Reliability Issues for Color Measurement in Quality Control Applications

John C. Briggs, David J. Forrest and Ming-Kai Tse
QEA, Inc.
755 Middlesex Turnpike, Unit 3, Billerica MA 01821 USA
Tel: (978) 528-2034 · Fax: (978) 528-2033
e-mail: info@qea.com
URL: www.qea.com

Paper presented at the IS&T's NIP14
International Conference on Digital Printing Technologies
October 18-23, 1998, Toronto, Ontario, Canada

Reliability Issues for Color Measurement in Quality Control Applications

John C. Briggs, David J. Forrest, and Ming-Kai Tse
QEA, Inc.
Burlington, MA/USA

Abstract

As color digital printing expands into the business and consumer markets, the importance of "color accuracy" has increased. As a result, more and more people have had to become familiar with concepts such as color difference equations and *DE**. While many of the concepts of color measurement are clearly explained in several current books, the myriad practical issues that contribute to color measurement error are seldom discussed. This leaves the practitioner of color measurement with tools for measuring color but little understanding of the key variables that affect the measurement process.

This paper examines critical factors affecting accuracy in color measurements and sheds light on issues involved in integrating color measurement into on-line or automated applications. The focus of the paper is to bring out key calibration issues, including frequency calibration, gain adjustment, dark reference and light reference. The repeatability and accuracy of color measurements are examined. Key sources of error are investigated, and suggestions are made for reducing sensitivity to them.

Introduction

Color in consumer products has become a distinctive feature of our modern society. From the clothes we wear, the cars we drive, and the houses we live in, to the shows we watch, books we read, and catalogues we shop from, color plays a central role. But it is not enough for an object to be colorful. The color must be "right," and in many cases it must "match." The blouse must match the skirt or shoes, and the television and catalogue advertisements for these items must closely represent the products delivered to our front door.

However, accurate reproduction of color is not an easy task, and part of the reason for this is that accurate color measurement is quite difficult. Accurately measuring an object's color is far more difficult than accurately measuring its weight or size. There are many reasons for this uncertainty. First, many different standards and methods of measuring color are currently in use. Secondly, compatibility between instruments from different manufacturers is frequently poor. Thirdly, many artifacts and confounding factors complicate color measurement. Lastly, instrument repeatability is sometimes poor. To make color measurement accurate, suitable calibrations of the instrumentation are required.

Many researchers have worked on the problem of color measurement and a number of fine articles and books have been written[1-4]. The details of calculating tristimulus values (XYZ) and L*a*b* have been clearly laid out by CIE[5] and others. Many subtleties involved in color measurement have been investigated and reported.[6-10] Articles in the literature discuss the repeatability of individual instruments and compare instruments from different manufacturers. But despite all that has been written, many of the most basic and important calibration issues involved in color measurements are difficult to find in the literature.

The spectrophotometer is a basic tool of color measurement, quantifying the spectrum of light reflected by or transmitted through an object. In this paper, we will go through the basics of setting up and calibrating a spectrophotometer. Key calibration issues will be discussed, including frequency calibration, gain adjustment, dark reference and light reference. The effects of these calibrations on color accuracy (*DE**) will be discussed. Sources of measurement error such as bulb warm-up, ambient temperature change, measurement adapter spacing, and line voltage variation will also be reviewed.

Rather than considering these issues in the abstract, we will look at how they affect one particular spectrophotometer. For this purpose, we will use an OEM-type spectrophotometer — that is, one designed for integration into other products — because it will allow us to control all the low-level functions of the color measurement process, making it possible to investigate independently the impact of various sources of error. The types of errors investigated can have significant impact in an increasingly common spectrophotometer application — on-line or automated color measurement.

Application of a Spectrophotometer

General Configuration of a Spectrophotometer

Designers of handheld spectrophotometers typically bundle all the components in a package roughly the size of a brick. While this arrangement is very portable, a device this large can be difficult to integrate into an on-line application. For this reason, OEM-type spectrophotometers are sometimes configured with fiber optic bundles to carry light to and from the sample, allowing the other components to be kept in a more remote location. This arrangement is shown in Figure 1. The fiber bundles are typically joined by a measurement adapter that controls the lighting geometry. The small size of the measurement adapter (usually less than 25 mm in all its dimensions) makes it easy to integrate into on-line applications.

