

sixth-switch-2ac.notion.site /Hailuo-AI-tutorial-1446c20a98eb802c94c5d9ebcb497401

Hailuo AI tutorial

minimax_hailuo : 21-27 minutes

Hailuo AI's Subject Reference Model enables the recognition of facial features from photos uploaded by users. It can then generate a character with corresponding facial traits and create video scenes featuring that character based on text prompts.

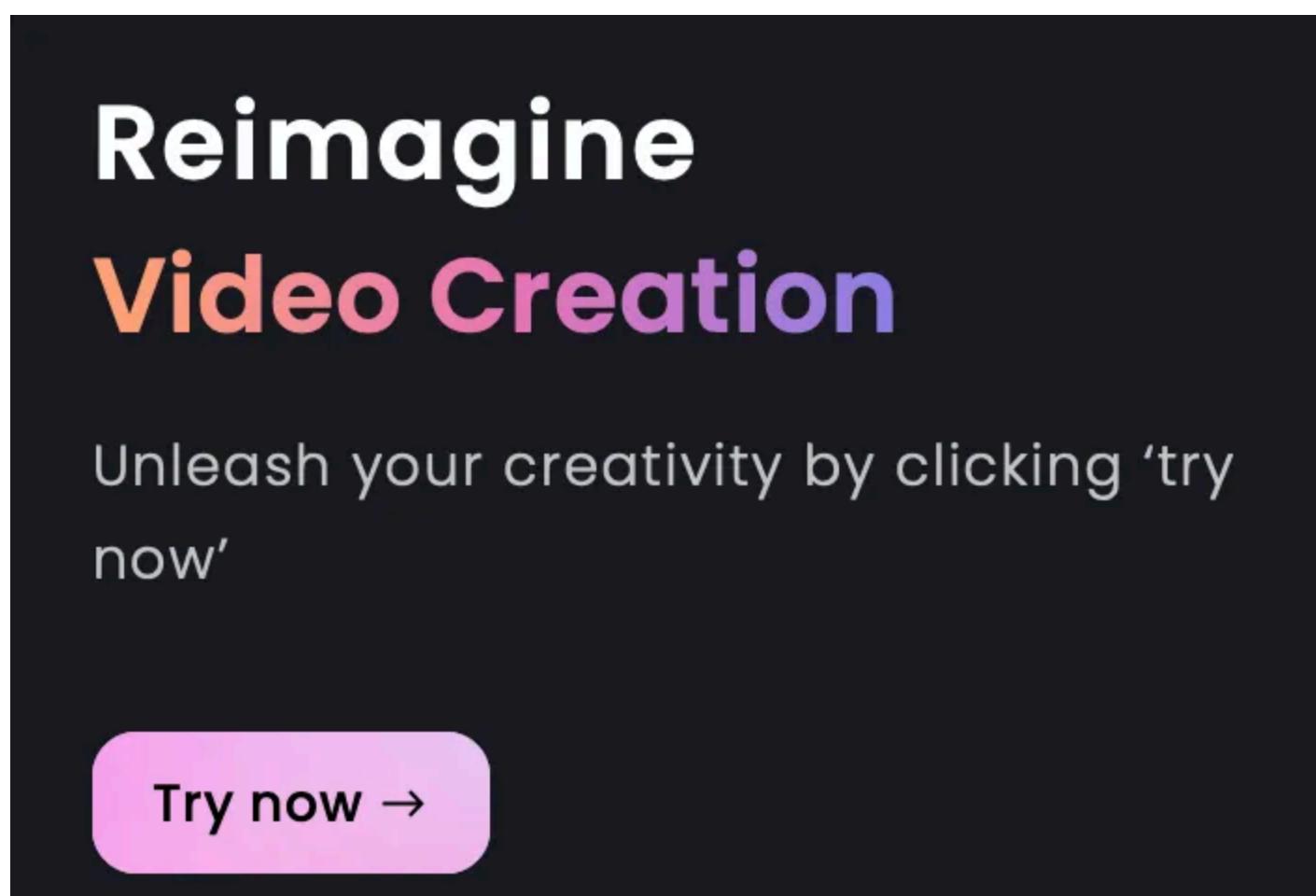
Subject Reference Model Tutorial

To generate videos using Hailuo AI's Subject Reference Model, users must provide two key inputs:

