A Tale of the Indian Naga Sadhu

(compilation video)

Vivek Karthikeyan

MAT255 - Final Project

For my final project, I wanted to work on a time-based work that draws on my filmmaking background while incorporating learnings from my experiments with generative imaging tools this quarter. For an earlier project, I had tried using the platform Runway but my efforts with it this time around were frustrated by the pricing/licensing of the product. Apparently I had exhausted all my onetime "free" credits during that project and no longer could use the platform without purchasing additional credits. I did some research on other generative AI platforms for which trials were available and after testing out half a dozen products finally settled on Stable Video (a video platform from the Stable Diffusion folks) and Kling.ai that another student had recommended during their presentation. Finally I also added a soundtrack to the video doing editing featuring various Indian percussion instruments that I downloaded from an online royalty-free music database.

Being a filmmaker and trained cinematographer, I decided to make a short compilation film based on a singular theme and character. Although I did not care for a narrative or plot line as such, the idea was to capture a somewhat mystical space like an enchanted forest that is often featured in mythical tales about Indian sadhus or ascetics who are (in)famous for leading solitary lives away from civilized society often indulging in occult practices and rituals. The film would essentially be an edited sequence of multiple different shots all centered around this core concept, emphasizing the cinematic elements of visual storytelling (shot types, framing, camera techniques, etc.). I started with a rough draft of shot descriptions with details of composition, camera angles, look and texture of image, etc. and kept refining them and playing around with the text prompts.

I was especially impressed with the cinematic quality of the shots generated by Kling.ai. What I found especially interesting was the ability to specify beginning and end frames using image upload. I was pretty blown away by the results of the resulting videos that had quite complicated camera movements and character motion notwithstanding several noticeable visual artifacts in many cases which were distracting. It also had options for Camera Movement which was pretty impressive. I particularly toyed with zooming in, tilt angles and Roll. Coupling this with specific image motifs like "smoke wafting" or "incense smoke everywhere" I was quite satisfied with the resulting atmospherics of the visuals. The platform also seemed to handle

character motions within the clips quite accurately. For instance, with the image upload option used to provide an Indian sadhu's starting image I was amazed to see how the system simulated his motion within the frame through a combination of visual techniques such as scaling down of objects, adding eye movements, etc.

