



The Prompt AI Handbook

A complete guide to writing image & video prompts — the formula, per-style recipes, bad→good examples, and how to fix AI-slop. Beginner to advanced.

SON SUNNY · [SON-SUNNY.COM](https://son-sunny.com)

A prompt is not magic

It is a **formula**. Every good AI image you have seen comes from a structured prompt — and once you know the structure, you can recreate anything, reliably, not by luck.

This book packs the exact formula we use daily: the 8-field order for images, the timeline for video, per-style recipes, bad→good examples with the reason, how to fix "AI-slop", and one full build from a blank page.

FASTEST PATH

Don't want to learn yet? Drop any picture into **Prompt Mirror** (son-sunny.com/mirror.html) — it returns exactly the formula in this book. The book teaches you to *understand* what the machine does, so you can write it yourself.

Contents

1 • Foundations: what a prompt is	the formula mindset
2 • The IMAGE formula (8 fields)	each field + example
3 • Four golden rules for images	specific · safe-zone · negatives · seed-lock
4 • Recipes by STYLE	music cover · portrait · product · food
5 • The VIDEO formula	timeline · camera · motion
6 • Bad → good library	learn by contrast
7 • Fixing AI-slop	6 traps + the fix
8 • Which model for what	a chooser
9 • Build a prompt from scratch	full worked example
10 • Copy-paste appendix	use now

Foundations: what a prompt is

An AI model does not "understand your intent." It reads words and builds whatever best matches those words. So **vague words = vague image**. The whole craft fits in one line: say exactly what can be SEEN.

Three root principles

- **Specific beats fancy.** "Beautiful", "high quality", "masterpiece" tell the model nothing. "Yellow raincoat under purple neon" does.
- **Order matters.** Models weight the start most. Put the subject first; push technical constraints (camera, format) to the end.
- **One variable at a time.** When tuning, change exactly one thing per pass so you know what caused what.

THE FORMULA MINDSET

Don't ask "how do I get a nice image." Ask "what COMPONENTS is this image made of" — subject, light, angle, color — then fill each in. That is the whole secret.

The IMAGE formula — 8 fields

Order them like this. The most important thing (subject) goes first; constraints (camera, format) go last.

1 • Subject	Specific subject, visible detail. <i>a girl in a yellow raincoat</i>
2 • Action	Action / pose. <i>looking up at a sign, holding her hat</i>
3 • Style	Style + medium. <i>cinematic photo, 90s Hong Kong film look</i>
4 • Composition	Layout + subject placement. <i>subject left, street receding right</i>
5 • Light	Specific light. <i>purple neon backlight, reflecting on wet asphalt</i>
6 • Grade	Color / grade. <i>deep teal-orange, light film grain</i>
7 • Camera	Lens + shot size. <i>85mm, close-up at shoulder level</i>
8 • Format	Aspect ratio. <i>9:16 (vertical) · 16:9 (wide) · 1:1 (square)</i>

PUT TOGETHER

a girl in a yellow raincoat, looking up at a sign, cinematic 90s HK film photo, subject left with street receding right, purple neon backlight, teal-orange grade with film grain, 85mm close-up, 9:16.

Four golden rules for images

1 • Specific, no empty adjectives

BAD

a beautiful girl, high quality, masterpiece

GOOD

a girl in a yellow raincoat under purple neon, wet street after rain

2 • Safe-zone for covers/thumbnails

Will text sit on top? Declare the aspect ratio + reserve deliberate empty space. E.g. "*subject left, leave negative space top-right for the title, 16:9*". Skip it and text lands on the face.

3 • Negative prompt — ban the AI tells

Keep it SHORT, one real problem each:

negative: watermark, deformed hands (extra fingers), gibberish text, blurry

NOTE

Some VIDEO models (and a few image ones) actually attend TO what you ban — say "no dog" and a dog appears. For those, don't ban — describe the positive thing you DO want, so the unwanted has no room.

4 • Seed-lock — tune one variable

Got a frame you like? Lock the seed and change exactly one field per pass (lens / grade / pose / time of day). You will know precisely what caused each change — instead of re-rolling and losing the good frame.

Recipes by STYLE

Drop your subject into these ready templates.

Music cover — phonk / night lofi

[subject], cinematic night look, [color] neon backlight, haze, high-contrast teal-orange grade, 35mm film grain, subject offset [left/right] leaving room for the title, 1:1
negative: watermark, gibberish text, blurry

Portrait — cinematic

portrait of [person], one-source Rembrandt light from the [left], dark background separating the subject, natural detailed skin, 85mm f1.8 shallow focus, soft warm grade, 4:5
negative: deformed hands, extra fingers, plastic skin, blurry

Product — for sale

[product] on a minimal [color] background, studio softbox light, subtle reflection underneath, 3/4 angle, tack sharp, clean tone, leave empty space on the right for the price, 1:1
negative: watermark, clutter, harsh shadow, blurry

Food

[dish] steaming hot, shot from a 45-degree top angle, side window light, visible steam, rustic wood surface, warm appetizing color, close-up, 4:5
negative: gibberish text, plastic look, blurry

The VIDEO formula

Video adds a TIME axis. Three backbone rules:

1 • Timeline — beats with marks

Write it as marked beats **[0s] ... [4s] ... [7s]**, declaring total length first. Budget: each beat needs ~2 seconds of breathing room — a 5s clip fits 2–3 beats. More and the model compresses it into mush.

2 • Camera vocabulary — one move per beat

Fixed set: **push-in • pull-out • orbit • pan • tilt • dolly • crane • handheld • locked-static** + a speed word (slow/steady/whip). Stacking moves in one beat = drift soup.