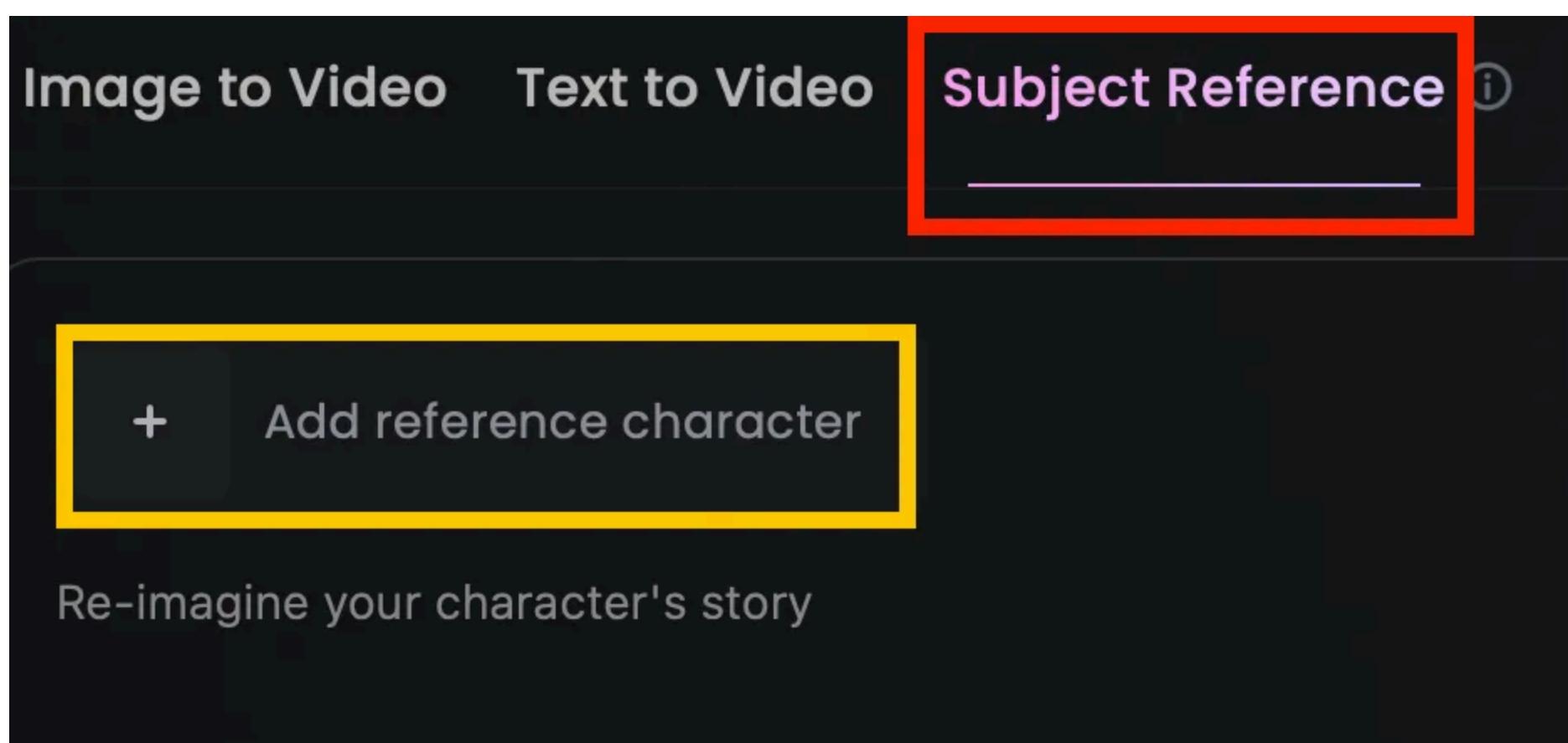
Subject's Facial Photo

Prompt for Video Generation

The first step is to click "Try Now" to enter the creation page.



The second step is to select the Subject Reference Model, then click "Add Reference Character," upload the image, enter the prompt, click "Create", and after a brief wait, your creation will be ready.



Subject's Facial Photo The Subject Reference Model requires a clear, recognizable photo of the subject's face to serve as a reference for generating the character's face in the video. Currently, the model only supports facial recognition for human faces.

Hailuo AI is capable of accurately identifying facial features across various genders, ages, and ethnicities. To achieve optimal video generation results, ensure that the uploaded image meets the necessary specifications, including proper angles and lighting conditions.

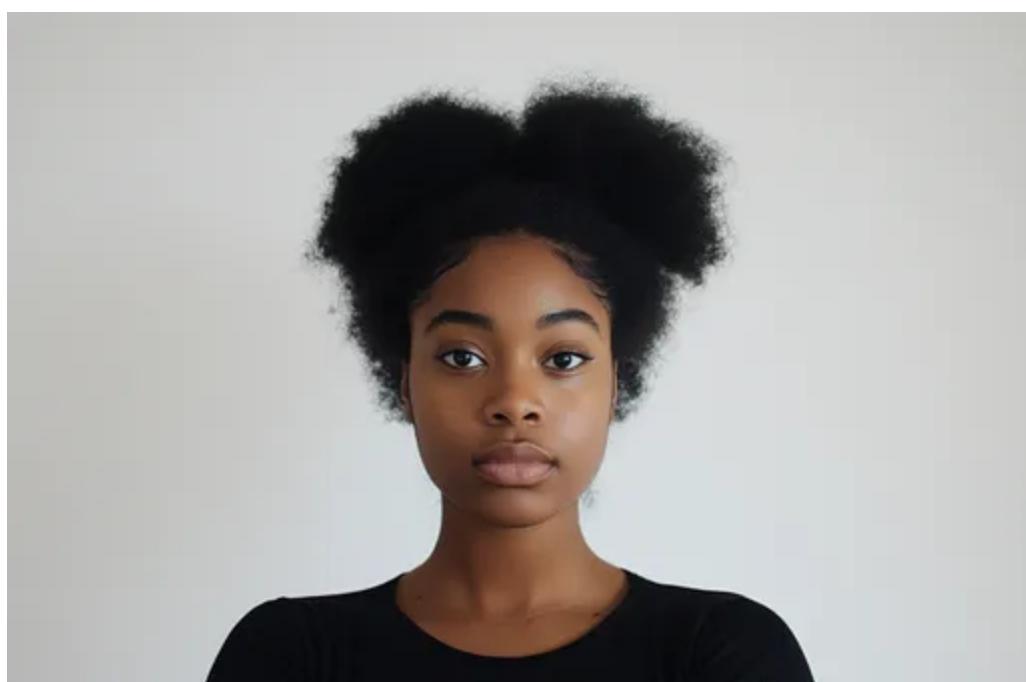
Face Count: The Subject Reference Model only supports uploading single-person photos for facial recognition.

