Diffusion-based Compression Quality Tradeoffs without Retraining

Supplementary Material

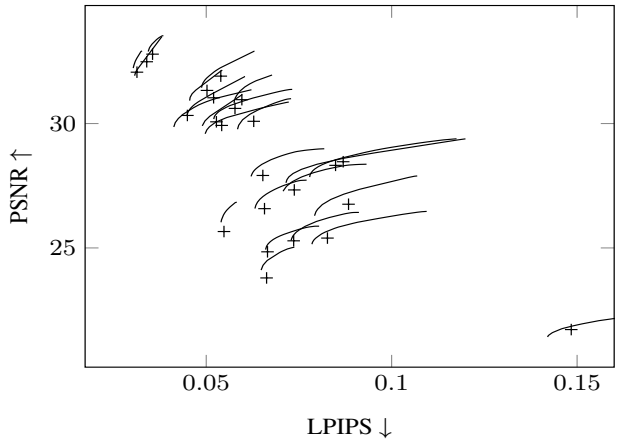


Figure 12. **Optimal configurations for CDC x_0 [40]** according to $\mathcal{T}$ with $\alpha \in [0, 1]$ and the corresponding default configuration for the 24 Kodak images. With the default configuration marked, and the best possible results shown as a black line.

A. Potential Quality Improvements on Kodak

In addition to Fig. 3, which only shows the first Kodak image, we provide results for all of the 24 Kodak images in Fig. 12, which shows the best configurations according to $\mathcal{T}$ for any α, as determined by the grid search, and the corresponding default configuration of the CDC x_0 model [40].

As discussed in Sec. 4.2, we find that it is possible to achieve performance improvements for any of the tested images. The achievable improvement strongly depends on the specific image and targeted metric. In many cases, the default configuration of CDC x_0 [40] already scores quite well on LPIPS, and only improvements in PSNR are possible.

B. Additional visual comparisons

To expand on Fig. 2, we provide a larger and more detailed version of the figure in Fig. 13.

In Fig. 14 we provide a visual overview of how the tradeoff parameter α affects the generated image for $0.9 \leq \alpha \leq 1.0$. Values $\alpha < 0.9$ lead to the same configuration during the optimization for the selected Kodak image and CDC checkpoint and will therefore generate the same image.

We provide additional visual comparisons for images generated with PerCo [12] (Fig. 15), as well as the lowest bitrate checkpoint of CDC x_0 [40] (Figs. 16 to 21). We show a selection of cropped reconstructions from the Kodak [15] dataset optimized for different tradeoffs (α).

As mentioned in the original PerCo paper, PSNR is not a suitable metric for very low bitrate scenarios, where realism is most important. Therefore, as expected, optimizing for PSNR does not yield visually pleasing results [12]. The images with $\alpha = 0$ are usually very washed out and lack contrast. Optimizing for LPIPS is much more reasonable in this case. However, other optimization targets might prove to be better suitable for this model.

For CDC, optimizing for PSNR with $\alpha = 0$ generally results in very smooth images, often lacking some details. When increasing $\alpha \to 1$ the model will generate sharper and more detailed images. However, this can sometimes lead to slightly noisy images.

