한국컴퓨터그래픽스학회
Korea Computer Graphics Society Vol. 32, No. 3, P. 143 ~153

웹툰의 영상화를 위한 생성형 AI 기반 제작 플랫폼

이영현° 박준형 윤대건 배종환 박상훈*

서강대학교 가상융합전문대학원

20hyun@u.sogang.ac.kr, {pjh7083, ajrjs9902, bjh0309, mshpark}@sogang.ac.kr

A Generative AI-Based Production Platform for Webtoon-to-Video Conversion

Yeonghyun Lee° Junhyeong Park Daegeon Yoon Jonghwan Bae Sanghun Park*

Graduate School of Virtual Convergence, Sogang University

요 약

본 논문은 웹툰 콘텐츠를 숏폼(short-form) 영상으로 자동 변환하는 생성형 AI 기반 통합 파이프라인 및 플랫폼을 제안한다. 제안 시스템은 제작자 플랫폼, 소비자 플랫폼, 공유 백엔드의 3 티어(tier) 구조로 설계되며, 제작자 플랫폼은 웹툰 입력으로부터 컷 분할, 대규모 언어 모델(LLM: large language model) 기반 스토리·감정 분석, 실사화 이미지 생성, 영상 클립 생성, 트랜지션·BGM 편집까지의 7단계를 대부분 자동으로 처리한다. 사용자는 컷 선별과 일부 수정 과정에만 개입한다. 특히 회차 전체의 스토리라인을 분석하여 후속 생성 단계마다 참조함으로써 고도화된 프롬프팅을 구현하고, 시퀀스의 첫 컷을 스타일 앵커(style anchor)로, 사용자가 제공한 캐릭터 레퍼런스 이미지를 가이드로 함께 활용하여 색감·캐릭터 일관성을 확보한다. 본 연구는 영상 제작 비전문가도 복잡한 프롬프팅 지식 없이 고품질 숏폼 콘텐츠를 제작·배포할 수 있는 엔드투엔드(end-to-end) 환경을 제안한다는 점에서 의의가 있다.

Abstract

This paper proposes an AI-based integrated pipeline and platform that automatically converts webtoon content into short-form videos. The system adopts a 3-tier architecture composed of a producer platform, a consumer platform, and a shared backend. The producer platform carries out a 7-stage pipeline—cut segmentation, LLM-based story and emotion analysis, photorealistic image generation, video clip generation, and transition/BGM editing—mostly automatically from a webtoon input, while the user intervenes only in cut selection and minor edits. In particular, it analyzes the storyline of an entire episode and references it at every subsequent stage to realize advanced automatic prompting, and uses the first cut of a sequence as a style anchor together with a user-provided character reference image to preserve color tone and character consistency. The significance of this work lies in proposing an end-to-end environment in which non-expert creators can produce and distribute high-quality short-form content without complex prompting knowledge.

키워드: 웹툰, 생성형 AI, 대규모 언어 모델, 프롬프팅, 스타일 앵커, 캐릭터 일관성
Keywords: Webtoon, Generative AI, Large Language Model, Prompting, Style Anchor, Character Consistency

*corresponding author: Sanghun Park / Graduate School of Virtual Convergence, Sogang University (mshpark@sogang.ac.kr)

Received : 2026.05.29./ Review completed : 1st 2026.06.07. 2nd 2026.06.25./ Accepted : 2026.06.29.
DOI : 10.15701/kcgs.2026.32.3.143
ISSN : 1975-7883(Print)/2383-529X(Online)

1. 서론

웹툰은 한국을 중심으로 전 세계적으로 확산된 디지털 만화 형식으로, 검증된 서사와 두터운 팬덤을 보유한 원천 지식재산권(IP: intellectual property)이다. 세로 스크롤 웹툰의 컷 구성은 영화 스토리보드에 준하는 시각적 연출 정보를 이미 내포하고 있어, 영상화 프리프로덕션의 비용과 리스크를 절감할 수 있는 구조를 갖는다. 나아가 TikTok, YouTube Shorts, Instagram Reels 등 숏폼 플랫폼의 확산은 웹툰의 단편 영상화를 새로운 콘텐츠 유통 전략으로 부각시키고 있다 [1].

이러한 매체 간 변환은 본 연구가 처음 시도하는 작업이 아니다. 텍스트(소설)에서 영상으로, 일반 만화에서 애니메이션으로, 영상에서 만화로 같은 서사를 옮기는 연구가 학술적·산업적으로 오랜 기간 누적되어 왔으며 [2], 최근에는 대규모 언어 모델과 영상 생성 모델의 발전과 함께 영상 제작 워크플로우 자체가 인간-AI 협업 구조로 빠르게 재편되고 있다 [3]. 본 연구는 바로 이 흐름 위에 자리하면서, 출발 매체로 한국형 세로 스크롤 웹툰을 채택하고 스토리 분석·실사화 이미지·영상 생성까지 최신 생성형 AI 모델군을 단일 파이프라인 안에 결합함으로써 인접 매체 변환 과제와 구분되는 새로운 변환 문제를 다룬다.

그러나 생성형 AI를 활용한 영상 제작은 여전히 비전문가에게 진입 장벽이 높다. 현 시점의 영상 생성 모델은 블랙박스 형태로 동작하여 동일한 입력에도 무작위에 가까운 결과를 산출하는 경향이 있으며, 원하는 품질의 영상을 얻기 위해서는 전문가 수준의 정교한 프롬프팅과 반복적인 시행착오가 요구된다. 컷 분할, 대사 추출, 장면 구성, 영상 편집 등 단계별 작업이 개별적으로 수행되어야 하고, 각 단계에서 요구되는 프롬프팅 역량과 누적되는 피로도는 일반 창작자의 진입을 어렵게 한다.

본 논문에서는 이러한 문제를 완화하기 위해, 웹툰 한 회차를 입력하면 대규모 언어 모델(LLM: large language model) 기반 스토리 분석과 생성형 AI API를 결합하여 숏폼 영상까지 자동 변환하는 엔드투엔드(end-to-end) 파이프라인과 운영 플랫폼을 제안한다. 본 시스템은 사용자가 컷 선별과 일부 수정 과정에만 개입하도록 설계되어, 비전문가도 최소한의 조작만으로 영상을 제작할 수 있다. 결과의 일관성은 (i) 회차 전체의 스토리 맥락을 추출하여 후속 단계마다 참조하는 자동 프롬프팅, (ii) 사용자가 제공한 캐릭터 레퍼런스 이미지, (iii) 시퀀스의 첫 컷을 기준으로 삼는 스타일 앵커(style anchor)의 세 장치를 통해 유지된다. 주요 기여는 다음과 같다.

