A Psychophysical Experiment Evaluating the Color Accuracy of Several Multispectral Image Capture Techniques

Ellen A. Day, Roy S. Berns, Lawrence Taplin, Francisco Imai
Munsell Color Science Laboratory
Rochester Institute of Technology, NY

Abstract

A paired comparison psychophysical experiment was performed to evaluate the quality of several imaging techniques. Images rendered for an LCD were compared with two- and three-dimensional objects viewed in a multi-source light booth. The test images were a combination of different spectral capture and image reconstruction techniques estimating spectral reflectance factor transformed into colorimetric images by imposing an illuminant and observer, in this case simulated daylight (6800K) and incandescent (2700K) for the 1931 CIE observer. Some images were also transformed directly to colorimetry. To compare the color accuracy with conventional RGB imaging, color-managed images from a typical consumer camera were included in the experiment. All image types captured using the research-grade camera performed very well, regardless of the illuminant, number of channels, or transformation method. A set of colorimetric images captured using previously optimized filters performed as well as the multispectral images. The commercial-grade images proved to be inferior to those created with the research-grade system, as expected.

Introduction

Most of today's imaging technologies are based on a three-channel system. This is made possible by metamerism.[1,2,3] Multispectral imaging can help or even correct many of the problems associated with three-channel systems.[1,4] This is especially important when exact reproductions are required, as in the replication of fine art pieces. This is true for both scientific applications of the reproductions, as well as respect for the artist's original intentions for the piece.[5] Multispectral imaging allows us to calculate the color of an object for any arbitrary observer and illuminant by capturing information about the spectral reflectance of every pixel in a given scene.[6] Of course this leads to a tremendously large amount of data, which must be handled efficiently in order to be useful.

This experiment is part of a larger project to develop a research-grade imaging system for use in museums, which will include a computer-controlled multispectral digital camera and associated software.[7] The purpose of this experiment is to evaluate the multispectral techniques that have been developed, thus far, at the Munsell Color Science Laboratory for their possible integration into the system.

The Imaging Process

The images for this experiment were acquired over two sessions. A Roper Scientific Photometrics Quantix monochrome camera fitted with a Cambridge Research and Instrumentation, Inc. liquid crystal tunable filter was used in the first session to create thirty-one-channel narrow-band images. In the second session, the Quantix camera was used again, along with six filters held in a filter wheel. First, six different filters were used to create the six channels: near infrared, red, yellow, green, turquoise, and blue. Second, a separate set of six channels were created using the red, green, and blue filters, and then the same three filters again with the addition of a Kodak Wratten No. 38 light-blue filter. In addition, three of the channels (red, green, and blue) were used separately to create a set of three-channel images. The purpose was to have a set of images with greater than three channels to compare with the images that use only the typical three channels, while still using the camera setup in this research. These three filters were previously chosen during extensive research on the design of optimal spectral sensitivities for a digital imaging system.[8-10] The filters were optimized for various illuminants and were evaluated with several metrics. This results in a filter set that performs as a colorimeter when used in conjunction with a well designed imaging system. These were the same red, green, and blue filters that were used in the six channel imaging. In addition, a set of images was taken using a Nikon D1 camera.

Targets

Targets were designed to amplify the camera system's vulnerabilities. The first target included a Gretag Macbeth ColorChecker and an original watercolor painting (this target is denoted *cc*). The second target included a Gretag Macbeth ColorChecker DC and a set of Gamblin Conservation colors (*ccdc*). These paints consist of many important pigments on an artist's palette. The third target consisted of large color-card samples distributed by Sherwin Williams (*paint*).

The next three targets were three-dimensional object set-ups. Three-dimensional objects are necessary in order to show defects in the system, especially relating to shading, gradients, and saturation. These effects are mainly related to the illumination of a three-dimensional surface. However, such objects were used to show that the system could be employed in every day scenes, and not just for two-dimensional imaging. These targets were *baby*, *fruit*, and *nature*. Figure 1 shows the targets.

All six of these targets included a Halon tablet which was used to help determine optimal exposure times. Two final targets were also imaged. These were used for spatial corrections and consisted of gray Color-aid paper, specifically, GRAY 4 and GRAY 6.5.

