

HYBRID-3D LDM \times Film Workflow

Abstract

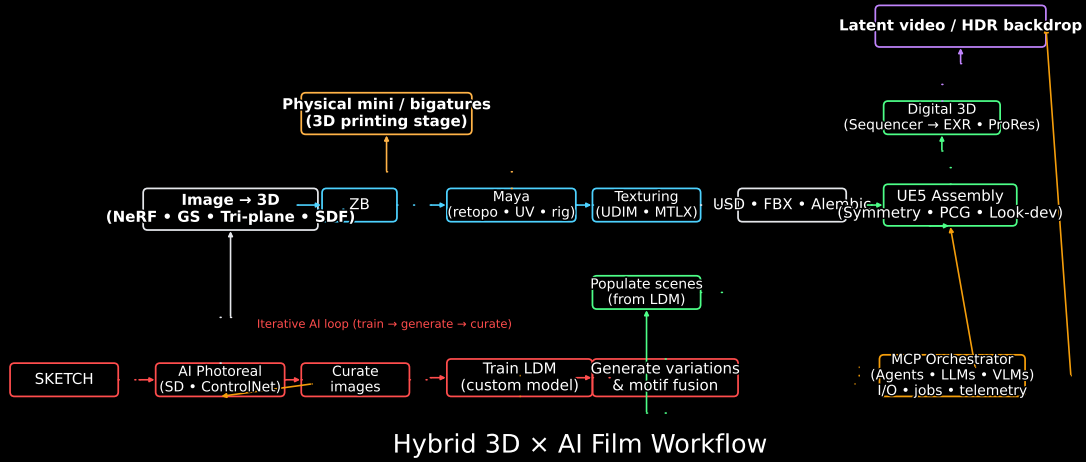
We describe a portable, film-first workflow that turns hand sketches into photoreal images with custom latent diffusion, converts them to editable 3D for ZBrush/Maya, and assembles large scenes in Unreal Engine 5. The loop is explicitly iterative: images refine the model; the model widens shape language; new variants are re-converted to 3D and populated into scenes. The final virtual cinematography stack branches into three complementary modalities: (i) digital 3D sets in UE5, (ii) physical mini/bigatures for a miniature stage, and (iii) latent video/HDR backdrops for diffusion-based scene extension. All assets use relative paths and *USD as the source-of-truth* for interchange (1 uu = 1 cm).

Pipeline (concise)

