Computational Photography Lab.

Find the project web page here:
https://yaksoy.github.io/VETI/

VETI: 3D-Guided Visual Effect Tuning on Images

Mingrui Zhao*
Simon Fraser University
Canada
mza143@sfu.ca

Duc Anh Nguyen*
Simon Fraser University
Canada
dan5@sfu.ca

Gnanavel Premnath
Simon Fraser University
Canada
gpa21@sfu.ca

S. Mahdi H. Miangoleh
Simon Fraser University
Canada
smh31@sfu.ca

Sebastian Dille
Simon Fraser University
Canada
sdille@sfu.ca

Yağız Aksoy
Simon Fraser University
Canada
yagiz@sfu.ca

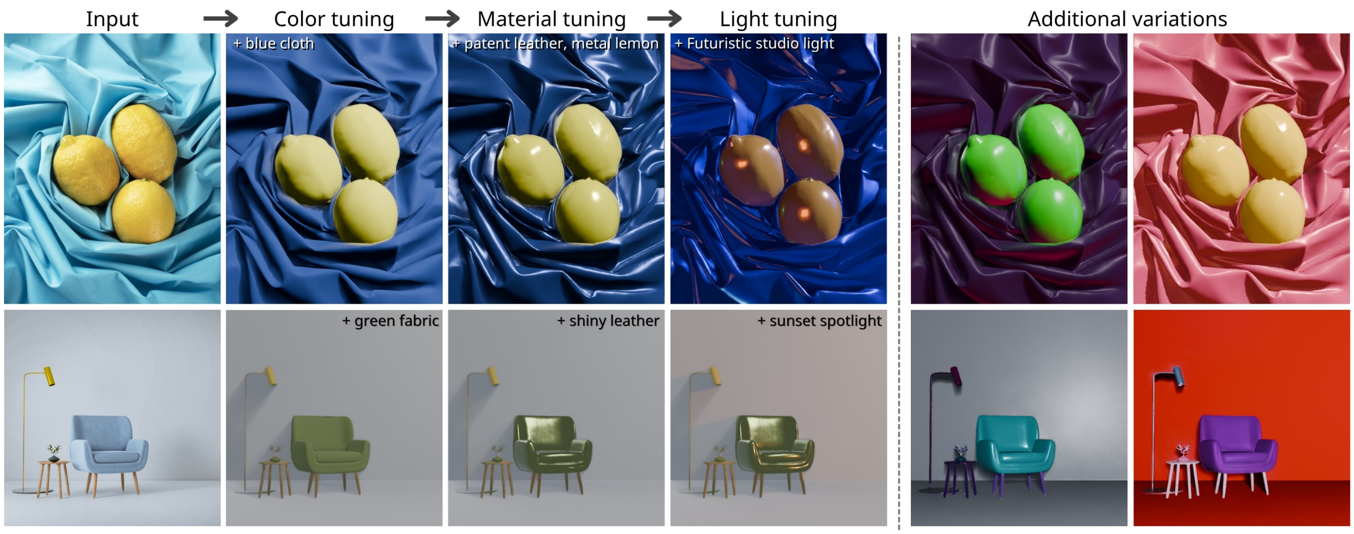

Figure 1: We propose an interactive photo editing toolbox within a physically based rendering engine, that allows *color tuning*, *Material tuning* and *Light tuning* allowing for controllable and physically meaningful edits. We show additional results on the right.

Image credits: @Magnific, @Virender Singh

ACM Reference Format:
Mingrui Zhao, Duc Anh Nguyen, Gnanavel Premnath, S. Mahdi, H. Miangoleh, Sebastian Dille, and Yağız Aksoy. 2026. VETI: 3D-Guided Visual Effect Tuning on Images. In *SIGGRAPH Asia 2026 Posters (SA Posters '26), December 01–04, 2026, Kuala Lumpur, Malaysia.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3829333.3847906

Photo editing is tied to the physical properties of the scene. Everyday tasks such as recoloring a piece of furniture, changing a product's finish, or shifting the key light involve the complex, real-world interplay of light, material, and color. For convincing results, both the physical plausibility of these parameters and the original scene content have to be preserved.

We present VETI, a framework that offers physically motivated image editing while remaining fully anchored to the input photograph. VETI poses photo editing as an interactive process within a rendering engine. It provides direct access to the physical properties of the scene by grounding lighting, object materials, and color palettes in an actual 3D representation of the image. Artists can edit real-world photographs with powerful tools typically reserved for synthetic assets, such as palette swatches, BSDF channels, and light rigs. Reflections and illumination changes are rendered realistically, replacing tedious painting of the approximated effects.

Given a single image, our framework assembles a 3D scene in which every visual property (reflectance, light and material) is readily available for interactive edits. As shown in Figure 1, users can: (i) **re-color** the scene's color pallette; (ii) **re-material** each objects by tuning physically-based surface properties (roughness, metallic, specular); and (iii) **re-light** the scene with any type of light source.

