Episode Book

EPISODE #6172

Recipes vs. Regeneration: AI Photo Editing in 2026

PUBLISHED	21 October 2026 at 13:54 UTC
DURATION	21:17
LISTEN	myweirdprompts.com/episode/ai-photo-editing-recipes-2026-06013/
HOSTS	Corn, Herman, Desmond, Hilbert
SCRIPT MODEL	anthropic/claude-sonnet-5.5
VOICES	chatterbox-regular (chatterbox-local)
PIPELINE	V5.11
BOOK GENERATED	22 October 2026 at 07:46 UTC

SYNOPSIS

Ask an AI to warm up your photo and one of two things happens: it hands you a recipe of exact slider values you can read and adjust, or it repaints every pixel and hopes you don't notice the logo grew a letter. This episode digs into the recipe camp — JarvisArt's agent over 200+ Lightroom tools, PhotoArtAgent's render-look-revise loop, and VeraRetouch's fully differentiable renderer — and the regeneration camp, where Photoroom's test of 850 products found the best instruction editor left the rest of the image untouched only 29% of the time. Plus why a recipe is reversible, resolution-independent, and batchable in a way a regenerated image can never be, and why a LoRA's single strength knob isn't the same thing as a tone curve.

TRANSCRIPT



Corn

...and that's what bugs me about the LoRA answer. You get one strength knob, and underneath it the model still repaints every pixel in the frame.



Herman

It's where everyone's hand goes first, though. Want a look, train a LoRA, ship it. I get the reflex.



Corn

Daniel's after what a photographer actually means by a preset. Exposure, white balance, a tone curve, a grade, a bit of denoise, all of it granular, none of it a regeneration. Which models do that in late twenty twenty-six, open or closed, and which of them hand you real controls instead of a vibe?



Herman

The nearest thing to a preset in the literature isn't a patch on a generator at all. It's a model that writes the recipe. Slider values, curve parameters, the numbers Lightroom would have taken from your own hands, and then an ordinary renderer applies them. You can read the recipe. You can change one number in it.



Corn

We did click-to-erase a while back, segmentation plus inpainting, and that always ended with a generator filling a hole. This is the other kind of edit, where nothing is invented.



Herman

The pixels come from your pixels. Take JarvisArt. It's a multimodal language model that acts as an agent over more than two hundred Lightroom tools. Two-stage training, first chain-of-thought supervised fine-tuning so it learns to reason about what the photo needs, then a reinforcement stage they call GRPO-R. It came out in June of last year and went into NeurIPS.



Corn

Two hundred tools. I use four.



Herman

You use four and you have still never found the right one. The paper also built its own benchmark, MMArt-Bench, and claims about sixty percent better pixel-level content fidelity than GPT four o. That's their number, against a model that regenerates the image, so I'd read it as "the recipe approach doesn't smear your picture", which is the point of it anyway.



Corn

Can you run it, though? Or is it a figure in an abstract?



Herman

There's a seven billion parameter version listed on a hosted platform from December. What I haven't checked is the licence or what base model it sits on, and the abstract page doesn't say. So I'd call it available to try, not cleared to build a business on.



Corn

Fair. What else writes recipes?



Herman

PhotoArtAgent is the one I find easiest to picture. A vision-language model looks at the photo, makes a plan, sends exact parameter values to Lightroom through its API, looks at what came back, and goes round again. And it explains its reasoning in plain text, which is the part a retoucher would want. The user studies in the paper say it beats automated tools and lands near human artists. I wouldn't lean on that last bit, it's their own study.



Corn

Of course it explains itself. It has to justify the warm shadows to somebody.



Herman

Then RetouchIQ, and I'm hedging here because all I've seen is the summary. Instruction-based retouching, the model emits executable Lightroom adjustments, trained with reinforcement learning against a generalist reward model, on a data set of roughly a hundred and ninety thousand instruction and reasoning pairs. It's a CVPR paper this year. Beyond that I'd be guessing.



Corn

And there's a personalised one, I thought. Region by region.



Herman

PerTouch, same story, I've only got the abstract. Personalised, region-level retouching, driven by an agent. I'd put it in the "watch this" pile.



Corn

Here's my doubt about all of them. They drive Lightroom. So I need Lightroom, and the model is a very fancy hand on somebody else's sliders.