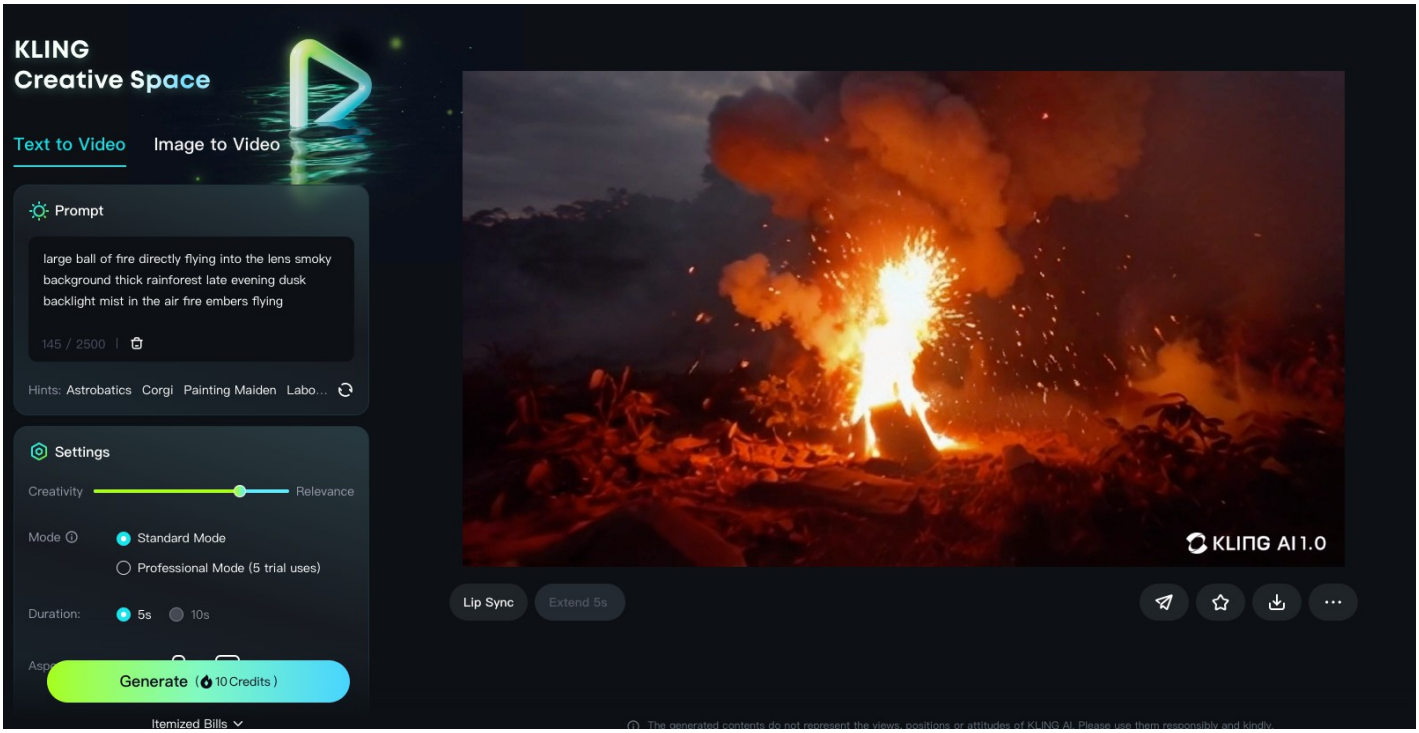
The videos generated by Stable Video in contrast seemed overly stylized and very AI-like for some reason. Although the resolution and quality of the images were arguably superior I found it harder to control for desired camera movements and angles, and often had to settle with the closest-next-best-thing for a lot of the shots. What I did like about the Stable Video footage though was the control I was able to exercise over the look. For instance, most of my prompts included descriptions of lighting requirements as well as ambient light conditions ("late

