

Remotion Masterclass: The VERY ADVANCED System Behind AI Motion Graphics

56 MIN · YOUTUBE · [HTTPS://WWW.YOUTUBE.COM/WATCH?V=PFHVXQ1S6F0](https://www.youtube.com/watch?v=PFHVXQ1S6F0)

<https://www.youtube.com/watch?v=PFHVXQ1S6F0>

SUMMARY

The episode discusses the innovative use of AI tools, particularly Claude Code and Gemini, to automate the creation of motion graphics and brand videos without traditional design skills. The speaker shares their journey of developing a system that segments videos into manageable pieces, allowing AI to recreate animations based on user-defined prompts and reference materials.

- The speaker transitioned from a data engineering background to creating a platform called creative.ai, focusing on user control over API keys.*
- They initially struggled with traditional motion graphics tools like After Effects, prompting the exploration of AI-driven solutions.*
- The process involves breaking videos into 5-second segments, using AI to analyze and recreate animations based on reference materials.*
- The speaker emphasizes the importance of context, using contact sheets to provide visual understanding for the AI models.*
- The system iteratively grades outputs, refining animations until they reach a quality score of 9.9 out of 10.*
- The combination of Claude Code and Gemini allows for efficient and high-quality motion graphics creation without needing extensive design knowledge.*
- The speaker plans to open-source parts of their project to help others leverage similar techniques in their own work.*

He's not a motion designer. So, he built a system where Claude code watches a brand video from 11 Labs or Higgs Field, breaks it into 5-second pieces, and recreates it animation by animation until two AIs graded a 9.9 out of 10. No After Effects, no keyframes, just plain English. By the end of this episode, you'll know his exact playbook and why the secret weapon isn't even Claude. This is the Remotion Masterclass. >> Let's maybe start really from the beginning here.

>> My background's actually in a lead data engineer and my background's also in data science. And recently, I've also got a very creative side. So, I thought maybe I can create platform called creative.ai. It's sim- similar to what you would Higgs Field, Magnificent, etc., etc. My vision for my platform was like users can bring their own API keys instead of using credits. You can do use credits, but I want people to have power and control of the money they spend.

For that, I was like, "Okay, how do I promote this very complicated app that I built?" Okay, I need a brand video. And I thought, "Okay, fine. Maybe I could whip up something with After Effects." I don't have a

background motion graphics and that proved to be really challenging. One of the first videos that I've been seeing recently was, "Oh my god, you can create brand videos, amazing motion graphics using AI." I was like, "Okay, cool. Let me see what this is about."

I think in general, some of the content creators on YouTube overblow what you can create with it. So, I started to create a basic project, which is like, "Okay, here is my brand. Let me give it some assets. Let's see what Claude code and Codex can create." It did brand video and it looked awful. And I'll show you what the first iteration of what Remotion and Claude code created. I think I can share screen as well. It was a simple project.

but I gave the context of Remotion skills. So, there was a new prompt or new skills that was released by Remotion themselves, which is like all of the animations, all of the useful things that an agent can use to create amazing animations. And this is what I created it, and I was disappointed. I don't think this is up to the standard that I expect for a professional brand video. It looks okay, but it doesn't look polished.

So, this is the first try, and I thought this is horrible. So, I decided to do something like look at other brand videos. And my first entry point to making interesting brand videos was Here's a good example of what kind of stuff I created with Remotion after that first failed attempt. Using other brand videos as a reference point. And I was able to just recreate prominent sort of UI motion graphics animations that you will see in some of the most popular brand videos online.

And this is simple animations to more complex. Like for instance, right now you're seeing like dynamic changes of size, the pictures and all the other details I'll get into. >> What I would suggest is let's do maybe a very quick explanation. Maybe you can open even a new project where you explain first of all what Remotion is, because again like for some people it's still very abstract what Remotion is.

And then you can maybe just like create on the spot like an animation, maybe with your own skill or with your own mental model here, so that people do understand like how it's actually looking not just like in theory but in in practice. First of all, what do they need? Again, think of them as non-technical. It's like too much. Even though Claude code and GPT Codex do write all the code, but if they see code, it's already oh, can I do that?

But what they don't realize is hey, you can actually write it in plain English. >> What is Remotion? Remotion is a allows you to and you give that power to Claude code or products, whichever agent that you like. And then you can think about what do you want to create? In this scenario, I've got loads of different animations. They're very complicated, very unique. And the way you can do all of these things is look at a reference image or a reference video.

So, you can see here right now there's four images with random text, different parts. All of this can be laid out by Claude code using Remotion purely in code without you thinking. If you give Claude code and this image and say I want it to be animated a particular way. Let's say let's take the simple example. Read this image. This is a very simple prompt and this is not how I generally prompt. So, if you change the model This is a very simple prompt.

