

Learning to Display High Dynamic Range Images

Jiang Duan and Guoping Qiu⁺

School of Computer Science, The University of Nottingham, United Kingdom

Graham Finlayson

School of Computing Science, The University of East Anglia, United Kingdom

Abstract

In this paper we present a novel method to map high dynamic range scenes to low dynamic range images for visualization. We formulate the problem as a quantization process and employ an adaptive learning strategy to ensure that the low dynamic range displays not only faithfully reproduce the original scenes but also are visually pleasing. This is achieved by the use of a competitive learning neural network that employs a frequency sensitive competitive learning mechanism. An L_2 objective function ensures that the mapped low dynamic image preserves the relative visual contrast impressions of the original scene. A frequency sensitive competitive mechanism facilitates the full and even utilization of the limited displayable values. We present experimental results to demonstrate the effectiveness of the method in displaying a variety of high dynamic range scenes.

1 Introduction

The human visual system is capable of perceiving scenes ranging five orders of magnitude, and can gradually adapt to brightness of over nine orders of magnitude. Advanced computer technology can produce high dynamic luminance maps of synthetic graphics and real world scenes in numerical form. However, the media used to display digital photographs and graphics, e.g., a video display or a printer, can only display data of no more than a few orders of magnitude.

With the rapid advancement of electronic imaging and computer graphics technologies, there have been increasing interests in high dynamic range imaging.¹⁻¹⁰ Figure 1 shows a scenario where high dynamic range photographic will be useful. This is an indoor scene with very high radiance contrasts. In order to see features in the dark areas, longer exposure intervals have to be used, but this renders the bright areas saturated. On the other hand, using shorter exposure intervals enables features in the bright area to be visible but obscures features in the dark areas. In order to make all features both in the dark and bright areas simultaneously visible in a single image, we can create a high dynamic range radiance map for the

scene that captures the full dynamic range of the scene's radiance. It turns out that it is relatively easy to create high dynamic range maps of high contrast scenes.³ Using the technique of Ref. [3], all one needs is a sequence of low dynamic range photos of the scene taken under different exposure intervals.

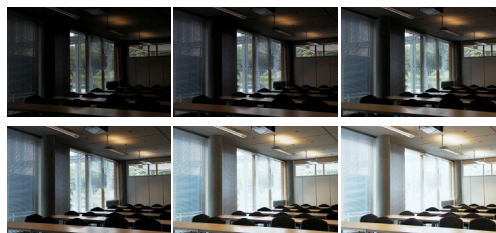


Figure 1. Low dynamic range photos of an indoor scene taken with different exposure intervals. Short exposure makes the picture too dark whilst long exposure renders the picture too bright.

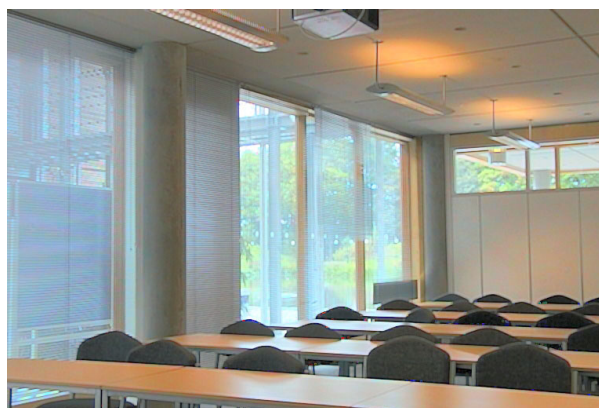


Figure 2. Low dynamic display of high dynamic range radiance map of the scene in Figure 1 using FSCL mapping.

However, because reproduction devices normally have much lower dynamic range than the radiance maps (or equivalently real work scenes), one of the key technical issues is how to map high dynamic range scene data to low

⁺ Corresponding Author (Email: qiu@cs.nott.ac.uk)

dynamic range display in such a way that the visual impressions (as perceived by the human observers) of the original real physical scenes are faithfully reproduced. In the literature, a number of methods have been proposed, ranging from simple quantization,² to medium complexity tone reproduction curve,⁴ and to computationally more demanding tone reproduction operators,⁷ for the display of high dynamic range scenes.