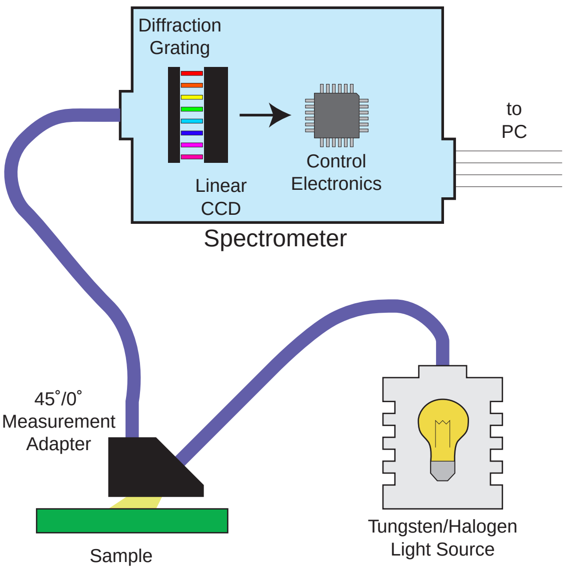

Figure 1: Schematic representation of a reflection spectrophotometer.

The components of an OEM-type spectrophotometer are a light source; fiber optic bundles; a measurement adapter (also known as illumination optics) for controlling lighting and sensing geometry; and a spectrometer for measuring spectra of emitted light. The light source typically uses a broad-spectrum tungsten/halogen bulb to provide light energy throughout the visible spectrum (approximately 380 nm to 780 nm). One of two fiber optic bundles carries the light from the light source to the measurement adapter. In the 45°/0° arrangement shown in Figure 1, the light comes in at 45° to the sample normal, and reflected light is collected normal to the sample. Other geometries (0°/45° and *integrating sphere optics*, discussed below) are also available.[2] The second fiber optic bundle carries reflected light from the measurement head to the spectrometer. The spectrometer splits the light into its frequency components using a diffraction grating. The light then falls on a linear CCD (charge coupled device) which measures the intensity of light in each frequency band. The measurement data are then sent to a computer (PC) for additional processing and analysis.

Frequency Calibration

The first step in calibrating the spectrophotometer is to determine which CCD elements in the spectrometer correspond to which wavelengths of light. The spectrometer calibration is performed using spectral calibration lamps, typically mercury and argon. The lamps have distinct peaks in their radiation spectra, traceable to NIST standards with an accuracy of ±1 nm.

Our spectrophotometer came with a calibration certificate consisting of a table with 16 pairs of CCD index values (0 to 2047) and their wavelengths (λ). The overall accuracy of this calibration was claimed to be ±3 nm.

Since it is necessary to know the wavelength represented by each CCD element, but the calibration process generates data at just 16 discrete wavelengths, some interpolation is needed. Figure 2 shows the calibration data with several different curve fit options. A simple linear relationship fits the data well, with a maximum error of 4.24 nm between the data and the curve. A better relationship can be obtained using a second order (parabolic) fit. This has a maximum error of only 0.43 nm. Unfortunately, the second order fit adds a significant complication to the calculation of the tristimulus values (X, Y, Z, CIE 1931) by making increments of λ non-uniform. We found that a good compromise was to use a linear curve fit between 400-700 nm only. This is the range of wavelengths that have the greatest weights in the CIE color matching functions, as shown in Figure 3. A linear fit over this range of frequencies results in a maximum error of only 2.34 nm. This approach was adopted for the measurements discussed here.

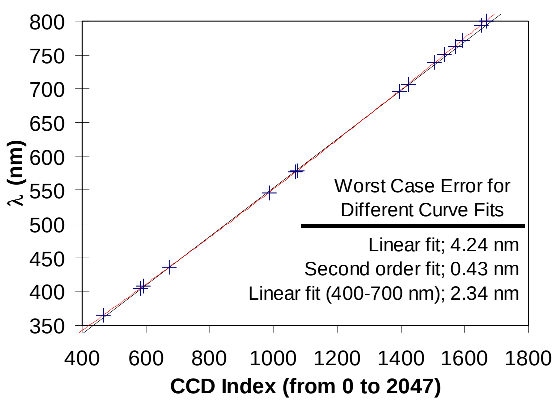

Figure 2: Frequency calibration data for the CCD element. The spectrometer has a resolution of approximately 0.3nm.

