Evaluating the effect of noise on 3D LUT-based color transformations

Zhaohui Wang, Anna Aristova, Jon Yngve Hardeberg; Gjøvik University College; Gjøvik, Norway

Abstract

Image reproduction suffers from several limitations in a color management system. In this paper, we have investigated artifacts resulting from the inherent characteristics of the color transformations by interpolation in three-dimensional look-up tables, and the unavoidable measurement noise of the color measurement done during device profiling. In our experiments, images were manipulated using three interpolation methods, and five levels of random noise. Psychophysical experiments were conducted to evaluate the quality of the reproduced images. Finally the experimental data were collected to analyze the color transformations, and test the performance of two color image difference metrics in this context.

Introduction

There has been increased demand to reproduce images using heterogeneous devices and media such as digital cameras, displays and printing systems. The employment of different color imaging devices results in a common problem that each device produces color differently. For example, to use the same values based on the device primaries, such as RGB for a display, would reproduce different colors by different printing systems. Hence, users are lacking color predictability and consistency to reproduce color images across different imaging media. This has been a driving force for the industry to develop technology to achieve successful cross-media color reproduction. Numerous attempts for the development of color management systems have been made to satisfy different image reproduction tasks from one medium to another. The most widely used systems are those based on the International Color Consortium (ICC) specifications.

The ICC specification version 4 [1] provides definitions of color management architecture, profile format, and data structure. A typical ICC-based color management system consists of four basic components: profile, profile connection space (PCS), rendering intent, and color management module (CMM). A profile is a standard formatted file describing the device characterization, which defines the relationship between a device's control signals and the actual color that those signals produce. The ICC profile often employs multi-dimensional look-up tables (LUTs) to store the desired values. A process known as device characterization (or profiling) serves for this purpose, which provides a reliable way for color communications between media, and it is sufficient in simple applications with well-specified viewing conditions. The CMM is simply a color engine or processing engine, which is typically built into operating system, application or output device. The CMM performs all calculations needed to translate from the color space of one device to that of another. Although the ICC specifies the format of color profiles and to some extent the types of transforms that must be taken place to match colors between profiles, much of the process is left up to the imagination of the CMM creators. While it is difficult to specifically evaluate a vendor's commercial secrets, CMMs can certainly be evaluated based on the results they produce, both objectively by analysis of measurements and subjectively for pleasing contents. However, because different profiling applications will generate slightly different profiles from the same set of measurement data, the choice of CMM makes far less difference than the choice of profiling device and software [2].

One of the most accurate numerical models for device profiling is achieved by the measurement of a large number of colors, which can be used to develop multi-dimensional LUTs with interpolation for any intermediate colors. The accuracy depends on color measurement. Lack of accuracy can lead to quantization effects. However, in practice, one must balance the time cost and the measurement. Thus, the selection of the number of measurement is a very challenging task in the design of LUTs. Nowadays, interpolation is widely used to decrease the number of measurements. Several interpolation methods, such as trilinear, prism, tetrahedral, etc., have been developed. Since there is more than one methods of interpolation, each with some errors, a situation arises where two CMMs, given identical input, can yield different results.

Over the last few years, considerable progress have been made in instrument design and manufacture, which have led to more reliable instruments, stable readings and devices that are faster, lighter, and easier to use. Systematic errors, due to factors inherent in the manufacture of the instruments and the measuring situation, remain constant in time with respect to the selection and calibration of instruments and well-controlled measuring environments. However, the random errors, due to unpredictable variations during color measurement, are somehow instantaneous and unavoidable in the course of measurement, and can only be optimized by using the average of a number of repeated and consecutive measurements. The precision and uncertainty of the color measurement are mainly represented in the LUTs, and affect the color transformation in a color management system.

Consequently, the quality of the reproduced image will suffer from both inaccurate color measurement and LUT transformation. In this paper, we have investigated the influences from both factors and compared three different interpolation models, trilinear, tetrahedral and prism, which are widely used for LUT interpolation, and analyzed the effects of random measurement noise on the transformation using LUTs. The qualities of reproduced images were evaluated by psychophysical experiments using categorical judgment method. An image quality algorithm [3], which was newly proposed for predicting color image difference, was also used to evaluate the quality of LUTs transformation and compared with sCIELAB [4] for this purpose.

