RD — AI CONTENT SYSTEM

A Practical Framework for Creating High-Quality AI Video Content

GPT · GPT Image · Higgsfield AI · Seedance 2.5 · WAN 3.0

From idea and references to structured prompts, controlled generations and repeatable AI video workflows.

Created by RD



Introduction

AI content creation becomes much easier when you stop treating every tool as a separate platform and start using them as part of one connected creative workflow.

This guide is designed to show you a practical system for creating high-quality AI video content — from the initial idea to the final generated scene.

The focus is not on theory or on learning every feature of every AI tool. Instead, you will learn how to build a repeatable creative process that gives you more control over the final result.

Throughout this guide, we will work with four core tools:

GPT(CLAUDE) — for developing ideas, structuring scenes, writing prompts and refining creative direction.

GPT Image — for creating source images and visual references when a scene requires stronger visual control.

Seedance — for text-to-video, image-to-video and reference-based video generation.

WAN — for alternative video generation workflows and different approaches to movement and scene dynamics. **Celebrities AI video.**

The system is built around three main methods:

1. Text to Video
Creating a video directly from a written prompt.

This is the most straightforward workflow and works well when you do not need strict control over a specific character, object or visual identity.

2. Image to Video
Creating a source image first and then using that image as the visual foundation for the video.

This method gives you more control over composition, styling, characters, environment and overall visual direction before animation begins.

3. Multi-Reference Text to Video
Creating a video directly from a text prompt while using multiple visual references within the same generation.

Instead of creating a source image first, you provide the video model with several references — for example a character, outfit, location, product, object or visual style — and define how each reference should influence the final scene.

This method is especially useful when you want to create a more detailed and visually controlled video while still generating the scene directly in the video model.

Each method gives you a different level and type of creative control.

In some cases, a strong text prompt is enough.

In others, it is more effective to establish the exact visual composition through a source image first.

And for more complex scenes, multiple references can be combined directly during video generation to give the model more visual information without creating an intermediate image.

The goal of this guide is to help you understand when to use each workflow and how to move from a simple idea to a more intentional and polished result.

You will learn how to:

- structure a visual idea before generating;
- create stronger and more precise prompts;
- generate source images with GPT Image;
- work with single and multiple visual references;
- assign a clear role to each reference;
- control camera movement and subject motion;
- identify why a generation is not working;
- refine a scene without rebuilding everything from scratch;
- create a repeatable workflow for AI content production.

This is not a collection of random prompts.

It is a practical framework designed to help you understand the logic behind the process, so you can adapt it to your own ideas, visual style and content.

MODULE 1 — THE AI CONTENT WORKFLOW

Before working with prompts, references or video generation, it is important to understand the overall structure of the process.

AI video creation is not just about writing a prompt and pressing generate.

The quality of the final result depends on a sequence of creative decisions:

Idea → Visual Direction → Prompt → References → Generation → Evaluation → Refinement

The tools may change, but this logic remains the same.

Once you understand the workflow, it becomes much easier to decide what information the model needs and which generation method is most suitable for a specific scene.

1.1 Start with the Idea, Not the Tool

A common mistake is to open an AI platform first and only then decide what to create.

A stronger approach is to define the visual idea before choosing the generation method.

Start with one simple question:

What do I want to see in the final video?

At this stage, the idea does not need to be technically detailed.

For example:

> **A woman walking through a hotel lobby while the camera slowly follows her.**

or:

> **A luxury perfume bottle standing on a stone surface while water moves around it.**

The purpose of this first step is to define the core visual concept.

Once the idea is clear, you can decide how much control the scene requires.

1.2 Choose the Right Workflow

This system uses three main workflows.

Each one is suitable for a different type of scene.

Workflow 1 — Text to Video

Idea → Prompt → Video

Use this workflow when the model can create most of the visual elements freely.

It works well when:

- you are testing a concept;
- you do not need an exact character;
- you do not need a specific product or object;
- the environment can be interpreted creatively;
- speed is more important than strict visual consistency.

This is the simplest workflow and often the best place to start.

Workflow 2 — Image to Video

Idea → Image Prompt → GPT Image → Source Image → Video Prompt → Video

Use this workflow when you want more control over how the scene looks before adding movement.

The source image allows you to establish:

- the character;
- outfit;
- composition;
- location;
- lighting;
- styling;
- colour palette;
- product placement;
- overall visual direction.

Once the image is right, the video model is used primarily to introduce movement.

This workflow is especially useful when the visual composition matters more than creative randomness.

Workflow 3 — Multi-Reference Text to Video

Idea → References → Prompt → Video

This workflow allows you to generate the video directly while providing multiple visual references at the same time.

There is no intermediate source-image generation stage.

Instead, the video model receives:

- a written description of the scene;
- several visual references;
- instructions explaining how those references should be used.

For example:

Reference 1 — character
Reference 2 — outfit
Reference 3 — location
Reference 4 — product
Reference 5 — visual mood or styling

The text prompt defines the scene and movement, while the references provide additional visual information.

This workflow is useful when a scene requires more detail or visual specificity, but you still want to generate the video directly.

1.3 Think in Terms of Control

The main difference between the three workflows is not simply the number of steps.

The real difference is the level and type of control you give the model.

Text to Video

You give the model primarily **written instructions**.

The model has more freedom to interpret:

- the character;
- environment;
- styling;
- composition;
- details.

Image to Video

You define the visual starting point before animation.

The model receives a much clearer visual foundation.

You have stronger control over:

- framing;
- styling;
- character appearance;
- environment;
- lighting;
- composition.

Multi-Reference Text to Video

You provide visual guidance without locking the generation to one specific starting frame.

The model can combine information from several references while creating the scene.

This gives you more control over individual visual elements while keeping more freedom in the final composition and motion.

1.4 More Control Does Not Always Mean Better Results

A more complex workflow is not automatically better.

If the scene is simple, adding multiple references can create unnecessary conflicts.

If the scene already looks exactly right as a still image, Image to Video may be more efficient.

If you only want to test an idea, Text to Video may be enough.

The goal is not to use the most advanced method every time.

The goal is to use the simplest workflow that gives you enough control.

1.5 The Role of Each Tool

GPT

GPT is used to help structure the creative process.

It can assist with:

- developing a rough idea;
- turning an idea into a structured scene;
- writing prompts;
- refining prompts;
- generating variations;
- defining camera movement;
- identifying missing details;
- troubleshooting unsuccessful generations.

The goal is not to let GPT make every creative decision for you.

Use it to translate your creative direction into clearer instructions for the generation model.

GPT Image

GPT Image is mainly used within the Image to Video workflow.

Its role is to create the visual foundation of the scene before animation.

It can help define:

- character appearance;
- clothing;
- environment;
- objects;
- products;
- lighting;
- composition;
- overall aesthetic.

The source image should already feel close to the final visual direction before you move to video generation.

Do not rely on the video model to completely redesign a weak source image.

Seedance 2.5

Seedance is used for video generation.

Depending on the workflow, it can receive:

- only a text prompt;
- a source image and a text prompt;
- multiple references and a text prompt.

The quality of the result depends heavily on how clearly you describe:

- the main action;
- camera movement;
- environmental motion;
- timing;
- subject behaviour;
- visual priorities.

WAN 3

WAN 3 is used as an alternative video-generation model and can be especially useful for scenes involving recognisable public figures or media personalities.

Depending on the platform and the type of reference used, Seedance may restrict or reject generations involving certain recognisable individuals, such as:

* actors;

* musicians;

* creators and influencers;

* models;

* public figures;

* other well-known media personalities.

In these cases, WAN 3 may offer a different generation workflow and may support references that are restricted by other platforms.

This also makes WAN useful when you want to test the same creative concept across different models, as each model interprets identity, movement, camera direction and visual references differently.

Important: Platform policies can change over time, and the ability to generate a recognisable person does not automatically give you the right to use their likeness commercially. Always make sure your use of reference material and identifiable individuals complies with the platform's current rules and applicable rights and regulations.

1.6 The Generation Is Only One Part of the Process

The first generation should not always be treated as the final result.

A more useful workflow is:

Generate → Evaluate → Identify the Problem → Adjust → Generate Again

When reviewing a result, avoid saying only:

> It looks bad.

Try to identify the exact issue.

For example:

- the camera moves too quickly;
- the character performs too many actions;
- the face changes;
- the environment becomes unstable;
- the references are conflicting;
- the motion feels unnatural;
- the model misunderstood the main action.

Once the problem is specific, it becomes much easier to correct.

1.7 Change One Variable at a Time

When a generation does not work, avoid changing the entire prompt immediately.

If possible, adjust one main variable at a time.

For example:

If the camera movement is wrong, change the camera instruction first.

If the character movement is unnatural, simplify the action.

If the references are being mixed incorrectly, clarify the role of each reference.

If the scene contains too much motion, remove secondary movement.

This makes it easier to understand what actually improves the result.

1.8 The Core Principle

The most important principle throughout this guide is simple:

Give the model enough information to understand your intention, but not so much information that the instructions start competing with each other.

More words do not automatically create better prompts.

More references do not automatically create better videos.

More movement does not automatically make a scene more cinematic.

Clarity, hierarchy and control are more important than complexity.

In the next module, we will look at how to break a visual idea into individual elements before writing the prompt.

MODULE 2 — HOW TO THINK IN SCENES

Before writing a prompt, you need to understand what you actually want the model to create.

One of the biggest mistakes in AI content creation is starting with a vague idea and immediately turning it into a prompt.

For example:

> **A beautiful woman walking through Paris.**

Technically, this is a prompt.

Creatively, however, it leaves almost every important decision to the model.

What does the woman look like?

What is she wearing?

Which part of Paris are we seeing?

What time of day is it?

How is the camera positioned?

How is she moving?

Is the scene elegant, cinematic, documentary-style or commercial?

The more important the final visual is, the fewer of these decisions you should leave completely undefined.

This does not mean that every prompt needs to be extremely long.

It means that you should understand the scene before you describe it.

2.1 Build the Scene Before You Build the Prompt

Instead of thinking:

“What should I write in the prompt?”

start by thinking:

“What exactly should happen in the shot?”

A useful way to do this is to separate the scene into several creative elements:

Subject
Who or what is the focus of the shot?

Environment
Where does the scene take place?

Appearance & Styling
What does the subject look like?

Action
What is the main subject doing?

Composition
How is the shot framed?

Camera
How does the camera behave?

Lighting
What type of light defines the scene?

Motion
What else is moving inside the frame?

Visual Direction
What should the overall scene feel like?

You do not need to describe every category in every generation.

These categories are a framework that helps you decide what information actually matters.

2.2 SUBJECT

The subject is the main visual focus of the scene.

It can be:

- a person;
- a product;
- a vehicle;
- an animal;
- an object;
- an architectural element.

