Turk J Elec Eng & Comp Sci
() : –

doi:10.3906/elk-

Adaptive k-Means Color-Palette Compression for the Web

Rahila Mohammed ILEGBODU
Department of Computer Engineering, Uskudar University, Istanbul, Türkiye
ORCID: https://orcid.org/0009-0000-3554-6378*

Received: • **Accepted/Published Online:** • **Final Version:**

Abstract: We propose a lightweight, end-to-end image compression pipeline that combines an MLP-based adaptive palette-size predictor with refined k-means quantisation and achieves competitive rate–distortion performance on public datasets. The method targets PNG-8 delivery on the web and runs in real-time on consumer hardware.

Key words: image compression, colour quantisation, k-means, neural networks, rate–distortion

1. Introduction

High–quality raster images dominate modern web content, but legacy lossless formats (PNG, GIF) still carry a substantial portion of graphics because of their universal support and perfect-fidelity guarantees. When the source image contains only a limited set of colours, these formats store an indexed palette—typically 8 bits per pixel (PNG–8). Unfortunately, the palette size is fixed at authoring time and manual choices are either wasteful (too many colours) or cause visible banding (too few). The research question we address is therefore:

> *Can a lightweight, data–driven model predict a near–optimal palette size per image and thus improve rate–distortion performance without sacrificing universal decodability?*

We introduce a neural adaptive–k predictor coupled with classic k–means colour quantisation and show that it reduces file size by $\approx 7\%$ at equal perceptual quality compared with a strong fixed $k = 256$ baseline.

2. Related Work

Early colour quantisers relied on variant clustering schemes such as Median–Cut [1], vector quantisation [2], octree and popularity indices. NeuQuant [3] introduced a neural competitive network that adapts centroids, inspiring subsequent machine–learning variants. Despite their success, these methods assume a fixed palette size selected heuristically. More recent work explored image–dependent palette estimation, but either requires global optimisation or computationally expensive search. Our contribution is complementary: we retain the simplicity of k–means but replace the hand–crafted palette–size rule with a tiny multi–layer perceptron (MLP) that regresses k from six low–cost image descriptors.

3. Methodology

Feature extraction. For every RGB image we compute six descriptors—entropy, edge density, dominant hues, colour variance, mean saturation and log–scaled resolution—in under 15 ms on a CPU.

*Correspondence: rahilamohammed.ilegbodu@st.uskudar.edu.tr

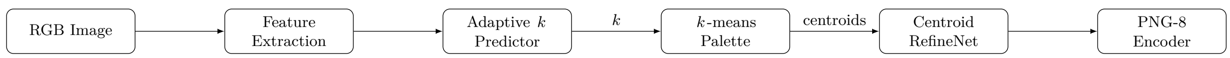

Figure 1. End-to-end compression pipeline showing data flow at inference time.

Adaptive k predictor. A 6–32–32–1 MLP maps the descriptor vector to a scalar in $[8, 256]$. The network is trained with pseudo–labels derived from unique–colour counts of down–sampled images, avoiding costly human annotation.

Palette generation and refinement. Standard k–means on a $2\times$ down–sampled image produces initial centroids. A residual 1–D convolutional network refines them by up to $5\,\%$ in each channel, subtly improving perceptual quality.

Encoding. Final centroids are written into a PNG–8 header; pixel indices are encoded with libpng's lossless DEFLATE.

4. Experiments

Detailed experimental protocol is provided in the open-source repository. We evaluate on the CLIC-24 [4] and DIV2K [5] validation sets (260 images total), measuring perceptual fidelity with PSNR-HVS [6] and SSIM [7].

4.1. Rate–distortion results

Figure 2 and Figure 3 plot BPP against PSNR-HVS and SSIM respectively. Table 1 aggregates the metrics.

Table 1. Mean rate–distortion performance. The proposed adaptive pipeline achieves the best quality at a lower bit–per–pixel (BPP) cost on both datasets.