Image Size: Resolution should be no lower than 120x120, and the image file should not exceed 20MB.

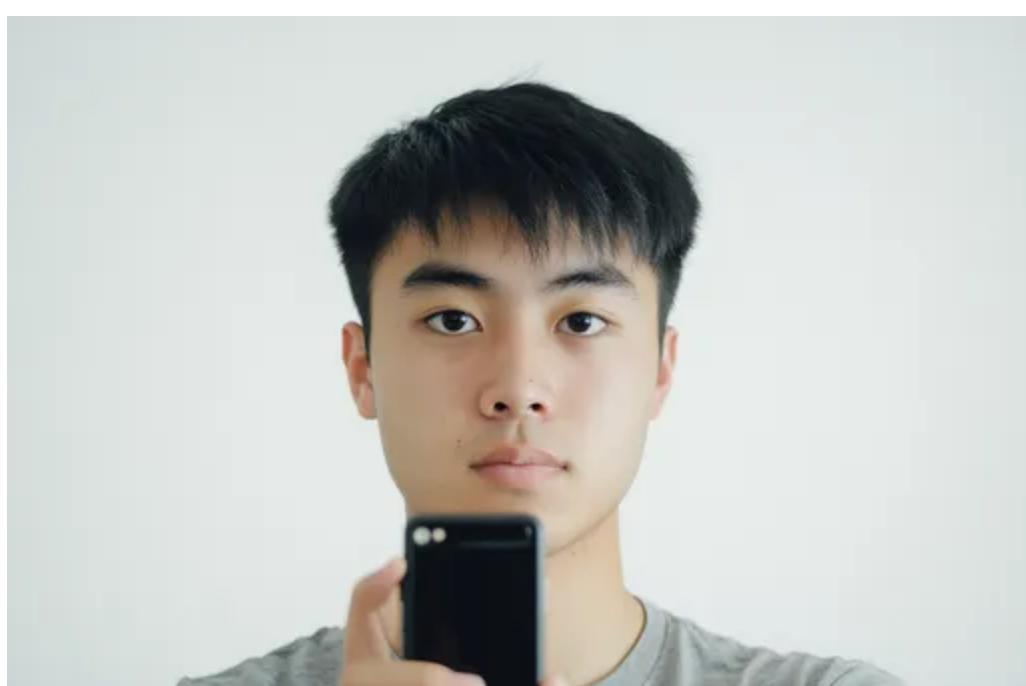
Other Requirements: The face must not be obstructed, and the photo should not be out of focus or contain excessive filters or special effects.

Standard Facial Photo :





Examples of photos not suitable for reference :



Facial obstruction

ALT



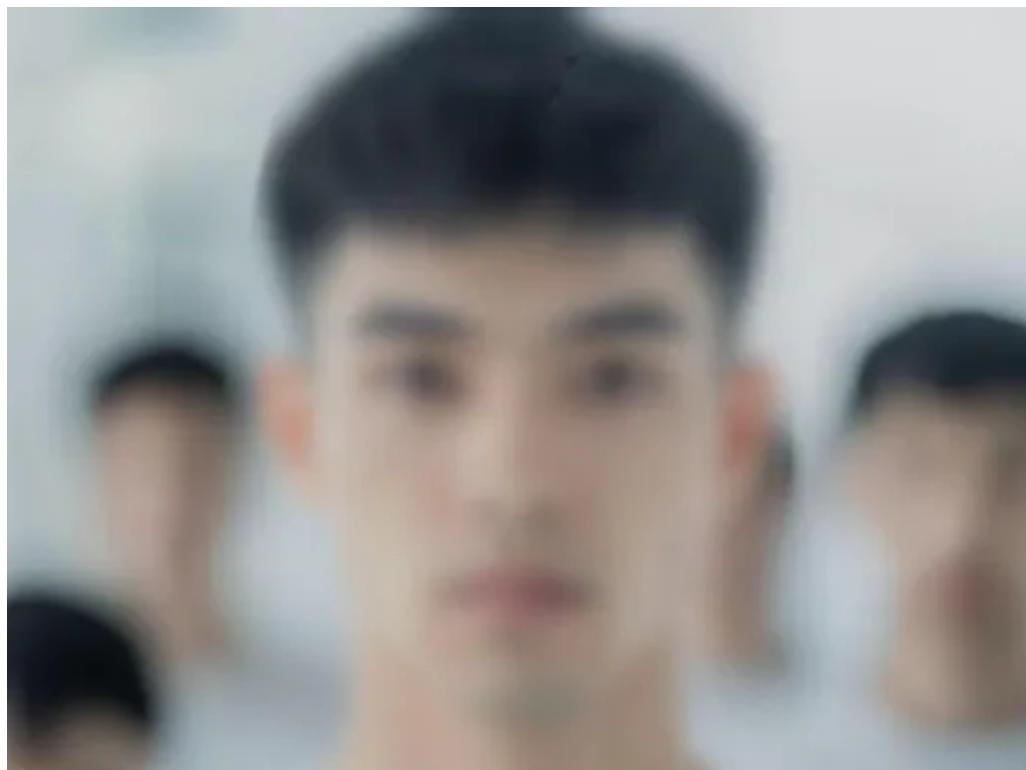
Multiple faces

ALT



Heavy filter effects

ALT



Excessive blurriness

ALT

The subject angle in the uploaded photo should be either directly facing the camera or slightly turned to the side.