Interpolation, Noise Generation and LUT Transformation for Images

The 3D LUTs employed in ICC profiles are generally built for transformation between device-dependent and device-independent color spaces. In this paper, the 3D LUTs were built based on establishing a mapping from 3D RGB color space of the input device and the selected 3D CIELAB color space. This is done by sampling the RGB values from 0-255 in different intervals and then converting to CIELAB using a Dell LCD display characterization model (see section – psychophysical experiment). This display will also be used for the later psychophysical experiment. The interval values were set to 32 (256/8) and 16 (256/16) which were generally used to build a medium size of LUTs (9x9x9, and 17x17x17 respectively) used in profiling.

We investigated three interpolation methods in this work. Trilinear interpolation is the extension of linear interpolation, which can be considered to apply the linear interpolation seven times in three dimensions [5, 6]. It approximates the value of the mapping point in the CIELAB color space by using the linear relationship between an intermediate point within the local RGB sub-cube and the data on the lattice point.

Prism interpolation is similar to trilinear [7]. However, the necessary lattice points used to build the linear relationship for the intermediate point are reduced by cutting the cube diagonally into two halves.

Kasson et al. [8] investigated several interpolation models and found that tetrahedral interpolation is capable of providing similar accuracy as trilinear interpolation but with less computational effort, which, in turn, faster in implementation than trilinear interpolation [9]. This interpolation method can be implemented in two steps: divide a sub-space into tetrahedrons and then interpolate in each tetrahedron. There are several ways to construct a tetrahedron [6].

Different random noise by increasing the signal-to-noise ratios (five levels were considered in this work, 0, 1, 5, 10 and 20) of Gaussian random noise function, were applied to the L*, a*, and b* channels in 3D LUTs for simulating the random noise from the color measurement. The value of noise ratio 0 corresponds to no noise. The value of noise ratio 20 represents the largest noise applied to the LUTs' entries in this study. The range of noise ratio was decided by a pilot experiment which two observers participated to give a rough evaluation of the image quality. The quantities of noise applied for the L* channel are plotted in Fig. 1 as an example, in which the LUTs were based on the interval value of 32 and only one segmentation (full range are 9 partitions according to the LUTs generated by interval value of 32) were selected to show how the random noise were applied.

