FDS: Frequency-Aware Denoising Score for Text-Guided Latent Diffusion Image Editing

Supplementary Material

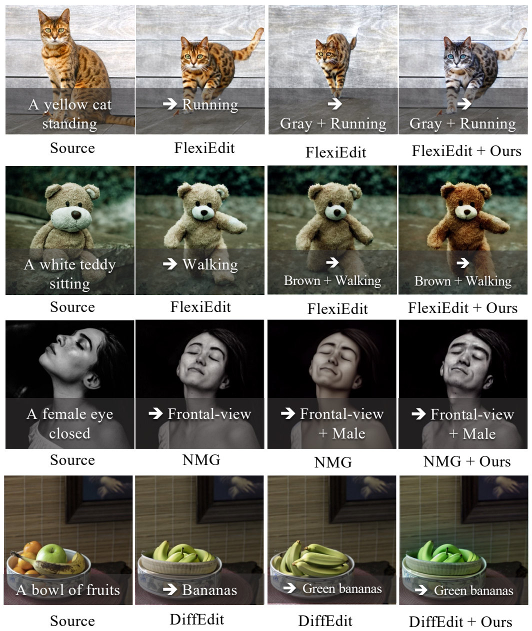

Figure 9. **Our method complements other editing methods.**.

Combining our method with other editing methods. As shown in Fig. 9, our method can be combined with other editing methods, i.e., FlexiEdit [25], NMG [7], and DiffEdit [10], to enable complex editing tasks.

Combining Ours with CDS. Our wavelet representation can also be integrated with CDS (see Fig. 19 and Fig. 10 (a)). In Fig. 19, both CDS and DDS fail to preserve the detailed latte art pattern across various hyperparameters and random seeds. However, when combined with our wavelet representation, both DDS+Ours and CDS+Ours demonstrate improved texture preservation. Fig. 10 (a) further highlights our color consistency benefits.

Challenging and Failure Cases. Fig. 10 (b) illustrates a challenging 3D texture editing scenario where the crystal exhibits severe self-occlusions. In this case, our method successfully changes its color despite the complexity. However, certain edits (see Fig. 10 (d)) are not physically plausible, and our method encounters difficulties when handling non-rigid editing scenarios.

Automatic Frequency Selection. Fig. 10 (c) demonstrates a pipeline that leverages a Large Vision Language Model to automatically select the appropriate frequency subband for editing based on the given prompt.

Additional Results in Detail Preservation. As shown in Fig. 11 and Fig. 12, our method preserves high-frequency details, such as intricate wood carvings, thin elements, and textures.

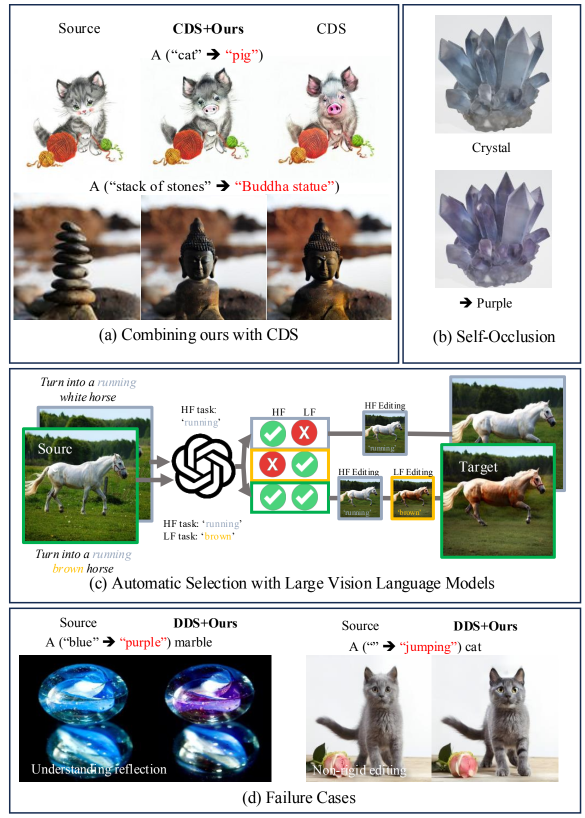

Figure 10. **Additional visualizations.**

Additional Results in Color Preservation. As shown in Fig. 12, baseline methods often fail to maintain source image colors (e.g., orange tones and cat colors), whereas our method achieves better color consistency compared to DDS and CDS.

Additional Texture Editing Results. We provide additional texture editing results in Fig. 14, Fig. 15, and Fig. 16. Compared to SDS, SDS+Ours achieves better detail preservation, such as the patterns on an owl's and a chicken's feathers, as well as improved color consistency, such as the color of a sofa. A video is attached to this supplementary material for better viewing quality.