It should be noted that the impact of errors in wavelength ($\Delta\lambda$) on ΔE^* depends on the nature of the color being measured. We know that for spectrally flat colors like black, white or gray, small changes in frequency should not affect the measured color. The impact of wavelength errors on the accuracy of other color measurements, however, is less well understood and will be the subject of future investigations.

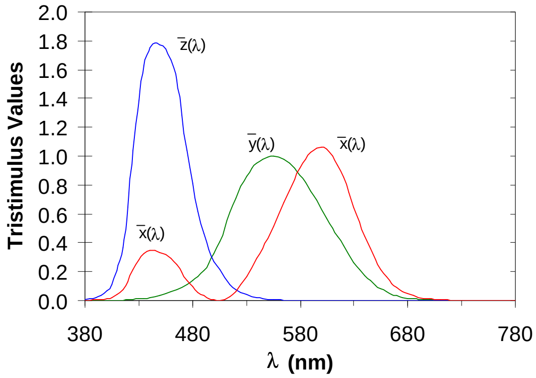

Figure 3: Tristimulus values for 2° observer.

Setting Illumination Level

As with most optical sensing devices, the exposure setting must be carefully adjusted for the spectrophotometer to work properly. This adjustment is made in two ways. First, the amount of light entering the spectrophotometer must be chosen or adjusted. For our sample spectrophotometer, we chose a 20W light source with no adjustment. Second, the *integration time* must be adjusted. This is the time during which the sensor collects light energy. It is adjusted electronically, typically through software control via the PC.

CCD elements are limited at the low end of the exposure range by dark current and at the high end by saturation. Even with no illumination, the CCD element will give a reading larger than zero. This is known as dark current, a form of electronic noise in the CCD element. At the high end of the exposure range, with excessive light energy, the CCD elements will no longer increase linearly as light intensity increases. This deviation from linearity is termed saturation.

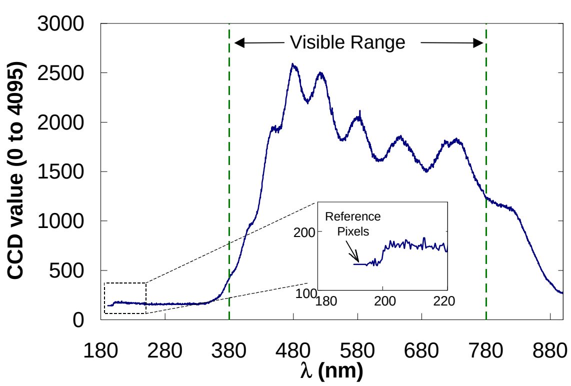

Figure 4: Spectrum of light from a white reference standard measured by the spectrophotometer with integration time of 100.

To check the exposure setting on our spectrophotometer, the measurement adapter was placed above a white reference standard. The light source was turned on and allowed to warm up, and spectra were acquired from the white reference over a range of integration times. Figure 4 shows a typical measured spectrum. The horizontal axis plots the wavelengths corresponding to the 2048 CCD elements. The vertical axis plots light intensities between 0 and 4095. The light intensities reported are a function of 1) the light source intensity, 2) the reflectance of the white standard, and 3) the sensitivity of the CCD elements. Each of these parameters is a function of frequency.

One of the requirements for the spectrophotometer is that there be sufficient light in the visible range, 380-780 nm. As can be seen from Figure 4, this condition is met except near 380 nm, where the light intensity is quite low. This deficiency is most probably a function of the light source. While it would be desirable to increase the light intensity in this range, it is probably not critical because the magnitudes of the color matching functions are quite small in that frequency range, as can be seen in Figure 3.

Also shown in Figure 4 are *reference pixels*, CCD elements that are shielded from illumination by a mask. The reference pixels provide a good measure of dark current from the CCD elements and can be used for dark reference calibration, as discussed below.

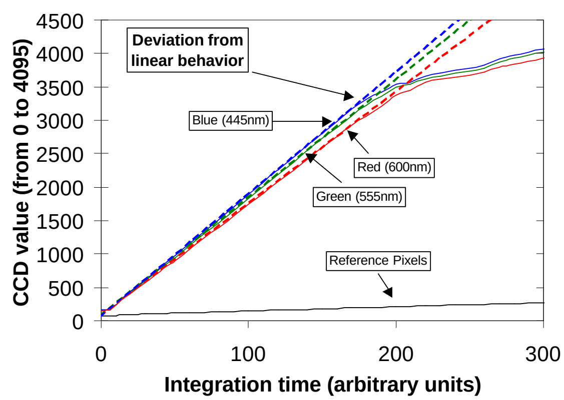

Figure 5: Linearity of CCD elements with integration time.