Directly facing the camera: The subject's face should be directly facing the camera, with clear facial features visible in the frame, providing the most complete facial information.

Slightly turned to the side: The subject's facial features and facial contours should still be fully visible in the photo, avoiding profiles or angles where the facial features are not within the frame.

Examples of quality camera angles :

Generated Video 1

Generated Video 2

Facial Photo 1



Prompt 1

On an overcast day, in an ancient cobbled alleyway, the model is dressed in a brown corduroy jacket paired with beige trousers and ankle boots, topped with a vintage beret. The shot starts from over the model's shoulder, following his steps as it captures his swaying figure. Then, the camera moves slightly sideways to the front, showcasing his natural gesture of adjusting the beret with a smile. Next, the shot

slightly tilts down, capturing the model's graceful stance as he leans against the wall at a corner. The video concludes with an upward shot, showing the model smiling at the camera. The lighting and colors are natural, giving the footage a cinematic quality.

Facial Photo 2



Prompt 2

On an overcast day, in an ancient cobbled alleyway, the model is dressed in a brown corduroy jacket paired with beige trousers and ankle boots, topped with a vintage beret. The shot starts from over the model's shoulder, following his steps as it captures his swaying figure. Then, the camera moves slightly sideways to the front, showcasing his natural gesture of adjusting the beret with a smile. Next, the shot slightly tilts down, capturing the model's graceful stance as he leans against the wall at a corner. The video concludes with an upward shot, showing the model smiling at the camera. The lighting and colors are natural, giving the footage a cinematic quality.

[Generated Video 3](#) [Generated Video 4](#)

Facial Photo 3



Prompt 3

In a softly lit modern living space, the model wears a cozy off-white knit sweater paired with high-waisted jeans and minimalist white sneakers, exuding a comfortable and casual vibe. Behind her is a large glass window letting in the sunlight, which spills onto the wooden floor. The model sits leisurely on the sofa, flipping through a book, with a cup of hot coffee beside her. The camera slowly zooms in, capturing the moment she gently lifts her head and gazes out the window, her expression calm and gentle. Then, the camera circles around to her side, showing her adjusting her hair.

Facial Photo 4



Prompt 4

In a softly lit modern living space, the model wears a cozy off-white knit sweater paired with high-waisted jeans and minimalist white sneakers, exuding a comfortable and casual vibe. Behind her is a large glass window letting in the sunlight, which spills onto the wooden floor. The model sits leisurely on the sofa, flipping through a book, with a cup of hot coffee beside her. The camera slowly zooms in, capturing the moment she gently lifts her head and gazes out the window, her expression calm and gentle. Then, the camera circles around to her side, showing her adjusting her hair.

Examples of subject angles not suitable for reference :

Facial features are not fully visible in the fram Profile view with insufficient facial features





Facial features are not fully visible in the frame

ALT



Profile view with insufficient facial features

ALT

To achieve the most accurate video representation of the face from the uploaded photo, the subject's face should avoid large areas of dark shadows or localized overexposure. The lighting should be natural, with some subtle shadow details on the face for better depth.

Examples of standard facial lighting :

[Generated Video 5](#) [Generated Video 6](#)

[Facial Photo 5](#)

**Prompt 5**

A cowboy rides a horse through a snowy landscape, with a rifle strapped to his back. He wears a wide-brimmed hat, and his rugged, striking facial features stand out prominently. In the distance behind him is a tall spruce forest, beyond which rise the breathtaking, snow-covered mountains. The ground is blanketed in pristine white snow as the cowboy slowly rides through the cold, open wilderness. His gaze is fixed ahead, exuding a sense of determination and bravery.

Facial Photo 6**Prompt 6**

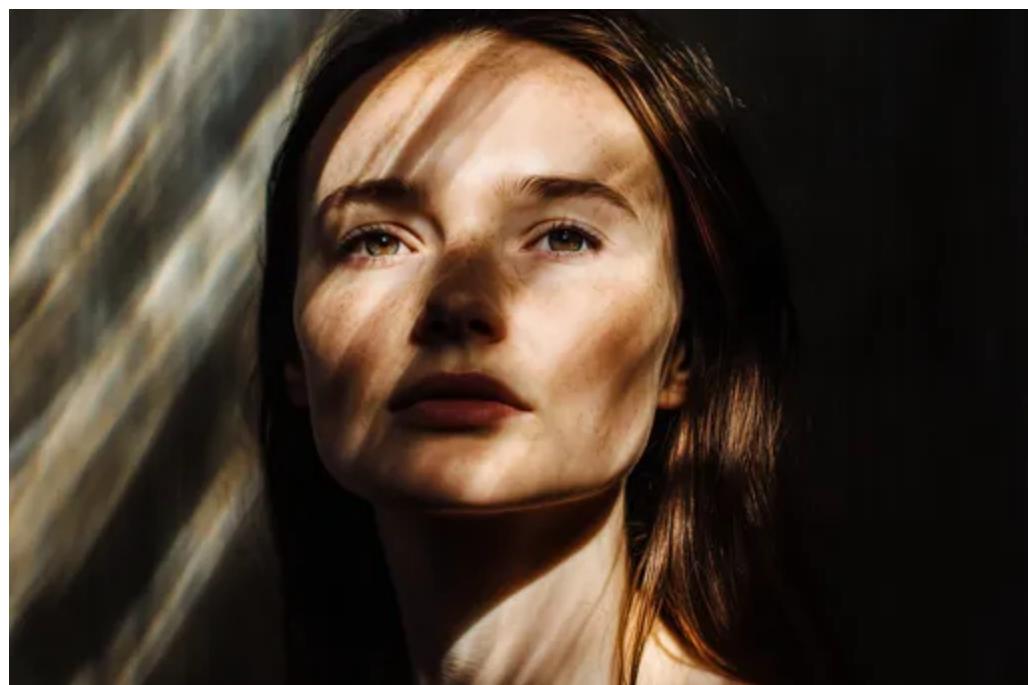
A model, dressed in a vintage apron with a ribbon tied around her neck and high-heeled ankle boots, stands behind the dimly lit wooden bar of a tavern. She holds a glass filled with whiskey in one hand, while the background features a wooden shelf lined with bottles and oil lamps. Her curly hair cascades over her shoulders, and a mischievous smile plays on her face as the camera moves dramatically.