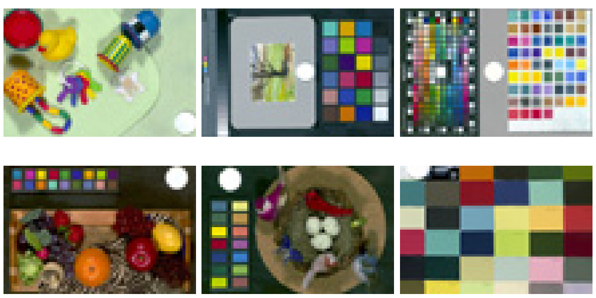

Figure 1. Targets used in this research (top left to bottom right): baby, cc, ccdc, fruit, nature, paint.

Image Transformations

After applying a spatial correction to all images to take into account the non-uniformity of illumination, the images were transformed into colorimetric images. All transformation matrices were created using the ColorChecker DC target. An illuminant and observer were imposed upon the multispectral images when the transformation matrices were applied. The CIE 1931 standard observer was always used. An incandescent light source and a filtered tungsten daylight simulator were used as the illuminants.

Information on the image transformations can be found in reference 11. References 2 and 12 through 15 contain information on other useful transformations. The transformations used in this research are summarized below.

For the thirty-one-channel images (narrow band, denoted *tf_pinv*), a simple pseudo-inverse transformation was used to create the transformation matrix. Equation 1 shows this transformation:

$$\boldsymbol{M}_{(m,m)} = \boldsymbol{R}_{(\lambda,p*n)}(\boldsymbol{DC}_{(m,p*n)})^T[(\boldsymbol{DC}_{(m,p*n)})(\boldsymbol{DC}_{(m,p*n)})^T]^{-1} \qquad (1),$$

where **M** is the (31 x 31) transformation matrix, **R** is the matrix of known reflectances of the original ColorChecker DC target, and **DC** is the matrix of the patch digital counts following a spatial correction. The subscript *m* represents the number of channels, in this case, 31 channels. The number of pixels per patch and the number of patches are represented by *p* and *n*, respectively. *T* denotes matrix transpose and *-1* denotes matrix inversion.

A similar transform was used for two of the transformations of the wide band images: once for the six-filter images (*pinv6*) and once for the three-filter images with the Wratten No.38 filter (creating six channels, *pinv6W*). The resulting transformations matrices were (31 x 6).

The two sets of six-channel images were also transformed with a two-step process using eigenvector analysis (denoted *pca6* and *pca6W*). First, a set of eigenvectors was derived from the spectral reflectances of the target. Based on preliminary analyses, six eigenvectors were always used.[6] The second part of the process included a pseudo-inverse calculation to compute a transformation matrix. Equation 2 shows this calculation:

$$\boldsymbol{M}_{(q,m)} = (\boldsymbol{E}_{(m,q)})^T[(\boldsymbol{E}_{(m,q)})(\boldsymbol{E}_{(m,q)})^T]^{-1} \ldots$$

$$\boldsymbol{R}_{(\lambda,p*n)}(\boldsymbol{DC}_{(m,p*n)})^T[(\boldsymbol{DC}_{(m,p*n)})(\boldsymbol{DC}_{(m,p*n)})^T]^{-1} \qquad (2),$$

where **E** is the matrix of eigenvectors and the subscript *q* is the number of eigenvectors (six, in this case).

The three-channel RGB image transformation was a simpler pseudo-inverse (*RGB*). First, tristimulus values were calculated from the known reflectances, using traditional equations. Next, a pseudo-inverse was applied, as in Equation 1, using tristimulus values instead of reflectances to create a transformation matrix. The result was a (3 x 3) transformation matrix.

The Nikon D1 images were manipulated using a different technique (*D1*). This was necessary since consumer grade digital cameras have their own built-in gamma function. The digital signals from the camera were linearized using a two-degree polynomial. A pseudo-inverse between the linearized digital counts and the calculated tristimulus values was used to create the transformation matrix.

Psychophysical Experiment

A forced-choice paired comparison psychophysical experiment was performed on a colorimetrically characterized[15]Apple Cinema Liquid Crystal Display. Twenty-seven observers were required to choose which of two images on the screen more closely matched the color of the original target that was placed in an adjacent Macbeth Spectralight II light booth. The experiment was performed using two different light sources: simulated daylight (6800K) and incandescent (2700K). Figure 2 shows the experimental set-up.