Herman

You've put your finger on exactly what VeraRetouch argues. Its authors say an agent that talks to external software is stuck, because the software isn't differentiable. You can only reward the agent after the fact, you can't push a gradient back through the sliders. So they built their own renderer, fully differentiable, and a tiny model, half a billion parameters, that writes the plan as control latents. Lighting, global colour, specific colours.



Corn

Half a billion. That runs on a phone.



Herman

It's meant to. Mobile-sized is the pitch. They trained on a data set they call AetherRetouch, over a million images, the paper went up at the end of April and a second version in May. The paper says code and models are public. I haven't checked the code licence, and the paper itself is non-commercial.



Corn

So the recipe isn't Lightroom's recipe anymore. It's a private dialect.



Herman

Which is the trade. You gain a model small enough to carry and trainable end to end, and you lose the thing where your edit opens in the software you already own. Would you take that trade?



Corn

Depends who's holding the recipe. If I can read the numbers and drag them, I don't care whose renderer it is. If it's latents I can't read, I've bought a worse LoRA.



Herman

Fair, and I don't know yet how legible those latents are in practice. The paper says lighting, global colour, specific colours, so there's some structure to them. Whether you'd want to nudge one by hand, I can't say.



Corn

Why do I care about reading them, though? If the result looks right, I'm done.



Herman

Because a recipe has three properties a regenerated image can't. It's reversible, you drop one slider back to zero and the edit is gone, no history of what the model painted over. It's resolution independent, the same numbers work on a phone preview and a sixty-megapixel file. And it's batchable. Say you shot eight hundred frames at a wedding under one set of lights. You tune the recipe on three of them and run it on the rest, and every frame gets the same treatment.



Corn

That last one is the real preset. A preset is a promise that frame seven hundred looks like frame one.



Herman

Yes, and an instruction editor can't make that promise. Ask it for the same warm grade eight hundred times and you get eight hundred slightly different interpretations. Corn, you shoot a lot of the family stuff. What do you actually do when the light changes halfway through a roll?



Corn

Fix the first one by hand, copy it across, then grumble at the ones that went wrong. Which is a recipe, I suppose. I've been a recipe model this whole time.



Herman

A slow one.



Corn

A thorough one. Back to the recipe models. Say JarvisArt works as promised. What does a bad edit look like? With a generator I can see a melted face. With sliders, what's the failure?



Herman

Quieter. It crushes the shadows a touch, it pushes saturation past where you'd go, it picks a white balance that's technically neutral and emotionally wrong. The good news is that every one of those has a slider you can drag back. The bad edit is a few numbers off, not a different picture.



Corn

Then the review step is easier. I look at the numbers and the histogram, not at pixels for a dragon.



Herman

And the agents are built around that loop. PhotoArtAgent literally renders, looks, and revises. What nobody has told me is how often the revising converges to something a person would sign off on. Their user studies say close to human artists, but taste is the hard part, and a study with a handful of judges only goes so far.



Corn

Now the other camp. The ones that look at your picture and draw a new one.



Herman

The instruction editors. Nano Banana Two, Nano Banana Pro, GPT Image Two, Flux Two and its smaller Klein model, Qwen-Image-Edit. You say make it warmer and they re-render everything. Photoroom ran a test on product shots, eight hundred and fifty products at one K, and counted how often an edit left everything else untouched.



Corn

One K?



Herman

One K, I believe, in the post. The best of them, Nano Banana Two, came out at twenty-nine percent clean. Nano Banana Pro just under, GPT Image Two a little under that, and the Flux Klein at about seventeen.



Corn

Twenty-nine percent. That's the winner. A winner who changes your picture seven times in ten.



Herman

There's a reason for it, and I'm reasoning from how these systems are built, not from a paper. A diffusion editor doesn't move your pixels, it compresses the picture into a latent space, edits there, and decodes a fresh image. Even if the edit itself is perfect, the decode is a new rendering. Fine texture, small print, skin grain, all of it is reconstructed rather than carried over.



Corn

So asking for a hair more exposure is like asking a forger to make your painting a bit brighter.



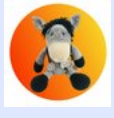
Herman

The forger is a decent picture, actually. On a good day the forgery is so close you can't see it, and on a bad day the logo has grown an extra letter.



Corn

An extra letter on the logo is how I lose a client.



