Quantification and Correction of Iris Color

Shaohua Fan[1], Charles R. Dyer [1], Larry Hubbard[2]
[1] Department of Computer Science, University of Wisconsin-Madison,
1210 West Dayton Street, Madison, WI 53706, USA
{shaohua, dyer}@cs.wisc.edu
[2] Fundus Photograph Reading Center, Department of Ophthalmology & Visual Sciences,
University of Wisconsin-Madison, 406 Science Drive, Madison, WI 53711, USA
hubbard@rc.ophth.wisc.edu

Technical Report 1495

December 2003

Abstract

A system has been developed that automatically extracts the iris region from photographs, computes the iris color in CIE $u'v'$ diagram color space, and corrects the color based on a standard calibration target. This system has advantages over previous manual methods in that it (1) is much less time-consuming because it is fully automatic, (2) corrects for the variability inherent in a film-based photographic process, (3) allows for greater objectivity and reproducibility, and (4) uses a mathematical foundation that enables quantification and statistical analysis. Preliminary experimental results show that color correction significantly improves grading accuracy.

We would like to thank Dennis Hafford and Jane Armstrong for providing assistance and the iris images. The support of the National Eye Institute of the National Institutes of Health under Grant No. N01-EY-0-2130 is gratefully acknowledged.

1 Introduction

The relationship between iris color and some eye diseases has been studied [Moss *et al*., 1987]. It has also been reported that some ocular medications may cause changes in iris color. So, iris color quantification is useful for epidemiological description of the distribution of iris color in selected populations and for clinical detection of iris color changes in individuals over time.

Traditionally, the measurement of iris color has relied on manual comparison of iris photographs with a series of iris standard photographs [Seddon *et al*., 1990; AREDES]. While useful, the manual method has several disadvantages: (1) Being a subjective process, the manual method shows large variability among graders (different people have different spectral sensitivity) and by the same grader over time. (2) Manual methods are time-consuming. (3) As commonly utilized, the traditional method has limited capacity to account for the variability inherent in taking and developing photographs, such as exposure time, development time, and variability in the film emulsion and the processing chemistry. (4) Manual grading only gives qualitative results that are difficult to use for iris color change detection and statistical analysis. (5) Although a human can discriminate subtle differences in relative color, there is limited ability to classify color into absolute categories.

Recently, there have been some attempts to measure iris color quantitatively and automatically. German *et al*. [1998] studied the response of an ocular drug based on the color in Red-Green-Blue (RGB) color space. While RGB is a quantitative color model, it is not uniform in that the perceptual difference between colors is not proportional to distance in color space. Takamoto *et al*. [2001] developed a method that automatically computed changes in iris features (area and intensity) under different exposures. However, no measurement of iris color was performed. Melgosa *et al*. [2000] used a CIE LAB color model to quantify iris color. Using a set of 72 synthetic iris samples as standards, 40 real iris images and 25 visual prostheses were classified into one of 72 categories based on the nearest standard in CIE LAB color space. No iris color correction was reported in their work.

The goal of this work is to develop a method to objectively describe iris color, color-correct iris images obtained under different conditions, and automatically classify iris images with high accuracy.

2 Materials and Photography Protocol

The iris photographs from living subjects, one eye each, were selected because they constituted a range of iris color – blue, hazel and brown. These subjects

were of particular interest because photographs of their eyes provide the standards currently used in the Wisconsin Classification, adopted by two large studies, the Beaver Dam Eye Study and the Age-Related Eye Disease Study [Klein *et al.*, 1998], to categorize iris pigmentation. For each of the three standard iris images, 18 copies were taken with different color filters so that the images were color-skewed compared with the images under normal conditions. These photographs were used for two purposes: to see how the measured colors of iris standards were distributed across the scale, and to test whether our color-correction technique could adequately correct for deliberate variations in the photographic process.

The living human subjects were photographed using a protocol developed for taking fundus reflex photographs in epidemiological and clinical trials such as the Age-Related Eye Study. To test the calibration procedure, photographs of the same human subjects were considered to be test subjects and copied on a Sickles Chromapro slide duplicating machine. This system contains three color wheels in the illumination path, allowing for deliberate and controlled introduction of different color biases. Slides were photographed with no color bias, with yellow bias, magenta bias, and cyan bias. At each setting, a standard calibration target (Kodak rectangle 6-step gray scale) was also photographed to provide a reference for correcting iris color during subsequent image processing. The photographs were digitized using a Nikon LS-3510AF film scanner.

