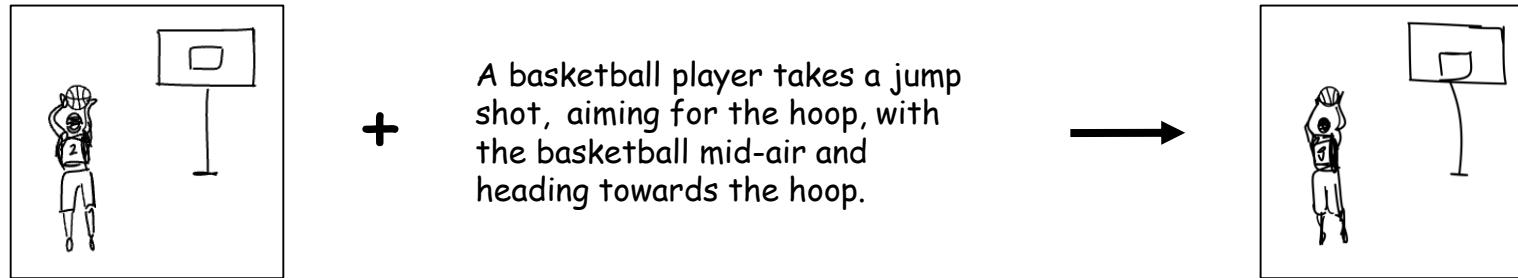


# Multi-Object Sketch Animation by Scene Decomposition and Motion Planning

Jingyu Liu, Zijie Xin, Yuhang Fu, Bangxiang Lan, Ruixiang Zhao, Xirong Li<sup>†</sup>



Project Page

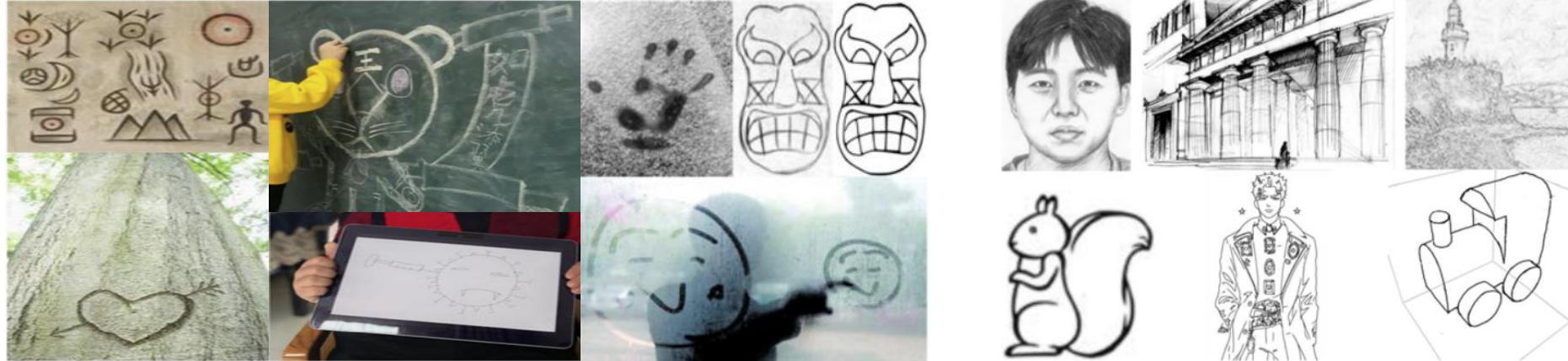
Speaker: Jingyu Liu (刘靖宇)

Renmin University of China, AIMC-Lab

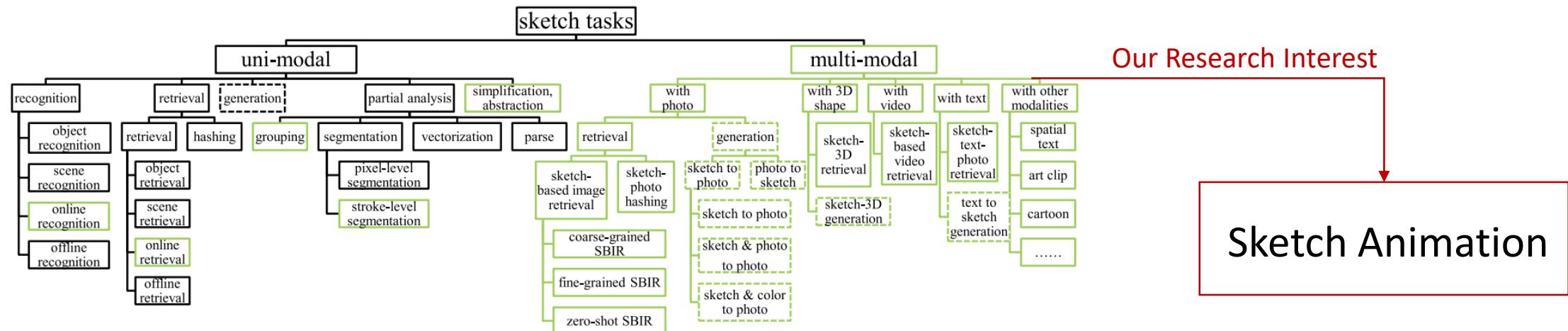
[liujingyu2023@ruc.edu.cn](mailto:liujingyu2023@ruc.edu.cn)

# Background

## ➤ Sketches are Everywhere in Daily Life



## ➤ Sketch-based Tasks are Numerous

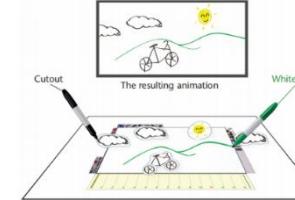


# Background

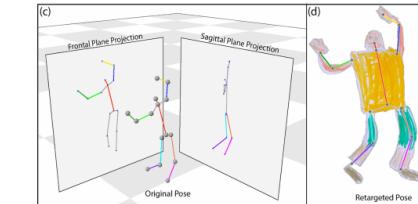


## ➤ A Brief Review of Sketch Animation Development

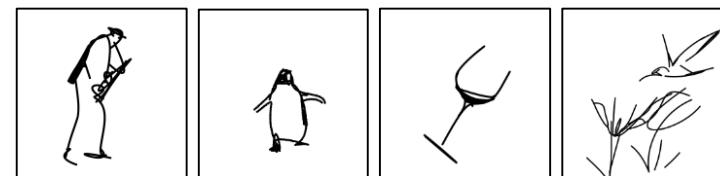
- Early Works
  - Focus on developing **tools** for human-driven creation
  - Need massive manual intervention



