Transforming 3D Colour Histograms of Images

Ján Morovic and Pei–Li Sun
Colour & Imaging Institute, University of Derby, Derby, United Kingdom

Abstract

The aim of this paper is to describe a method for transforming an image's 3D colour histogram so as to accurately match a predetermined target state. The method proposed here consists of a colour indexing stage followed by the determination of a histogram transformation matrix on the basis of the Earth Mover's Distance histogram difference metric.[1] In addition to describing this method, the paper also analyses the results of transforming an image's 3D colour histogram in terms of its effect on the image's spatial characteristics. Finally, examples of using the technique to transform an image's histogram to match that of another image are shown. The purpose of having developed this approach is to be able to perturb this image characteristic in an accurate, direct and controlled way as this can be of use in studies that aim to study the impact of image characteristics on various imaging contexts, like colour reproduction or database indexing and searching. Being able to modify the 3D colour histograms of an image then allows for the generation of image test sets in which images have this characteristic in arbitrary states.

Introduction

Being able to transform image characteristics in an arbitrary but accurate and direct way can be a powerful tool in the study of their impact on various imaging contexts.2 For example, to understand the impact of an image's mean colour on some application that is behaving in an image–dependent way it is useful to have a set of images whose mean colours have various values but whose other unrelated characteristics are identical or at least very similar. What is meant here by unrelated image characteristics are those whose states do not necessarily change when then given image characteristic is changed. In the case of the mean image colour characteristic unrelated characteristics would, for example, be image content, image gamut, spatial properties, etc., as the mean colour can be changed without resulting in a change of these unrelated characteristics.

Clearly it is very difficult to find a set of natural images that differ only by their average colours, whereby what is meant here by natural images are images that are directly the result of some image capture or generation process and that would look unperturbed to an observer (iDE. they include not only images captured from nature but also computer graphics that look 'normal' to observers). A first step then in understanding this question can be to chose a single natural image and to generate a set of artificial images from it by perturbing it in a way that changes the mean colour but keeps other unrelated image characteristics as constant as possible.

Methods for perturbing some image characteristics in this way have already been introduced in some previous work[3–5] and the aim of this paper is to describe an approach that can be used for changing an image's 3D colour histogram so as to make it match a predetermined target state accurately.

Using this method, it is then possible to transform natural images and generate sets of test images that have predetermined 3D colour histogram properties. Such sets of test images can then be used in studying the impact of this image characteristic on any imaging application that exhibits image dependent behaviour and to see what role the 3D colour histogram plays in this.

Overview Of 3D Histogram Matching

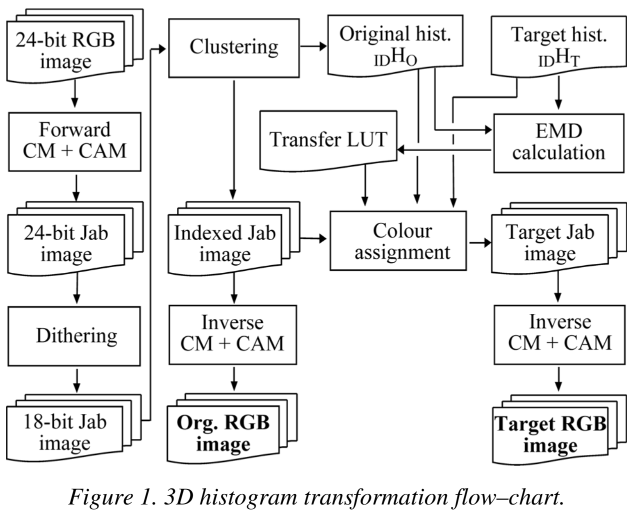

Figure 1. 3D histogram transformation flow–chart.

The method of 3D histogram matching proposed here is shown as a flow chart in Figure 1. The kernel of the method is the Earth Mover's Distance (EMD) metric[1] which provides an optimised relationship between two histograms by minimising the sum of cross-bin distances between original and target histograms. Furthermore, the ΔE formula can be used to evaluate cross-bin distances for any colour pair between a pair of 3D image histograms. Ideally, the 3D image histogram should have as many bins as there are possible colour combinations, iDE. 2^{24} for three–channel 8 bits/channel images. However, the computational cost would be extremely high as a matrix with 2^{24} x 2^{24} (iDE. 2^{48} = 2.8 x 10^{14}) members in float format would be used by EMD. In order to make computation more manageable, images can first be converted to an indexed colour format, which reduces the number of colours in an image using a quantisation technique, and only this smaller set of colours is then used for building the histograms.