In this paper, we propose a novel learning based method to map high dynamic range scene to low dynamic image to be displayed in low dynamic range devices. This is the first method that adaptively learns from the high dynamic range data to produce a mapping model. We use frequency sensitive competitive learning (FSCL)¹³ to ensure that the mapping not only takes into account the relative brightness of the pixel values, but also fully exploits all displayable values, thus reproducing a copy of low dynamic image that preserves the visual impression of the original scene and is also visually pleasing.

The organization of the paper is as follows. In the next section, we describe our method. In section 3, we present experimental results to demonstrate the effectiveness of the new method. Section 4 concludes our presentation.

2 Learn to Display High Dynamic Range Images

The process of displaying high dynamic range image is in fact a quantization and mapping process as illustrated in Figure 3. Because there are too many (discrete) values in the high dynamic scene, we have to reduce the number of possible values, this is a quantization process. The difficulty faced in this stage is how to decide which values should be grouped to take the same value in the low dynamic display. After quantization, all values that will be put into the same group can be annotated by the group's index. Displaying the original scene in a low dynamic device is simply to represent each group's index by an appropriate display intensity level. In this paper, we only concerned ourselves with the first stage, i.e., to develop a method to best group high dynamic range values.

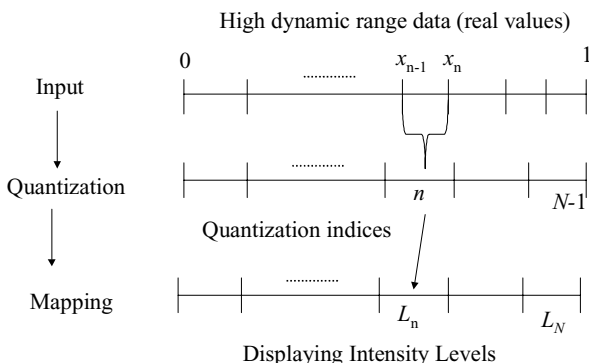


Figure 3. The process of display high dynamic range scene (from purely a numerical processing's point of view)

Quantization, also known as clustering, is a well-studied subject in signal processing and neural network literature. Well known techniques such as k-means and various neural networks based methods have been

extensively researched.¹¹⁻¹⁴ One particular method that is well suited to our application, is the frequency sensitive competitive learning (FSCL) method.¹³ We will first briefly explain the technique, and explain why this is a good method for our application in Section 3.

A vector quantizer is described by an encoder \mathbf{Q} , which maps the k-dimensional input vector X to an index $n \in N$ specifying which one of a small collection of reproduction vectors (codewords) in a codebook $C = \{C_n; n \in N\}$ is used for reconstruction, and there is also a decoder, \mathbf{Q}^{-1} , which maps the indices into the reproduction vectors, i.e., $X' = \mathbf{Q}^{-1}(\mathbf{Q}(X))$.

There are many methods developed for designing the VQ codebook. The k-means type algorithms, such as the LBG algorithm,¹¹ and neural network based algorithms, such as the Kohonen feature map¹² are popular tools. In this work, we used a specific neural network-training algorithm, the frequency sensitive competitive learning (FSCL) algorithm¹³ to design a quantizer for the mapping of high dynamic range scenes to be displayed in low dynamic devices. To understand why FSCL is ideally suited for such an application, we first briefly describe the FSCL algorithm:

FSCL VQ Design Algorithm

1. Initialise the codewords, $C_i(0)$, $i = 0, 2, \dots, N-1$, to random numbers and set the counters associated with each codeword to 1, i.e. $F_i(0) = 1, i = 0, 2, \dots, N-1$
2. Present the training sample, $X(t)$, where t is the sequence index, and calculate the distance between $X(t)$ and the codewords,

$$D_i(t) = \|X(t) - C_i(t)\|$$

and modify the distance according to

$$D_i^*(t) = F_i(t)D_i(t)$$

3. Find j , such that

$$D_j^*(t) \leq D_i^*(t) \quad \forall i$$

4. Update the codeword and counters as

$$C_j(t+1) = C_j(t) + \eta(X(t) - C_j(t))$$

$$F_j(t+1) = F_j(t) + \alpha$$

where $\alpha > 0$ and $0 < \eta < 1$

5. If converged, then stop, else go to 2.
-

The philosophy behind the FSCL algorithm is that, when a codeword wins a competition, the chance of the same codeword winning the next time round is reduced, or equivalently, the chance of other codewords winning is

increased. The end effect is that the chances of each codeword winning will be equal. Put it the other way, each codeword will be fully utilized.