- **자동화 7단계 파이프라인**: 웹툰 입력에서 영상 편집까지 대부분을 자동 처리하고 사용자 개입을 최소화하는 통합 워크플로우.
- **스토리 맥락 기반 자동 프롬프팅**: 회차 전반의 서사 정보를 후속 생성 단계에 일관 주입하여 비전문가도 고품질 결과를 얻도록 함.
- **일관성 보장 기법**: 캐릭터 레퍼런스 이미지와 스타일 앵커의 결합으로 다중 컷에 걸친 색감·캐릭터 일관성 확보.
- **통합 품질 지표 기반 게이트**: 의미적 정합 지표를 결합한 종합 점수를 시스템 내부에서 산출하여, 임계값 미달 시 재생성을 트리거함으로써 출력 안정성을 높임.

2. 관련 연구

본 절에서는 (1) 서사로부터 일관성 있는 영상을 생성하는 연구, (2) 생성 품질을 좌우하는 LLM 기반 자동 프롬프팅 연구, (3) 만화와 영상 간 상호 변환 연구, (4) 생성 결과의 품질을 평가하는 정량 지표들을 차례로 검토한다.

2.1 서사 기반 영상 생성

텍스트 서사로부터 일관성 있는 영상을 생성하는 연구는 생성형 AI의 발전과 함께 급속히 성장해 왔다. 초기 연구는 텍스트-이미지(T2I) 지식을 텍스트-영상(T2V) 생성으로 전이하거나, LLM을 프레임 수준의 연출자(director)로 활용하는 방향으로 전개되었다 [4, 5]. 최근에는 LLM이 다단계 파이프라인을 오케스트레이션하여 장면 분해·스토리보드·영상 합성을 자동으로 처리하는 시스템, 키프레임을 먼저 계획한 뒤 시간적 동작을 합성하는 분리형 접근, 단일 캐릭터 스케치로부터 스토리 주도 애니메이션을 생성하는 멀티모달 LLM 기반 시스템 등이 제안되어 시간적 일관성과 캐릭터 일관성 측면에서 향상된 성능을 보고하고 있다 [6, 7, 8].

2.2 LLM 기반 자동 프롬프팅

생성형 AI의 출력 품질은 입력 프롬프트의 질에 크게 의존하므로, 이를 자동화하는 연구가 활발하다. 대표적으로 APE(Automatic Prompt Engineer)는 LLM으로 지시 프롬프트를 자동 생성·선택하여 인간 작성 수준에 견주는 성능을 달성하였고, 이후 연구들은 비전 단서를 결합한 멀티모달 프롬프트 최적화 등으로 발전해 왔다 [9, 10]. 이러한 흐름은 서사 맥락에 기반한 자동 프롬프팅이 이미지·영상 생성 품질 향상에 효과적임을 일관되게 시사한다.

2.3 만화-영상 상호 변환

영상과 만화 사이의 변환 연구는 양방향으로 축적되어 있다. 영상에서 만화를 추출하는 방향에서는 자막과 대화 기반의 감정 인식 말풍선·다중 페이지 레이아웃 생성, 콘텐츠 인식 레이아웃 최적화로 망가(manga: 일본식 만화 스타일) 스타일을 자동 구성하는 시스템이 제안되었다 [11, 12]. 다만 기존 만화→영상 연구는 대체로 페이지형 코믹스를 가정하여 패널의 독해 순서를 먼저 복원해야 하는 반면, 본 연구가 대상으로 하는 한국형 세로 스크롤 웹툰은 컷의 독해 순서가 위에서 아래로 결정되고, 각 컷이 단일 컷 단위로 이미 화면을 가득 채우는 구도여서 영상 한 클립의 시작 프레임으로 거의 그대로 사용 가능하며, 회차 단위 발행

구조가 영상 합성의 자연스러운 단위와 일치한다. 본 연구는 이 매체 고유 특성을 영상 합성의 제약이자 신호로 적극 활용한다.

2.4 생성 품질 평가 지표

생성 결과의 품질을 정량적으로 측정하기 위해 이미지·텍스트·영상을 임베딩하거나 모델에 질의해 점수를 산출하는 다섯 가지 대표 지표가 알려져 있다. 본 절에서는 각 지표의 작동 방식과 강점·한계를 검토한 뒤, 본 연구의 평가 대상에 가장 부합하는 조합을 식별한다.

(1) **CLIP**: 이미지와 텍스트를 동일한 임베딩 공간에 사상하고 두 임베딩의 코사인 유사도(cosine similarity)로 의미적 정합도를 측정한다 [13]. 이미지-텍스트 평가의 사실상 표준으로 의미 정합에 강하지만, 픽셀 수준의 미세 차이는 반영하지 않는다.

(2) **CLIP-I**: 같은 CLIP 이미지 인코더로 두 이미지를 비교하는 변형으로, DreamBooth가 생성 이미지의 대상(subject) 보존성을 평가하기 위해 정착시킨 프로토콜이다 [14]. 픽셀이 아닌 의미 수준에서 보존성을 측정하므로 도메인 격차가 큰 변환(예: 일러스트→실사)의 비교에도 유효하나, 격차가 클수록 절대값은 낮게 형성되어 상대 비교 지표로 활용되는 것이 일반적이다.

(3) **LPIPS**: 사전학습 CNN의 중간 특성으로 두 이미지 간 지각 거리를 산출한다 [15]. 픽셀-수준 시각적 유사성을 잘 반영하지만, 시각 양식이 근본적으로 바뀌는 의미적 변환(웹툰→실사)에서는 변별력이 낮다.

(4) **ViCLIP**: InternVid 데이터셋과 함께 공개된 영상-텍스트 임베딩 모델로, 영상과 텍스트를 각각 벡터화해 코사인 유사도를 측정한다 [16]. 영상-텍스트 평가의 대표 지표이나, 32 토큰의 텍스트 길이 제한이 보고된다.

(5) **VQAScore**: 비전-언어 모델에게 "이 시각 자료가 주어진 내용을 보여주는가"를 질의하고 응답 확률(0∼1)을 점수로 사용한다 [17]. 인물·장소·표정·동작이 결합된 복합 서술과의 정합 평가에서 강점을 보이며, 영상의 다속성 내용 검증에 적합하다.

