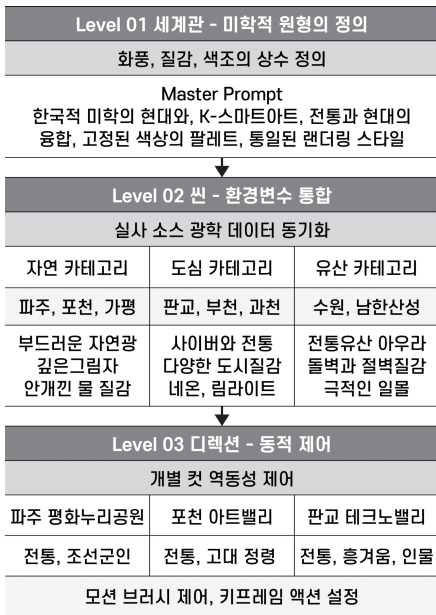
[그림 1] 프롬프트텔링 아키텍처의 구조

[그림 2]는 계층형 프롬프트 아키텍처를 실제 영상 제작 과정에 적용하는 흐름을 나타낸 것이다. 본 연구에서는 실사 배경 데이터와 AI 생성 요소를 단계적으로 결합하기 위해 동일한 AI 엔진과 공통 배경 데이터를 기반으로 프롬프트 구조를 적용하였다. 특히 상위 단계에서 설정된 시각 정보가 하위 단계로 공유되는 구조를 통해 장소 변화에도 영상의 스타일과 연출 방향이 유지될 수 있도록 설계하였다. 이와 같은 계층형 프롬프트 구조는 생성형 AI 영상 제작 과정에서 발생하는 시각적 변동성을 완화하고, 실사 공간과 AI 기반 가상 에셋 간의 정합성을 유지하기 위한 제작 구조로 활용될 수 있다. 또한 장소성, 서사성, 연출 요소를 단계적으로 관리함으로써 하이브리드 영상 제작 과정의 일관성과 연출 제어 가능성을 높이는 기획 구조로 기능한다.⁸⁾

6) 김은아, 안종혁, 디지털 영상 미디어의 하이브리드적 표현 양상에 대한 연구(2024), 커뮤니케이션디자인학회, pp.38-40.

7) <https://brunch.co.kr/@tkdan1239/309>

8) 박좌용, 생성형 AI 기반 텍스타일 디자인의 정확성 향상을 위한 프롬프트 엔지니어링 연구(2025),



[그림 2] 생성형 AI 비디오 프롬프트 엔지니어링의 3단계 계층적 설계도

3-2. 서사 기반 프롬프트 라이브러리 구축

계층형 프롬프트 아키텍처를 실제 하이브리드 영상 제작에 적용하기 위해, 경기도 주요 랜드마크 10개소를 대상으로 서사 기반 프롬프트 라이브러리를 구축하였다. 프롬프트 라이브러리는 장소의 물리적 특성과 문화적 맥락을 시가 반복적으로 참조할 수 있는 제작 데이터로 전환하기 위한 구조이다. 이를 통해 지역별 장면이 서로 다른 배경을 가지더라도 캐릭터의 디자인, 영상의 스타일, 조명 톤이 일정하게 유지되도록 하였다.

라이브러리 구축 과정은 세 단계로 진행하였다. 첫째, 경기도 내 주요 랜드마크 10개를 선정하고 각 장소의 실사 영상 및 이미지 데이터를 수집하였다. 둘째, 장소별 역사적·문화적 특징, 공간 분위기, 색채, 조도, 질감 등을 분석하여 서사 키워드와 시각 키워드로 분류하였다. 셋째, 분류된 요소를 유니버스, 씬, 디렉션 레벨에 맞추어 프롬프트 모듈로 정리하였다. 유니버스 레벨에서는 전체 영상의 공통 시각 기준이 되는 화풍, 색채 톤, 캐릭터 스타일, 시대적 분위기를 설정하였다. 씬 레벨에서는 각 랜드마크의 장소성, 조도, 색온도, 공간 질감, 배경 분위기를 반영하였다. 디렉션 레벨에서는 카메라 움직임, 캐릭터 동작, 장면 전환, 조명 효