Additional Gradient Visualizations. We provide additional gradient visualizations for both detail preservation and color fidelity cases in Fig. 17 and Fig. 18, respectively.

Figure 11. **Qualitative results on detail preservation.** Our method preserves detailed structures during editing, such as the intricate carvings in wood (rows 1 and 5) and small details like the petiole and leaf veins (last row). In contrast, CDS does not follow instructions well, and other diffusion sampling-based methods either distort the details or fail to adhere to instructions. (Best viewed on a screen when zoomed in)

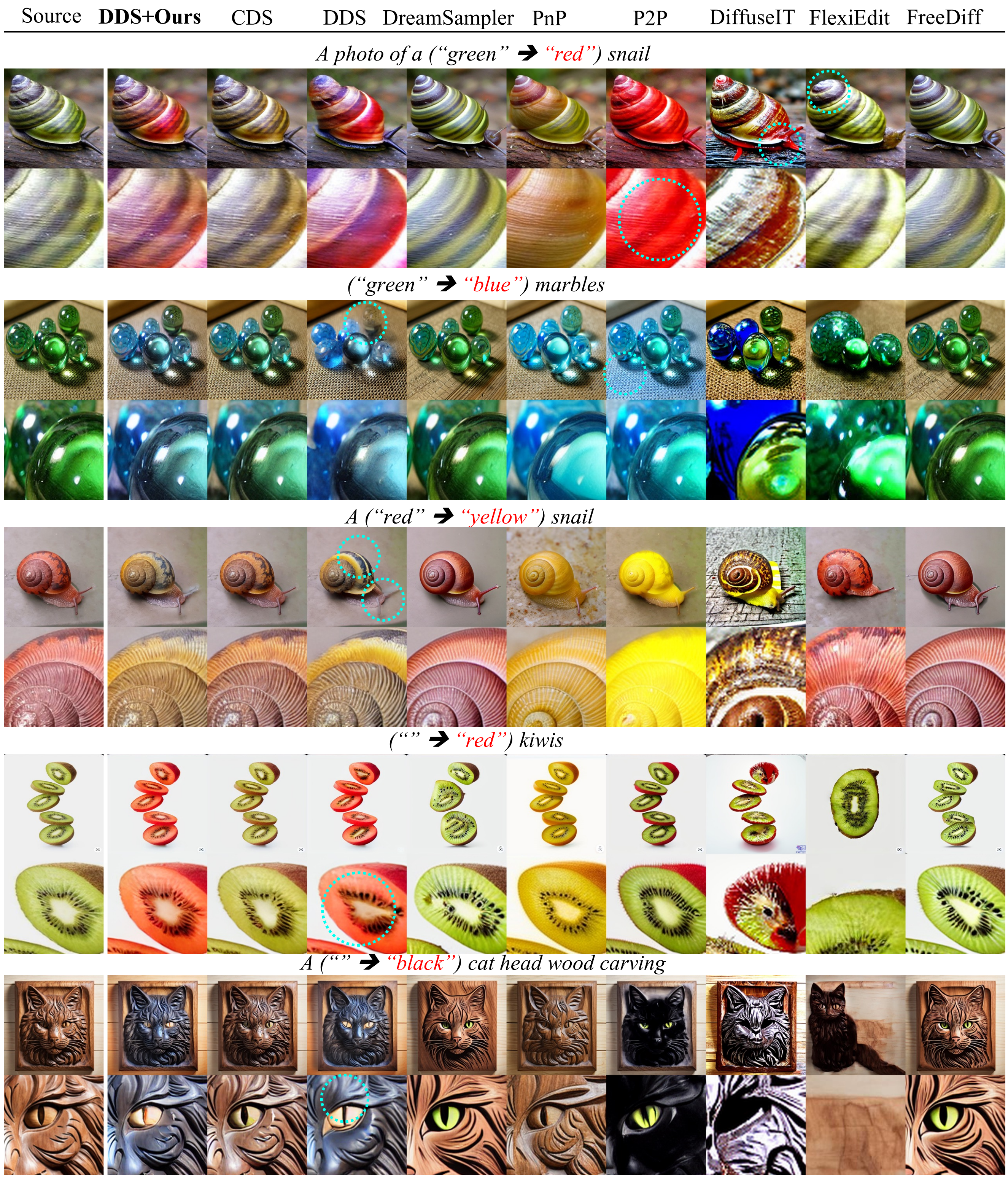

Figure 12. **Qualitative results on detail preservation.** Our method effectively preserves detailed textures, such as the texture on a snail's shell and the reflection highlights on marbles. CDS, FlexiEdit, and FreeDiff do not follow instructions well, and other methods distort local details. (Best viewed on a screen when zoomed in)