Avoid descriptions that are too generic when the appearance of the subject matters.

Instead of:

> **a woman**

you may define:

> **a woman in her late twenties with long dark hair**

or:

> **a fashion model wearing a structured black evening dress**

The goal is not necessarily to create a longer description.

The goal is to remove ambiguity where it matters.

2.3 ENVIRONMENT

The environment establishes where the scene takes place.

Compare:

> inside a hotel

with:

> inside an elegant European hotel lobby with marble floors, tall arched windows and warm ambient lighting

The second version gives the model a much clearer visual direction.

Environment can include:

- architecture;
- interior design;
- landscape;
- weather;
- time of day;
- background activity;
- materials;
- atmosphere.

For some scenes, the environment is only supporting context.

For others, it is one of the most important visual elements.

2.4 APPEARANCE & STYLING

If your scene contains a person, product or recognisable object, define the visual details that matter.

For a character, this could include:

- hairstyle;
- clothing;
- accessories;
- makeup;
- materials;
- silhouette.

For a product, it could include:

- material;
- colour;
- finish;
- packaging;
- proportions;
- surface details.

For example:

> She wears a floor-length black satin dress with a minimal silhouette and subtle silver jewellery.

This is much more useful than:

> She wears a beautiful dress.

Try to describe visible characteristics rather than subjective words whenever possible.

2.5 ACTION

Action describes what the main subject is doing.

For AI video, this is one of the most important parts of the scene.

Examples:

> She walks slowly towards the entrance.
>
> He turns his head towards the camera.
>
> She steps out of the car.
>
> The perfume bottle slowly rises through the water.
>
> The car drives past the camera.

Whenever possible, define **one clear primary action**.

A common mistake is trying to fit an entire sequence of events into one generation.

For example:

> She gets out of the car, closes the door, walks towards the hotel, looks at the camera, fixes her hair and enters the building.

This gives the model multiple actions that need to happen correctly in sequence.

In many cases, the result will be more stable if this is divided into separate shots.

2.6 COMPOSITION

Composition defines how the subject appears inside the frame.

Examples include:

- close-up;
- medium shot;
- full-body shot;
- wide shot;
- low-angle composition;
- high-angle composition;
- centred composition;
- symmetrical framing;
- subject positioned on the left or right side of the frame.

Composition becomes especially important when generating source images for Image to Video.

The initial frame determines much of the visual structure the video model will work with.

Ask yourself:

What should the viewer see first?

and:

How much of the environment needs to be visible?

2.7 CAMERA

The camera determines how the viewer experiences the scene.

There are two separate decisions:

Camera position

and

Camera movement

Camera position can include:

- eye level;
- low angle;
- high angle;
- close to the subject;
- distant;
- behind the subject;
- over the shoulder.

Camera movement can include:

- static shot;
- dolly in;
- dolly out;
- tracking shot;
- pan;
- tilt;
- orbit;
- handheld movement.

You usually do not need several camera movements at once.

One clear camera instruction often produces a cleaner result.

We will cover camera movement in much more detail later in the guide.

2.8 LIGHTING

Lighting can completely change the visual character of the same scene.

Consider the difference between:

> **a woman standing outside a hotel**

and:

> **a woman standing outside a hotel in soft blue-hour light, with warm light coming from the windows behind her**

The subject and location are almost identical.

The atmosphere is completely different.

Useful lighting directions can include:

- soft natural daylight;
- golden-hour sunlight;
- overcast diffused light;
- direct midday sunlight;
- soft studio lighting;
- dramatic side lighting;
- cinematic backlighting;

- warm practical lighting;
- blue-hour ambient light.

Avoid adding complicated lighting terminology simply to make a prompt sound professional.

Use details that actually contribute to the visual result.

2.9 MOTION

In video generation, movement does not only come from the main subject.

There can also be secondary motion inside the scene.

Examples:

- hair moving naturally in the wind;
- fabric moving as the character walks;
- curtains moving slightly;
- leaves moving in the background;
- reflections changing across a car;
- smoke drifting through the frame;
- water moving around a product;
- background pedestrians walking.

Secondary motion can make a scene feel more alive.

However, too many simultaneous movements can also make the generation unstable.

Think in terms of hierarchy:

Primary motion — the main action.

Secondary motion — supporting movement.

The primary motion should always remain clear.

2.10 VISUAL DIRECTION

Visual direction defines the overall creative language of the scene.

This may include terms such as:

- luxury campaign;
- fashion editorial;
- cinematic;
- documentary;
- minimal;
- surreal;
- intimate;
- polished commercial;
- raw handheld;

- high-fashion;
- natural lifestyle.

Visual direction should support the scene rather than replace its description.

For example:

> cinematic luxury fashion campaign

can be useful as an overall direction.

But it should not be expected to define the character, location, action, lighting and camera by itself.

2.11 THE SCENE FRAMEWORK

Before creating a prompt, you can use the following structure:

SUBJECT
Who or what is in the scene?

ENVIRONMENT
Where is it happening?

APPEARANCE / STYLING
What should the subject look like?

ACTION
What is the main action?

COMPOSITION
How is the scene framed?

CAMERA
What is the camera doing?

LIGHTING
How is the scene lit?

MOTION
What secondary movement is happening?

VISUAL DIRECTION
What should the scene feel like?

You will not always need all nine elements.

The purpose of the framework is to help you identify which details are important for the specific scene you are creating.

2.12 FROM A SIMPLE IDEA TO A DEFINED SCENE

Let us take a very simple idea:

> A woman gets out of a car in front of a hotel.

Now we can define the scene.

Subject
A woman in her late twenties with long dark hair.

Environment
The entrance of an elegant European hotel.

Styling
A long black satin evening dress, minimal jewellery.

Action
She slowly steps out of a black vintage car.

Composition
Medium full-body shot.

Camera
The camera slowly tracks backwards as she steps out.

Lighting
Soft evening light with warm illumination coming from the hotel entrance.

Secondary Motion
Her dress and hair move subtly in the evening breeze.

Visual Direction
A refined cinematic luxury campaign.

We now have a clear scene.

Only after this stage do we turn the information into a generation prompt.

2.13 DO NOT DESCRIBE EVERYTHING

A detailed scene does not mean that every visible object needs to be included in the prompt.

Prioritise the details that affect the result.

A useful hierarchy is:

1. What must happen

2. What must look a specific way

3. What supports the atmosphere

Everything else can often be left to the model.

The goal is controlled direction, not micromanagement.

2.14 THE KEY PRINCIPLE

Before every generation, ask yourself:

What decisions do I want to make myself, and what decisions am I comfortable leaving to the model?

The answer will determine:

- how detailed your prompt needs to be;
- whether you need visual references;
- whether you should create a source image first;
- which of the three workflows is most appropriate.

In the next module, we will move from scene planning to the actual structure of a strong AI video prompt.

MODULE 3 — BUILDING A STRONG AI VIDEO PROMPT

A strong AI video prompt is not necessarily a long prompt.

It is a prompt that gives the model clear priorities.

The goal is to describe:

- what we see;
- what happens;
- how the camera behaves;
- what should remain visually important;
- how the scene should feel.

The clearer the hierarchy, the easier it is for the model to interpret the scene correctly.

3.1 START WITH THE VISUAL, THEN ADD THE MOTION

When writing a video prompt, separate two questions:

What does the scene look like?

and

What happens in the scene?

For example:

> **An elegant woman in a black satin dress stands beside a vintage black car outside a luxury European hotel.**

This establishes the visual scene.

Then we add movement:

> **She slowly steps away from the car while the camera tracks backwards in front of her.**

Now the model understands both the visual setup and the action.

This is usually stronger than mixing every detail into one unstructured sentence.

3.2 THE CORE VIDEO PROMPT STRUCTURE

A useful structure for video prompts is:

Subject + Environment + Main Action + Camera + Secondary Motion + Lighting + Visual Direction

This is not a rigid formula.

You can remove or expand individual elements depending on the scene.

The purpose of the structure is to make sure the most important information appears clearly.

3.3 SUBJECT

Start by defining the main subject.

For example:

> **A woman in her late twenties with long dark hair, wearing a floor-length black satin dress.**

or:

> **A transparent glass perfume bottle with a minimal silver cap.**

The description should focus on details that are visually important.

Avoid overloading the first sentence with information that does not affect the scene.

3.4 ENVIRONMENT

Next, establish where the scene takes place.

For example:

> **She stands outside an elegant European hotel with a stone façade and warm light coming from the entrance.**

or:

> **The bottle stands on a dark natural stone surface surrounded by shallow water.**

The environment gives the subject context.

It can also establish scale, depth and atmosphere.

3.5 MAIN ACTION

For video generation, the main action should be one of the clearest parts of the prompt.

Examples:

> She slowly walks towards the hotel entrance.
>
> She turns her head slightly towards the camera.
>
> The car moves slowly past the camera.
>
> The perfume bottle remains still while water flows gently around its base.

Whenever possible, keep the primary action simple.

A model usually handles one controlled action better than several consecutive actions.

3.6 CAMERA MOVEMENT

After defining the subject action, describe what the camera does.

Examples:

> The camera slowly tracks backwards in front of her.

The camera remains completely static.

The camera performs a slow dolly in towards the product.

The camera follows the car from a low side angle.

The camera instruction should support the main action.

It should not compete with it.

For example, if the character is already performing a complex movement, a simple camera movement may produce a more stable result.

3.7 SECONDARY MOTION

Secondary motion adds realism and depth.

Examples:

Her hair and dress move subtly in the breeze.

Curtains move softly in the background.

Reflections travel across the surface of the car.

Water creates small natural ripples around the bottle.

Secondary motion should remain secondary.

Do not let it become more complex than the main action.

3.8 LIGHTING

Lighting can be added when it is important to the visual direction.

Examples:

Soft evening light with warm illumination coming from the hotel entrance.

Diffused daylight entering through large windows.

Strong directional sunlight creating hard architectural shadows.

Soft studio lighting with controlled highlights on the glass surface.

You do not need to describe lighting in extreme technical detail unless the scene depends on it.

3.9 VISUAL DIRECTION

Finish by defining the overall visual language when necessary.

Examples:

> Refined cinematic luxury campaign.
>
> Minimal high-fashion editorial.
>
> Natural lifestyle commercial.
>
> Dark atmospheric beauty film.
>
> Clean premium product campaign.

These descriptions help establish the overall tone.

However, visual direction should support the prompt rather than replace concrete visual instructions.

3.10 A COMPLETE EXAMPLE

Let us take the scene from the previous module.

Simple Idea

> **A woman gets out of a car in front of a hotel.**

Defined Scene

Subject
A woman in her late twenties with long dark hair, wearing a long black satin evening dress.

Environment
An elegant European hotel entrance with a stone façade.