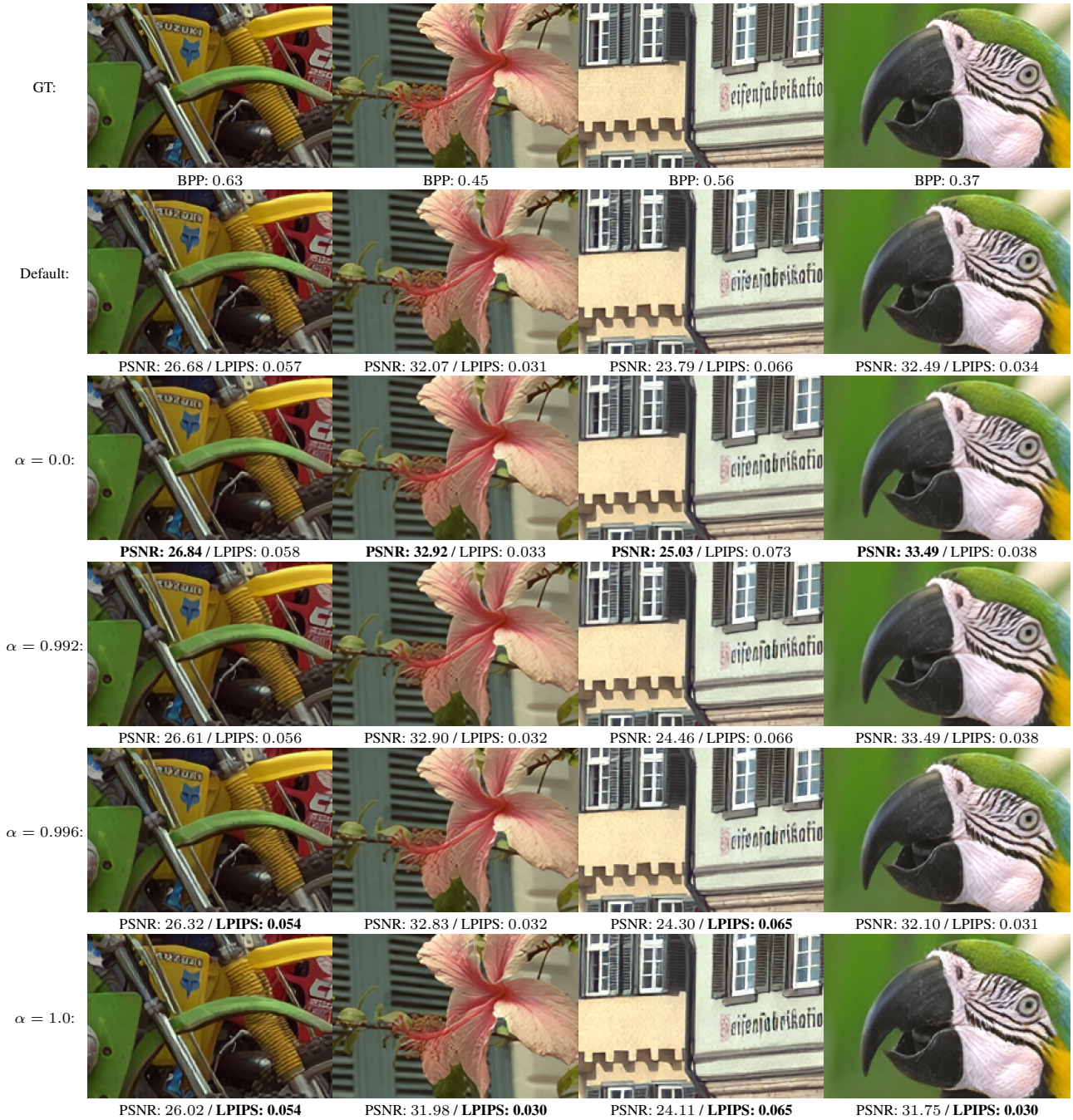


Figure 13. **Qualitative comparison of different tradeoffs between LPIPS and PSNR as generated by CDC x_0 [40]** for different crops of Kodak images. This is a more detailed version of Fig. 2. Images optimized for PSNR tend to be more blurry while optimizing for LPIPS can result in more noisy images.