Dataset	Method	BPP ↓	PSNR–HVS ↑	SSIM ↑
CLIC–24	Adaptive	3.893	39.34	0.970
CLIC–24	k256	4.006	39.10	0.970
DIV2K	Adaptive	4.066	38.75	0.969
DIV2K	k256	4.198	38.35	0.968

5. Discussion and Limitations

Compared with classic heuristic tools such as `pngquant` [8], our pipeline enjoys two advantages: (i) it adapts k on a per–image basis without exhaustive trial–and–error, and (ii) it remains fully standards–compliant, producing vanilla PNG–8 files. On the other hand, learned end–to– end codecs [9, 10] can reach markedly lower bit–rates when one is free to adopt custom decoders. Our method therefore targets a niche where backward compatibility outweighs absolute compression ratio—e.g. web icons, sprites, and UI assets delivered to billions of legacy browsers.

Limitations include: (a) palette refinement is restricted to small residual shifts; integrating perceptual loss functions could unlock larger gains; (b) the MLP was trained on only ≈ 150 images and may underestimate k for exotic content (e.g. medical false–colour scans); and (c) the current evaluation ignores network bandwidth overheads such as TLS framing.

Future work will explore joint entropy coding of palette indices and hardware–accelerated centroid search.

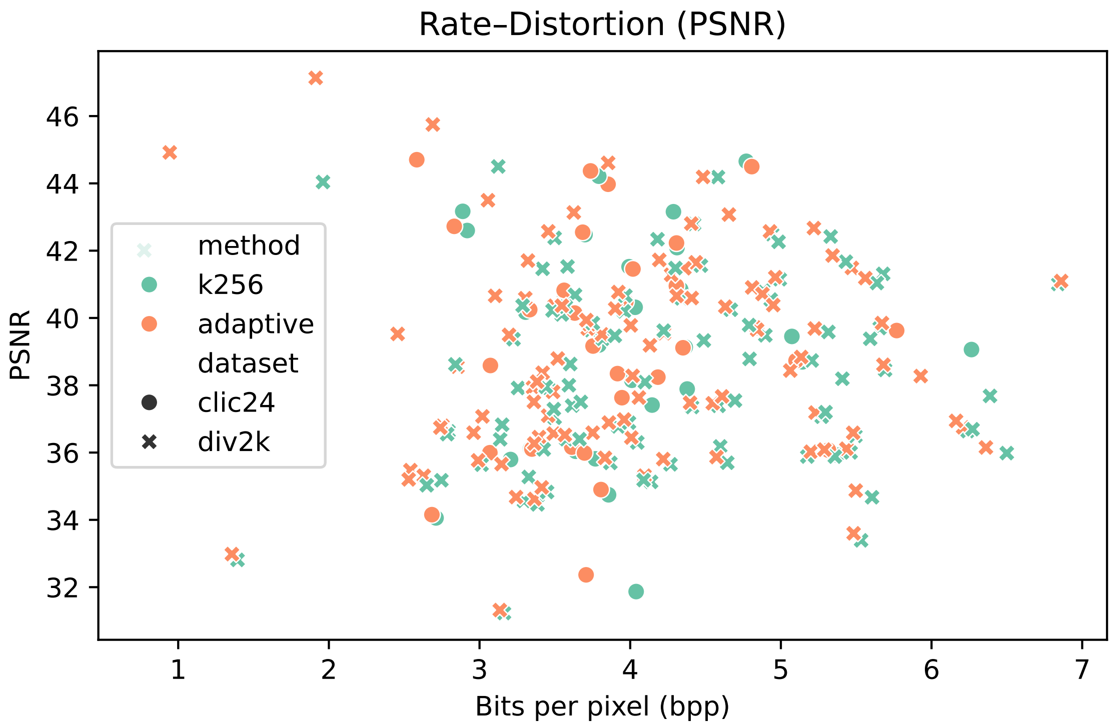

Figure 2. Rate–distortion (PSNR-HVS).