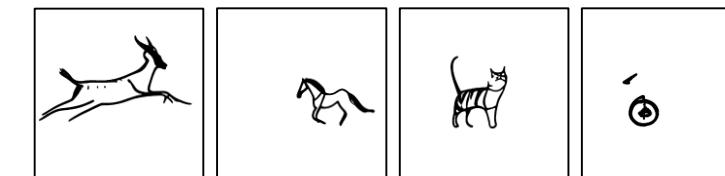
- Later Works
  - Focus on incorporating **additional inputs** (video, skeleton)
  - Reduce manual intervention



- Latest Works (**texts and sketches as inputs**): Live-Sketch, FlipSketch



Live-Sketch (**Vector Sketch**)



FlipSketch (**Raster Sketch**)

Our Goal

Poor Performance in  
**Multi-Object** Scenarios

[1] Sohn, et al. Sketch-n-Stretch: sketching animations using cutouts. CG&A 2012

[2] Yu, et al. VideoDoodles: Hand-Drawn Animations on Videos with Scene-Aware Canvases. TOG 2023

[3] Smith, et al. A Method for Animating Children's Drawings of the Human Figure. TOG 2023

[4] Gal, et al. Breathing Life Into Sketches Using Text-to-Video Priors. CVPR 2024

[5] Bandyopadhyay, et al. FlipSketch: Flipping Static Drawings to Text-Guided Sketch Animations. CVPR 2025

## ➤ Two Challenges in Multi-Object Scenarios

### Object-aware Motion Modeling



- Motion Planning for All Objects (LLM)
- Object-aware Motion Network

### Complex Motion Optimization



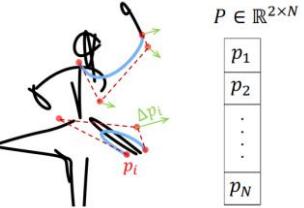
- Decompose Complex Motions into Simple Ones (LLM)
- Optimize Them Separately (SDS)

**Divide and Conquer !**

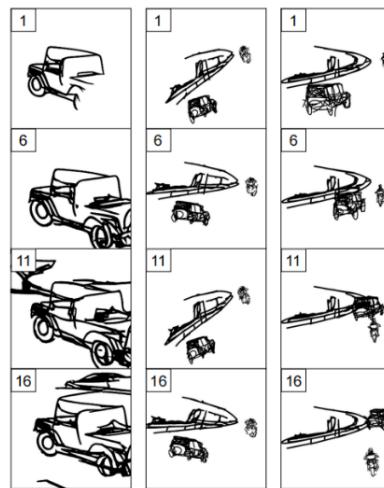
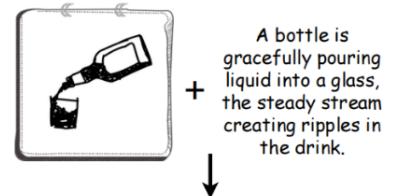
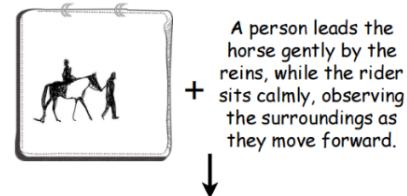
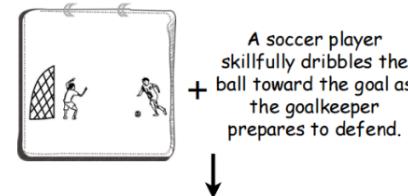
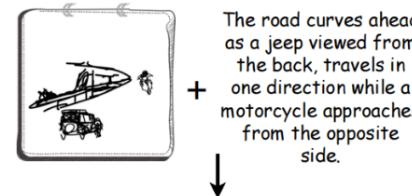
# Motivation



## ➤ Vector Sketch VS Raster Sketch



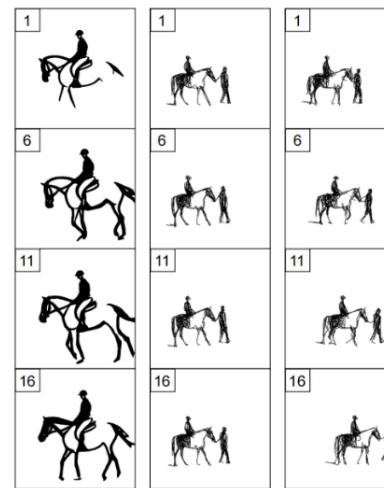
- Vector Sketch (composed of strokes, where each stroke is a cubic Bezier curve)
  - Resolution-Free
  - Easy to Edit
  - More Compact
  - More Stable



