Proxy Engine

State

Box · Facing · HP · ID

Actions

Attack · Skill · Timing

Rules

```
victim.health -= damage
```

Interaction Proxy Compilation

Entity State
BBox · facing · stance · health

Action / skill windows
Class · strength · onset · duration

Event Records
Victim · hit time · health

Spatial layout

Window Marks

HUD update

Deterministic Rasterizer

Shared Canvas

(t, x, y)

Interaction Proxy π

Video Renderer

Video Latents

Patch

Future noise only

Interaction Proxy

VAE

×

Entity Reference

VAE

Identity +

Arena Reference

VAE

×

Sequence Concat

Video · Entity · Arena · Proxy

DiT Block × N

Self-Attention → Cross-Attention → FFN

Text Encoder

Decode video tokens only

VAE

Generated Video

Appearance & identity binding

Coverage C_i · Identity Slots

e_1 · e_2

Identity field

Video · first latent

Proxy · all frames

Character reference

Same slot e_i

VAE → Concat

Character tokens

Arena Reference

VAE → Concat

Arena tokens

RoPE

Video · Proxy

Shared (t, x, y)

Reference slots

F+2 · F+3