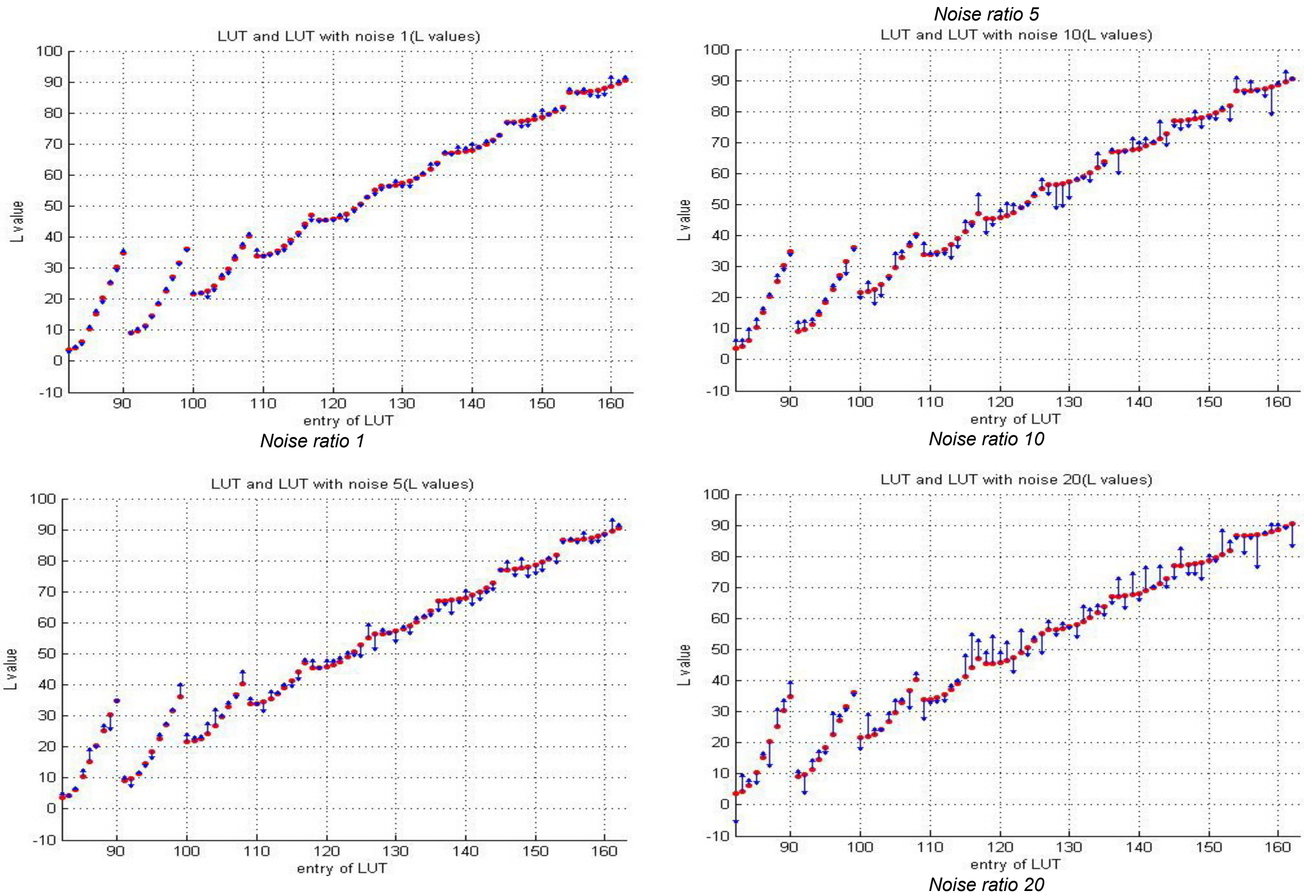

Figure 1 Samples of different noise ratio application on L channels from 1/8 segmentation of LUTs

Figure 2 experimental samples ([10],[11])

Thus the number of image transformations used in our experiments is the combination of two LUT sizes, three interpolation methods, and five random noise ratios (including noise ratio 0), a total of 30(=2x3x5) reproductions. We selected four images, as shown in Fig. 2, including one business graphic and 3 pictorial images. The original image states were set to sRGB color space in a resolution of 800x600 (96 pixels per inch) under D65 standard illuminant. The image reproduction using the LUT transformation is presented in Fig. 3.

A pre-process was applied to generate 3D LUTs of RGB-LAB by forward monitor characterization model in two sizes of 9x9x9 and 17x17x17. Then five simulating noise ratios were added to the 3D LUTs' entries. For any input image, the conversion from RGB space to CIELAB used the generated LUTs with different interpolation methods and then converted back to RGB space using inverse display characterization model. Totally, 120 testing images were prepared.

Psychophysical Experiment

Experiments were conducted in a dark room using a DELL 21-inch LCD display. The display was calibrated and characterized according to ISO3664 [12]. The GOG (Gamma-Offset-Gain) [13] characterization model was applied which gave a predictive error with a median ΔE of 0.43 for the forward characterization and a median ΔE of 0.97 for the inverse characterization in terms of CIELAB. The results provided a reliable translation from RGB to CIELAB, which is also used to build accurate LUTs of RGB-LAB transformation as mentioned before.

Four normal color vision observers joined in the experiment. Each observer was asked to evaluate the total image difference between an original image and a manipulated image using categorical judgment. Seven categories were used according to [3, 14]. Stages one and seven represent no difference and extreme difference separately and stage four stands for the average and acceptable difference.

