Quality assessment for visible and infrared color fusion images of typical scenes

Shaoshu Gao (高绍姝), Weiqi Jin (金伟其)*, and Lingxue Wang (王岭雪)

MoE Key Laboratory of Photoelectronic Imaging Technology and System, School of Optoelectronics, Beijing Institute of Technology, Beijing 100081, China

**Corresponding author: jinwq@bit.edu.cn*

Received February 7, 2012; accepted March 14, 2012; posted online June 15, 2012

Two comprehensive evaluation metrics, image perceptual quality based on target detectability (PQTD) and perceptual quality based on scene understanding (PQSU), are proposed to measure image quality for visible and infrared color fusion images of typical scenes. A psychophysical experiment is performed to explore the relationship between conventional quality attributes and the proposed evaluation metrics. The prediction models for PQTD and PQSU are derived by multiple linear regression statistical analyses. Results show that the variation of PQTD can be predicted by sharpness and perceptual contrast between the target and background, and that color harmony and sharpness can predict PQSU. The proposed evaluation metrics and their prediction models provide a foundation for further developing objective quality evaluation of color fusion images.

OCIS codes: 110.3000, 330.5020, 330.1720, 350.2660.

doi: 10.3788/COL201210.081101.

Visible and infrared color image fusion combines two source images into a single composite false-color image that is suitable for special visual tasks. Many fusion algorithms and systems have been successfully applied. At present, however, no generally accepted methods of color fusion image quality evaluation exist. Image quality assessment has been investigated using subjective and objective approaches. Subjective approaches evaluate the image quality based on subjective perception by observers, whereas the goal of objective image quality assessment research is to design computational models that can predict the perceived image quality accurately and automatically. The numerical measure of quality an algorithm provides should correlate well with human subjectivity[1].

Many factors are able to influence perceptual image quality. For true-color images, Choi *et al*. performed a psychophysical experiment in which colorfulness, contrast, and naturalness were the key attributes controlling image quality[2]. Pedersen *et al*. found that color, lightness, sharpness, contrast, physical attributes, and artifacts were the most meaningful attributes for the print image quality[3]. However, unlike true-color images, color-fused images contain dual-band information. Thus, the purpose of image fusion is not to obtain the color image that completely corresponds with the true-color image of the same scene, but to improve the suitability for special vision tasks (i.e., target detection and scene understanding)[4]. Until recently, most subjective evaluations have been investigated based on different visual tasks, such as target detection and recognition, scene recognition, and situational awareness[5−7]. Shi *et al*. presented three influence factors (i.e., target detection, detail, and colorfulness) to evaluate the color fusion image quality[8]. Thus far, however, no agreement on which quality metric should be used to evaluate the quality of visible and infrared color fusion images has been reached. For the same image, perceptual quality assessment results differ according to different visual tasks[9]. Therefore, to evaluate the color fusion image quality comprehensively, we propose two evaluation metrics: image perceptual quality based on target detectability (PQTD) and perceptual quality based on scene understanding (PQSU). A psychophysical experiment was performed to explore the relationship between common quality attributes (QAs) and the proposed evaluation metrics. Multiple regression statistical analyses were conducted to establish prediction models of QAs for PQSU and PQTD.

Both objective and subjective evaluations of image quality are dependent on a number of QAs related to perception, such as sharpness, contrast, and naturalness[10]. These QAs influence the overall image quality in different ways, and studying their relationships would be helpful to optimize the image quality model. The QAs selected in the psychophysical experiment must be convenient for modeling and quantization in order to further evaluate the image quality objectively. According to specific applications of visible and infrared color fusion images and the characteristics of fusion images, such as false color and low resolution, four QAs were used in our subjective assessment experiment. These four QAs have been proposed as evaluation metrics in previous studies.

The perception contrast of the target and the background (PCTB) in a color image refers to the visibility of color variation between the target and background regions. Target detection for human eyes depends on perception contrast of the target and background to a great extent. The target can be detected easily by enhancing the color contrast between the target and background[11]. The target detectability of color fusion images can be predicted by measuring the contrast[12,13]. Many factors influence perception contrast at different levels, such as human visual system, perceptual luminance variation, chroma difference, and hue difference between target and background. In this experiment, the target refers to the hot target in the infrared source image. If a image contained more than one target, observers were asked to measure PCTB for all the targets comprehensively.

