Evaluating the smoothness of color transformations

Anna Aristova, Zhaohui Wang, Jon Yngve Hardeberg
The Norwegian Color Research Laboratory
Gjøvik University College, Gjøvik, Norway

ABSTRACT

Multi-dimensional look up tables (LUTs) are widely employed for color transformations due to its high accuracy and general applicability. Using the LUT model generally involves the color measurement of a large number of samples. The precision and uncertainty of the color measurement will be mainly represented in the LUTs, and will affect the smoothness of the color transformation. This, in turn, strongly influences the quality of the reproduced color images. To achieve high quality color image reproduction, the color transformation is required to be relatively smooth. In this study, we have investigated the inherent characteristics of LUTs' transformation from color measurement and their effects on the quality of reproduced images. We propose an algorithm to evaluate the smoothness of 3D LUT based color transformations quantitatively, which is based on the analysis of 3D LUTs transformation from RGB to CIELAB and the second derivative of the differences between adjacent points in vertical and horizontal ramps of each LUT entry. The performance of the proposed algorithm was compared with a those proposed in two recent studies on smoothness, and a better performance is reached by the proposed method.

Keywords: smoothness, 3D LUT, color transformation, quality, ICC profile

1. INTRODUCTION

Cross-media color reproduction typically aims to accurately reproduce the original scene or image from an original medium to a reproduction medium according to a predetermined purpose. An original scene is described in terms of a specific medium which is often referred to medium dependent. For transforming colors from one medium to another, the medium dependent data are generally converted into medium independent descriptions. A process known as device characterization serves for this purpose, which provides a reliable way for color communications between media. The International Color Consortium (ICC) specifies a standard file format, an ICC profile, to describe the device characterization, which has been widely employed in industrial color managed workflows.

The method of multi-dimensional LUTs is often used for device characterization to store the desired values for the purpose of accurate color transformation. The accuracy is achieved by the measurement of a large number of color samples. Inaccuracies of color measurement can lead to artifacts in the transformation. However, accurate color measurement is a quite difficult task [1]. The precision and uncertainty of the color measurement are mainly represented in the LUTs, and affect the color transformation in a color management system.

Smoothness, in mathematics, is a property of functions which have derivatives of all orders. The smoothness is a desirable property of a color transformation [2], for which abrupt changes in monotonically variations are very detrimental [3]. When lacking smoothness, the color transformation will evoke several artifacts on the results of image reproduction. Olson [4] analyzed smoothness artifacts in image reproductions and their originations in ICC profile based color transformations. Several issues influencing smoothness artifacts include luminance errors, size of LUTs, influences of medium, etc. Fig. 1 shows a comparison of an original image (left) and its reproduction (right) by a LUT transformation using an ICC printer profile, in which discernable contours are visually present.

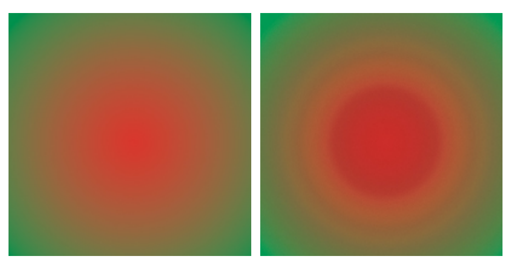


Figure 1. Visual artifacts generated by LUT color transformation

Estimating the smoothness of color transformation directly might be a challenge because of the nonlinear color mapping and noise introduced by measurement. Investigation [2] has been done on second derivatives of the color differences between each adjacent point in the output space of the LUT transformations. The derivatives are then described using statistics to quantify the degree of the smoothness. Kitoh et al. [5] proposed a method using a characterized scanner based instrument to quantitatively evaluating the smoothness of color transforms. This method is based on the generation of ideally smooth curves by averaging 20 smooth gradations using spline interpolation and the difference between the ideal and given gradation is used. An extension of [2] has been done by focusing the perceptual variations of abrupt changes at adjacent points [6], which compares the differences between the 1^{st} and the 2^{nd} derivatives, and weighted 5^{th} percentile of the 1^{st} derivative and 95^{th} percentile of the 2^{nd} derivative according to the perceptual estimation.