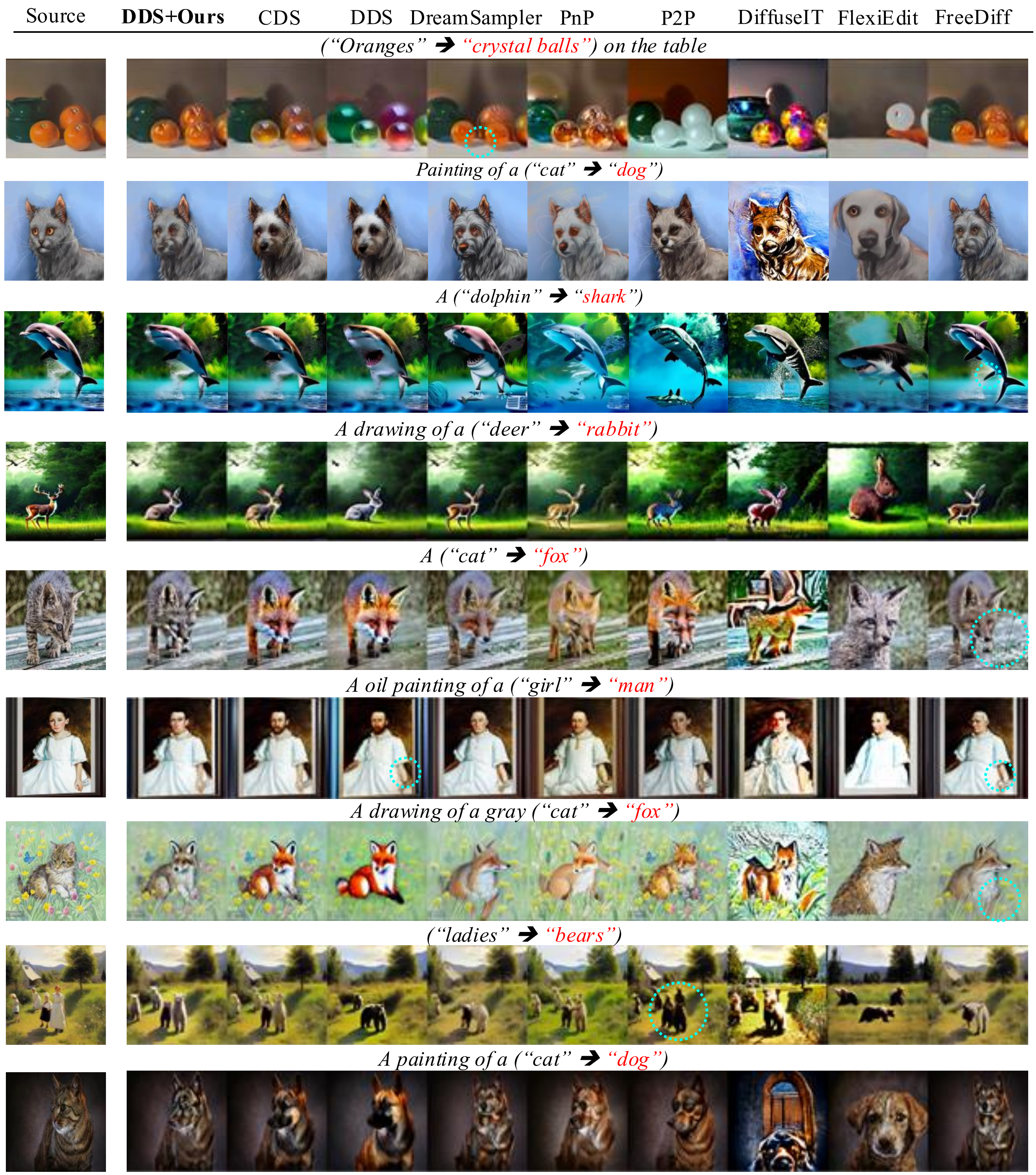

Figure 13. **Qualitative results on color preservation.** Our method maintains the colors of the source image, such as the orange color of oranges and the fur color of animals, resulting in edits that not only follow the editing instructions but also appear more consistent with the source image. While FreeDiff also preserves colors, it fails to maintain geometric structures, such as altering the head position when turning a cat into a fox. (Best viewed on a screen when zoomed in)