Sharpness (S) is frequently used to describe the image quality[3]. It consists of two concepts: resolution and acutance. Resolution involves resolving detail, while acutance involves transition of the edges. The perception of sharpness is related to the clarity of detail and edge definition of an image[14]. In literature, sharpness is linked with color[2], noise[15], contrast[16], and so on.

Color harmony (CH) has a high correlation with image color preference evaluation for true-color images[17]. A generally accepted understanding of color harmony is given by Burchett: "Colors seen together to produce pleasing affective response are said to be in harmony"[18]. Color harmony can be influenced by many factors, such as shape, area, and type of combination[18,19]. Pseudo-colorization makes the fused image appear to be more or less different from the real scene. In order to avoid visual fatigue and negative psychological effects, color harmony should be considered as a QA of great importance.

Color naturalness (CN) is defined as the degree of correspondence between the image colors and memory colors of real-life scenes. It is interpreted as the subjective impression of the fidelity of color rendering[20]. The strong link between image quality and naturalness found in experiments suggests that naturalness is an important quality attribute in the color reproduction of natural scene images[21]. People regard image quality as better if the image colors are close to their long-term memory colors[22].

In order to improve the target detectability of color fusion images, color contrast between the target and background is often enhanced by making the target color bright (intense red, for example). Therefore, color harmony and naturalness in this letter refer to color harmony and naturalness of background in the fusion image without consideration of the target color.

Clearly, the above four QAs describe image quality from different aspects. The visible and infrared color fusion image should be suited for special-purpose applications applied to improve target detection and scene understanding. Previous experimental results showed that, compared with individual image modalities, the appropriate color fusion images could contribute more to target detection and scene understanding[5−7]. Therefore, this letter proposes a comprehensive evaluation metric, perceptual quality based on different visual tasks, which contains two aspects: PQTD and PQSU.

a) PQTD. The experimental task required observers to measure the detectability of hot targets in fusion images. The suitability of the fusion image for fast and accurate target detection is considered the most important factor to measure the perceptual quality of the color fusion image. If a image contained more than one target, observers were asked to measure PQTD for all the targets comprehensively.

b) PQSU. The experimental task required observers to evaluate the perceptual quality based on scene understanding according to sharpness and color, without considering target detectability. The suitability of the fusion image for direct and accurate understanding of the image scene is considered the most important factor to measure the perceptual quality of the color fusion image.

To produce the test images, IR and visible images were fused using the following eight methods: TNO fusion scheme[23], MIT fusion scheme[24], BIT fusion scheme (in YUV color space)[25], linear fusion scheme[26], steerable pyramid based color fusion scheme[27], and three color-transfer methods after the linear fusion (i.e., linear color transfer in YUV space[28], and multi-resolution based color transfer in YUV and RGB color space)[29]. The three color-transfer methods were used to select three different target images to obtain rich colors in the test fusion images.

Improvement of scene perception for observers is one of basic applications of color fusion images. The fusion images were classified by typical scene for perceptual quality evaluation to select suitable fusion schemes according to different types of image scenes. In this study, the test images were classified into three classes: Plants, Sea and Sky, and Towns and Buildings. For each of the three typical scenes, seven sets of images were selected in the experiment. Each set contained the resulting images of five fusion methods and three color-transfer methods. Total 168 test images were contained in this study. In the 21 pairs of visible and infrared source images, 7 pairs were provided by Morris *et al*[30]. The rest of the source images were obtained by our own fusion system, which is comprised of an 8–14-μm wavelength band infrared detector and a visible CCD. These images contained obvious hot targets and specific background that have characteristics of the typical scene.

The experiments were conducted in a dark room, using a characterized cathode ray tube (CRT) display [with resolution of $1\,024\times768$ (pixel)] for presenting the test fusion images. The CRT, with a display peak white luminance of 100 cd/m^2, provided the only light source in the room. The x, y chromaticity coordinates of the display's white point were (0.314, 0.329), which are quite close to that of the D65 (0.313, 0.329). To ensure precise color reproduction on the CRT display, the GOG model[31] was implemented for conversion between the tristimulus values and monitor RGB values. The GOG model was tested for the display by means of CIELAB color difference using 27 test colors (0, 128, and 255 digital counts for each channel of red, green, and blue) with average color difference values of 1.8 for the CRT display. Two minutes of adaptation time were given to each observer. The viewing distance was 50 cm. The image size was set to 320×240 pixels.