I use a specific more complicated prompt to do my animations, but as a starting point, let's just take that example of this image. The animation that I want to create is what I've done before. A simple fade effect and the text pops in. Nothing complicated, very simple. And that animation is kind of nice for a lot of videos that you want to create. So, I've got something called Remotion skills added to my Claude code and you can see that's being used by Remotion best practices.

I just wanted to show you that is just a way for Claude code to understand what Remotion what animations Remotion has access to. Nothing more than that. It's like these are the types of animations where it's like rotation, fade, different types of complicated parts. We're just giving that to Claude code. So, now Claude code is going to go off and create a Remotion project. I think this will be like at least a minute before it gets to a creating the project.

>> Because people are confused. There are so many frameworks now. You can use like straight LLMs, After Effects, Hyperframe, Remotion, and like they're asking me like those questions like where to start, what to do. >> I think genuinely what the unlock for me with Remotion was how easy it was. Like I don't have to think about motion graphics, at least professional level motion graphics with ease. With it After Effects also is trying to enter into sort of agentic work, Claude code, and whatnot. But with Remotion I can just give that framework, which does animations to an agent.

And I'm not thinking about how that design looks. This is because Remotion has existed for some time. So, the code base, the projects, the nuances of how it works is somewhat learned knowledge by Claude code and Codex. As a result as a consequence it does animations in Remotion quite in a to a decent level. Where I'm not prompting it, "Can you do this? Can you make it move this way?" And I just get it right. In After Effects, I've done animations through Claude code as well. It's a lot a lot complicated.

It's a lot more intensive on the computer. It's slower. You're dealing with layers, complexities of attaching objects to fixed positions, maintaining camera camera was another challenge. Having Let's just take this example you're seeing this movement effect, these transitions, all of this it's complex. It looks complex, but Remotion can do it in under 10 minutes. To do the same effect in After Effects if you were doing it manually, it's For a professional, it's probably easy, but if you're a newbie, >> Yeah.

>> this would take so much time. >> Yes. >> And that's that's where the unlock for me was, "Oh my god, this is the first time I can do motion graphics without being a motion graphic designer. Of course, it's nowhere near professional grade to the level of what motion graphic designer would be like, "This is perfect." But for someone that needs a decent brand video, decent animations, a sort of you sort of interesting visuals that are trying to convey my UI, my product functionality, this is sufficient. This is more than sufficient. And that's where the unlock was.

Maybe I should show you the other one cuz I don't That was called That was called code. This is Codex. So, I didn't Codex was better at animations some animations. Just this animation. This looks so pleasing. I wouldn't use After Effects to do this thing cuz I can't I don't know how to do this. That's why I use Remotion. It was so

much easier to recreate a video purely through code code in a matter of minutes and be like, "Oh, wow, I have I have an output."

than it would have been if I did it on in After Effects, which would have taken me a long time. But I guess those skills and trade-off is that of course, you can learn After Effects. It will take you what a couple of days and sometimes more complex animations, maybe a month. And then you've got those gra- basic skills to do the amazing animations, but then if you're only doing it animations or videos repetitively, >> Yeah.

>> you don't need to do that. You can just use Cloud Code and Remotion. It's powerful. If If let's say you're doing an explainer video of from new tool and you just have an image of that and you want to animate it and you want to do cool visuals while you're talking about it, Remotion can do that. Because again, it's all through code. Agent with Remotion is such a powerful combination. Don't But like I said, everything is under your control. Like you can't expect it to just recreate everything for you in the best way. You have to be the driver.

>> Yes. >> The motion graphic designer with who has this in their hands can do amazing things. And a beginner that has no idea of what they want to create and they just go off and do something, well, you will end up with results like my first try, which was like here is Remotion skills. Go and do some animations. To some other users, this might be sufficient. But again, like even if you want to do this, you can. This is To me, this is very basic.

And this was done by just saying here is Remotion skills. Do some cool animations. I was like, here is my UI at my platform, creatively.ai. Go and do it. But when you want to think about more complex stuff, and the way I improved was, let's say, in here, when this animation effect came where there's a prompt that is typed and then you see a other block, which is like oh, here's a suggested generation. I had to prompt what my app was very carefully. And I was like, here's a picture of my app UI, use that. Those details you will never lose. You have to do it. But this is powerful. It's a powerful because it's like, I don't know how much time it would have taken just to do this animation.

>> Yeah. >> And now I've got a base a general framework to do more animations. Let's say I add a new feature. I don't want to go into off fits. I can be like, "Code code, can you do this? Here's a new UI, go and do it." I'm not thinking about it and that that's where Remotion became super powerful. Let's see what we actually did in the past 3 minutes. >> >> It's doing the packages, like I find.