Examples of unsuitable facial lighting for reference :



Shadows obscure the facial features

ALT



Complexed facial shadows

ALT



Excessive or unnatural facial lighting colors

ALT



Overexposure

ALT

The generation logic of the Subject Reference Model is similar to that of text-to-video generation. For more details, please refer to the Text-to-Video Prompt tutorial.

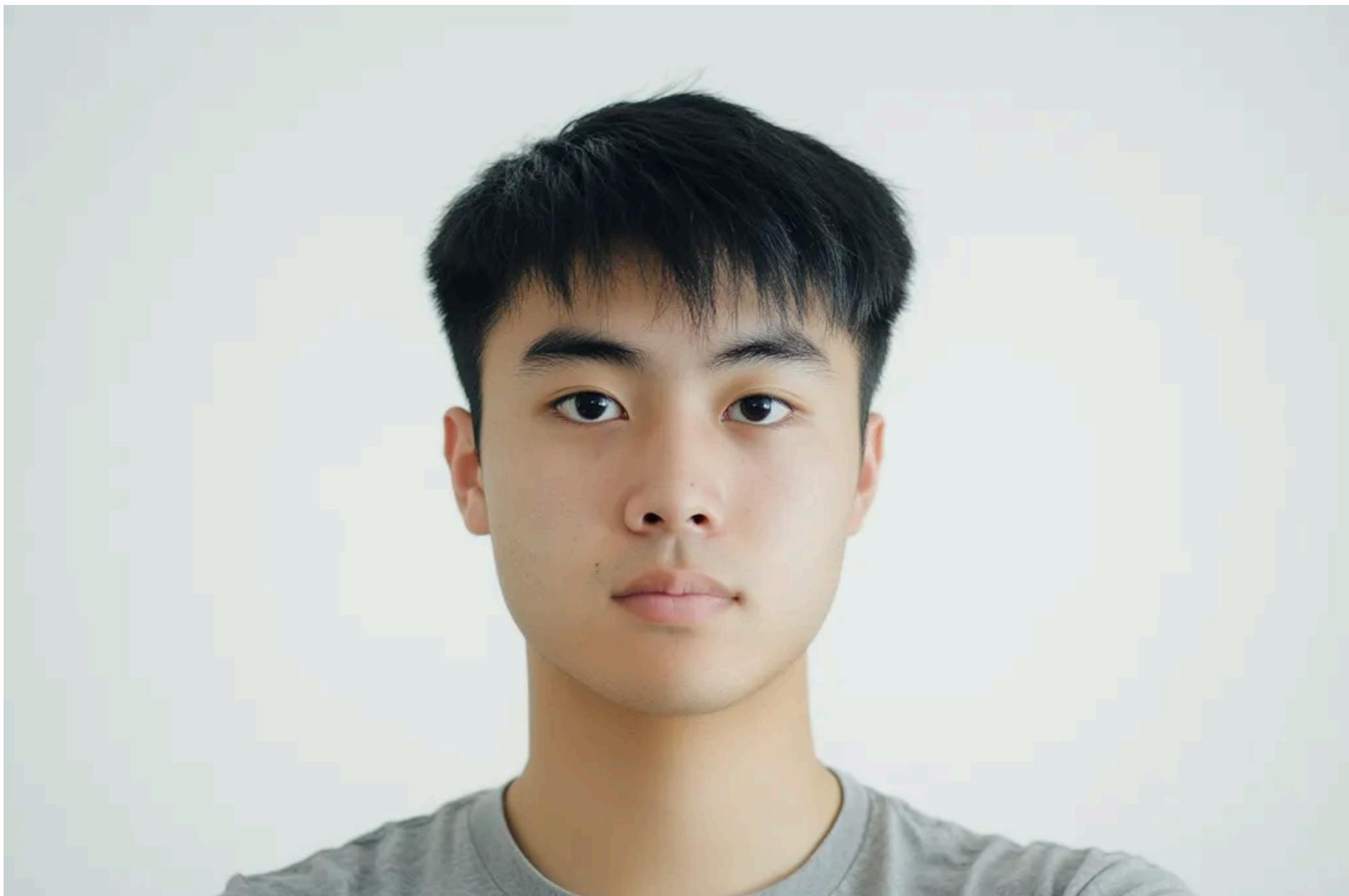
The prompt used for the Subject Reference Model should focus on a single subject to achieve the most accurate results.

If the prompt involves multiple subjects, additional descriptions of the specific subjects can be included. Hailuo AI can identify the corresponding facial features based on age and gender and generate the appropriate subject accordingly.

Adding gender information of the face in the prompt can help Hailuo AI lock onto the correct subject and generate the expected scene.

Generated Video 7 Generated Video 8

Facial Photo 7



Prompt 7

In the scene, a black-haired man wearing a suit sits at a European-style dining table, looking at the camera. Next to him is a woman with brown hair, smiling and gazing at the camera.