*Equal contribution

SA Posters '26, Kuala Lumpur, Malaysia

ACM ISBN 979-8-4007-2839-6/2026/12
https://doi.org/10.1145/3829333.3847906

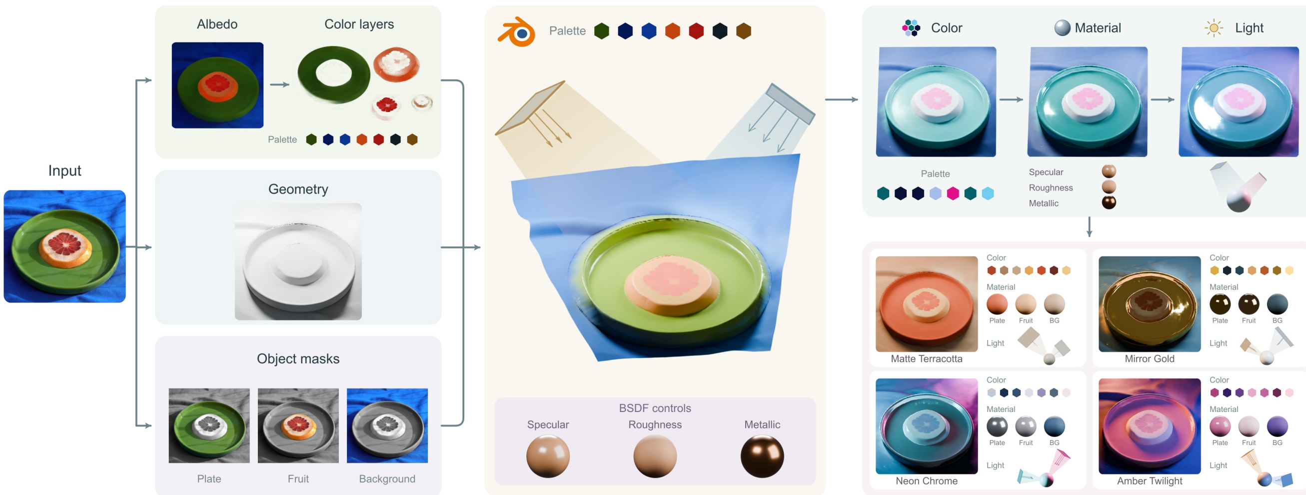

Figure 2: Our pipeline: intrinsic decomposition [Careaga and Aksoy 2024] followed by soft color segmentation [Aksoy et al. 2017] gives a palette and per-pixel alpha maps; MoGe [Wang et al. 2025] estimates geometry; and SAM 2 [Ravi et al. 2025] produces per-object masks. Then a Blender scene with Principled BSDF shader graph that exposes color, material, and light controls is created. Each component is then exposed for the user to edit. Image credits: @Mustafa akın

By keeping light, material, and color strictly disentangled, VETI ensures every adjustment is physically plausible and precisely under the user's control. This enables a diverse range of complex edits and stylization, as we demonstrate in Figure 1 and 3. With its flexibility, VETI gives artists an important advantage over traditional pixel-level algorithms, which are limited to low-level operations and demand excessive manual labor, or generative editors that sacrifice fine-grained, predictable control for indirect influence. Our implementation is available in the supplementary material, ready for testing and experimentation.

1 Building an Editable Scene

From an input photograph, our pipeline creates a structured Blender scene [Blender Foundation 2025] in which color, material, and lighting can be manipulated meaningfully and independently. Figure 2 summarizes the four stages of our pipeline, and intermediate results for each stage are visualized in Figure 1 and 3.

Re-coloring. To enable intuitive and global image recoloring, we extract a color palette using soft color segmentation. Our pipeline first isolates the diffuse surface reflectance from the input photograph using the off-the-shelf intrinsic decomposition method of Careaga and Aksoy [2024]. Removing shadows and highlights ensures that our subsequent color analysis operates purely on true surface albedo rather than observed illumination. We then apply the sparse color unmixing framework of Aksoy et al. [2017] to decompose this albedo into N distinct palette swatches, each paired with a per-pixel alpha map. Users can modify individual swatches while the soft alpha masks guarantee smooth transitions.

Material adjustment. To enable material editing, we lift the photograph into 3D by reconstructing a pixel-aligned mesh from the input image using MoGe-2 [Wang et al. 2025]. We texturize the mesh with the colors generated from the previous stage. We show the pixel-aligned 3D reconstructions before any edits in the second column of Figure 3. We use SAM 2 [Ravi et al. 2025] for object segmentation of the image and then assign principled BSDF parameters on a per-object basis. This allows for object-specific material adjustment. Users can modify the surface structure of individual objects and, for example, turn the matte cloth of the chair into shiny leather without affecting the matte wall in Figure 1. Our lifted 3D representation creates realistic reflections for the geometry without having to paint highlights manually (e.g shiny stones and plates in Figure 3).

Re-lighting. The reconstruction in Blender further allows for intuitive light setting in 3D. Users can add, move, and adjust light sources with the software interface, using the established tools for full 3D scenes. The reconstructed geometry and principled BSDFs will faithfully reflect the changing light in the rendered scene.