Okay, let's see. Let's see what a naive attempt without any prompting Code code did with that image. And I didn't actually turn it I didn't specify the animations so far. I just said very basic with no direction. Very basic. It's some photos fading and use public images. I think on blacks. Oh, wow. Oh my god. Okay, so this is way beyond what I expected, but you know, this is what I mean. Like this is a genuine reaction of what Code code did with an image. I gave the simplest prompt you can imagine.

what did I say? I gave this basic image of what I want to recreate and create a Remotion project with some animation where the photo is fading. Use public images. Go and do it. and the Pixels and Pixabay are like websites that give free images that you can use. And the output is this. This is very clean. But of course we I am

using 4.8 Opus as the model. It's a very powerful reasoning model and if it's Code X you'd use the latest and greatest. I would recommend even for any sort of animations, just use the bigger models because you might spend more time with the smaller models or like maybe Sonar etc.

whatever. It might take longer for you to get the kind of output that I'm seeing, but do play about because if it's very token hungry that you can just use in a video somewhere where you're trying to promote something. And the those images that you're seeing, they're public images that I didn't gather. This was Claude code going out finding the relevant images by searching. My guess is look at makeup images or what was the image or the image that yeah.

Has some video or whatever. Go for the channel run products and add the pit. And the animation of the generating videos is something that I didn't prompt either. That was also by Claude code. And it's a very cool fade effect that you're seeing all this. Again, all Claude code. Decisions that it makes, this is where you can refine. So now I've got this. I don't maybe like the way generating video shows up. Well, okay, that's fine. All I need to do is like maybe I want it and it transitioning from the right side and comes in.

And I can just prompt can you make generating videos come from the right side? So this again, when you think about After Effects Okay, the combination is not complex if you know what you're doing, but I have to know have to have the software installed. It's complicated. It's very intensive on the computer. You load it and then you do the key frame animation to do this and Claude code did it in a second. So I've just changed the animation from going from the bottom to the right side and there you go. The transition was nice.

Again, with a single prompt. >> This is looking amazing. How long did it took you to create something like that in Remotion? >> In Remotion, so what I started with was I picked a good brand video that as a starting point. The first ever brand video was magnificent. And this is not very good. I don't know how well this will show up. And then Remotion, sorry, the pre-picked brand video was a simple test, which is okay, this is the feasibility or how powerful an agent and re-motion can be if you give it a reference. And our reference being the video, which is the original brand video.

It's not great, but it is pretty cool. This is purely So, the bottom right is the original video by FreePik. >> Yeah. >> And what you're seeing on the screen is the recreation I had. So, I was like, this worked well. This one is using Claude code and re- my brand video that I showed you earlier. So, I open-sourced the entire video I showed you earlier, which was like all the animations, starting point, whatever. And I fed this to Claude code and said, "Do something better."

And I The reference was, "Use this open-source project and then use this FreePik video to recreate the brand video." Which was not too bad as a first try. >> Yeah. >> What I ended up with was a set of rules as I was building it and iterating on it on the first day. So, I had my generic animation template that I created, which wasn't great, and I have a reference brand video, and I kept prompting it, "Do this, execute this, refine this, do that" to get to as close as the original brand video.

Then I realized, "Oh, wait, I can do other brand videos. This is a great starting point." As I started to think about what other cool brand videos are there that I can use to repurpose when I'm doing my main brand video. I think I did But, I'm not going to get into like the cost and whatnot. I'll get into that later. I did Hixfeld, ElevenLabs, and another Hixfeld one, and Adobe Firefly. These are all like interesting, unique, dynamic videos. Animations are also unique.

So, I decided I can't keep prompting it, refining it, cuz it's just complicated. So, what I did was I built a project with Cloud Code and CodeX called the video recreation kit. And that sort of laid out the entire skeleton structure for I have a video. >> Yeah. >> I want it in real motion. >> Yeah. Got it. >> Go and do it. I've got good taste. Like, I know what good looks like, but I don't know how to execute it. Like, do move this, I don't know, chatbot in a certain way in the animation, move the cursor in a certain way to create I can't be asked to do all of that. But, there's good reference.

And then I can then refine it using my own application. Say, replace the elements. You use my images. You brand design. This is essentially the framework. What I do is two main agents, Cloud Code, CodeX, as the driving force to recreate a video. They're given this entire template. And the core driving powerhouse in this whole thing is actually Gemini. >> Okay.

>> tool to actually create motion graphics is not Cloud Code, it's not CodeX, it's Gemini. >> That's interesting. >> Gemini has, I think, best in class when it comes to JavaScript animations that I've seen. I've tested with different tools. So, I used Cloud Code to be like, can you recreate this HTML animation? This is before, a couple of months ago. Can you use I used Jet Chat GPT, X High mode, the latest model at the time, and I used Gemini. Gemini created the cleanest animations. And as a model, it is the only model that can understand video.

>> Yeah. Yeah. >> So, Cloud Code and CodeX, they can't understand video. They can read an image, but they can't understand video. So, this is why Gemini was used primarily because it creates the best animations in terms of UI and HTML, and that can neatly translate into Remotion. So, the process is let's just take a random video or the the FreePik or the Hicks field video. It's fed in, split up into 5-second segments. Each segment is then just fed into Gemini.