At least 15 observers should be used according to recommended ITU-R of the subjective assessment methodology[32]. Seventeen Chinese observers, comprising nine males and eight females with normal vision and who have passed Ishihara's Tests for Color Deficiency, participated in the experiment. Considering the characteristics of night-vision-system users, the average age of observers was 29 years (raging from 20 to 46). The observers, who included three military personnel, had different professional backgrounds but all have some knowledge or experience of color night vision.

Before the experiment, instructions were given to the observers, including the purpose and process of the psychophysical experiment, detailed description and definitions of all evaluation metrics, and the target and scene content of each set of test images. Thus, observers could rate the test images without any bias.

In order to avoid the observer fatigue and interaction between these QAs, the experiment was divided into six sessions. In each session, observers assessed one of the six aspects (PCTB, S, CH, CN, PQTD, and PQSU). For each set of test images, a pair of visible and infrared source images was presented first to help observers to identify hot targets and understand scene contents, after which the eight corresponding fusion images were simultaneously displayed in the middle of a gray background (L^*=50) with random arrangement. To scale observer perceptions of the six aspects, a categorical judgment technique was adopted. Observers were asked to rate each image using a seven-point verbally labeled category scale. For example: "7" corresponded to "highest quality imaginable," "4" to "average quality," and "1" to "lowest quality imaginable."

The different observers rated these six attributes for fusion images of different scenes with different scale ranges. To ensure that all the experimental data fell within a specified rang, the scores of one set with eight test images were handled as a unit for each attribute. They were transformed into [0, 1] by normalization as

$$z_{ij} = \frac{y_{ij} - y_i^{\min}}{y_i^{\max} - y_i^{\min}}, \quad (1)$$

where y_{ij} (i=1, 2, $\cdots$, 21, j=1, 2, $\cdots$, 8) represents the experiment data; $y_i^{\min}$ and $y_i^{\max}$ are the lowest and highest value in the unit i, respectively; and z_{ij} is the normalized score.

We collected and averaged the normalized scores of the six attributes given by the observers and analyzed their correlation. Table 1 shows the Pearson correlation coefficient (r)[33] between each pair of attributes and the result reaches significance level (P=0.000<0.05). The result indicates a very high significance of correlation between the scores of PQTD and PCTB, as well as a high correlation between PQTD and S. The perceptual quality based on scene understanding correlated strongly to S, CH, and CN. It is proved a certain level of validity to explain image perceptual quality by the four attributes. This study adopted the analytic method of multiple linear regression to explore the influence of the four QAs on comprehensive image quality based on vision tasks, PQTD, and PQSU.

Stepwise regression[34] was adopted to obtain the best combination of the least variables for predicting image perceptual quality. In stepwise regression, at each step, the best variable is added to the model if its corresponding F-test is significant (P$\leqslant$0.05). Before the next variable is added, however, the stepwise method takes an additional look-back step to check all variables included in the current model, and deletes any variable that has a P-value greater than or equal to the 0.10 significance criterion. Only after the necessary deletions are accomplished can the procedure move to the next step of adding another variable into the model. The stepwise search continues until every variable in the model is significant and every variable not in the model is insignificant.

For explaining PQTD in all image categories, the basic stepwise regression procedure is as follows. Firstly, select the PCTB most correlated with PQTD and obtain the linear regression model 1 (Table 2). Check this variable by F-test and determine whether it is significant (P=0.000<0.05). Then, enter the independent variable S, which is the best remaining variable. Check all variables included in the current model and determine whether each reaches significant level. The other attributes with no statistical significance are excluded, and the regression model 2 (see Table 2) is obtained. The t statistics can help determine the relative importance of each variable in the model. PCTB was determined to be the most important variable, followed by S. The variance inflation factor (VIF) of each included variable was smaller than 10, implying the absence of any significant multicollinearity problems. The regression equation can be written as

$$\mathrm{PQTD} = 0.714\mathrm{PCTB} + 0.314\mathrm{S} - 0.025 (R^2 = 0.921), \quad (2)$$

where R^2 is coefficient of determination that indicates how much of the dependent variable, PQTD, can be explained by the independent variables, PCTB and S. In this case, 92.1% can be explained. The F-test was highly significant (F=1080.487, P=0.000<0.05), indicating that the model, as a whole, is significantly good in predicting the variable PQTD.