본 연구의 평가 대상은 (i) 컷 이미지↔생성 스크립트(이미지-텍스트), (ii) 원본 컷↔실사화 이미지(이미지-이미지, 도메인 격차 큼), (iii) 생성 영상↔컷 서술(영상-내용, 복합 서술)의 세 갈래로, 단일 지표가 모두를 다루지 못한다. 이에 본 연구는 이미지 보존성 평가에 강한 **CLIP-I**와, 단락 단위 긴 서술·다속성 복합 서술 평가에 강한 **VQAScore**를 조합한다(VQAScore는 스토리보드와 비디오 단계 모두에 적용). **CLIP**은 표준 이미지-텍스트 정합 지표이나 77 토큰의 길이 한도가 스토리보드의 단락 서술을 단일 벡터로 표현하기 어렵게 만들어 채택하지 않는다. **LPIPS**는 의미적 변환 평가에 변별력이 낮아 채택하지 않으며, **ViCLIP**은 텍스트 길이 한도로 주 지표가 아닌 보조 지표로만 산출한다.

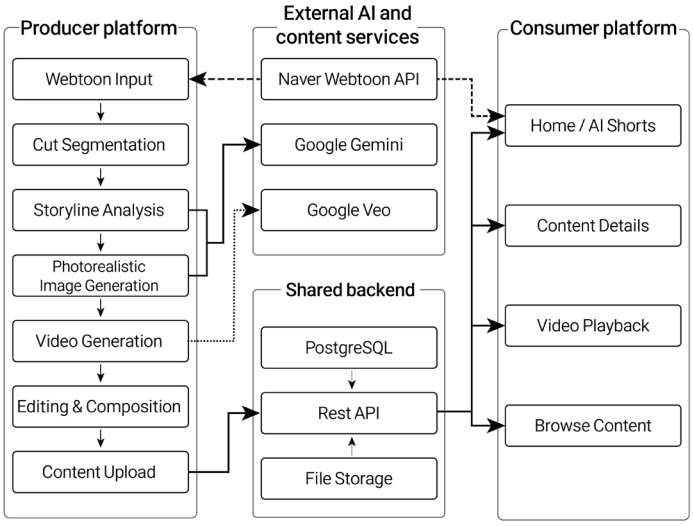

Figure 1: Overall system architecture: a 3-tier design (producer platform, consumer platform, shared backend) and the 7-stage processing pipeline of the producer platform.

3. 시스템 개요

제안 시스템은 서로 독립적으로 개발·운용되는 세 개의 컴포넌트로 구성된 3 티어(tier) 아키텍처를 채택한다(Figure 1). 제작자 플랫폼은 웹툰을 입력받아 7단계 파이프라인 전체를 수행하는 영상 제작 도구이고, 소비자 플랫폼은 생성된 AI 영상을 사용자가 감상하는 모바일 앱이며, 공유 백엔드는 두 플랫폼을 잇는 REST API 서버로 콘텐츠·에피소드·영상 파일을 단일 진실 공급원(SSOT: single source of truth)으로 관리한다. 본 시스템은 다음 네 가지 설계 원칙을 따른다.

- **맥락 보존성**: 회차 스토리라인 분석 결과를 구조화된 형태로 유지하여 후속 생성 단계에 일관 주입.
- **사용자 최소 개입**: 자동화 파이프라인 전반에 걸쳐 사용자의 개입은 컷 선별과 일부 수정에 한정.
- **숏폼 최적화**: 모든 단계에서 9:16 세로 비율(1080×1920)을 적용.
- **일관성 유지**: 스타일 앵커와 캐릭터 레퍼런스 이미지의 결합으로 다중 컷에 걸친 색감·캐릭터 일관성 확보.

모델 및 플랫폼 선택 근거. 제안 시스템은 LLM·이미지·영상의 3축으로 구성되며, 본 구현에서는 각 축의 모델로 Google Gemini 2.5 Flash, Gemini 2.5 Flash Image, Veo 3.1 Lite(동영상+오디오)를 채택하였다. Table 1은 본 시스템의 3축에서 채택한 모델과 대표 대안 모델들의 API(application programming interface) 단가를 정리한 것이다. 선정은 최근 학술 벤치마크의 정량 성능, API 단가, 단일 벤더 통합 결제·멀티모달 호환성을 종합적으로 고려하였다. **LLM 축**에서 Gemini 2.5 기술 보고서 [18]는 Flash가 조절 가능한 추론 예산(controllable thinking budget)으로 품질·비용·지연의 균형을 조절할 수 있는 하이브리드 추론 모델임을

Table 1: Representative API pricing across the three pipeline axes (snapshot 2026-05; subject to change). ⋆ = adopted model.

Axis	Model	Price (USD)
LLM	⋆ Gemini 2.5 Flash	0.30 / 2.50 per 1M tok (in/out)
	GPT-5.4	2.50 / 15.00 per 1M tok
	Claude Sonnet 4.6	3.00 / 15.00 per 1M tok
Image	⋆ Gemini 2.5 Flash Image	0.039 / image
	Imagen 4 Standard	0.040 / image
Video	⋆ Veo 3.1 Lite (video+audio)	0.05 / 0.08 per sec (720p / 1080p)
	Veo 3.1 (Standard 1080p)	0.40 / sec
	Runway Gen-4.5	0.12 / sec
	Kling V3-Omni std/pro	0.10 / 0.15 / sec

명시하여, 회차 단위 배치 처리에서의 비용 제어 근거를 제공한다. **이미지 생성 축**에서는 후속 벤치마크 GenEval 2 [19]가 Gemini 2.5 Flash Image의 GenEval 인간 평가 96.7%를 근거로 기존 GenEval이 사실상 포화되었음을 보고하여, 웹툰 컷→실사 변환 품질의 학술적 근거를 갖추었다. **영상 생성 축**에서는 VBench-2.0이 동세대 SOTA(state-of-the-art) 영상 생성 모델군이 화질·시간 일관성 등 표면적 충실도에서 포화 수준에 근접하였음을 보고하여 [20], 영상 생성 모델 전반이 상용 파이프라인 요구를 충족하는 단계에 진입했음을 시사한다. Veo 3.1 Lite는 동영상과 오디오를 동기화하여 한 번에 생성하므로 별도 음향 트랙 합성 없이 숏폼 영상의 기본 음향 컨텍스트를 제공하며, 단가는 $0.05/초(720p)∼$0.08/초(1080p)로 동세대 SOTA 티어에 속하는 Kling V3-Omni($0.10∼$0.15/초) 대비 가격 경쟁력을 갖춘다. 여기에 단일 벤더 통합 결제와 안전 필터·재시도 메커니즘의 안정성이 결합되어 운영 효율성 측면에서도 합리적인 조합이며, 파이프라인은 동일 구조에서 타 모델로 교체 가능하도록 추상화되어 있다.