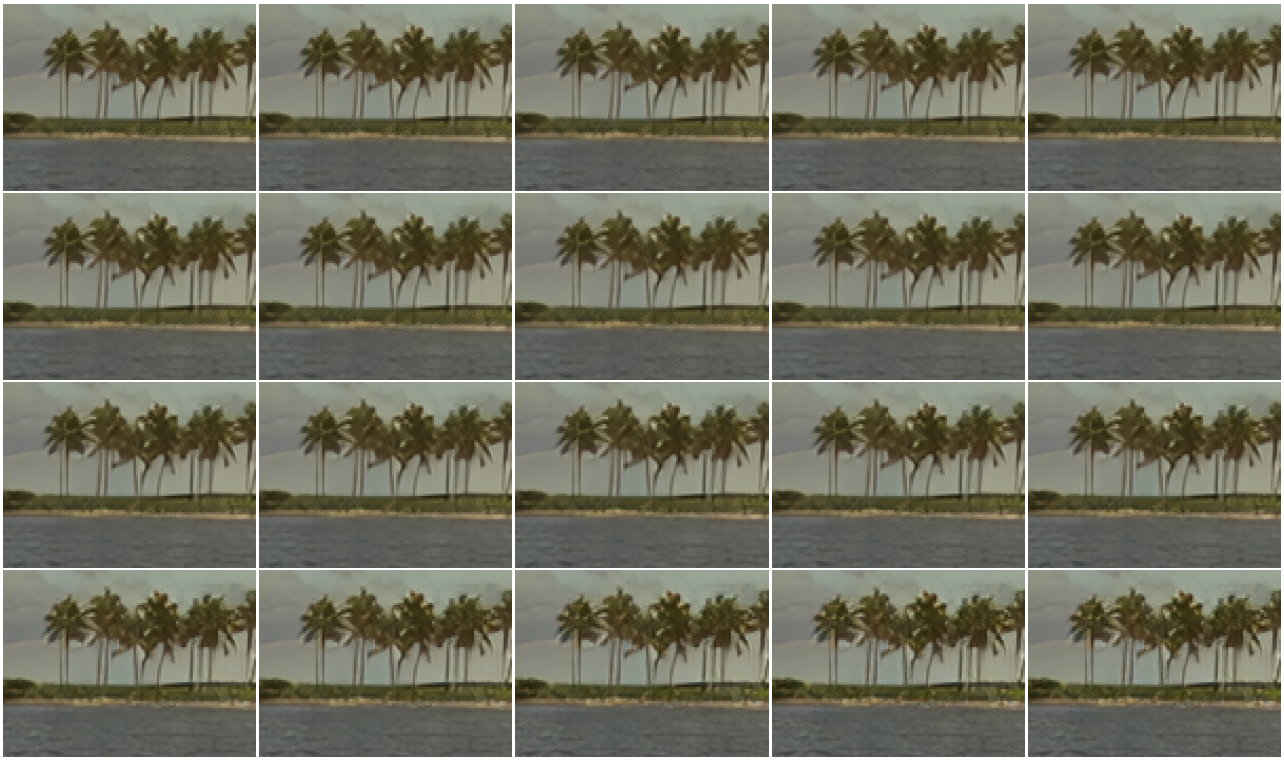


Figure 14. **Qualitative comparison of tradeoffs uniformly distributed from $\alpha = 0.9$ (top-left) to $\alpha = 1.0$ (bottom-right) as generated by CDC x_0 [40].** α increases from left to right and top to bottom. As is the case for most Kodak images for CDC x_0 [40], optimizing with $\alpha < 0.9$ results in the same configuration as when optimizing for $\alpha = 0.0$. The tradeoff parameter α is most sensitive near 1.0.

(a) Default configuration. PSNR: 19.18, LPIPS: 0.33, BPP: 0.032

(b) Tradeoff $\alpha = 0.0$. **PSNR: 19.73**, LPIPS: 0.31, BPP: 0.032

(c) Tradeoff $\alpha = 0.98$. PSNR: 19.46, **LPIPS: 0.30**, BPP: 0.032

(d) Tradeoff $\alpha = 1.0$. PSNR: 19.19, **LPIPS: 0.30**, BPP: 0.032

Figure 15. **Qualitative comparison of different tradeoffs for PerCo [12].** Optimizing this method for PSNR (b) leads to images with very low contrast. When optimizing for LPIPS (c, d) the generated images increase in sharpness and contrast.

(a) Default configuration. PSNR: 28.91, LPIPS: 0.094, BPP: 0.29

(b) Tradeoff $\alpha = 0.0$. **PSNR: 29.95**, LPIPS: 0.102, BPP: 0.29

(c) Tradeoff $\alpha = 0.985$. PSNR: 29.59, LPIPS: 0.094, BPP: 0.29

(d) Tradeoff $\alpha = 1$. PSNR: 28.91, **LPIPS: 0.089**, BPP: 0.29

Figure 16. **Qualitative comparison of different tradeoffs for CDC x_0 [40].** When optimizing this method for PSNR (b) the generated images tend to be more blurry. However, the generated image improves over the default configuration (a) by 1dB in PSNR. With increasing α (c, d), the sharpness of the image increases and texture details, such as the structure in the wood, become clearer.