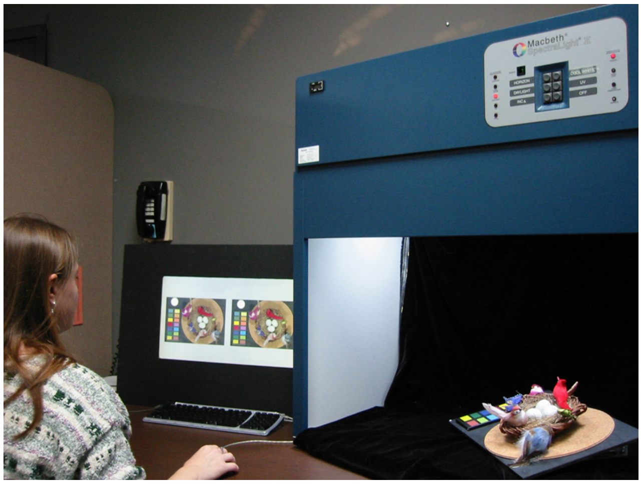

Figure 2. Experimental set-up. Please note that the room lights were turned off during actual experiments.

Results and Discussion

Thurstone's Law of Comparative Judgments (case V) was used to transform the observer data into interval scales.[16] The results for the daylight experiment are shown in Figure 3. The image types are shown on the x-axis. The y-axis shows the perceived color reproduction quality in interval scale units. The six plots represent the six different targets. The error bars on these plots were calculated in terms of interval scale units for a 95% confidence interval.

For two image types to be significantly different, their errors bars must not overlap. The plot for the fruit target in Figure 3 shows this well. In this plot, the maximum error for the D1 image type is below the minimum error for the other six targets. Therefore, the perceived color accuracy is significantly different.

All of the image types were judged equivalent to each other, with the exception of the D1 images. This was true for all targets with some variation in the degree of uncertainty. The significance of this result is that observers could not distinguish differences in the color reproduction accuracy of the various image types for the Quantix sensor, irrespective of the number of channels or types of transforms. The three-channel image type performed as well as the image types with a greater number of channels showing that a well designed three-channel system can achieve a high degree of color reproduction accuracy. The consumer camera had significantly lower accuracy.

The spectral reflectance of each patch on the original targets was measured and the corresponding colorimetric values were calculated using the light booth spectral power distributions. The patches included those on the ColorChecker DC from the *ccdc* target, the paint chips from the *paint* target, the Gamblin patches also from the *ccdc* target, and the Macbeth ColorChecker from the *cc* target. The colorimetric values were calculated similarly using the estimated spectra for the six- and thirty-one-channel images or the direct transforms from the three-channel and D1 images. The measured and estimated values were compared using the CIEDE2000 color difference equation for each of the four sets of patches used and the seven different image types. Because these calculations were made on color patches and not pictorial images, spectral-based color difference metrics were not required.

The plots in Figure 3 for the *cc*, *ccdc*, and *paint* targets show the results of the color difference evaluation. Since the set of Gamblin paints was on the same target as the ColorChecker DC, the CIEDE2000 values seen in the *ccdc* plot are the average of the color difference results for the two color sets combined. The triangles denote average CIEDE2000 values for each image type and the circles denote maximum values. The dotted lines only help to visualize the pattern of color differences in comparison to the paired comparison data. Also note that the CIEDE2000 axes are reversed, so that the larger color differences are at the bottom of each plot.

The trend of the maximum color differences seems to enhance the trend seen in the paired comparison analysis. The maximum color differences for the D1 image type are higher than those of the other six image types, all created with the Quantix camera. The trend can also be seen in the average color difference values; however, it is not as pronounced.

