Very Simple Tone Curves

James Bennett and Graham Finlayson, University of East Anglia, Norwich, UK.

Abstract

Tone mapping algorithms are used to compress dynamic range, make image details more conspicuous and generally enhance the image for preference. Global tone mapping manipulates the brightnesses of pixels by applying a single function - or tone curve - to every pixel in the image. Tone curve generation algorithms often constrain the shape of their tone curves and it has been argued that tone curves should be simple, meaning they have one or zero inflexion points. In this work, we investigate whether tone curves should be simplified even further. We present our method which finds the zero inflexion tone curve - which we call a Very Simple (VS) curve - that best approximates a potentially complex tone curve. For the MIT-Adobe FiveK dataset, comprising 25,000 expert tone adjustments, we calculate the best VS approximations and find these curves produce visually similar images compared with more complex counterparts.

Introduction

Arguably, lighting is a critical aspect of producing good images where details are conspicuous and represent the observer's perception of the scene. Accordingly, professional photographers attempt to control the light when taking photos in the studio. Yet cameras capture high dynamic range scenes where not all the details are illuminated well. A single image often contains deep shadows and strong highlights - i.e. a very high dynamic range of brightnesses - and it is impossible to physically reproduce these on a display. Relighting the image so that all the detail is *pleasingly* illuminated is a large area of research and there are algorithms ranging from modern image-to-image trained networks [1, 2], to theories based on modelling human vision [3] to tone-mapping algorithms [4] which *effectively* rebalance the dynamic range via tone-mapping operations.

For illustration, let us consider an example image taken from *TM-DIED* [5] - The Most Difficult Image Enhancement Dataset - displayed in Figure 1a (images in this data set are all default tone-mapped poorly by the camera). We see that the image foreground has a very dark appearance with the building and the area in front barely conspicuous but detail in the sky is readily observable. Because the dynamic range of this scene is high the image processing pipeline in the camera hasn't been able to map the shadow and highlight detail to be simultaneously visible. We will now, for illustrative purposes, consider how we might make the shadows brighter and, in so doing, make a new output that is preferred to the input.

An enhanced image, $\underline{O}(x,y)$ is shown in Figure 1b where we can see much more detail in the foreground which, to a large extent, has remedied the illumination problem that was present in the initial image. In Figure 1c we show a light adjustment map, $\underline{L}(x,y)$. The brightest pixel value in this image denotes 'multiplying by 6' and a pixel value of zero means 'multiplying by 1'. Figure 1b is the result of multiplying the input image by the illu-