The following are then the steps of this method that can be used for transforming an image displayed on a CRT to have a predetermined 3D colour histogram in terms of the CAM97s2 Jab colour space[6]:

1. Transform a 24-bit original image from RGB to a uniform colour space (e.g. CAM97s2 Jab) using a forward CRT characterisation model (CM), e.g. GOG,[7] in conjunction with a forward colour appearance model (CAM) transform.
2. Transform the Jab image from 24-bits (8-bits per channel) to 18-bits (6-bits per channel) by performing Floyd–Steinberg error diffusion.[8] The reason for using this algorithm is to reduce sample size for the next step while preserving most of the colour information from the original. Since the (J,a,b) intervals of the resulting 18-bit image were only (1.57,4,4) CAM97s2 units, the image colour differences between the original and the image obtained after this step were almost imperceptible.
3. Transform the 18-bit image to N indexed colours (N values of 256, 512, 1024 and 3072 were investigated in this study) using a modified Foray clustering algorithm[9] whereby the resulting N colours are chosen differently for each image. In the clustering process, colours are added to the cluster having the smallest Euclidean distance in Jab colour space.
4. Compute image histograms (${}_{ID}H$s) in terms of the N indexed colours. The data structure of the ${}_{ID}H$ consists of index number, corresponding Jab values and colour frequency as a percentage of total number of pixels.
5. Use EMD to calculate optimal cross–bin distances between the histogram of an original image's (${}_{ID}H_O$) and the target histogram (${}_{ID}H_T$). The output of EMD is then a LUT indicating the net–like relationship of all bins between the two histograms.
6. Compute the final 3D-histogram matched image by reading the Jab value of each pixel from the original image, checking the index value of the pixel by referring to ${}_{ID}H_O$, calculating the corresponding index value in ${}_{ID}H_T$ based on the EMD–generated LUT and finally assigning the corresponding Jab value to the resulting pixel in the output image. A random colour assignment technique was performed in this step by randomly accessing the LUT.[4]
7. Both the indexed original and resulting target images are finally converted to RGB using the inverse CAM97s2 and the inverse CRT characterization model. When displayed on the CRT, these RGB images then have the same 3D colour histograms in terms of CAM97s2 Jab.

Using the above process, the output image has exactly the same 3D colour histogram as ${}_{ID}H_T$. However, when the original and target histograms are very different, the output image can show strong artefacts and has a more ample power spectrum in the high spatial frequency range. In order to minimise such artefacts and to reduce the power of high spatial frequencies, an attempt can be made to optimise the steps of the above procedure. The success of the attempts to optimise some parts of the process was assessed in terms of a local colour spatial frequency metric and this will be introduced next.

Local Colour Spatial Frequency

After performing 3D-histogram matching, the high spatial frequencies in the power spectrum of the histogram–perturbed image normally have higher amplitudes than corresponding ones in the original. It would be ideal if the histogram–perturbed image could maintain not only the image content of the original image but also the power spectrum distribution of local colour spatial frequencies (LCSFs) in it. As the original image changes its colours totally after histogram matching, no colour difference metric can estimate the variation of its power spectrum and a separate metric for measuring these errors has therefore been developed.

To evaluate the variation of the power spectrum, both the original and histogram–perturbed Jab images are first subdivided into blocks of 16 x 16 pixels and 50 units are subtracted from all the J values. Each block is then transformed to a 16 x 16 spatial frequency matrix F(u,v) using a 2D fast Fourier transform (FFT) for J, a and b respectively. The logarithmic power spectrum logP(u,v) of F(u,v) is computed and averaged across all orientations (ψ) and neighbouring spatial frequencies ($w\pm1$) to yield a discrete 1-D function logP(w) of radial spatial frequency w. The differences between each logP(w) pair of original and perturbed images are then averaged for the whole image and the resulting difference function ΔlogP(w) can further be reduced to a single value $\Delta \log P_{Jab}$ by averaging the discrete values of ΔlogP(w) (where w = 1,2...11) and finally averaging the results of the three colour channels (J, a, and b). Since the spatial frequency of the DC component of the Fourier spectrum is zero (w=0), the corresponding error values (ΔlogP(0)) should not be used for evaluating LCSF.