과 등 개별 장면의 연출 요소를 구성하였다. 이러한 구조는 각 지역의 고유성을 유지하면서도 영상 전체의 시각적 일관성을 확보하기 위한 것이다.⁹⁾



[그림 3] 경기도 랜드마크 서사 라이브러리 데이터셋

[그림 3]은 경기도 랜드마크를 기반으로 구축한 서사 프롬프트 라이브러리 데이터셋을 나타낸 것이다. 본 라이브러리는 장소 정보, 캐릭터 설정, 조명 조건 및 연출 방향을 통합한 제작 모듈로 구성된다. 역사·문화 유형에는 전통성과 상징성을, 자연·생태 유형에는 자연광과 계절감을, 예술·관광 유형에는 색채 변화와 판타지적 연출을, 미래·혁신 유형에는 디지털 이미지와 역동적 카메라 연출을 반영하였다. 이를 통해 장면별 프롬프트를 반복 작성하는 비효율성을 줄이고, 실사 배경과 AI 생성 캐릭터 간의 장소성, 서사성, 조명 및 시각 스타일을 통합적으로 관리할 수 있는 제작 기준을 구축하였다. 프롬프트텔링 기반 계층형 프롬프트 구조의 적용 가능성을 다양한 환경에서 검증하기 위해 경기도 주요 랜드마크 10개소를 연구 대상으로 선정하였다. 연구 대상은 역사·문화,

9) 김재모, 장수민, 조승우, 서지연, 안희준, 생성형 AI 모델들을 사용한 전과경 영상제작 워크플로우 설계 및 평가(2025), 한국방송·미디어공학학회, pp.324-327.

자연·생태, 예술·관광, 미래·혁신 분야를 균형 있게 반영하였으며, 지역 대표성, 문화적 상징성, 시각적 차별성 및 콘텐츠 활용 가능성을 기준으로 선정하였다.

3-3. 실험 환경 및 분석 지표 설정

생성형 AI 기반 하이브리드 영상 제작에서 장면 간 시각적 일관성과 실사 배경 및 가상 에셋 사이의 정합성을 분석하기 위해 평가 지표를 설정하였다. 생성형 AI 영상은 동일하거나 유사한 프롬프트를 사용하더라도 장면마다 캐릭터 형태, 색채, 조명, 질감 등이 달라지는 특성을 가지므로, 영상 전체의 일관성을 유지할 수 있는 분석 기준이 필요하다. 이에 디자인 항상성, 조도 동기화율, 지형 접지 정밀도, 서사적 적합성을 주요 분석 지표로 설정하였다. 디자인 항상성은 장면 전환 이후에도 캐릭터의 외형과 시각적 스타일이 유지되는 정도를 의미한다. 조도 동기화율은 실사 배경의 광원과 AI 생성 에셋의 조명 표현이 얼마나 자연스럽게 연결되는지를 평가하기 위한 지표이다. 지형 접지 정밀도는 캐릭터가 실사 공간의 지면과 원근 구조에 자연스럽게 결합되는 정도를 분석하기 위해 설정하였다. 또한 서사적 적합성은 장소의 분위기와 캐릭터 연출이 영상의 맥락 안에서 조화롭게 연결되는지를 평가하기 위한 기준이다. 아울러 정량적 분석만으로 판단하기 어려운 시각적 완성도와 연출 적합성을 보완하기 위해 전문가 평가를 병행하였다. 전문가 평가는 시각적 정합성, 서사적 적합성, 미학적 완성도를 중심으로 구성하였으며, 세부 평가 항목은 [표 4]와 같다.