So, my backend is creatively creative.ai's backend. I can either use that or my APIs I've got in the backend using Gemini 3.1 Pro, by the way. Not the older models. You have to use anything higher than 3.1 Pro. And that model is really powerful when it comes to reasoning. And it has a very specific prompt, which is like, "When you read this video" I can't remember where the Oh, yeah. So, this is the agent prompt, which is, "When you read the video, how do you respond so that Claude Code and Codex understands that 5-second video?"

It reads the video, understands nuances of that video, returns a detailed explanation segment by segment. And then I also extract frames from the 5-second video. Every No, obviously not every frame, but just like in between into video. That can be fed into both models, Claude Code and Codex. So, now it has context of an animation prompt design from Gemini. Uses the images as reference.

And all Codex and Claude Code are doing is I'll start with a rough template using the prompt that Gemini has created. Then look at the image. All right. Refine. Look at the Remotion project. Refine. And here are, for instance, let's say as an example, let's say it started off with maybe three blocks with an image and looked ugly. It looks at the source reference image and the context of the prompt that Gemini has created on what it's supposed to look like. And the animation and the movement, obviously 0 to 5 seconds at the 5 second mark, there there might be another animation sequence.

So, there was like interesting challenges regarding overlap of animations, challenges of like when should it stop, what is the motion graphic in that sequence? You can see here some things are simple. It's just a pan, a camera pan. some things are more complex. This one's not complex, it's just blocks popping up. I think the more complex one where it's stuff like this. This one took ages to get right. It seems simple, it's just text. What is this? Just scrolling up.

And an image is popping up. This was challenging. Even little details like this. For some reason code could could get this right. This very complex animation. Yeah. It's nice, right? But it just both agents failed to understand that there is an image behind text. So it was challenging, but that reference of here is a five segment five second video segment, understand it, refine with an actual image reference of at second one, second four, what does it look like?

And they keep updating. They are the graders, they are the refiners of code. And now we have a project. I've got a template of that five second. And I'm like, okay, this is and I've got what I call a like in the playbook I keep I give it guidelines on what quality score that I should expect. And each agent grades by 9 out of 10. Only if it reaches 9 out of 10 and it assesses itself, what is 9 out of 10, does it continue to the next segment? So whenever you when you see all of these projects, me, I haven't prompted it to do this.

I've given it context of a brand video that I liked. I thought you like even this very simple animation I quite like this one I think. There's a prompt by that shows up. >> Generate a couple of my own. >> Yeah. >> Okay. >> So there are two variations there. So this is Claude Code's version and this is Code X interpretation. And there's so many iterations that I've done to change what I want. There are parts like this is my video editor in my platform. Again, I had to change the UI completely to match my the brand video I chose was Another Labs. They have a video editor that is very unique.

But it also looked quite similar to mine so I was like okay fine I can use this brand video and then make it my own. This animation is quite cool. Like you have three graphic images that are then that the kind timeline elements. Like I couldn't prompt that. I could not prompt an agent to do that. >> Let me like like first of all a lot of people are not very technical. Watching my videos they are like more coming like from the content creation side. But there is like one thing I would be interested in.

Creating a reference image and like the image understanding of each model is quite good. And now Gemini is the only model that can understand video as well. But still I found like when I tried to ask Gemini to reconstruct the movement, I still had problems that the movement were very like it was going into the direction but it I had

problems that it was that specific. Now how did you make it that specific in a way that like every movement, those transitions, those let's stay on this example here. You have three clips and it's transforming into this timeline. Is it like really the understanding from German did you do it in a way that Gemini did understand stand like all the transitions or how did you do it?

>> I would say the biggest thing that changed what kind of motion graphics I can do with Remotion, giving it the Remotion skills that they released recently. I think it was like a month ago or two months ago. And in combination with my open source brand video project helping it to understand that here is a reference of some generic animations. These agents Gemini know they've got wide depth of knowledge in terms of what they can execute, but they also know how what kind of animations or limitations that Remotion supports. Like I didn't know that Remotion can do this very cool animation where is it? This effect of bouncing cursors.

This was only done because of Remotion skills. >> Yeah. >> And then when I created the brand video, I always asked Gemini to use this reference to see the art of the possible using Remotion. I think only when you have some complex set of animations as a guideline, do these models work well. Otherwise, they don't know what they're doing. You're right. Gemini wouldn't have done this effect especially as well unless it had a generic template of oh, it can move a block and make it smaller with the timeline. It can change with the time. And as I was grading it, it there was a None of the process is fully automated.

I am still grading it after checkpoint. So, let's say I go I say go and recreate this video. And it looks horrible. I'm like, start over. This is the worst thing ever. I give it clear guideline on where it went wrong. And I think that's what sort of refined my video recreation kit playbook. So, this playbook was probably over 3 4 days. And then after the fourth day, I was like, "Oh, okay, fine. It's it's doing a good job without me inputting any more sort of comments on how it should do it. Now, I can actually just give it any video and be like, "Oh, it will do a decent job."