Through the literature, it can be seen that perceptual evaluations have been widely employed due to the visual artifacts brought by the unsmooth transformation. The assessment of image quality and difference between original and reproduced images has been in the discussion for a number of years [7]. Different algorithms have been proposed and tested for different applications. To apply an image difference metric to evaluate the smoothness of color transformation objectively is therefore a logical option.

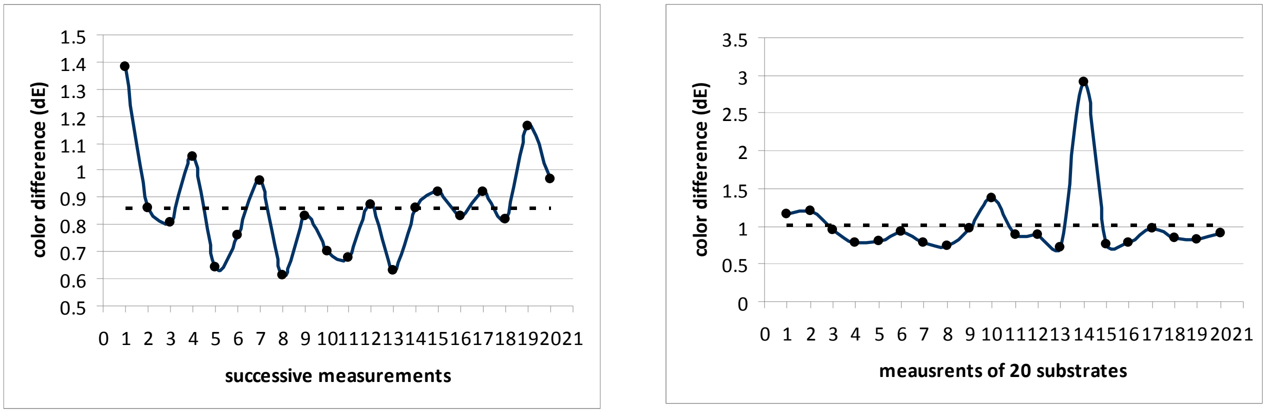
In this paper, we first in Section 2 introduce our proposed algorithm to evaluate the smoothness of 3D LUT color transforms based on the analysis of the 3D LUT entries in the output space. The vector of 2^{nd} derivatives of lightness differences between adjacent points on each vertical and horizontal lattice is calculated to quantify the smoothness. Psychophysical experiments were conducted (Section 3) to compare the performances of the proposed algorithm and the methods of [2] and [6]. A number of image difference evaluation matrices were investigated as an alternative the visual judgments involved in the psychophysical experiments. In Section 4 we present and discuss our experimental results, before concluding in Section 5.

2. PROPOSED METHOD

The LUT in the ICC profile is built upon a large number of color measurements. Consequently, the accuracy of the LUT transform depends on the measurement. However, in practice, one must balance the time cost and the measurement. Thus, the selection of the number of measurements is a very challenging task in the design of LUTs. Nowadays, interpolation is widely used to decrease the number of measurements. A previous study [8] has shown that the number of LUT entries has more effects than the method of interpolation on the LUT transformation.

A GretagMacbeth/X-Rite i1 Pro spectrophotometer was used in spectral reflectance scanning measurement mode to build the ICC profile using LUT structure. The i1 Pro employs diffuse illumination and normal viewing. An UV- cutoff filter at 400 nm was built in to provide high precision in printer profiling, which removes the UV part of spectrum of light reflected from papers and prevents it from disturbing measurements in the visible spectrum part. To generate printer profiles, we used a Xerox Phaser 7760 printer in RGB mode. TC9.18 RGB test chart was created using GretagMacbeth ProfileMaker5 based on the reference text associated with i1 Pro, which includes twenty-seven 9-step colour ramps covering the colour spectrum repeated three times used to be averaged in profiling software to produce a more accurate set of numbers, and seven 17-step narrow gamut and grayscale ramps for linearising. This chart was printed in the size of 27×21.5*cm* on normal office paper (each colour patch is 0.7×0.7*cm*). The printer resolution was set to 600×1200*dpi*. The software ProfileMaker5 was used through the measurement and profiles generation. Two profile sizes are provided by the software, which correspond to the size of LUTs. In this study, profiles were generated in LUTs size of 33×33×33 which is interpolated from the measurement data of the TC 9.18 RGB chart.