In the same way, for explaining PQSU in all image categories, CH was the first attribute selected, S was the second, and CN was the final attribute selected in the created regression models (see Table 3). In Model 3, because the VIF values of CH and CN were 18.845 and 14.871,

Table 1. Pearson Correlation Coefficients between Each Pair of Attributes

	PQTD	PQSU	PCTB	S	CH	CN
PQTD	1					
PQSU	0.550	1				
PQTB	0.900	0.256	1			
S	0.670	0.938	0.403	1		
CH	0.479	0.973	0.184	0.895	1	
CN	0.480	0.957	0.196	0.865	0.966	1

Table 2. Regression Models for PQSU in All Image Categories

Model		B	t	Sig. (P)	VIF	F	Sig. (P)	R^2
1	PCTB	0.855	28.002	0.000	1.000	784.132	0.000	0.807
2	PCTB	0.714	33.265	0.000	1.194	1 080.487	0.000	0.921
	S	0.314	16.307	0.000	1.194			

Table 3. Regression Models for PQSU in All Image Categories

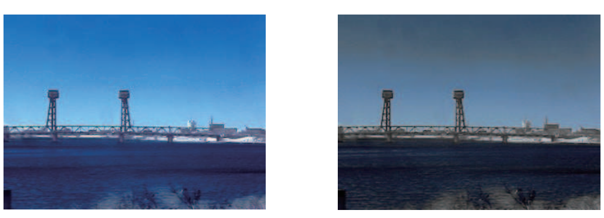
Model		B	t	Sig. (P)	VIF	F	Sig. (P)	R^2
1	CH	0.980	57.670	0.000	1.000	3 325.826	0.000	0.947
2	CH	0.675	23.345	0.000	5.032	2 692.790	0.000	0.970
	S	0.341	11.804	0.000	5.032			
3	CH	0.426	8.214	0.000	18.845	2 308.502	0.000	0.974
	S	0.341	12.711	0.000	5.032			
	CN	0.256	5.604	0.000	14.871			

respectively, which are greater than 10, multicollinearity suggesting that a high degree of multicollinearity is present. Furthermore, the collinearity analysis of Model 3 can be measured by a condition index (CI) greater than 15 (CI=21.9); therefore, multicollinearity is a concern. A very strong correlation existed between CH and CN (r=0.966). When observers evaluated color harmony of the false-color fusion images, they considered, not only the harmony of color combinations in the image, but also the harmony between the color and the corresponding scene content. Inevitably, color harmony of fusion images was influenced by image content. By analyzing the scores of CH and CN, the correspondence between the image colors and memory colors of real-life scenes was found to be one of important influential factors of CH. However, the meaning of color harmony was more abundant, which includes other factors besides color naturalness, such as visual comfort[35]. Figure 1 presents two fusion images of approximately equal mean scores for the CN grading (G_{CN}) but greatly different mean scores for the CH grading (G_{CH}). Figure 1(b) appears similar to the scene of the previous evening and obtained high score in CN grading; however, it is not as harmonious as Fig. 1(a). Figure 1(a) not only appears similar to the real-life scene, but also produced pleasing affective response; therefore, its CH grading score was very high. Compared with CN and S, CH had the strongest correlation with PQSU, suggesting that CH is the most important attribute for explaining PQSU. In this letter, Model 2 in Table 3 was employed to predict PQSU, which can avoid the multicollinearity problem and simplify the regression model. The t statistics indicated that CH was the most important influencing factor, followed by S, and that both reached significant level (P=0.000<0.05). The VIF of each included a variable is smaller than 10, implying that there was no significant multicollinearity problem. The regression equation can be written as

$$\mathrm{PQSU} = 0.675\mathrm{CH} + 0.341\mathrm{S} - 0.014(R^2 = 0.970), \quad (3)$$

where $R^2 = 0.970$ indicates that 97.0% of the dependent variable, PQSU, can be explained by the independent variables, CH and S, which are very large. The F-test was highly significant (F=2692.790, P=0.00<0.05), which indicates that the model, as a whole, is statistically significant.

Because images of different categories have different characteristics, each attribute plays a different role in different image categories. In order to obtain more suitable prediction models, the same statistical analysis was conducted to each image category. The prediction models used for PQTD and PQSU of the three typical scenes were



(a) G_{CH}=6.4 G_{CN}=6.1 (b) G_{CH}=4.8 G_{CN}=6.0

Fig. 1. Two fusion images of approximately equal mean scores of G_{CN} but greatly different mean scores of G_{CH}.