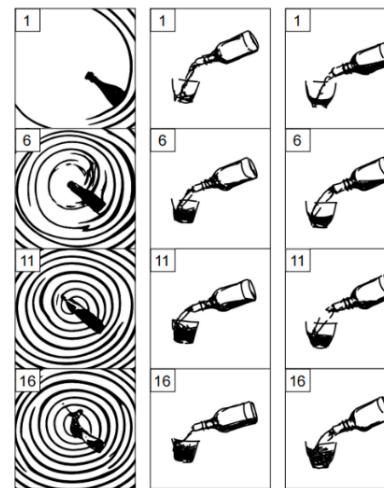
FlipSketch   Live-Sketch   MoSketch



FlipSketch   Live-Sketch   MoSketch



FlipSketch   Live-Sketch   MoSketch



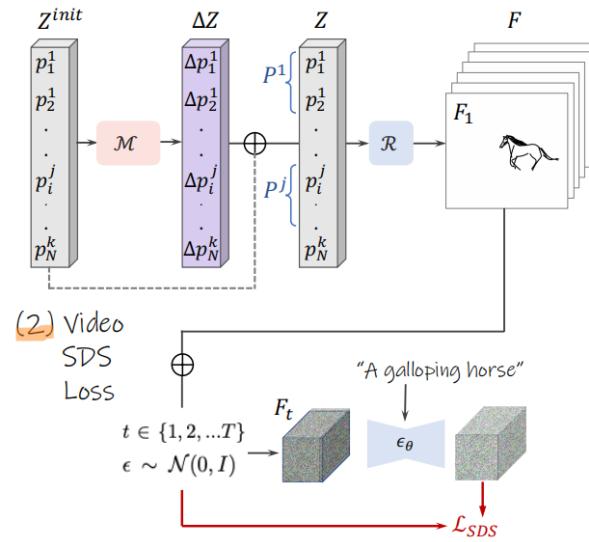
FlipSketch   Live-Sketch   MoSketch

● Raster Sketch

● Vector Sketch

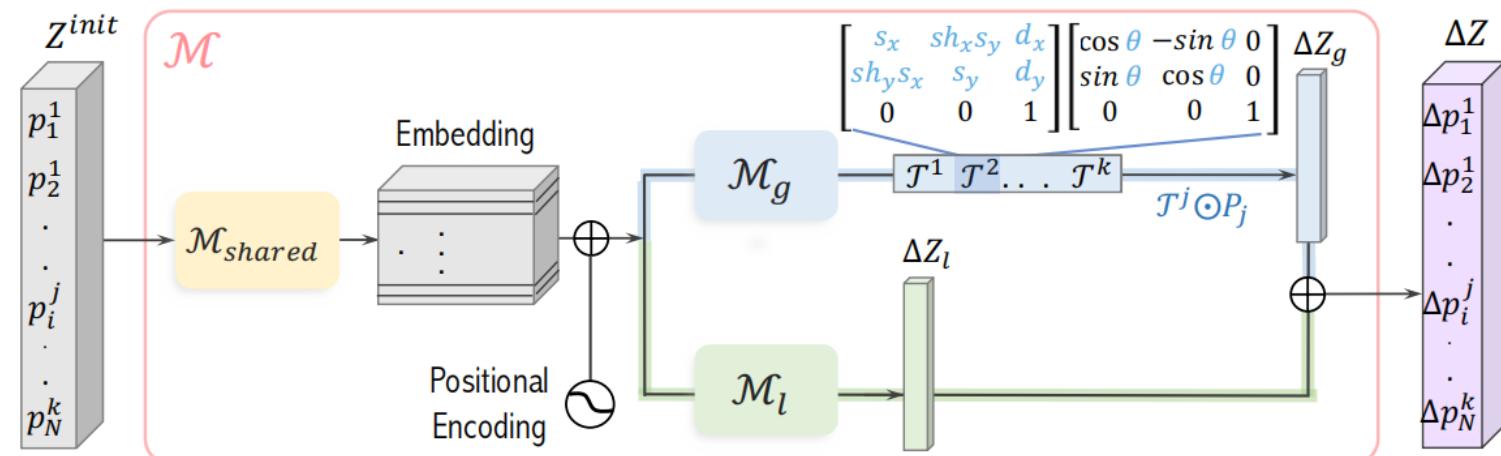
## ➤ Live-Sketch in a Nutshell (Training-Data Free)

## (1) Displacement and Rendering



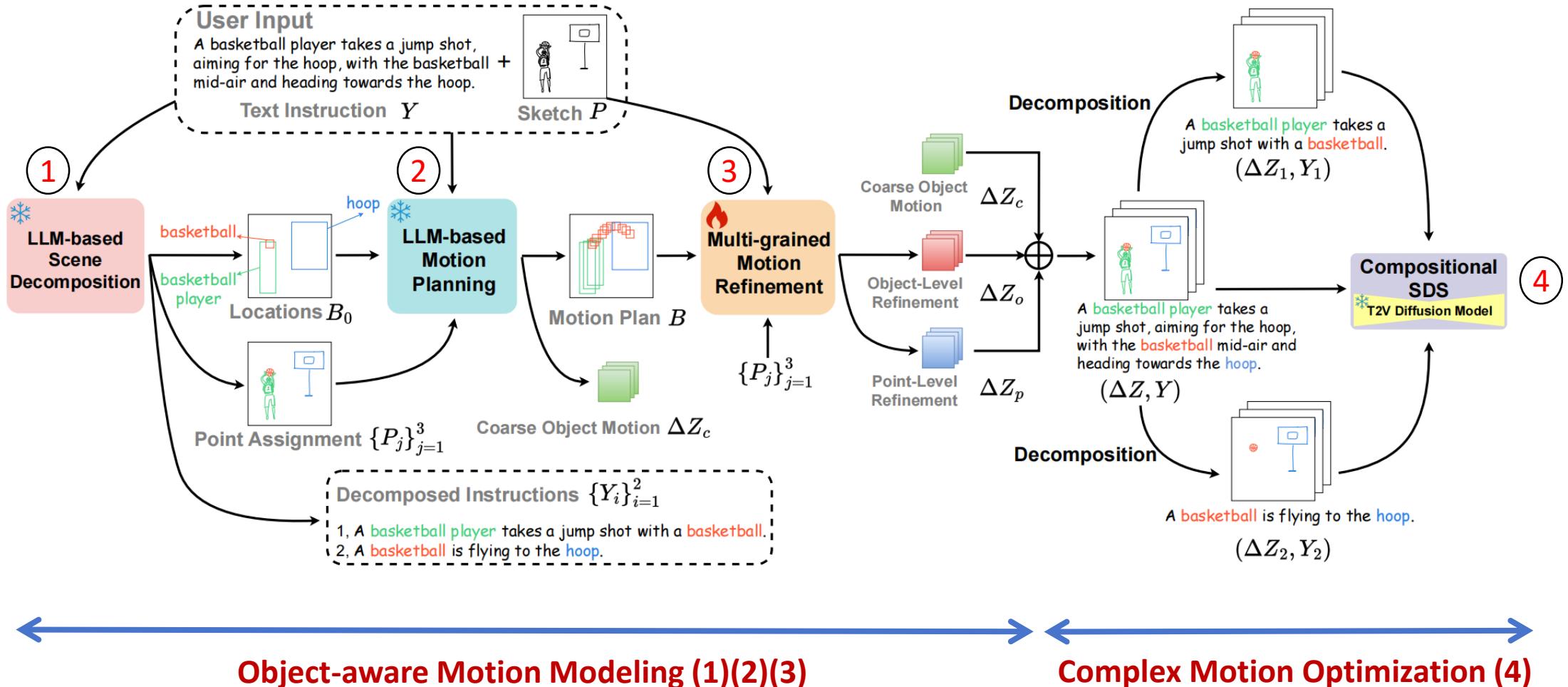
$$\nabla_{\phi} \mathcal{L}_{SDS} = \left[ w(t)(\epsilon_{\theta}(x_t, t, y) - \epsilon) \frac{\partial \epsilon}{\partial \phi} \right]$$