Facial Photo 8



Prompt 8

In the scene, a black-haired man in a suit sits at a European-style dining table, looking at the camera. Beside him is a woman with brown hair, smiling and gazing at the camera.

2. Adding age information of the face in the prompt can help Hailuo AI lock onto the correct subject and generate the expected scene.

[Generated Video 9](#) [Generated Video 10](#)

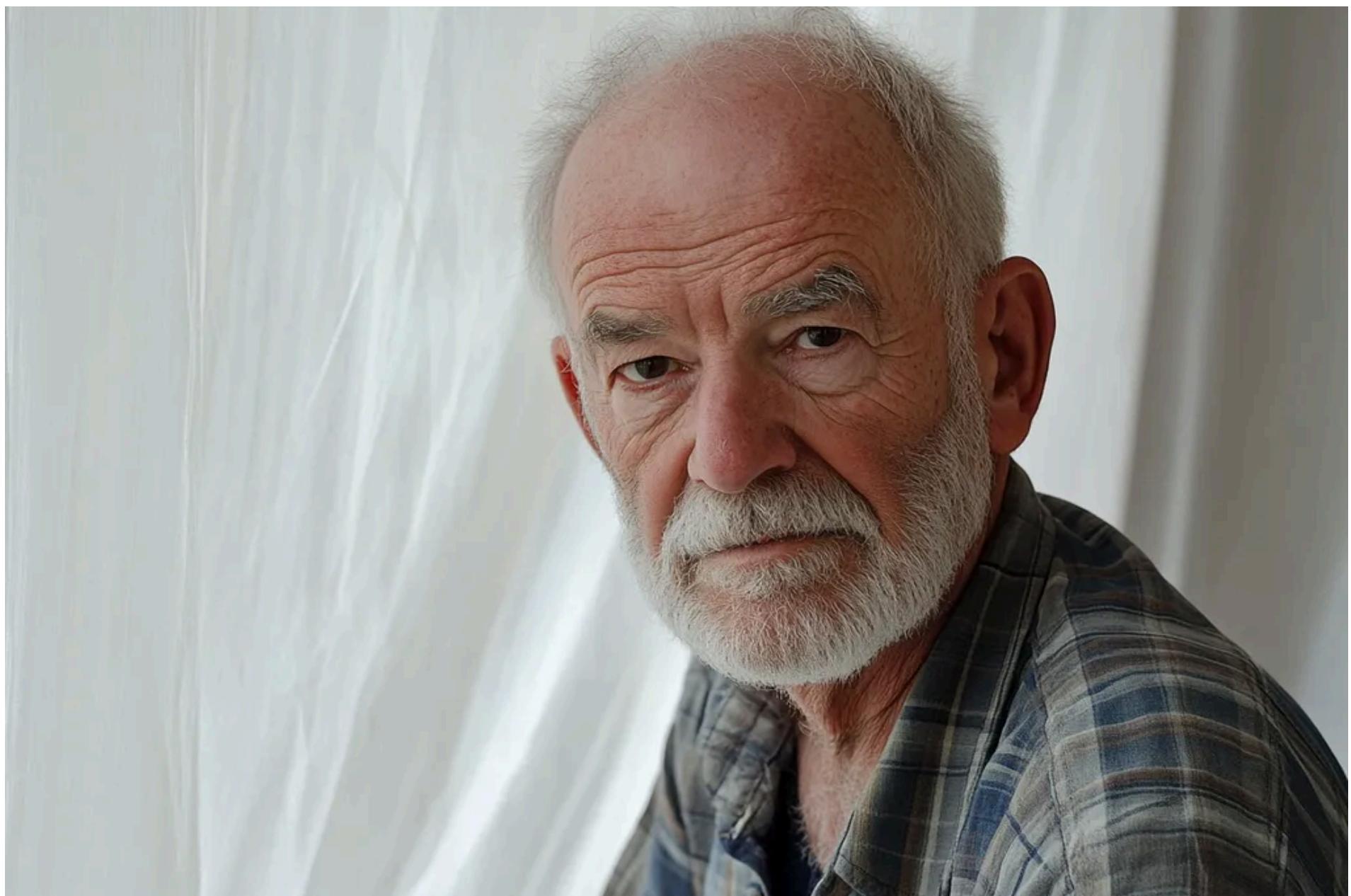
Facial Photo 9



Prompt 9

An elderly man in a shirt is holding a young girl in a red dress. They are sitting together in a cozy indoor setting, embracing and smiling at each other.

Facial Photo 10



Prompt 10

An elderly man in a shirt is holding a young girl in a red dress. They are sitting together in a warm, cozy indoor setting, embracing each other and smiling.

Prompt is the key to the world of generated AI, constructed from your inspiration and creativity.

Whether you are a newcomer to AI or a veteran who is already proficient in using AI to empower productivity, you can easily master the skills of writing prompts through a clear structure and a flexible way of expression. The vision of "one person + one AI = one professional crew" is right around the corner.

To improve the quality of video results, we have developed two Prompt formulas of your use:

Basic Prompt Formula is designed for those who are uncertain about the exact videos they wish to create, or those who seek to use Hailuo AI as a source of inspiration. This type of prompt will offer you a range of more incentive visual ideas.

Basic Prompt Formula= Main Subject + Scene + Motion

Main Subject: This core message of the video can be a person, an animal, another object or an imaginary object that does not exist.