Plants

$$\mathrm{PQTD} = 0.768\mathrm{PCTB} + 0.246\mathrm{S} - 0.030(R^2 = 0.935), \quad (4)$$

$$\mathrm{PQSU} = 0.532\mathrm{CH} + 0.463\mathrm{S} - 0.012(R^2 = 0.944). \quad (5)$$

Sea and Sky

$$\mathrm{PQTD} = 0.531\mathrm{PCTB} + 0.447\mathrm{S} - 0.010(R^2 = 0.928), \quad (6)$$

$$\mathrm{PQSU} = 0.823\mathrm{CH} + 0.201\mathrm{S} - 0.015(R^2 = 0.982). \quad (7)$$

Towns and Buildings

$$\mathrm{PQTD} = 0.815\mathrm{PCTB} + 0.224\mathrm{S} - 0.024(R^2 = 0.952), \quad (8)$$

$$\mathrm{PQSU} = 0.679\mathrm{CH} + 0.347\mathrm{S} - 0.015(R^2 = 0.988). \quad (9)$$

Each of independent variables reaches significant level (P=0.000<0.05). No significant multicollinearity problem exists in the above models. Each model, as a whole, is statistically significant, and their prediction powers are all above 92%. Thus, PQTD and PQSU can be predicted by these independent variables very well.

For fusion images of three typical scenes, PCTB was the most important influencing attribute, followed by S in the prediction models for PQTD. For the category of Sea and Sky, S represented considerable higher anticipation in the PQTD model compared with the other scene categories because the fusion images of Sea and Sky contained large uniform background areas; thus, observers could detect the edges of targets easily. Furthermore, the perception of sharpness is related to the clarity of detail and edge definition in images, which is extremely helpful in improving target detectability. Therefore, sharpness also had a great influence on PQTD for fusion images in the category of Sea and Sky.

For predicting color fusion image PQSU in the three categories, CH was the first important attribute and S was the second in the created regression models. Specifically, in the category of Sea and Sky, CH represented observably higher anticipation to the scores of PQSU compared with S. Because fusion images in the category of Sea and Sky contained fewer details, image scenes were easily recognized by their color and observers were insensitive to sharpness changes. CH was much more important than S for predicting color fusion image PQSU in the category of Sea and Sky. Whereas, in the category of Plants, CH represented only a little higher anticipation to the scores of PQSU compared with S. Because fusion images in the category of Plants contained more details, the influence of S on PQSU was nearly equal in significance as CH.

In conclusion, perceptual quality based on visual tasks is proposed to measure color fusion image quality comprehensively, including two submetrics: PQTD and PQSU. A psychophysical experiment is performed to explore the relationship between the QAs and the proposed metrics. Multiple regression statistical analyses are conducted to derive prediction models for PQTD and PQSU. The result showed that PCTB and S are significantly good in predicting PQTD. CH and S are included in the model that predicted PQSU very well. In the three image categories, the influence levels of PCTB, S, and CH on PQTD and PQSU are different. The proportional coefficients in prediction models for PQTD and PQSU are different; however, the basic forms of prediction models are unchanged.

PQTD and PQSU are able to measure color fusion image quality comprehensively; however, they are difficult to model and quantize directly. The four QAs (i.e., PCTB, S, CH, and CN) are able to describe image quality from different aspects successfully. These QAs are not only suitable for image quality subjective evaluation, they are also convenient for computation and quantization. Therefore, the result of our research provides a way to solve the difficult problem on how to evaluate color fusion image quality based on visual tasks objectively. In other words, through establishing objective evaluation models of PCTB, S, and CH, objective evaluation of color fusion image quality based on vision tasks can be achieved in combination with the corresponding prediction model proposed in this letter according to different types of image scenes.

This work was supported by the National Natural Science Foundation of China (No. 60971010) and Pre-research Foundation of General Armament Department of China (No. 40405030302). The authors would like to thank all the observers that participated in the experiment. They would also like to thank Mr. Nigel J. W. Morris for providing the visible and infrared source images used in this work.

