Evaluating Preference for Images Enhanced with Simple Tone Curves

James Bennett[a*], Graham Finlayson[a]

[a]*University of East Anglia, Norwich, United Kingdom*
[*]*Corresponding author: James Bennett, james.r.bennett@uea.ac.uk*

ABSTRACT

Tone curves are central to camera pipelines and image editing workflows, describing how tonal values are mapped to produce visually appealing images. While tone curves can be arbitrarily complex, recent work has suggested that simple tone curves – defined as monotonic functions with at most one inflexion point – are sufficient to approximate expert adjustments. This raises an important perceptual question of whether human observers prefer images enhanced with simple tone curves. We address this by conducting a psychophysical study on the MIT-Adobe FiveK dataset. Expert tone curves were approximated using a constrained optimisation method from the prior art to derive the best simple curve in the primal domain, and additionally in a logarithmic domain. For each of 24 selected images (drawn from well-, less-well- and poorest-fitted simple curves), we compared four renditions: the unenhanced input, the expert output, the primal-domain simple rendition, and the log-domain simple rendition. Eight naïve observers completed pairwise preference tests under controlled viewing conditions. Our results show that both primal and log-domain simple tone curves produce images with perceptual quality comparable to expert renditions. Observers did not show a significant preference between expert and simple renditions, indicating that simplified curves provide a suitable approximation, with the log-domain simple renditions being marginally favoured. These findings demonstrate that simple tone curves are not only objectively similar but also perceptually competitive to expert renderings, offering a practical alternative to complex expert adjustments in image enhancement.

Keywords: Tone mapping, Tone curves, Image enhancement, Psychophysical experiment, Pairwise preference

INTRODUCTION

Tone mapping is widely used in camera pipelines and image editing workflows to manipulate brightnesses and enhance the perception of images. When applied to real world luminance, tone maps describe camera response functions and the conversion from high dynamic range [1,2]. Tone mapping can also be applied within post processing [3,4] to enhance the perception of an image according to one's preference [5]. A tone curve describes the function that maps input tonal values to output tonal values. When the same function is applied to every pixel in the image, the tone mapping is said to be global. Tone curves may have a fixed shape, be custom-defined per image, or be computed algorithmically. They can take specific functional forms, such as power laws or sigmoids, or be free-form curves created by interpolating control points. While tone curves are continuous and usually increasing, their shape is typically unconstrained.

Previous work [6] has argued that tone curves should be simple – that is, they have at most one inflexion point, in contrast to complex, "wiggly" curves. This gives rise to four possible cases, depicted in Figure 1. Case 1 and Case 2 have zero inflexion points: in Case 1 the gradient increases monotonically, while in Case 2 it decreases monotonically. Case 2 corresponds to common tone curve shapes, such as gamma encoding. Case 3 and Case 4 each have one inflexion point, the position of which can vary across the domain. Case 3 is the commonly seen "Sshaped" curve, where the gradient increases to the inflexion point and then decreases, boosting midtone contrast. Case 4 is the opposite, with the gradient decreasing to the inflexion point and then increasing.

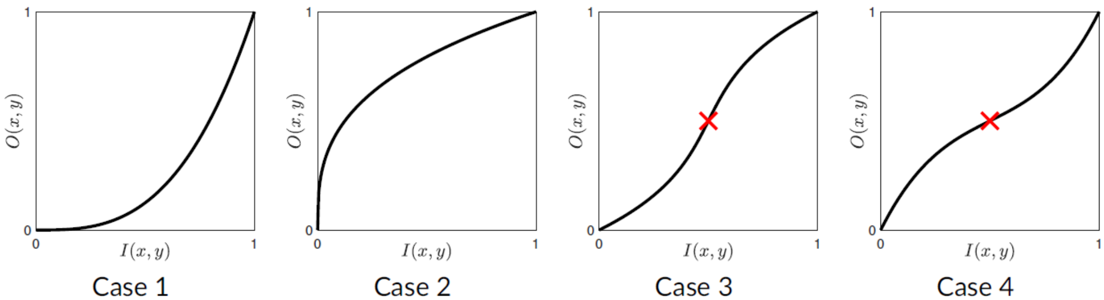


Figure 1: Depiction of four simple cases

In [6], a mathematical optimisation method was proposed to derive a simple curve that approximates a more complex tone curve. The objective function seeks a simple tone curve that closely matches a target curve, subject to constraints enforcing monotonicity and the allowed inflexion case. The problem is formulated as a quadratic program with linear constraints and can be solved efficiently and a global minimum is found [7]. For the MIT-Adobe FiveK dataset, expert tone curves were found to be either inherently simple or well approximated with a simple curve.