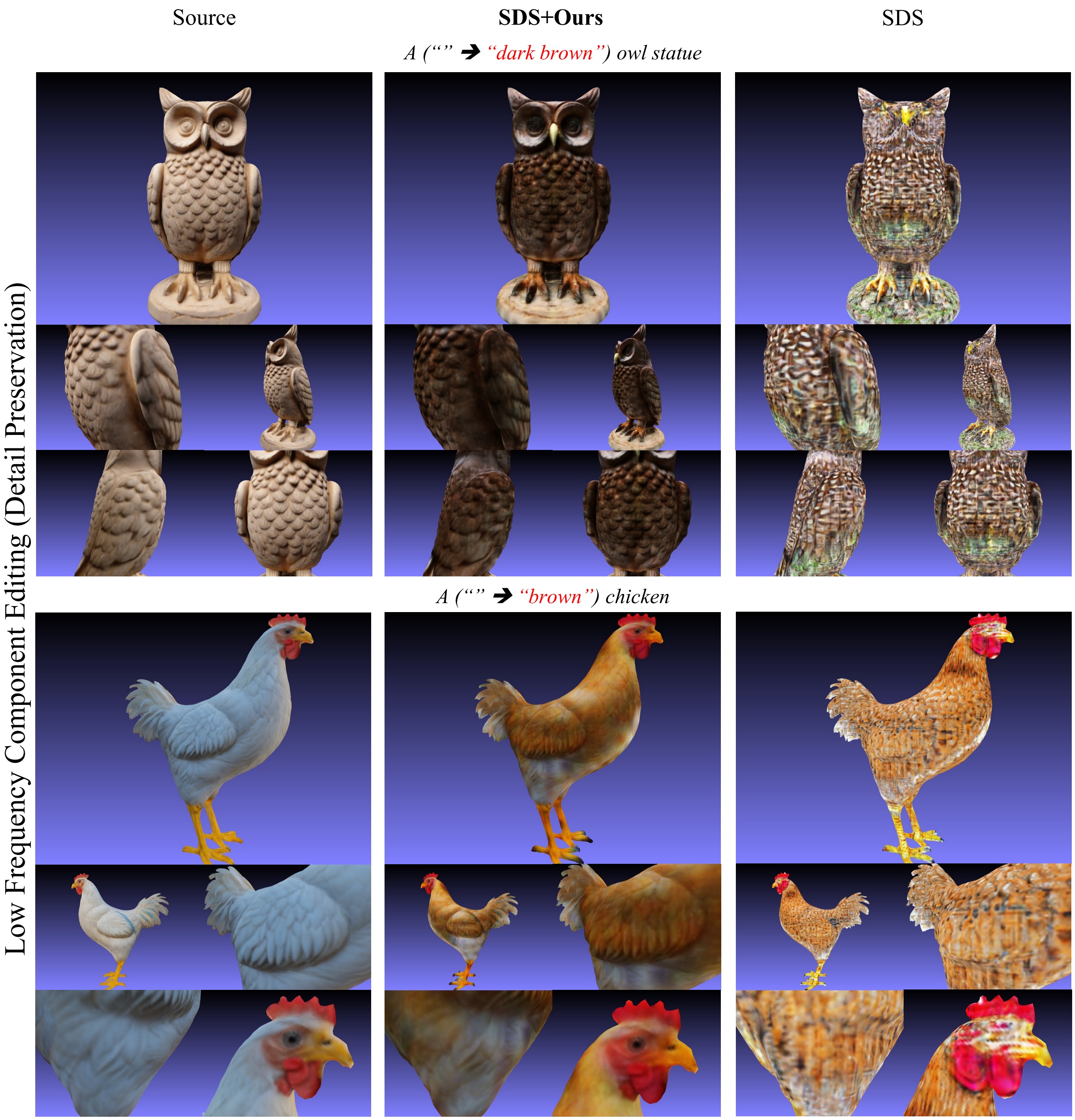

Figure 14. **Qualitative results on texture low frequency editing.** Compared to vanilla SDS, which largely ignores the detailed patterns of the original texture, SDS+Ours produces edits that preserve these details, such as the feathers of a chicken. Notably, SDS nearly removes the eyes of the owl and the chicken, as they are high-frequency details, while our method preserves them well. (Best viewed on a screen when zoomed in and also see in attached videos)