$\Delta \log P_{Jab}$ can be regarded as a metric for evaluating the difference between images in terms of LCSF. The reason for using a logarithmic scale is because it gives the values for each spatial frequency a similar weight and is therefore more suitable for calculating overall errors. This way of calculating LCSF is partly based on Thomson and Foster's application of higher-order image statistics.10 Note that zero $\Delta \log P_{Jab}$ means that both images have the same distribution of spatial frequency variation.

The LCSF of an image also can be illustrated by plotting spatial frequency w against the logarithmic power spectrum (logPJab(w)). Figure 2 shows a series of images with different spatial treatment (blurring, sharpening and adding noise to the original) and their corresponding LCSFs.

Optimising 3D Histogram Matching

As could be seen from its description, this method first obtains an indexed–colour version of a given 24 bit image. Here the four elements that determine the performance of the transformation are the colour space, the dithering and clustering methods and the number of indexed colours used. As the first three of these are well understood and as the impact of the number of indices used on the level of artefacts is simple, this section will focus on the optimisation of the subsequent 3D histogram transformation itself.

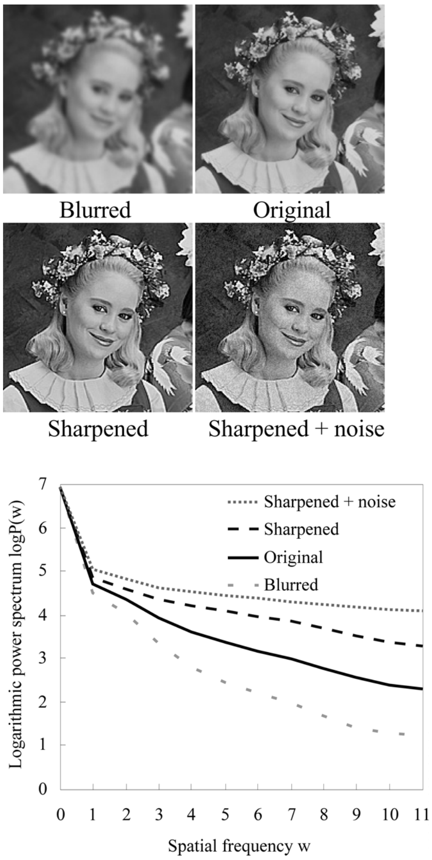

Figure 2. Images with different spatial treatment and their LCSFs.

In the process of transforming the 3D histogram of an image there are a number of parameters that can be adjusted and it is therefore important to try and set them so as to reduce artefacts in the resulting images. The most important of these parameters are: (a) the power to which ΔE values are raised in EMD (i.e. whether EMD minimises ΔE or ΔE^n), (b) the weights of ΔJ, ΔC and ΔH in the ΔE formula, (c) the number of indices used for colour quantisation.

To evaluate the effect of varying the power of ΔE in the EMD calculation, ΔE, ΔE^2 and ΔE^3 were used for generating all the histogram matching images by using images with 512 indexed colours and overall results in terms of $\Delta \log P_{Jab}$ were 0.69, 0.64 and 0.68 respectively. This means that in general the LCSF difference between original and histogram-perturbed image pairs was reduced by using the ΔE^2 function. As some studies suggest that weighted ΔE formulæ are preferred for gamut clipping11 a $\Delta E^2_{wt}(1{:}2{:}1)$ function was also evaluated with EMD where the weights dividing [ΔJ,ΔC,ΔH] were [1,2,1]. The testing process was the same as used above and the resulting $\Delta \log P_{Jab}$ for $\Delta E^2_{wt}(1{:}2{:}1)$ was 0.65, however, as this metric was lower for the unweighted ΔE2 the latter should be used.

Different numbers of indices were also tested (256, 512, 1024 and 3072) and this showed that using 1024 bins gives the best result (Table 1). As can be seen, using 3072 indexed colours didn't improve the quality of resulting images. A possible reason for this is that colours of uniform areas in the original were transformed to have a larger number of target colours when increasing the index numbers. Hence, the uniformity of some areas in the original can be reduced more dramatically if there is a greater number of target indexed colours assigned to them and if these indexed colours are from different parts of colour space.

Table 1. Effect of changing number of indexed colour used in terms of $\Delta \log P_{Jab}$. Refer to Figure 3, CG-to-MUS here, for instance, means the LCSF difference between CG(org.) and CG-to-MUS images.