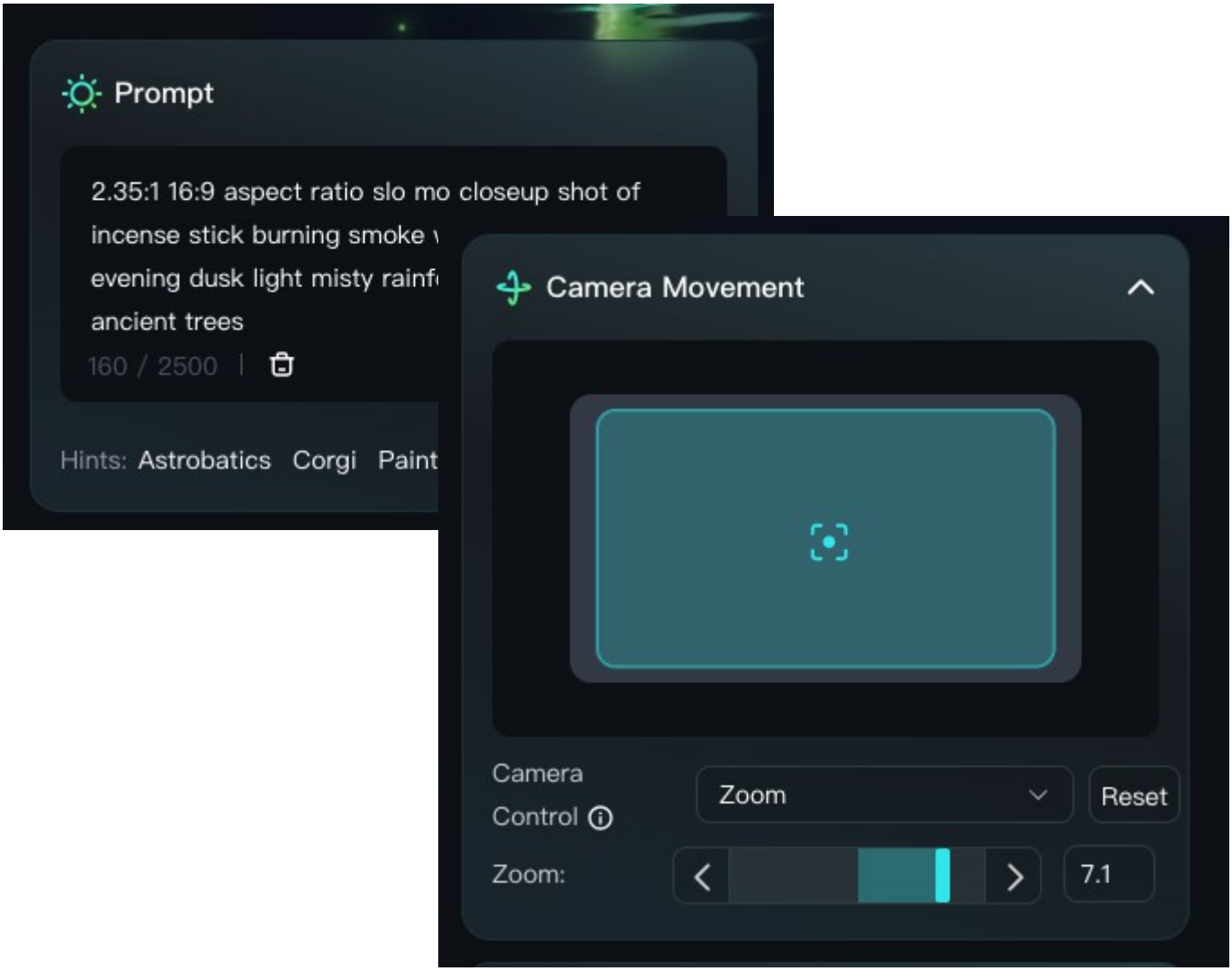
evening", "dusk", "strong backlight") which I thought was rendered quite accurately. The backlighting of characters is a standard technique in special effects cinematography to ensure continuity between fast motion shots such as action sequences that often involve body doubles that may not share 100% facial similarities. I was especially interested in seeing whether I can create a consistent look cinematographically speaking to produce a visually coherent sequence of multiple clips. To this end, I used specific tags like "rainforest" "ancient trees" and "incense smoke" in all prompts to steer the model toward the type of atmospherics I wanted to achieve in the visuals. At a technical level, I am not sure if the model was playing off of previously generated visuals in the series to create a consistent "latent space" as it were but the results were very much in the ballpark of the visual setting I had in mind.