But, that took some time where I was like, "Here is a template. Here is the mistakes you made. Here are the challenges." Remotion, you one of your questions you said was, "Can anyone do this?" With good good templates. I'm thinking of actually open sourcing one of the projects. Even if you think that's helpful for engagement on your channel, like I can open source it. Like, I can I can give that as a project on my own GitHub repo and then they can use it as a starting point.

>> Yeah. >> A starting point can help them realize the potential of Remotion. And then the model because of the model has context of oh, it can do these things when there are images. It can do these types of animations. I'm just using it by to execute an action. But, when I got to this point of the playbook, I had to even create custom scripts that Cloud Code reuses every time it needs to perform an action because I need predictable outcomes.

Predictable meaning for segmenting and for extraction of that like images from the video. All of those things, I created them as tools. Again, all by Cloud Code is like, "I need this to be predictable every time I run it. I need extraction to be exact same with that. I need it to be You If you could go through your playbook because I think like high-level speaking, the mental model and then what are the steps really here? So, explaining maybe your

playbook here.

Like, just speaking from the steps itself. Like, high-level. Not going into the details, but just like so that people do understand like, okay, this is how it could look like after a really lot of trial and error. The process is like you take a video, you segment it, meaning you break up the video that you're trying to recreate in re-motion. the tool that you use in this case is the segment script I've created.

Then I'm extracting individual frames from each of the segment so that I have a little more understanding of what I'm trying to recreate in something called a contact sheet. The contact sheet is essentially this. It's a panel of images from a scene so that Claude Code or Codex has even more visual understanding. And now it has three unique sets of information. One is a video, two is individual frames if it needs to sample or if it needs to look at some parts of it, an entire scene collage where it has more context.

And now Gemini is just reading that information and returning one thing. And that one thing is a detailed overview of the animation, the images that are present, and that's also critical by the way. A lot of motion graphics and why this looks as good as it does is it has interesting images. Like this only looks good because it has interesting images. So even the type of images that are present in that scene is then what Gemini tries to understand.

Gemini then responds with here's a very detailed overview of this scene, this 5-second segment. Then what I'm doing is Claude Code just takes this gigantic prompt. I can actually show you how that prompt looks. So it returns a scene-by-scene playbook per video. And that detailed prompt is just executed by Cloud Code or Codex. And when it does that, it iterates. You can see it's iterating. So, over here, it says each agent owns one segment. So, Cloud Code takes that detailed overview of what a video looks like in code language.

And it just runs a first pass of let me try to recreate it. >> Yeah. >> Let me create some of the images cuz I use my Creatively backend. Let me create some of the images that I'm also seeing on this video. And then it checks the ground truth. Meaning ground truth meaning what the source is. All it does is oh, I've I've created a video. Now, let me compare. Create a video, let me compare. That iterative loop is done autonomously by the agent until until until it has a grading. So, both Gemini and I think I forgot to mention another detail.

Gemini and Cloud Code are grading it. So, I asked Gemini and Cloud Code to grade this output. Take an image frame, take an image frame. Does it look similar? Yes or no? And refine refine refine. Does the animation look smooth? Refine refine refine. Until until until the grade is 9.9. And this is sort of how I've sort of done it. Like until the >> genius. >> rating is the highest rating possible by two agents grading it.

And once it has a score of 9.9, stop. Render the MP4 file. And we're done. And that's essentially the overall playbook on how I've recreated different brand videos and so I can repurpose and reuse them for my own brand video. Cuz some of these animations, they're not all great. Some are clean, like this was clean enough that I could just take out some parts of it, repurpose it to what I like, and reuse it. So, even this, this cool animation, I can change the colors.

I can change the overall layout. I can change the text. See, you you can see it says creatively, but obviously this was 11 Labs. >> Yeah. >> And those details are me then guiding it me then guiding it to my own design. And I like the level of projects that I was able to recreate, it just it's it takes probably about 10, 15 minutes for the entire video piece to run, and I'm just back and looking at Oh, here is the video.

Like, even this, this This is very clean. These subtle shader effects, and the fact that the background has a nice gradient, all of those decisions were made by me. This was made by another motion graphic designer. >> Yeah. >> And I'm recreating so that I can then use my own knowledge and my own vision so that I'm not thinking about motion graphics. I'm like, execute go and execute. Here's a nice prompt, go and execute. >> You know, like I can see already some of the comments like is stealing okay?

What do you say? >> No, no, this is the thing. That that is This is primarily because I can't pay a motion graphic designer a lot of money. Like, I would love to be in a position where I can afford to pay someone to do this professionally, but I don't have it. Like, I want to do cool animations, cool brand videos to promote what I'm creating and that I'm passionate about, but I I don't have the resources.