Action
She slowly steps away from a vintage black car.

Camera
The camera tracks backwards in front of her.

Secondary Motion
Her hair and dress move subtly in the evening breeze.

Lighting
Soft blue-hour ambient light with warm light coming from the hotel entrance.

Visual Direction
Refined cinematic luxury campaign.

Video Prompt

> **A woman in her late twenties with long dark hair, wearing a floor-length black satin evening dress, steps slowly away from a vintage black car outside an elegant European hotel with a stone façade. She moves naturally and confidently towards the entrance while the camera tracks backwards in front of her at a slow, controlled pace. Her hair and dress move subtly in the evening breeze. Soft blue-hour ambient light contrasts with the warm light coming from the hotel entrance. Refined cinematic luxury campaign, elegant and realistic movement.**

The prompt contains the same information as the scene breakdown.

The difference is that it is now written as one clear instruction for the video model.

3.11 ORDER MATTERS

The order of information can influence how easy the prompt is to understand.

A useful priority is:

1. Main subject

2. Main action

3. Camera

4. Environment

5. Supporting motion

6. Lighting

7. Visual direction

The most important information should not be hidden inside a long paragraph of styling details.

If a specific action is essential, make it clear early in the prompt.

3.12 ONE SHOT, ONE MAIN IDEA

Think of each generation as one shot rather than an entire film.

A prompt such as:

> She gets out of the car, walks through the entrance, enters the lobby, approaches the reception desk and turns towards the camera.

contains several separate shots or actions.

Instead, divide the sequence.

Shot 1

She steps out of the car.

Shot 2

She walks towards the hotel entrance.

Shot 3

She moves through the lobby.

Shot 4

She stops and looks towards the camera.

This approach gives you more control and makes it easier to build a sequence later.

3.13 AVOID CONFLICTING INSTRUCTIONS

Some prompts fail because different instructions compete with each other.

For example:

> The camera remains static while slowly orbiting around the subject.

These instructions contradict each other.

Another example:

> She walks quickly in slow motion.

The intended result is unclear.

Whenever possible, check that:

- camera instructions do not conflict;
- motion speed is consistent;
- the subject is not asked to perform incompatible actions;
- the visual style does not contradict the lighting or environment.

Clarity is more important than creative-sounding language.

3.14 DO NOT OVERLOAD THE PROMPT

Adding more details can sometimes improve a result.

But after a certain point, additional information creates noise.

Avoid describing:

- invisible details;
- unnecessary background objects;
- several camera movements;
- multiple actions;
- too many visual styles;
- details that are already clearly established by a reference image.

Every part of the prompt should have a reason to be there.

3.15 USE NATURAL LANGUAGE

You do not need to write prompts as disconnected keywords.

Instead of:

> woman, hotel, black dress, cinematic, dolly, luxury, evening, wind

write a clear instruction:

> An elegant woman in a long black satin dress walks slowly towards the entrance of a luxury European hotel while the camera tracks backwards in front of her.

Natural language makes the relationship between the different elements much clearer.

3.16 PROMPT LENGTH

There is no perfect prompt length.

A simple scene may need only a few sentences.

A complex scene may require more detail.

Do not measure prompt quality by word count.

Ask instead:

Does the model have enough information to understand the scene?

and:

Is every important instruction easy to identify?

If yes, the prompt is probably detailed enough.

3.17 WHEN TO USE GPT

You do not need to manually write every prompt from scratch.

GPT can be used to turn a rough creative idea into a structured generation prompt.

For example, you can provide GPT with:

- your idea;
- the subject;
- the desired action;
- camera movement;
- environment;
- references;
- visual direction.

GPT can then organise this information into a cleaner prompt.

The important part is that **you remain responsible for the creative decisions**.

GPT should help structure your direction, not replace it.

Later in this guide, you will receive ready-to-use GPT templates for each workflow.

3.18 THE PROMPT CHECK

Before generating, quickly check the following:

Subject
Is it clear who or what the scene is about?

Action
Is there one obvious main action?

Camera
Is the camera behaviour clearly defined?

Environment
Is the location clear enough?

Motion
Are secondary movements necessary and controlled?

Lighting
Does the lighting support the scene?

Visual Direction
Is the intended aesthetic clear?

Conflicts
Do any instructions contradict each other?

If the answer is clear across these points, the prompt is ready for generation.

3.19 THE KEY PRINCIPLE

A good prompt should feel less like a list of keywords and more like a clear direction given to a production team.

You are telling the model:

what the shot is, what needs to happen and how you want the viewer to experience it.

Once you start thinking this way, prompt writing becomes much easier.

The next section moves from understanding prompts to using the first complete generation workflow.

PRACTICE 01 — TEXT TO VIDEO IN HIGGSFIELD AI

This is where the practical part of the system begins.

From this point forward, we are going to work inside **Higgsfield AI**.

This is the platform I personally use for most of my video generations because it allows me to access different video models and workflows from one place.

For this exercise, we will use:

Higgsfield AI → Video → Seedance 2.5

The workflow is:

Idea → Seedance Master Prompt Guide → GPT / Claude → Final Prompt → Higgsfield AI → Seedance 2.5 → Review → Refinement

---

BEFORE YOU START — DOWNLOAD THE SEEDANCE MASTER PROMPT GUIDE

Before generating anything, download the:

SEEDANCE MASTER PROMPT GUIDE - https://drive.google.com/file/d/1xBrPSDTgX1bGvCgaBds12BgkTQcloT3Y/view?usp=drive_link

This file is one of the most important resources included in this product.

I recommend using it every time you create a new Seedance prompt.

Instead of trying to remember every technical rule yourself, upload the guide into **GPT or Claude** and ask the model to use it as the foundation for your prompts.

Think of it as giving GPT or Claude a specialised instruction manual for Seedance.

HOW TO SET IT UP

Open a new conversation in GPT or Claude.

Upload the Seedance Master Prompt Guide.

Then send:

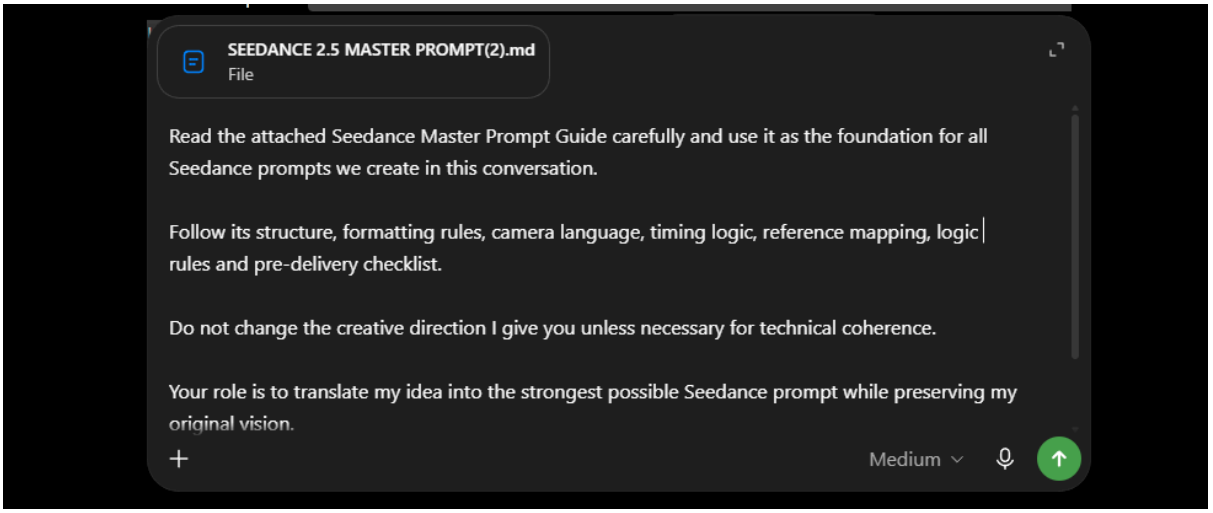
> Read the attached Seedance Master Prompt Guide carefully.
>
> From now on, use it as the foundation for every Seedance prompt we create in this conversation.
>
> Follow its structure, timing logic, shot construction, camera language, reference mapping, continuity rules, negative prompting and final checklist.
>
> Preserve my creative direction. Do not replace my idea with your own.
>
> Your role is to translate my concept into a precise, cinematic and AI-readable Seedance prompt.
>
> I will send you the first concept next.

You can then continue working inside the same conversation.

This means you do not have to explain the entire Seedance prompt structure every time.

You describe the idea.

GPT or Claude applies the framework.



WHY I USE THIS METHOD

There is a major difference between asking:

> Write me a video prompt.

and giving an AI a complete framework for how that prompt should be constructed.

The Master Prompt Guide helps the model think about:

- duration;
- pacing;
- number of shots;
- subject;
- wardrobe;
- environment;
- props;
- mood;
- camera framing;
- lens choice;
- camera movement;
- sound;
- continuity;
- physical logic;
- negative prompting;
- common generation failures.

Most importantly, it helps prevent one of the biggest mistakes in AI video:

trying to make too much happen in too little time.

STEP 1 — OPEN HIGGSFIELD AI

Go to **Higgsfield AI** and open the Video generation workspace.

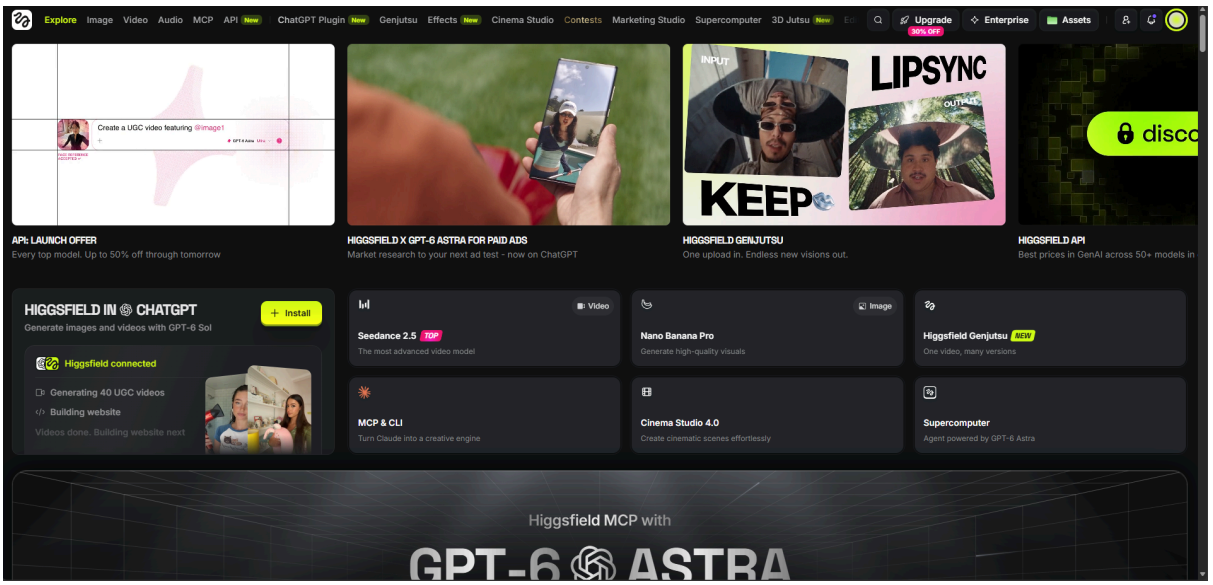
For this workflow, select:

Seedance 2.5

Use the screenshots in this guide to follow the same setup I use.

