Efficient Progressive Image Compression with Variance-aware Masking - Additional material

Alberto Presta[1] Enzo Tartaglione[2] Attilio Fiandrotti[1,2] Marco Grangetto[1] Pamela Cosman[3]
[1]University of Turin, Italy
[2] LTCI, Télécom Paris, Institut Polytechnique de Paris
[3] Dept. of Electrical and Computer Engineering, UC San Diego, CA, USA
alberto.presta@unito.it

A. Additional results

Fig. 12 compares our method with other progressive models regarding *Multiscale structural similarity index measure* (MS-SSIM) [1] on Kodak. While slightly outperformed by Jeon overall, we maintain complexity improvements noted in the main manuscript. Part of the future work will be to focus on perceptual metrics. Fig. 13 shows the distribution of the bpps against the quality of the reconstruction in terms of the mean squared error (MSE) when varying q. As is evident, at lower bitrates, the length of the bitstream increases exponentially with the chosen percentile values q used to generate progressive masks. This trend is logical because initially we include points with higher standard deviations, which contain more information but require more bits for storage. In contrast, as q increases and approaches the maximum value (100), only points with a lower standard deviation are omitted, which however have minimal impact on the final bitstream. Figures 14 (a,b) show GFLOPs and the time to decode using a CPU on Kodak.

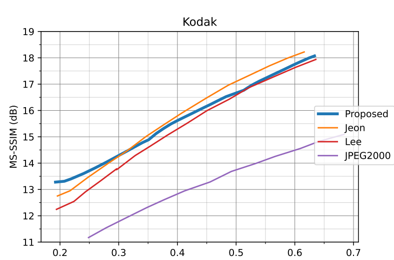


Figure 12. bbp vs. MS-SSIM curve against existing progressive sotas models on Kodak: Proposed, Jeon, Lee, and JPEG2000.

Figures 15,16 are image reconstruction from Kodak. Figures 17,18 are from CLIC validation dataset, with some latent representations across various qualities.

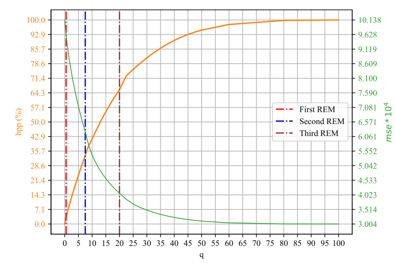
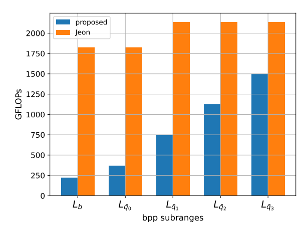
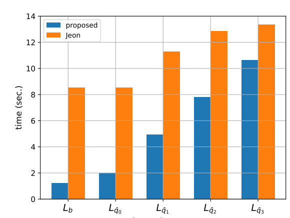


Figure 13. Distribution of the bpps (orange) and MSE (green) when varying q on Kodak.



(a)



(b)

Figure 14. GFLOPs and decoding time complexity on CPU (a,b) vs. Jeon *et al.*, considering different subranges on Kodak.

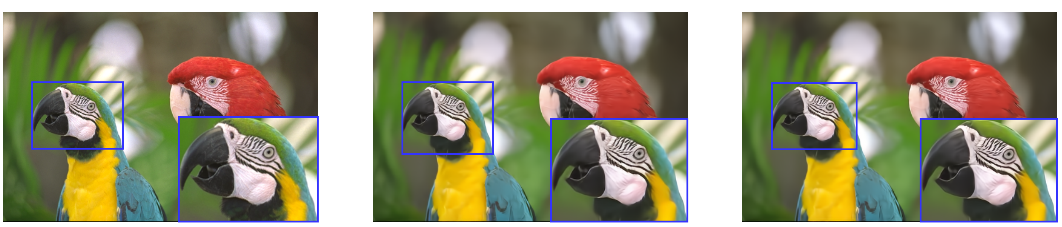
References

[1] Zhou Wang et al. Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers*. Ieee, 2003. 1

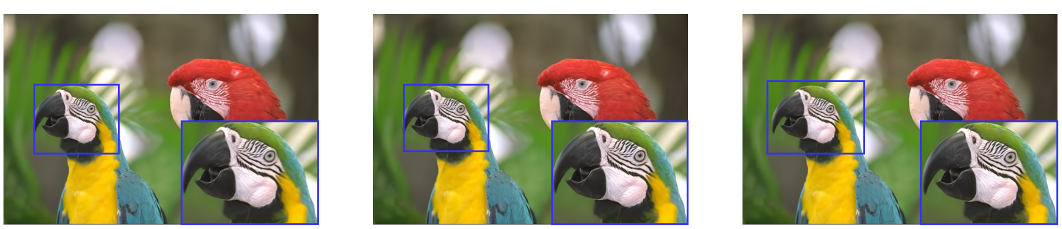
(a) Input: bpp/PSNR/ % of bpp (b) *base*: 0.16/32.93/37 (c) q=0.5:0.17/32.25/39.5

(d) q=10:0.32/35.62/74 (e) q=25: 0.38/36.84/88 (f) q=10: 0.43/37.48/100

Figure 15. Reconstruction of Kodim07 from kodak dataset using proposed method.