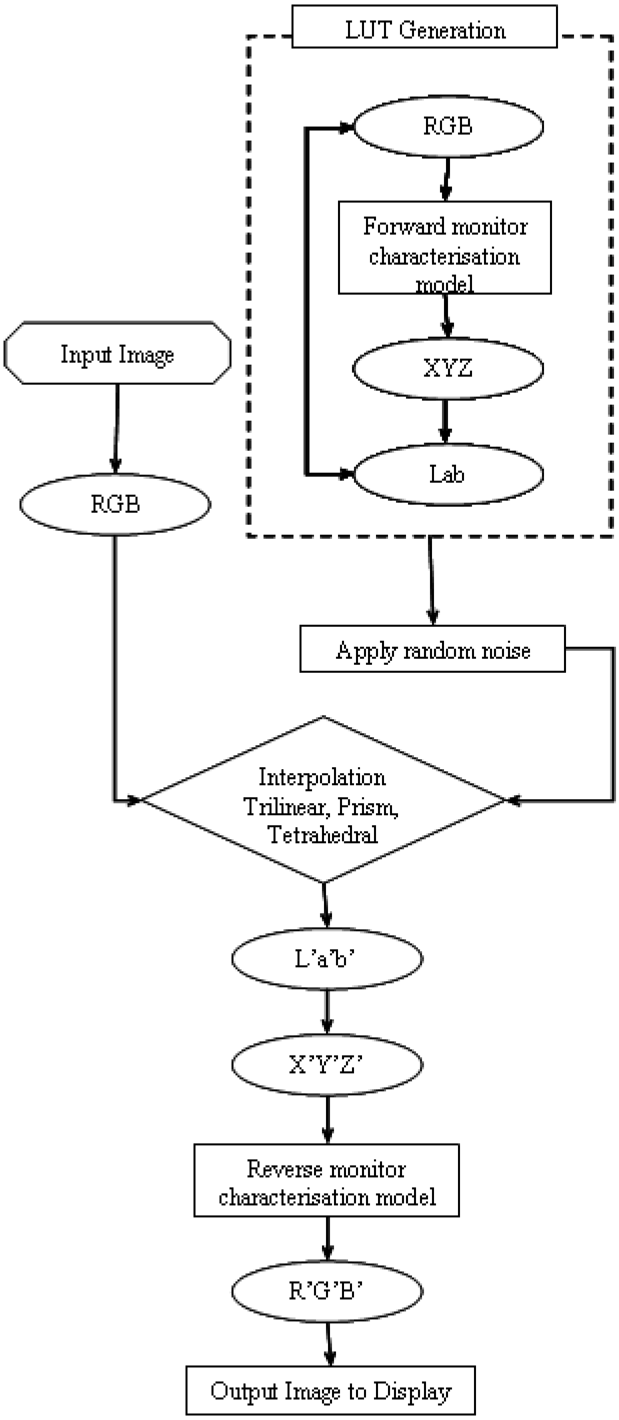

Figure 3 workflow of image reproduction by LUT transformation

Discussion and Results

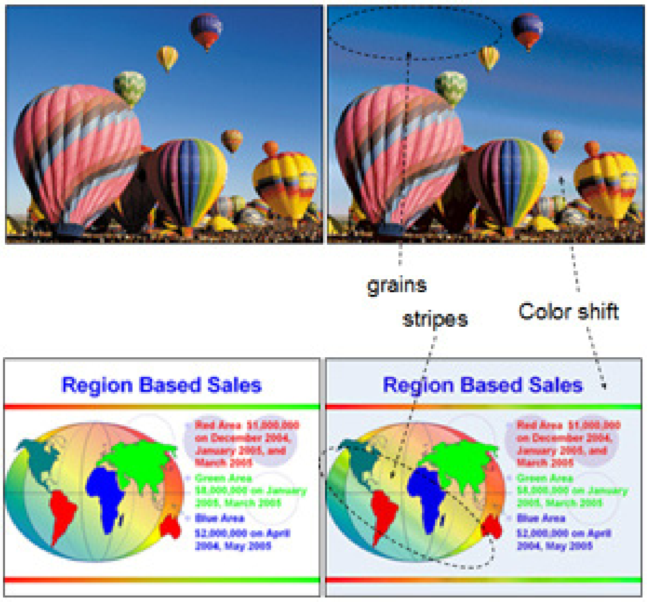

Figure 4 The artifacts examples compared between original images (Left) and reproduced samples (Right)

The artifacts resulting from interpolation and random noise could be found in several aspects: the stripes and grains in a large uniform area, the contours, and the color shift which is more visible for the white background with a large noise ratio. Examples are shown in Fig. 4.

The color shift is caused due to the high random noise ratio. Fig. 5 shows an example of how white is shifted according to different noise ratios. Note that the color shift should not be constant due to the random noise.