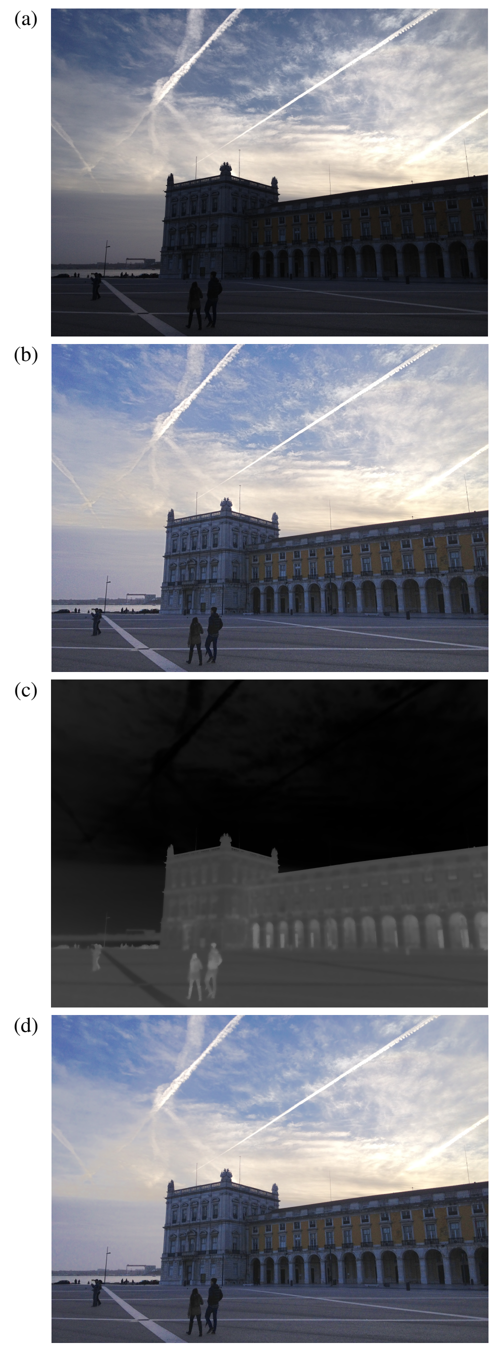


Figure 1: Demonstration of tone mapping a poorly illuminated image. (a) Input, (b) Enhanced output, (c) Illumination map from (a) to (b), (d) Enhanced with a tone curve.

mination map. Written as a formula,

$$\underline{O}(x,y) = \underline{I}(x,y) * \underline{L}(x,y). \tag{1}$$

Importantly, implementing light change as multiplication by a light adjustment map is physically accurate. Adding more light to the shadowed area of a scene would cause all the pixel values recorded for that region to scale by the same multiplicative factor. This said, often light adjustment is cast not as a multiplicative process but as a tone map. In Figure 1d we show the tone mapped output image $\underline{O}_T(x,y)$ where T is the tone mapping function:

$$\underline{O}_T(x,y) = T(\underline{I}(x,y)). \tag{2}$$

A tone map is an increasing function of brightness that is applied to an image. Tone maps effectively serve many purposes: they attempt to map dynamic range (and relight an image), they account for display non-linearities, and they enhance an image in some respects. Whilst tone mapping functions can, in principle, be arbitrary - except the function is increasing - in practice their slopes are constrained [6]. They are not allowed to be too high or too low. Equally, if there are good 'whites' and 'blacks' in an image these are typically preserved after tone mapping [7].

Recently, compelling evidence has been given that tone curves should also be simple: they should have no more than one inflexion point [8]. The gradient of an S-shaped tone-curve, for example, monotonically increases to an inflexion point and then it monotonically decreases (and, so, has one inflexion point). A *wiggly* tone curve is not simple: there will be many points where the slope increases then decreases and vice versa (there are many inflexion points).

In this paper, we ask the question whether tone curves might be simplified even further. Rather than there being one or zero inflexion points, we constrain tone curves to have zero inflexion points. An example of a zero inflexion point curve is one where the slope monotonically decreases across the function's domain, such as a fractional *gamma* curve. We call zero inflexion point curves **Very Simple** tone curves or **VS** curves.

We present a simple optimisation method to approximate an arbitrary tone curve with a VS counterpart. Where the new method is a restriction of the previous simple tone curve formulation [8] and has the advantage that it is much faster to compute. In making our formulation we take care to generate tone curves that do not have sharp gradient changes [9].

We carry out experiments with the MIT-Adobe FiveK dataset. This dataset comprises 5,000 images that are then tone adjusted by 5 experts giving a set of 25,000 tone mapped output images to consider. Previous work has shown that almost all of these output images can be well approximated using a simple tone curve. In this paper, we wish to evaluate how well a VS curve can work. Surprisingly, we find that VS-curves provide a very good approximation to the more wiggly curves sometimes used by photographers.

Finally, the results reported here are metric-based: we say two images look equivalent if they meet a mean CIELAB ΔE threshold (following from [10]).

Background

The MIT-Adobe FiveK (FiveK) Dataset: The FiveK dataset [11] has been widely used in creating and evaluating automatic tone mapping algorithms. The dataset comprises 5,000 images that have been retouched by five experts who each made adjustments according to their preference. The result is 25,000 input-enhanced image pairs containing a diverse range of adjustments which can broadly be approximated by a single global tone curve. Let $\underline{I}$ denote an input image and $\underline{P}$ denote the expert modified output, expressed as a global tone curve adjustment [8]. For each given pixel, at location (x,y), these images are denoted,

$$\begin{aligned} \underline{I}(x,y) &= [L_I^* \; a_I^* \; b_I^*]^\top \\ \underline{P}(x,y) &= [L_P^* \; a_P^* \; b_P^*]^\top \end{aligned} \tag{3}$$

The photographer's tone mapping operation only affects the brightness aspect of an image (only L^* is mapped): $L_P^* = T(L_I^*)$. In this paper we are interested in finding a tone map $\hat{T}$ from input to output subject to additional constraints, $\hat{T}(L_I^*) \approx L_P^*$.