To investigate the variation of the LUTs transformation in profiles, a series of profiles were generated. A total of 20 profiles were generated by 20 successive profiling operations using an identical paper. Then another five profiles were generated in a period of five hours with even time interval. To further investigate the variability due to the different substrates, 20 profiles were generated by using 20 papers from the same batch. Totally, 45 profiles were prepared for evaluating the smoothness of LUTs transformation. Fig. 2 plots the average color difference in the 20 successive measurements (Fig. 2(a)) and 20 measurements with different substrates (Fig. 29b)). The color difference is the average of all patches in the chart and calculated using the function provided by ProfileMaker5. From Fig. 2, it might be concluded that the quality of different profiles is quite similar except the 14^{th} item.



(a) 2 average color differences of 20 successive profiles (b) average color differences of 20 papers' profiles

Figure 2. Investigation of two sets of profiles

Investigating the structure of the generated ICC profiles helps to understand how color is mapped by using the LUTs. Fig. 3 shows an example of a LUTs using the data extracted from one of the ICC profiles. The device dependent color space, RGB, is equally sampled and interpolated as shown in the left, and the corresponding measurements and interpolated, LAB values, construct the irregular output space of the LUTs.

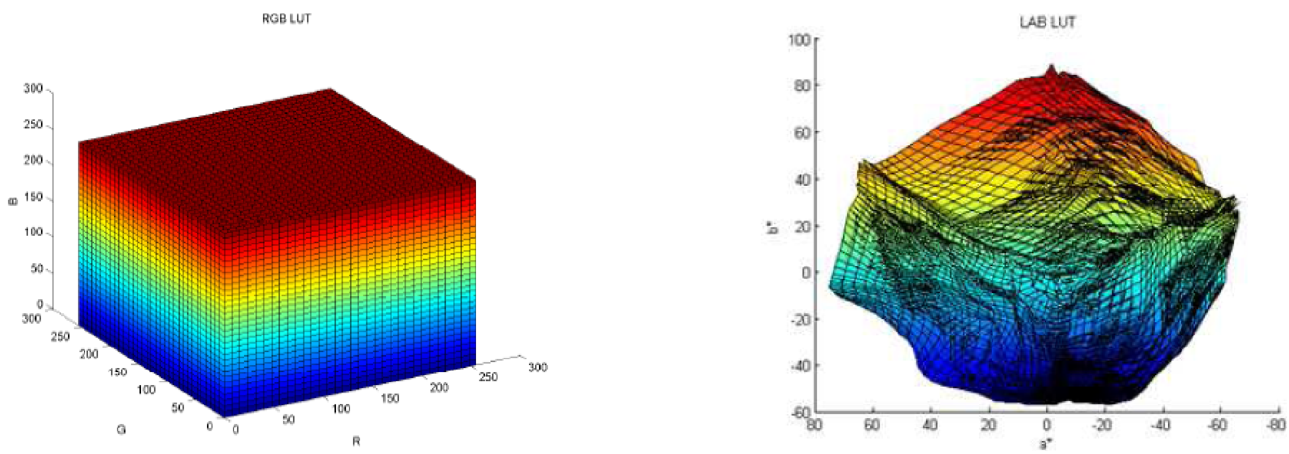
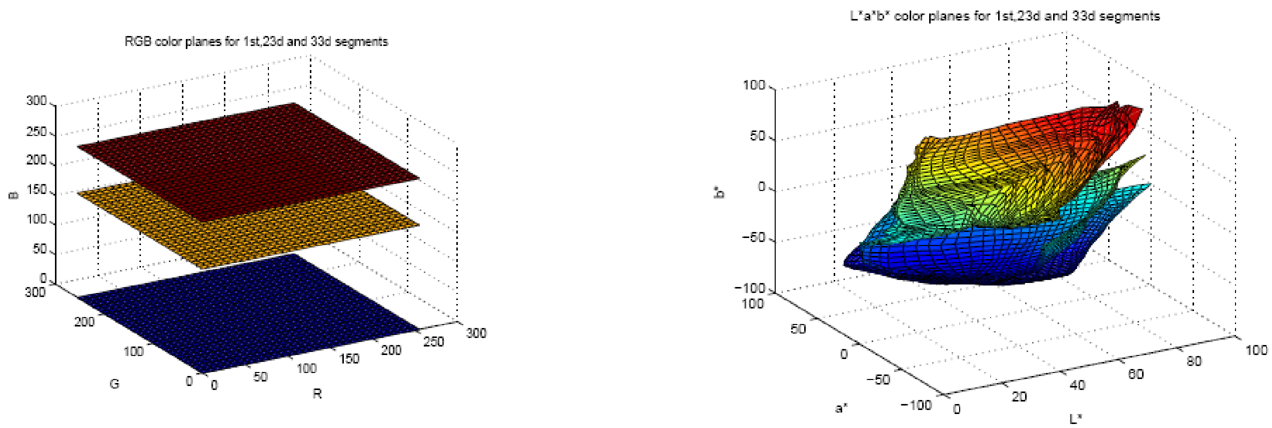