References

1. Z. Wang and A. C. Bovik, *Modern Image Quality Assessment* (Morgan & Claypool Publishers, USA, 2006).
2. S. Y. Choi, M. R. Luo, M. R. Pointer, and P. A. Rhodes, J. Imaging Sci. Technol. **25,** 040904 (2008).
3. M. Pedersen, N. Bonnier, J. Y. Hardeberg, and F. Albregtsen, J. Electron. Imaging **19,** 011016 (2010).
4. X. Gu, S. Sun, and J. Fang, Chin. Opt. Lett. **5,** 274 (2007).
5. E. A. Essock, M. J. Sinai, J. S. McCarley, W. K. Krebs and J. K. DeFord, Human Factors **41,** 438 (1999).
6. A. Toet, N. Schoumans, and J. K. Ijspeert, in *Proceedings of Fusion 2000*, TUD3-17-TUD3-23 (2000).
7. A. Toet and E. M. Franken, Displays **24,** 25 (2003).
8. J. Shi, W. Jin, and L. Wang, J. Infrared and Millimeter Waves **24,** 236 (2005).
9. W. K. Krebs and M. J. Sinai, Human Factors **44,** 257 (2002).
10. P. G. Engeldrum, J. Imaging Sci. Technol. **48,** 446 (2004).
11. S. Yin, L. Cao, Y. Ling, and G. Jin, Infrared Physics & Technology **53,** 146 (2010).
12. Y. Yuan, J. Zhang, B. Chang, H. Xu, and Y. Han, Chin. Opt. Lett. **9,** 011101 (2011).
13. J. Shi, W. Jin, L. Wang, and H. Chen, in *Proceedings of Infrared Components and Their Applications* (ed.) 594 (2005).
14. J. Caviedes and F. Oberti, Signal Process: Image Communication **19,** 147 (2004).
15. B. W. Keelana and E. W. Jin, in *Proceedings of Image Quality and System Performance VI* 72420Z-1 (2009).
16. A. J. Calabria and M. D. Fairchild, J. Imaging Sci. Technol. **47,** 479 (2003).
17. S. Guan and P. Hung, Color Res. Appl. **35,** 213 (2010).
18. K. E. Burchett, Color Res. Appl. **27,** 28 (2002).
19. L. C. Ou, P. Chong, M. R. Luo, and C. Minchew, Color Res. Appl. **36,** 355 (2011).
20. S. N. Yendrikhovskij, F. J. J. Blommaert, and H. de Ridder, Color Res. Appl. **24,** 52 (1999).
21. H. de Ridder, F. J. J. Blommaert, and E. A. Fedorovskaya, in *Proceedings of Human Vision, Visual Processing, and Digital Display VI* 51 (1995).
22. P. Bodrogi and T. Tarczali, Color Res. Appl. **26,** 278 (2001).
23. A. Toet and J. Walraven, Opt. Eng. **35,** 650 (1996).
24. A. M. Waxman, A. N. Gove, D. A. Fay, J. P. Racamato, J. E. Carrick, M. C. Seibert, and E. D. Savoye, Neural Networks **10,** 1 (1997).
25. L. Wang, S. Shi, W. Jin, S. Zhang, Y. Zhao, and J. Huang, Trans. Beijing Institute of Technol. **28,** 1 (2008).
26. L. Wang, Y. Zhao, W. Jin, S. Shi, and S. Wang, in *Proceedings of Signal Processing, Sensor Fusion, and Target Recognition* 65671G1 (2007).
27. S. Shi, L. Wang, W. Jin, and Y. Zhao, J. Optoelectron. Laser **20,** 1552 (2009).
28. S. Shi, L. Wang, W. Jin, and Y. Zhao, in *Proceedings of International Symposium on Photoelectronic Detection and Imaging* 66230B (2008).
29. S. Shi, L. Wang, W. Jin, and Y. Zhao, Acta Photonica Sinica **39,** 553 (2010).
30. Nigel J. W. Morris and S. Avidan, "Database of IR and visible image pairs", http://www.dgp.toronto.edu/∼nmorris/data/IRData/ (June 24, 2007).
31. R. S. Berns, Displays **16,** 173 (1996).
32. ITU-R Recommendation BT.500-12, "Methodology for the subjective assessment of the quality of the television pictures," http://www.itu.int/rec/R-REC-BT.500-12-200909-I/en (September, 2006).
33. J. L. Rodgers and W. A. Nicewander, The American Statistician **42,** 59 (1988). 7
34. N. R. Draper and H. Smith, *Applied Regression Analysis* (John Wiley & Sons, Inc, New York, 1998).
35. H. Arnkil, in *Proceedings of Interim Meeting of the International Colour Association* 15 (2008).