Tuning. Once the scene is assembled, users can edit parameters manually for fine-grained creative control. Tunable parameters include the N swatch colors, per-object roughness, metallic and specular values, and the position, color and intensity of each light. Individual swatches can be adjusted for fine-grained creative control, or the whole palette replaced at once with a harmonic set for global style tuning. Materials can be edited per-object. Lights can be added, moved, and animated, or an HDRI loaded for environment lighting.

Alternatively, users can connect VETI to use an LLM control. The LLM takes in the assembled blender file, the initial render, and a prompt of desired style (e.g "Gothic"), and returns the full list of edited values. In both cases, only the attribute values change and the photograph itself is never resynthesised. Figure 3 shows the resembled 3D geometry, progressively edited results across all three stages, and the LLM-prompted bundle edits. Both tuning

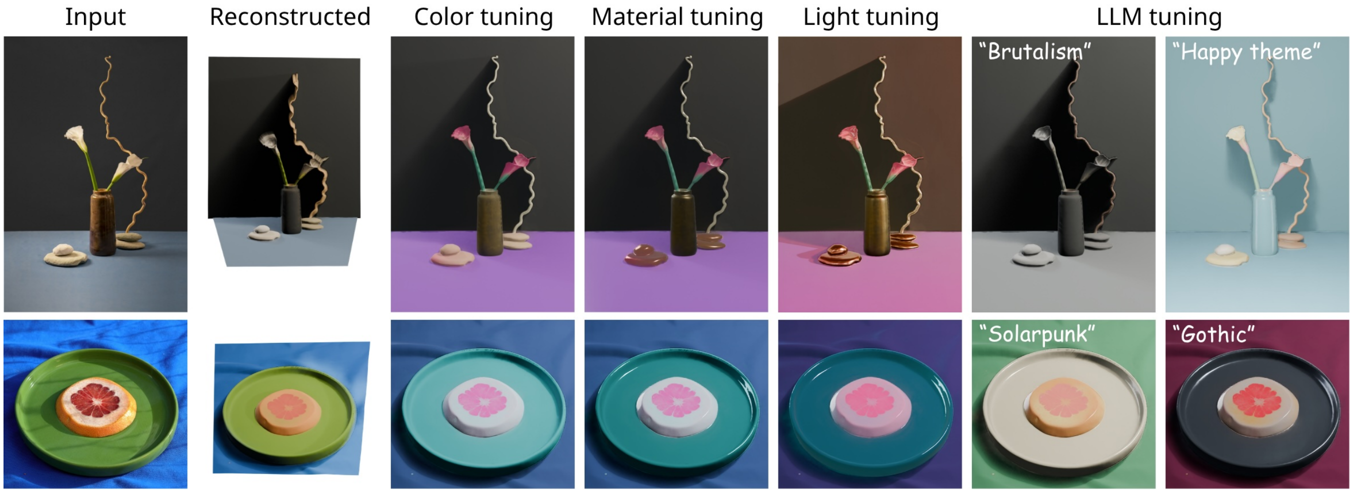

Figure 3: Given an image, we first assemble a 3D Blender scene which we then edit iteratively. color palette, per-object BSDF material , and lighting. Our pipeline also accepts LLM-instructed edits. Given a natural-language prompt describing a target style, an LLM returns the full parameters list to be applied to the scene. **Image credits: @Magnific, @Mustafa akın**

modes compound light, material, and color effects that are difficult to combine in 2D but stay disentangled with VETI.

2 Conclusion

We presented VETI, a pipeline that turns a single photograph into an editable 3D scene in which color, material, and light are independent tunable controls. Limitations include realism at occlusion edges: the reconstruction is only 2.5D, leaving surfaces to stretch where nothing was observed behind them. Future work will bridge this gap and pursue more photorealistic rendering.

References

Yağız Aksoy, Tunç Ozan Aydın, Aljoša Smolić, and Marc Pollefeys. 2017. Unmixing-Based Soft Color Segmentation for Image Manipulation. *ACM Trans. Graph.* (2017).

Blender Foundation. 2025. Blender – a 3D Modelling and Rendering Package. https://www.blender.org.

Chris Careaga and Yağız Aksoy. 2024. Colorful Diffuse Intrinsic Image Decomposition in the Wild. *ACM Trans. Graph.* 43, 6, Article 178 (2024), 12 pages.

Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson, Eric Mintun, Junting Pan, Kalyan Vasudev Alwala, Nicolas Carion, Chao-Yuan Wu, Ross Girshick, Piotr Dollar, and Christoph Feichtenhofer. 2025. SAM 2: Segment Anything in Images and Videos. In *Proc. ICLR.*

Ruicheng Wang, Sicheng Xu, Yue Dong, Yu Deng, Jianfeng Xiang, Zelong Lv, Guangzhong Sun, Xin Tong, and Jiaolong Yang. 2025. Moge-2: Accurate monocular geometry with metric scale and sharp details. In *Proc. NeurIPS.*