3.1 제작자 플랫폼

제작자 플랫폼은 프레젠테이션·비즈니스 로직·데이터·외부 서비스의 4계층 구조로 설계된다. 프레젠테이션 계층은 7단계 단계별 UI를 제공하고, 비즈니스 로직 계층은 크롤링·컷 분할·스토리 분석·이미지 생성·영상 생성·편집 모듈을 담당하며, 데이터 계층은 디스크 캐시와 프로젝트 메타데이터를 관리한다. 외부 서비스 계층은 Gemini 2.5 Flash(스토리 분석), Gemini Image(실사화), Veo 3.1 Lite(영상+오디오 생성), Naver Webtoon API(메타데이터)와 연계된다. 단계별 사용자 개입은 컷 선택, 프롬프트 수정, 객체 대체, 트랜지션 조정, BGM 삽입에 한정된다. 제작자 플랫폼 UI 예시는 Figure 2(a)에 제시한다.

3.2 소비자 플랫폼

소비자 플랫폼은 Flutter 기반 모바일 앱으로, 반응형 상태 관리와 라우팅 위에 홈, 콘텐츠 상세, 풀스크린 숏츠 플레이어, 틱톡식 세로 스크롤 AI 피드, 개인화 추천 화면을 제공한다. 핵심 기능은 회차 번호를 기준으로 Naver 웹툰의 라이브 메타데이터(제목·썸네일)와 자체 생성 AI 영상을 융합하여, 실시간 웹툰 정보와 AI 영상이 하나의 화면에서 자연스럽게 통합되도록 하는 이종 데이터 융합이다. 소비자 플랫폼 UI 예시는 Figure 2(b)에 제시한다.

3.3 공유 백엔드

공유 백엔드는 Node.js·Express 기반 REST API 서버로 GCP VM에서 컨테이너로 운영되며, 콘텐츠 목록, 에피소드, 파일 업로드, 인증 엔드포인트를 제공한다. PostgreSQL은 콘텐츠와 에피소드를 단일 진실 공급원으로 관리하고, 콘텐츠 ID 체계를 통해 두 플랫폼 전반의 식별 일관성을 유지한다.

4. 파이프라인 상세 설계

제작자 플랫폼의 7단계 파이프라인은 입력 → 컷 분할 → 스토리라인 분석 → 컷 선택 → 자동 프롬프팅·실사화 → 영상 생성 → 편집·업로드 순으로 진행되며, 각 단계 사이는 구조화된 JSON 인터페이스로 연결된다. 본 파이프라인은 비전문가 사용자의 자원·분량·프롬프팅 부담을 단계별로 분담·흡수하도록 설계되어 사용자 개입을 컷 선택과 일부 수정에 한정하며, 각 절은 해당 단계가 왜 필요한지와 어떤 가이드라인을 따르는지를 함께 기술한다.

4.1 웹툰 입력

입력 단계는 사용자가 영상화하려는 웹툰 회차를 시스템에 제공하는 단계이다. 사용자가 원본 웹툰 이미지를 항상 손에 들고 있다고 가정할 수 없으므로, 본 시스템은 보유 자원의 형태와 무관하게 회차를 투입할 수 있도록 이미지 파일 직접 업로드, 네이버 웹툰 URL 기반 자동 크롤링, 웹툰 API 연동 선택의 세 가지 입력 경로를 단일 인터페이스로 통합 제공한다. URL이 입력되면 헤드리스 브라우저가 해당 회차 이미지를 자동으로 크롤링하며, 복합키 기반 캐시를 두어 동일 회차의 재크롤링을 방지한다. 이렇게 수집된 회차 이미지는 해상도 정규화 전처리를 거쳐 후속 분석 모듈로 전달된다.

4.2 컷 분할

회차 이미지를 통째로 영상화하면 의미가 중첩되고 영상 길이를 제어할 수 없으며, 특히 영상에서는 컷별 객체의 참조 이미지를 이후 단계에 걸어두는 것이 외형 일관성 유지에 필수적이다. 따라서 본 단계는 회차 이미지를 의미 단위인 패널로 자동 분리하여 후속 단계가 다룰 수 있는 단위로 정제하는 역할을 담당한다. 효율적인 처리를 위해 본 시스템은 외부 API에 의존하지 않는 경량 수치 알고리즘을 채택한다. 구체적으로 각 행의 픽셀 표준편차(row standard deviation)로 단색 여백 구간을 검출하고 그 중앙을 분할 좌표로 삼는 무손실 패널 분리를 수행하며, 외부 API 호출이 없는 순수 수치 연산이므로 비용·속도 측면에서 우위를 갖는다. 또한 비정형 패널 레이아웃에 대비해 딥러닝 기반 패널 추출을 보조적으로 결합하고, 검출된 컷은 해시 기반 디스크 캐시에 저장하여 중복 처리를 방지한다.

Table 2: Per-cut storyboard script excerpts produced by the LLM analysis stage for the five selected cuts (Samples 1–5 in Table 3). Texts are condensed from the system's English output; full structured fields (situation, mood, character, dialogue) are kept verbatim.

Cut	Scene & mood summary	Dialogue / inner voice / SFX
1	A young woman in a grey hoodie stands outdoors, looking at something with a serious expression and suspecting stalker activity. **Mood**: suspicion, anxiety, alertness.	**Inner**: "Is this also a stalker…?"
2	The same woman holds a small yellow-and-blue object (possibly a lip balm) with a slightly troubled expression. **Mood**: serious, quietly bewildered.	**SFX**: "Swish"
3	A light-skinned hand holds a "super-strong self-defense spray" with a teddy-bear sticker attached. **Mood**: curiosity, mild puzzlement about the sticker.	**Inner**: "Today is pepper spray…did someone put this sticker on?"
4	A woman stands indoors in front of rows of storage lockers, pressing a small yellow pill-like object. **Mood**: calm, focused, absorbed.	**SFX**: "Press"
5	A character with short black hair stands in front of mailboxes at a building entrance, arms crossed, lost in thought. **Mood**: reflective, frustrated, slightly irritated.	**Inner**: "How long is this person going to keep doing this anyway?"