For this first exercise, we are creating a video directly from text.

You do not need to generate a source image first.



STEP 2 — START WITH THE IDEA

Do not start by writing the final prompt.

Describe the concept naturally.

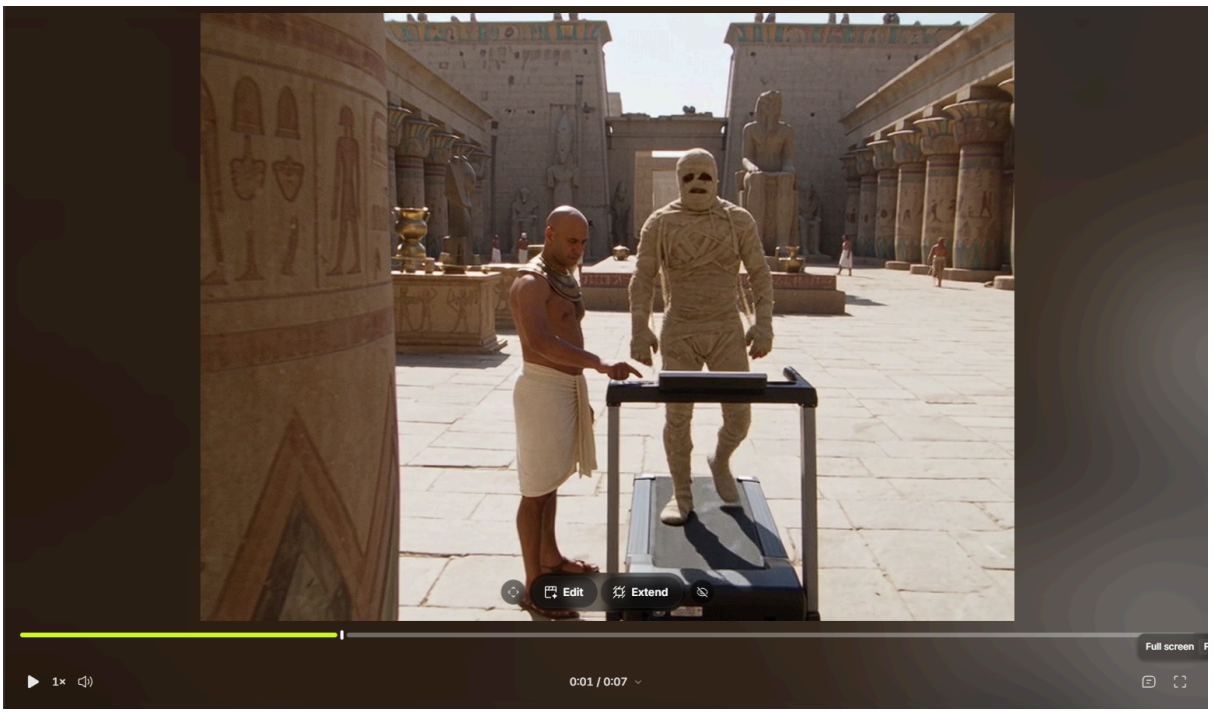
For example:

> I want accidental-looking footage from ancient Egypt where a mummy is exercising on a completely modern treadmill. An Egyptian priest becomes curious and keeps increasing the treadmill speed until the mummy is desperately trying to keep up.

This is the creative idea.

We will let GPT or Claude convert it into a much more detailed generation prompt.

REAL EXAMPLE — ONE OF MY VIRAL VIDEO PROMPTS



One of my strongest-performing videos was based around an absurd Ancient Egypt concept.

The idea was simple:

A fully wrapped mummy is exercising on a modern treadmill inside an ancient Egyptian temple. A priest becomes curious about the controls and gradually increases the speed until the mummy is desperately running.

The humour comes from treating the situation completely seriously.

The characters are not performing a joke.

The scene should feel like someone accidentally captured something impossible on an old consumer camera.

Below is one of the actual prompts I used to create a shot from this video.

You can study it, adapt it or use its structure as a starting point for your own ideas.

EXAMPLE PROMPT — ANCIENT EGYPT TREADMILL

FORMAT

Duration: exactly 7 seconds.

Aspect ratio: landscape 4:3.

Resolution appearance: clear 1440×1080 consumer-video footage.

Structure: ONE continuous uninterrupted take.

Capture style: accidental footage recorded by a random eyewitness.

Camera character: old consumer Sony Handycam / early-2000s MiniDV.

Setting: open courtyard of a wealthy ancient Egyptian temple complex, hot sunny day.

Tone: absurd, escalating, deadpan and physically believable.

NO cuts.

NO transitions.

NO montage.

NO slow motion.

NO cinematic camera movement.

CORE CONCEPT

A random eyewitness is casually filming activity inside a magnificent ancient Egyptian temple courtyard.

Among the priests, columns, statues and ordinary temple activity stands ONE completely modern motorised treadmill.

A fully wrapped ancient Egyptian mummy is walking seriously on the moving treadmill as though exercising.

Beside the treadmill stands an Egyptian priest.

The mummy initially walks comfortably at a slow pace.

The priest becomes curious about the modern control panel and presses the SPEED-UP button several times.

The treadmill accelerates progressively.

The mummy changes naturally from:

slow walking → fast walking → awkward jogging → desperate running

The mummy clearly does not want to run this fast.

It eventually grabs both treadmill handrails and desperately tries to keep up.

One loose strip of linen begins streaming backwards from the movement.

The eyewitness notices the escalating chaos and zooms tighter.

End while the mummy is still desperately running.

The mummy never falls.

CAMERA

The operator stands several metres away, partially concealed behind one of the temple columns.

The footage should feel completely accidental.

The camera operator did not arrange the scene.

Camera behaviour includes:

- natural handheld tremor;
- small body sway;
- imperfect framing;
- slightly crooked horizon;
- late reactions;
- manual panning;
- consumer optical zoom;
- brief autofocus hunting.

The camera should react to the event rather than anticipate it.

Do not make the footage feel professionally choreographed.

VISUAL CHARACTER

Authentic late-1990s / early-2000s consumer camcorder look.

4:3 frame.

MiniDV / small CCD sensor feeling.

Natural handheld movement.

Subtle micro-shake.

Limited dynamic range.

Slight highlight clipping in bright sunlight.

Natural motion blur.

Mild chromatic fringing.

Restrained sensor noise.

Subtle early-digital compression.

NO modern HDR.

NO cinematic stabilisation.

The footage should remain relatively clear but unmistakably amateur.

MOTION PROGRESSION

The treadmill speed and mummy movement must follow this logical progression:

SLOW WALK

↓

FASTER WALK

↓

FAST WALK

↓

AWKWARD JOG

↓

RUN

↓

DESPERATE RUN WHILE HOLDING THE HANDRAILS

Do not begin with fast running.

Do not instantly accelerate.

The humour depends on watching the situation gradually become worse.

PHYSICAL REALISM

Correct human anatomy.

Correct running biomechanics despite the wrapping.

Feet contact the treadmill belt naturally.

Correct treadmill belt direction.

Correct body inertia.

Correct hand contact with the rails.

Correct fabric physics.

Correct shadows.

The treadmill remains firmly planted.

No foot sliding.

No levitation.

No object morphing.

No duplicated characters.

CONTINUITY

Exactly ONE mummy.

Exactly ONE treadmill.

Exactly ONE main priest controlling it.

The mummy remains on the treadmill throughout.

The priest remains beside the console.

The treadmill remains in the same location.

The treadmill speed only increases.

The mummy never falls.

The mummy never leaves the treadmill.

The mummy never looks at the camera.

The mummy remains fully wrapped.

One long outer linen strip becomes loose only near the end.

One continuous camera take.

No camera teleportation.

No cuts.

AUDIO

Synchronized diegetic sound only.

Temple courtyard ambience.

Wind.

Birds.

Distant voices.

Sandals on stone.

Quiet treadmill motor at the beginning.

The motor pitch gradually rises as the speed increases.

Footsteps become progressively faster.

Natural treadmill belt friction.

Fluttering linen.

No music.

No soundtrack.

No narration.

No laugh track.

No cartoon sound effects.

NEGATIVE PROMPT

No horror mummy.

No zombie appearance.

No gore.

No exposed body.

No mummy attacking anyone.

No mummy looking at the camera.

No instant treadmill acceleration.

No treadmill movement across the courtyard.

No floating objects.

No mummy falling.

No complete unraveling.

No duplicated characters.

No modern buildings.

No smartphones.

No gym environment.

No professional cinematography.

No gimbal.

No dolly.

No drone.

No perfectly smooth camera movement.

No montage.

No slow motion.

No CGI appearance.

No glossy AI look.

FINAL FEEL

A completely believable piece of impossible found footage.

Someone with an old consumer camcorder happens to be filming inside a functioning ancient Egyptian temple courtyard and notices a fully wrapped mummy seriously exercising on a modern treadmill.

A priest curiously increases the speed.

The mummy gradually goes from calmly walking to awkward jogging and finally desperately running while gripping both handrails.

The scene remains completely serious and physically believable.

The recording ends abruptly while the situation is still escalating.

WHY THIS PROMPT WORKS

Do not focus only on how long the prompt is.

Look at how the information is organised.

There are several important layers.

1. THE IDEA IS EXTREMELY EASY TO UNDERSTAND

Ancient Egypt.

Mummy.

Modern treadmill.

Priest accidentally increases the speed.

You can understand the concept in a few seconds.

The detailed prompt does not replace the idea.

It protects it.

---

2. THE MOTION HAS A CLEAR PROGRESSION

The mummy does not immediately start running.

The situation develops.

Walk → faster walk → jog → run → desperate run.

This creates both physical logic and comedic timing.

---

3. THE CAMERA HAS A CHARACTER

The camera is not simply described as:

> handheld camera

It has a reason for behaving imperfectly.

The person filming is an eyewitness who notices the situation as it happens.

This explains:

- why the framing is imperfect;
- why the zoom happens late;
- why autofocus briefly struggles;
- why the camera does not perfectly follow the action.

