

**TL;DR** First and last frames help AI video models start and end in the right place. They do not guarantee continuity through the middle. The practical rule is simple: match the camera, framing, lighting, object count, and rigid geometry first. Then write the prompt around the desired steady state, not around every visual event you imagine happening between the anchors.

---

## 1. The short answer

Anchor frames are useful because they turn a loose image-to-video prompt into a constrained shot. The model now knows what the first frame should look like, and sometimes what the last frame should look like too.

But an anchor pair is not a contract. The model still has to invent motion, intermediate states, camera behavior, object persistence, and temporal physics. That is where continuity breaks.

In our Seedance titration workflow, anchor frames did real work:

- A new first-frame anchor fixed a missing purple potassium manganate(VII) solution in the burette.
- A first+last anchor pair fixed the endpoint colour from too-saturated magenta to a pale pink endpoint.
- The same workflow also showed the traps: mismatched anchor compositions created a visible opening morph, and prompt wording about a "deep purple drop" overrode the intended pale-pink interpolation.

So the right mental model is:

| <b>Anchor type</b> | <b>What it helps with</b>                           | <b>What it does not solve</b>                           |
|--------------------|---|---|
| First frame        | Initial composition, subject identity, object state | Middle-frame drift, end-state accuracy                  |
| Last frame         | Desired final state                                 | How the model gets there                                |
| First+last         | Start and end constraints                           | Clean interpolation, prompt conflicts, camera stability |
| Shot-per-clip      | Maximum editorial control                           | Higher assembly effort                                  |

## 2. Why first and last frames fail in production

Most anchor-frame failures are not mysterious. They come from asking the model to solve too many changes at once.

### Failure mode 1: mismatched composition

If the first frame is a close-up and the last frame is a wider shot, the model has to change both the subject state and the camera composition. That is exactly what happened in the Seedance endpoint segment. The first anchor came from an earlier Seedance close-up. The last anchor came from a fresh GPT Image frame with a wider composition.

Seedance honored both anchors, but the first second had to morph between them. The fix was not a new prompt. We trimmed the first second off the generated clip and held the final frame for the remaining duration with `ffmpeg`.

Production lesson:

- If the state change is important, keep the camera boring.
- If the camera move is important, keep the subject state simple.
- Do not ask the same 5-second render to change both composition and scientific state unless you are willing to accept morphing.

### Failure mode 2: prompt text fights the anchors

In the endpoint segment, the first and last anchors both pointed toward a pale pink solution. The first render still failed because the prompt mentioned a freshly added "deep purple" drop dispersing into the liquid.

Seedance latched onto the colour words. The middle frames became fully purple before returning to the pale endpoint anchor.

The successful rerun removed the contrasting colour narrative and described only the steady-state target:

The solution is barely visible pale pink throughout, almost colourless.

The pink stays consistent and gentle from start to end.

The solution never becomes purple, magenta, saturated, or intense.

Production lesson:

- Use anchors for start and end state.

- Use text for motion and constraints.
- Avoid naming colours or objects that should not dominate the generated middle frames.

### **Failure mode 3: rigid objects morph between anchors**

Scientific and product videos are unforgiving because viewers notice rigid geometry. A burette, tripod, bottle label, phone, watch, logo, or product pack cannot slide, bend, or resize casually.

Anchor pairs need side-by-side review before generation. Check:

- camera height
- focal length feel
- crop and subject scale
- lighting direction
- number of hands
- number of objects
- position of rigid parts
- label, dial, graduation, or text placement

If those do not match, the video model will often "solve" the difference by warping the object.

### **Failure mode 4: the model inserts cuts inside one clip**

Seedance can turn a single 5-second image-to-video request into multiple sub-shots. In one titration segment, frame inspection showed three internal compositions: a close-up dropper shot, a wide dark-purple overshoot shot, then a wide pale-pink swirling shot.

This was not requested by the prompt or anchors. It was the model's own composition choice.

Sometimes that is useful. In this case, the quick wrong-example to right-example rhythm worked pedagogically. But if you need deterministic timing, do not rely on a single generated clip. Generate one beat per clip and concatenate in post.

## **3. The Seedance continuity arc**

The useful part of the titration workflow is not that one model eventually produced a good result. The useful part is the iteration pattern.

### **v14: first-frame anchor fixed the opening state**

The previous version had a continuity issue: the burette did not clearly contain the purple titrant at the start. A new GPT Image anchor edited the opening frame so the burette was visibly filled with deep purple potassium manganate(VII).

Seedance then rendered the shot with a prompt that reinforced the visible state and explicitly avoided the bad states:

- the burette starts filled
- the liquid is saturated royal violet
- it does not fade, lighten, or empty

The evaluator panel agreed that this solved a major logic problem. That is the best use of a first-frame anchor: lock an important initial state before motion begins.

### **v15 attempt 1: anchor pair failed because the prompt leaked colour**

The next blocker was endpoint colour. The solution should finish as a barely visible pale pink, not a saturated magenta.

We used:

- first anchor: faint pink endpoint frame
- last anchor: newly generated barely-pink endpoint frame
- Seedance image-to-video with both anchors

The first render failed because the prompt narrated a deep-purple drop dispersing. That was chemically reasonable, but visually dangerous. The model turned the middle of the clip purple.

