Multi-Object Sketch Animation by Scene Decomposition and Motion Planning

Supplementary Material

In this supplementary material, we report:
- More results on the created sketches (Fig. 9).
- Evaluation on freehand sketches (Sec.S1).
- Evaluation on Single-object sketches (Sec.S2).
- GPT-4 instructions in MoSketch (Sec.S3).
- The created sketches with instructions (Figs. 12 to 17).

S1.Evaluation on Freehand Sketches

To evaluate MoSketch's external validity, we selected a diverse set of 30 multi-object freehand drawings from FS-COCO [1], a human-drawn vector scene sketch dataset. Fig. 10 shows the animation results of MoSketch, and Tab. 4 provides the quantitative comparison with Live-Sketch and FlipSketch. MoSketch achieves vivid multi-object sketch animation on freehand drawings, notably improving the drawing imperfections of freehand sketches.

S2.Evaluation on Single-object Sketches

We evaluate MoSketch on single-object sketches from Live-Sketch, as shown in Fig. 11. Tab. 5 shows that MoSketch is less effective in single-object scenario.

S3.GPT-4 Instructions in MoSketch

We employ GPT-4 for scene decomposition and motion planning, and the relative instructions are listed as follows.

S3.1.Instructions for Scene Decomposition

> You are an intelligent Scene Decomposition Assistant for Multi-object Sketch Animation. I will give you a sketch and a complex instruction to animate it. We want to use a divide-and-conquer method. You should decompose a complex instruction for Multi-object Sketch Animation to no more than five simple ones, and each instruction involves no more than seven objects, one or two are preferring. Objects should be used for grounding in next process, so too small and abstract objects could be ignored. Reasonable imagination is fine.
> **Input:** a sketch, a complex instruction
> **Output:** objects, simple instructions: [(instruction1, object_set1),...]

Method	Text-to-Video Alignment	Sketch-to-Video Alignment	Motion Smoothing	Dynamic Degree
FlipSketch	0.181	0.757	0.823	-
Live-Sketch	0.173	0.732	0.827	0.500
MoSketch	**0.197**	**0.927**	**0.940**	**0.633**

Table 4. **Evaluation on 30 freehand drawings from FS-COCO**.

Method	Text-to-Video Alignment	Sketch-to-Video Alignment	Motion Smoothing	Dynamic Degree
FlipSketch	0.211	**0.936**	0.858	**1.000**
Live-Sketch	**0.217**	0.884	0.827	0.392
MoSketch	0.209	0.914	**0.968**	0.571

Table 5. **Performance comparison in a single-object scenario**.

S3.2.Instructions for Motion Planning

> You are an intelligent Motion Planning Assistant for Multi-object Sketch Animation. A sketch, an instruction to animate it and each object 's bounding box are provided. You should predict the bounding box of each object in 16 frames according to the reasonable inference. Note that the movement should follow the laws of physics such as inertia and gravity. If the sketch is in the first person, then the rule that objects far away are small and objects near are large should also be considered. Don't forget considering the interaction or relationship of objects. The image size is 256 * 256, and objects should appear in the image as far as possible. Show me the reasoning process before planning.
> **Input:** a sketch, a complex instruction, objects: [(object1,[x1, y1,w1,h1]), ...]
> **Output:** the reasoning process, motion plan: [(object1: [[x1,y1,w1,h1],...,[x16,y16,w16,h16]]),...]

References

[1] Pinaki Nath Chowdhury, Aneeshan Sain, Ayan Kumar Bhunia, Tao Xiang, Yulia Gryaditskaya, and Yi-Zhe Song. Fs-coco: Towards understanding of freehand sketches of common objects in context. In *ECCV*, pages 253–270, 2022. 1