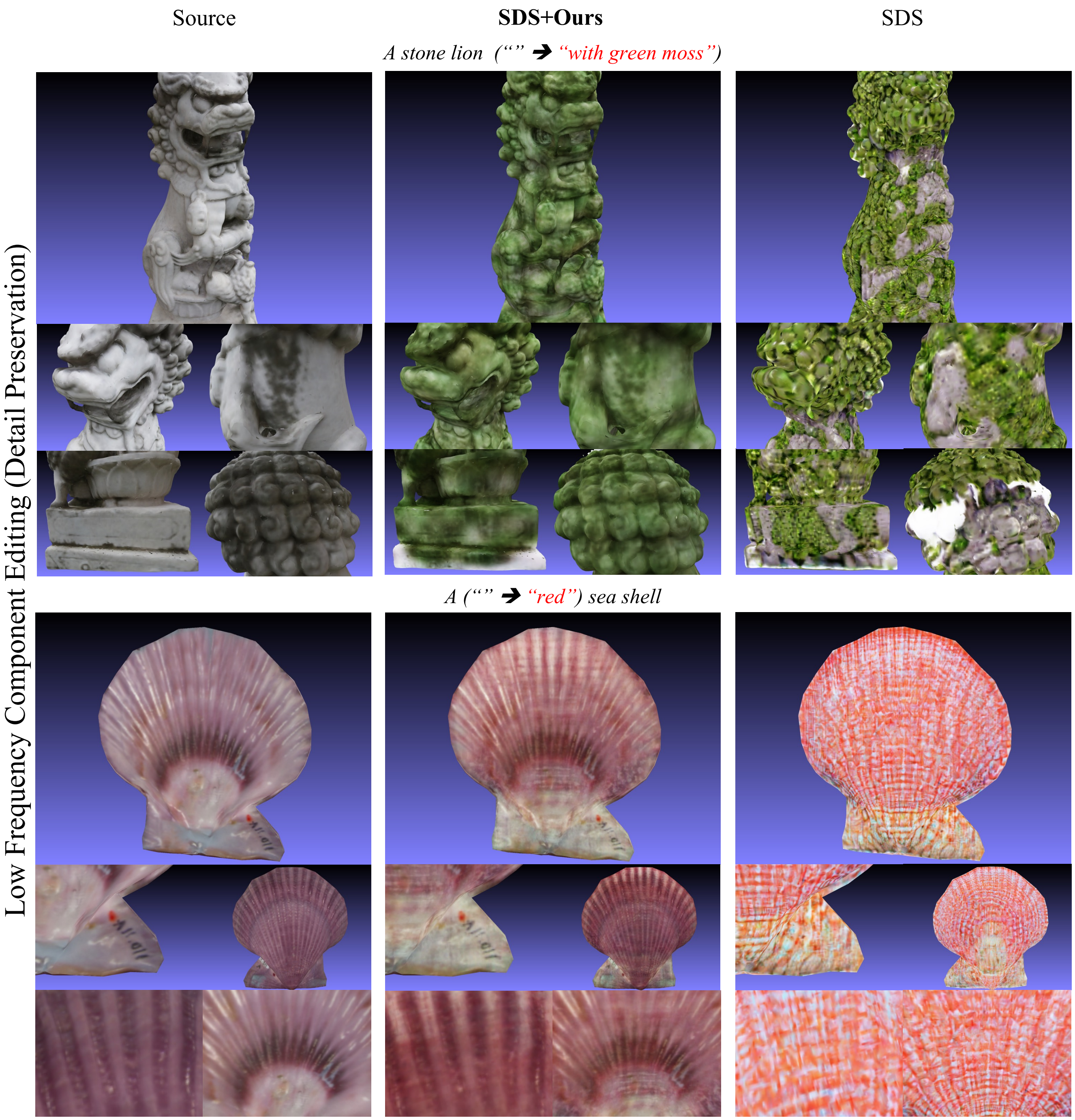

Figure 15. **Qualitative results on texture low frequency editing.** Our method preserves high-frequency details, such as the intricate texture on the shell and the fine lines on the stone. In contrast, SDS alone drastically alters these details, resulting in unrealistic-looking objects (as in the stone lion case) or entirely different objects (as in the sea shell case). (Best viewed on a screen when zoomed in and also see in attached videos)