No. of indices	CG	SKI	STR	CG	MUS	STR	Overall
	to-MUS			to-SKI			
256	0.40	0.12	0.64	0.63	0.95	1.05	**0.633**
512	0.43	0.11	0.62	0.67	0.95	1.09	**0.645**
1024	0.41	0.12	0.66	0.66	0.95	0.99	**0.632**
3072	0.40	0.33	0.85	0.75	1.12	1.16	**0.766**

Applying 3D Histogram Matching

One way of using the 3D histogram matching technique described above is to take a set of natural images and transform each of them so as to give it the histogram of the other images in the set. This will result in a matrix of images as shown in Figure 3, where all the images in a given row have the same image content (i.e. they show the same scene) and all the images in the same column have the same 3D colour histogram. In other words, all images in the same column are made up of the same pixels – the only difference between them is the spatial arrangement of the pixels. A set like this could, for example, be used for understanding what the relative importance of the 3D colour histogram is compared to image content and this could be seen by looking at the variation of some imaging application's performance for the images in the columns *versus* the rows of the matrix.

Figure 4 then shows the $\log P_{Jab}(w)$ characteristics of the images in column three of Figure 3. Comparing this figure with Figure 2 shows that the difference of spatial characteristics introduces to the images due to histogram transformation is similar in magnitude to the difference between the original image and the image that has been sharpened and had noise add to it in Figure 2.

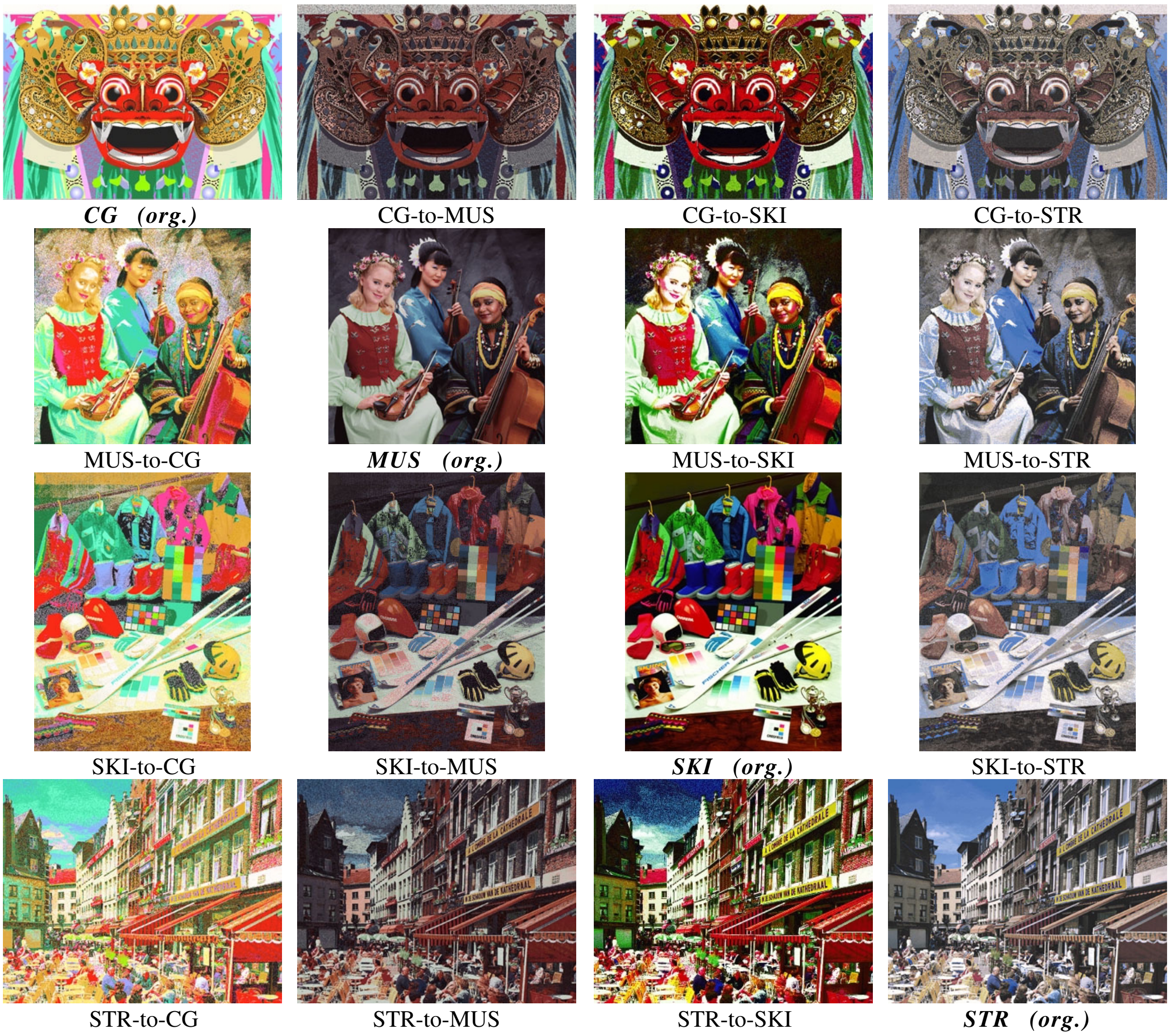

Figure 3: Image matrix showing 3D-histogram matching.