[표 4] 전문가 평가 지표 및 질문지 구성

평가 지표	세부 평가 문항 (질문 내용)	비고
시각적 정합성	실사 배경의 광원(조도, 색온도)과 가상 에셋의 질감이 이질감 없이 통합되었는가?	기술적 정합
서사적 적합성	각 장소의 역사적·문화적 맥락이 가상 캐릭터의 외형 및 동작 서사와 개연성 있게 결합되었는가?	기획적 정합
미학적 완성도	생성형 AI의 무작위적 노이즈가 통제되어 상용화 가능한 수준의 고품질 시각 결과물을 도출하였는가?	디자인 완성도

정량 분석은 디자인 항상성, 조도 동기화율, 지형 접지 정밀도 및 서사적 적합성을 중심으로 수행하였다. 디자인 항상성은 캐릭터의 외형, 복식 및 색채의

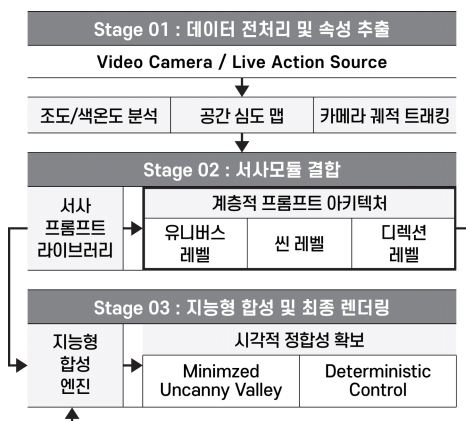
유지 정도를, 조도 동기화율은 실사 배경과 AI 생성 캐릭터 간 광원 및 명암 표현의 일치 정도를, 지형 접지 정밀도는 캐릭터와 실사 공간의 자연스러운 결합 정도를 기준으로 평가하였다. 서사적 적합성은 전문가 평가를 바탕으로 측정하였으며, 영상디자이너 및 AI 콘텐츠 분야 전문가 10인을 대상으로 리커트 5점 척도를 활용하여 분석하였다.

4. 하이브리드 제작 파이프라인 구축과 구현

4-1. 실사 기반 하이브리드 워크플로우 설계

생성형 AI 영상 제작 과정에서 발생하는 시각적 변동성과 실사가상 에셋 간의 이질감을 최소화하기 위해 실사 기반 프롬프트 라이브러리를 중심으로 한 하이브리드 제작 워크플로우를 설계하였다. 기존 단일 프롬프트 방식은 장면 전환 시 캐릭터 형태, 조명, 색채 표현이 불규칙하게 변화하고 실사 배경과 AI 생성 요소 간 정합성이 저하되는 한계를 보였다. 이에 프롬프트를 단순 생성 명령이 아닌 제작 구조로 활용하여 생성 과정을 단계적으로 제어하고자 하였다.

워크플로우는 실사 데이터 분석, 프롬프트 라이브러리 적용, AI 생성 및 합성의 세 단계로 구성하였다. 실사 데이터 분석 단계에서는 경기도 주요 랜드마크 10개소의 조도, 색온도 및 공간 분위기 정보를 추출하였으며, 이를 실사 공간과 가상 캐릭터의 정합성을 위한 환경 데이터로 활용하였다. 프롬프트 라이브러리 단계에서는 유니버스 레벨에서 공통 화풍과 캐릭터 스타일을 설정하고, 씬 레벨에서 장소별 공간 및 광원 특성을 반영하였으며, 디렉션 레벨에서는 카메라 움직임과 장면 연출 요소를 구성하였다.



[그림 4] 워크플로우 설계도

[그림 4]는 본 연구의 하이브리드 영상 제작 워크플로우를 나타낸 것으로, 실사 환경 데이터와 프롬프트 라이브러리를 연계하여 생성형 AI의 출력 방향을 단계적으로 제어하는 구조를 보여준다.