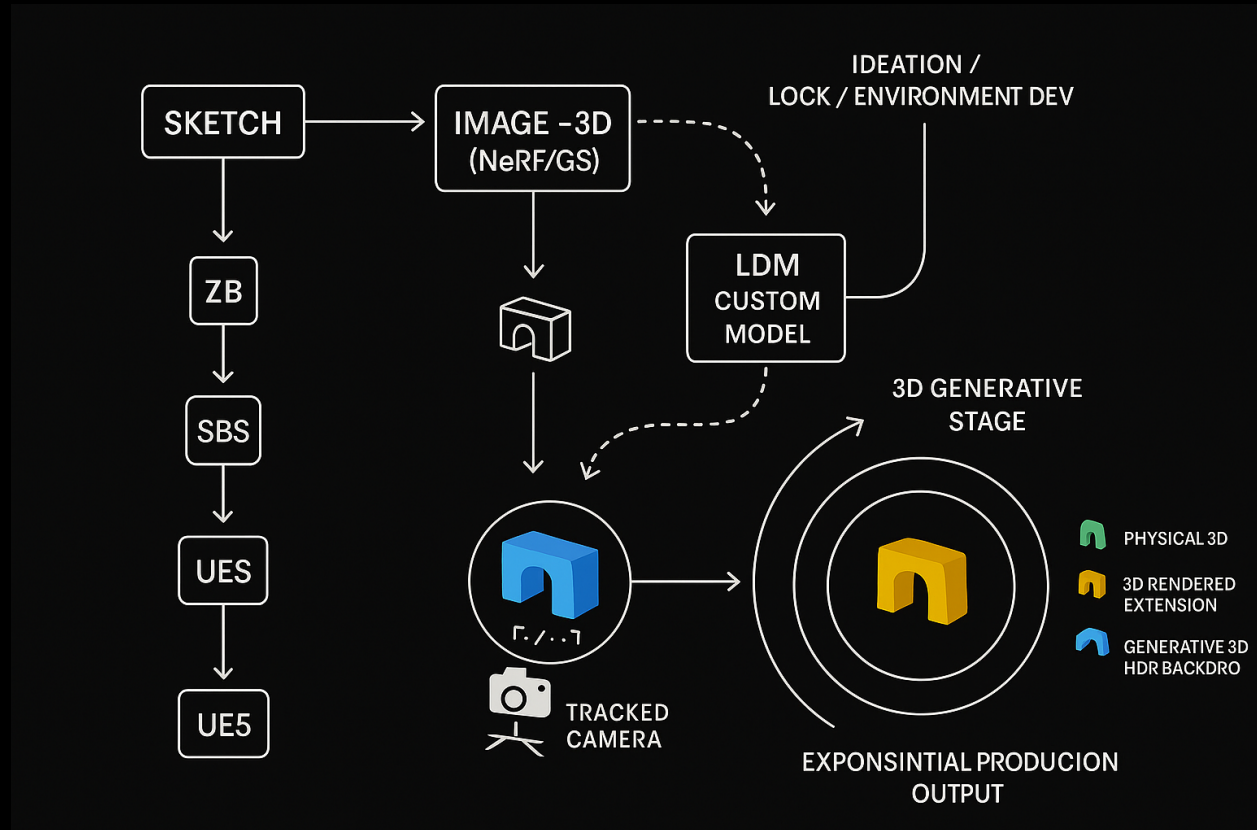
1. **Sketch** by hand (paper/tablet). Curate silhouettes and motifs.
2. **Photorealize** via img2img + ControlNet (materials, lighting, style).
3. **Curate** the best frames; clean tags and metadata.
4. **Train a custom LDM** on refined images; mix with auxiliary models for motif fusion.
5. **Generate variations** to broaden shape language; choose exemplars.
6. **Image** \rightarrow **3D**: depth normals \rightarrow mesh; or NeRF/GS/tri-plane/SDF \rightarrow marching cubes.
7. **ZBrush sculpt** (form); **Maya** (retopo/UV/rig); **Texture** (Substance/MaterialX, UDIM bakes).
8. **Interchange**: USD/FBX/Alembic; enforce film scale (1 uu = 1 cm) and metadata.
9. **UE5 assembly**: Symmetry/PCG widgets, Geometry Script ops; look-dev & lighting.
10. **Populate** scenes with variants from the trained LDM; iterate (train \rightarrow generate \rightarrow curate).
11. **Branch to three modalities**:
 - (a) Digital 3D sets in UE5 (Sequencer \rightarrow EXR/ProRes \rightarrow comp).
 - (b) Physical mini/bigatures (3D print; HDR dome/LED stage; camera match; plate shoot).
 - (c) Latent video/HDR backdrops (diffusion extension; projection/LED or comp).

Figure 1. Overview diagram of the hybrid workflow.