# Global (scaling, shear, rotation, translation)



## Local (Internal Motion Modeling)

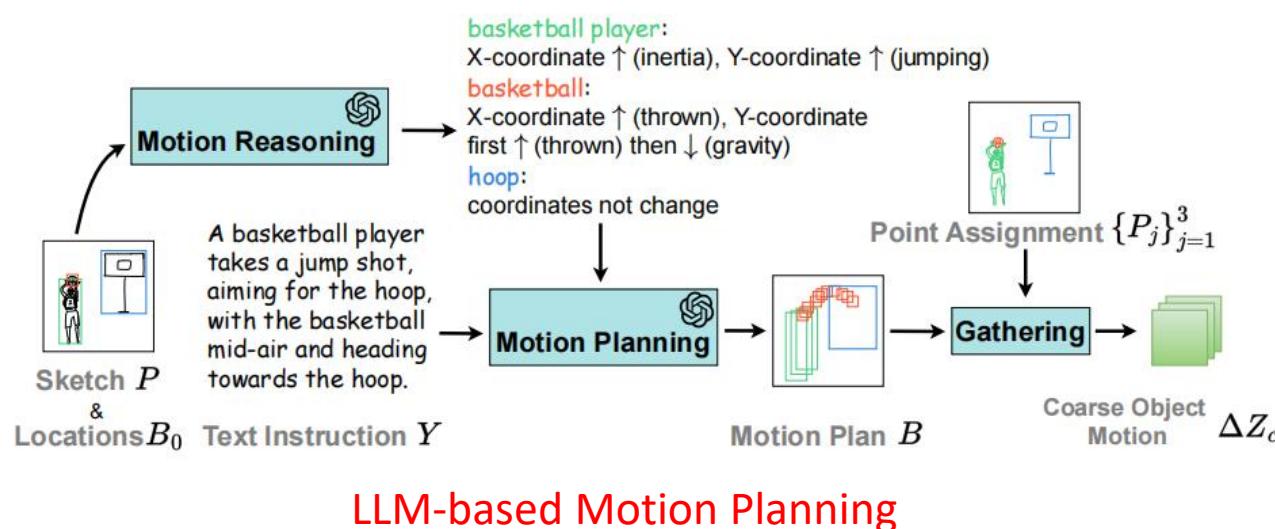
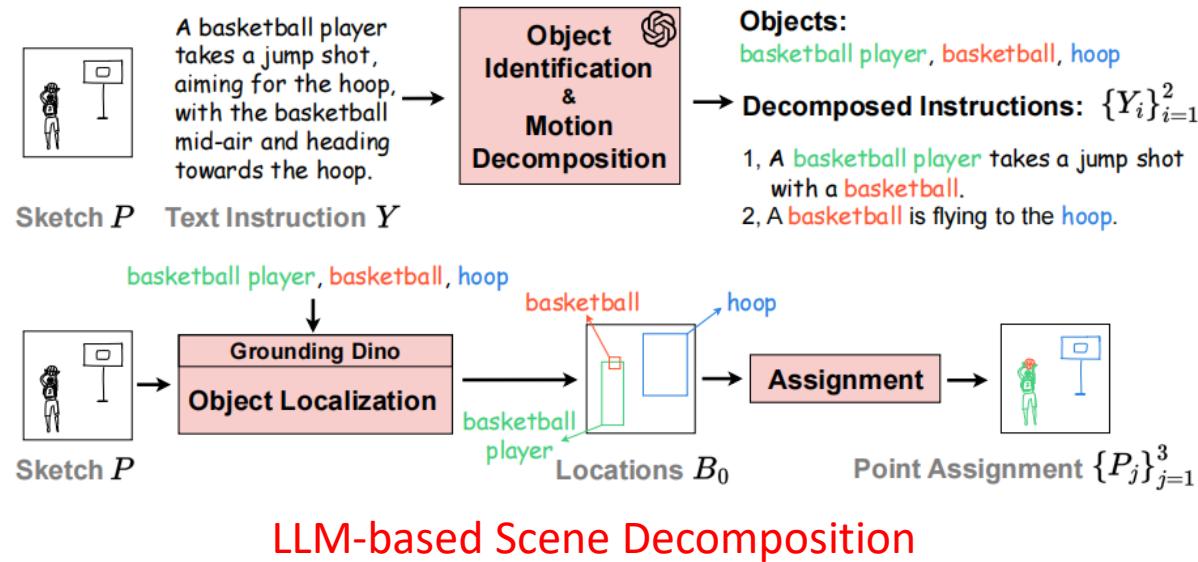
## ➤ Our Solution——MoSketch (Training-Data Free)