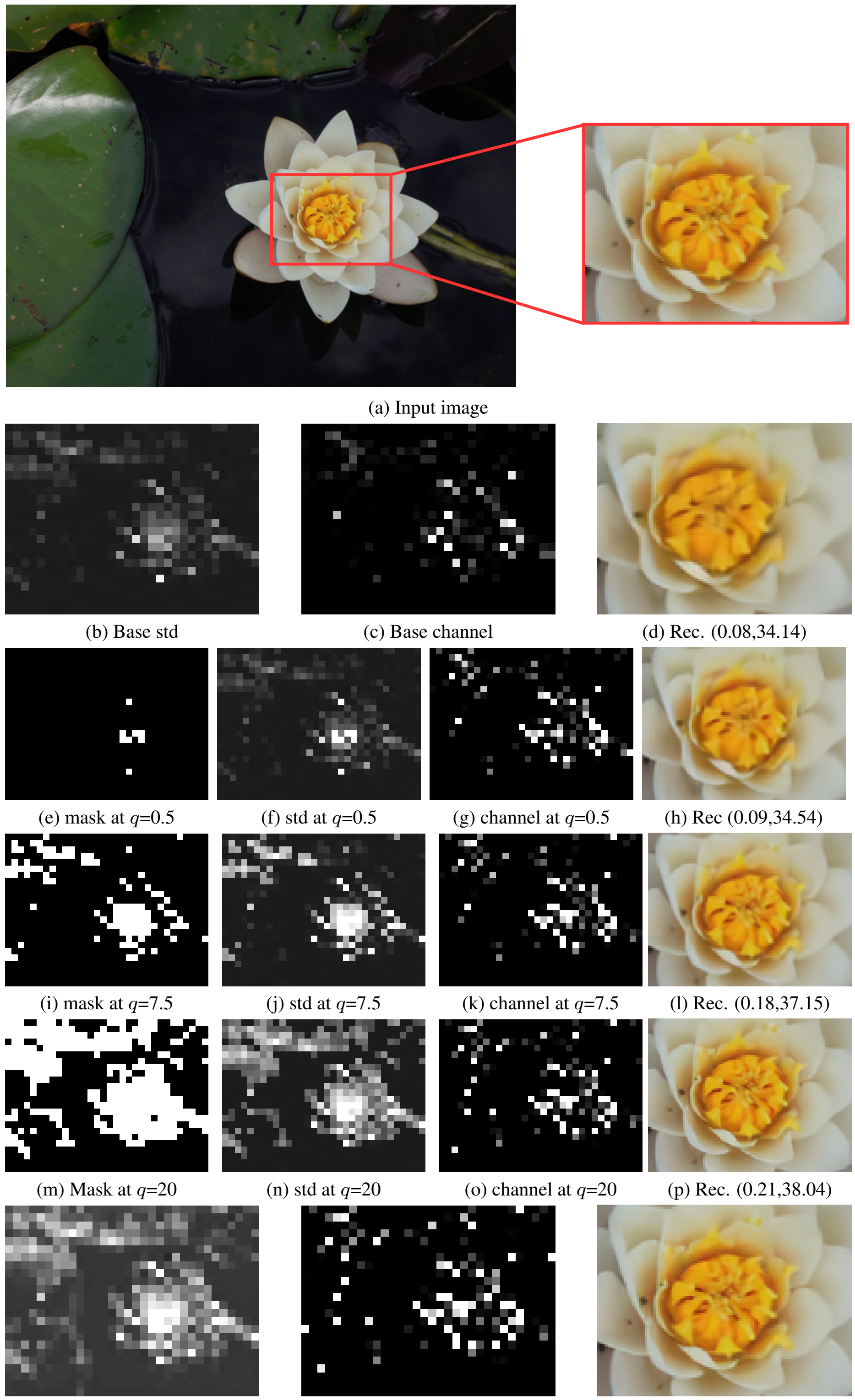


(a) Input: bpp/PSNR/ % of bpp (b) *base*: 0.096/33.33/34 (c) q=0.5:0.109/33.79/41.9



(d) q=10:0.22/37.40/84.6 (e) q=25: 0.24/37.89/92.3 (f) q=10: 0.26/37.96/100

Figure 16. Reconstruction of Kodim23 from kodak dataset using proposed method.