성을 유지하며, 원본 웹툰 컷의 배경과 구도를 보존한다. 또한 실사(photorealistic) 스타일을 지정하여 2D 애니메이션·일러스트를 배제하고, 이미지 내 텍스트 렌더링은 명시적으로 차단하는 가이드라인을 일관 적용한다.

자동 결과가 기대에 미치지 못할 경우 사용자는 **프롬프트 수동 수정**으로 재생성을 요청하거나, **객체 대체(object replacement)** 기능으로 인페인팅 기반 부분 편집을 수행할 수 있다.

4.6 실사화 이미지 기반 영상 생성

텍스트만으로 영상을 직접 합성하는 방식은 고품질 앵커 프레임을 사전에 확보하기 어려워 장면 구성이 불안정해진다. 본 단계는 이러한 한계를 우회하기 위해 직전 단계에서 확정된 실사 이미지를 시작 프레임(starting frame)으로 활용하여 이미지-영상(image-to-video) 변환을 수행하며, 다음 가이드라인을 적용해 원본 서사와의 정렬과 파이프라인의 연속성을 함께 확보한다. 컷당 5초 분량의 영상 클립을 생성하고, 회차 메타 서사와 컷별 스크립트로부터 산출된 모션 프롬프트로 동작과 장면 전환을 원본 서사에 맞게 정렬한다. 또한 컷별 스크립트의 대사·나레이션·효과음을 JSON 음성 트랙으로 함께 입력하여 Veo 3.1 Lite가 동영상과 동기화된 오디오까지 한 번에 출력하도록 구성하며, 안전 필터에 의해 생성이 차단되는 경우에는 자동으로 무해한 대체 프롬프트로 재시도하여 파이프라인 중단을 방지한다.

본 모듈은 고품질 앵커 프레임을 먼저 확정한 뒤 이를 조건 신호로 시간적 동작을 합성하는 분리 방식을 채택함으로써 장면 구성의 정확성을 높인다.

4.7 영상 편집 및 자동 업로드

사용자가 영상 편집을 위해 별도의 외부 편집기로 빠져나가야 한다면 엔드투엔드 자동화의 의미가 약화된다. 본 단계는 이를 방지하기 위해 컷별로 생성된 영상 클립의 결합·트랜지션(transition)·배경음악(BGM: background music) 삽입까지를 모두 동일 플랫폼 내부에서 처리하도록 FFmpeg를 모듈로 내장하였고, 출력 영상은 곧바로 소비자 플랫폼과 자동 연동된다. 본 단계는 다음 절차로 진행된다. (1) 개별 영상 클립을 순서대로 결합하고 교차 디졸브(cross dissolve)·가속 페이드(accelerate fade) 등 트랜지션과 BGM을 FFmpeg 기반 모듈로 삽입한다. (2) 미리보기로 검토한 뒤 공유 백엔드 업로드 엔드포인트로 전송하여 회차 번호를 기준으로 소비자 플랫폼과 자동 매핑한다. (3) 콘텐츠 ID 기반 식별 체계를 통해 향후 버전 관리·시리즈 연결·통계 연동이 용이하도록 한다.

5. 실험

본 절에서는 제안 파이프라인의 핵심 단계가 원본 웹툰의 내용을 얼마나 충실히 반영하는지, 그리고 비전문가 사용자가 이를 어떻게 경험하는지를 평가하기 위한 실험 설계를 기술한다. 평가는 단계별 입력·출력 간 내용 일치도를 임베딩·VLM 기반 객관 지표로 측정하는 정량 평가와, 실제 사용자를 대상으로 한 사용성·효용성 설문을 병행한다.

5.1 평가 대상과 표본

정량 평가는 파이프라인에서 가장 핵심이 되는 세 단계, 즉 (1) 스토리보드 추출, (2) 실사화 이미지 생성, (3) 비디오 생성을 대상으로 한다. 동일한 웹툰 1회차(약 150개 컷)를 입력으로 사용하고, 사용자가 선별한 컷 5개를 1차 표본으로 삼았다. 다만 5개 표본만으로는 플랫폼 전반의 성능을 일반화하기 어려우므로, 동일 회차에서 표본을 확대하여 컷 번호 10 이하(7개)·20 이하(14개) 구간에 대해 동일 지표를 추가로 측정하였다(6.3절). 이때 확장 표본의 컷은 특정 결과가 유리한 구간을 고르는 대신 회차의 컷 순서를 따라 차례대로 취하되, 컷 분할 단계에서 빈 여백으로 분류된 구간만 제외하였다. 이를 통해 사용자가 임의 선별한 1차 5개 표본에서 우려될 수 있는 선택 편향(selection bias)을 완화하고자 하였다.

Table 3: Per-sample quantitative content-fidelity results for the five selected cuts (Sample 1–5 with their mean), together with the mean over an expanded cut set taken in reading order for generalization (cut index $\leq$10, n=7; $\leq$20, n=14).

Metric (comparison)	5 selected cuts						Expanded mean	
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Mean	$\leq$10	$\leq$20
Storyboard VQAScore (per-cut text $\leftrightarrow$ cut image)	0.847	0.872	0.707	0.917	0.721	0.813	0.798	0.805
Photoreal CLIP-I (webtoon cut $\leftrightarrow$ photoreal image)	0.713	0.745	0.736	0.711	0.667	0.714	0.673	0.652
Video VQAScore (generated video $\leftrightarrow$ cut description)	0.843	0.724	0.902	0.834	0.808	0.822	0.756	0.756
Video ViCLIP (generated video $\leftrightarrow$ cut description)	0.349	0.224	0.241	0.249	0.257	0.264	0.173	0.175
Composite quality score S	0.801	0.780	0.782	0.821	0.732	**0.783**	**0.742**	**0.738**

Table 4: Storyboard faithfulness (VQAScore) by narrative category. *N* denotes the total number of descriptor sentences extracted from the storyline of a single episode (*N*=42 in total).

