

1 - The core problem (why this matters)

A model has NO memory between generations; a text prompt is too loose to pin a face.
Consistency = feed an EXTRA signal (reference / LoRA / shared attention / camera path).

2 - The four levers (read 'best for' first)

Lever	Train?	Best for
Reference conditioning	No	Fast default - same face/place from an image
Fine-tuning (LoRA)	Yes	Recurring host / mascot across many shots
Attention sharing	No	A storyboard generated together in one go
Explicit camera control	No	A repeatable dolly / orbit / crane move

3 - 2026 model feature to reach for

Need	Reach for
Same face, a few shots, fast	Runway Gen-4 ref / Veo 3.1 ingredients / Sora 2 cameo
Recurring host, thousands of shots	Open-weights model + a trained character LoRA
Matching storyboard in one go	Kling 3.0 multi-shot / StoryDiffusion (open)
Same location / object	Runway Gen-4 locations & objects, or an object LoRA
A specific repeatable camera move	CameraCtrl (open) or a model camera-path control

4 - Production order (layer the levers)

- Lock identity once - a saved reference set (closed model) or a trained LoRA (open model).
- Fix the world - reference-condition the location and key props so the place won't drift.
- Change one variable per shot - hold lighting and wardrobe wording constant.
- Specify the camera per shot with an explicit move, not a typed phrase, where the tool allows.
- Batch shots that must match so attention sharing works; fall back to the saved identity later.

5 - Pitfalls to check before you build

- Drift with a locked reference is almost always the PROMPT contradicting the reference - let the reference describe identity, the prompt only action and scene.
- Lighting changes regenerate faces; change one variable per shot, not five.
- Reference strength is a dial: too high locks identity but ignores the prompt; too low loosens the face.