Scene: The scene describes the environmental information of the video, etc a library, a coffee shop or even a fantastical fictional space.

Motion: Motion is the description of the main object's moving status in the video, such as stillness, movement; the transformation of the environment and space, and so on.

Example

A puppy runs in the park.

A woman with an umbrella walks down a rainy street.

A small stream flows quietly in a valley.

The Prompt Precision Formula is designed for users requiring advanced camera movement or image presentation, aiming for a more professional-grade video production. Precise and detailed prompts can enhance the visual quality of the footage.

Precise Prompt Formula = Main Subject + Scene + Motion + Camera Movement + Aesthetic Atmosphere

Camera Movement: You can use common camera movements, such as tracking, zooming, panning, handheld and other camera movements to restrict the final presentation.

Aesthetic Atmosphere: You can describe the visual style and the sense of atmosphere to get a better result.

For Example :

A couple sit on a park bench communicating. The camera maintains a fixed shot of the couple. The color tone of the picture is warm, the atmosphere is cozy.

A lamb is grazing in a meadow. The camera slowly pushes forward the lamb. The color tone of the picture is natural and realistic.

A man in a suit eats noodles in a noodle shop. The camera gradually pulls away to show the noisy environment of the noodle shop. The picture has a natural color tone.

There's no right or wrong to write a prompt. But overall, prompt writing follows 2 main basic principles:

The more precise the expression is, the more accurate the result will be.

The richer the expression gets, the better the overall effect will be.

Hailuo AI is adept at responding effectively to various camera movements. It can provide professional video users with well-defined video footage.

Add detailed orders to the camera movement

For example, "Lens rotating down" can be optimized as :

"The camera descends slowly and then wraps around to the right as it descends."

Add detailed descriptions to the change made with camera movement

For example, "The camera shoots the giraffe from the bottom up" could be optimized as:

"The camera starts with a close-up of the giraffe's feet, slowly rising to capture the giraffe's body, and finally stops at the giraffe's face, showing giraffe's facial expression in close-up."

Limit the length of the camera movement to a reasonable 5-6 seconds: avoid overly complex shots that can lead to generation failures

The following example shows the result of adding [precise camera movement] and [detailed changes in the frame] in the prompt.

Example:** The green highlight** is the output responding to the change of the prompt

Generated Video 1

Generated Video 2

Prompt 1

The camera begins with an overhead shot of a mysterious forest clearing, with a female warrior in armor standing alone in the center of the clearing. Her armor shimmers with a cold silver light in the moonlight, and the trees around her cast dappled shadows. The camera begins to slowly descend and hover around her, as the perspective changes, the longsword in her hand can be seen trembling slightly, with undried blood remaining on the blade, the camera continues to wrap around, sliding from side to side, gradually drawing in closer to capture the side of her determined face. Her slightly clenched lower lip could be seen, and a few loose strands of hair on her face were soaked in sweat and plastered to her forehead. As the camera circled around to her front, the image became tighter, the focus gradually coming to rest on her face. By now, the forest in the background has blurred into a dark shadow. The camera eventually stops at a close-up of her front, focusing on her sharp eyes. The moonlight hits her face, outlining every detail of her face.

Prompt 2

The camera looks down on a moonlit forest clearing. The camera rotates around a female warrior in silver armor holding a longsword, after which the camera drops down to capture the side of the female warrior, who has blood on her longsword. The camera continues to circle, sliding past the side and gradually

pulling in closer to capture the side of the warrior woman's face, after which the camera wraps around to her front, with the focus gradually settling on a close-up of her face.

Generated Video 3

Prompt 3

The camera wraps around to capture a female warrior in silver armor standing in the jungle at night.

In addition to precise camera movement, Hailuo AI is able to accurately recognize aesthetics and ambience descriptions in the prompt and generate beautiful graphics and dynamic designs accordingly.

The following example shows the results of adding [Aesthetic and Atmospheric Descriptions] to the prompt, with an increase in output quality.

Example 1:** The highlighted portions of the Prompt are additional aesthetic descriptions.

Generated Video 4

Generated Video 5

Prompt 4

On the streets of a modern city, the camera first follows a motorcycle as it speeds between skyscrapers. The wheels leave deep imprints on the wet ground as the surrounding buildings sway in the ever-moving wind. The camera pans quickly right to catch a police car following close behind, its lights flashing, reflecting the neon lights and traffic on the street in brilliant colors. The bustling cityscape blurs into a flow of light in the background, and the movements of the motorcycle and police car seem unusually hurried and tense.

Prompt 5

In the modern city, the camera first follows a motorcycle as it speeds through the buildings, then pans quickly right to capture the police car in hot pursuit behind it, with the busy streets, traffic and lights as a blurred backdrop.

Generated Video 6

Prompt 6

The motorcycle sped through the city with police cars in hot pursuit.

Example 2:** The same basic prompt can be rendered very differently with different aesthetic ambience descriptions.

Generated Video 7

Generated Video 8

Prompt 7

The camera captures a woman sitting in a cafe, the woman looks up in thought, the camera slowly moves to capture the street outside the window, the image is gray in tone, the colors are undersaturated and the atmosphere is somber.

Prompt 8

The camera shoots a woman sitting in a cafe, the woman looks up at the window, the camera slowly moves to capture the street outside the window, the image shows warm tones, rich colors and a relaxing atmosphere.

Generated Video 9

Generated Video 10

Prompt 9

A male in a trench coat walks down a street in a modern city, the camera follows the male's movement from behind, the image has a grayish tone and a somber atmosphere.