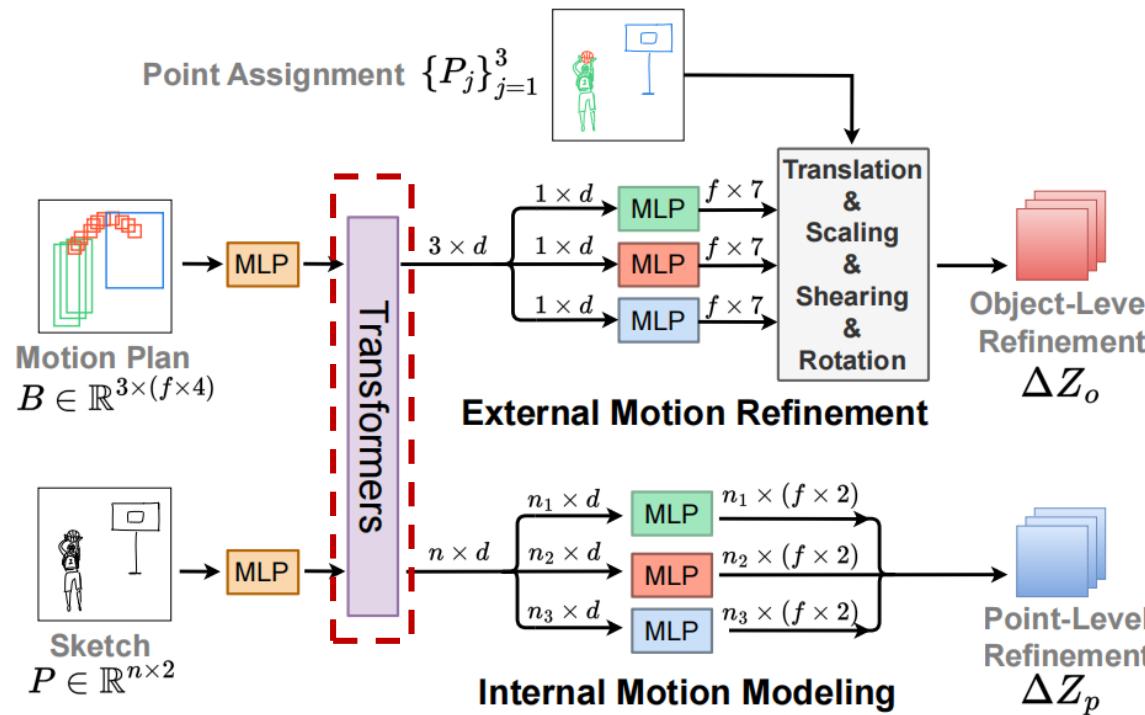
# Method



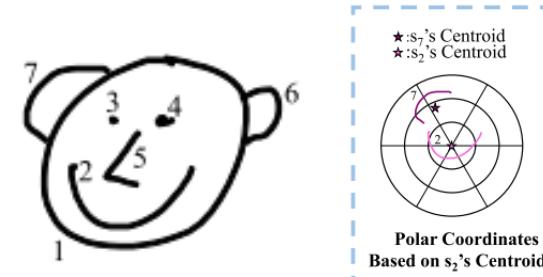
## ➤ LLM's Power



## ➤ Multi-grained Motion Refinement



Positional Encodings Specially  
Designed for Vector Sketches



- eye
- face
- mouth
- hair

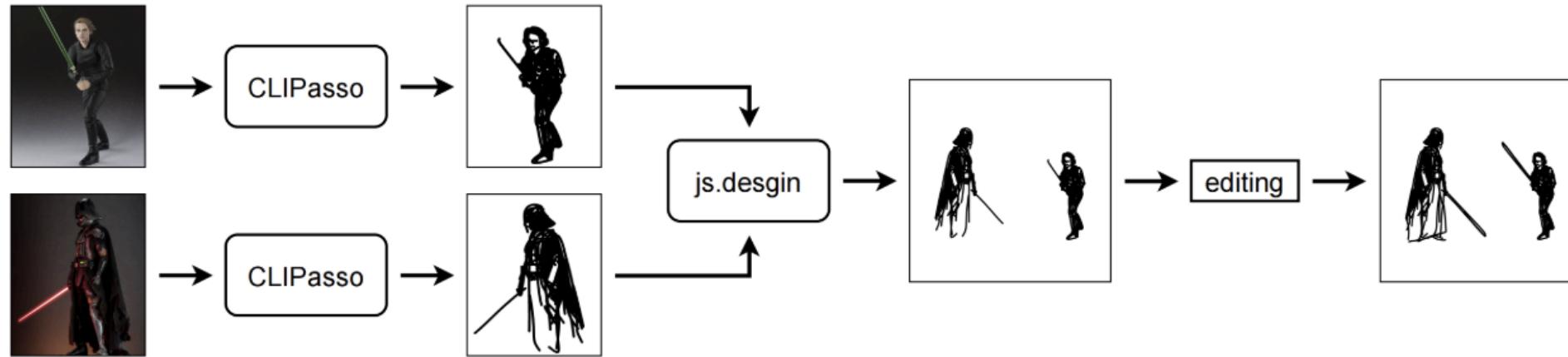


Sketch Segmentation with the  
Positional Encodings

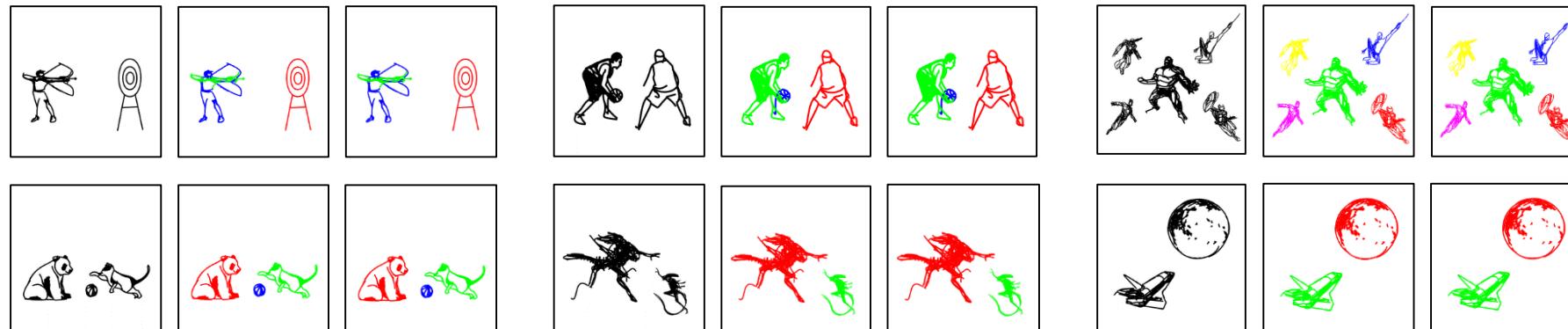
# Data Creation for Evaluation



➤ CLIPasso + SVG Designer → 60 Scene Sketches (“human”, “animal”, “object”)



➤ We provide more 500 sketches !