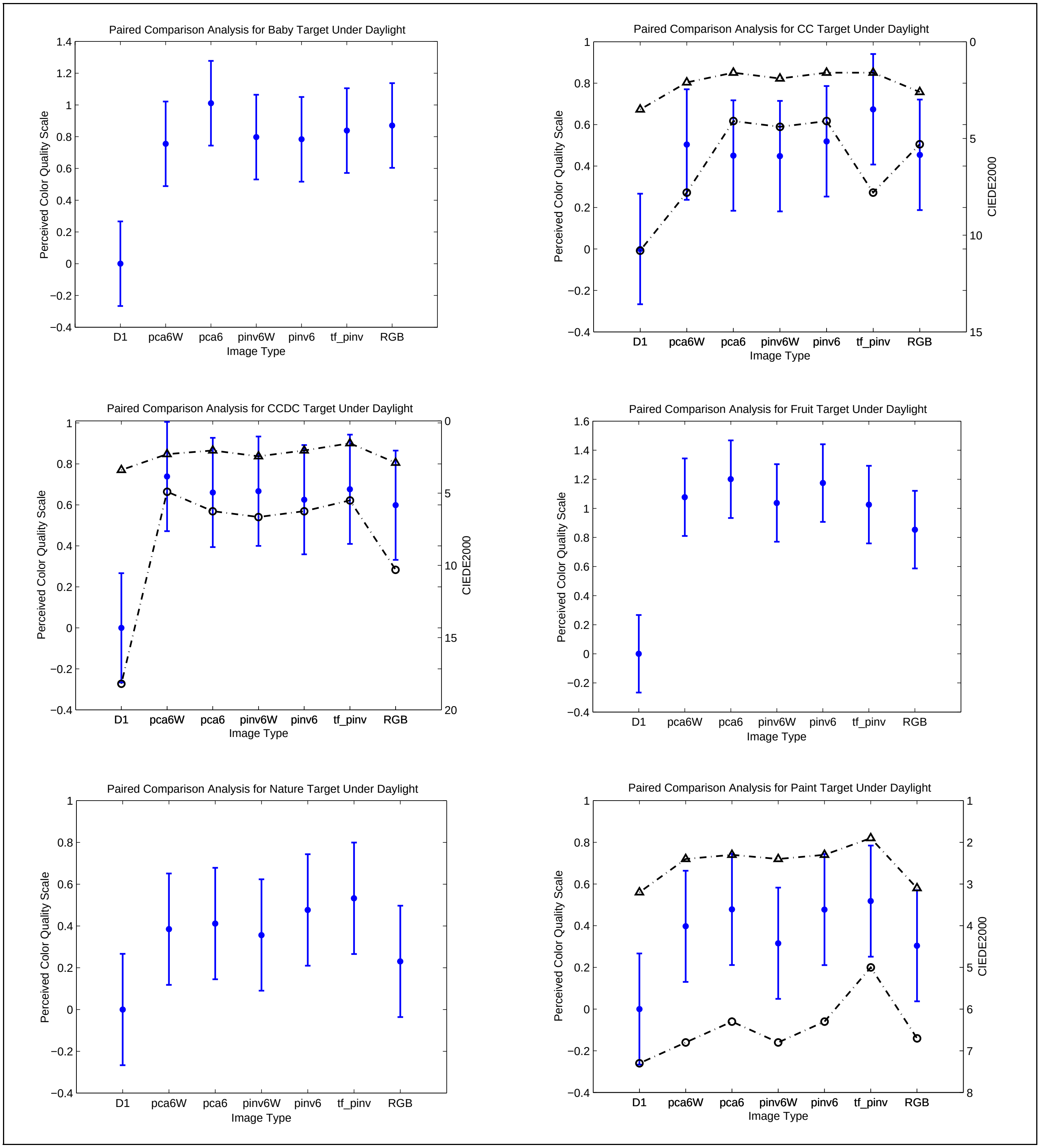

Figure 3. Paired comparison results for daylight experiment. In three plots, color difference results are also shown (CIEDE2000). The triangles denote average values and the circles denote maximum values.

Four of the plots resulting from the analysis of the experiment performed under incandescent illumination are shown in Figure 4. Two redundant plots were left out to conserve space. A visual analysis of this set of plots shows that under incandescent illumination, observers, as a whole, were more ambiguous in their judgments than under daylight. This ambiguity is shown by the error bars that overlap to a greater extent. In this case, it is not possible to conclude if any of the image types might perform superiorly to any others under incandescent illumination. The targets were originally imaged under an incandescent light source. While the imaging illuminant has no effect on the multi-channel images, this improved the color accuracy of the D1 images, in some cases. The nearly matched camera-taking and display illuminations improved the colorimetric performance.

The color difference trends for incandescent illumination are also seen in Figure 4. As in the daylight analysis, the maximum color difference trends enhance those of the paired comparison analysis. The maximum color difference clears up the ambiguity that was illustrated in the paired comparison analysis under incandescent illumination. Specifically, the D1 again shows a larger maximum color difference than the color differences for most of the other image types.