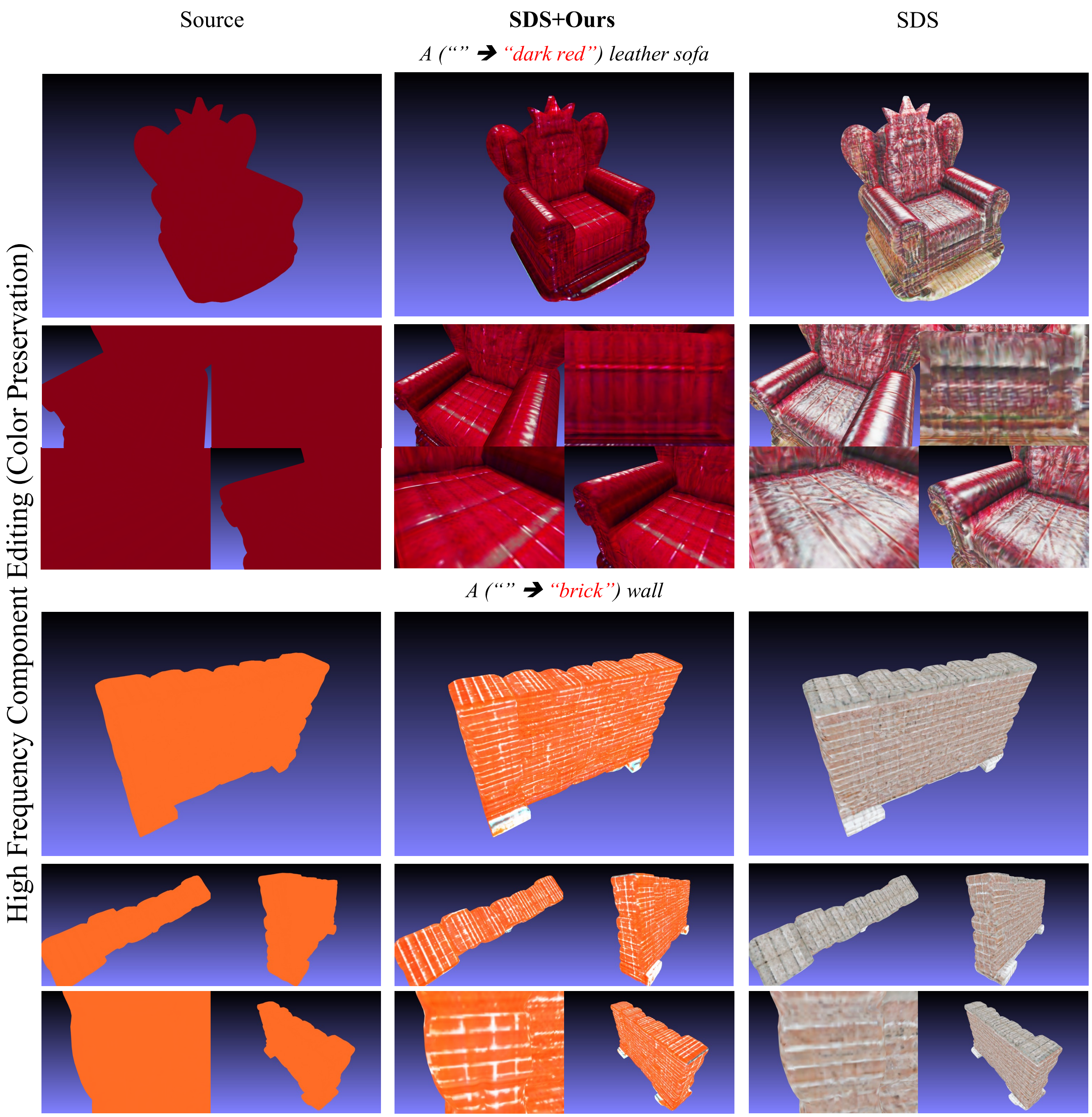

Figure 16. **Qualitative results on texture high frequency editing.** The goal of high-frequency texture editing is to make modifications while preserving the original object's color. As shown, SDS drastically alters the texture, almost entirely disregarding the original texture map. In contrast, our method adjusts high-frequency details while maintaining the original color of the sofa and the wall. (Best viewed on a screen when zoomed in and also see in attached videos)