In this work, we investigate whether human observers prefer images enhanced with simple versus complex tone curves. We conduct a psychophysical experiment to evaluate preference between the ground-truth expert images and their simplified counterparts. The remainder of this paper details the creation of images for the subjective experiment using the optimisation method, describes the experimental setup, and presents the results. Our findings indicate that the simplification method produces images that are as preferred as expert renditions, with a marginal perceptual advantage when the simple curves are solved in log space.

METHODOLOGY

Generating the images

This study uses the MIT-Adobe FiveK dataset [8], provided as an Adobe Lightroom catalogue from which pairs of input–output images can be extracted. Following the procedure of [6], adopted from [9], we export input images from the collection Input minus 1.5 in ProPhoto 16-bit TIFF format, with the long edge scaled to 640 px. Because the Lightroom adjustments are numerous and undisclosed, the images are not strictly related by a single global tone curve. The expert images are generated according to [6] ensuring a global mapping. Informally, the most global fit was found by mapping the L* channel of CIELAB. The input, expert and simple images, per (x, y) pixel, are respectively $\mathcal{I}(x, y)$ and $\mathcal{P}(x, y)$:

$$\mathcal{I}(x, y) = [L_I^* \; a_I^* \; b_I^*]^\top$$

$$\mathcal{P}(x, y) = [L_P^* \; a_P^* \; b_P^*]^\top = [T(L_I^*) \; a_P^* \; b_P^*]^\top$$

$$\hat{\mathcal{P}}(x, y) = \left[\hat{T}(L_I^*) \; a_P^* \; b_P^*\right]^\top$$

Function T is the expert's global tone curve mapping L_I^* to L_P^*. For each input-expert pair, we use the prior optimisation to solve for a simple tone curve $\hat{T}$ (constrained to be increasing with at most one inflexion point) that best approximates the expert tone curve, generating $\hat{\mathcal{P}}(x, y)$. Simple tone-mapped images are generated by applying $\hat{T}$ to the input L^*. We also solve for the simple tone curve in logarithmic space. Applying a logarithmic transformation to the L^* channel emphasises darker values in the optimisation. The simple tone curve is then solved in this transformed space, and subsequently mapped back to the original domain. For each image, then, we have: (1) the input version, (2) the expert rendition (where the expert tone curve may be complex), (3) an image tone mapped with a simple approximation of the expert curve, (4) a simple tone mapped rendition where the simple curve is solved for in log-space. These are exported from MATLAB in the sRGB colour space for display.

Conducting the preference experiments

We conducted a pairwise psychophysical experiment using a subset of 24 images from the MIT-Adobe FiveK dataset. The selection comprised the 8 images with the highest ΔE colour difference between simple and expert renditions, 8 chosen randomly from above the 99th percentile, and 8 chosen randomly from below it.

The experiment compared four image renditions: input, expert, simple, and log-space simple. A pair of renditions were shown side by side at a time. The observer was asked to choose which image they preferred by clicking a button. For four renditions there are six unique pairings. Each pair was shown twice with reversed display order, giving 12 comparisons per image and 288 comparisons per observer.

Eight naïve observers with normal colour vision (verified using Ishihara plates) participated in the study. Experiments were conducted in a grey room with controlled lighting and a calibrated monitor, following ISO 3664:2009. Images were displayed in random order (with the restriction that the same image was not shown simultaneously). Observers viewed the images and recorded preferences through a MATLAB graphical user interface, with colour profiles matched to the monitor.