Figure 3. An example of LUT color transforms between RGB and CIELAB.

Considering the size of the LUTs, 33×33×33, we could separate each color space into 33 segments. For example, the segments shown in Fig. 4(a) represent the constant value of blue primary in 1st, 23rd, and 33rd segments, and the corresponding CIELAB segments are shown in Fig. 4(b).

For each entry in any segments, it is identified by the L*, a* and b* values. The differences of adjacent entries are calculated in terms of L*, a* and b* in horizontal and vertical directions as shown in Fig. 5, which result in vectors of ΔL*, Δa* and Δb*. The values of 2nd derivatives are computed for these three vectors to represent the variations. The maximum of 95 percentile of these three vectors of 2nd derivatives is used to quantify the smoothness of the certain LUT transformation. The maximum is calculated here as we assume that the variations (or gradients) are not only from the abrupt changes of Lightness but also from that of chromatic channels.



(a) (b)

Figure 4. An example of segments separated from the LUT

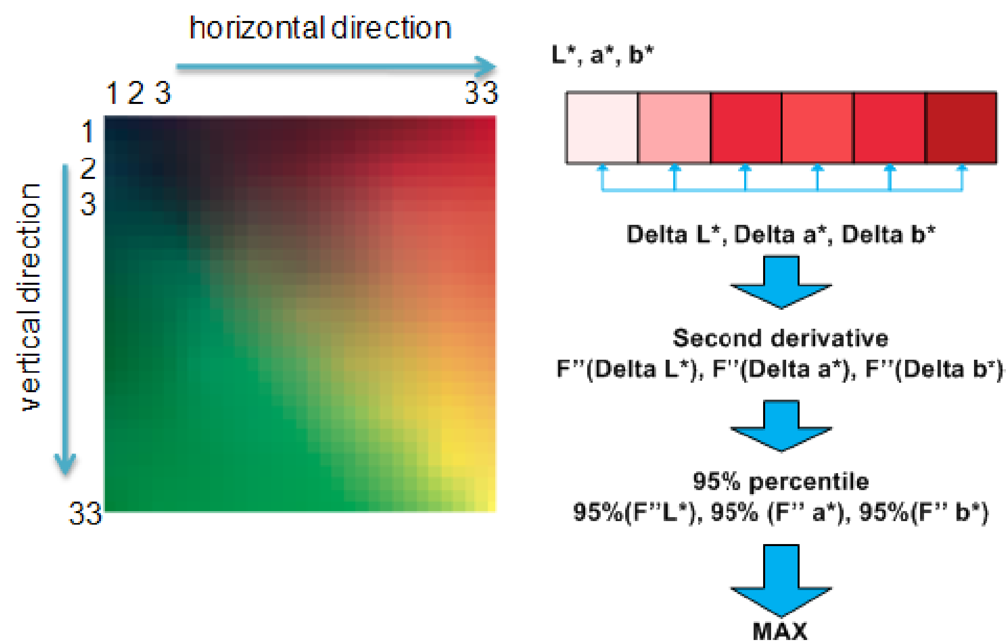
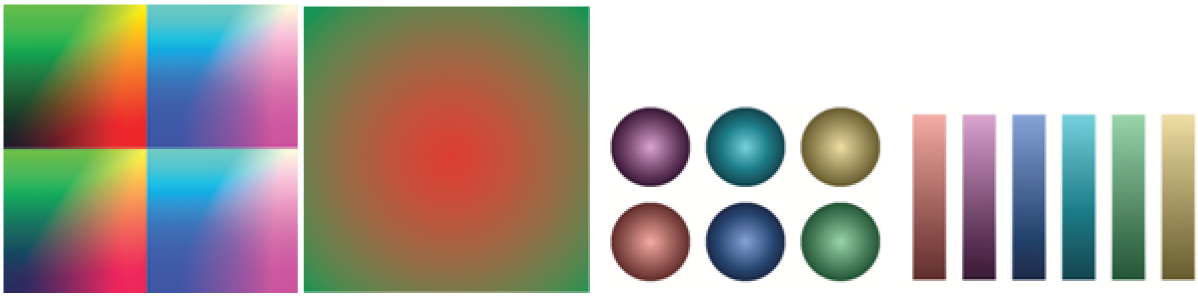