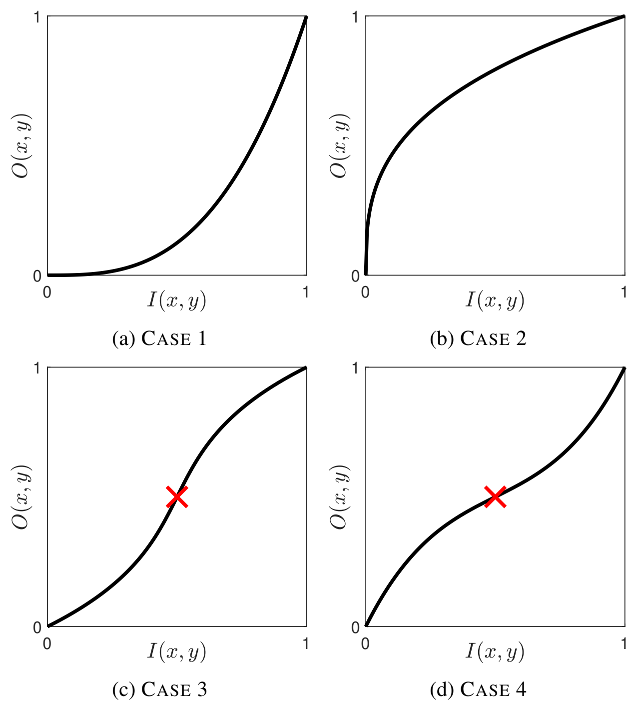


Figure 2: Depiction of the four simple cases, red cross marks the inflexion point

Simple Tone Curves: Previous work [8] has introduced the concept of simple tone curves which are curves that have at most one inflexion point. This leads to four cases of tone curve as illustrated in Figure 2. CASE 1 and CASE 2 have zero inflexion points with the former having an increasing gradient over the whole curve and the latter a decreasing gradient. CASE 3 and CASE 4 curves have one inflexion point which is marked with a red cross in our illustration and whose location can vary over the entire domain. A CASE 3 curve's gradient increases up to the inflexion point then decreases afterwards and CASE 4 decreases to and increases from the inflexion point.

In the context of this paper VS curves are CASE 1 and CASE 2 tone maps. That is, VS curves have zero inflexion points and either have a monotonically increasing or decreasing gradient.

Importantly, previous work [8] analysed the 25,000 expert manual tone adjustments in the FiveK dataset and found that all expert tone curves were inherently simple or could be closely approximated (without any visually important difference).

Method

In Figure 3, we plot the actual tone curve that generated Figure 1d (red) and the simple (one inflexion point) approximation in blue. The curves are almost identical but they are a little different (see inset). In green we show the VS curve that the closest approximation. The images created by tone mapping with the simple and VS curves are shown in panels (a) and (b). Even though the tone curves appear quite different the resulting tone mapped outputs look similar.

Now let us develop a method that ‘solves for’ the best VS approximation to a given tone curve. Let us sample the $[b_1, b_n]$ input brightness domain using in n uniformly spaced values. We can write this as the vector $\boldsymbol{b} = [b_1, b_2, ..., b_n]^\top$, where the superscript $^\top$ denotes the transpose operator. In this discrete representation, when we apply a tone function T the output is also a vector: $\boldsymbol{t} = [T(b_1), T(b_2), ..., T(b_n)]^\top$. We seek a VS tone curve $\hat{\boldsymbol{t}} \approx \boldsymbol{t}$. The derivative of the VS curve must be monotonically increasing or decreasing (see CASE 1 or 2 in Figure 2) as this property enforces the zero inflexion point requirement.