RESULTS AND DISCUSSION

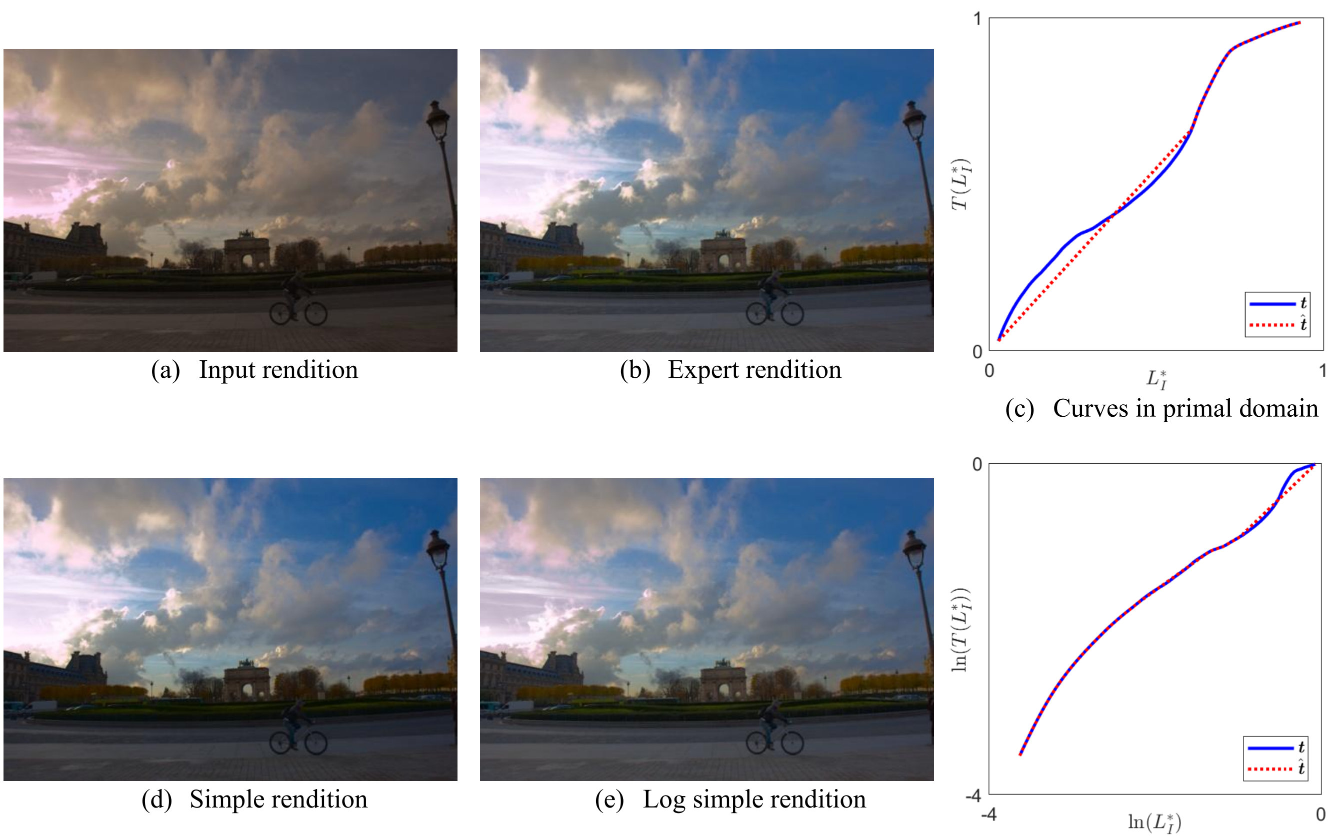
Table 1 lists the images used in the psychophysical experiment by their expert and input number. For each image we report the objective ΔE colour difference between the expert rendition and the simple renditions (solved in both primal and logarithmic domains), along with the ΔE rank out of 25,000 images. The table also shows the number of times each rendition was voted as preferred. Each row (of preference votes) sums to 96, corresponding to 12 comparisons per image for each of the eight observers.

Table 1: Psychophysical experiment statistics and preference votes.

	Primal solved		**Log solved**		**Preference votes**			
Image	ΔE	**Rank**	ΔE	**Rank**	**Input**	**Expert**	**Simple**	**Log simple**
C2704	< 0.01	29	0.0658	18479	0	39	30	27
B3531	< 0.01	7982	< 0.01	9079	16	27	20	33
A286	< 0.01	10503	0.0176	13091	1	31	29	35
A2485	< 0.01	11408	< 0.01	2738	0	31	29	36
E3626	< 0.01	11446	< 0.01	7141	0	30	27	39
A2213	0.0503	19062	< 0.01	3335	12	26	23	35
B2473	0.0551	19423	0.274	22666	0	32	32	32
B1373	0.151	22671	0.180	21634	3	28	28	37
A707	0.827	24780	2.64	24982	0	28	31	37
C1620	0.848	24800	1.54	24896	0	37	34	25
B2079	0.890	24815	1.00	24709	12	23	27	34
C823	1.00	24858	1.33	24863	0	32	30	34
A1543	1.06	24869	1.11	24772	1	28	36	31
E604	1.12	24887	0.610	24176	0	31	32	33
C4215	1.62	24957	2.01	24949	6	29	29	32
C374	1.68	24962	1.63	24905	0	30	33	33