Figure 5. A computational workflow in each segment of LUTs

3. PSYCHOPHYSICAL EXPERIMENTS

Four samples, as shown in Fig. 5, were designed [9] to test performances of the proposed algorithm. These samples were processed with 45 printer profiles using Adobe Photoshop. The perceptual rendering intent was chosen to preserve the overall image appearance. A number of 180(=4×45) test images were thus reproduced in total.



(1) (2) (3) (4)

Figure 5. Testing samples

Perceptual experiments were conducted in a dark room using a Dell 21-inch LCD display. The display was calibrated and characterized according to ISO3664 [10], which gave a predictive error with a median ΔE_{ab} of 0.43 for the forward characterization and a median ΔE_{ab} of 0.97 for the inverse characterization. The original images were provided and compared with the reproduced images on screen to give observers clear conception how smooth an image could be.

Twenty normal color vision observers participated in the experiments to evaluate the smoothness level of the reproduced images using the category judgment method. The categories were defined from 1 to 5 corresponding to the smoothness level as shown in Table 1. Totally, 3600 (=20×180) visual judgments were collected. As the image is only processed by profiling, we considered the visual judgment of each image pair relates to the profile embedded. The visual judgments were converted to z-scores according to Torgerson's Law of Categorical Judgment. The z-scores were averaged for each profile.

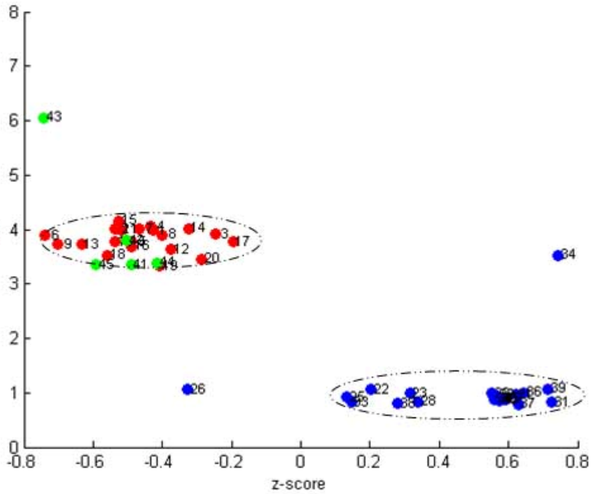
Table 1. Description of categories

Category	Match in Smoothness
1	Perfect match
2	Slightly different
3	Acceptable match
4	Moderate match
5	Worst match