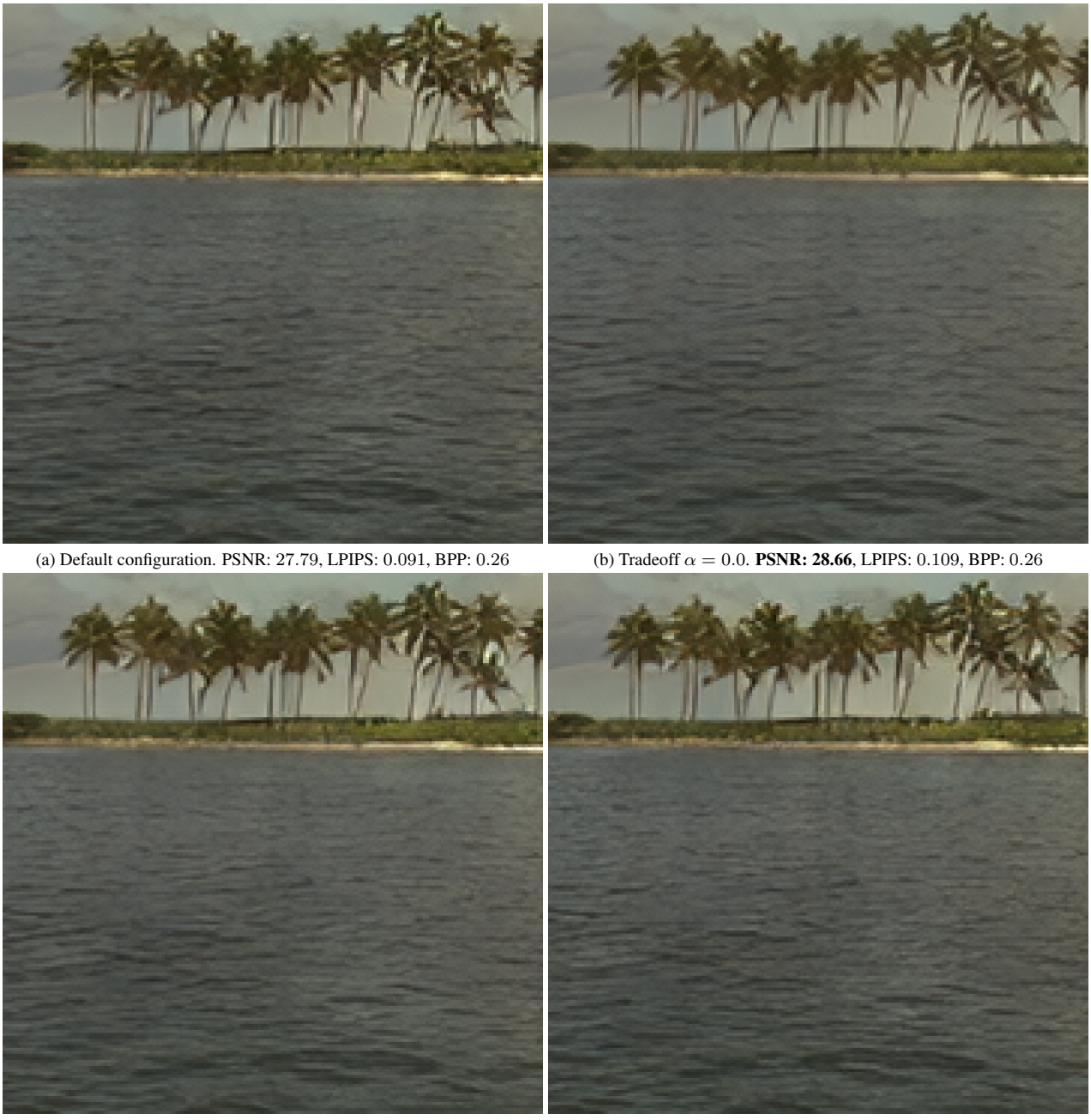
(a) Default configuration. PSNR: 28.47, LPIPS: 0.080, BPP: 0.22