---

4. THE PROMPT DEFINES PHYSICS

For unusual scenes, this becomes extremely important.

The model is explicitly told how the treadmill works, how the mummy moves against the belt, how the fabric reacts and what must remain physically consistent.

When your concept contains something unusual, do not assume the model will automatically understand the physics you have in mind.

Explain the important relationships.

---

5. CONTINUITY IS LOCKED

The prompt explicitly defines:

- one mummy;
- one treadmill;
- one priest;
- one location;
- one continuous take.

These rules reduce unwanted duplication and visual changes.

---

6. THE NEGATIVE PROMPT PROTECTS THE CONCEPT

Without additional direction, the word “mummy” could easily push the generation towards horror.

But that is not the concept.

So the prompt explicitly removes:

- horror;
- zombies;
- gore;
- exposed anatomy;
- monster behaviour.

The goal is absurdity, not horror.

This is a good example of using negative instructions to protect the creative direction rather than simply listing random things you dislike.

STEP 3 — CREATE YOUR OWN VERSION

Now do not simply copy my Egypt concept exactly.

Use the structure behind it.

Think of an idea that contains:

A clear world

-

One unexpected element

-

A simple escalating action

For example:

> Medieval castle + modern vending machine
>
> Renaissance banquet + food delivery courier
>
> Ancient Rome + airport security scanner
>
> Victorian ballroom + robotic vacuum cleaner

The combination itself should be easy to understand.

Then let GPT or Claude use the Seedance Master Prompt Guide to develop it.

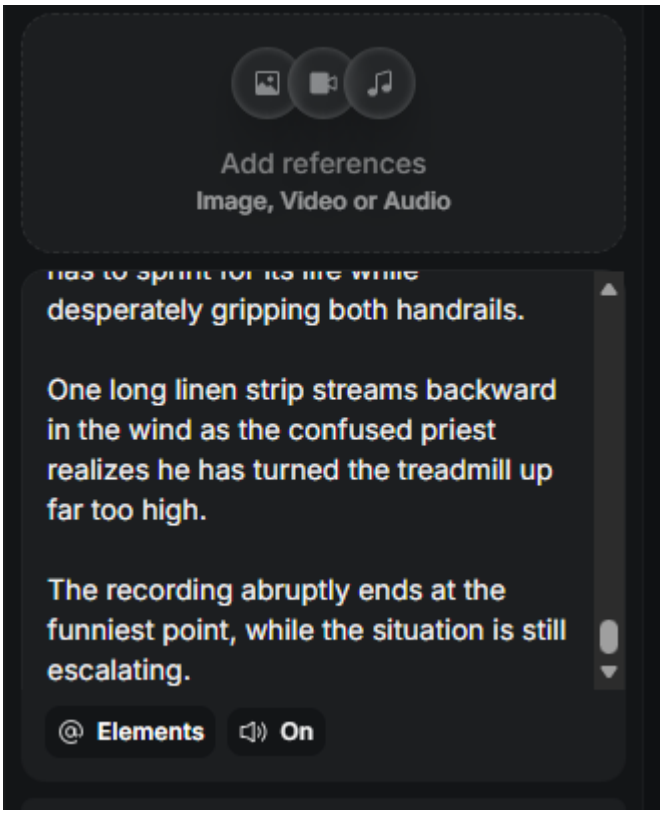
STEP 4 — GENERATE IN HIGGSFIELD

Once GPT or Claude gives you the final prompt:

Return to:

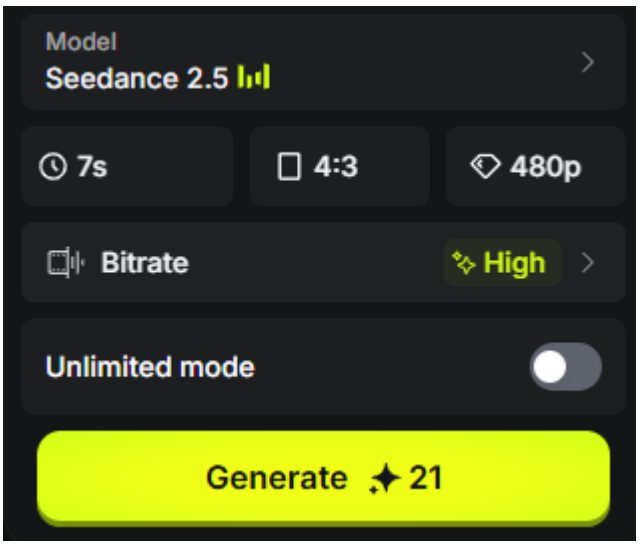
Higgsfield AI → Video → Seedance 2.5

Paste your prompt into the prompt field.



Choose the appropriate:

- duration;
- aspect ratio;
- resolution;
- audio setting.



Then generate.

STEP 5 — WATCH THE ENTIRE VIDEO

Do not judge only the first frame.

Watch the complete generation.

For an escalating scene like the Egypt example, pay particular attention to:

Beginning
Does the scene establish itself clearly?

Middle
Does the action develop naturally?

End
Does the escalation reach the intended point?

A strong first frame does not matter if the motion falls apart halfway through the clip.

STEP 6 — IDENTIFY THE MAIN FAILURE

Be specific.

For example:

> The mummy starts running too early.
>
> The priest presses the controls, but the treadmill speed does not visibly change.
>
> The camera movement looks too professional.
>
> The mummy begins to unravel too much.
>
> The treadmill moves instead of remaining fixed.
>
> The environment looks like archaeological ruins instead of a functioning ancient temple.

Each one requires a different fix.

STEP 7 — RETURN TO GPT OR CLAUDE

Stay inside the same conversation where you uploaded the Master Prompt Guide.

Describe exactly what happened.

For example:

> The overall result is strong, but the mummy begins running almost immediately.
>
> I need the escalation to be much more gradual:
>
> slow walk → fast walk → awkward jog → desperate run.
>
> Keep the visual style, environment, camera behaviour and all other parts of the prompt unchanged.
>
> Make only the minimum changes necessary.

This is one of the most important parts of the workflow.

Do not rebuild everything when only one element is broken.

YOUR PRACTICE TASK

Create one Text to Video generation inside Higgsfield AI using Seedance.

Your workflow is:

1. Create the idea

↓

2. Upload the Seedance Master Prompt Guide to GPT or Claude

↓

3. Describe the scene

↓

4. Generate the structured Seedance prompt

↓

5. Paste it into Higgsfield

↓

6. Generate the video

↓

7. Identify the main problem

↓

8. Return to GPT / Claude and refine only that part

↓

9. Generate again

↓

10. Save the final prompt and result

REMEMBER

You are not trying to memorise a huge prompt.

You are learning how to build a repeatable creative system.

Your idea provides the creative direction.

The Master Prompt Guide provides the structure.

GPT or Claude turns that direction into a detailed prompt.

Higgsfield gives you the generation environment.

Seedance creates the video.

Your feedback improves the next version.

That is the workflow.

PRACTICE 02 — IMAGE TO VIDEO

Image to Video is the workflow to use when you want stronger control over the visual starting point before animation begins.

Instead of asking the video model to invent the character, environment, composition, lighting and movement at the same time, you split the task into two stages:

Stage 1 — Create the visual

Stage 2 — Animate the visual

The workflow is:

Idea → GPT Image Prompt → Source Image → Motion Plan → Video Prompt → Higgsfield AI → Seedance 2.5 → Review → Refinement

The key advantage is control.

You decide what the scene looks like first.

Only after the visual foundation is correct do you introduce motion.

---

1. THE CORE PRINCIPLE

The most important rule in Image to Video is:

Do not use the video model to fix a weak image.

If the source image has:

- the wrong composition;
- the wrong outfit;
- the wrong object;
- the wrong lighting;
- the wrong environment;
- an awkward body position;

fix those issues before animation.

The video model should primarily solve:

- movement;
- pacing;
- camera behaviour;
- continuity;
- physical realism.

This separation makes the workflow much more stable.

---

2. THINK OF THE IMAGE AS FRAME 1

A source image should not only look good as a photograph.

It should also work as the first frame of a video.

Before generating the image, ask:

What needs to happen next?

If the character needs to walk, the body should be positioned in a way that allows believable movement.

If the camera needs to pull back, the composition should contain enough environment to reveal.

If the subject needs to turn, the framing should leave space for that motion.

If an object needs to move, its position should make physical sense.

This is one of the biggest differences between generating a beautiful image and generating a useful video source frame.

---

3. BUILD THE IMAGE PROMPT FIRST

The image prompt defines the visual state of the scene.

It should answer:

What is visible before the motion begins?

A reliable structure is:

Subject + Appearance + Wardrobe / Product + Environment + Composition + Camera Perspective + Lighting + Material Detail + Visual Direction

You do not need every category in every prompt.

Use only the details that influence the final image.

---

4. SUBJECT

Define the main visual subject clearly.

Avoid generic descriptions if identity, styling or object design matters.

Weak:

> a woman

Stronger:

> a woman in her late twenties with long dark hair and a refined editorial appearance

Weak:

> a perfume bottle

Stronger:

> a transparent rectangular glass perfume bottle with a minimal brushed-metal cap

The goal is not excessive detail.

The goal is visual specificity.

---

5. APPEARANCE AND WARDROBE

When a person is central to the frame, define the visible styling.

This can include:

- hairstyle;
- clothing;
- accessories;
- makeup;
- silhouette;
- fabric;
- colour;
- texture.

Use concrete visual language.

Weak:

> elegant clothing

Stronger:

> a floor-length black satin dress with a clean minimal silhouette and subtle silver jewellery

Specificity gives the image model clearer visual constraints.

---

6. ENVIRONMENT

Describe the setting with enough detail to establish:

- location;
- architecture;
- materials;
- atmosphere;
- time of day;

- visible background elements.

Avoid writing long decorative descriptions that do not change the image.

Focus on what the model needs to construct.

For example:

> an elegant European hotel entrance with pale stone architecture, tall windows and warm interior light

is usually more useful than:

> a breathtaking luxurious dreamlike hotel full of elegance and sophistication

Concrete detail is stronger than vague adjectives.

---

7. COMPOSITION

Composition is especially important in Image to Video.

Define how the subject appears inside the frame.

Examples:

- close-up;
- medium shot;
- medium full-body shot;
- wide shot;
- centred composition;
- symmetrical framing;
- low-angle composition;
- subject positioned slightly off-centre.

Composition should support the movement you want later.

A very tight portrait is a poor starting frame for a full-body walking sequence.

A wide frame may be better if the camera needs to move or reveal the environment.

---

8. CAMERA PERSPECTIVE

Even before video generation, define the camera viewpoint.

Useful directions include:

- eye-level camera;
- slightly low angle;
- high angle;
- side perspective;
- frontal perspective;

- three-quarter angle.