Figure 5 shows the effect of increasing the integration time for three different wavelengths (red, green and blue). Increasing the integration times causes linear increases in the output of the CCD elements until saturation is reached. As the figure shows, the spectrophotometer starts deviating from linear behavior long before it reaches its maximum value of 4095. At about 3200, the gain of the CCD elements starts dropping significantly. It is therefore important to adjust the light and the integration time so that no region of the spectrum exceeds 3200.

Figure 5 also shows the effect of integration time on the reference pixels. Even though these pixels have no illumination, their values increase with increasing integration time due to the dark current of the CCD elements.

The effect of saturation on *DE** is quite dramatic but difficult to quantify. If the spectrometer is saturating, then it will probably saturate during light reference calibration; and because the light reference is used in all color measurement calculations, these will be affected by the saturation. Errors due to saturation will probably show up as color shift, because some parts of the spectrum will saturate at lower intensities than others.

Dark Reference and Light Reference

With the frequency calibrated and the illumination level and integration time set, the percent reflectance must be calibrated at each frequency. Since the spectra acquired by the spectrophotometer are the product of the CCD sensitivity, the illumination, and the sample reflectance, a calibration must be performed to extract the sample reflectance.

The basic strategy for doing this is to perform a two-point calibration at each frequency. This calibration scheme assumes a linear relationship between percent reflectance and CCD values. Figure 5 shows this to be a good assumption, as long as the CCD values are below saturation (3200 for our spectrometer). As with any two-point calibration scheme, it is generally better to interpolate than extrapolate. The two calibration points should be at the extreme ends of the range of percent reflectance, using a black surface for dark reference and a white surface for light reference.

Dark Reference

There are several different ways to perform a dark reference. The simplest way is to turn off the light source and measure the spectrum. Using this method, it can be assumed that all frequencies will measure the CCD values corresponding to 0% reflectance. This assumption may actually be a source of error, however, since there may be "stray light" (that is, light not reflected off the sample) passing between the light source and the light collection point of the measurement adapter. The magnitude of stray light depends on the design of the measurement adapter. One way to account for stray light is to perform a dark reference using a commercially available black reflectance standard, which comes with a certification giving its percent reflectance values over a range of frequencies, traceable to recognized standards. Using this method, we determined that stray light was not a significant factor for our measurement adapter. (It should be noted that the black reflectance standard itself introduces a small but real degree of uncertainty.)

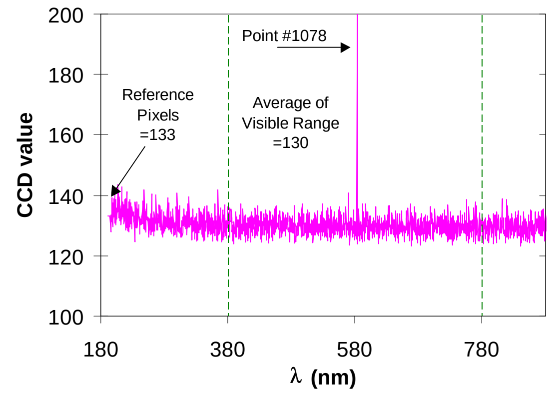

Figure 6: Dark reference spectrum. Integration time = 100. Light switched off.

Figure 6 shows a typical dark reference measurement spectrum. The spectrum is basically flat across the visible range. The one exception is CCD element #1078 at 584 nm which has a substantially larger response than the other elements – undoubtedly a manufacturing defect of some kind. This does not affect the dark reference measurement, however, because the two-point calibration accounts for this type of variation.

Once the dark reference data are acquired, they are stored in an array, say $D(l)$, in the computer. The dark reference may have to be checked periodically, but in principle it should remain constant as long as the operating environment of the temperature-sensitive CCD array does not change.