A21	2.68	24993	0.630	24228	3	32	27	34
A596	2.69	24994	0.357	23268	31	29	16	20
A1692	2.72	24995	2.75	24984	0	25	37	34
A2628	2.86	24996	0.840	24576	0	33	28	35
B1999	2.88	24997	2.60	24978	0	30	33	33
A2049	3.12	24998	2.14	24957	10	28	34	24
A2823	3.13	24999	1.92	24942	0	37	27	32
A2673	3.22	25000	1.52	24893	1	38	29	28

From Table 1 we see that the selected images span a wide range of objective ΔE values, in line with the selection methodology. To illustrate the renditions and their tone curves, we display image A2823 in Figure 2, one of the most challenging cases. The top images compares the input (left) with the expert rendition (right). The expert is an enhanced version of the input produced through a complex tone curve. The bottom row shows two simplified alternatives: the left image is generated by applying the best primal-domain simple curve, while the right image applies the best simple curve solved in log-space. Visually, the log-domain solution more closely resembles the expert rendition due to closer similarity of the dark tones.



(a) Input rendition (b) Expert rendition (c) Curves in primal domain

(d) Simple rendition (e) Log simple rendition (f) Curves in log domain

Figure 2: Results for image 2823 showing image renditions and expert and solved for simple curves.

To better understand this behaviour, we plot the corresponding tone curves for $L^*$and for $log(L^*)$. In both cases, the non-simple expert curve is shown in solid blue and the best simple approximation in dotted red. Consistent with the visual results, the simple curve solved in the log domain more faithfully follows the expert curve, especially in the dark tones, suggesting that the logarithmic transformation improves the approximation for this image.

For most images, Table 1 shows that the expert, primal, and log-simple renditions were preferred more often than the unenhanced input. Nevertheless, in some cases the input rendition still attracted votes, indicating that observers do not universally reject unenhanced images. Preferences among the expert, primal, and log-simple renditions varied by image, making it difficult to draw strong conclusions from raw vote counts alone.

To aggregate and compare preferences across all images, we applied Thurstone's Case V model [10] using a pairwise comparison toolbox [11]. This method psychometrically scales the pairwise votes into *Just Objectionable Difference* (JOD) units, where higher scores indicate stronger preference and 1 JOD unit corresponds to a 75% discrimination threshold. Although the units are arbitrary and not comparable across experiments, they provide a consistent scale for analysing relative differences. Figure 3 shows the scaled results, with 95% confidence intervals computed via bootstrapping (1,000 samples).

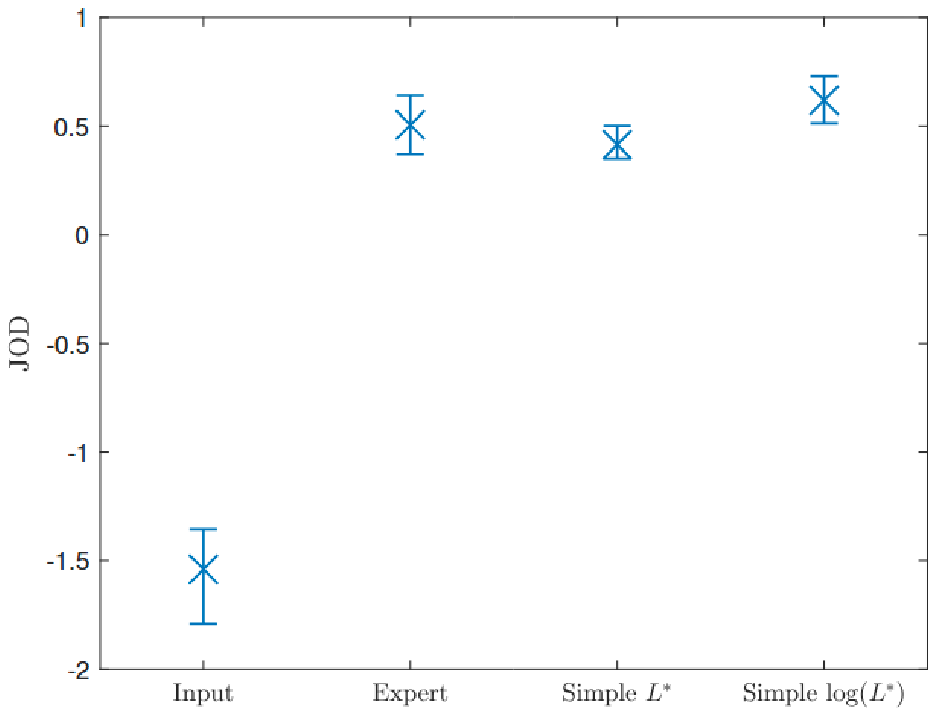


Figure 3: Scaled preference scores, input (-1.54), expert (0.51), simple (0.42), and log-space simple (0.62).