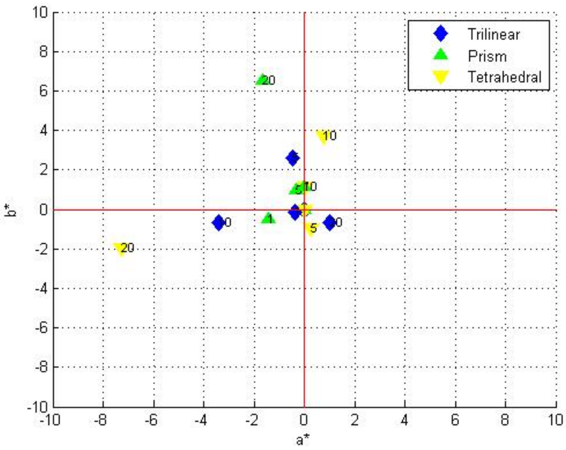

Figure 5 An example of white shift according to different random noise

The effects of two intervals (32 and 16), which decided the number of LUTs entries, were analyzed. In Fig.6 we have plotted the average visual judgments in terms of categories of all reproduced images. It can be found from Fig. 6 that the visual judgments of intervals value 16 (17x17x17 LUT) are slightly higher than that of 32 (9x9x9 LUT). However, the main influences are from noise which can be found by the separation of 5 noise ratio (0, 1, 5, 10, and 20).

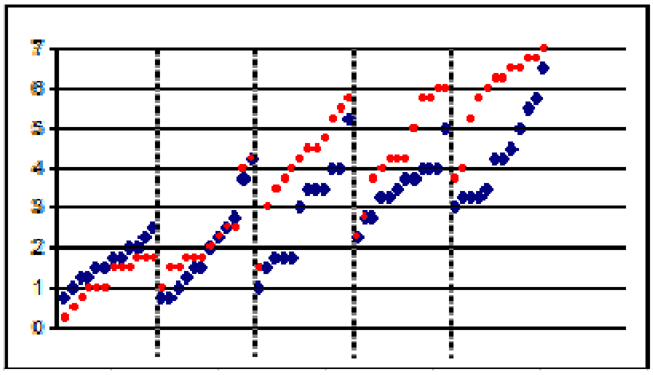

Figure 6 The difference between two LUT intervals (results of interval 16 in red and that of interval 32 in blue)

Three interpolation methods, prism, trilinear, and tetrahedral, have been described and discussed in details in [5, 6, 7, 8, 9 and 15]. In Fig. 7 we have plotted the average visual judgments in terms of categories of all reproduced images according to different interpolation methods. Results show similar performances between different interpolation methods.

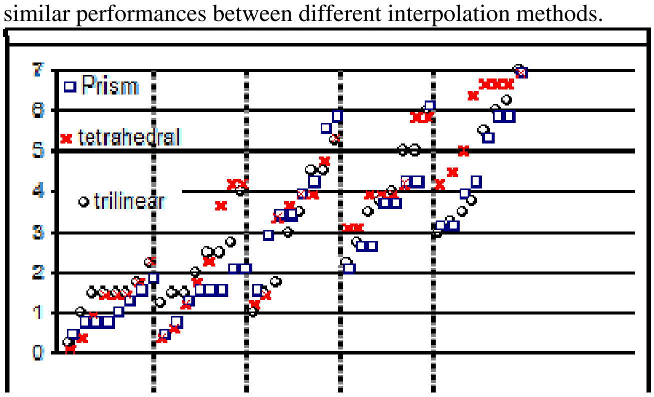

Figure 7 The difference between three interpolation models: prism, tetrahedral and trilinear

The random noise ratios were designed to simulate the different levels of random noise in color measurement procedure. In Fig. 8 we show the average visual judgments over all the images, against the noise ratios. It can be seen that the tolerance of acceptable difference (category 4) is achieved by noise ratio smaller than 5 roughly, which need smaller noise ratio interval to refine the result.