Another way to perform a dark reference is to use the reference pixels. Because the CCD elements are masked off, they are not exposed to light. As shown in Figure 6, the reference pixels in our spectrophotometer gave a reading of 133, very close to the average CCD reading of 130 in the visible range. Our analysis has shown that the difference of 3 CCD units would result in a *DE** of about 0.2. Despite the small error, this approach has some significant advantages. Among these, as we will see, are automatic update of dark reference, avoidance of thermal variation effects on the spectrometer, and elimination of the requirement for dark calibration.

Light Reference

Light reference is somewhat more complex than dark reference. Light reference for reflectance measurement requires the use of a reference standard with known reflectance properties. Such reference standards are commercially available.[11] The reference standard should be highly reflective so that it is brighter than the brightest object to be measured. If it is not, measurement data will have to be extrapolated, which is not desirable.

For these experiments, we used a white reference standard made from Albrillon, a pressed polymer powder. The sample came with a calibration certificate showing percent reflectance as a function of wavelength, traceable to NIST standards.

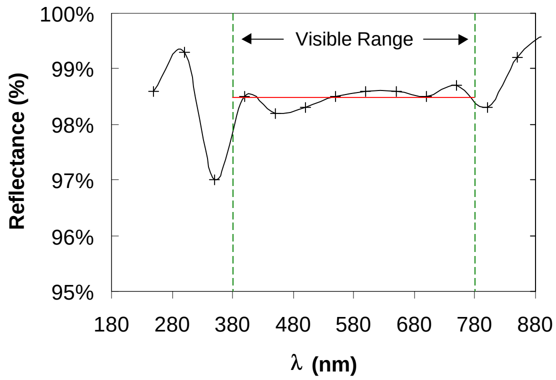

Figure 7: White reference standard reflection data.

The data from the calibration certificate is shown in Figure 7. The reference standard has an average reflectance of 98.5% over the visible range. The magnitude of reflection from the standard will need to be accounted for during the light reference calibration.

One simple and accurate calibration implementation is a look-up table that interpolates the reflectance values for each frequency. A simpler, though slightly less accurate, approach is to assume that the percent reflectance is a constant 98.5% over the visible range. Figure 7 shows that this assumption results in a maximum error of 0.3% at 450 nm. This is a small error, and the error for mid-tone colors would be still smaller. For our measurements, the white standard reflectance was assumed to be 98.5% across the visible spectrum.

To perform the light reference, the light is turned on and the measurement adapter is placed above the white reference standard. The results of a typical measurement are shown in Figure 4. Clearly, these data are far from spectrally flat, an effect due to the combined effects of the light source spectral power distribution and the spectral sensitivity of the spectrometer. The lack of flatness does not adversely affect color measurements, because the two-point calibration accounts for the observed variability. The measurement data are stored in an array, $L(l)$, in the computer. They will be used to calculate percent reflectance values, as shown in the next section.

In situations where light reference measurement is difficult to perform, another method, called a *dual-beam measurement*, may be preferable. For this method, a second spectrometer is used, and the light from the light source is carried through a bifurcated light fiber bundle. A portion of the light is reflected off the sample and carried to the first spectrometer, as in Figure 1. The remaining light goes directly to the second spectrometer. The value of the second spectrometer is that it directly measures the light output from the light source. In this way, changes in light spectral power distribution can be accounted for in real time, making it possible to eliminate them from the color measurement. Although this is clearly a more expensive method, it may be necessary in some critical applications.

For accurate color measurements, the spectral power distribution of the light source must be the same during color measurement as it was during the light reference measurement. A significant source of shift in spectral power distribution is bulb warm-up, a process that typically takes several minutes. To study the effect of bulb warm-up on spectral power distribution, a number of measurements were taken with the spectrophotometer. The results are shown in Figure 8. The data are normalized so that the value at time=0 is 100%. On average, the light output drops by about 2% during the first 10 minutes. This is a significant error. Even more troubling is that the change in light output is not uniform with respect to wavelength. As the figure shows, light output drops 1.7% for red and 1.9% for green; blue, however, drops by a much larger 3.4%. If this change were to occur between light calibration and sample measurement, there would be a significant shift away from blue for all colors measured. For our system, simply letting the bulb warm up for 10 minutes prior to calibration eliminated these problems.