Narrative category	*N*	Mean	Min	Max
Plot summary	3	0.708	0.679	0.723
Character description (Char-A)	7	0.723	0.565	0.851
Character description (Char-B)	5	0.732	0.632	0.899
Character description (Char-C)	6	0.621	0.449	0.695
Character relations	7	0.743	0.626	0.898
Emotion arc	7	0.745	0.695	0.869
Key events / inferences	7	0.734	0.597	0.836
Overall	**42**	**0.717**	0.449	0.899

6.1 단계별 결과

스토리보드 단계. 선별된 5개 컷의 컷별 VQAScore는 평균 **0.813**(표준편차 0.094, 최저 0.707, 최고 0.917)으로 모든 표본이 $\tau = 0.70$을 상회하였다(Table 3). 보조적으로 회차 전체 서사 분석을 42개 서술 표본 단위로 검증한 평균은 **0.717**(표준편차 0.087, 최저 0.449, 최고 0.899)이며(Table 4, Figure 5), 감정 흐름(0.745)·관계 묘사(0.743)가 높고 인물 C 묘사(0.621)는 "예측 불가능한 성격"과 같이 시각적으로 단정하기 어려운 추상 묘사에서 가장 낮았다. 분포는 τ 이상에 약 62%(26/42), 0.65 이상에 약 79%(33/42)가 위치한다.

실사화 단계와 비디오 단계. 5개 컷에 대한 평균 CLIP-I는 **0.714**(표준편차 0.027, 최저 0.667, 최고 0.745)로, DreamBooth가 보고한 CLIP-I 영역(생성-참조 약 0.800, 실제 이미지 간 약 0.885)보다는 낮으나 웹툰 일러스트와 실사 이미지 간의 큰 도메인 격차를 감안하면 의미적 보존성을 확인하는 보조 지표로서 충분히 의미 있는 수치로 해석된다(Figure 4에서 인물 외형·구도·배경의 실사 톤 일관 보존을 확인할 수 있다). 비디오 단계의 평균 VQAScore는 **0.822**로 완전 일치를 의미하는 1.0에 비교적 근접하여 생성 영상이 원본 컷의 내용을 대체로 충실히 반영함을 시사하며, 보조 지표 ViCLIP은 평균 0.264로 32 토큰 텍스트 길이 제한 탓에 절대값은 낮지만 표본 간 분산이 작아 일관된 경향을 보였다.

6.2 통합 품질 지표 적용

컷-수준 평균값($S_{\text{story}} = 0.813$, $S_{\text{img}} = 0.714$, $S_{\text{vid}} = 0.822$)을 Eq. (1)에 대입하면 통합 품질 지표는 $\bar{S} =$ **0.783**으로 산출된다. Table 3 마지막 행에서 보듯 5개 표본의 컷별 S 값은 모두 운영 임계값 $\tau = 0.70$을 상회하며, 가장 낮은 표본 5도 0.732로 게이트를 통과한다. 보조 지표인 회차-수준 서사 분석도(평균 0.717)는 약 38%의 서술이 τ 미만이며 시각적으로 단정하기 어려운 인물 묘사에 집중되어 있어, 향후 서술 범주별 차등 가중치 부여나 시각 단서가 약한 서술의 별도 평가 채널 도입을 검토할 수 있다.

6.3 확장 표본에 대한 일반화

5개 표본만으로 플랫폼 전반의 성능을 단정하기 어렵다는 점을 보완하기 위해, 동일 회차에서 표본을 확대하여 컷 번호 10 이하(7개)·20 이하(14개) 구간에 대해 동일한 정량 지표를 재측정하였다(Table 3의 마지막 두 열). 통합 품질 지표 S는 1∼10 구간에서 평균 **0.742**(표준편차 0.035), 1∼20 구간에서 **0.738**(표준편차 0.033)로, 5개 표본의 0.783보다 다소 낮으나 두 구간 모두 운영 임계값 $\tau = 0.70$을 안정적으로 상회한다. 스토리보드 VQAScore(0.80 내외)와 비디오 VQAScore(0.76 내외)는 표본을 늘려도 평균이 크게 흔들리지 않고 표준편차가 작아, 5개 표본에서 관찰된 경향이 더 큰 표본에서도 유지됨을 시사한다. CLIP-I는 표본 확대 시 0.65 내외로 다소 낮아졌는데, 이는 웹툰→실사 도메인 격차가 큰 다양한 컷이 추가로 포함된 결과로 해석된다.

6.4 사용성 평가 결과

사용자 설문(*N*=10) 결과를 Table 5에 제시한다. 시스템 사용성은 SUS 점수 **81.7**(표준편차 19.6)로 통상 우수 등급 기준선(약 80점)을 상회하여, 비전문가도 큰 어려움 없이 플랫폼을 사용할 수 있음을 보였다. 효용성 및 수용 의도(E)는 7점 만점에 평균 **6.17**로 가장 높아, 전문 지식 없이도 쓸 만한 결과물을 제작할 수 있다는 점이 가장 높게 평가되었으며, 제작 부담(W, 역코딩 평균 5.97)이 낮게 나타나 자동화가 제작 피로를 실제로 경감함을 뒷받침한다. 결과물 품질 측면에서도 이미지 품질(5.90)·영상 품질(5.65)·스토리라인 반영도(5.49)가 모두 척도 중앙값(4)을 크게 상회하였다. 자유 응답에서는 "쉽고 간편하게 콘티부터 완성본까지"의 자연스러운 제작 흐름과 컷 분할·실사화 품질에 대한 긍정 평가가 많았던 반면, 일부 컷에서 선택 장면과 생성 영상의 불일치, 컷 선택 UI의 가독성 개선 요구가 제기되었다. 다만 본 설문은 AI·가상융합 전공 대학원생 10명의 동질적 소표본을 대상으로 하므로, 일반 창작자 모집단으로의 일반화를 위해서는 더 크고 다양한 표본의 후속 검증이 필요하다.

Table 5: User study results (N=10). Construct scores are on a 7-point Likert scale (higher is better; reverse items recoded). SUS is the 0–100 System Usability Scale score.

Construct (items)	Mean	SD
Storyline reflection (S1–S8)	5.49	1.32
Image quality (I1–I6)	5.90	0.82
Video quality (V1–V6)	5.65	0.97
System usability (U1–U10)	5.90	1.18
Utility & acceptance (E1–E9)	6.17	1.06
Low workload (W1–W3, recoded)	5.97	0.96
Overall satisfaction (G1–G2)	5.45	1.79
SUS score (0–100)	**81.7**	19.6

Table 6: Ablation of the two consistency mechanisms at the photorealization stage (mean over the five cuts). VQAScore (realism) is the primary indicator; CLIP-I is shown for reference, as it favors minimal transformation (cf. 2.4).