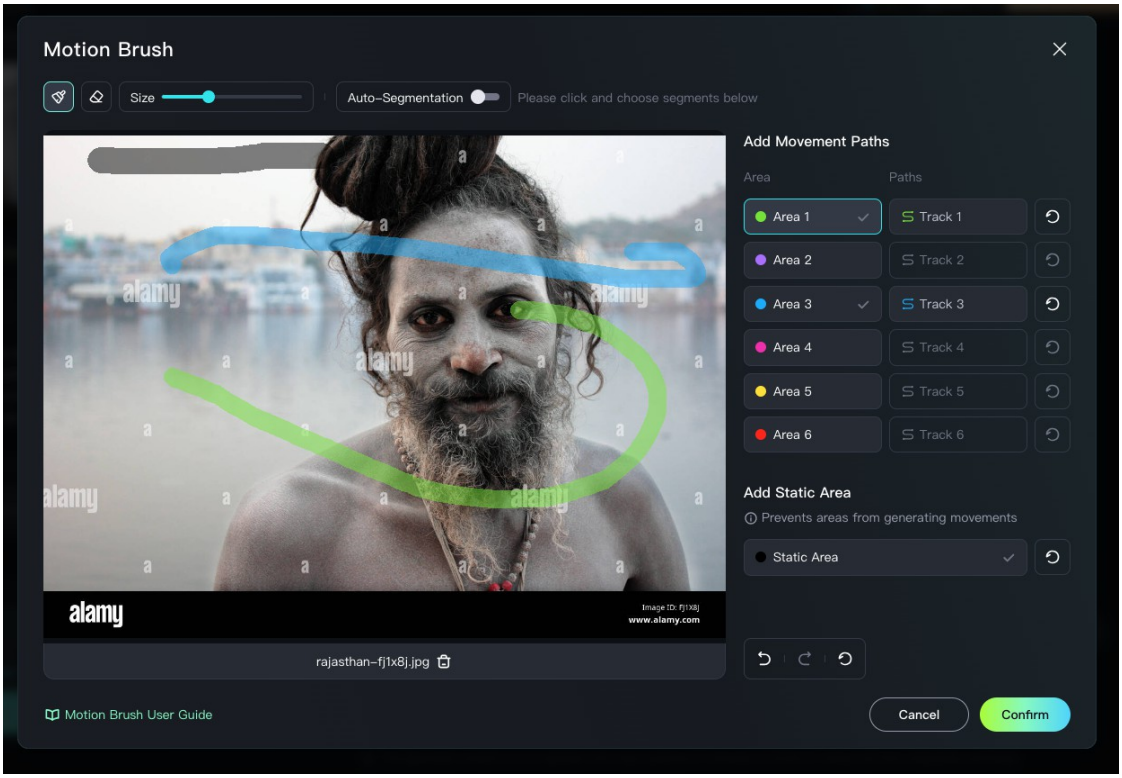
Overall, I felt pleasantly surprised at the final result because it felt like I was able to produce a coherent enough visual look for the film stringing together footage from different platforms which is often what happens in motion pictures too (cinematographer shoots on multiple film stocks, uses multiple digital LUTs, to design the overall look of a film.) I also felt a lot more comfortable getting the system to do what I wanted, meaning "prompt engineering"felt a lot less of a struggle than when I started out in the course.

https://vimeo.com/1037879371/2f08bea5ae?share=copy

Kling AI video editor screenshots




Prompt
2.35:1 16:9 aspect ratio slo mo closeup shot of
incense stick burning smoke
evening dusk light misty rainf
ancient trees
160 / 2500
Hints: Astrobatics Corgi Paint
Camera Movement
Camera Control
Zoom
Reset
Zoom:
7.1


Motion Brush
Size
Auto-Segmentation
Please click and choose segments below
Add Movement Paths
Area
Paths
Area 1
Track 1
Area 2
Track 2
Area 3
Track 3
Area 4
Track 4
Area 5
Track 5
Area 6
Track 6
Add Static Area
Prevents areas from generating movements
Static Area
rajasthan-fj1x8j.jpg
Motion Brush User Guide
Cancel
Confirm

slow mo 120 fps closeup 75mm telephoto lens soft image with bokeh 500 ASA film grain following brown

Generated from Text • December 4, 2024, 12:18:44 PM

View Prompt

Generate from another image

Image | Text

Image Upload Upload an image to generate a video.

slow mo 120 fps closeup 75mm telephoto lens soft image with bokeh 500 ASA fil...

Experimental

Camera Motion

Add camera motions to your scene

More will be enabled soon!

Camera: Locked | Shake

Tilt: Up | Down

Orbit: Orbit

Pan: Pan

Zoom: In | Out

Dolly: In | Out

Move: Up | Down

Advanced

Generate 10

Advanced

Generate

Seed

Sets a starting point for random generation

0 Use seed

Steps

Select more steps for higher quality or less steps for faster generations.

0 40

Most Efficient Highest Quality

Motion Strength

Encourages the AI to increase or decrease the amount of motion.