(a) Input image

(b) Base std (c) Base channel (d) Rec. (0.08,34.14)

(e) mask at q=0.5 (f) std at q=0.5 (g) channel at q=0.5 (h) Rec (0.09,34.54)

(i) mask at q=7.5 (j) std at q=7.5 (k) channel at q=7.5 (l) Rec. (0.18,37.15)

(m) Mask at q=20 (n) std at q=20 (o) channel at q=20 (p) Rec. (0.21,38.04)

(q) Std at q=100 (top). (r) channel at q=100 (top) (s) Rec. (0.23,38.3)

Figure 17. Final reconstruction with latent representations for different qualities, which varies through the raws.

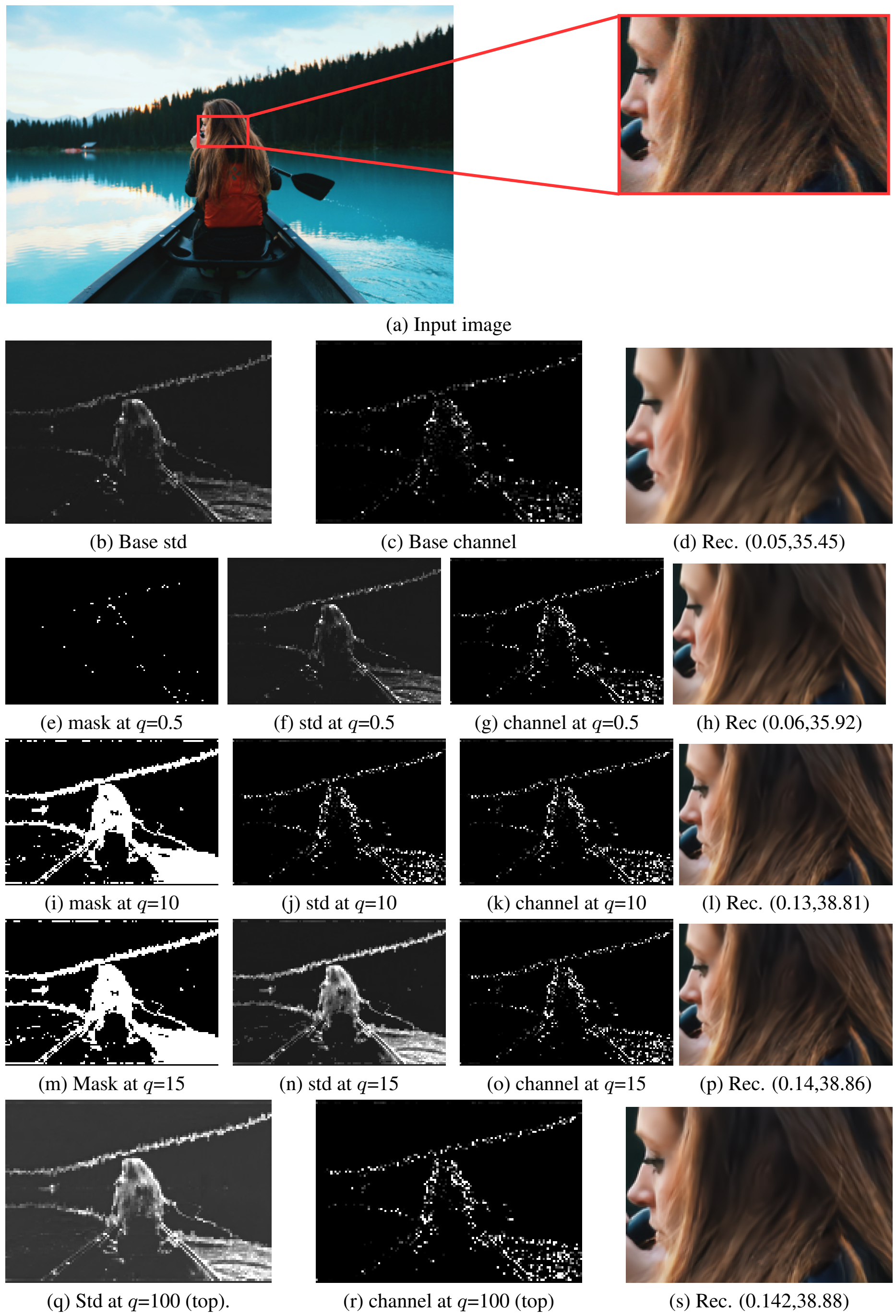


Figure 18. Final reconstruction with latent representations for different qualities, which varies through the raws.