(b) Tradeoff $\alpha = 0.0$. **PSNR: 29.22**, LPIPS: 0.090, BPP: 0.22

(c) Tradeoff $\alpha = 0.985$. PSNR: 28.14, LPIPS: 0.070, BPP: 0.22

(d) Tradeoff $\alpha = 1$. PSNR: 27.92, **LPIPS: 0.068**, BPP: 0.22

Figure 17. **Qualitative comparison of different tradeoffs for CDC x_0 [40].** When optimizing this method for PSNR (b) the generated image is lacking detail in areas such as the water. However, the reconstruction improves over the default configuration (a) by almost 0.8 dB in PSNR. With increasing α (c, d) sharpness increases and the water has a much more realistic look. However, especially (d) introduces some noise artifacts, for example in the faces.



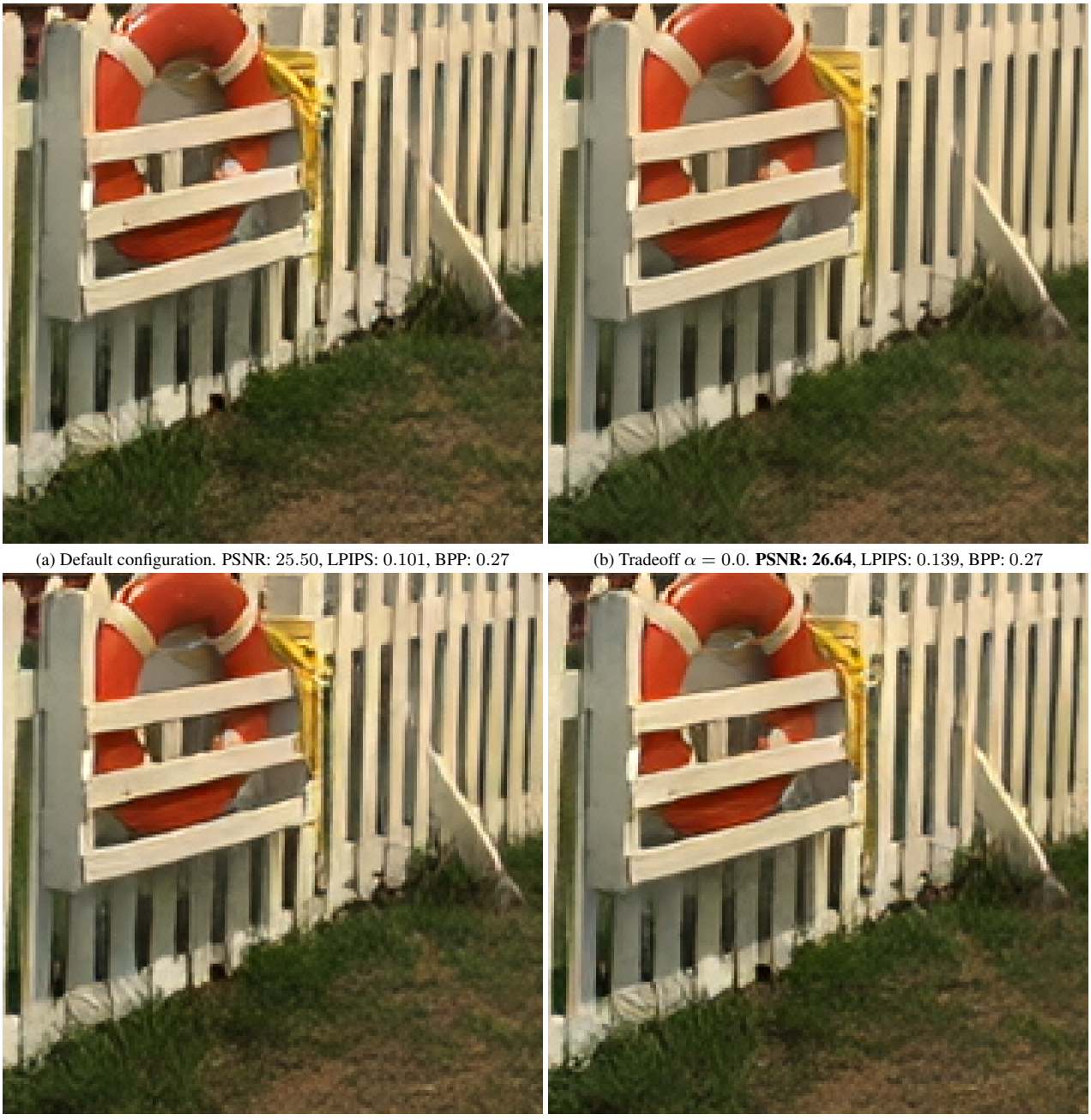
(a) Default configuration. PSNR: 27.79, LPIPS: 0.091, BPP: 0.26