You can also add a lens feel when necessary:

- 35mm for a wider natural perspective;
- 50mm for a neutral look;
- 85mm for a more compressed portrait feel.

Do not add technical language unless it changes the visual result.

---

9. LIGHTING

Lighting should be described in visual terms.

Examples:

- soft natural daylight;
- warm evening light;
- blue-hour ambient light;
- hard direct sunlight;
- diffused studio light;
- soft backlighting;
- strong architectural shadows.

Lighting should support the mood and preserve visibility.

Avoid extreme descriptions unless you actually want extreme results.

For example, if you write:

> almost completely dark

the model may remove detail you later need for animation.

---

10. MATERIAL DETAIL

Material description is especially useful for:

- fashion;
- beauty;
- product;
- automotive;
- luxury content.

Examples:

- satin;
- brushed aluminium;
- polished glass;
- matte leather;
- natural stone;

- reflective chrome;
- textured linen.

Material detail helps the model create believable surfaces.

This also improves the way those surfaces behave once animated.

---

11. VISUAL DIRECTION

Use visual direction as the final layer.

Examples:

- luxury fashion campaign;
- editorial photography;
- understated commercial realism;
- cinematic lifestyle;
- minimal product advertising;
- documentary realism.

Do not rely on style labels alone.

“Luxury cinematic” is not a substitute for describing the actual frame.

---

12. A STRONG IMAGE PROMPT STRUCTURE

A useful prompt can follow this order:

1. Main subject

2. Appearance / wardrobe / product design

3. Environment

4. Composition

5. Camera perspective

6. Lighting

7. Materials

8. Visual direction

This order creates a clear visual hierarchy.

The most important elements appear first.

---

13. DO NOT ADD VIDEO MOTION TO THE IMAGE PROMPT

This is a common mistake.

Do not write:

> She walks slowly forward while the camera tracks backwards.

inside the image prompt.

The image model only needs to create the starting state.

Instead, describe the position before movement begins.

For example:

> She stands near the hotel entrance with one foot slightly forward, body naturally aligned as if about to begin walking.

This creates a better transition into motion later.

---

14. GENERATE THE SOURCE IMAGE

Generate the image in GPT Image.

Do not immediately choose the most visually impressive result.

Choose the one that is easiest to animate.

Evaluate:

- body posture;
- hand placement;
- facial consistency;
- object geometry;
- environment stability;
- framing;
- available movement space;
- visual balance.

The best image for video is often the one with the clearest physical logic.

---

15. CHECK THE IMAGE FOR ANIMATION READINESS

Before moving to video, ask:

Can the subject move naturally from this position?

Are all important body parts visible?

Are important objects clearly defined?

Is there enough space for the camera movement?

Is the environment stable and readable?

Does the image already contain the correct styling?

If not, fix the image first.

---

16. MOVE FROM VISUAL PROMPTING TO MOTION PROMPTING

Once the image is approved, your job changes.

You are no longer designing the frame.

Now you are directing movement.

This is where the video prompt begins.

---

17. THE VIDEO PROMPT HAS A DIFFERENT JOB

The video prompt should answer:

What should happen after Frame 1?

A strong motion prompt focuses on:

Main Action + Camera Movement + Secondary Motion + Pacing + Behaviour + Continuity + Physical Realism

This is a different prompt-engineering task from image generation.

---

18. MAIN ACTION

Define one clear primary action.

Examples:

> She begins walking slowly towards the entrance.
>
> He turns his head slightly towards the camera.
>
> The bottle remains stationary while water moves around it.
>
> The car begins moving forward at a controlled speed.

Avoid combining too many actions unless the duration genuinely allows it.

---

19. CAMERA MOVEMENT

Choose one dominant camera movement.

Examples:

- static;
- slow push-in;
- slow pull-back;
- tracking;
- pan;
- orbit;
- handheld.

The camera should support the action.

Do not add multiple camera instructions simply to make the prompt sound cinematic.

Complexity usually reduces control.

---

20. SECONDARY MOTION

Secondary motion makes the scene feel alive.

Examples:

- hair reacting to wind;
- fabric moving;
- reflections changing;
- curtains moving;

- dust drifting;
- water rippling;
- background movement.

Keep secondary motion subordinate to the main action.

The model should always understand what matters most.

---

21. PACING

Pacing is one of the most important control variables.

Words such as:

- slowly;
- gradually;
- gently;
- controlled;
- deliberate;
- restrained;
- energetic;
- sudden;

can completely change the motion.

Do not assume the model understands your intended speed.

State it explicitly when pacing matters.

---

22. BEHAVIOUR

Behaviour describes how the subject performs the action.

Compare:

> She walks forward.

with:

> She walks forward with calm, natural body movement, relaxed shoulders and a controlled pace.

The second version gives the model a clearer performance direction.

Use behavioural detail when human motion matters.

---

23. CONTINUITY

The video model should preserve the source image.

Explicitly protect the elements that matter.

For example:

> Preserve the subject's facial identity, hairstyle, wardrobe, body proportions, environment, lighting and object design from the source image.

Continuity instructions are especially useful when the scene contains:

- faces;
- products;
- vehicles;
- architecture;
- branded objects;
- distinctive clothing.

---

24. PHYSICAL REALISM

If the action contains physical interaction, describe the logic.

For example:

- feet remain in natural contact with the ground;
- hands maintain realistic contact with the object;
- fabric responds naturally to motion;
- body weight shifts realistically;
- the object remains fixed in place;
- shadows remain consistent.

Do not overuse this for simple scenes.

Use it where the model is likely to misunderstand the mechanics.

---

25. A STRONG VIDEO PROMPT STRUCTURE

A useful order is:

1. Main action

2. Subject behaviour

3. Camera movement

4. Secondary motion

5. Pacing

6. Continuity

7. Physical constraints

This structure keeps motion instructions readable.

---

26. USE THE SEEDANCE MASTER PROMPT GUIDE

Keep the Seedance Master Prompt Guide uploaded in GPT or Claude.

For Image to Video, give the model additional context:

> I am working with an existing source image in Seedance 2.5.
>
> Use the attached Seedance Master Prompt Guide as the foundation for the video prompt.
>
> The source image already defines the character, wardrobe, environment, composition, lighting and visual style.
>
> Do not redesign the scene.
>
> Focus primarily on motion, camera behaviour, pacing, continuity and physical realism.
>
> Preserve all important visual information from the source image unless I explicitly ask for a change.

Then provide the motion you want.

---

27. WORK INSIDE HIGGSFIELD AI

Open:

Higgsfield AI → Video → Seedance 2.5

Choose the Image to Video workflow.

Upload the final source image.

Paste the motion prompt.

The important point is that the image and prompt have different responsibilities.

The image tells Seedance:

This is what the scene looks like.

The prompt tells Seedance:

This is how the scene behaves.

---

28. START WITH SIMPLE MOTION

For the first generation, use:

one primary action

-

one camera movement

-

one or two secondary movements

This makes it easier to identify what Seedance understands correctly.

You can add complexity later.

---

29. REVIEW THE RESULT SYSTEMATICALLY

After generation, do not judge the clip only as:

good / bad

Break the result into categories.

SUBJECT

Did the subject remain visually consistent?

ACTION

Did the intended action happen?

CAMERA

Did the camera behave as instructed?

PACING

Was the movement too fast or too slow?

CONTINUITY

Did wardrobe, face, objects and environment remain stable?

PHYSICS

Did the movement make physical sense?

This makes refinement much easier.

---

30. IDENTIFY THE PRIMARY FAILURE

Do not try to fix everything at once.

Choose the most important issue.

For example:

> The face changes when the subject turns.
>
> The walk is too fast.
>
> The camera moves too aggressively.
>
> The dress moves unnaturally.
>
> The architecture changes.
>
> The object changes shape.
>
> The model introduces unnecessary gestures.

Each problem should produce a different prompt correction.

---

31. USE MINIMAL ITERATION

Return to the same GPT or Claude conversation.

Give precise feedback.

A strong correction looks like:

> The overall generation is correct, but the subject moves too quickly.

> Reduce the walking speed and make the movement more deliberate and controlled.
>
> Keep the camera behaviour, composition, visual identity, wardrobe, environment and lighting unchanged.
>
> Make only the minimum necessary change.

This is much stronger than:

> Rewrite the prompt and make it better.

---

32. DO NOT REWRITE WHAT ALREADY WORKS

If the camera is correct, keep it.

If the lighting is correct, keep it.

If the identity is correct, protect it.

If only one motion instruction failed, change only that part.

Prompt iteration should be surgical.

---

33. WHEN THE PROMPT IS NOT THE PROBLEM

Sometimes no amount of prompt editing will fix the result.

This usually happens when the source image conflicts with the desired motion.

For example:

A character is seated tightly in frame.

You ask the model to stand them up and walk away.

The model now has to invent:

- the lower body;
- additional environment;
- new geometry;
- new spatial relationships.

The prompt is not necessarily wrong.

The source frame is wrong for the task.

In that case, go back to GPT Image and create a better starting image.

---

34. THE IMAGE-TO-VIDEO DECISION RULE

Use Image to Video when:

- the visual identity matters;
- the composition matters;
- the character must look specific;
- the wardrobe must stay specific;
- the environment needs strong control;
- the product or object must remain recognisable;
- the first frame needs to look exactly right.

Use Text to Video when:

- visual freedom is acceptable;
- speed matters more than precision;
- the scene is concept-driven;
- you do not need a fixed starting composition.

---

35. THE MAIN LESSON

Image to Video works because it separates two different creative problems.

Image generation solves appearance.

Video generation solves behaviour.

GPT Image creates the visual foundation.

You approve the visual direction.

Higgsfield AI provides the generation environment.

Seedance 2.5 creates the movement.

Your video prompt should not redesign the image.

It should direct what happens next.

That separation is what gives you control.

PRACTICE 03 — MULTI-REFERENCE VIDEO

So far, we have worked with two types of generation:

Text to Video — the model creates the visual world and the motion.

Image to Video — you create the visual frame first, then animate it.

Now we move to a more advanced workflow:

Multi-Reference Video.

Instead of relying on one image, you provide several visual references and assign a specific purpose to each one.

This allows you to control different parts of the scene independently.

For example:

Reference 1 — character

Reference 2 — wardrobe

Reference 3 — location

Reference 4 — vehicle

Reference 5 — product

The key is not simply uploading more images.

The key is telling the model exactly:

what to take from each reference

and

what NOT to take from each reference.

PART A — MULTI-REFERENCE VIDEO IN HIGGSFIELD

For this workflow, we continue working with:

Higgsfield AI → Seedance 2.5

The workflow is:

Idea → References → Reference Mapping → Prompt → Higgsfield → Seedance 2.5 → Review → Refinement

There is no intermediate GPT Image generation unless you need to create one of the references yourself.