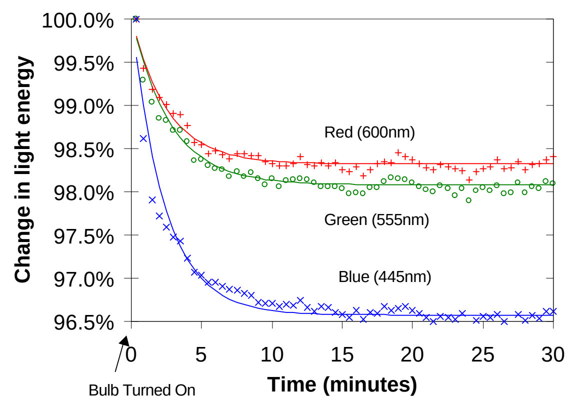

Figure 8: Effect of warm up time on light energy. Vertical axis is (CCD value)/ (CCD value at time 0)

Other issues can also affect the constancy of spectral power distribution of the light source. Temporal fluctuations in line voltage, for example, are a common problem. It is important that the voltage supplied to the light source be constant, and use of a well-regulated power supply is advisable. Another issue is bulb aging. Bulbs have a limited lifetime, and aging changes spectral power distribution. This happens slowly over time, so periodic checking of light reference should prevent significant measurement error due to bulb aging. Finally, some materials exhibit thermochromic effects, that is, color variations with temperature. With a continuous (i.e., not pulsed) light source of the type used here, significant sample heating can occur. The 20W light source used in these sample measurements caused significant thermochromic effects in some measurement samples. Thus, it may be necessary to choose a lower intensity light source.

Use of Dark and Light Reference

Having acquired the dark and light reference data, it is possible to convert measured spectra into percent reflectance values. Figure 9 illustrates the concept graphically. For each frequency, λ, the dark and light reference data are used to create a line. The equation of the line can be shown to be:

$$R(\lambda) = \frac{W(\lambda)}{L(\lambda) - D(\lambda)} CCD(\lambda) - \frac{D(\lambda)W(\lambda)}{L(\lambda) - D(\lambda)}$$

where

- $R(\lambda)$= percent reflectance value
- $CCD(\lambda)$ = CCD reading (from 0 to 4095)
- $W(\lambda)$= reflectance of the white standard (e.g. 98.5%)
- $D(\lambda)$ = dark reference value (in CCD units)
- $L(\lambda)$ = light reference value (in CCD units)

This calculation must be performed at each frequency to generate a plot of percent reflectance versus frequency. Once this has been done, the data can be used to compute the tristimulus values.

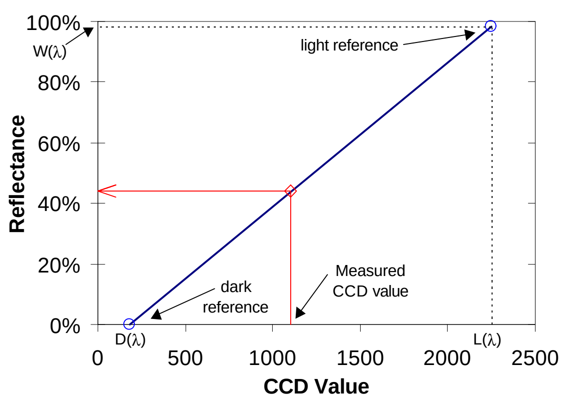

Figure 9: Example of how dark and light reference data is used to convert CCD values into reflectance data.

Use of Averaging

One point that should be considered when making spectral measurements for dark and light reference is how much averaging should be used. Averaging is used to reduce effects of measurement noise. Figure 10 shows the effects of averaging on dark reference values. For each data point on the graph, the specified number of spectra was acquired. For example, to obtain the data point derived from 32 averages (32 on the x axis), the following steps were taken. Thirty-two spectra were acquired and averaged, and the results were reported as a single spectrum. This process was repeated ten times, acquiring a new set of 32 spectra each time. The standard deviation (STD) of the resulting ten summary spectra was calculated, and the result was plotted on the graph.

Using a small number of averages (up to about 16) for this process was found to produce a significant reduction in the standard deviation. Beyond that, however, additional averaging was found to have little effect. In deciding how much averaging is appropriate, measurement time must be taken into account. With no averaging, the measurements took about 2 seconds. With 16 averages, they took about 7 seconds. This is probably not a problem for dark and light calibrations, which are done infrequently, but it could significantly decrease throughput of actual sample measurements. Ultimately, the decision is probably application dependent.