(b) Tradeoff $\alpha = 0.0$. **PSNR: 28.66**, LPIPS: 0.109, BPP: 0.26

(c) Tradeoff $\alpha = 0.98$. PSNR: 28.08, LPIPS: 0.087, BPP: 0.26

(d) Tradeoff $\alpha = 1$. PSNR: 27.69, **LPIPS: 0.083**, BPP: 0.26

Figure 18. **Qualitative comparison of different tradeoffs for CDC x_0 [40].** When optimizing this method for PSNR (b) the generated image is lacking detail in areas such as the water and vegetation. However, the reconstruction improves over the default configuration (a) by about 0.8 dB in PSNR. With increasing α (c, d), the sharpness of the image increases and the water and vegetation have a much more realistic look.



(a) Default configuration. PSNR: 25.50, LPIPS: 0.101, BPP: 0.27

(b) Tradeoff $\alpha = 0.0$. **PSNR: 26.64**, LPIPS: 0.139, BPP: 0.27

(c) Tradeoff $\alpha = 0.98$. PSNR: 25.63, LPIPS: 0.095, BPP: 0.27

(d) Tradeoff $\alpha = 1$. PSNR: 25.43, **LPIPS: 0.093**, BPP: 0.27

Figure 19. **Qualitative comparison of different tradeoffs for CDC x_0 [40].** When optimizing this method for PSNR (b) the generated image is very smooth, and especially the vegetation is missing detail. However, the reconstruction improves over the default configuration (a) by about 1.1 dB in PSNR. With increasing α (c, d), the sharpness of the image increases and the vegetation has a much more realistic look. Additionally, some smaller details, such as the shadows on the ring, are more visible.