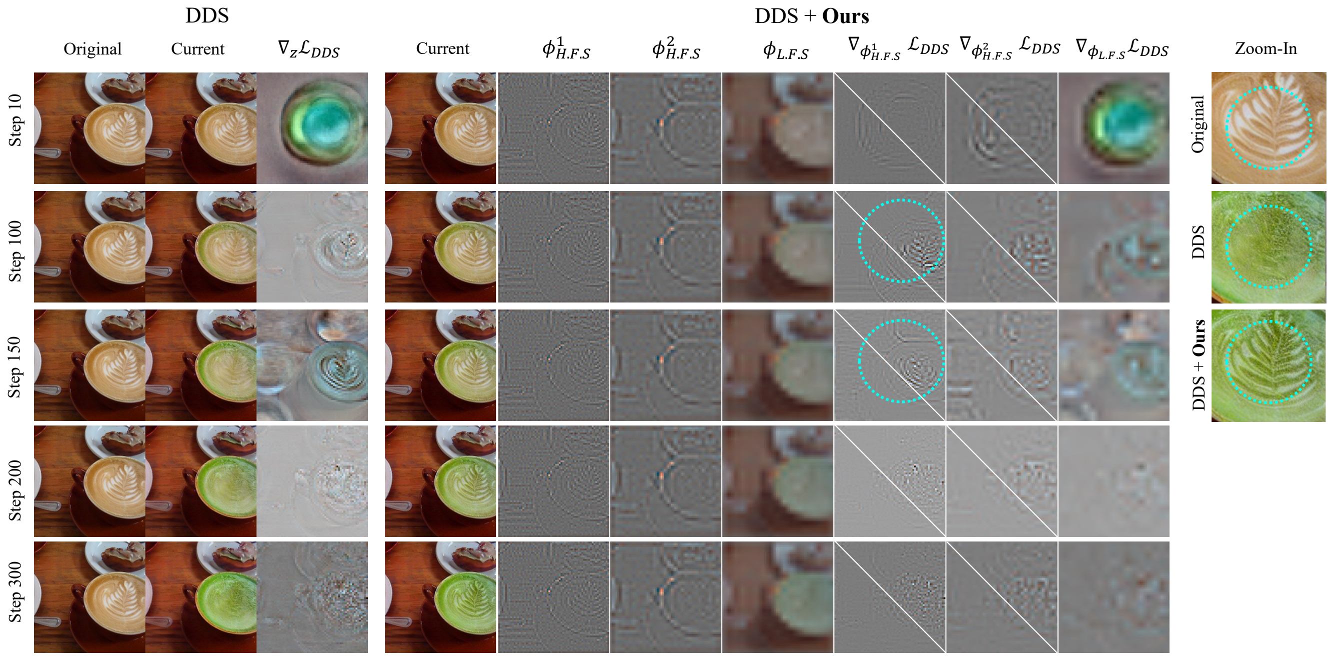

Figure 17. **Additional gradient visualization during optimization.** Edit with "A cup of (coffee → matcha)". The high frequency gradient $\phi^1_{\text{H.F.S.}}$ and $\phi^2_{\text{H.F.S.}}$ in DDS distort the detail pattern of latte art. With frequency awareness, our method is able to preserve these details.

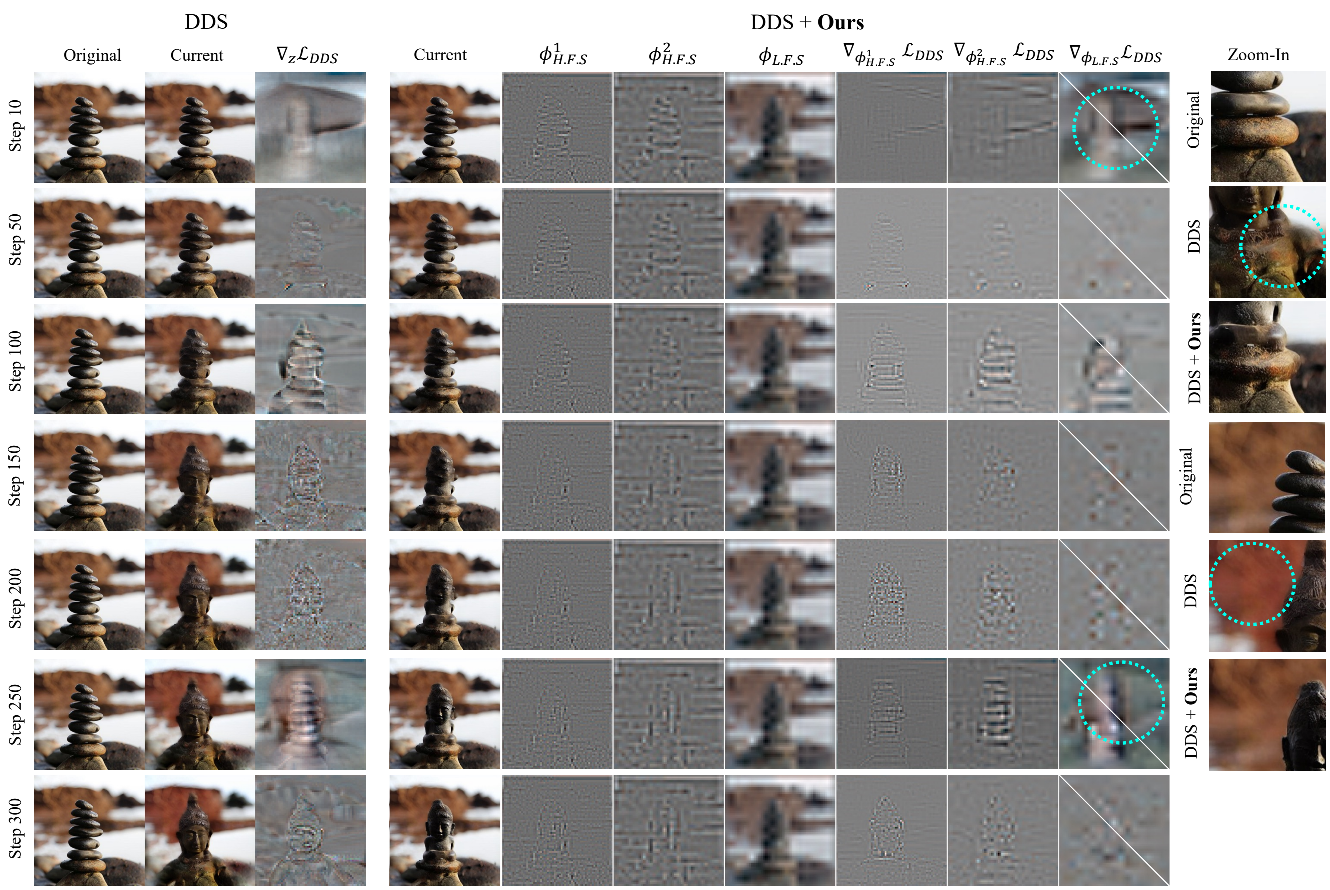

Figure 18. **Additional gradient visualization during optimization**. Edit with "A (stack of stones → Buddha statue)". The low frequency component of $\phi_{\text{L.F.S.}}$ is changed modified and thus the color of the stone is turned into yellowish. (Best viewed on a screen when zoomed in)