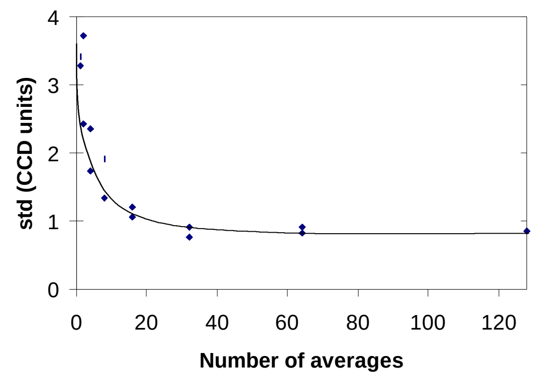

Figure 10: Effect of averaging on dark reference values. Standard deviation (std) of CCD value reading vs. the number of averages.

Consistency of Calibration

For accurate color measurements, the spectrophotometer must be calibrated correctly and stay in calibration. We have seen that changes in the calibration data can be caused by changes in the light source due to line voltage fluctuations and by bulb warm-up or aging. Another major source of calibration data variation is temperature variation affecting the spectrometer.

For accurate measurements, the operating environment of the CCD elements in the spectrometer must be free from significant temperature fluctuation. To illustrate the importance of this, Figure 11 shows the output values of the reference pixels at different temperatures. As the temperature increases, the CCD value increases across the whole spectrum. When the temperature increases from 23.5°C to 25°C, the CCD values increase an average of 24. This represents a 1% error in the reflectance reading. An increase from 23.5°C to 30°C changes the reading by 87, an error in reflectance of about 4%.

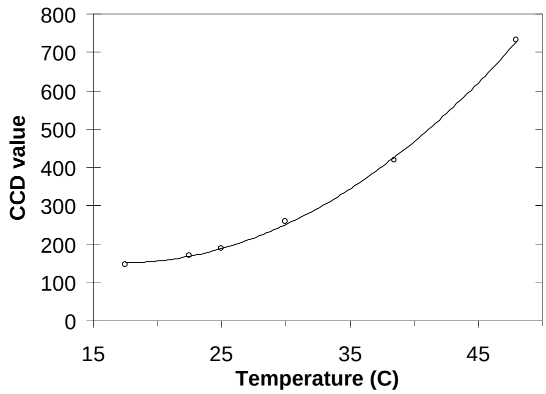

Figure 11: Effect of temperature on reference pixel measurement. Integration time = 150. Temperature is measured near the spectrometer.

These readings from the reference pixels, reflecting changes with temperature, can be used to reduce the effect of temperature on color measurement. By monitoring the reference pixels over time for any change in temperature, it can be determined whether new dark and light references should be performed. Dark and light reference updates can be done automatically.

Measurement Adapter

There are several common geometries for the measurement adapter. One common geometry places the light source at 45° to the surface normal and the optical sensor at 0°. This is referred to as a 45°/0° configuration. The reverse of this arrangement, with the light at 0° and the sensor at 45°, is known as a 0°/45° configuration. Another common geometry is known as *integrating sphere optics*. For the color measurement, integrating sphere optics can be configured either to include or to exclude specular reflections from the sample surface. These reflections arise from surface gloss and can have a significant effect on color measurement. Due to their geometries, 45°/0° and 0°/45° can not be configured to include specular reflections.

Another consideration in the measurement adapter design is aperture size. Aperture size controls the size of the area that can be measured. The choice of aperture size is application dependent. For small or curved objects, a small aperture may be required. For broad, flat areas, a larger aperture may be preferable for averaging color and reducing measurement variability. On the other hand, if the goal is to measure color non-uniformity over a broad area, a small aperture may be desirable.

Although most spectrophotometers are placed in contact with the surface to be measured, some applications may require non-contact measurement. For these applications, the measurement adapter must be designed for a specified “stand-off,” or gap, between the measurement head and the sample. If deviations in the stand-off distance occur during the measurement process (e.g. due to variations in part thickness), errors will be introduced into the color measurement.

To evaluate the magnitude of this effect, measurements were made on our spectrophotometer, whose measurement adapter was designed for a nominal 3.175 mm stand-off. The spectrophotometer was calibrated and measurements were made on the white reference standard over a range of actual stand-off spacings. Figure 12 shows that large errors may be introduced when the stand-off varies. If the stand-off gets very large or very small, the percent reflectance drops uniformly across the visible spectrum. As can also be seen, the optimum stand-off distance for this measurement head is actually -0.6mm from nominal. At this spacing, maximum light is received by the sensor and sensitivity to stand-off variation is reduced to a minimum.