0 255

Text Prompt

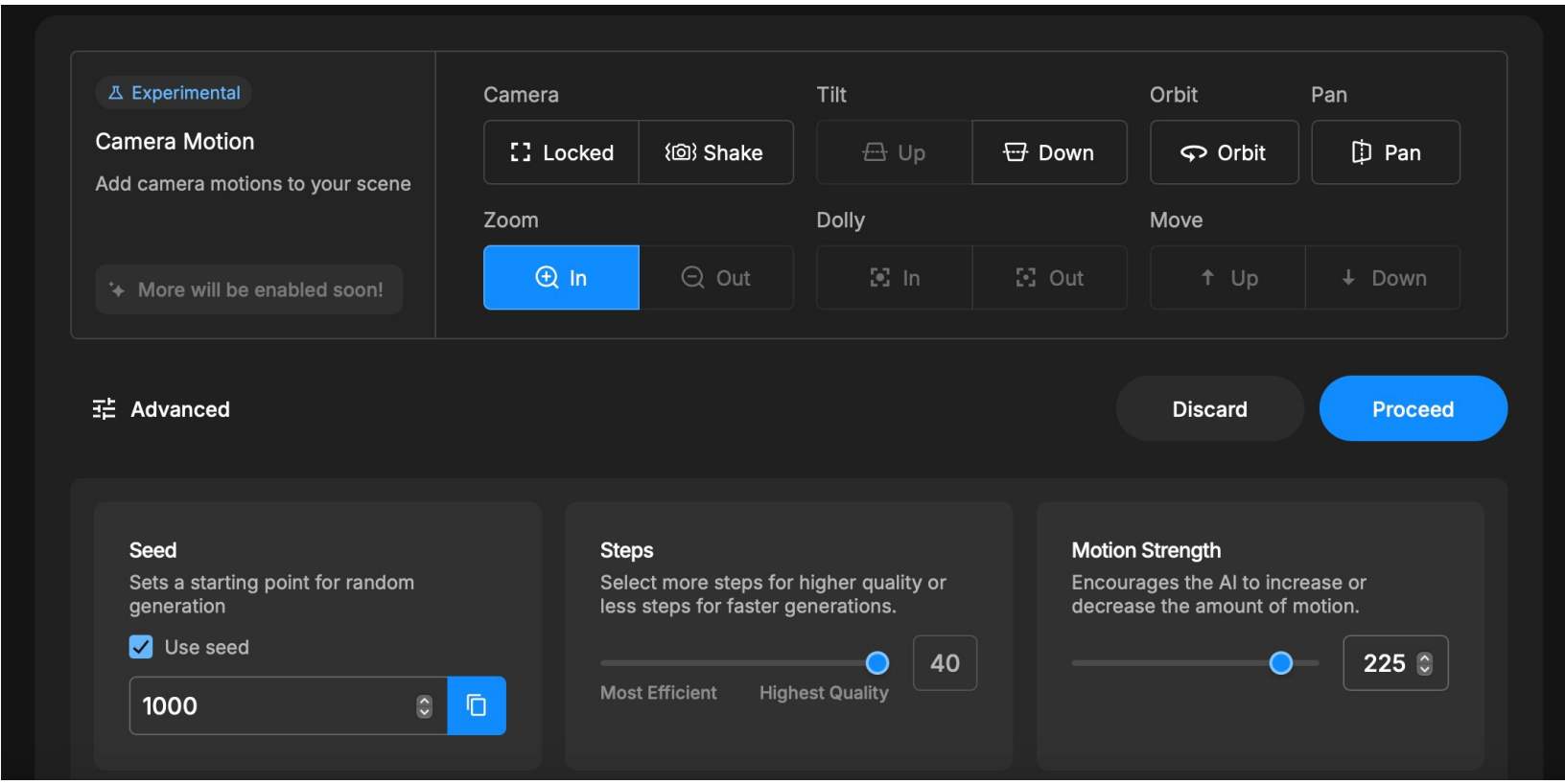
slo-mo 120fps closeup over the shoulder shot behind indian sadhu leaping in air flying hair foregrou...

Aspect Ratio

16:9

Style

Analog Film


Experimental
Camera Motion
Add camera motions to your scene
More will be enabled soon!
Camera
Locked
Shake
Tilt
Up
Down
Orbit
Pan
Orbit
Pan
Zoom
In
Out
Dolly
In
Out
Move
Up
Down
Advanced
Discard
Proceed
Seed
Sets a starting point for random generation
Use seed
1000
Steps
Select more steps for higher quality or less steps for faster generations.
40
Most Efficient
Highest Quality
Motion Strength
Encourages the AI to increase or decrease the amount of motion.
225

Midjourney generated reference images