To find the derivative of an n-vector in the discrete domain we apply an $n \times n$ matrix $\mathbf{D}$ defined as

$$\begin{aligned} &\mathbf{D}_{1,1:2} = [-1\ 1] \\ &\mathbf{D}_{i,i-1:i} = [-1\ 1]\ ,\ i \in \{2, 3, ..., n\} \end{aligned} \tag{4}$$

and elsewhere $\mathbf{D}$ is zero. The meaning of the index $_{i,i-1:i}$ is the components in the ith row and the columns $(i-1)$ and i. Note that the first two rows of $\mathbf{D}$ are the same and this implies we expect the derivative to be constant at the boundary of the domain. An increasing or decreasing derivative corresponds to a positive or negative second derivative respectively. Hence, we calculate the second derivative by differentiating twice,

$$\mathbf{D}^2 = \mathbf{D}\mathbf{D}. \tag{5}$$

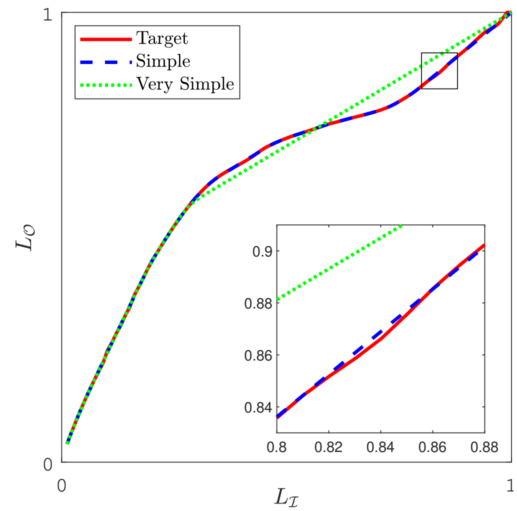
We now find the CASE 2 VS curve that approximates a target curve $\boldsymbol{t}$ by minimising,

$$\arg\min_{\hat{\boldsymbol{t}}} \|\hat{\boldsymbol{t}} - \boldsymbol{t}\| \tag{6a}$$

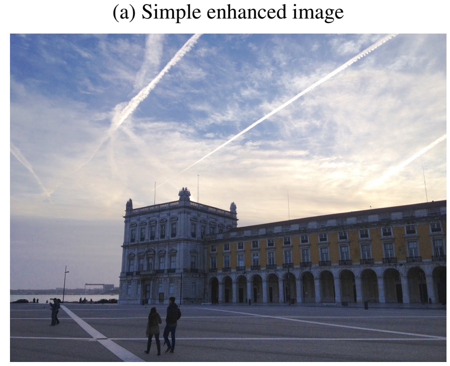
$$\text{s.t.} \begin{cases} \mathbf{D}\boldsymbol{z} \geq \mathbf{0} & \text{(6b)} \\ \mathbf{D}^2\boldsymbol{z} \leq \mathbf{0} & \text{(6c)} \\ \hat{t}_1 = t_1 & \text{(6d)} \\ \hat{t}_n = t_n & \text{(6e)} \end{cases}$$

This minimisation has a single quadratic objective and two inequality and two equality constraints and therefore has a global minimum that can be found using quadratic programming [12, 13]. Constraint (6b) requires the curve we are solving for to be increasing. Note the new variable $\boldsymbol{z}$ that has been introduced which is related to $\hat{\boldsymbol{t}}$ by,

$$\hat{\boldsymbol{t}} = \mathbf{G}\boldsymbol{z}, \tag{7}$$



(a) Simple enhanced image



(b) Very simple enhanced image

Figure 3: Target, Simple and Very Simple tone curve shown in top panel. The simple (a) and very simple (b) renditions are similar to one another.

where the $n \times n$ matrix $\mathbf{G}$ is a Gaussian smoothing matrix which is defined in [9] (the standard deviation $\sigma = 2$ is used in the context of the fact that we are representing tone curves as 100-component vectors) to eliminate sharp corners from the tone curve which could otherwise induce Mach bands in the tone mapped images [14, 15]. The vector $\boldsymbol{z}$ can be thought of as a tone curve equivalent to $\hat{\boldsymbol{t}}$, except that when multiplied by $\mathbf{G}$, it smooths the curve whilst preserving simplicity.