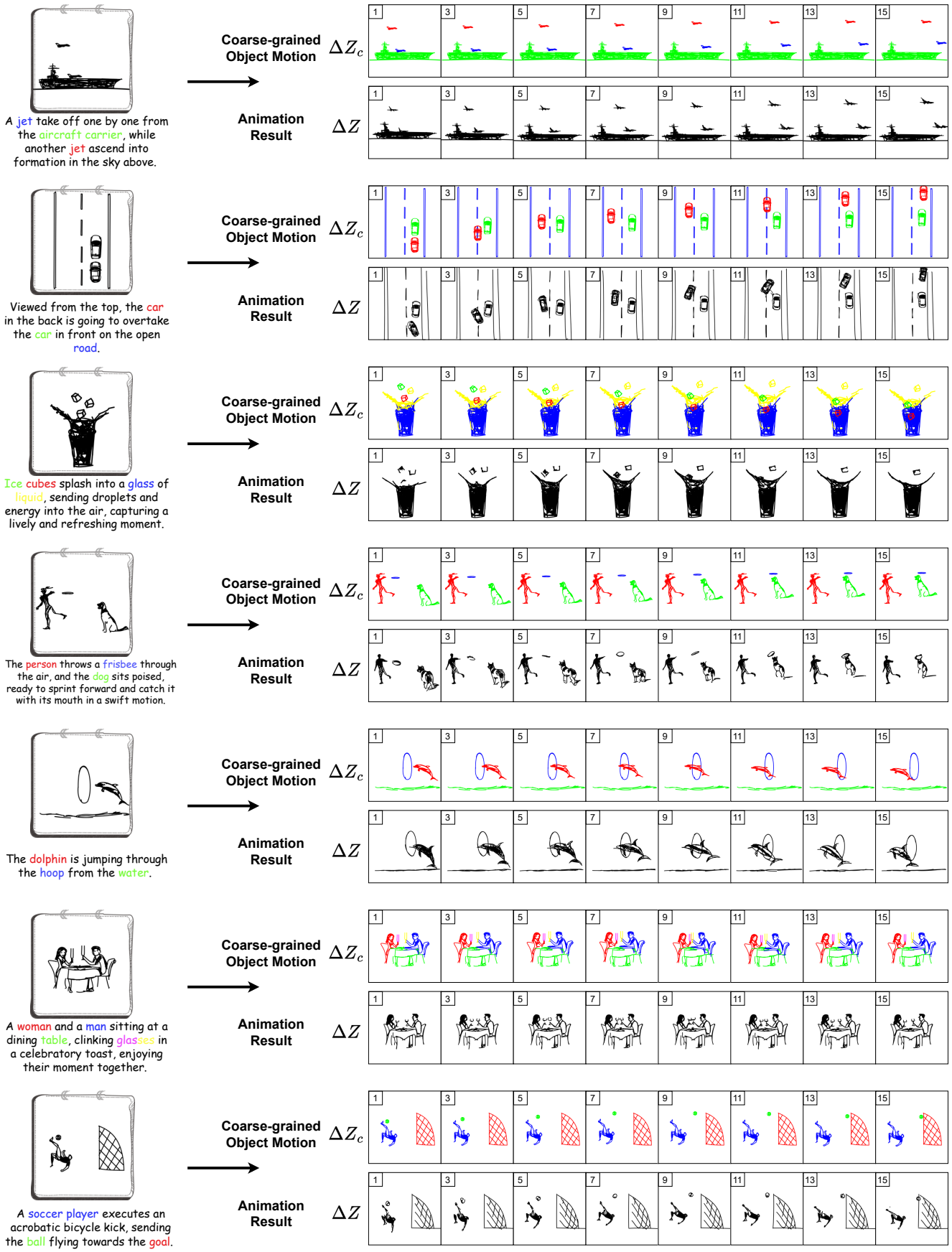


Figure 9. **More results on the created sketches**.