For the ColorChecker target under both illuminants, the maximum color difference is relatively high for the thirty-one-channel image type (*tf_pinv*), and even slightly higher than that of the D1 image type under incandescent A. It should be noted that this large color difference occurs only for the black patch of the ColorChecker. Examination of the *average* color differences, in this case, reassures us that overall, the color differences for the 31 channel image type are acceptable. In addition, the maximum color difference did not occur for the same ColorChecker patch for the D1 and 31 channel image types. For the D1, the maximum color difference occurred for the *purplish-blue* patch on the target. This confirms prior knowledge that the D1 camera cannot reproduce blue hues well.

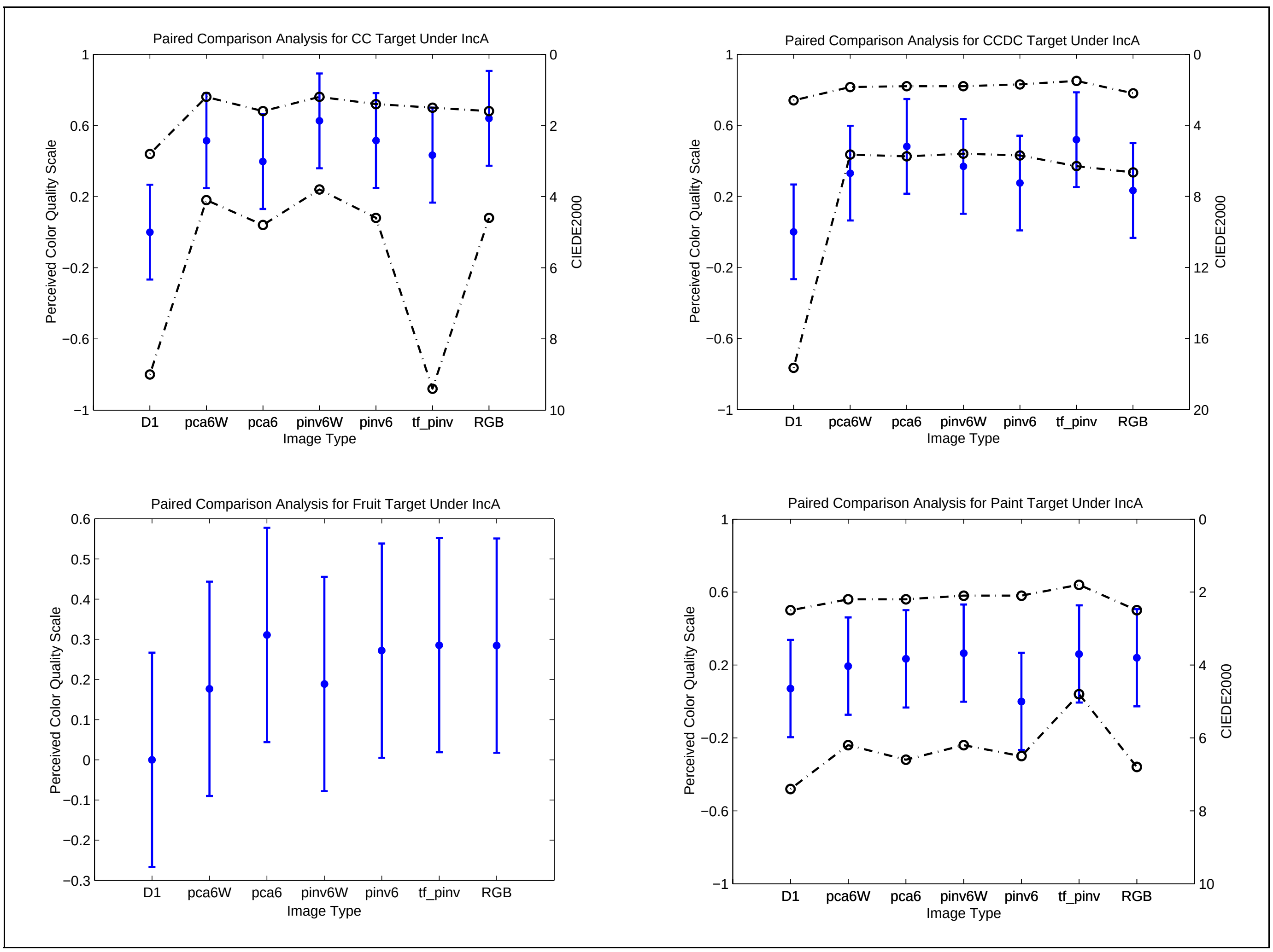

Figure 4. Paired comparison results for incandescent A experiment.