As expected, observers preferred the expert and simple renditions over the input, confirming that the enhancements provide a perceptual benefit. The differences among expert, primal, and log-simple renditions were within the margin of uncertainty, with overlapping confidence intervals. This suggests that observers do not exhibit a strong preference between these three methods, and that the simplified tone curves approximate the expert adjustments effectively.

A point of interest is the slightly higher preference for log-space simple curves. While we must be cautious about over-interpreting results with overlapping confidence intervals, the data suggest that observers may favour log-space solutions. Whilst the primal-domain simple renditions scored just below the expert renditions, the log-domain renditions scored slightly above. This indicates that observers may even prefer a simple tone enhancement to the expert-designed one.

CONCLUSION

Simple tone curves – defined as monotonic increasing functions with at most one inflexion point – can effectively enhance images while preserving visual quality comparable to more complex expert-designed curves. Our study shows that human observers do not exhibit a significant preference for expert tone curves over simple ones. Moreover, solving for the simple tone curve in log space appears to provide a slight perceptual advantage. These results support the use of simple tone curves as a practical tool for image enhancement. By reducing complexity without sacrificing perceptual quality, simple curves offer a compelling alternative to expert adjustments and may, in some cases, even be preferred.

ACKNOWLEDGEMENTS

This work was supported by the Engineering and Physical Sciences Research Council, grants EP/S023917/1 and EP/S028730/1.

REFERENCES

[1] Grossberg, M., & Nayar, S. K. (2003). Determining the camera response from images: What is knowable? IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(11), 1455–1467. https://doi.org/10.1109/tpami.2003.1240119

[2] Salih, Y., bt. Md-Esa, W., Malik, A. S., & Saad, N. (2012). Tone mapping of HDR images: A review. 2012 4th International Conference on Intelligent and Advanced Systems (ICIAS2012), 1, 368–373. https://doi.org/10.1109/ICIAS.2012.6306220

[3] Qi, Y., Yang, Z., Sun, W., Lou, M., Lian, J., Zhao, W., Deng, X., & Ma, Y. (2021). A Comprehensive Overview of Image Enhancement Techniques. Archives of Computational Methods in Engineering, 29(1), 583–607. https://doi.org/10.1007/s11831-021-09587-6

[4] Wang, D. C. C., Vagnucci, A. H., & Li, C. C. (1983). Digital image enhancement: A survey. Computer Vision, Graphics, and Image Processing, 24(3), 363–381. https://doi.org/10.1016/0734-189x(83)90061-0

[5] Sing Bing Kang, Kapoor, A., & Lischinski, D. (2010, June 1). Personalization of image enhancement. IEEE Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/cvpr.2010.5539850

[6] Bennett, J., & Finlayson, G. (2023). Simplifying Tone Curves for Image Enhancement. Color and Imaging Conference, 31(1), 108–114. https://doi.org/10.2352/cic.2023.31.1.21

[7] Lawson, C. L., & Hanson, R. J. (1974). Solving Least Squares Problems. Prentice-Hall.

[8] Vladimir Bychkovsky, Paris, S., Chun, E., & Durand, F. (2011, June 1). Learning photographic global tonal adjustment with a database of input/output image pairs. CVPR 2011. IEEE Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/cvpr.2011.5995413

[9] Hu, Y., He, H., Xu, C., Wang, B., & Lin, S. (2018). Exposure: A White-Box Photo Post-Processing Framework. ACM Transactions on Graphics, 37(2), 1–17. https://doi.org/10.1145/3181974

[10] Thurstone, L. L. (1927). A law of comparative judgment. Psychological Review, 34(4), 273–286. https://doi.org/10.1037/h0070288

[11] Perez-Ortiz, M., & Mantiuk, Rafal K. (2017). A practical guide and software for analysing pairwise comparison experiments. ArXiv.org. https://arxiv.org/abs/1712.03686