Prompt 10

A male in a trench coat walks down a street in a modern city, and the camera follows the male's movements from behind, with a brilliant sci-fi atmosphere and colorful tones.

Generated Video 11

Generated Video 12

Prompt 11

The camera captures a seaside wilderness with grass blowing in the sea breeze and a vast ocean in the distance, with somber tones and low color saturation.

Prompt 12

The camera shoots a seaside wilderness, the grass is blown by the sea breeze, and in the distance is a vast ocean, the color tone of the picture is warm, and the picture has a warm atmosphere.

Generated Video 13

Generated Video 14

Prompt 13

Close up shot of a boy reading a book at a desk in a room with strong contrasts of light and dark, gray tones, and a tense, depressing atmosphere.

Prompt 14

Close up shot of a boy reading a book at a desk in a room with warm colors and a cozy, warm atmosphere.

Generated Video 15

Generated Video 16

Prompt 15

Two women driving a blue Cuban convertible sedan on a seaside highway, the camera looks down on the car from the air, following the car's movement, the image has a somber tone, very low color saturation, and a depressing atmosphere.

Prompt 16

Two women driving a blue Cuban convertible sedan on a seaside highway, the camera looks down on the car from the air, following the car's movement, the image has bright, rich tones and a relaxed atmosphere.

Hailuo AI's precise lens control and aesthetic understanding can help users generate more responsive outputs

Generated Video 17

Generated Video 18

Original Prompt 17

A man on the rooftop of a tall building at night looks down at the city.

Optimized Prompt 18

Under the deep blue night sky, the neon lights of the city are reflected in puddles of freshly rained water on the rooftops of tall buildings. The camera slowly rises from the reflection in the water, through the rain, to reveal a black figure standing on the edge of a rooftop. The camera then continues to rise to show a sprawling modern cityscape. Finally, the camera freezes and the lone figure standing on the edge stares at the city, the interplay of light and shadow creating a cold atmosphere.

Generated Video 19

Generated Video 20

Original Prompt 19

The camera goes through the glass and focuses on a pensive lady in the window.

Optimized Prompt 20

The camera shoots a modern glass-look high-rise building through a clear glass, the building shimmering with reflected light in the sunlight. The camera then closes in on a window, where reflections on the glass show a bustling cityscape outside. The camera slowly pulls back and through the glass, eventually focusing on a woman in a suit inside the window. She sits in a minimally decorated office, staring intently at a computer screen. Her face appears pensive and introspective in the soft light, contrasting with the hustle and bustle outside the window. The office is decorated in a simple and modern way, emphasizing her focus and introspection.

Generated Video 21

Generated Video 22

Original Prompt 21

In a forest covered in dark clouds, the young boy sprints through the mud with his sword in his hand, his sleeves swinging in the wind. The camera pushes forward and lightning illuminates his resolute face.

Optimized Prompt 22

In a forest shrouded by dark clouds, which were low in the sky, the young boy held a sharp sword with a cold glint in his hand, and walked quickly on the muddy ground with firm footsteps. His sleeves hunt in the cold wind, as if adding a touch of vigor to his determination at this moment. The camera slowly advances from his side, and the trees in the background sway in the wind, making it look particularly eerie. Suddenly, a bolt of lightning breaks through the gloomy sky, instantly illuminating the young boy's determined and resolute face as he looks ahead with determination in his eyes.

Generated Video 23

Generated Video 24

Original Prompt 23

When the boy plays the game, the character on the screen suddenly comes to life and interacts with the boy, blending reality with the virtual.

Optimized Prompt 24

In a dimly lit room, a young boy is captured in close-up, his eyes glued to the glowing screen of a game console. The camera is slightly above the horizon, focusing on his concentrated expression as his fingers deftly manipulate the controller. Suddenly, the mood changes as the figure of a game character suddenly appears and breaks through the screen. The camera has a shallow depth of field and pans smoothly between the boy and the now three-dimensional figure, blurring the edges of the console and the wall behind it, which is plastered with posters for various games. The boy interacts with the character in surprise, his movements reflecting the excitement and disbelief of the moment. The characters are vivid and detailed, jumping and dancing around the room and playing with the boy. The camera captures the fantastical scene with dynamic angles and fluid movements, enhancing the sense of wonder and movement through a skillful combination of slow motion and fast editing. Subtle changes in lighting enhance the realism of the characters' movements, with shadows and highlights interplaying on the boy and the characters.

The image used in image-to-video generation appears as the first frame of the video, defining the subject's appearance and the basic aesthetic style of the video. Therefore, compared to text-to-video prompts, image-to-video prompts can appropriately reduce the amount of information.

Basic Prompt Formula=Main Subject in the first frame + Motion/Change

Main Subject in the first frame : The objects in the scene, including key details and environmental information, are precisely recognized by Hailuo AI, which can identify characters, objects, and other visual elements in the image and generate videos based on the prompt requirements.

Motion/ Change : Motion is the description of the main object's moving status in the video, such as stillness, movement; the transformation of the environment and space, and so on.

Specific Cases : (The image was generated using MidJourney.)

Generated Video 25

Generated Video 26

Image 25

**Prompt 25**

In the scene, a small dog emits a blue light from its eyes. The messy clothes in front of the dog also begin to glow blue and slowly float into the air, folding neatly on their own before landing in front of the dog. Afterward, the blue light in the dog's eyes fades away.

Image 26**Prompt 26**

A blue furry creature in the scene is constantly stirring a steaming pot of soup. Then, the blue creature blows on the pot, causing the soup in the bowl in front of it to freeze into a block of ice.