3 Methods and Algorithm

3.1 Measurement of Iris Pigmentation

It is important to select a good color model to represent colors in the application problem because different color spaces convey different information about color [Wyszecki, 1982]. For iris color analysis an ideal color space should have the following properties: 1) the color space is objective and an accurate representation of color. 2) distance in color space is uniform, both with respect to perceptual difference and with respect to amount of iris pigmentation. 3) color comparison is invariant to luminance (lightness) differences.

RGB color space does not meet the above criteria because it is not a uniform color space and, furthermore, some colors have negative coefficients. The Commission International de l´Eclairage (CIE) XYZ color system does not have negative coefficients and is defined relative to RGB by a linear transformation [Fairchild, 2001]. The *X, Y* and *Z*, when normalized, become *x, y,* and *z*. The variables *x,y,z* are chromaticity values, which depend on the hue and saturation of the color only. If two colors differ only in luminance, they have the same *x,y,z* values. Because the sum of the *x,y,z* components is 1, only two of the three values need to be explicitly expressed. Normally, the *x* and *y* chromaticity values are

used, defining the CIE *xy* color space. While widely used, *x,y* chromaticity has the disadvantage of being non-uniform. To overcome this weakness, CIE defined a uniform chromaticity scale diagram (also known as the CIE 1976 UCS diagram) called the *u´,v´* diagram [Hunt, 1998]. We chose the CIE *u´v´* color space to describe iris color because it is a uniform color space and is independent of luminance. Using this color space, the two color components can be plotted on a 2D plane, providing computational and graphical display advantages over 3D color spaces.

3.2 Iris Segmentation using the EM Algorithm

Since iris photographs contain more than just the iris, other parts such as the pupil and sclera had to be separated from the iris "region of interest" (ROI). This segmentation problem requires partitioning the pixels in the image into iris and non-iris clusters based on their color components and locations.

For each pixel in the image, if we associate with it a vector x specifying its color, we can form a feature vector space V. Assume the image can be divided into m segments based on a similarity metric defined in the feature vector space V. The distribution of the feature vectors in each segment can be modeled as a Gaussian distribution. Assume the parameters for i^{th} segment are $\theta_i = (\mu_i, \sigma_i)$, and the probability of a pixel chosen from this segment is π_i when drawing a pixel from the image. Then the probability of a feature vector x being generated from the image is $P(x) = \sum_i p(x \mid \theta_i) \pi_i$. For a given image, combining all the probabilities for the pixels, we can have the data likelihood function $L(x, u) = \Pi p(x,u)$. Thus, the segmentation problem has been formulated to find the unknown parameters θ_i for all segments and assign every pixel in the image into one segment, which will have the maximum data likelihood function value.

Considering this as an incomplete data problem, what is missing is information indicating which cluster each pixel was drawn from. The Expectation-Maximization (EM) algorithm [McLachlan and Krishnan, 1996] can be used to solve problems of this type.

The EM algorithm recursively estimates both the missing data and the unknown model parameters. First, given an initial guess for the missing values, i.e., assignment of each pixel to a cluster, the method computes the maximum likelihood estimate of the model parameters for each cluster from the pixels assigned to the cluster. Second, based on the estimated model parameters for every cluster, the probability of each pixel belonging to that cluster is calculated. Each pixel is then reassigned to the cluster with highest probability.

One disadvantage of the EM algorithm is that it can get "stuck" at a local maximum when searching the parameter space. We get around this problem by carefully selecting initial values of the parameters. We can first detect edge points in the image in order to find the approximate boundaries of the iris, and then fit ellipses to these edge points. The two ellipses form a ring and the area inside the

ring is the ROI for the iris. The parameters (mean and standard deviation) estimated from the pixels in the ROI are good initial values for the EM algorithm.

Since iris color changes over time can be very small, we need to compute iris color as accurately as possible. The edge detection-based method alone is not enough for this purpose for the following reasons: (1) We would like the whole process be automatic, but illumination and exposure variations make it difficult to set all the parameters for edge detection automatically. (2) The boundaries of the iris may not fit the ellipse shape model well. (3) The images contain noise and the ROI may include other features such as crypts.

The algorithm we developed to segment the iris region contains the following steps:

1) Detect edge points and select those that approximately bound the inner and outer contours of the iris.
2) Fit the inner and outer edge point sets with two ellipses, forming an annulus that partitions the image into two sets of pixels, iris and non-iris.
3) Compute parameters for each cluster based on the color vector values of the pixels in the cluster. The mean of each cluster is the centroid of the cluster in feature space.
4) Based on the current cluster parameters, at each pixel compute its distance in feature space to the centroid of each cluster.
5) Reassign each pixel to its closest cluster, taking into account the image coordinates of the pixel.
6) Repeat steps 3-5 until convergence is achieved.