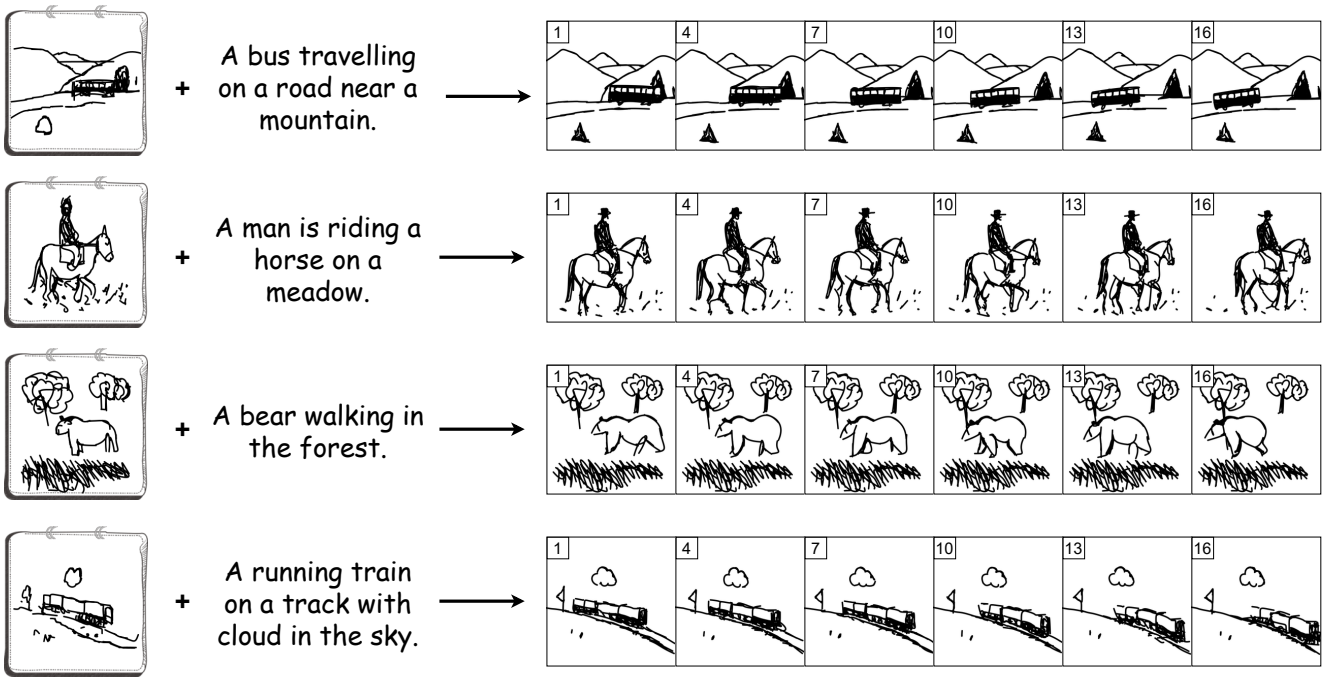


Figure 10. **Animating FS-COCO samples by MoSketch**.