제작 과정의 편차를 최소화하기 위해 동일한 생성형 AI 엔진과 공통 캐릭터 조건을 유지하였으며, 동일한 시드 값을 적용하여 프레임 간 불필요한 스타일 변형을 줄이고자 하였다. 이를 통해 장소가 변화하더라도 캐릭터의 시각적 정체성과 장면의 연속성이 안정적으로 유지될 수 있도록 하였다. 이와 같은 워크플로우는 생성형 AI의 우연적 생성 구조를 반복 생성 중심의 방식에서 벗어나, 서사와 연출을 통합적으로 관리할 수 있는 제작 프로세스로 확장하기 위한 목적을 가진다. 특히 실사 공간과 AI 생성 캐릭터를 결합하는 과정에서 요구되는 시각적 연속성과 공간적 정합성을 보다 안정적으로 유지할 수 있도록 설계되었다.¹⁰⁾

4-2. 경기도 랜드마크 영상 콘텐츠 구현

앞서 설계한 계층형 프롬프트 아키텍처와 서사 기반 프롬프트 라이브러리를 실제 영상 제작 과정에 적용하여 경기도 주요 랜드마크를 중심으로 하이브리드 영상 콘텐츠를 구현하였다. 구현 과정에서는 장소의 물리적 환경과 문화적 맥락이 AI 생성 캐릭터 및 연출 요소와 자연스럽게 결합될 수 있도록 장소별 서사 전략을 차등 적용하였다.

구현 대상은 역사·문화, 자연·생태, 예술·관광, 미래·혁신 유형의 공간으로 구성하였다. 각 장소는 단순 배경 이미지가 아니라 장소 고유의 분위기와 상징성을 포함하는 서사 공간으로 설정되었으며, 프롬프트 구조 역시 장소 특성에 맞추어 조정하였다. 역사·문화 유형의 공간은 전통성과 시간성을 중심으로 구성하였고, 자연·생태 유형은 자연광과 공간의 유기적 흐름을 강조하였다. 예술·관광 유형은 색채 변화와 판타지적 분위기를 중심으로 연출하였으며, 미래·혁신 유형은 디지털 조명과 역동적인 공간 이미지를 활용하여 시각적 대비를 강화하였다.

[그림 5]는 경기도 10개소를 대상으로 구축한 서사 프롬프트 전략 및 구현 결과를 나타낸 것이다. 각 지역은 서로 다른 공간 특성과 서사 구조를 가지지만, 유니버스 레벨에서 설정한 공통 화풍과 캐릭터 스타

일을 유지하도록 설계하여 영상 전체의 시각적 통일성이 유지될 수 있도록 하였다.

랜드마크	서사 테마 및 상징성	기술적 제어 모듈	정합성 결과
파주 평화누rian공원	평화의 확산과 바람의 흐름	Wind-swaying texture	능선과 바람개비의 곡선미가 캐릭터와 유기적 조화
포천 아트밸리	기암괴석과 신비로운 영산	stone-dust particles	암벽의 거친 질감과 캐릭터의 신비로운 분위기가 정합
가평 자라섬	강변의 평온과 자연의 생명력	River-mist effect	수변 풍경과 캐릭터의 온화한 톤이 자연스럽게 일치
과천 서울랜드	축제의 활기와 원색적 즐거움	Vibrant primary colors	테마파크 환경과 캐릭터의 높은 대비 확보
광명 동굴	지하의 신비와 빛의 파동	Subterranean rim light	인공 광원과 캐릭터 금속성 장식 정합
구리 타워	수직적 지향성과 도심의 비상	Skyscraper perspective	타워의 수직 구조물과 캐릭터의 현대적 비례감
남한산성	역사적 요새와 장군의 기개	Fortress stone texture	성곽의 역사적 무게감과 캐릭터의 강인한 서사 결합
다산 생대공원	실학의 정신과 선비의 풍류	Traditional scholar attire	벚꽃과 전통 모티프 등의 시각적 통일성 정합
부천 영상단지	디지털 영상과 미래형 전통	Cyber-traditional remix	미래형 건축 디자인과 캐릭터의 지능적 융합 강조
수원 통영거리	로컬 미식 문화와 시민의 행복	Realistic food texture	서민적 공간의 일상성과 캐릭터의 해학적 표현 정합

[그림 5] 경기도 10개소 서사 프롬프트 현황 및 실증 결과

[표 5]는 지역 유형에 따른 주요 서사 키워드와 조도 동기화 결과를 정리한 것이다. 역사·문화 유형의 남한산성은 전통 색조와 안정적인 카메라 구성을 통해 장소의 역사성을 강조하였으며, 자연·생태 유형의 포천 아트밸리는 자연광과 반사광 표현을 활용하여 공간의 개방감을 강화하였다. 미래·혁신 유형의 광명동굴은 디지털 조명과 공간 대비 효과를 중심으로 연출하여 높은 조도 동기화율을 나타냈다.