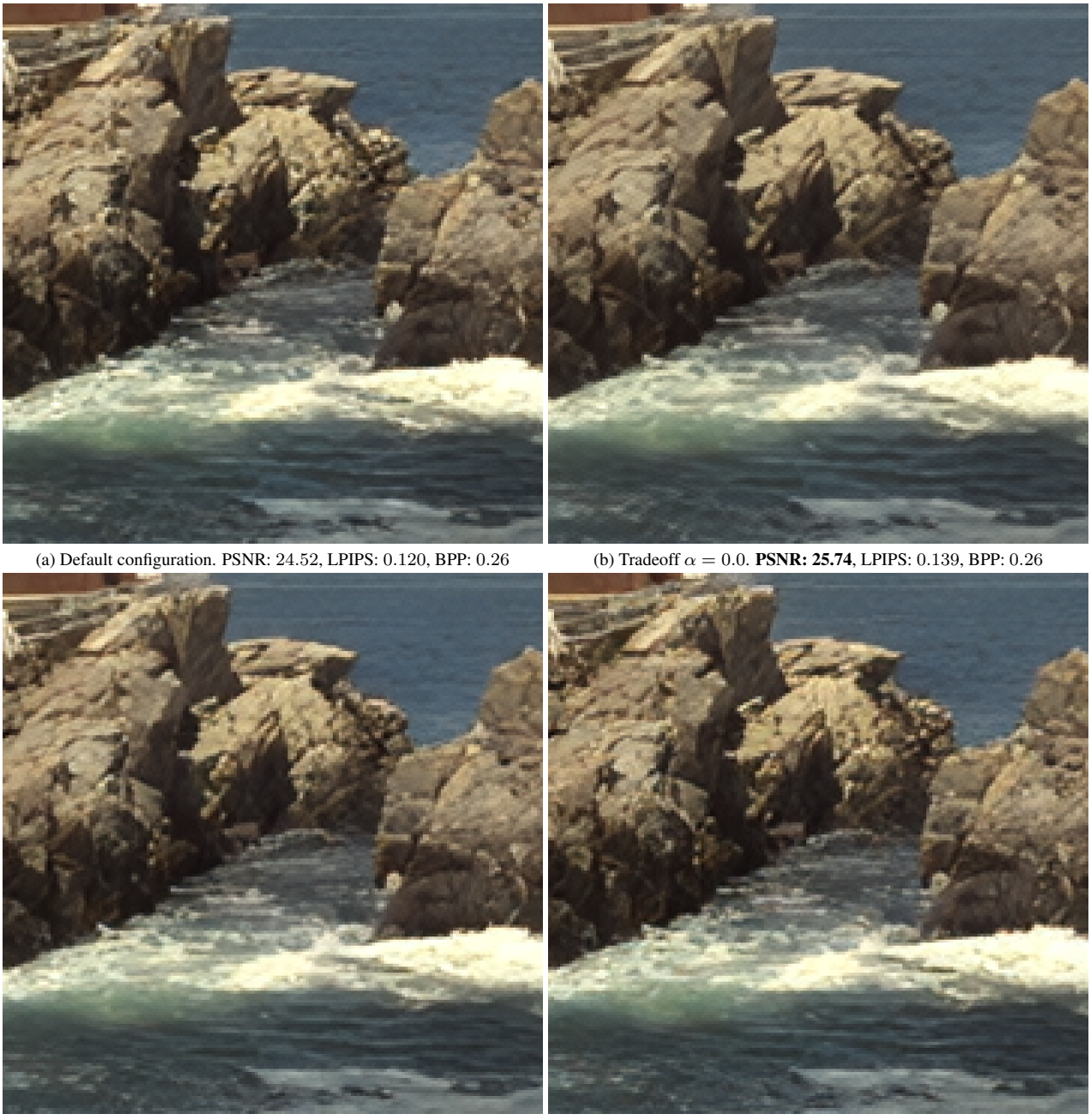
(a) Default configuration. PSNR: 28.10, LPIPS: 0.064, BPP: 0.18

(b) Tradeoff $\alpha = 0.0$. **PSNR: 29.30**, LPIPS: 0.085, BPP: 0.18

(c) Tradeoff $\alpha = 0.98$. PSNR: 28.67, LPIPS: 0.070, BPP: 0.18

(d) Tradeoff $\alpha = 1$. PSNR: 27.73, **LPIPS: 0.061**, BPP: 0.18

Figure 20. **Qualitative comparison of different tradeoffs for CDC x_0 [40].** When optimizing this method for PSNR (b) the generated image is very smooth, and missing some details in areas. However, the reconstruction improves over the default configuration (a) by about 1.2 dB in PSNR. With increasing α (c, d), the sharpness of the image increases. However, for high values of α (d), this comes at the cost of some noisy artifacts, especially noticeable in the lettering on the side of the plane.



(a) Default configuration. PSNR: 24.52, LPIPS: 0.120, BPP: 0.26

(b) Tradeoff $\alpha = 0.0$. **PSNR: 25.74**, LPIPS: 0.139, BPP: 0.26

(c) Tradeoff $\alpha = 0.98$. PSNR: 24.88, LPIPS: 0.110, BPP: 0.26

(d) Tradeoff $\alpha = 1$. PSNR: 24.48, **LPIPS: 0.106**, BPP: 0.26

Figure 21. **Qualitative comparison of different tradeoffs for CDC x_0 [40].** When optimizing this method for PSNR (b) the generated image is very smooth, and missing texture details for the rocks and water. However, the reconstruction improves over the default configuration (a) by about 1.2 dB in PSNR. With increasing α (c, d), the sharpness of the image increases, which is especially noticeable on the rock face.