This allows you to take ownership of something that's creative, and you can have your own spin with it. Like >> Yeah. >> Yes, I am taking source material, but I'm still using my UI for the inspiration of how it's going to look in the end. >> And and I think when I'm you know, like at the end of the day, you are training basically. Now you have basically the secret sauce and that's the starting point where you can start to create actually something exceptional based from that, you know, like And if we look >> Exactly. Exactly.

>> You know, if we look at all the artists and basically like all of the great artists are are using references. Like me like when I've directed commercials, the first thing an ad agency wanted to see is always like show me your mood board. So it's basically already like part of the process stealing. I mean, not really stealing, but because you have to recreate something. >> Inspiration. Exactly. >> Yeah. So I think but still like a lot of people don't don't really appreciate, you know, like they don't understand that that's kind of like the starting point and now that you have the framework basically, that's when you can build on top and that's where you can and as you mentioned like afford to create also some really nice brand videos.

>> Exactly. Exactly. And that's that's the vision. Like now that I have an idea that I want to execute, which is an animation of a part of a new feature I'm releasing. I don't need to recreate all the videos. I can just use these amazing templates I've recreated and cherry pick interesting animations, interesting visual language and just adopt it to my own brand vision. And I just do the animation and then show it showcase it on my platform.

Sheets are you can just think of it as like a collage of the entire scene. This this kind of visual understanding is also very helpful because the model then sees a start to an end frame of what you're trying to recreate. Instead of cuz some of some of these models they're not smart they they'll they're not they can't take in a video input and understand reason like Gemini and I still want to use Claude code so I give it a contact sheet. Which is >> That's that's actually framed. Nice.

>> Because I think you solved the problem that I'm encountering explaining basically the motion. So if you are like having here a sheet the motion is explained through the sheet. So you give it like one sheet >> Exactly. >> and and now the model is understanding the motion. >> Exactly and like the transition from the left of text kind of fading away to some images flashing to another UI showing up. I can't prompt that out. I to prompt that out. I just want an agent to reason about it and then just do it in Remotion animation.

Remotion SDK. Like all of these parts I don't want to think about so I've just created contact sheets. And Claude code reasons and all code acts reasons about these contact sheets. >> Yeah. >> And then uses Gemini to then create a visual prompt which is like what are most likely the motion graphics from Remotion that it needs to use to achieve that effect where for instance let's say you can see there's this tilt effect of the prompt bar.

>> Yeah. >> Gemini can understand that. It's seen this contact sheet. It's seen the video. It's seen the frames. And I'm not giving I'm not giving you everything because you know you explode your token count but I'm giving it things that are relevant. And at this point when it gets the visual understanding of the scene well then Claude Code is amazing. I think of Claude Code as the best executor. Like if it has good understanding of what you're trying to accomplish, it can do amazing things.

It can do it super fast. It can do it to a really high quality and when I say high quality, I mean like things like this. These things were created with Claude Code. All of these animations using that framework of do the sequence of events and I only got to this point by iterating. Like I didn't know I had to break up a video into a segment or a contact sheet. These things only became important so that simple MD file that Claude Code created for the first one where it was like use an image, recreate it. Two of my very complicated approach or when I say complicated, I mean like the agents has complex instructions to take a video, to go to my end goal of recreate it.

I had to iterate on it and be like, oh, I want to look like this. How can I solve it? And Claude Code would just give me a solution. I wouldn't know some of the solutions or I would know some of the solutions. Breaking up a video is something that I can think of. I didn't know that you could do contact sheets. This was not my invention. This was Claude Code telling me contact sheets might be helpful in this scenario. And it created amazing. And it created my script. Let me open up Gemini Omni agent. So, this is a skeleton of when Claude Code tries to recreate a video.

Yeah. All I have to do this is I put a source video in this template. So, Claude Code creates the skeleton folders and I can put one video. What it does then is uses the first script which is so which is the segment. Take that video and it breaks it up into segments. Yeah. So, that video is split into 5-second segments is then broken down into per segment frames, if I need it.

>> Let me ask a more technical question. Why do you split the clips into 5-second segments? Have you ever tried something like a scene cut detection? I think there is like a Python library that can measures the like or is it like in motion graphics not really practical? I'm not sure. Have you ever considered like scene cut detection or >> Honestly, that might be an improvement to this now. Like the reason I don't I didn't want to invest too much into the nuances of where scenes transition. That was a challenge. So, the number five is arbitrary. I found that it

was the only one where I had more visual control over scene. If it was it was it was 3 seconds, it was too simplistic and the animation >> Yeah. It will work great.

>> If it was 10 seconds, it just gobbled up the animations because there were too many complex scenes. So, I chose a very arbitrary number, but scene detection might actually improve it. Like just just take a video and say take the timestamps of which scene ends where it starts and this cut point might be perfect because then I don't need to deal with some of the complexities cuz you can see that the first three are clean. There's one >> Yeah.