The second constraint (6c) requires that the derivative is decreasing (CASE 2 curve). Constraints (6d) and (6e) respectively ensure the approximate tone curve maps to the same range. We can solve for a CASE 1 tone curve by substituting constraint (6c) for $\mathbf{D}^2 z \geq \mathbf{0}$.

From Equation 3 we have the input images $\underline{I}$ and expert adjusted images $\underline{P}$ and we preprocess all the images by applying a clamping function. The whites of both $\underline{I}$ and $\underline{P}$ are clamped to the 0.999 quantile of their L^* value which removes spurious brightnesses affecting the tone curve. The blacks of $\underline{P}$ are clamped to 0.01 (where $L^* \in [0,1]$). We then use the methodology of [8] to extract the global tone curve $\boldsymbol{t}$. The tone curves used in the optimisation only concern the non-clamped values.

For each tone curve $\boldsymbol{t}$, the optimisation of Equation (6) is then solved. The VS curves $\hat{\boldsymbol{t}}$ and are interchanged with function notation, yielding $\hat{T}$. This draws attention to the fact that we have to interpolate our discrete tone curve in some way when we apply a tone curve to an image. Here we use shape-preserving monotone Piecewise Cubic Hermite Interpolating Polynomials (PCHIP) [16]. When this curve is applied to every pixel of an input image $\underline{I}$, it generates $\underline{\hat{P}}$, where

$$\underline{\hat{P}}(x,y) = [\hat{T}(L_I^*)\ a_P^*\ b_P^*]^\top \tag{8}$$

Results and Discussion

We have 25,000 input and photographer enhanced pairs where each enhancement is encoded as a ground-truth tone curve. We denote a ground-truth tone curve as $\boldsymbol{t}$ and for each of these we solve for its VS counterpart $\hat{\boldsymbol{t}}$ using the optimisation set forth in the last section. The photographers' own adjustments can result in CASE 1 through 4 tone curves (0 or 1 inflexion point) or the adjustment can require multiple inflexion points (and we call these 'wiggly'). In blue in Figure 4, we show the counts of each type of photographer made adjustments. It is noteworthy that only slightly more than a quarter are very simple and that about a third are neither simple nor very simple. The VS curves, by definition, are all CASE 1 and 2. The frequency of the approximate VS curves are shown in red. One can see that almost all of them are CASE 2 (monotonically decreasing derivatives).

Of course, we should only use VS curves - as a proxy for a photographer's own more complex adjustments - if the images they produce look similar to the experts' output images. It is often assumed that a pair of photographic images (of complex scenes) are visually indistinguishable from one another if their mean CIELAB ΔE is less than 3 [10, 17]. Quantitively, we compare the 25,000 $\underline{\hat{P}}$ images to their corresponding expert adjustment $\underline{P}$ by calculating their mean ΔE [18] colour difference. Figure 5 summarises these by showing the the histogram of the 25,000 mean ΔE values with the 50th, 90th and 99th quantiles shown. We observe that 99% of adjustments have a mean ΔE of 3.52 or less which is unlikely to be readily noticeable. There are just 63 adjustments where $\Delta E > 5$ and the maximum is 10.4.

In Figure 6 we show the input $\underline{I}$, expert adjusted image $\underline{P}$, and VS tone mapped image $\underline{\hat{P}}$, alongside the tone curves $\boldsymbol{t}$ and $\hat{\boldsymbol{t}}$ for several example images. The mean ΔE and case classification for each of the images are recorded in Table 1.