[표 5] 지역 유형별 서사 전략 및 구현 결과

지역 구분	랜드마크 샘플	서사 테마	조도 동기화율
역사/문화	남한산성 등	전통의 현대적 재해석 및 호국 서사	95.2%
자연/생태	다산 생대공원 등	자연과의 공존 및 서정적 서사	93.8%
예술/관광	포천 아트밸리 등	예술적 판타지와 창의적 변주	92.5%
미래/혁신	광명동굴 등	디지털 전환 및 미래 지향적 서사	94.9%

장소 유형에 따라 서로 다른 서사 키워드와 연출 전략이 적용되었음에도 불구하고, 실사 배경과 AI

10) 임장한, AI 영화 제작을 위한 캐릭터 일관성에 관한 연구 - 얼굴과 의상을 중심으로(2024), 한국방송-미디어공학회, pp.478-480.

생성 캐릭터 사이의 조도 및 색온도 연결성은 비교적 안정적으로 유지되었다. 이는 계층형 프롬프트 구조가 장소별 환경 변화 속에서도 공통 시각 기준을 유지하는 데 효과적으로 작용하였음을 보여준다.



[그림 6] 남한산성과 포천 아트밸리



[그림 7] 경기도 내 8개 랜드마크

[그림 6]부터 [그림 7]까지는 실제 구현된 하이브리드 영상 장면 일부를 정리한 것이다. 남한산성 구현 사례에서는 전통 의상과 저채도 색조를 활용하여 역사적 분위기를 강조하였으며, 포천 아트밸리 구현 사례에서는 자연 암벽과 반사광 효과를 활용하여 공간의 입체감을 강화하였다. 또한 광명동굴 구현 사례에서는 실사 공간 내부의 조명 구조를 AI 캐릭터의 반사광 표현에 반영함으로써 실사 배경과 가상 에셋 간의 시각적 이질감을 최소화하였다. 이와 같은 구현 결과는 프롬프트 라이브러리가 단순한 명령어 집합이 아니라 장소 정보와 연출 요소를 통합적으로 관리하는 제작 데이터로 기능할 수 있음을 보여준다. 또한 계층형 프롬프트 구조를 기반으로 장소별 환경 정보

를 단계적으로 제어함으로써, 실사 공간과 AI 생성 캐릭터 사이의 시각적 연속성과 공간적 정합성을 보다 안정적으로 유지할 수 있음을 확인하였다.

4-3. 대조군 · 실험군 비교 분석

계층형 프롬프트 아키텍처의 효과를 확인하기 위해 단일 프롬프트 방식으로 생성한 대조군과 본 연구에서 제안한 계층형 프롬프트텔링 방식을 적용한 실험군을 비교 분석하였다. 대조군은 장소별 장면 생성 시 하나의 통합 프롬프트를 입력하는 방식으로 제작하였으며, 실험군은 유니버스, 씬, 디렉션 단계의 프롬프트 구조를 적용하여 영상 전체의 공통 스타일과 장소별 환경 정보, 장면별 연출 요소가 단계적으로 반영되도록 하였다. 비교 분석은 3-3에서 설정한 디자인 향상성, 조도 동기화율, 지형 접지 정밀도, 서사적 적합성의 네 가지 지표를 중심으로 진행하였다. 분석 결과, 단일 프롬프트 방식은 장소가 전환될 때 캐릭터의 복식, 색채, 렌더링 스타일이 불규칙하게 변화하는 경향을 보였다. 이는 장면마다 시가 프롬프트를 개별적으로 해석하면서 캐릭터와 배경의 시각적 기준이 일관되게 유지되지 못했기 때문으로 판단된다. 반면 계층형 프롬프트 아키텍처를 적용한 실험군은 상위 단계에서 설정한 캐릭터 스타일과 조명 기준이 하위 단계에 지속적으로 반영되면서 장면 간 시각적 변동이 상대적으로 낮게 나타났다. 특히 유니버스 레벨에서 고정한 화풍과 캐릭터 기준은 디자인 향상성 유지에 영향을 주었고, 씬 레벨에서 반영한 장소별 조도와 색온도 정보는 실사 배경과 AI 생성 에셋 간의 광학적 이질감을 줄이는 데 기여하였다. 또한 디렉션 레벨에서 카메라 움직임과 장면 전환 요소를 설정함으로써 캐릭터가 실사 공간 안에 보다 자연스럽게 결합되는 결과를 보였다.