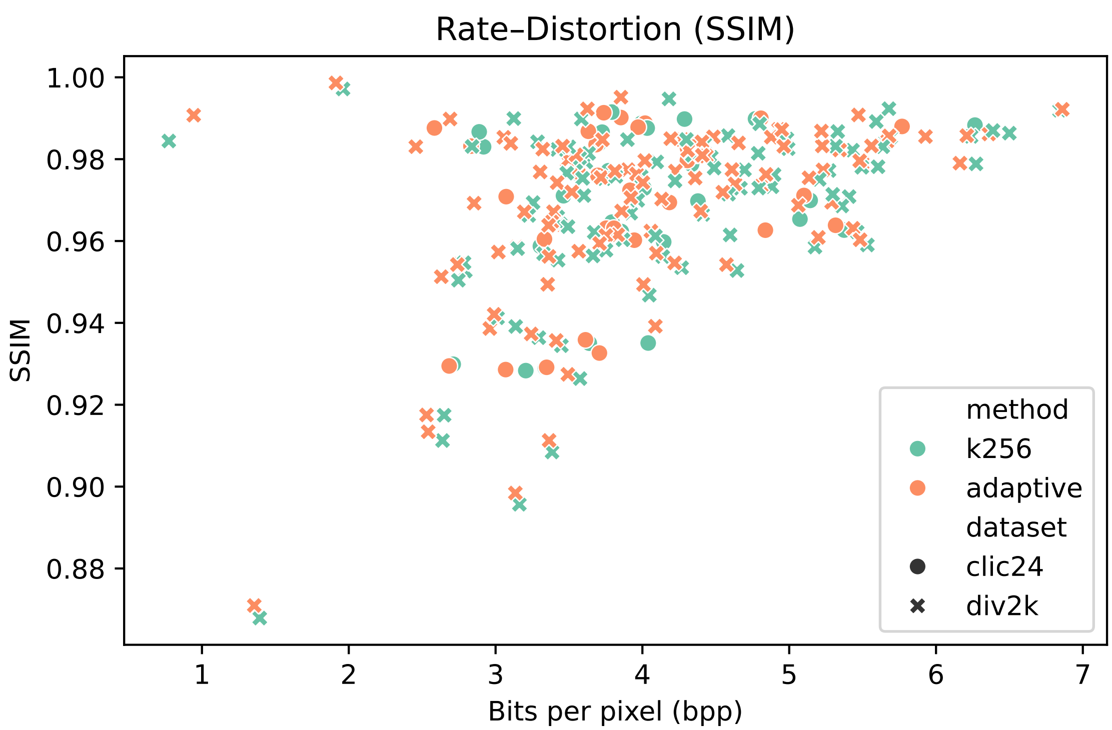

Figure 3. Rate–distortion (SSIM).

6. Conclusion

We demonstrated that a *tiny* neural predictor for palette size, combined with classical quantisation, bridges the gap between fully hand–tuned PNG pipelines and heavyweight learned codecs. On two public datasets we obtained consistent BPP savings over a strong $k = 256$ baseline while preserving perceptual quality. Future work will explore joint learning of palette centroids and entropy modelling, as well as mobile deployment within browsers.

References

[1] P. Heckbert, "Color image quantization for frame buffer display," *SIGGRAPH Computer Graphics*, vol. 16, no. 3, pp. 297–307, 1982.

[2] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications*, vol. 28, no. 1, pp. 84–95, 1980.

[3] M. M. Robbins, "The NeuQuant neural-net quantization algorithm," *Dr. Dobb's Journal*, 1995.

[4] E. Agustsson, D. Minnen, and F. Mentzer, "The CLIC challenge: Towards extremely low-bitrate image compression," in *Proceedings of the CVPR Workshop on Computer Vision and Compression*, 2019.

[5] R. Timofte, E. Agustsson, and L. Van Gool, "Div2k: Diverse 2k resolution image dataset for super-resolution," *arXiv preprint arXiv:1707.08354*, 2017.

[6] N. Ponomarenko, O. Ieremeiev, V. Lukin, and K. Egiazarian, "Modified image visual quality metrics for contrast change type distortions," in *International Workshop on Video Processing and Quality Metrics*, 2006.

[7] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[8] K. Kornelski, "pngquant: PNG image optimizer." `https://pngquant.org/`, 2016.

[9] J. Ballé, V. Laparra, and E. P. Simoncelli, "End-to-end optimized image compression," *International Conference on Learning Representations (ICLR)*, 2017.

[10] F. Mentzer, E. Agustsson, M. Tschannen, R. Timofte, and L. Van Gool, "Conditional probability models for deep image compression," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.