For image D2158 the photographer themself makes a VS (CASE 2) adjustment. Unsurprisingly, we find a VS curve works very well. The mean ΔE in Table 1 is really very small and the photographer's own and the VS adjusted output (middle and

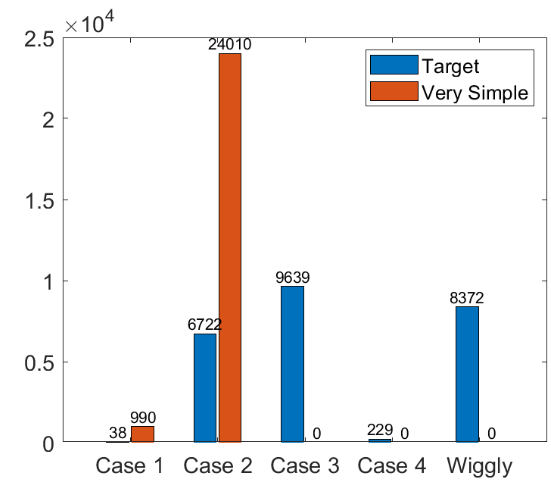


Figure 4: Distribution of case classification for ground truth curves T and VS curves $\hat{T}$.

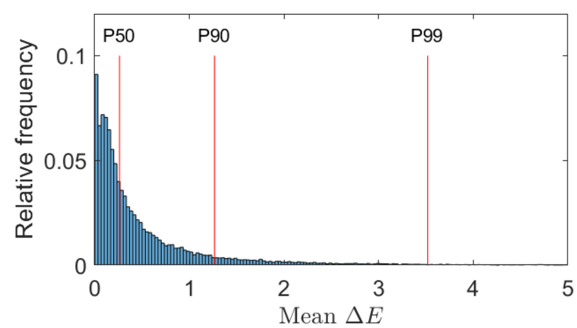


Figure 5: Histogram of mean ΔE for all 25,000 tone adjustments with 50th, 90th and 99th percentile shown. There are 63 results $\Delta E > 5$, with maximum at 10.4.

right) in Figure 6 are identical. Note, we do not find the precisely the same curve as the photographer as we have the additional Gaussian smoothing constraint in our problem formulation. Images E4830 and A3067 are wiggly (photographer adjustments) replaced by a VS curves of CASE 2 and 1. The mean ΔE is small and the VS-curve and photographers' adjustments look identical.

Image C3220 is represents the 90th percentile mean ΔE error and the photographer's adjustment is a CASE 3 curve which is approximated by a VS CASE 2 curve. Even for this 90th percentile *hard* case the mean ΔE difference is just 1.26 and the images look visually almost the same. There are a few images, such as A3905 which is the 15th worst image as ranked by the mean ΔE where a VS curve does not approximate the ground truth target satisfactorily. The best VS curve is CASE 2 but $\underline{\hat{P}}$ and $\underline{P}$ are 8.04 ΔE apart. Here, the VS curve enhanced image is visually quite different to that produced by the photographer's own adjustment.

Each of the 5,000 input images are adjusted by 5 experts and each expert creates an image rendition that can be notably different from another expert. Per input image, we calculate the median ΔE error to summarise how well a VS curve works across the 5 expert renditions. We now, as before, calculate the histogram