Conclusions

Overall this paper describes a method for transforming the 3D colour histogram of an image so as to exactly match any predetermined target state. The method is based around the EMD histogram difference metric and has a number of parameters that can be optimised so that it results in images with fewer artefacts. This paper also showed how to attempt such an optimisation and how to quantify changes to the spatial characteristics of the transformed images.

The ability to perturb the 3D colour histograms of images in a controlled and accurate way can then be used in any context where one needs to modify the colour histograms of images. For example, it can be used for generating sets of test images for studying the influence of various statistical image characteristics on different imaging applications. Furthermore, this method could also be used for enhancing the appearance of images (if one knew what statistical image characteristics are preferred) or for removing variation from sequences of images depicting the same scene.

Acknowledgements

The authors would like to thank Tadashi Yoshida and Todd Newman (Canon R&D Center Americas, Inc.) for supporting this study.

References

1. Y. Rubner and C. Tomasi, Texture Metrics, *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 4601–4607, (1998).
2. J. Morovic and P. L. Sun, Methods for Investigating the Influence of Image Characteristics on Gamut Mapping, *Proc. IS&T/SID 7th Color Imaging Conf.*, 138–143, (1999).

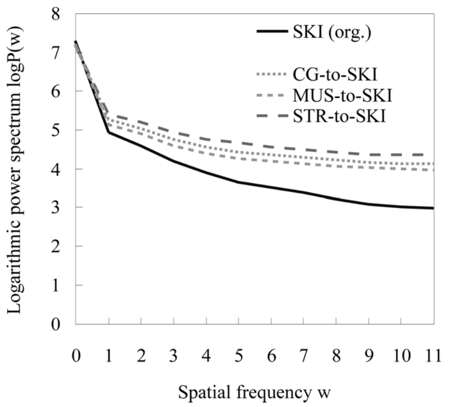

Figure 4. Effect of histogram matching on image spatial characteristics.

3. J. Morovic and P. L. Sun, The Influence of Image Gamuts on Cross–Media Colour Image Reproduction, *Proc. IS&T/SID 8th Color Imaging Conference*, 324–329, (2000).
4. J. Morovic, J. Shaw and P. L. Sun, A Fast, Non–Iterative and Exact Histogram Matching Algorithm, *PATREC*, **23**/1-3:127–135, (2001).
5. P. L. Sun and J. Morovic, The Influence of Image Histograms on Cross–Media Colour Image Reproduction, *PICS 2001*, Montrèal, Canada, 363-367, (2001).
6. C. J. Li, M. R. Luo and R. W. G. Hunt, The CAM97s2 Model, *Proc. IS&T/SID 7th Color Imaging Conference*, 262–263, (1999).
7. R. S. Berns, Methods for Characterising CRT displays, *Displays. Special Issue: 'To Achieve WYSIWYG Colour'*, 173–182, (1996).
8. R. Ulichney, *Digital Halftoning*, MIT Press, 241, (1993).
9. E. Gose, R. Johnsonbaugh and S. Jost, *Pattern Recognition and Image Analysis*, Prentice Hall Inc., 211, (1996).
10. M. G. A. Thomson and D. H. Foster, Role of Second- and Third-order Statistics in Discriminability of Natural Images, *J. Opt. Soc. Am. A*, **14/** 9:2081-2090, (1997).
11. N. Katoh and M. Ito, Gamut Mapping for Computer Generated Images (II), *Proc. 4th IS&T/SID Color Imaging Conference,* 126–129, (1996)

Biography

Dr. Ján Morovic received a PhD in Colour Science at the Colour & Imaging Institute (CII) where the topic of his research was "To Develop a Universal Gamut Mapping Algorithm." He now works at the CII as lecturer in Digital Colour Reproduction and is module leader for three modules on the MSc in Imaging Science. Further he also serves as chairman of the CIE's technical committee 8-03 on Gamut Mapping.