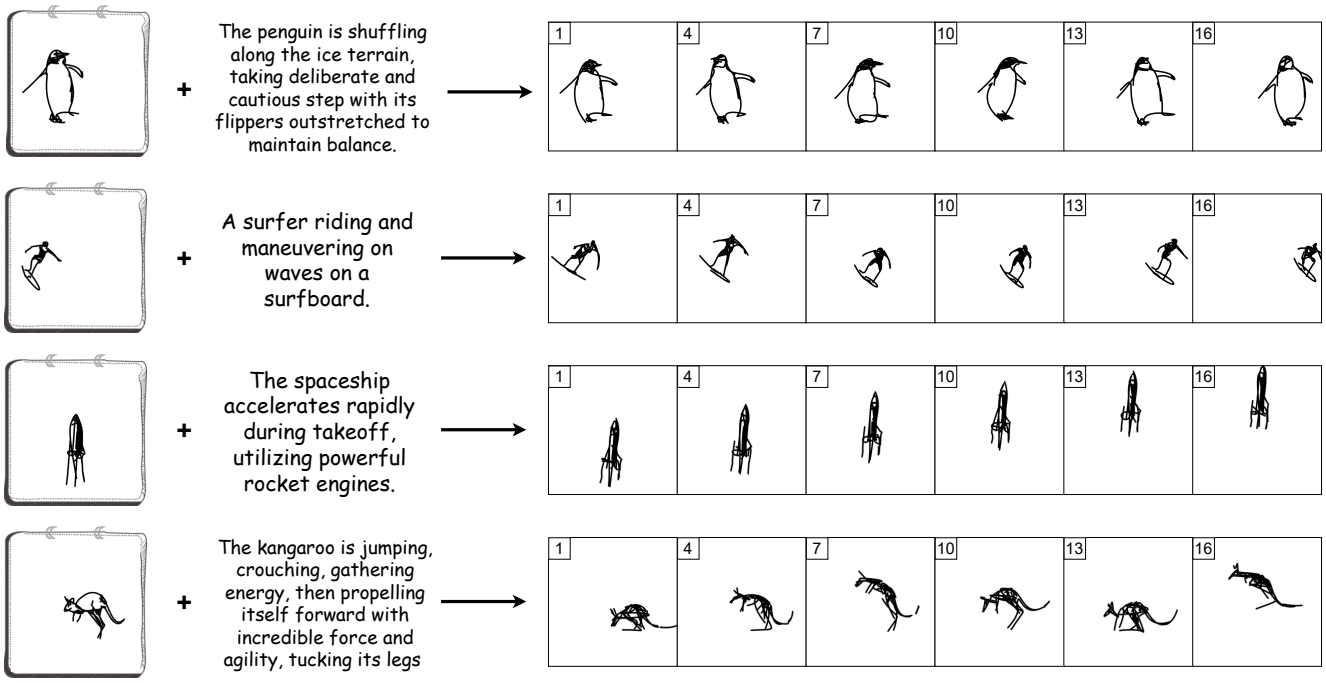
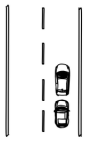


Figure 11. **Animating results on single-object sketches**.

A jet take off one by one from the aircraft carrier, while another jet ascend into formation in the sky above.

Viewed from the front, an airplane lifts off from the runway, its wheels leaving the ground as it ascends into the sky.

A shell bursts from the cannon, leaving a trail of smoke as it hurtles through the air.



Viewed from the top, the car in the back is going to overtake the car in front on the open road.

Viewed from the back, a car navigates a curving road at the base of majestic mountains.

The road curves ahead as a jeep viewed from the back, travels in one direction while a motorcycle approaches from the opposite side.

A lone car maneuvers with steep cliffs and dense vegetation enclosing the winding road.

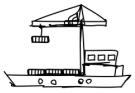
Two cyclists maneuver a curvy road, one leading while the other follows closely, capturing the thrill of the ride.

The motorcycle will jump over the oncoming car.

A person holding a spoonful of food close to their lips, ready to take a bite.

Figure 12. **The created sketches with text instructions for the "object" class**.

An excavator digging and scooping soil, forming a pile nearby.



A large crane lowers a container onto a heavily loaded cargo ship.

Two people rappel down a rope from a hovering helicopter.

Ice cubes splash into a glass of liquid, sending droplets and energy into the air, capturing a lively and refreshing moment.

A bottle is gracefully pouring liquid into a glass, the steady stream creating ripples in the drink.

Rollercoaster cart at the peak of a rollercoaster drops, bracing for the adrenaline-fueled descent.

The satellite shifts position as the Earth rotates, adjusting its pose to gather information from different angles.

The space shuttle begins its ascent, tilting slightly as it gains altitude, leaving a trail of flames near the launch pad.

A tank is firing shells towards a distant rectangular target, with a burst of energy smoke.

A vintage steam locomotive with an attached carriage is traveling along the tracks, exuding a sense of historic charm and industrial innovation.

Figure 13. **The created sketches with text instructions for the "object" class**.

A stealthy cat crouches low, its eyes locked on a tiny mouse, ready to spring forward with calculated precision. The mouse is going to escape.

Two playful cats stand on their hind legs, batting at a ball suspended between them in mid-air, engaged in a lively and fun interaction.

A curious cat jumps up to a table, reaching toward a bowl of food on the table with curiosity.