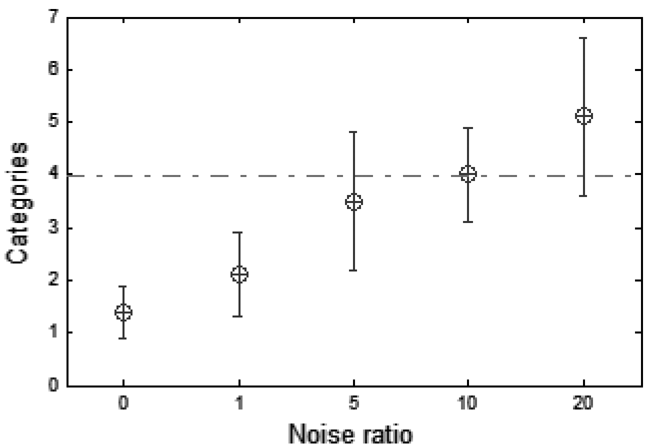

Figure 8 The relationship between random noise ratios and visual judgments

Many efforts have been spent to build image quality metrics to replace the time-consuming visual psychophysical judgments. In this work, two metrics, sCIELAB [4] and an adaptive bilateral filter [3] were investigated based on the experimental results using Pearson correlation.

We compared both metrics here because they have similar mechanisms. The algorithm, sCIELAB, was first introduced by Zhang et al. based on the analysis of spatial characterization of the human visual system and later recommended by the CIE TC 8-02 [16]. The sCIELAB workflow was carried out using the procedure provided by the CIE TC8-02 report.

The adaptive bilateral filter was based on the bilateral filter [17] which is adaptive to the viewing conditions and image itself. The difference between sCIELAB and adaptive bilateral filter is that the adaptive bilateral filter preserves the edges while filtering the image and sCIELAB not.

Given a color image *f(x)*, the bilateral filter [17] can be expressed as:

$$h(x) = k^{-1}(x)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi)c(\xi,x)s(f(\xi),f(x))d\xi,$$

where $k(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c(\xi,x)s(f(\xi),f(x))d\xi$,

and where the function $c(\xi,x)$ measures the geometric closeness between the neighborhood centre x and a nearby point ξ:

$$c(\xi,x) = e^{-\frac{(\xi-x)^2}{2\sigma_d^2}}$$

The function $s(\xi,x)$ measures the photometric similarity between the neighborhood centre x and a nearby point ξ:

$$s(\xi,x) = e^{-\frac{(f(\xi)-f(x))^2}{2\sigma_r^2}}$$

The domain spread σ_d is determined by the viewing distance, which is based on the conduction of contrast sensitivity functions [18].

$$\sigma_d = \frac{n/2}{180/\pi \cdot \tan^{-1}(l/(2m))}$$

where an image whose width is n pixels corresponds to l meters of physical length and is viewed from m meters away.

For a certain observer, the visual acuity is proportional to the viewing distance [19]. When the viewing distance is increased, the decrease in sensitivity at higher frequencies has been attributed to image blur. When the viewing distance is kept constant, smaller images displaying on a certain screen will be blurred more and larger images on the same screen will be blurred less.

The range spread is adapted to the image itself by:

$$\sigma_r = K/E,$$

where E is the image entropy defined by [20] and constant K is used to rescale the image entropy into an optimized value and entropy E is larger than zero for images.

Using the adaptive bilateral filter, the experimental image pairs were filtered in the CIELAB color space and then the average pixelwise differences were calculated using the CIELAB color difference formula.

Fig. 9 shows the performance of two metrics on each image. The Pearson's correlation values were calculated between visual judgments and the predicted values by each metric. The closer Pearson's correlation value is to +/- 1, the higher the performance. The average Pearson's correlation value is 0.66 for adaptive bilateral filter and 0.64 for sCIELAB. The error bars indicate 95% confidence interval which is calculated by $95\%CI = 1.96/\sqrt{2N} = 0.06$, where N represents the number of overall observations.

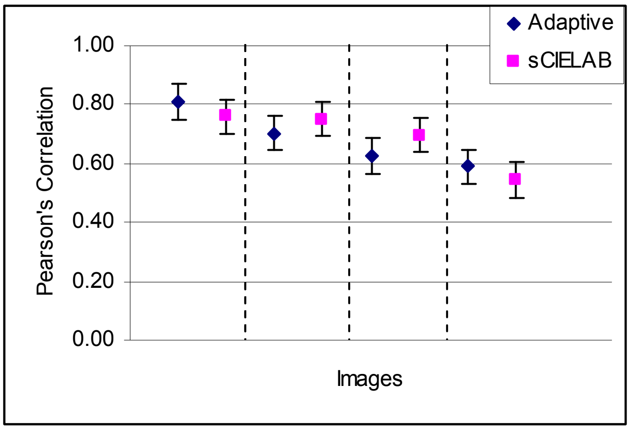

Figure 9 Comparison of the performance of sCIELAB and Adaptive bilateral filter based on Pearson's correlation (images from left to right: balloons, threads, picnic, sales)