Configuration	CLIP-I	VQAScore
Full (story prompt + style anchor)	0.669	**0.524**
w/o story-level prompt	0.648	0.446
w/o style anchor	0.679	0.473
w/o both	0.677	0.489

6.5 구성 요소 절제 분석

제안한 일관성 보강 장치의 개별 기여를 분리하기 위해, 동일한 5개 컷에 대해 **스토리 맥락 자동 프롬프팅**과 **스타일 앵커**를 각각 켜고 끄는 2×2 절제 실험을 수행하였다(Table 6). 실사화는 만화 양식을 사진 양식으로 의도적으로 변형하는 과제이므로, 외형 보존을 재는 CLIP-I는 변형이 약할수록 오히려 높아지는 구조적 편향을 갖는다. 실제로 스타일 앵커를 끈 조건에서 CLIP-I가 더 높게 나타나(0.679 대 0.669) 이 편향이 확인되므로, 본 분석은 결과물이 실제 사진처럼 보이는 정도("a realistic photograph"와의 정합)를 직접 측정하는 VQAScore를 주 비교 지표로 삼고 CLIP-I는 참고로만 제시한다. VQAScore 기준으로 두 장치를 모두 적용한 전체 구성이 평균 **0.524**로 가장 높았으며, 스토리 프롬프팅을 제거하면 0.446, 스타일 앵커를 제거하면 0.473으로 각각 하락하여 두 장치가 모두 실사화 품질에 기여함을 확인하였다.

7. 결론 및 향후 연구

본 논문은 종래에 모델링·렌더링·라이팅·텍스처링 등 노동집약적 전문 인력의 협업이 요구되던 영상 제작을, 비전문가가 단일 플랫폼 안에서 컷 선별과 일부 수정만으로 수행할 수 있도록 통합한다는 점에서 의의가 있다. 회차 스토리 맥락을 후속 단계에 일관 주입하는 프롬프팅, 레퍼런스 이미지·스타일 앵커·외형 고정 프롬프트의 결합, 통합 품질 지표 기반 임계값 게이트의 세 장치를 통합한 결과, 컷-수준 스토리보드와 비디오 단계 VQAScore가 평균 0.8을 상회하였고 통합 품질 지표 또한 5개 표본 전반에서 운영 임계값 $\tau = 0.70$을 안정적으로 상회하여 자동 재생성 게이트가 실제 운영 환경에서 동작 가능함을 확인하였다. 아울러 컷 표본을 20개까지 확대해도 통합 품질 지표가 임계값을 안정적으로 유지하였으며, 사용자 10명을 대상으로 한 설문에서 SUS 81.7의 우수한 사용성과 높은 효용성(7점 만점 6.17)을 확인하여 비전문가의 활용 가능성을 뒷받침하였다. 또한 구성 요소 절제 분석을 통해 스토리 맥락 프롬프팅과 스타일 앵커가 각각 실사화 품질(VQAScore)에 기여함을 확인하여, 제안한 일관성 보강 장치의 효과를 정량적으로 검증하였다.

다만 본 시스템은 두 가지 명확한 한계를 함께 보인다. 첫째, 본 시스템은 프로페셔널 제작이 아닌 개인 창작자의 활용을 가정하는데, 본문에서 정리한 외부 AI API 단가가 절대적으로 낮지 않아 회차 단위 반복 사용 시 누적 비용이 개인 부담의 핵심 허들로 작용한다. 둘째, 다중 컷에 걸친 외형·서사 일관성(consistency) 문제는 현재 AI 영상 생성 연구에서 가장 활발히 다루어지는 미해결 과제로, 본 연구는 세 가지 일관성 보강 장치로 이를 완화하였으나 미세한 외형 흔들림과 안전 필터에 의한 폴백 시 미세한 서사 이탈, 시각 단서가 약한 추상 인물 묘사(특히 인물 C 범주 평균 0.621)에서의 평가 저하가 잔여 한계로 남는다. 후자는 모델 자체의 발전과 함께 점진적으로 해소될 것으로 전망되며, 본 연구가 함께 제시한 임계값 게이트와 통합 품질 지표가 그 진전을 정량적으로 추적할 수 있는 기반을 제공한다.

향후에는 LoRA(low-rank adaptation) 파인튜닝과 텍스추얼 인버전(textual inversion) 등 캐릭터 임베딩 고정 기법으로 외형 일관성을 강화하고, 시각 단서가 약한 추상 서술을 위한 별도 평가 채널과 서술 범주별 차등 가중치를 도입하며, 망가와 서양 코믹스를 포함한 다양한 만화 형식으로 적용 범위를 확장할 계획이다.

감사의 글

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 대학ICT연구센터사업(IITP-2026-RS-2023-00259099), 메타버스융합대학원(RS-2022-00156318), 문화체육관광부 및 한국콘텐츠진흥원의 문화기술 연구개발사업(RS-2023-00219237)의 지원으로 수행되었음.

References

[1] A. Deshmane and X. Barriola, "Frame by Fame: Content Creation on Short Video-Format Platforms," *Manufacturing & Service Operations Management*, 2024, https://pubsonline.informs.org/doi/10.1287/msom.2021.0332.

[2] A. Azzarelli, N. Anantrasirichai, and D. R. Bull, "Intelligent Cinematography: A Review of AI Research for Cinematographic Production," *Artificial Intelligence Review*, vol. 58, p. Article 108, 2025, https://doi.org/10.1007/s10462-024-11089-3.

[3] K. Kim, Y. Lee, and D. Hyun, "Exploratory Study on the Structural Transformation of Video Production Using Generative AI: Focusing on Nonlinear Production Workflows and Human-AI Collaboration Dynamics," *The Korean Journal of Arts Education*, vol. 23, no. 4, pp. 429–450, 2025, https://doi.org/10.34004/kjae.2025.23.4.025.

[4] U. Singer, A. Polyak, T. Hayes, X. Yin, J. An, S. Zhang, Q. Hu, H. Yang, O. Ashual, O. Gafni, D. Parikh, S. Gupta, and Y. Taigman, "Make-A-Video: Text-to-Video Generation without Text-Video Data," arXiv:2209.14792, 2022, https://arxiv.org/abs/2209.14792.

[5] S. Hong, J. Seo, H. Shin, S. Hong, and S. Kim, "DirecT2V: Large Language Models are Frame-Level Directors for Zero-Shot Text-to-Video Generation," arXiv:2305.14330, 2023, https://arxiv.org/abs/2305.14330.