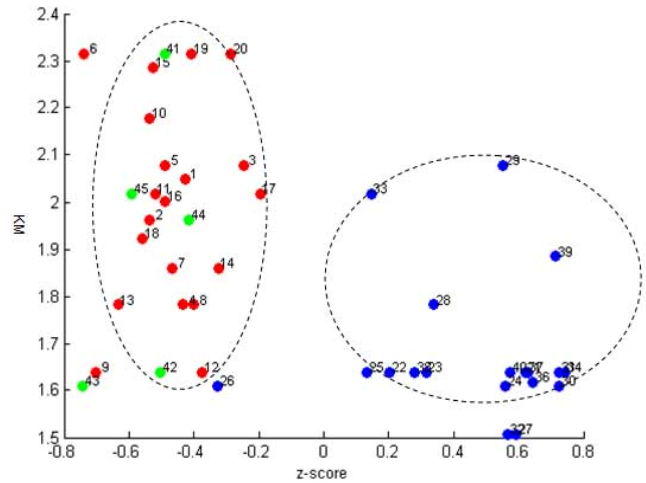
4. RESULTS AND DISCUSSION

The experimental results were used to compare the performances of three algorithms, including the proposed method (PM), algorithm of [2] (PG), and method of [6] (KM), in terms of Pearson's correlation. The Pearson's correlation value indicates the degree of linear relationship between two variables and ranges from -1 to 1. The Pearson's correlation values were calculated between the average z-scores of each profile and the results by each algorithm on each profile.

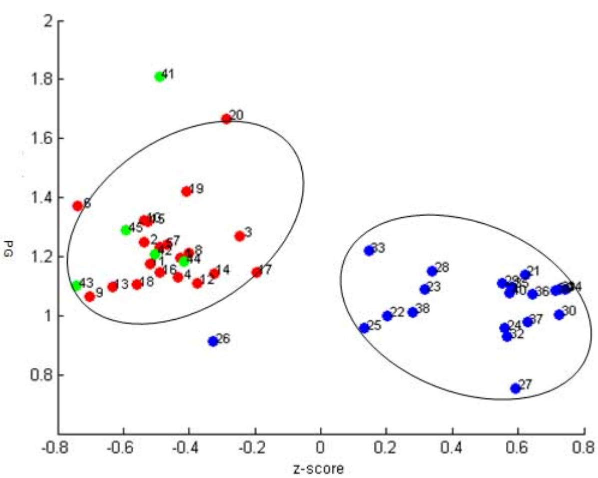
Fig. 6 plots the values of z-score against the performances of each algorithm based on the 45 profiles. As shown in the figure, we could find that the plots are categorized into two groups which are corresponding to the generation of profiles, the right category presents the profiles generated using different substrates and the left includes profiles generated from the same substrate. The proposed algorithm (PM) results in a higher Pearson's correlation of 0.62 comparing with the Pearson's correlation of the other two algorithms, 0.51 of PG and 0.47 of KM.



(a) the performance of algorithm of proposed method



(b) the performance of KM [6]

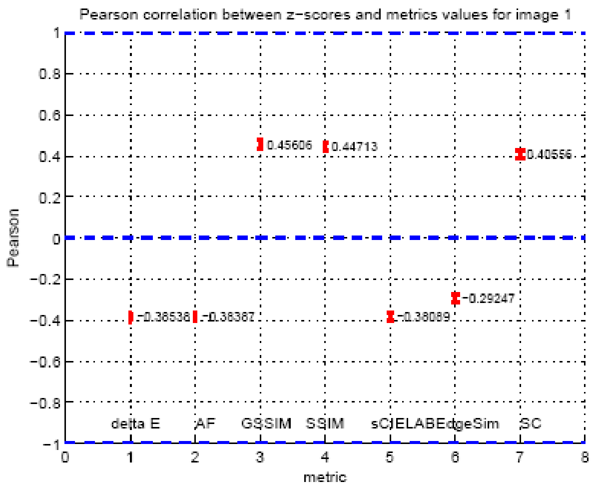


(c) the performance of PG [2]