Conclusions

An experiment was conducted under two illuminants in order to evaluate the color reproduction accuracy of various imaging techniques and transformations. A research-grade Quantix camera captured four six-channel image types, a 31 channel image type, and a three-channel colorimetric image. A Nikon D1 captured images for comparison as a commercial-grade digital camera.

All image types captured using the Quantix camera performed equivalently, regardless of the illuminant, number of channels, or transformation method. The three-channel image type performed as well as the image types created using more channels, signifying that an imaging system with carefully designed spectral sensitivities can perform as well as multi-channel systems. The D1 images proved to be inferior to those created with the research-grade system, as expected.

To evaluate if physical results correlated with psychophysical results for the color difference experiment, trends in color difference values were compared to the results of the paired comparison analysis. Overall, the results of the color difference evaluation mimic and enhance those of the paired comparison analysis showing that a psychophysical experiment proves useful in the evaluation of color accuracy.

Acknowledgements

The authors would like to thank the National Gallery of Art, Washington, D.C., the Museum of Modern Art, New York, and the Andrew W. Mellon Foundation for their financial support of the Art Spectral Imaging (Art-SI) Project.

References

1. König F, Werner P, A Multispectral Scanner, in L. MacDonald and M.R. Luo, Eds., Colour Imaging: Vision & Technology, John Wiley & Sons, Chichester, 1999, pg. 129-143.
2. Hill B, Multispectral Color Technology: A Way towards High Definition Color Image Scanning & Encoding, Proc. SPIE, pg. 2-13. (1998).
3. Berns, RS, Billmeyer and Saltzman's Principles of Color Technology, 3rd edition, New York, NY, 2000.
4. Berns RS, Challenges for Color Science in Multimedia Imaging, in L. MacDonald and M.R. Luo, Eds., Colour Imaging: Vision & Technology, John Wiley & Sons, Chichester, 1999, pg 99-127.
5. Maitre H, et al, High Quality Imaging in Museum: from Theory to Practice, Proc. SPIE, 30-39 (1997).
6. Hardeberg JY, et al, Multispectral Image Acquisition & Simulation of Illuminant Changes, in L. MacDonald and M.R. Luo, Eds., Colour Imaging: Vision & Technology, John Wiley & Sons, Chichester, 1999, pg. 145-161.
7. Art-SI.org: Art Spectral Imaging. http://www.art-si.org. Accessed 27 January 2003.
8. Quan S, Evaluation and Optimal Design of Spectral Sensitivities for Digital Color Imaging, RIT Ph.D. Thesis 2002.
9. Quan S, et al, Optimal Design of Camera Spectral Sensitivity Functions Based on Practical Filter Components, IS&T/SID 9th Annual Color Imaging Conference, pg. 326-331. (2001).
10. Quan S, et al, Unified Measure of Goodness and Optimal Design of Spectral Sensitivity Functions, J. Imaging Sci. Techn., 46, 485-497 (2002).
11. Imai FH, Taplin LA, Day EA, Comparison of the Accuracy of Various Transformations from Multi-band Images to Reflectance Spectra (As Part of End-to-End Color Reproduction from Scene to Reproduction Using Spectral Imaging), Munsell Color Science Laboratory Technical Report 2002, http://www.art-si.org/publications.html.
12. Keusen T., Praefcke W, Multispectral Color System With An Encoding Format Compatible to Conventional Tristimulus Model, IS&T/SID 3rd Color Imaging Conference, pg. 326-331. (1995).
13. König F, Herzog PG, On the Limitations of Metameric Imaging, IS&T PICS Conference, pg. 163-168. (1999).
14. König F, et al, Displaying Multispectral Images, International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives Proceedings, pg. 123-126. (1999).
15. Day EA, Berns RS, Taplin, LA, Colorimetric Characterization of a Computer-Controlled Liquid Crystal Display, Manuscript in preparation, 2003.
16. Engeldrum, P, Psychometric Scaling: A Toolkit for Imaging Systems Development, Imcotek Press, Winchester, MA, 2000.

Biography

Ellen A. Day is a M.S. student of color science in the Center for Imaging Science at Rochester Institute of Technology. Her thesis topic is *The Effects of Multi-channel Visible Spectrum Imaging on Perceived Spatial Image Quality and Color Reproduction Accuracy*. She graduated from Rochester Institute of Technology in November 2000 with a B.S. degree in Imaging and Photographic Technology.