[6] M. O. Atalay, M. N. Alpdemir, A. G. Özkan, O. Karaman, Y. K. Akyüz, F. T. Bülbül, E. K. Açıkgöz, and E. Arısoy, "Automated Rawstory-to-Video Generation from Nasreddin Hodja Tales via an Expert-Inspired Multi-Stage Transformation Pipeline," in *IEEE*, 2026, https://ieeexplore.ieee.org/document/11418405/.

[7] J. Xiao, C. Yang, L. Zhang, S. Cai, Y. Zhao, Y. Guo, G. Wetzstein, M. Agrawala, A. Yuille, and L. Jiang, "Captain Cinema: Towards Short Movie Generation," arXiv:2507.18634, 2025, https://arxiv.org/abs/2507.18634.

[8] J. Zheng and X. Cun, "FairyGen: Storied Cartoon Video from a Single Child-Drawn Character," in *ACM SIGGRAPH Asia*, 2025, https://dl.acm.org/doi/10.1145/3757377.3764007.

[9] Y. Zhou, A. I. Muresanu, Z. Han, K. Paster, S. Pitis, H. Chan, and J. Ba, "Large Language Models Are Human-Level Prompt Engineers," arXiv:2211.01910, 2022, https://arxiv.org/abs/2211.01910.

[10] L. Franklin, A. Boonmee, and K. Wongsuwan, "Vision-Driven Prompt Optimization for Large Language Models in Multimodal Generative Tasks," arXiv:2501.02527, 2025, https://arxiv.org/abs/2501.02527.

[11] X. Yang, Z. Ma, L. Yu, Y. Cao, B. Yin, X. Wei, Q. Zhang, and R. W. H. Lau, "Automatic Comic Generation with Stylistic Multi-page Layouts and Emotion-driven Text Balloon Generation," *ACM Trans. on Multimedia Computing, Communications, and Applications*, vol. 17, no. 2, pp. 1–19, 2021, https://dl.acm.org/doi/10.1145/3440053.

[12] G. Jing, Y. Hu, Y. Guo, Y. Yu, and W. Wang, "Content-Aware Video2Comics With Manga-Style Layout," *IEEE Transactions on Multimedia*, vol. 17, no. 12, pp. 2122–2133, 2015, http://ieeexplore.ieee.org/document/7226841/.

[13] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Learning Transferable Visual Models From Natural Language Supervision," in *Proc. International Conference on Machine Learning (ICML), PMLR 139*, 2021, pp. 8748–8763, https://arxiv.org/abs/2103.00020.

[14] 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 *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, https://arxiv.org/abs/2208.12242.

[15] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, https://arxiv.org/abs/1801.03924.

[16] Y. Wang, Y. He, Y. Li, K. Li, J. Yu, X. Ma, X. Li, G. Chen, X. Chen, Y. Wang, *et al.*, "InternVid: A Large-scale Video-Text Dataset for Multimodal Understanding and Generation," in *Proc. International Conference on Learning Representations (ICLR)*, 2024, https://arxiv.org/abs/2307.06942.

[17] Z. Lin, D. Pathak, B. Li, J. Li, X. Xia, G. Neubig, P. Zhang, and D. Ramanan, "Evaluating Text-to-Visual Generation with Image-to-Text Generation," in *Proc. European Conference on Computer Vision (ECCV)*, 2024, https://arxiv.org/abs/2404.01291.

[18] G. Comanici, E. Bieber, M. Schaekermann, *et al.*, "Gemini 2.5: Pushing the Frontier with Advanced Reasoning, Multimodality, Long Context, and Next Generation Agentic Capabilities," arXiv:2507.06261, 2025, https://arxiv.org/abs/2507.06261.

[19] A. Kamath, K.-W. Chang, R. Krishna, *et al.*, "GenEval 2: Addressing Benchmark Drift in Text-to-Image Evaluation," arXiv:2512.16853, 2025, https://arxiv.org/abs/2512.16853.

[20] D. Zheng, Z. Huang, H. Liu, *et al.*, "VBench-2.0: Advancing Video Generation Benchmark Suite for Intrinsic Faithfulness," arXiv:2503.21755, 2025, https://arxiv.org/abs/2503.21755.

〈저자소개〉

이 영 현

- 2024 성결대학교 차세대미디어제작학과 예술공학사
- 2025 ~현재 서강대학교 가상융합전문대학원 메타버스엔터테인먼트 전공 메타버스 엔터테인먼트석사 재학중
- 관심분야: 인공지능, 디지털 컨텐츠, HITL, AI-인간 협업, AI 영상, HCI 등
- https://orcid.org/0009-0008-2408-9413

박 준 형

- 2025 서강대학교 미래교육원 멀티미디어학사
- 2025 ~현재 서강대학교 가상융합전문대학원 메타버스테크놀로지 전공 공학석사 재학중
- 관심분야: 인공지능, 시뮬레이션, Unity
- https://orcid.org/0009-0003-6501-0728

윤 대 건

- 2025 순천향대학교 한국문화콘텐츠학과 인문학사
- 2025 순천향대학교 컨버전스디자인학과 컨버전스디자인학사
- 2025~현재 서강대학교 가상융합전문대학원 메타버스엔터테인먼트 전공 메타버스엔터테인먼트 석사 재학 중
- 관심분야 : 인공지능, 확장현실, HCI
- https://orcid.org/0009-0003-8074-8050

- 배 종 환
- 2016 경희대학교 건축학과 건축학사
- 2025 서강대학교 메타버스전문대학원 메타버스테크놀로지 전공 공학석사
- 2025~현재 서강대학교 가상융합전문대학원 메타버스테크놀로지 전공 공학 박사 재학 중
- 관심분야: 인공지능, 메타버스, 확장현실, 건축/건설, 디지털트윈
- https://orcid.org/0009-0007-7520-2319

박 상 훈

- 1993 서강대학교 수학과 학사
- 1995 서강대학교 컴퓨터학과 석사
- 2000 서강대학교 컴퓨터학과 박사
- 2002~2005 대구가톨릭대학교컴퓨터정보통신공학부 조교수
- 2001 University of California, Davis 방문연구원
- 2005 ~2023 동국대학교 멀티미디어학과 교수
- 2023 ~현재 서강대학교 가상융합전문대학원 교수
- 관심분야: 실시간 렌더링, 사실적 렌더링, 과학적 가시화, 고성능 컴퓨팅 등
- https://orcid.org/0000-0001-5383-7005