The person throws a frisbee through the air, and the dog sits poised, ready to sprint forward and catch it with its mouth in a swift motion.

The dog prepares to enjoy a delightful meal near the table, using its paws to carefully handle the chopsticks.

The dog reaches up to the table, anticipating a chance to grab the food within its reach.

The dog races up the stairs in pursuit of the cat, eager to catch up to its swift companion.

The seal on a stage prepares to leap toward the ball. The ball is throw to the seal.

The seal seats on a stage, juggling the ball skillfully on its nose, preparing to toss it into the air for an impressive trick.

The dolphin is mid-leap, heading toward the ball, preparing to nudge it forward with its snout, creating a playful splash.

Figure 14. **The created sketches with text instructions for the "creature" class**.

The dolphin is jumping through the hoop from the water.

The eagle is in pursuit of the smaller bird, its wings stretched wide as it closes the distance in an intense aerial chase.

The larger fish opens its mouth wide as it chases the smaller fish, preparing to close the gap and capture its prey in an instant.

The giraffe is approaching the tree, preparing to eat from its high branches.

A goat grazes peacefully on the grass, its head lowered as it feeds on the vegetation.

Godzilla and the other monster face off, the rubble of the city forming a chaotic battlefield between them.

A horse-drawn carriage moves steadily forward, the driver urging the horses onward as the wheels turn across the terrain.

A person leads the horse gently by the reins, while the rider sits calmly, observing the surroundings as they move forward.

The horse gracefully soars over the obstacle, with the rider maintaining perfect balance and focus in mid-air on the horse.

The tiger leaps through the air, claws extended, as the fox darts away swiftly, trying to escape the predator's pursuit.

Figure 15. **The created sketches with text instructions for the "creature" class**.

The drummer pounds the cymbals, bringing the song to a crashing halt. The guitarist strums a final chord, while the keyboardist focused on the melody.

The player soars through the air with a basketball, arm extended for an electrifying slam dunk to the hoop.

A climber dangles mid-air, navigating the rugged face of a towering mountain.

A climber ascends a rock wall, dynamically reaching for the next hold with determination.

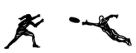
A woman and a man sitting at a dining table, clinking glasses in a celebratory toast, enjoying their moment together.

A man gently feeds a woman across a dining table, offering her a spoonful of food with a caring expression.

A man pulls himself up on a horizontal bar, while a sporter lifts a barbell overhead in a deep squat position, showcasing strength and stability.

The scene depicts two individuals engaging in exercises: a man is performing dips on parallel bars, and a woamn is holding a plank position on the floor, demonstrating core strength.

The boxer throws a powerful punch at the heavy sandbag. The sandbags hoisted with brackets is swing.



A woman throws a frisbee with force while a man dives to catch it mid-air.

Figure 16. **The created sketches with text instructions for the "human" class**.

A soccer player executes an acrobatic bicycle kick, sending the ball flying towards the goal.

A soccer player skillfully dribbles the ball toward the goal as the goalkeeper prepares to defend.

Three athletes in action during a hurdle race: one running towards the hurdle, another mid-air clearing a hurdle, and the third already sprinting ahead, showcasing their athleticism and agility.

Three children enjoying a jump rope activity: two on either end turning the rope, while one jumps enthusiastically in the middle.

A worker climbing an inclined ladder, wearing protective gear and holding onto the ladder rungs securely, moving steadily upwards.

A bride and groom stand facing each other, holding hands. The bride's flowing gown and veil trail elegantly behind her, while the groom leans slightly forward, creating an intimate and heartfelt moment.

A person carefully stacking boxes, aligning them into a neat pile with focused attention.

A person sits alone, holding a fork poised over their food, savoring a moment of quiet dining.

Superman flies swiftly through the air, arm extended, racing to catch the falling man just moments before he hits the ground.

The workers in unison: one shovels debris, another operates a power tool, and a third transports materials with a wheelbarrow.

Figure 17. **The created sketches with text instructions for the "human" class**.