[표 6] 대조군과 실험군의 지표별 정량적 비교 분석

분석 지표	대조군 단일 프롬프트	실험군 본 아키텍처	개선 수치 Gap
디자인 향상성	57.5%	92.2%	+34.7%
조도 동기화율	48.3%	94.1%	+45.8%
지형 접지 정밀도	61.2%	88.5%	+27.3%
서사적 적합성	45.0%	91.0%	+46.0%

[표 6]에서 확인할 수 있듯이, 디자인 향상성은 대조군 57.5%에서 실험군 92.2%로 향상되었으며, 조도 동기화율은 48.3%에서 94.1%로 증가하였다. 지형 겹치 정밀도는 61.2%에서 88.5%로 개선되었고, 서사적 적합성은 45.0%에서 91.0%로 향상되었다.

분석 결과는 영상디자인 및 AI 콘텐츠 분야 전문가 10인의 평가와 장면별 시각 정합성 분석 결과를 종합하여 산출하였다. 결과는 계층형 프롬프트 구조가 단일 프롬프트 방식에 비해 장면 간 시각적 기준을 안정적으로 유지하는 데 효과적임을 보여준다. 특히 조도 동기화율과 서사적 적합성의 개선 폭이 크게 나타난 점은, 프롬프트텔링 방식이 단순히 캐릭터의 외형을 유지하는 데 그치지 않고 실사 공간의 환경 정보와 장소의 서사적 맥락을 함께 반영할 수 있음을 의미한다. 따라서 본 연구의 계층형 프롬프트 아키텍처는 하이브리드 영상 제작에서 시각적 일관성과 공간적 정합성을 확보하기 위한 유효한 제작 방법으로 해석할 수 있다.

4-4. 전문가 평가 결과 분석

대조군·실험군 비교 분석과 함께, 생성 결과의 시각적 완성도와 실무적 활용 가능성을 검토하기 위해 전문가 평가를 실시하였다. 정량 지표는 디자인 향상성, 조도 동기화율과 같은 수치적 차이를 확인하는데 유효하지만, 영상 콘텐츠의 완성도는 장면의 분위기, 시각적 몰입감, 서사적 설득력과 같은 정성적 요소도 함께 고려되어야 한다. 이에 본 연구에서는 영상디자인 및 AI 콘텐츠 분야 전문가 10인을 대상으로 리커트 5점 척도 평가를 진행하였다.

평가 항목은 앞서 3-3에서 설정한 시각적 정합성, 서사적 적합성, 미학적 완성도를 기준으로 구성하였다. 시각적 정합성은 실사 배경과 AI 생성 캐릭터가 하나의 장면 안에서 자연스럽게 결합되었는지를 확인하기 위한 항목이며, 서사적 적합성은 장소의 역사적·문화적 맥락과 캐릭터 연출이 개연성 있게 연결되는지를 판단하기 위한 항목이다. 미학적 완성도는 생성형 AI 특유의 왜곡이나 시각적 노이즈가 어느 정도 통제되었는지를 검토하기 위한 기준으로 활용하였다.