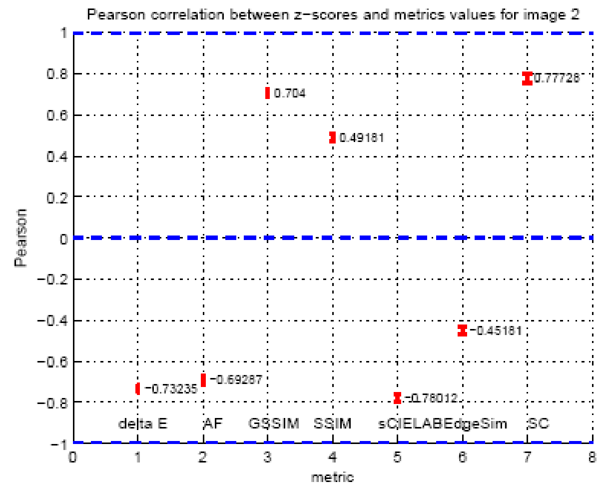
Figure 6. The comparison on the performances of three smoothness evaluation algorithms

To collect the visual judgments from a series of psychophysical experiments is a time-consuming work, this gives a strong motivation to investigate the use of a suitable quality metric to simulate the judgments. In our work, a number of image difference metrics were involved and compared to find the proper one. These metrics include, CIELAB [11] color difference formulae (deltaE), Adaptive Bilateral Filter [12] (ABF), Gradient Structural Similarity Index metric [13] (GSSIM), Structural Similarity Index Metric [14] (SSIM), spatial CIELAB [15] (sCIELAB). These metrics are widely used to evaluate the quality or difference of complex images. Another two mathematical functions are employed to calculate edges' similarity (EdgeSim) and the structural content (SC). The edge similarity is based on Minkowski distance of the edge information between image pairs. The structural content is based on the comparison of each pixel between image pairs.

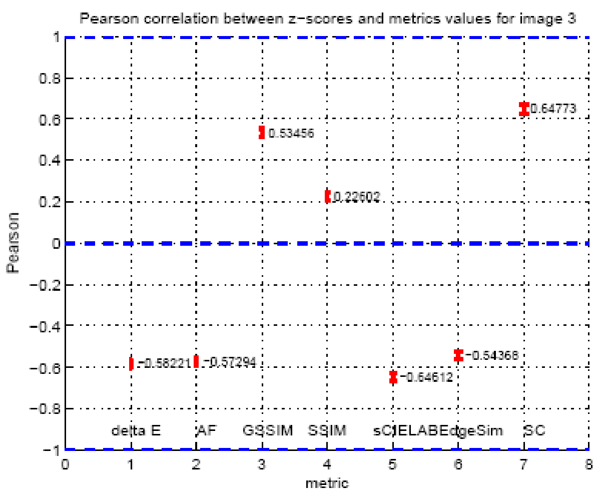
The performances of image difference metrics are compared in terms of Pearson's correlation as well. Fig. 7 plots the Pearson's correlation of each metric based on each testing image (as shown in Fig. 5). It could be found that methods of GSSIM, SSIM and SC result in positive Pearson's values for each image and the others turn out negative values, SSIM based methods and SC function are based on pixels values comparison. CIELAB, ABF, and sCIELAB compute the Euclidean difference and the EdgeSim can be considered as Euclidean distance mixing with Manhattan distance. T-test is implemented in these two different categories according to the Pearson's value to test whether the performances are similar or not. The differences between performances of SSIM based methods and SC function are not significant ($M_{difference}$=0.11, t(4)=1.15, $t_{critical}$=3.18, two-tail). Also, the differences between performances of color difference based methods (CIELAB, ABF and sCIELAB) and EdgeSim are not significant ($M_{difference}$=0.10, t(4)=1.69, $t_{critical}$=3.18, two-tail). According to the results, it is hardly to conclude which type of algorithm is particularly suitable in this case. However, SC function is the simplest method to implement. Experiment on large testing samples, including natural images, is necessary to have a concrete result.