Conclusion

In this paper, we have investigated factors inherited from 3D LUT based color transformation, including the number of entries (or size of the LUT), interpolation algorithms, and measurement noise which is always present in practical color measurement. These three factors have been analyzed based on visual psychophysical judgments. The results clearly show that random noise will put the main effects on the LUTs transformation which, in turn, affects the image quality reproduced by the color management system. Two image quality metrics, sCIELAB and an adaptive bilateral filter, were evaluated using the experimental results. Both metrics achieved high Pearson's correlation between the visual results and the metric's predictions.

References

[1] ICC, ICC File Formats for Color Profiles Version 4, ICC Specification, (2001).

[2] R. M. Adams II, "Choosing a Color Management Module", GATFWorld, 1999.

[3] Z. Wang and J. Y. Hardeberg, "An Adaptive Bilateral Filter for Predicting Color Image Difference," Proc. IS&T/SID, the 17th CIC, 27-31 (2009).

[4] X. Zhang and B. A. Wandell, "A Spatial Extension of CIELAB for digital Color-image Reproduction," Proc. SID Symposiums, pg. 731-734 (1996).

[5] I. Amidror, "Scattered data interpolation methods for electronic imaging systems: a survey, " J. Electronic Imaging, 11(2), 157-176 (2002).

[6] H. R. Kang, Color Technology for Electronic Imaging Devices, Bellingham: SPIE Optical Engineering Press, 64-101 (1997).

[7] Katsuhiro Kanamori et al., "A Color Transformation Algorithm Using Prism Interpolation", IS&T'S 8th International Congress on Advances in Non-Impact Printing Technologies, pp. 477-482, (1992).

[8] J. M. Kasson, S. I. Nin, W. Plouffe and J. L. Hafner, "Performing Colour Space Conversions with Three-dimensional Linear Interpolation," J. Electronic Imaging, 4(3), 226-249 (1995).

[9] R. Balasubramanian, "Reducing the Cost of Lookup Table based Color Transformations", Proc. IS&T/SID, the 7th CIC, 65-68 (1999).

[10] Z. Wang and R. Luo, "Experimental Filter for Estimating Images Differences", Proc. IS&T/SID, CGIV08, 112-115 (2008).
[11] IEC, "Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space – sRGB", IEC 61966-2-1 ed1.0, (1999)
[12] ISO, "ISO 3664:2000, Viewing conditions --- Graphic technology and photography," (2000).
[13] R. S. Berns, "Methods for characterizing CRT displays, " Displays, 16(4), 173-182 (1996).
[14] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," Psychological Review, 63, 81-97 (1956).
[15] P. Green and L. MacDonald, "Colour Engineering – Achieving Device Independent Colour", John Wiley & Sons, Ltd., 127-141, (2002).
[16] CIE, "CIE TC8-02 Technical Report: Methods for Deriving Colour Differences in Images - Draft Version 0.12," Central Bureau of the CIE, Vienna (2007).
[17] C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Images," in IEEE International Conference on Computer Vision, Bombay, India, pg. 839-846 (1998).
[18] J. A. Movshon and L. Kiorpes, "Analysis of the development of spatial sensitivity in monkey and human infants," J. Opt. Soc. Am. A, 5, pg. 2166- (1988).
[19] R. T. Hennessyt and W. Richards, "Contrast sensitivity and viewing distance", J. Opt. Soc. Am. A, 65(1), 97-98 (1975).
[20] C. E. Shannon, "A Mathematical Theory of Communication," Bell System Technical Journal, 27, pg. 379-423, 623-656 (1948).

Author Biography

Zhaohui Wang received his MSc in color imaging from the Color & Imaging Institute at the University of Derby (2004) and his PhD in color imaging from the Department of Color Science at the University of Leeds (2008). Currently, he is a postdoctoral researcher in the Norwegian Color Research Laboratory at Gjøvik University College. His work has focused on the development of a perceptual image difference metrics, and is sponsored by the Research Council of Norway