3 Results and Discussions

Similar to work of other authors, in our experiments, we only mapped the dynamic range of the luminance, and color correction for the images was not performed. In fact, FSCL was used to design a scalar quantizer for the mapping of pixels in the luminance channel only. Our method belongs to the global (spatially invariant) tone reproduction curve (TRC) category of techniques.⁴ Even though our method can be easy used alongside spatially variant tone reproduction techniques such as layered methods,¹⁹ we found that by simply sharpening the resultant low dynamic range images of our method produces equally good results.

3.1 Results

We have applied our method to a number of high dynamic range rendered graphics and photographs. Some of the original data are publicly available from the web pages of the creators and others were kindly supplied to us by various authors as acknowledged in the figures. Many authors have also put their results on their web pages to allow easy comparison of algorithms. All our results in this document plus more are available on this web site address.*

Examples of our results and those of several other methods are shown in Figures 2, and 4-11. Note that some results are mapped using the models trained on their own data, some results are mapped using models trained using other images. It turns out that the models can be pre-trained, that is, the mapping models trained using one image (or a group of images) can be used to mapped those images that are not part of the training data[†]. Visual inspection of our results indicates that they are comparable to those in public literature.

3.2 Discussions

In the context of using the FSCL for mapping high dynamic range data to low dynamic display, the limited number of low dynamic levels (the codewords) will be fully utilized. The intuitive result of full utilization of display levels is that the displayed image will have good contrast, which is exactly what is desired in the display. However, unlike histogram equalization, the contrast is constrained by the original scene intensity values. The overall effect of such a mapping is therefore that, the original scenes are faithfully reproduced, while at the same time, the low dynamic range displays are well contrast.

The cost function that the learning process minimizes can be expressed as follows¹⁴:

$$E = \sum_t Y_t(j) \|X(t) - C(j)\|$$

* <http://www.cs.nott.ac.uk/~qiu/LearningIP/LearningIP.html>

[†] In the neural networks literature, this is known as the generalization capability, which is one of the essential properties of neural networks that enables a model trained on training samples to function on unseen novel data.

where

$$Y_t(j) = \begin{cases} 1, & \text{if } \|X(t) - C(j)\| \leq \|X(t) - C(i)\|, \forall i \\ 0, & \text{otherwise} \end{cases}$$

Informally, the FSCL process can be viewed as a constrained optimization operation that minimizes a cost function of the following form:

$$J = E + \lambda \sum_j \left(\sum_t Y_t(j) - \frac{\|X(t)\|}{N} \right)^2$$

where λ is a nonnegative Lagrange multiplier constant, $\|X(t)\|$ represents the total number of training samples (pixels), N is the number of codewords (display levels).

Minimizing the first term ensures that pixel values that are close to each other will be group together and will be displayed as the same single brightness. The mapping also ensures that brighter pixels in the original image are mapped to brighter displaying levels and darker pixels are mapped to a darker display levels[‡] thus ensuring the correct brightness order to avoid visual artifacts such as the “halo” phenomenon. Minimizing the second term facilitates that each available displayable level in the mapped image will be allocated similar number of pixels thus creating a well contrast display. However, it is important to note that this well-contrast-ness is constrained by the minimization of E .

Because the mapping not only desires an equally distributed pixel distribution but also is constrained by the distances between pixel values and the mapped codeword values, the current method is fundamentally different from both traditional histogram equalization and simple quantization based methods. Histogram equalization[§] only forces the mapped pixels to be evenly distributed regardless of their relative brightness. Simple quantization based mapping (including many linear and nonlinear mapping techniques) only take into account the pixel brightness value whilst ignore the populations in each mapped level. As a result, histogram equalization will create false contrast resulting many saturated areas and simple quantization-based method will create mappings with under utilized display levels which often resulting in many features invisible.