---

1. ONE REFERENCE = ONE RESPONSIBILITY

The most important principle of multi-reference prompting is:

Give every reference a clear job.

Do not simply upload five images and write:

> Combine these into one video.

The model now has to guess which information matters.

Instead, create explicit mapping.

For example:

> Reference 1 — use for character identity only.
>
> Reference 2 — use for wardrobe only.
>
> Reference 3 — use for location and architecture only.
>
> Reference 4 — use for vehicle design only.

This immediately reduces ambiguity.

---

2. TELL THE MODEL WHAT TO IGNORE

Every reference contains information you do not need.

For example, an outfit reference may also contain:

- another person's face;
- a different hairstyle;
- an unrelated location;
- a pose;
- lighting;
- furniture.

If you only need the clothing, say so.

Example:

> Use Reference 2 only for the clothing: garment design, colour, silhouette, fabric and accessories.
>
> Ignore the face, hairstyle, body identity, pose, environment and composition from Reference 2.

This is one of the most useful techniques in multi-reference prompting.

---

3. BUILD A REFERENCE HIERARCHY

Not every visual element has the same importance.

Decide what the model must preserve most accurately.

A useful hierarchy is:

PRIMARY REFERENCES

Elements that must remain highly consistent.

For example:

- character identity;
- hero product;
- distinctive vehicle.

SECONDARY REFERENCES

Important supporting visual information.

For example:

- clothing;
- architecture;
- interior.

STYLE REFERENCES

References used only for:

- lighting;
- colour;
- texture;
- atmosphere.

If references conflict, the model should know which one wins.

---

4. CREATE A REFERENCE MAP

Before writing the actual video prompt, define the reference map.

For example:

> REFERENCE 1 — CHARACTER
> Use only for facial identity, hairstyle, body proportions and general appearance.
>
> REFERENCE 2 — WARDROBE
> Use only for clothing design, silhouette, materials and accessories.
>
> REFERENCE 3 — LOCATION
> Use only for architecture, environment and spatial design.
>
> REFERENCE 4 — VEHICLE
> Use only for the vehicle's body design, colour, proportions and exterior details.

Then add:

> Do not blend identity, clothing, environment or object information between references unless explicitly instructed.

This creates clear boundaries.

---

5. THE TEXT PROMPT STILL DIRECTS THE SCENE

References tell the model:

what things should look like.

The text prompt tells the model:

what happens.

This distinction is essential.

A character reference may show someone standing still.

That does not mean your video needs to reproduce the pose.

Your prompt may say:

> The character walks towards the vehicle, opens the driver's door and gets inside.

The reference defines identity.

The prompt defines performance.

---

6. CONTROL REFERENCE BLEEDING

Reference bleeding happens when information from one image affects another element.

Examples:

The wardrobe reference changes the face.

The location reference changes the clothing.

The style reference changes the character.

The vehicle reference creates an additional car.

When this happens, strengthen the boundary.

For example:

> Reference 2 controls wardrobe only. Do not use facial features, hairstyle, body shape, pose, lighting or background information from Reference 2.

More precise boundaries usually work better than simply making the whole prompt longer.

---

7. CONTROL IDENTITY

When identity matters, make it the highest-priority reference.

For example:

> Reference 1 is the primary identity reference.
>
> Preserve the same facial structure, facial proportions, hairstyle, hair colour, skin tone and overall recognisable appearance throughout the complete shot.
>
> Do not blend facial features from any other reference.

Avoid extreme motion when identity preservation is critical.

Fast head rotations, heavy motion blur and severe lighting changes can make identity less stable.

---

8. CONTROL OBJECTS AND VEHICLES

Objects should also have identity.

If you are using a car reference, do not simply say:

> Use this car.

Instead:

> Reference 4 defines the vehicle.
>
> Preserve the same body shape, proportions, colour configuration, grille, headlights, wheels and overall exterior design throughout the entire shot.
>
> Exactly one vehicle matching Reference 4.
>
> No alternative version of the vehicle.

The same logic applies to:

- products;
- bags;
- furniture;
- watches;
- jewellery;
- architecture.

---

9. CONTINUITY RULES

For multi-reference generation, continuity rules are extremely useful.

Examples:

> Exactly one main character.
>
> Same identity throughout the full shot.
>
> Same wardrobe throughout.
>
> Exactly one hero vehicle.
>
> Vehicle design remains unchanged.
>
> Environment remains consistent.
>
> No duplicated characters.
>
> No duplicated objects.
>
> No identity blending.
>
> No wardrobe morphing.
>
> No background replacement.

Think of continuity rules as visual locks.

---

10. MULTI-REFERENCE PROMPT STRUCTURE

A strong multi-reference prompt can follow this structure:

REFERENCE MAP

Define the exact function of every reference.

FORMAT

Duration, aspect ratio, shot structure.

SUBJECT

Who is in the scene.

ENVIRONMENT

Where the scene takes place.

ACTION

What happens.

CAMERA

Framing and camera movement.

MOTION

Primary and secondary movement.

CONTINUITY RULES

What must remain unchanged.

NEGATIVE CONSTRAINTS

What must not happen.

This structure allows the model to understand both the visual assets and the scene logic.

PART B — RECOGNISABLE CHARACTER WORKFLOW WITH WAN

There is one workflow I keep separate from my standard Seedance process.

When I work with recognisable character references or media personalities, I may use **WAN** for the video generation stage.

The creative logic remains the same:

References + Reference Mapping + Scene Direction + Continuity Rules → WAN

The difference is the generation model.

Always follow the current rules of the platform you are using and make sure you have the appropriate rights for any commercial use of identifiable people or protected characters.

THE CHARACTER CARD METHOD

One of the strongest references you can create for a recurring character is a:

CHARACTER REFERENCE SHEET

or

CHARACTER CARD.

Instead of giving the video model one random photograph, you create one reference image containing several consistent views of the same character.

A strong character card can include:

- full-body front view;
- three-quarter view;
- side profile;
- back view;
- facial close-up;
- eye and facial-detail close-ups;
- torso / wardrobe details;
- sleeve or accessory details;
- footwear.

This gives the generation model significantly more visual information about the same identity.

WHY A CHARACTER CARD IS SO USEFUL

One portrait gives the model information about:

one angle.

A character card gives information about:

identity from multiple angles

-

body proportions

-

clothing

-

materials

-

accessories

-

small visual details.

This becomes especially useful when the character needs to:

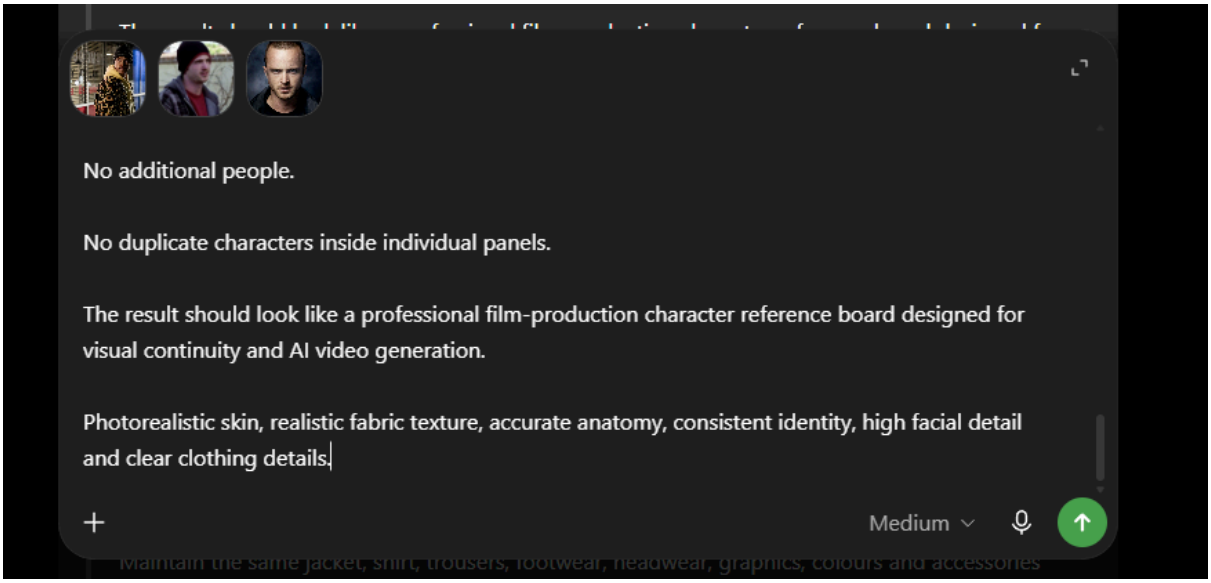
- turn;
- walk;
- enter or exit a vehicle;
- appear from different angles;
- move through multiple scenes.

HOW TO CREATE A CHARACTER CARD

Start with one clear reference photograph of the character you want to reproduce.

Upload it into GPT Image.

Then use a prompt based on the following template.



CHARACTER CARD MASTER PROMPT

Create a highly consistent professional character reference sheet based exclusively on the person in the uploaded reference image.

Preserve the same facial identity, facial proportions, eye shape, nose, mouth, jaw structure, skin tone, hairstyle, hair colour and overall recognisable appearance across every view.

Do not redesign or beautify the character.

Do not create different versions of the same person.

The entire sheet must depict ONE consistent individual.

Create a clean studio character sheet containing:

1. full-body front view;
2. full-body three-quarter view;
3. full-body side profile;
4. full-body back view;
5. large facial three-quarter close-up;
6. frontal facial close-up;
7. detailed eye / upper-face close-up;
8. torso and clothing detail;
9. sleeve / accessory detail;
10. footwear detail.

The character must wear exactly the same clothing in every panel.

Preserve all clothing colours, materials, graphics, patterns, accessories and footwear consistently.

> Use neutral standing poses for the full-body views.
>
> Keep body proportions identical across all views.
>
> Use a simple neutral studio background with soft even lighting.
>
> No cinematic environment.
>
> No dramatic shadows.
>
> No props unless they are part of the character's permanent outfit.
>
> No text labels.
>
> No captions.
>
> No additional people.
>
> No duplicate characters inside individual panels.
>
> The result should look like a professional film-production character reference board designed for visual continuity and AI video generation.
>
> Photorealistic skin, realistic fabric texture, accurate anatomy, consistent identity, high facial detail and clear clothing details.

IF THE CHARACTER HAS A SPECIFIC OUTFIT

Add:

> Preserve the outfit from the uploaded image exactly.
>
> Do not reinterpret the clothing.
>
> Maintain the same jacket, shirt, trousers, footwear, headwear, graphics, colours and accessories throughout every view.

If some part of the outfit is not visible in the original image, tell GPT Image what you want rather than allowing it to invent wildly different clothing.