This is the rule that matters most for operators:

In image-to-video prompts, a colour word can become a visual target even if you meant it as a transient event.

### **v15 attempt 2: steady-state prompt succeeded**

The successful prompt described only the desired endpoint state. It did not mention the contrasting purple drop. The result stayed in the pale-pink zone through the motion segment and reached the intended final frame.

This is where first+last anchoring paid off. The anchors carried the endpoint target, and the prompt stopped fighting them.

## **v15 post-fix: trimming beat re-rendering**

The output still had a short opening morph because the first and last anchors were not composition-matched. Instead of re-rolling the whole segment, we trimmed the first second and held the final frame to preserve duration.

That decision matters. A production pipeline should not treat re-rendering as the only repair tool.

Use the cheapest repair that preserves the approved parts:

| <b>Problem</b>                    | <b>First repair to try</b>                  |
|-----------------------------------|---|
| Bad first 0.5 to 1.0 seconds      | Trim head, extend tail, preserve audio sync |
| Bad final colour only             | Local grade or regenerate last segment      |
| Prompt leaked an unwanted colour  | Rewrite prompt and rerender                 |
| Object geometry morphs throughout | Rebuild matched anchors                     |
| Timing of one action is wrong     | Split into separate clips and concatenate   |

## **4. Anchor-frame checklist before generation**

Before you submit a first+last-frame job, inspect the two stills as if they are animation keyframes.

### **Composition**

- Same camera angle.
- Same crop.
- Same subject scale.
- Same focal-length feel.
- Same background layout.
- Same lighting direction.

## Object continuity

- Same number of hands.
- Same number of objects.
- Same rigid-object positions.
- Same labels, logos, ticks, graduations, and UI elements where possible.
- Same liquid vessel, product pack, instrument, or prop geometry.

## State change

- Only one meaningful state should change.
- Write that state change in one sentence.
- If you cannot describe the difference in one sentence, split the shot.

## Prompt

- Describe the desired steady-state look.
- Describe the motion plainly.
- Avoid naming visual states you do not want to dominate.
- Put negative constraints around known failure states.
- Keep camera instructions compatible with both anchors.

## 5. When to use one anchor, two anchors, or multiple clips

Use one first-frame anchor when the opening state matters more than the exact ending. This works for product reveals, presenter openings, hero compositions, and shots where motion can be loose.

Use first+last anchors when the final state matters. This works for chemistry endpoints, before/after transformations, product state changes, UI transitions, and controlled camera moves where start and finish must be recognizable.

Use multiple clips when you need beat-level control. This is the right choice when:

- an object must stay rigid
- a physical process must follow a real sequence
- narration has word-level timing
- a wrong-example and right-example need specific durations

- a model keeps inserting its own cuts

For many production jobs, the robust answer is boring: generate short controlled clips, then assemble them with deterministic editing.

## 6. QA workflow

Do not approve an anchor-frame render from stills alone. First and last frames can look correct while the middle frames are wrong.

Use this review sequence:

1. Check the input anchors side by side.
2. Generate a short draft.
3. Sample the output at 4 fps or similar.
4. Count distinct camera positions.
5. Check the first frame, middle frames, and last frame separately.
6. Watch the full clip with narration.
7. Decide whether to trim, grade, overlay, rerender, or split the shot.

For high-stakes work, use a panel rather than one model judge. In the titration workflow, the useful split was:

- omni-modal review for audio-visual coherence
- reasoning-heavy review for domain-specific errors
- vision-only review for visual inventory errors that narration might mask

That combination catches different failures. It also prevents a single evaluator from turning one noisy observation into a production decision.

## 7. Practical decision tree

If the first frame is wrong:

- Fix the first anchor.
- Do not expect the prompt to repair a bad opening state.

If the last frame is wrong:

- Add or regenerate a last-frame anchor.
- Keep the prompt focused on the final state.

If the middle frames are wrong but both anchors are right:

- Remove conflicting prompt language.
- Check whether the model is interpolating through a named but unwanted state.
- If timing matters, split into separate clips.

If the camera morphs:

- Rebuild the anchors with matched composition.
- Or trim the morph if it is confined to the head or tail.

If the model inserts cuts:

- Accept it only if the cuts help the story.
- Otherwise shorten the clip or generate one beat per clip.

## 8. What this means for AI video teams

The professional version of anchor-frame prompting is not "upload two nice images and hope". It is shot design.

The operator has to decide:

- what is allowed to change
- what must remain rigid
- what the prompt is allowed to mention
- what repair tool should be used before paying for another render

That is also why anchor frames belong inside a larger production pipeline. The render is only one step. The system needs still review, frame sampling, full-video QA, revision logs, and deterministic assembly.

For the broader pipeline frame, see:

- [What a Production-Grade AI Video Pipeline Actually Needs](#)
- [Quality Control for AI-Generated Video](#)
- [How to Run an AI Video Model Bakeoff Without Turning It Into Vibes](#)

## References

- Kling AI: [Start and End Frames](#)

- ComfyUI: [WanFirstLastFrameToVideo node documentation](#)
- ComfyUI: [Wan2.2 video generation workflow](#)
- Runway: [Controlling Aleph edits with a Reference Image](#)
- ByteDance Seed: [Seedance 1.0 technical report](#)