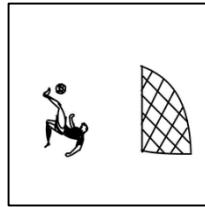
# Experiments



## ➤ Comparison

Method	Text-to-Video Alignment	Sketch-to-Video Alignment	Motion Smoothing	Dynamic Degree
CogVideoX	0.141	0.610	0.747	-
DynamiCrafter	0.184	0.771	0.868	-
FlipSketch	0.199	0.704	0.839	-
Live-Sketch	0.207	0.897	0.956	0.266
<b>MoSketch</b>	<b>0.218</b>	<b>0.914</b>	<b>0.977</b>	<b>0.283</b>

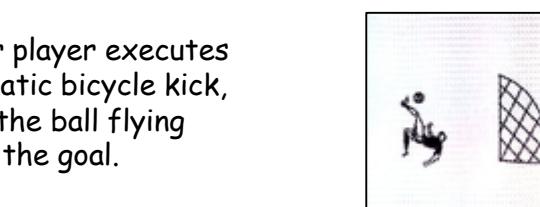
Sketches



+

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

Text



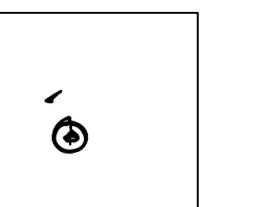
CogVideoX



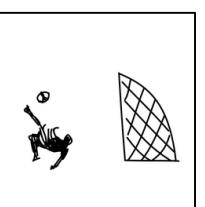
Dynamicrafter



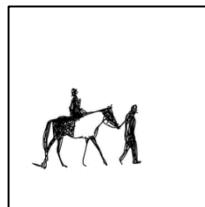
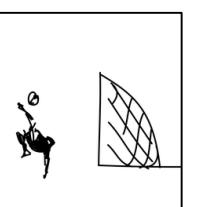
FlipSketch



Live-Sketch

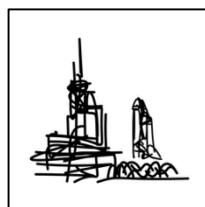
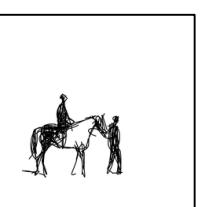
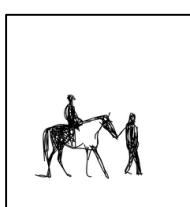
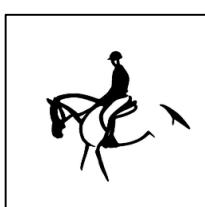


**MoSketch**



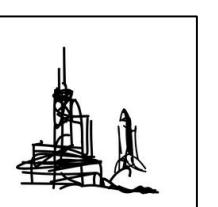
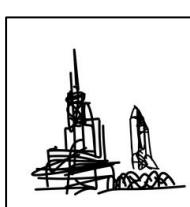
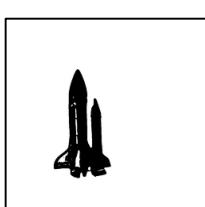
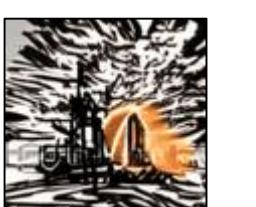
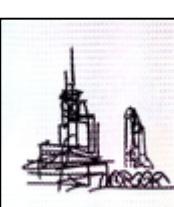
+

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



+

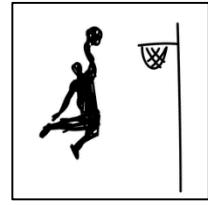
The space shuttle begins its ascent, tilting slightly as it gains altitude, leaving a trail of flames near the launch pad.



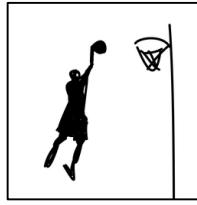
# Experiments



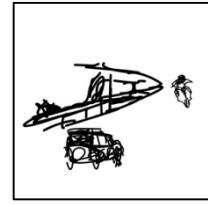
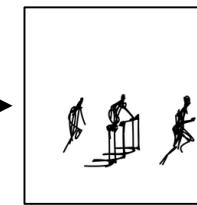
## ➤ Gallery



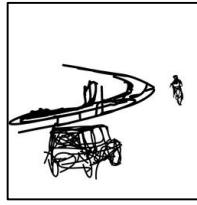
+ The player soars through the air with a basketball, arm extended for an electrifying slam dunk to a hoop.



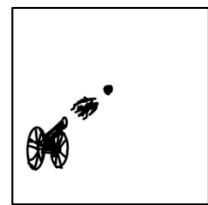
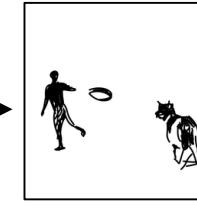
+ Three athletes in action during a hurdle race: one running towards the hurdle, another mid-air clearing a hurdle, and the third already sprinting ahead, showcasing their athleticism and agility.



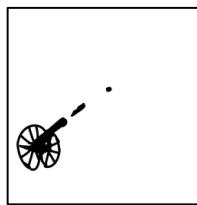
+ The road curves ahead as a jeep viewed from the back, travels in one direction while a motorcycle approaches from the opposite side.



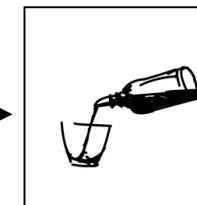
+ The person throws a frisbee through the air, and the dog sits poised, ready to sprint forward and catch it with its mouth in a swift motion.