Herman

Colour shifts in about eight percent of the generations, subtle but visible, and logos and text distorted in about one in five. Now, Photoroom sells a fidelity layer, so they have a stake, and the task was editing and restyling product images, not a pure slider move. Even



Herman (continued)

so, it matches what people running these say, that you ask for a little warmth and the grey card changes.



Corn

Is it only warmth? Because I can imagine a user asking for plus one stop and the model deciding what a stop means.



Herman

That's where CameraMaster comes in. It's still diffusion, but the exposure and white balance are numbers fed in as embeddings, trained on about seventy-eight thousand annotated pairs. The claim is that the response is close to linear, monotonic, and that parameters combine. And it says ordinary diffusion editors shift colour in non-monotonic ways when you change colour temperature, so you turn the dial one way and the cast wobbles both ways.



Corn

It's the nearest a generator gets to a slider. It's a slider on a painter.



Herman

Yes. Numeric control, but the pixels are still regenerated. There's also a paper on whether editors understand lighting, around a thousand pairs of high dynamic range light probes. From the search snippet, Nano Banana Pro ranked first and Qwen-Image-Edit third of six. I haven't confirmed that from the paper itself, so take it loosely.



Corn

And the LoRAs people actually reach for?



Herman

There are community relight LoRAs for Qwen-Image-Edit, lighting direction, colour temperature, intensity. They do work as a look. I haven't checked the weights or licences, and they're hosted-platform pages. But the point stands: a LoRA steers the generator. There's no tone curve in there to open.



Corn

I'm still not convinced the LoRA gets a fair hearing. For a look, a film stock, a mood, it does the job.



Herman

For a look it does, and I'd never argue otherwise. Where it stops is the knob. A LoRA has a strength, one number, and the strength doesn't mean "plus a third of a stop". It means "more of whatever this was trained on". You can't tell it to leave the sky alone and lift the shadows.



Corn

So the LoRA is a preset the way a cover band is the original.



Herman

It has the same tunes. You can't ask it to change the key.



Corn

We now join Desmond for his regular segment.

D

Desmond

Desmond's Money Tip. Cut the end off the toothpaste tube. There is about a week of paste in the shoulders, and over a year that is roughly one and a half tubes saved, which at two pounds a tube is three pounds a year, or two hundred and forty pounds across a lifetime of eighty years.

D

Desmond

That is not nothing. That is two hundred and forty pounds, spread thin across a very long time, which is how all money actually arrives.

D

Desmond

I keep a pair of scissors in the bathroom cabinet for exactly this, and my wife has asked me to stop, and I have not stopped, because the tube is not finished when the tube says it is finished.

D

Desmond

Money is mostly just refusing to let things end early.



Corn

I have scissors somewhere.



Herman

Corn, no. That man has been wrong in public for years and he's sharing it with a microphone.



Corn

Back to pixels, then. Is there a cleaner version of a preset than a recipe? A LUT is literally a colour function.



Herman

A LUT is the plain preset, yes. A lookup table that maps every input colour to an output colour. The research that interests me is NILUT, a neural implicit LUT. A small multilayer network standing in for a three D LUT, continuous instead of a grid, and it can emulate real LUTs and blend several styles inside one little network. AAAI two years ago, from Conde and colleagues. I haven't found a big follow-up since.



Corn

So "apply this look" has a solution that fits in a kilobyte and never invents a pixel.



Herman

Nothing invented, nothing to smear, and you can interpolate between two grades. For a colourist that's closer to the real thing than anything diffusion offers. What it can't do is look at your photo and decide what it needs. A LUT has no idea what's in the picture. Apply a warm film grade to a shot that already has a colour cast and you get a worse cast, with total confidence.



Corn

Which is where I'd put an agent. Not to paint, to choose. Look at the shot, pick the grade, set the strength, show me the numbers.



Herman

That's the combination I'd bet on. A model for judgement, a deterministic renderer for the pixels. Nobody has shipped it as one tidy open package that I know of, as of today. You'd be stitching a recipe model to a renderer yourself. The split I'd draw: if you want a decision made for you, with the working shown, that's the agents. If you want a look applied exactly, that's a LUT, learned or otherwise. If you want something new in the frame, only then the generators.



Corn

And Adobe, who owns the sliders to begin with?