3.3 Color Calibration

In the photographic process there are many factors such as film lot, illumination, exposure, and development that affect a photograph. Consequently, in order to reliably describe iris pigmentation and make images taken at different times comparable, we need to correct the iris images to remove these effects. Color correction is important for two major purposes: iris color classification and iris color change detection.

In this section a method is described for color correction based on a standard calibration target image associated with each iris image. A standard calibration target provided by Kodak was used. This calibration target consists of six wedges from dark to light. Each wedge is approximately uniform with known RGB intensity values. Color changes are assumed to be independent of position in a photograph, so color correction requires estimation of a function $\boldsymbol{y} = f(\boldsymbol{x})$, where $\boldsymbol{x}$ is the color at a pixel in an input image and $\boldsymbol{y}$ is its color-corrected value. To compute f, the calibration target was photographed using different filter settings in order to create a set of known, color-skewed images. In addition, at each color

filter setting, photographs were taken of the three Wisconsin Iris Standards. The color-skewed images of the Standards were used to verify f.

For each red, green, and blue color component, let the intensity values of the six wedges in a color-skewed calibration target image and the non-color-skewed calibration target image be x_i and y_i, respectively, where $i = 1, \ldots, 6$. The color correction function f was modeled as a power function $y = f(x) = ax^b$, and the Lavenberg-Marquardt algorithm was used to compute the least squares error estimate of the parameters a and b [Press *et al.*, 1992]. To speed up the computation, f can be approximated by a look-up table (LUT), where x determines the index into the LUT containing the value $f(x)$.

4. Experimental Results

The three Wisconsin Standard Iris Photographs (blue, hazel, and brown) were taken without a color bias using a (60-0-80) filter (i.e., yellow 60, cyan 0, magenta 80). In addition, 18 color-skewed photographs were taken using various settings of the three color wheels.

4.1 Iris Segmentation

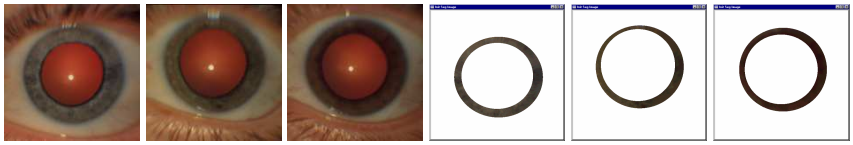


Fig. 1. Iris Standards under Normal Imaging Conditions and their Iris Segmentation. (a) is Standard 1 (blue), (b) is Standard 2 (hazel), (c) is Standard 3 (brown). (d), (e) and (f) are Iris Segmentation for Iris Standards Standard 1, 2 and 3 respectively

To robustly compute iris color, only the middle-third annulus of the iris was used to calculate its mean CIE $u'v'$ value (Fig. 1). Our experiments showed that the width of the ring used was not critical for dilated eyes, which are used in our study. The boundary of the iris was quite reliably detected. However, in case edge positions were mis-detected, for example due to the light reflex at the top of Standard 2, the middle third of the annulus gave a better estimate. This strategy also is suitable for eyes with arcus seniles, which tends to occur first and most at the outer periphery.

4.2 The Three Iris Standards in CIE u´v´ Color Space

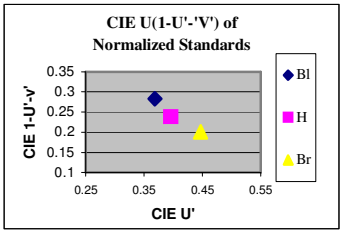


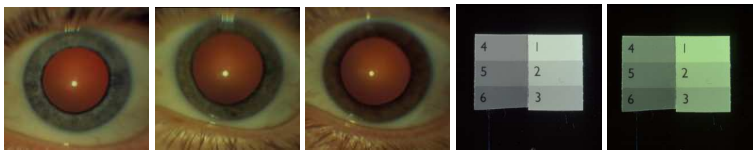
Fig. 2. (*u´,1-u´-v´*) color coordinates for the three Wisconsin Iris Standards in CIE *u´v´* color space. The diamond is for Standard1, the rectangle for Standard2, and the triangle for Standard3

Fig. 2 shows that the three standard irides are well distributed and separated in this color space. The linear relationship of iris color ranging from blue to brown is due to the fact that color is proportional to one physical parameter, density of the pigment melanin. The coordinates shown are (*u´,1-u´-v´*) because this representation makes the relationship between iris colors approximately linear.