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



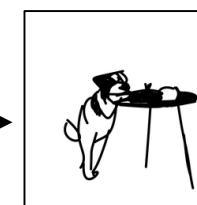
+ A bottle is gracefully pouring liquid into a glass, the steady stream creating ripples in the drink.



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



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

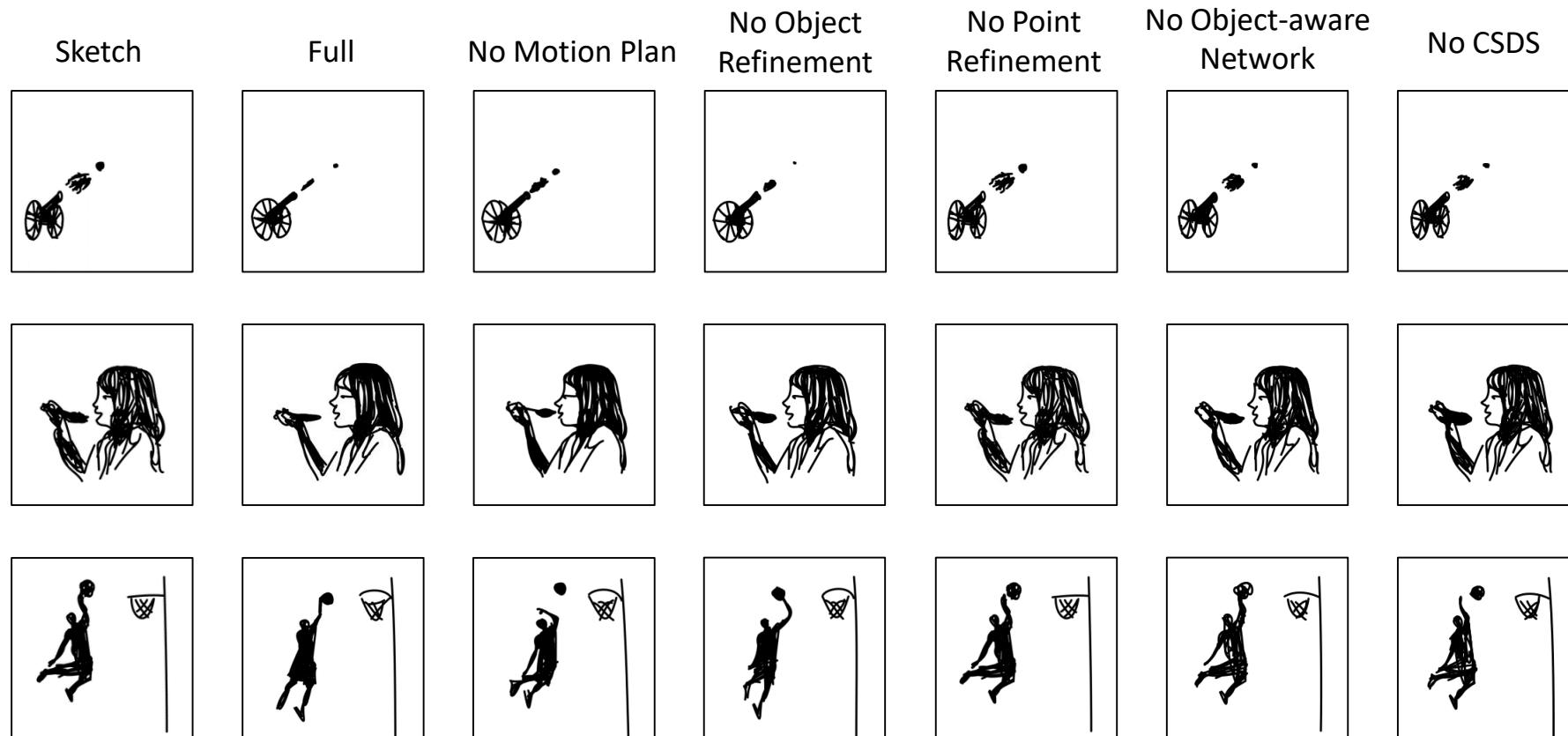


# Experiments



## ➤ Ablation Study

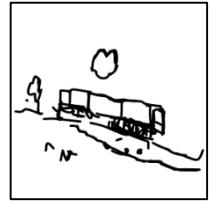
#	setup	Text-to-Video Alignment	Sketch-to-Video Alignment	Motion Smoothing	Dynamic Degree
0	Full	<b>0.218</b>	0.914	<b>0.977</b>	<b>0.283</b>
1	w/o $\Delta Z_c$	0.212	0.955	0.959	0.083
2	w/o $\Delta Z_o$	0.212	0.909	0.964	0.266
3	w/o $\Delta Z_p$	0.203	<b>0.971</b>	0.971	0.200
4	w/o Object-aware	0.205	0.932	0.968	0.266
5	w/o CSDS	0.207	0.911	0.966	0.267



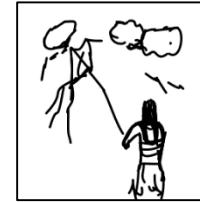
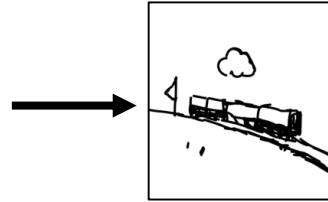
# Experiments



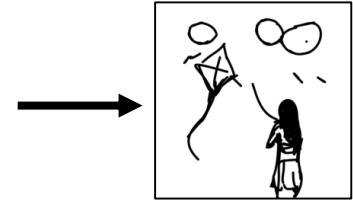
## ➤ Performance on Freehand Scene Sketches (FSCOCO Dataset)



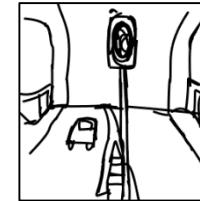
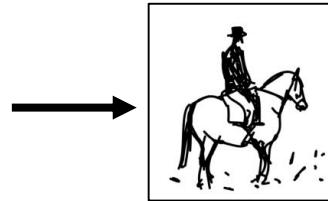
+ A running train on a track with cloud in the sky.