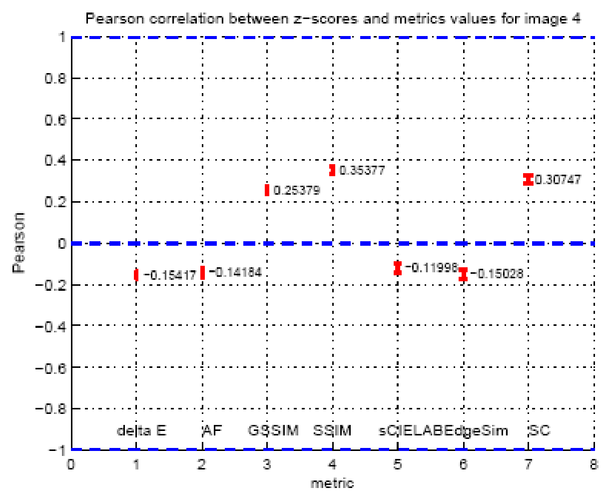
(a) Performances on image 1



(b) Performances on image 2



(c) Performances on image 3



(d) Performances on image 4

Figure 7. Performance comparison of different algorithms on each testing image

5. CONCLUSION

The ICC white paper [16] raises several issues to determine a profile's quality including the accuracy of colorimetric profiles, assessing the quality of non-colorimetric profiles, and the smoothness of color transformation. In this study, we addressed the smoothness measurement of 3D LUTs' color transformation in ICC profiles. An algorithm is proposed based on the analysis of the LUT structure in a profile. Perceptual experiments were conducted to test the performance. Testing images were purposely designed to reflect the smoothness of color transformation conveniently. The proposed method shows a better performance comparing with the two recent studies in terms of Pearson's correlation. An idea was motivated to use image difference metrics instead of perceptual experiments.

REFERENCES

[1] P. Nussbaum, A. Sole, and J. Y. Hardeberg, "Consequences of using a number of different color measurement instruments in a color managed printing workflow," TAGA Proceeding, (2009).

[2] P. Green, "A smoothness metric for colour transforms," in [Color Imaging XIII: Processing, Hardcopy, and Applications], SPIE Proceeding, 6807, 1-5 (2008).

[3] J. Morovič, A. Albarran, J. Arnabat, Y. Richard, & M. Maria, "Accuracy-preserving smoothing of color transformation LUTs." in [Sixteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications], 243-246, IS&T, Portland, Oregon, USA (2008).

[4] T. Olson, "Smooth Ramps: Walking the Straight and Narrow Path through Color Space." in [Seventh Color Imaging Conference: Color Science, Systems and Applications], 57-64, IS&T, Scottsdale, Arizona, USA (1999).

[5] S. Kitoh, and P. Hung, "A new method to quantitatively evaluate gradation smoothness of output media." in [Thirteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications], 89-94, IS&T, Scottsdale, Arizona, USA, (2005).

[6] Y. Kim, Y. Bang, and H. Choh, "Gradient approach to quantify the gradation smoothness for output media," J. of Electronic Imaging, 19(1), 011012-011012-9 (2010).

[7] M. Pedersen, and J. Y. Hardeberg, [Survey of full-reference image quality metrics] The Norwegian Color Research Laboratory, Gjøvik University College, (2009).

[8] Z. Wang, A. Aristova, and J. Y. Hardeberg, "Evaluating the effect of noise on 3D LUT-based color transformations". In [CGIV 2010 -- Fourth European Conference on Colour in Graphics, Imaging, and MCS/10 Vision 12th International Symposium on Multispectral Colour Science], 88-93, IS&T, Joensuu, Finland, (2010).

[9] E. Garcia, and M. Gupta, "Building Accurate and Smooth ICC Profiles by Lattice Regression." in [Thirteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications],101-106, IS&T, Albuquerque, New Mexico, USA (2009).

[10] ISO 3664:2000, [Viewing conditions --- Graphic technology and photography], ISO, (2000).

[11] CIE 15:2004, [Colorimetry] Central Bureau of the CIE, Vienna (2004).

[12] Z. Wang, and J. Y. Hardeberg, "An Adaptive Bilateral Filter for Predicting Color Image Difference." In [Seventeenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications], 27-31, IS&T, Albuquerque, New Mexico, USA (2009).

[13] G. H. Chen, C. L. Yang, and S. L. Xie, "Gradient-based structural similarity for image quality assessment." In [Image Processing, 2006 IEEE International Conference on], IEEE, 2929–2932 (2006).

[14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity." IEEE Transactions on Image Processing, 13(4):600–612 (2004)..

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

[16] ICC, [The role of ICC profiles in a colour reproduction system], (2004).