

Research on Generative AI-based Reconstruction of Cinematic Images Guided by 3D Spatial Information in Unreal Engine

A case study of Spider-Man: Brand New Day

Zhang Yuxin
Zhou Jiani
Dongseo University

| Why use UE to guide AI?

AI images can look good, but still miss the space and camera view we want.

People and buildings

Their positions, distance and depth

Camera placement

Camera direction, angle and framing

| Research question and scope

Can UE 3D scene construction help AI match the space in a film frame?

Study material

Four high-rise shots from the trailer

Experimental scope

Still-image generation and visual comparison

| Experimental comparison

A UE-guided

UE spatial reference + text prompt

Image generation

Comparison with the target frame

B AI-only

Text prompt

Same generation model

Comparison with the target frame

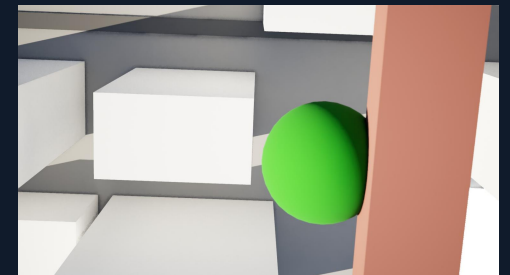
We compare spatial layout, not character identity or costume.

Building the shot in UE

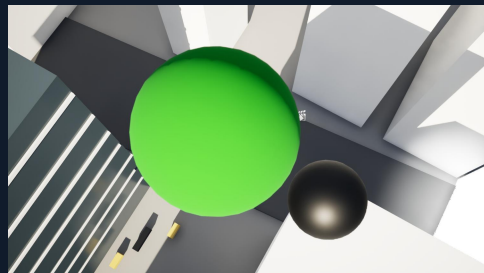
Shot 1 Target frame / UE proxy



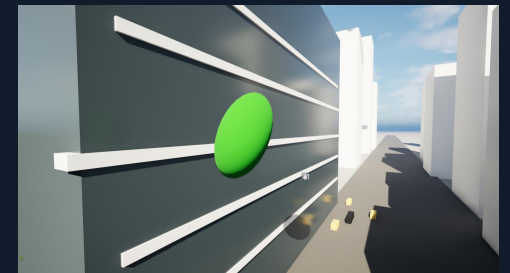
Shot 2 Target frame / UE proxy



Shot 3 Target frame / UE proxy



Shot 4 Target frame / UE proxy



Reconstruction workflow

01 Analyze the frame

Look at the subject, buildings and scene depth

02 Build the UE scene

Set simple shapes and the camera view

03 Generate images

Use the UE scene reference and the shared text prompt

04 Compare outputs

Compare each result with the original frame

| What do we compare?

Spatial relationships

Where the subject and buildings are placed

Composition

The subject's position and size in the image

Perspective

Depth and the direction of building lines

Camera angle

The direction and height of the camera view

Shot 1: a stronger result

Target frame



AI-only



UE-guided



UE guidance gives a closer layout and composition.
Perspective and camera angle match in both results.

Shot 2: a more difficult view

Target frame



AI-only



UE-guided



UE helps with the subject–building positions.
Perspective and camera angle are still hard to match.

Shots 3 and 4: two more examples

Shot 3

Target



AI-only



UE-guided



Improved:
spatial relations
and perspective

Shot 4

Target



AI-only



UE-guided



Improved:
composition
and camera angle

| What did we learn?

Main finding

UE helps with spatial layout and composition.

Still difficult

Extreme camera views can still cause errors.

Current scope

Four still-image cases, compared visually

Next step

Try more shots and check results across repeated runs.

Conclusion and future work

UE 3D scene construction can help AI follow the spatial layout we want.

Next steps

Use UE images to guide image-to-video generation

Use UE videos to guide video-to-video generation

Check spatial layout and motion over time

Thank you. Questions?