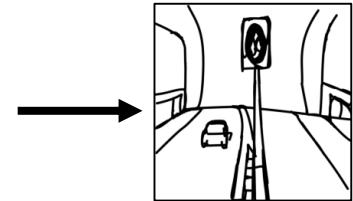
+ A girl flying the kite high.



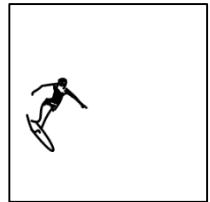
+ A man is riding a horse on a meadow.



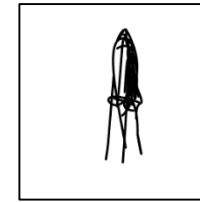
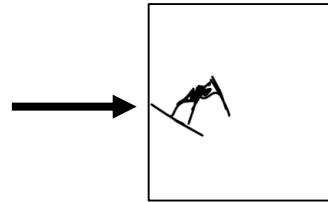
+ A vehicle is driving on the road.



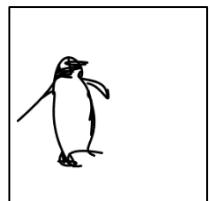
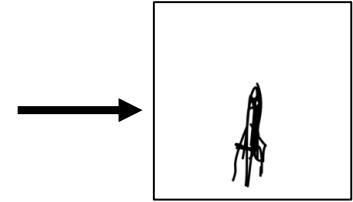
## ➤ Performance on Single-Object Sketches (Live-Sketch)



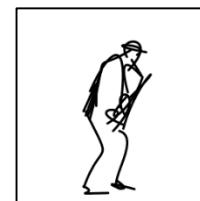
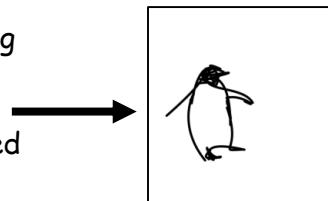
+ A surfer riding and maneuvering on waves on a surfboard.



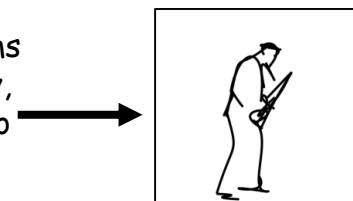
+ The spaceship accelerates rapidly during takeoff, utilizing powerful rocket engines.



+ The penguin is shuffling along the ice terrain, taking deliberate and cautious step with its flippers outstretched to maintain balance.



+ The jazz saxophonist performs on stage with a rhythmic sway, his upper body sways subtly to the rhythm of the music.

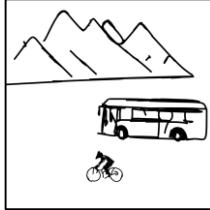


# Experiments

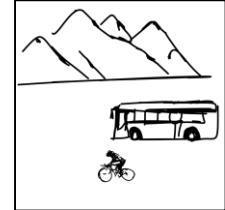
## ➤ Varying the texts



Sketch



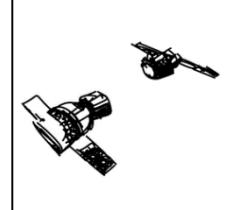
The cyclist overtakes the bus.



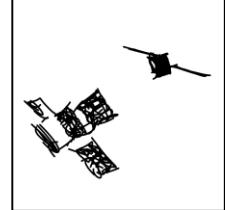
The bus and the cyclist travels forward.



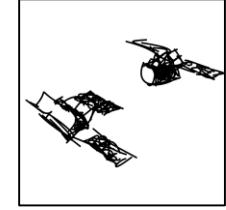
The Bus follows the cyclist.



Sketch



Two satellites prepare for a precise docking maneuver amidst the vastness of space.

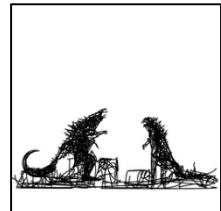


The satellites maintain their positions, floating separately in the vast expanse of space.

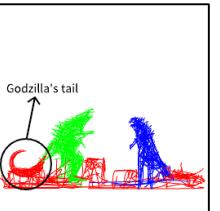
## ➤ Limitations

### Incorrect point assignment

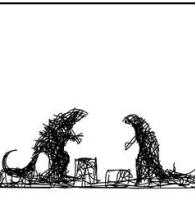
(Godzilla's tail is incorrectly assigned to "city")



Sketch



Point Assignment

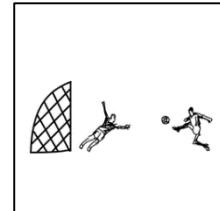


Output

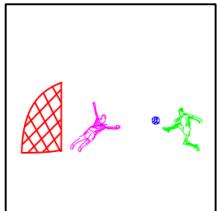
Godzilla and the other monster face off, the rubble of the city forming a chaotic battlefield between them.

### Incorrect motion planning

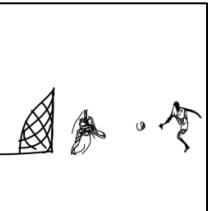
(the goalkeeper should move towards the football)



Sketch



Object Motion



Output

A soccer player kicks a football toward the goal, while the goalkeeper dives in an attempt to make a save.

### Failed to generate specified motion

('fight' could not be generated successfully)



Sketch



Object Motion

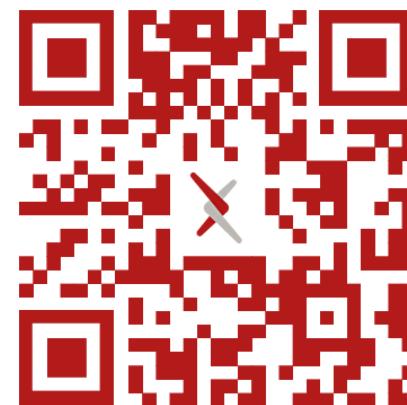


Output

The superman charges through the air with determination to fight with who counters with his overwhelming Hulk, strength in the ensuing clash.



# Thanks for listening!



Welcome to star if useful!

Feel free to contact us if there are any questions!