All units: 1 uu = 1 cm • USD is source-of-truth • All paths relative



Virtual cinematography stage — plates



Hybrid 3D × AI Workflow

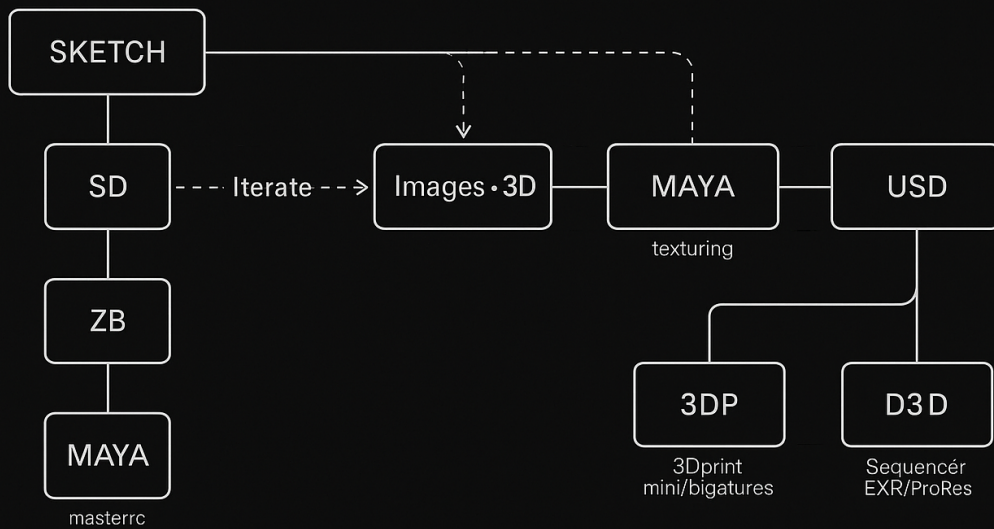


Figure 1: Plate A — Stage in perspective with circle-packing layout; largest physical set pieces (blue) at the center, smaller toward the boundary. Camera is tracked for virtual cinematography.