>> animation that's happening. It's just it was just popping Texas fading away. If I just cut the scene there, Claude Code or Codex using my sort of workflow of segmentation, it would do an amazing job. But, the introduction of the second panel it then tries to do more things and it can mess up. So, scene detection is another improvement. So, Theo just gave me an excellent example of how I now can improve this. It's like, "Oh, there's another animation layer now."

Where and they would go into this point, it would just be here. Before segmentation, I would add scene detection. And then segmentation so happens but based on scene detection. So, Claude code would have just helped you with that as well. It's like, "Can you create a script to do scene detection?" >> But I can just imagine like scene cut detection is working fantastic like on like cinematic clips but on motion graphics where yeah, spikes are going up and down.

I'm not sure if maybe like the splitting into in into like defined segments or maybe it must be a combination but yeah, I guess like really playing around and then figuring out like that's the way to go, yeah. >> literally a trial and error process. Yeah, I would say a lot of this is like trial and error. Like, I would even I would try the scene detection just to see if I'm >> I'm like, "This is it. This is good enough. Let it go and do its thing.

Here's a video recreated." >> Yeah. >> It'll help the model understand these are the things that I can recreate later if when a user asks me X and Y. So, I've got about 20 projects that I've recreated. I don't use all of them. I only use two, three projects that I liked and the reason why that I sort of show I showcase here. Those templates are the rough starting point when I want to iterate. Like in my brand video that what I ended up using So, all of this all of the animations I'm I'm going to show you here. This is all in Remotion. I've just added it.

But I was catering it towards what my platform supports. So, a user can do it but they just need a template >> Yeah. >> for an agent to look into. Oh, this is all the possible. Claude code, I would just say, "Can you create an MD file with the learnings or can you create an MD file on how to create such cool Remotion projects?" And what Claude could do is whatever effect that you've seen here, it'll capture that because you've specified something and you like the effect and you've also changed the prompt to say to go from bottom to a top to right to left, it'll capture that in MD file.

>> So, this is basically your playground. >> what I did. Essentially, yes. >> I mean, of course, like after 4 4 days of hard work. but >> Well, before days of hard work, so before >> Oh, sorry. >> Before days of hard work, yeah. This this is it. This is it. And like iterate iterate iterate iterate until I've got the visual language that I am happy

with. Like like I don't like this. This looks boring. But I'm happy with this.

So, now I've got a template I can reuse. >> Yeah. >> So, this is now saved project that I can ask Claude could to reference when it's done creating another video as a style reference or animation reference or whatever you want. After that painful process of I don't like this. I like this. I don't like this. I like this, you've got some guidelines on what your taste looks like and what what your visual language, what your idea of good looks like. And you're just capturing those motion graphics, animation, simple transitions into a simple prompt structure. Which you're like, I like clean fade effects. That could just be a prompt. But then it's a little more detailed and you just capture that.

let's see what Claude could did on how to build something. So, this is Claude code writing what it considers as my template. So, again, the very complicated template you saw earlier was created like this. It's just guidelines. It should say here you can see it's talking about how to do a fading grid, which is this effect here. But this is something that you now have for future reference. And this is done automatically by Cloud Code because it's a React code and it Cloud Code is powerful, especially when it comes to code. It can do this in a matter of seconds or minutes.

Yeah. >> If it's still abstract for some people, let's open maybe like the learnings starting MD file versus your file so that they kind of have an understanding how it's like can develop over time. Like maybe your playbook versus how to build MD here. >> Yeah. Yeah. So, this obviously was a very basic one. What did I do? I said read an image, recreate it, and I gave some guidelines on the animation. That's great. That that that works for something simple.

but it doesn't work when I'm thinking about multiple transitions, different segments of a motion graphics video. A lot of the brand videos for big companies, Higgsfield, they're complicated. There is not one animation. There's not just like a simple effect. There are dynamic elements, sizes, shapes changing. For that to go from sort of this basic template or here's an image, recreate it with some basic effects, to go to this more detailed playbook and how I sort of get got to this point is as I've mentioned before is literally iterating.

Iterating meaning I've now got a starting point of an animation. Let's say I've got another animation that I want this to transition to. So, let's make these square grid elements move out and just expand. I could either use a reference or prompt it more. So, it's whichever way works for you. You could do a start image and end image as your starting point of here is what I want it to look like. Earlier I gave this as a starting point.

You can use AI to generate this next image, and then feed that as the end frame for Claude code to recreate this next step. I mean, like that's what that's another approach that you could take if you don't have a reference and you want to be in control. So, if you want the >> Now, did you actually some more complex learnings? Do you have a built-in hook that like took maybe the learnings automatically, or did you really went there and give it and ask Claude to save the learnings manually, or did you have any special technique there, or how did you do that learning stuff?

>> Nothing complicated. as I was sort of recreating the MD file. So, this wasn't In its way it's initial state, maybe it was only maybe 200 lines. Mhm. And over time it kept improving, improving. Once I realized these are the guardrails, the things I needed to do. So, earlier I showed about I take an image, generate an animation, great. But it's it's simple. To get to this point of something more complex where I'm recreating a video or scene of a video, I had to think about what tools it needs to execute the to execute what I want to achieve. In this case, tools like tool that can cut up video out into multiple segments.