3 • Motion dial — the right dose

subtle (gentle drift) • **moderate** (one clear action) • **dynamic** (fast, layered). Match the scene: subtle for a portrait, dynamic for a race.

clip 10s, 9:16, motion: moderate —
 [0s] a girl opens the cafe door, bell rings, camera: push-in slow
 [4s] coffee being poured, steam rising, camera: locked-static
 [7s] she looks up and smiles, warm light blooms, camera: tilt slow

EMOTION = ACTION

Don't write "very sad" — the model can't draw the word "sad". Describe body action: "shoulders trembling, hands gripping the cold cup." One or two actions — don't stack them into a twitching puppet.

Bad → good library

The fastest way to learn is contrast. Each pair: same idea, different writing.

BAD

a nice logo for a coffee shop

GOOD

minimalist logo, the word "Suong", a single brush stroke, deep brown on cream, balanced, clean vector, 1:1

BAD

a majestic mountain landscape

GOOD

misty morning mountains, sun just rising behind the peak, pine forest foreground, low golden sidelight, cool grade, 16:9

BAD

a professional product photo

GOOD

a glass perfume bottle on white marble, softbox from the left, subtle reflection, 3/4 angle, tack sharp, room for text on the right, 1:1

BAD

a cool anime character

GOOD

a silver-haired man in a long coat, standing in neon city rain at night, modern anime style, blue-purple light, dynamic framing, 9:16

THE COMMON LAW

The "Good" column always adds: concrete detail (color/material), light, angle/framing, and aspect ratio. The "Bad" column is only adjectives. Learn to TURN adjectives into detail.

Fixing AI-slop

Check the spec BEFORE spending a render. Six common traps and their fix:

SYMPTOM	CAUSE	FIX
Video mush	Too many beats for the seconds	Fewer beats, each $\geq 2s$
Torn in half	Contradictory pair (slow-mo + high-speed)	Drop one
Mime / pointless flailing	Wrong motion dose	Match dial to scene
Melty morph	Bracket endpoints too far apart	Same place, same light
Drift soup	Stacked camera moves in one beat	One move per beat
Wooden face	Naming the emotion	Describe body action
Deformed hands	Missing negatives	Add "deformed hands, extra fingers"
Gibberish text	Model can't render text	Add "gibberish text" to negatives, or add text in post

CHARACTER DRIFT (MULTI-SHOT VIDEO)

To keep the SAME character across scenes: repeat the description block VERBATIM (or use a seed / reference image) in every shot. Writing "roughly similar" each time turns the character into someone else over the sequence.

Which model for what

There is no "best" model — only the right one for the job. A quick chooser (principles, not a price list):

JOB	PREFER	WHY
Photoreal, portraits	Strong photoreal image model	Skin, light, natural detail
Images WITH text (covers, posters)	A model that renders text	Most render text poorly — pick a text-capable one, or add text in post
Illustration / anime style	Illustration/anime model	Linework, color, shapes fit the style
Seamless single-gesture video	First-last-frame model	Smoothly interpolates two endpoints
Multi-beat video	Timeline/beat model	Control pacing second by second
Editing one region	Inpaint model	Change one part, keep the rest

HOUSE RULE

Own the FORMULA (the prompt), rent the ENGINE (the model). Models get renamed, raise prices, shut down — your formula stays intact and runs on the next model. Don't chain yourself to one model.

Model names change constantly — this is a CHOOSING principle, not a product list. Check what's available before you run.

Build a prompt from scratch

Brief: **a lofi album cover about a rainy night, melancholy.** Step by step.

Step 1 — subject + action

"a small room looking out at a rainy night window, a cup of tea steaming on the desk"

Step 2 — style

+ "warm lofi style, soft digital painting"

Step 3 — composition + safe-zone (it's a cover!)

+ "desk in the lower foreground, window filling the upper half, leave empty space top-right for the album title"

Step 4 — light + grade

+ "warm yellow desk lamp as the key light, cold blue outside, warm-cool contrast, soft grade"

Step 5 — camera + format

+ "eye-level angle, 35mm lens, 1:1"

Step 6 — negatives

+ "negative: watermark, gibberish text, blurry"

FINISHED PROMPT

a small room looking out at a rainy night window, a cup of tea steaming on the desk, warm lofi soft digital painting, desk in lower foreground with window filling the upper half leaving empty space top-right for the album title, warm yellow desk lamp key light with cold blue outside, warm-cool contrast, soft grade, eye-level 35mm, 1:1. negative: watermark, gibberish text, blurry

Then seed-lock: swap "desk lamp" → "candle" next pass if you want it warmer. One variable at a time.

Copy-paste templates

Image template

[specific subject], [action], [style + medium],
[composition + placement], [light], [color/grade],
[lens + shot size], leave negative space [where] for the title, [16:9/9:16/1:1]
negative: watermark, deformed hands, gibberish text, blurry

Video template (10s)

clip 10s, [9:16], motion: [subtle/moderate/dynamic] —
[0s] [action], camera: [one move]
[4s] [action], camera: [one move]
[7s] [action], camera: [one move]

Shared negative set

image: watermark, deformed hands, extra fingers, gibberish text, blurry
portrait: + plastic skin, extra limbs
product: + clutter, harsh shadow

Put the formula to work

Prompt Mirror — drop any picture, get the 8-field formula + a video idea: **son-sunny.com/mirror.html**

Learn Lab — practice writing prompts, auto-graded per lesson: **son-sunny.com/learn/**

Prompt Gallery — real images with real prompts, copy and use: **son-sunny.com/gallery.html**