[‡] Notice that the values of the codewords have been sorted in increasing order such that $C(0) < C(1) < \dots < C(N-1)$ to ensure correct brightness order in the mapped image.

[§] In fact histogram equalization is inappropriate for the display of high dynamic range images in general. Histogram equalization is only useful when the original data has low contrast and histogram equalization can be used to *enhance* such data. However, high dynamic range data itself is normally well contrast, we do not want to enhance its contrast but instead the purpose is to compress the dynamic range and preserve the original contrast. Given that mapping to lower dynamic range will inevitably reduce the contrast, a mechanism has to be introduced to ensure good contrast display to maintain visual pleasantness.

As a result of the two constraints, the new method achieves a mapping that is both faithful to the original scenes in terms of preserving relative brightness orders and renders visually pleasant display. Such an idea can be served as a general principle for mapping high dynamic range data for low dynamic range display. In the current paper, such idea is elegantly realized by an adaptive learning method. We conjecture that other techniques can also be developed, so long as they follow this principle, good results can be expected.

Although not a focus of this paper, it is important to mention how we render our quantized images for display. So far, we have taken an image and quantized it into 8 bits. How should we use these quantized values to drive a CRT? Well in fact we choose to do nothing at all! The rationale for this comes from what is known about lightness perception in our own vision. Using the CIE standards, it is known that if we are to display a uniform grey scale on a monitor then we choose uniform steps in Lightness and then display these Lightness values (by carrying out the correct calibration).

Let us consider our quantised values to be quantized 'Lightness' values (since we wish each level to be visually of equal importance). In the CIE standard, Lightness is defined to be the cube root of a linear Luminance signal. It follows then, that to linearise our processed data we would have to raise it to the power of 3. However, we note that the transfer function of a typical monitor is often modelled as a power around 2.2 to 2.4. That is, the monitor transforms our perceptual measures into almost the correct linearised signal for display.

Of course if the target for images is a medium other than a CRT monitor then we may have to carry out a calibration and convert our processed images accordingly.

4 Conclusions

In this paper, we have presented a new method, which learns to map high dynamic range image data to be displayed in low dynamic range devices. The frequency sensitive competitive learning (FSCL) algorithm, which ensures that all the available levels in the low dynamic range display are fully utilized and preserves the lightness orders of the original scene, has been shown to give very good performances. As a general principle for mapping high dynamic range scenes for low dynamic range display, it is important to ensure that the mapping should be

faithful to the original scene by preserving their relative visual contrast and to ensure that the mapped image fully utilizes all available display levels to counter the inevitable loss of contrast caused by dynamic range compression.

References

1. K. Chiu, et al, "Spatially nonuniform scaling functions for high contrast images", in Proceedings of Graphics Interface'93, pp. 245 -253
2. C. Schlick, "Quantization techniques for the visualization of high dynamic range pictures", in Proceedings of 5th Eurographics Workshop on Rendering, June 1994
3. P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs", In Proceedings of SIGGRAPH'97, pp. 369 - 378
4. G. Ward, H. Rushmeier and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes", IEEE Trans on Visualization and Computer Graphics, vol. 3, pp. 291 - 306, 1997
5. J. Tumblin and H. E. Rushmeier, "Tone reproduction for realistic images", IEEE Computer Graphics and Applications, vol. 13, pp. 42 - 48, 1993
6. J. Tunblin, J. K. Hodgins and B. K Guenter, "Two methods for display of high contrast images", ACM Trans on Graphics, vol. 18, pp. 56 - 94, 1999
7. R. Fatah, D. Lischinski and M. Werman, "Gradient domain high dynamic range compression", in Proceedings of SIGGRAPH 2002
8. E. Reinhard, M. Stark. P. Shirley and J. Ferwerda, "Photograph tone reproduction for digital images", in Proceedings of SIGGRAPH 2002
9. F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high dynamic range image", in Proceedings of SIGGRAPH 2002
10. M. Ashikhmin, "A tone mapping algorithm for high dynamic contrast images", in Proceedings of 13th Eurographics Workshop on Rendering, 2002
11. A. Gersho, R. M. Gray, Vector quantization and signal compression, Kluwer Academic Publishers, Boston, 1992
12. T. Kohonen, Self-organization and associative memory, Berlin: Springer-Verlag, 1989
13. S. C. Ahalt, et al, "Competitive learning algorithms for vector quantization", Neural Networks, vol. 3, pp. 277-290, 1990
14. S. Haykin, Neural Networks: A comprehensive foundation, Prentice Hall International, Inc, 1999