I had to use my back end, different parts of it I was building over Oh wait, I need to do this. I So, like the different parts where it's a let's just say I had to use create specific scripts that were created by Claude code. Over here, I'm not doing any scripts. It's just a It's just a guide. >> You have built a system that yeah, you that that has kind of like a self-improving mechanism. How does it work like in in your playbook? Like you give it a rating. Is it like how is it working?

>> So, the way I select the agent there >> I maybe >> Yeah, sure. So, the way I let the agent grade it so was through multiple program refinements on what good looks like. So, this was over multiple attempts. I've done it over 4 days. It was like, "Here's a general starting point. Here is the full prompt." So, there's also learnings I've created when you do different stages, what it shouldn't do, do. And when there's something called per scene agent prompt, which is like, "How do you do the extraction?" So, like when you feed this information, all of this were guidelines that had to drill into the model when it reads it.

So, that as it's grading a scene and output, so it does looks at the scene, breaks it up into images, looks at the video, tries to recreate first pass. Horrible. And the horrible in the grading happens where the agent what it does is it spits out and I'm asking the agent to look at it again. Like, here's the scene that you created. Here's the original scene. There's a mismatch. This is crap. Grade is five. And so on and so forth until that grade is nine. And the nine itself I'm pretty meaning it just knows what good looks like. It knows these models are good enough to know that, let's say.

Those four images the way they look. They are laid out nicely. They have some unique design to it. And if they're all over the place, then the grade is five. And that's prompting that I've given to the playbook. Make sure general concepts of like all of these are just like UI elements, right? So, it understands from that perspective as well. It knows how it's supposed to look as well. And then it grades it and it tries it until it's is Once it reaches a score of nine, it stops. The agent stops that scene recreation. So, just like another question, do you write this all in one MD file, or do you split it into basically like multiple smaller MD files where you are like, okay, I have maybe a categories, and now you are How do you save context basically? Because I guess it gets like after working for four days on this one, you have probably a lot of information.

Now, how do you optimize that one here? >> Yeah, I think but because I've got the pro plan, like I had both for Codex and Cold Code because I I need it when I'm doing the platform. I don't think too much about context optimization or like using tokens really well, just because it they don't burn through barely half of what I need in the day when you're looking at images. So, this entire prompt, I've got maybe like four MD files. So, in the video recreation kit, there's a playbook, which is like set of rules, what it needs, information. And information

meaning guidelines.

Guidelines on what it's trying to execute. It's you don't want to constantly recreate, write a big prompt, or maintain large pieces of text. And the only way to capture what you're trying to accomplish is typically either in the text file, or in this case, an MD file. MD file is just prettier because it captures the bold, italic, tables, and it has nice functionality like like a Word document. All I'm doing is capturing a gigantic prompt. If I want to break it down, it's a gigantic prompt that the model uses.

So, that when it starts any pieces of work any piece of work, it follows the rules that have outlined. It's just a certain actions that I'm giving Claude code or Codex in a very detailed manner because I don't want to mess up because they do. These models are stupid. They are not stupid. They are smart in the sense of they're only smart if you give it detailed instructions. In this scenario, I'm giving this gigantic piece of text so that it knows first do this.

What I'm saying is look at the look at the video file. Break it up. Get images. And then store it. Get a understanding of what the video is. Done. Then this is the context of how you can do it. Use these tools only. There are some tools that can break up a video. So, FFmpeg is a CLI tool that can take in a video and just break it up up. And there's other parts to it. And that's those are just guidelines for the model to be like, do this, do that, do this in sequence.

And then once it understands that starting point or structure, it executes. >> First of all, thanks. I think it's a very amazing framework that you have developed in specific that the model like can counter check. And I think it's just genius that you have created this contact sheets first and then let the model kind of refine the graphics and until it hits the score basically. I think this is this is like incredible that you have shared this because I think for a lot of people this is absolutely valuable. And very glad that you yeah, didn't gate keep it but share it. And this is like the secret sauce. I mean >> Thank you. Thank you. Thank you.

>> But but but but people really don't understand this like you could gate keep this but you shared really the mental model. Amazing. Thanks. Thanks a lot. >> Yeah, I mean like I I guess if a quick shout out to I guess creative.ai on it's free. You can test it out. Like you can just generate some random images right now that you're not charged for it. Any feedback and I'm also open sourcing one project that you guys want like any of them I told there's four projects that I recreated in my channel. So if you want any of these open sourced just put a comment. I guess if there's more likes on one I can open source one of them like whichever one you want like there's the magnificent index fields chatbot and Adobe Firefly. So if you want to if you want me to open source any of them, I can one of them at least.

I guess like I want users to try on my platform. And if you have any questions about emotion video recreation kit like I can answer any questions in the comments as well.