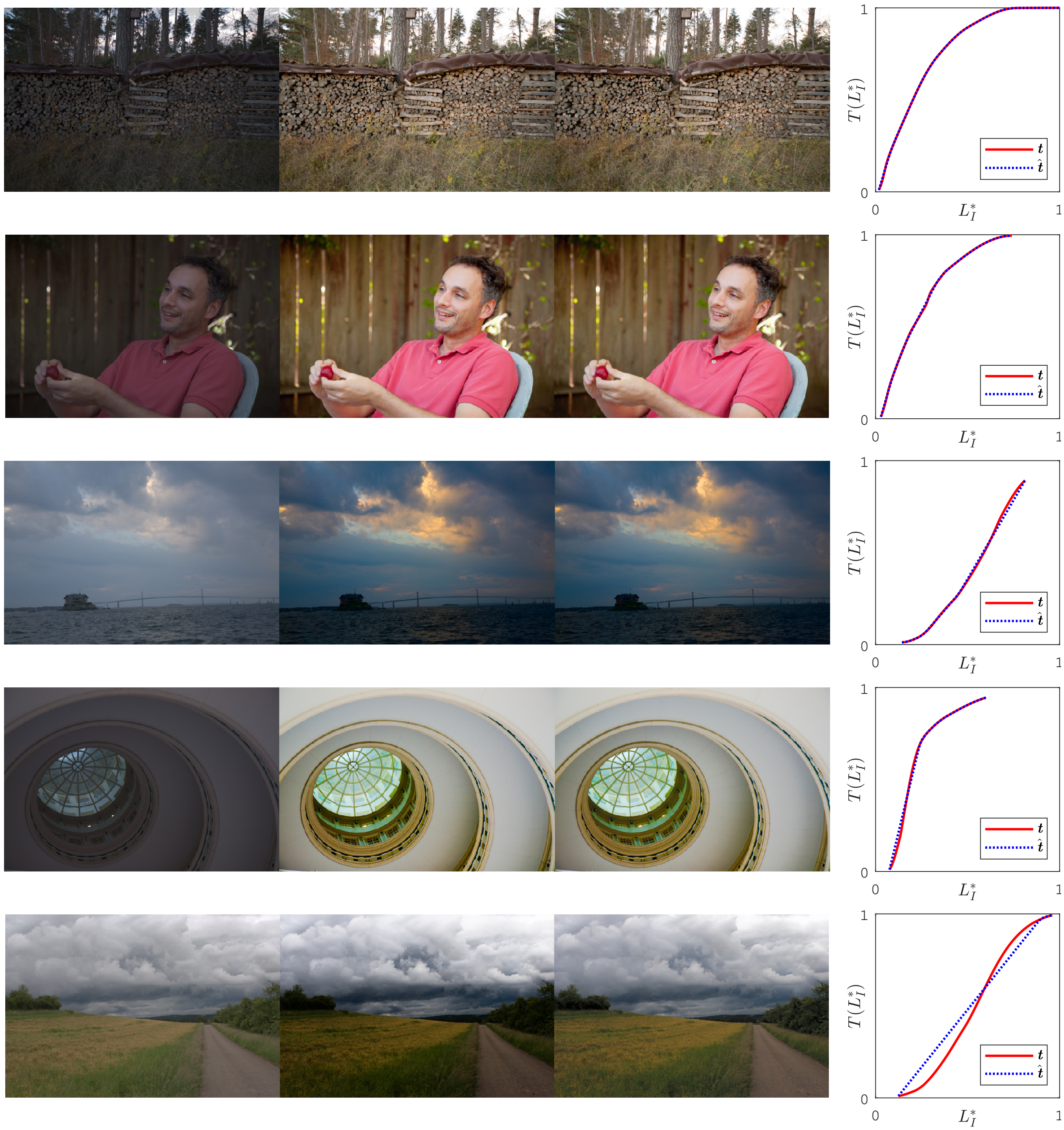


Figure 6: Example results showing images (top-bottom) D2158, E4830, A3067, C3220 and A3905. From left to right, displays the input image, a photographer's own tone mapped output and an VS curve approximation.

Table 1: Statistics for images in Figure 6.

Image	Mean ΔE	M. ΔE Rank	$\underline{P}$ case	$\hat{P}$ case
D2158	0.0762	1129	2	2
E4830	0.279	13029	Wiggly	2
A3067	0.699	19549	Wiggly	1
C3220	1.264	22494	3	2
A3905	8.04	24995	3	2

for the 5,000 mean ΔE errors and this is shown in Figure 7. The error statistics are even more favourable than before as almost all images are less than the mean $\Delta E = 3$ criterion [10, 17]. The largest error is just 4.95. In summary, Figure 7 teaches that for all images there are VS adjustments that produces a visually similar result, on average.

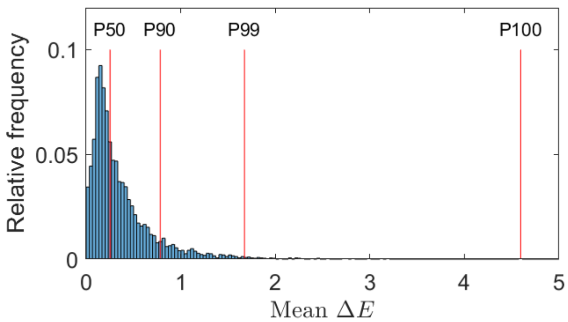


Figure 7: Histogram of mean ΔE for the for median expert adjustment, showing 50th, 90th, 99th and 100th percentile.