전문가 평가 결과, 시각적 정합성은 평균 4.58점, 서사적 적합성은 평균 4.55점으로 나타났다. 이는 실험군 결과물이 실사 배경과 AI 생성 캐릭터 간의 결합 측면에서 비교적 높은 완성도를 확보하였음을 의미한다. 특히 전문가들은 장소별 조명 톤과 캐릭터의 색채 표현이 자연스럽게 연결되었고, 지역별 서사 키

워드가 장면 분위기 형성에 효과적으로 반영되었다는 점을 긍정적으로 평가하였다.

[표 7] 전문가 10인 리커트 척도 평가 데이터 (5점 만점)

구분	전문가 그룹 구성	시각적 정합성	서사적 적합성
전문가1	영상디자인 전공 교수	4.8	4.7
전문가2	영상디자인 전공 교수	4.5	4.4
전문가3	AI 콘텐츠 테크니컬 디렉터	4.7	4.6
전문가4	AI 시각예술가	4.6	4.8
전문가5	문화콘텐츠 기획 전문가	4.4	4.6
전문가6	미디어아트 디렉터	4.6	4.5
전문가7	인터랙티브 디자인 전문가	4.5	4.4
전문가8	디자인학 박사	4.7	4.5
전문가9	광고영상 프로듀서	4.4	4.5
전문가10	디지털 콘텐츠 디자인 연구원	4.6	4.5
평균	항목별 전체 평균 점수	4.58	4.55

시각적 정합성 항목에서는 실사 배경의 조도와 캐릭터의 렌더링 톤이 비교적 안정적으로 조율되었다는 평가가 높게 나타났다. 이는 씬 레벨에서 장소별 광원과 색온도 정보를 반영한 결과로 해석할 수 있다. 서사적 적합성 항목에서는 남한산성, 포천 아트밸리, 광명동굴 등 각 장소의 성격에 맞추어 전통성, 자연성, 미래성 등의 키워드가 캐릭터 연출과 결합된 점이 긍정적으로 작용하였다. 다만 일부 전문가들은 캐릭터의 세부 표정, 손동작, 움직임 전환 등에서 생성형 AI 특유의 미세한 왜곡이 나타나는 장면이 있다고 지적하였다. 이는 생성형 AI가 전체 분위기와 스타일을 형성하는 데는 효과적이지만, 세부 동작의 정밀 제어나 장시간 서사 흐름 유지에는 여전히 한계가 있음을 보여준다. 따라서 향후 제작 과정에서는 프롬프트 구조뿐 아니라 후반 편집, 프레임 보정, 세부 모션 제어 과정이 함께 보완될 필요가 있다.

전문가 평가는 계층형 프롬프트텔링 방식이 단순한 시각 효과 생성에 머무르지 않고, 실사 공간과 AI 생성 캐릭터를 하나의 장면 안에서 비교적 안정적으로 통합할 수 있음을 확인해준다. 또한 정량 분석에서 나타난 디자인 향상성과 조도 동기화율의 개선 결과가 실제 전문가의 시각적 판단에서도 긍정적으로 수용되었음을 보여준다. 이는 제안한 프롬프트텔링 기반 하이브리드 영상 제작 방식이 실무적 적용 가능성을 가진 제작 방법론으로 활용될 수 있음을 시사한다.

5. 결론

본 연구는 생성형 AI 영상 제작 과정에서 발생하는 장면 간 시각적 불일치와 실사가상 에셋 간의 정합성 문제를 개선하기 위해 프롬프트텔링 기반의 계층형 프롬프트 아키텍처를 제안하고, 이를 활용한 하이브리드 영상 제작 방식을 연구하였다. 기존 생성형 AI 영상 제작은 단일 프롬프트 중심의 생성 구조로 인해 장면 전환 과정에서 캐릭터의 형태, 색채, 조명, 렌더링 스타일 등이 불규칙하게 변화하는 한계를 보였다. 이에 본 연구는 프롬프트를 단순 입력문이 아닌 서사와 연출 구조를 포함하는 제작 체계로 확장하여 생성 과정의 시각적 변동성을 제어하고자 하였다. 이를 위해 유니버스, 씬, 디렉션 단계로 구성된 계층형 프롬프트 구조를 설계하고, 경기도 주요 랜드마크를 기반으로 서사 프롬프트 라이브러리를 구축하였다. 또한 실사 영상 데이터와 AI 생성 캐릭터를 결합하는 하이브리드 제작 워크플로우를 구성하여, 단일 프롬프트 방식과 계층형 프롬프트텔링 방식의 결과를 비교 분석하였다.