Figure 4. Left: Result of FSCL mapped image. Right: Results of Ashikhmin's tone mapping method.¹⁰ Data courtesy of Greg Ward⁴



Figure 5. Top: Result of FSCL mapped image. Bottom: Results of Schlick's rational sigmoid quantization technique.² Data courtesy of Sumant Pattanaik

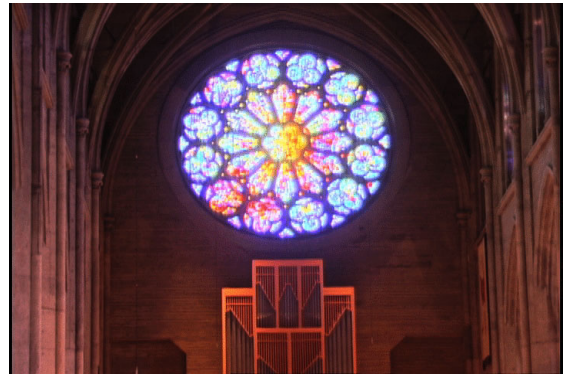


Figure 6. Top: Result of FSCL mapped image. Bottom: Result of Reinhard et al.⁸ Data courtesy of Paul Debevec USC³



Figure 7. Left: Result of FSCL mapped image. Right: Results of Ashikhmin's tone mapping method.¹⁰ Data courtesy of Paul Debevec USC³



Figure 8. Top: Result of FSCL mapped image. Bottom: Results of Durand and Dosey.⁹ Data courtesy of Fredo Durand, MIT.⁹

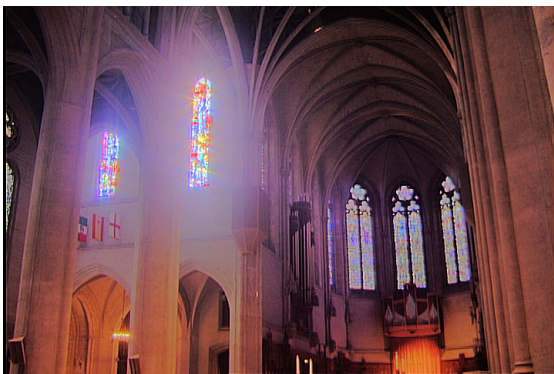
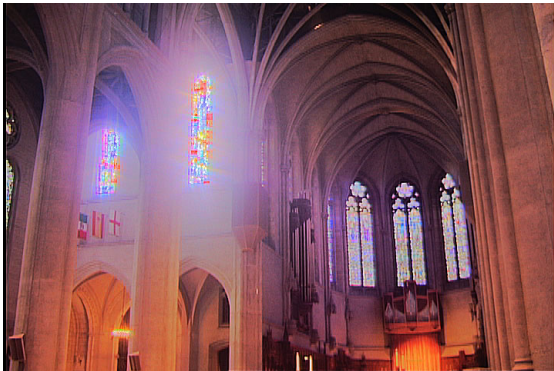


Figure 9. Top: Result of FSCL mapped image, mapping model trained on its own data. Bottom: Result FSCL mapped image, mapping model trained on data from the image in Figure 6. Data courtesy of Paul Devecic USC.³



Figure 10. Result FSCL mapped image, mapping model trained on data from the image in Figure 11. Data courtesy of Greg Ward.



Figure 11. Result FSCL mapped image, mapping model trained on data from the image in Figure 10. Data courtesy of Greg Ward.