Conclusion

Tone curves are an important tool in image processing pipelines and can be used to manipulate brightnesses in images, compressing dynamic range, making details more conspicuous and the image more pleasing. Previous work has proposed that tone curves should be simple (not wiggly) which is technically expressed by the tone curve having no more than one inflexion point (informally this means tone curves cannot be wiggly). In this paper, we have proposed that only zero inflexion point tone curves might be used - we call these very simple (VS) curves. An optimisation was presented to find the VS tone curve that best approximates an arbitrarily shaped target tone curve. With this algorithm in hand, we can approximate a given tone curve (from a user's own adjustments or from an algorithms) with its best VS curve approximation. We perform experiments on the MIT-Adobe FiveK dataset which comprises 25,000 input and tone curve enhanced image pairs. We solve for the VS tone curve that best approximates each of the 25,000 expert tone adjustments. Our experiments show that a very simple VS tone curve can well approximate almost all of the tone adjustments made by experts. In almost all cases - even when a photographer used a wiggly tone curve - the best VS approximation looks visually almost indistinguishable from the photographer's own enhancement.

Acknowledgments

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

References

[1] L. Jiao and J. Zhao, "A survey on the new generation of deep learning in image processing," *IEEE Access*, vol. 7, p. 172231, 2019.

[2] R. Archana and P. E. Jeevaraj, "Deep learning models for digital image processing: a review," *Artificial Intelligence Review*, vol. 57, no. 11, 2024.

[3] E. H. Land and J. J. McCann, "Lightness and retinex theory," *Journal of the Optical society of America*, vol. 61, no. 1, pp. 1–11, 1971.

[4] Y. Salih, A. S. Malik, N. Saad *et al.*, "Tone mapping of HDR images: A review," in *International Conference on Intelligent and Advanced Systems*, 2012, pp. 368–373.

[5] V. Vonikakis, "TM-DIED: The most difficult image enhancement dataset," 2021. [Online]. Available: https://sites.google.com/site/vonikakis/datasets/tm-died

[6] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, pp. 355–368, 1987.

[7] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Transactions on image processing*, vol. 18, no. 9, pp. 1921–1935, 2009.

[8] J. Bennett and G. Finlayson, "Simplifying tone curves for image enhancement," in *Color and Imaging Conference*, 2023, p. 108.

[9] ——, "Even simpler tone curves," in *London Imaging Meeting*, 2024, pp. 124–128.

[10] M. Stokes, M. D. Fairchild, and R. S. Berns, "Precision requirements for digital color reproduction," *ACM Transactions on Graphics*, vol. 11, pp. 406–422, 1992.

[11] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, "Learning photographic global tonal adjustment with a database of input/output image pairs," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2011, pp. 97–104.

[12] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*. Academic Press Inc., 1999.

[13] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Prentice-Hall, 1974.

[14] F. Ratliff, *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. Holden-Day, 1965.

[15] J. P. Thomas, "Threshold measurements of mach bands," *Journal of the Optical Society of America*, vol. 55, no. 5, pp. 521–524, 1965.

[16] F. N. Fritsch and R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM Journal on Numerical Analysis*, vol. 17, no. 2, pp. 238–246, 1980.

[17] M. K. E. Mahy, L. V. Eycken, and A. Oosterlinck, "Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV," *Color Research & Application*, vol. 19, no. 2, 1994.

[18] "Colorimetry — Part 4: CIE 1976 L*a*b* colour space," CIE International Commission on Illumination, Standard, Jun. 2019.

Author Biography

James Bennett is a PhD student at the University of East Anglia in the Colour & Imaging Lab. He is a member of the AgriFoRwArdS CDT and is working in collaboration with Antobot on imaging in agriculture. Professor Graham Finlayson is the Director of the Colour & Imaging Lab. He has published over 200 conference papers, over 75 journal papers and is the inventor of 30+ patents. His interests span perception, colour image processing and physics-based computer vision.