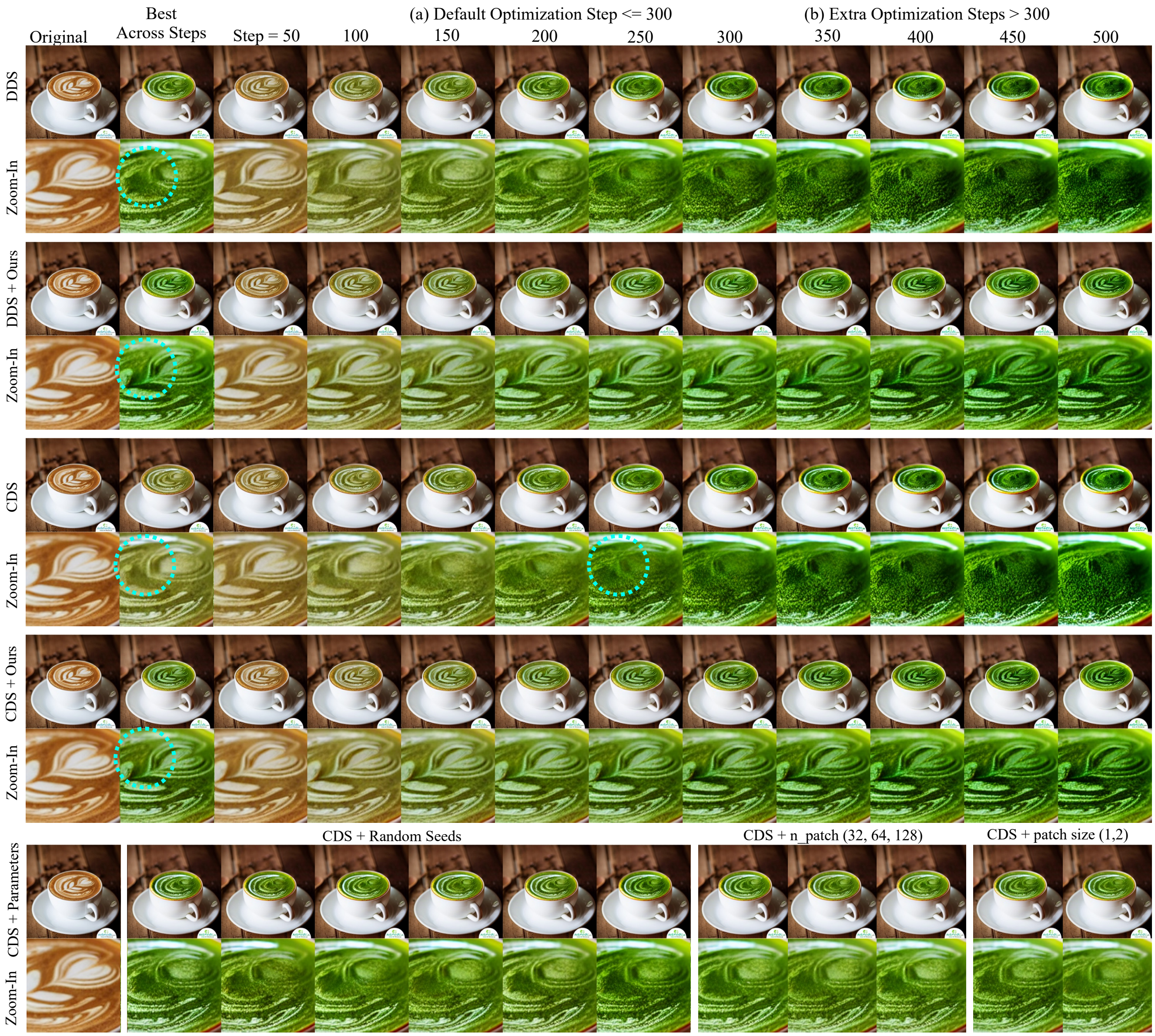

Figure 19. **Combining CDS with our wavelet representation.** Neither vanilla DDS nor vanilla SDS successfully preserves the intricate patterns of latte art. Notably, the distortion of details occurs simultaneously with the color change to green. By incorporating our frequency-aware representation, we achieve superior detail preservation, even when the number of steps is significantly increased, such as to 500. Additionally, we provide further results demonstrating the performance of CDS under varying hyperparameters and random seeds. (Best viewed on a screen when zoomed in)