실험 결과, 계층형 프롬프트 구조를 적용한 실험군은 디자인 향상성, 조도 동기화율, 지형 접지 정밀도, 서사적 적합성 항목에서 모두 향상된 결과를 나타냈다. 특히 장면 전환 이후 캐릭터의 시각적 스타일과 실사 배경의 조명 환경이 비교적 안정적으로 유지되었으며, 전문가 평가에서도 시각적 정합성과 서사적 적합성 측면에서 높은 평가를 기록하였다. 이는 프롬프트를 계층 구조로 설계할 경우 생성형 AI의 확률론적 특성을 일정 부분 제어할 수 있으며, 실사 공간과 AI 생성 에셋 사이의 시각적 연결성을 강화할 수 있음을 보여준다. 또한 본 연구는 생성형 AI 기반 영상 제작에서 프롬프트를 단순 생성 명령이 아닌 제작 구조와 연출 기준을 통합적으로 관리하는 기획 체계로 활용하였다는 점에서 의미를 가진다. 특히 장소 기반 서사와 시각 연출을 연결하는 프롬프트 라이브러리 구조를 제안함으로써, 생성형 AI 환경에서도 영상 제작자의 기획 의도와 연출 방향을 보다 안정적으로 유지할 수 있는 가능성을 제시하였다. 다만 특정 지역과 제한된 생성 환경을 중심으로 실험을 진행하였다는 한계를 가진다. 또한 생성형 AI 특유의 세부 모션 왜곡과 장면 전환 과정의 미세한 불안정성이 일부 확인되었으며, 이는 향후 후반 보정 기술과 모션 제어 연구를 통해 보완될 필요가 있다. 향후 다양한 생성형 AI 모델과 실시간 인터랙션 기반 콘텐츠 환경을 활용한 후속 연구가 이루어질 필요가 있다.

본 연구는 생성형 AI 기반 하이브리드 영상 제작에서 프롬프트텔링 기반 계층형 구조가 장면 간 시각적 일관성과 실사가상 에셋 간의 정합성을 향상시키는 데 효과적으로 활용될 수 있음을 확인하였다. 이는 생성형 AI 영상 제작이 단순 자동 생성 중심에서 벗어나, 서사와 연출을 구조적으로 제어하는 설계 기반 제작 방식으로 확장될 수 있음을 시사한다.

참고문헌

1. 김은정, 김정윤, 감성형 AI 게임 캐릭터 시스템 설계 및 구현 연구(2025), 차세대컨버전스정보서비스기술논문지
2. 이연우, AI 기반 이미지 생성 프롬프트 활용을 통한 영상스토리텔링 교육 모델 연구(2025), 멀티미디어학회논문지
3. 주종우, 2035년까지의 글로벌 사진 산업 전망과 생성형 AI(2025), 현대사진영상학회논문집
4. 김은아, 안종혁, 디지털 영상 미디어의 하이브리드적 표현 양상에 대한 연구(2024), 커뮤니케이션디자인학회
5. 박좌용, 생성형 AI 기반 텍스타일 디자인의 정확성 향상을 위한 프롬프트 엔지니어링 연구(2025), 학위논문
6. 김재모, 장수민, 조승우, 서지연, 안희준, 생성형 AI 모델들을 사용한 전과정 영상제작 워크플로우 설계 및 평가(2025), 한국방송-미디어공학회
7. 임장한, AI 영화 제작을 위한 캐릭터 일관성에 관한 연구 - 얼굴과 의상을 중심으로(2024), 한국방송-미디어공학회
8. brunch.co.kr
9. rarang.com
10. wikidocs.net