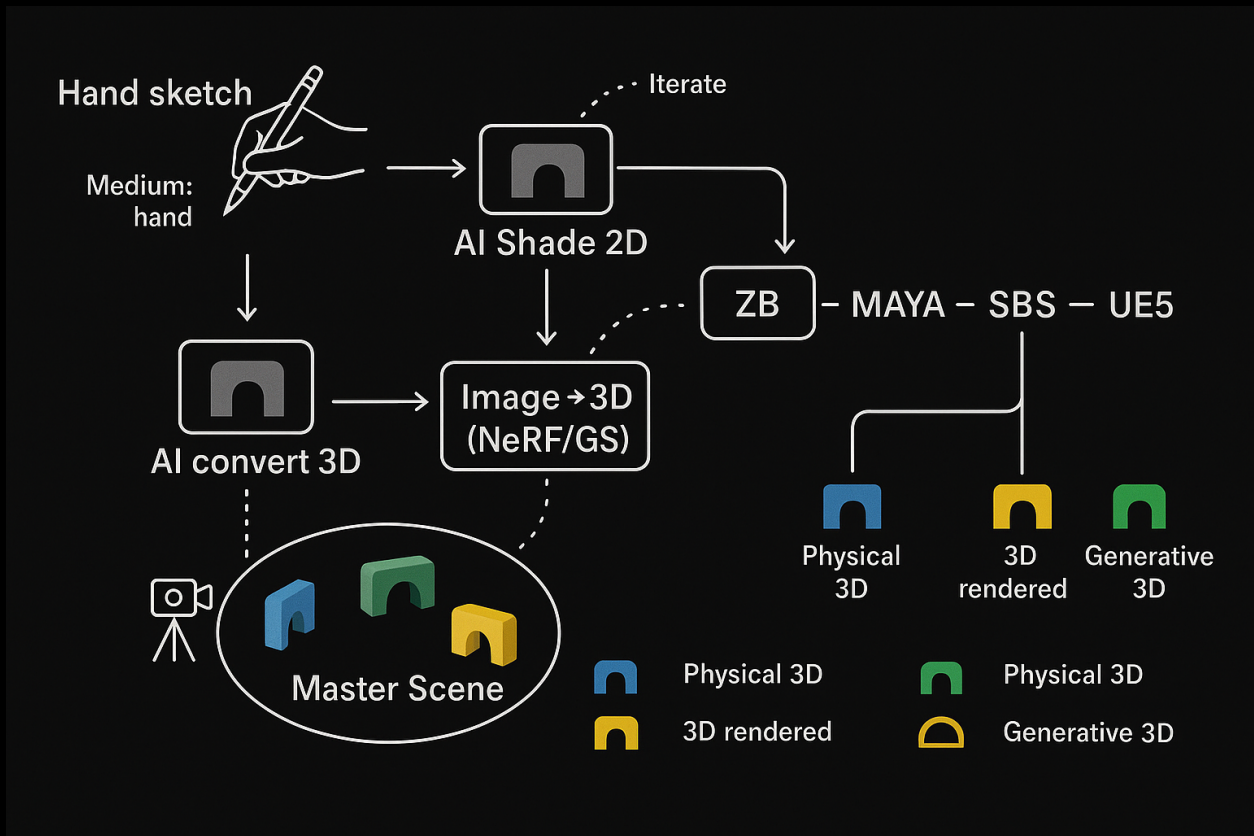


Figure 2: Plate B — Concentric resolution/medium rings around the archway hub: **Blue** = physical 3D mini/bigatures, **Yellow** = 3D rendered extensions, **Green** = generative HDR/backdrop.

VIRTUAL CINEMATOGRAPHY STAGE

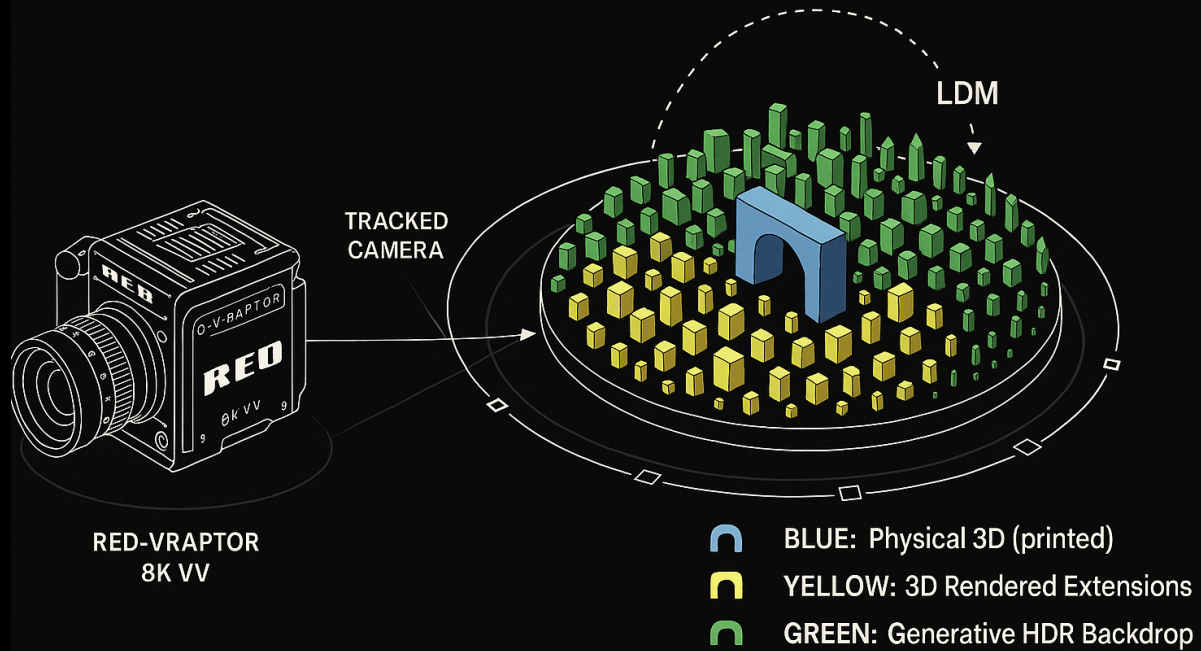


Figure 3: Plate C — Non-linear feed: images from the custom LDM loop ideation feed back up the chain to the 3D-generative stage, accelerating look/env development and scene population.