Herman

Adobe's putting a language front end on the slider stack. There's a Photoshop assistant, and a Firefly assistant described as acting across Photoshop, Lightroom and other apps, with an example prompt like brighten the subject, blur the background slightly, make the tones warmer. Whether the Lightroom side of that actually edits through the sliders, I haven't seen confirmed anywhere. The Lightroom Classic update in August is quieter: feather and edge sliders for the AI masks, render to DNG.



Corn

Quiet is when they're actually doing something.



Herman

There's also the old guard that's gone neural. DxO released PureRAW six in February, with its DeepPRIME XD three denoiser extended to Bayer sensors and smaller DNG output. Topaz renamed Photo AI to Topaz Photo and keeps adding tools, a noise-aware sharpen, a grain tool. Those are trained for one job, denoise or sharpen, and you get a strength slider. That is a preset, in the sense that it's one dial and the same result every time.



Corn

Same result every time sounds nice after twenty-nine percent.



Herman

Which tool would you actually reach for, Corn? Say you've a folder of two hundred raw files from a dim room.



Corn

Denoiser first, on the raw files, before anything else touches them. Then a LUT for the look. An agent if I'm feeling lazy, a generator never.



Herman

Never is strong. What about a distracting sign in the background?



Corn

That's the click-to-erase job, and I'd do that last, on one frame, and look at it hard. I'm not against generators. I just want them where invention is the job.



Herman

Same place for me. Though I'd love to see someone run a pure exposure move across those editors, one stop up and nothing else, because I haven't seen that test anywhere.



Hilbert

Two K. They ran the products at two K, Herman, not one. I only mention it because I read the card before I look at the picture, and Gerald reads the figures aloud, and he'd stop at one K. The Society of Tonal Honesty has fourteen members and nine of them are dead. The nine still vote. Gerald Pomfret holds their proxies in a Huntley and Palmers tin. Every first Tuesday the motion is that the bulb in the meeting room counts as reference white. The dead say yes, the living say no, nine to five, carried, and nothing happens, because the switch is on the chiropodist's side of the wall.



Herman

A filament lamp that small is very warm light, so anything you call white under it is going to come out orange on a card. Nine members voting for orange, then.



Hilbert

They voted for the bulb, Herman. You keep saying recipe. Ours is a card. Nine numbers and a date, written in pencil. Gerald has copied about four thousand of them by hand. Not once has he written down the bulb. Tuesday I brought the receipt. Hartley's, Holloway Road, six pounds forty. It says forty watt. The minutes have said sixty for years.



Corn

So the dead have been ratifying a bulb that never existed.



Hilbert

Mister Alderton objected. He is one of the nine. Gerald objected on his behalf and then seconded himself. The bulb itself went in the bin before I joined, so there is the receipt and there is nothing else. I have the receipt in a plastic sleeve. I find it more reliable than the bulb was.



Herman

Forty against sixty is a real difference in brightness. Every exposure number on every card he's copied is relative to a lamp nobody measured, in a room with no window.



Hilbert

The room has a window. It looks onto the chiropodist's. Gerald said the cards are still correct as cards, and the living agreed, five to nothing, with the dead abstaining. His brother-in-law wired that room and he is the only man who ever read the plate on the fuse box. He says it is rated for a sixty. Nobody has asked what that is for.



Corn

The fuse box says sixty and the receipt says forty. I'd have both framed.



Hilbert

The receipt is framed. The fuse box is on the chiropodist's side.



Herman

Rated for a sixty and wired by a brother-in-law. I'd want a light meter in that room before I trusted a single card.



Hilbert

Gerald has copied every card by hand and he has never got a number wrong. That is why people still turn up.



Herman

The number nobody writes on the card is the thing it was measured against. For the recipe models I'd want that attached: which renderer, which version, which colour profile the numbers assume. A recipe you can read is only worth something if you can tell what it was cooked in.



Corn

And if the renderer changes under you, the recipe changes with it. That's the Lightroom-agent problem and the VeraRetouch argument in one, and I still don't know which side I'd bet on. I'd probably start with a LUT and a denoiser I trust, and let an agent suggest the numbers.



Herman

That's a sensible kit for now. Thanks to our producer, Hilbert Flumingtop, for keeping it all on the rails.



Corn

This has been My Weird Prompts. If you want more, the website is my weird prompts dot com.



Herman

See you tomorrow.