IMPORTANT: IDENTITY CONSISTENCY

After generating the card, inspect it carefully.

Check:

- Do all facial views look like the same person?
- Does the side profile still preserve identity?
- Is the hairstyle consistent?
- Are facial proportions stable?
- Is the clothing identical?
- Are colours stable?
- Are shoes and accessories consistent?

If one panel looks like a different person, regenerate or correct the card.

A character sheet is only useful if all panels represent the same identity.

WHERE TO GENERATE WITH WAN

For this workflow, I use:

Magnific — https://www.magnific.com/

Magnific gives you access to multiple image and video models inside one platform, including WAN.

For the workflow in this section:

1. Open Magnific

Go to the Video Generator.

2. Select the model:

WAN 3.0 PRIME

3. Upload your references

The process is very similar to the multi-reference workflow we already covered in Higgsfield.

You upload the visual references required for the scene, for example:

- character reference;
- character card;
- vehicle reference;
- location reference;
- wardrobe reference;
- object reference.

4. Keep the reference order clear

If your prompt says:

Reference 1 = character
Reference 2 = vehicle
Reference 3 = house

your uploaded images must follow the same order.

The reference map and the uploaded assets must always match.

5. Paste the final prompt

Use the same principles we already covered:

- assign one role to each reference;
- define what each reference controls;
- define what the model should ignore;
- protect identity and object continuity;
- describe the action separately;
- define camera behaviour;
- add continuity rules;
- add negative constraints where necessary.

6. Choose your generation settings

Set the duration, aspect ratio and other available generation parameters according to the scene you are creating.

The exact interface may look different from Higgsfield, but the logic is almost identical:

References → Prompt → Settings → Generate

The platform changes.

The prompt-engineering principles do not.

IMPORTANT

Do not treat WAN as a completely different creative system.

Everything you learned earlier still applies:

Reference mapping

Reference hierarchy

Identity consistency

Object consistency

Camera direction

Motion logic

Continuity locks

Minimal iteration

The main difference is simply that you are generating through **WAN 3.0 PRIME inside Magnific** instead of Seedance 2.5 inside Higgsfield.

From this point, we can use the character + car + house example to see how the same multi-reference principles work inside WAN.

REAL WORKFLOW EXAMPLE — CHARACTER + CAR + HOUSE

Here is an example of how I used this system for a recognisable fictional character scene.

The concept was:

The character arrives at his house in his own car.

Instead of trying to describe everything from memory, I used three visual references:

Reference 1 — Vehicle



The character's red two-door coupe.

Reference 2 — Character Card



A multi-angle character sheet containing identity, clothing and body details.

Reference 3 — House



The exterior of the character's home.

Now the scene has three independently controlled visual components:

CHARACTER

-

CAR

-

LOCATION

REFERENCE MAPPING FOR THE SCENE

Use a mapping like:

> REFERENCE 1 — VEHICLE
>
> Use exclusively for the hero car.
>
> Preserve its red body colour, grey upper surfaces, two-door coupe proportions, rectangular front lighting, grille design, wheels and overall late-20th-century appearance.
>
> Exactly one hero vehicle matching Reference 1.
>
> Do not redesign or modernise the car.

REFERENCE 2 — CHARACTER

Use as the primary character identity reference.

Preserve the same face, facial proportions, hairstyle, headwear, body proportions, black layered streetwear, dark loose-fit trousers and black footwear.

Use the character card to maintain identity from different angles.

Do not blend identity or clothing information from the other references.

REFERENCE 3 — HOUSE

Use exclusively for the exterior location.

Preserve the white stucco exterior, tiled roof, arched architectural details, front lawn, mature trees and overall residential layout.

Do not reproduce people, cars or temporary objects from the reference unless explicitly requested.

WAN SCENE PROMPT — ARRIVING HOME

After mapping the references, build the scene.

A strong prompt could look like this:

FORMAT:

8-second continuous cinematic-realistic shot.

Natural late-afternoon residential light.

Grounded realism.

No montage.

No scene transition.

REFERENCE LOGIC:

Reference 1 controls the vehicle only.

Reference 2 controls the character's identity, body proportions and wardrobe only.

Reference 3 controls the house and immediate residential environment only.

Do not mix visual information between references.

SCENE:

The residential house from Reference 3 is seen from across the street in a natural wide shot.

The red coupe from Reference 1 enters frame at a believable residential speed and approaches the house.

The vehicle slows naturally and pulls up along the curb directly in front of the property.

Through the driver's side window, the character from Reference 2 is visible behind the wheel.

Preserve his facial identity, clothing and overall appearance from the character reference sheet.

The car decelerates realistically, the suspension settles slightly as it stops, and the vehicle remains parked in front of the house.

The character stays inside the vehicle at the end of the shot.

CAMERA:

Natural eye-level street perspective.

Mostly static camera with very subtle handheld human movement.

The camera makes a small restrained pan to follow the car as it enters and stops.

No orbit.

No crane movement.

No dramatic push-in.

No drone perspective.

VEHICLE PHYSICS:

Correct wheel rotation.

Correct forward direction.

Natural acceleration and braking.

Tires remain in contact with the road.

Correct vehicle scale.

Vehicle body remains rigid and consistent.

No morphing.

No sliding.

No floating.

CHARACTER CONTINUITY:

Exactly one character matching Reference 2.

Identity remains consistent throughout.

Same headwear.

Same clothing.

Same facial structure.

No duplicate character.

No face morphing.

VEHICLE CONTINUITY:

Exactly one hero car matching Reference 1.

Same body shape, colour configuration, wheels, grille and exterior design throughout the shot.

No additional version of the hero car.

LOCATION CONTINUITY:

The house remains visually consistent with Reference 3.

Preserve the architecture, roof, lawn, major trees and overall structure.

No transformation into another house.

No modern redesign.

FINAL FEEL:

The scene should feel like a believable moment captured in the real world:

the character simply arriving home in his familiar car.

Understated, natural and physically believable rather than overly cinematic.

WHY THIS SCENE IS A GOOD MULTI-REFERENCE EXAMPLE

The scene contains three things that need visual consistency:

the person

the vehicle

the location

Trying to describe all three only with text would give the generation model significant freedom to redesign them.

References remove much of that uncertainty.

But the references still need clear roles.

The model must understand:

> This image defines the person.
>
> This image defines the car.
>
> This image defines the house.

The prompt then defines:

> how those three elements interact.

That is the core principle behind advanced multi-reference generation.

BUILDING MORE SCENES WITH THE SAME CHARACTER

Once you have a strong character card, you can reuse it.

For example, the same character could:

- exit the vehicle;
- walk towards the house;
- enter a shop;
- sit inside a diner;
- stand beside the car;
- walk through another environment.

You do not need to recreate the identity every time.

The character card becomes your reusable identity asset.

You can then change:

LOCATION

VEHICLE

WARDROBE

ACTION

TIME OF DAY

while maintaining the same core character reference.

CHARACTER CARD + MULTI-REFERENCE = REUSABLE VISUAL SYSTEM

This is one of the most useful ways to think about advanced AI video.

Instead of creating every generation independently, build reusable visual assets.

For example:

CHARACTER CARD

-

WARDROBE REFERENCES

-

VEHICLE REFERENCES

-

LOCATION REFERENCES

-

PRODUCT REFERENCES

Then combine only the assets required for a specific shot.

This gives you a much more scalable production workflow.

THE MAIN LESSON

Multi-reference generation is not about uploading as many images as possible.

It is about separating visual information into controlled components.

A strong workflow looks like:

Character Reference

-

Object / Vehicle Reference

-

Location Reference

-

Clear Reference Mapping

-

Scene Prompt

-

Continuity Rules

-

Motion Logic

The character card gives you identity consistency.

The individual references give you control over the world.

The prompt determines what happens.

The generation model brings those components together into one moving scene.

FINAL — BUILD YOUR OWN AI CONTENT SYSTEM

You now have the complete workflow.

You know how to move from a simple idea to a structured AI video generation process.

More importantly, you now understand that strong AI content is not created by one perfect prompt.

It is created through a system.

The core process is:

Idea → Structure → Prompt → References → Generation → Review → Refinement

Depending on the scene, you can choose between three main workflows:

Text to Video
Use when you want speed, flexibility and more creative freedom.

Image to Video
Use when you want stronger control over the starting frame, character, styling, composition or environment.

Multi-Reference Video
Use when several specific visual elements need to exist inside the same generation.

For more advanced identity-based workflows, you can build reusable character cards and combine them with location, vehicle, wardrobe or object references.

KEEP THE PROCESS SIMPLE

You do not need to use every technique in every generation.

Use only what the scene actually needs.

A simple scene may only require:

Idea → Prompt → Generate

A more controlled scene may require:

Idea → GPT Image → Source Image → Motion Prompt → Generate

A complex scene may require:

Idea → Character Reference + Location Reference + Object Reference → Reference Mapping → Prompt → Generate

The goal is not complexity.

The goal is control.

YOUR CORE TOOLKIT

Your workflow can now be built around:

GPT / Claude
For structuring ideas, creating prompts and refining generations.

GPT Image
For creating source images, characters and visual assets.

Higgsfield AI + Seedance 2.5
For your main video generation workflows.

WAN
For selected advanced workflows where you need an alternative video-generation approach.

And most importantly:

The Seedance Master Prompt Guide

Use it as your prompt-writing framework rather than trying to memorise every technical detail yourself.

BUILD YOUR OWN LIBRARY

From this point forward, start saving:

- strong prompts;
- failed prompts;
- useful corrections;
- character cards;
- location references;
- wardrobe references;
- recurring continuity rules;
- camera instructions that work;
- motion instructions that work.

Do not build only a prompt library.

Build a **decision library**.

Document:

> What was the problem?
>
> What did I change?
>
> What improved?

This is how your workflow becomes faster over time.

FINAL PRINCIPLES

Before every generation, remember:

Start with the idea, not the tool.

One clear action is better than five competing actions.

References need roles, priorities and boundaries.

A strong source image makes Image to Video easier.

Specificity is more useful than unnecessary length.

Continuity should be protected explicitly when it matters.

Do not rewrite an entire prompt to fix one small problem.

If the problem is random, regenerate before rewriting.

If the foundation is wrong, rebuild the foundation.

And most importantly:

You make the creative decisions. AI executes them.

The better you become at communicating those decisions, the more control you gain over the final result.

FROM HERE

The best way to improve is to create.

Take one idea.

Choose the simplest workflow that can achieve it.

Generate.

Analyse the result.

Change one thing.

Generate again.

Repeat.

Over time, you will stop thinking in terms of:

> "What prompt should I use?"

and start thinking in terms of:

> "What information does the model need to create exactly what I want?"

That is the real skill behind AI content creation.