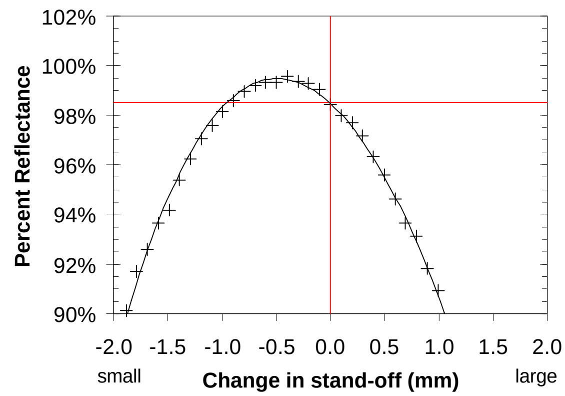

Figure 12: Effect of measurement head stand-off on color measurement. Sample measured is a white reference target. Reflectance data is averaged over the 380-780 nm range.

Conclusions

Measurement accuracy is an important aspect of any system for color measurement. Calibration procedures detailed in this study can significantly reduce color measurement error. A detailed description of how to set up a CCD-based spectrophotometer is given, including calibration information that is often difficult to find in the literature.

1) For frequency calibration, a linear relationship between CCD index and wavelength is desirable. The effect of frequency calibration on *DE** depends on the color measured.
2) The gain of the spectrophotometer must be set by illumination intensity and integration time to avoid saturation, a non-linear effect that introduces significant uncertainty in ΔE.
3) Dark reference measurements for frequency calibration can be made in several ways, but the use of the reference pixel method is shown to have some advantages. Reference pixels were shown to

be a good means for monitoring temperature changes in the spectrometer.

4) Bulb warm-up in a system using a tungsten-halogen light source is a highly significant source of variation in spectral power distribution, which was shown to change by as much as 3.4% during the warm-up process.
5) Averaging was shown to be a good means of reducing error, but reached diminishing returns after about 16 averages.
6) A 1.5º C change in spectrometer temperature was shown to cause a *DE** of 1.0, a very dramatic change.
7) Measurement head stand-off was shown to have an optimum setting that minimizes sensitivity to variations.

References

1. R.W.G. Hunt, The Reproduction of Colour, Fifth Edition, Fountain Press, (1995)
2. Fred W. Billmeyer, Jr. and Max Saltzman, Principles of Color Technology, Second Edition, John Wiley & Sons, (1981)
3. Fundamentals of Color and Appearance, Published by Macbeth, 800-MACBETH (USA), no date
4. Precise Color Communication, Published by Minolta, (1994)
5. CIE, International Commission on Illumination, Bureau Central de la CIE, Paris.
6. Danny C. Rich, "Critical Parameters in the Measurement of the Color of Non-Impact Printing," *IS&T Eight International Congress on Advances in Non-Impact Printing Technologies* (1992)
7. Günter Wyszecki and W.S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd Edition, John Wiley & Sons. (1982)
8. Toshio Konaka, et al, "Study of Colorimetric Errors with Different Printing Samples," *IS&T Eight International Congress on Advances in Non-Impact Printing Technologies* (1992)
9. J.Ingraham, P.Engledrum, J.Laskin, "Evaluating Colormetric Repeatability of Continuous Inkjet Images," *Color Hard Copy and Graphics Arts*, SPIE, **1670**, pp.217-223 (1992)
10. Andrea Raggi and Giancarlo Barbiroli, "Colour-Difference Measurement: The Sensitivity of Various Instruments Compared," *COLOR Research and Application*, volume 18, Number 1, February 1993.
11. Labsphere, North Sutton, NH

Biography

Dr. John C. Briggs joined QEA in January 1998. He is responsible for new product development. Previously at Iomega Corporation, he was a key contributor to the design and development of the Zip™ drive. Dr. Briggs holds two patents and has several patents pending. Between 1986 and 1991, he received his BS, MS, and Ph.D degrees in Mechanical Engineering from the Massachusetts Institute of Technology. His research focused primarily on non-destructive testing and acoustic emission measurements.

Contact at jcb@qea.com or briggsjc@ix.netcom.com