4.3 Color Correction Based on Calibration Targets

In order to determine the soundness and accuracy of using calibration targets as the basis for color correction, the three Standard irides and associated calibration targets were photographed using various color filters in order to introduce color biases. Fig.3 shows the Standard irides photographed with a color-skewing filter setting of 90-0-110, which introduces a color bias in yellow and magenta compared with the Standards.

To construct the color-correction function, *f*, first requires deciding on the type of function to use to model the range of iris colors. We evaluated three possible models: linear, power, and exponential functions. Data from the calibration target was used to estimate the model parameters, and each model was tested using the 18 color-skewed Standards. A good model should map the color values of the color-skewed Standards, e.g., (90-0-110), onto the Standards.



(a) (b) (c) (d) (e)

Fig. 3. (a),(b),(c) are Standards under color skewed with a 90-0-110 filter, (d), (e) shows images of the calibration target using the non-color-skewed filter (60-0-80) and the color-skewed filter (90-0-110)

Table 1. Predictions of the blue intensity values for the color-corrected Standards

	Standard 1		Standard 2		Standard 3	
	Pred.	Real	Pred.	Real	Pred.	Real
Linear	50.31	56.38	31.77	40.74	10.62	24.31
Exponential	68.75	56.38	59.74	40.74	50.89	24.31
Power	54.50	56.38	38.14	40.74	20.35	24.31

Table 1 shows the three candidate models computed from the calibration target data, and their predictions of the blue intensity values for the color-corrected Standards. All three models fit the wedge data well, with squared residual values greater than 0.99. However, for the three Standard iris color values, the linear model predicts values that are too low, and the exponential model predicts values that are too high. The reason is that the six wedges in the calibrate target have blue intensities ranging from 69 to 134, while the blue intensity of the three Standards are all less than 50, so predicting the values of the Standards requires extrapolation of the function *f*. The extrapolation error increases with distance from the wedge points. The power function model is the best predictor for the Standards.

4.4 The Iris Colors Before Color Correction and After Color Correction

Ideally, the color correction process should eliminate color-skew effects completely, so that the irides in all color-corrected images for a given Standard have identical CIE u´v´ coordinates. However, because of noise and modeling error, the color-corrected values will be clustered around the true value (Fig. 4). The clusters associated with the three Standards overlap one another significantly, showing that without color correction there is no reliable way to classify these images correctly. After color correction, however, the three clusters are well separated. To further verify this point, we computed the standard deviation of the u´ and (*1-u´-v´*) color coordinates for each cluster. The within-cluster standard deviation for each Standard after color correction is in the range of 1/3 to 1/9 of that before color correction.

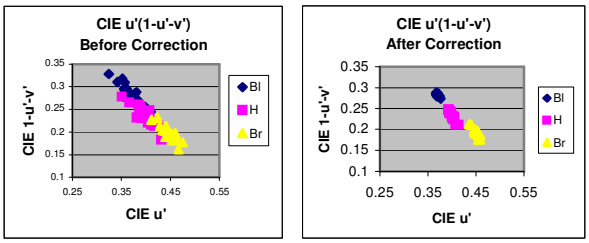


Fig. 4. The distribution of the three Standards in CIE u´v´ color space under 19 different filter settings (18 color-skewed settings plus 1 non-color-skewed setting). The left is before color correction, and the right is after color correction

As a final evaluation of the color-correction function, each color-skewed image was classified as one of the three Standards based on the minimum distance in color space to the Standards. We evaluated the classification results before and after color correction using an error matrix and kappa index [Bishop *et al.*, 1975].

Table 2. Error Matrix for Iris Color Classification. The rows represent the predicted class and the columns represent the correct class

	Before Color Correction			After Color Correction		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
Class1	16	3	0	19	0	0
Class2	2	14	3	0	19	0
Class3	0	1	18	0	0	19

The *kappa index* for the error matrix before color correction is 0.763. After color correction, all the irides were correctly classified, resulting in a kappa index value of 1.0. The *Z*-value in the *Z*-test for whether the difference between the accuracies in the two error matrices is significant is -3.2796. This value is less than -2.3263, which is the z value at 0.01 significance level, so the classification accuracy after color correction is significantly improved than before color correction.

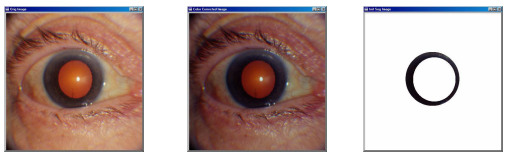


Fig. 5. The left image is the original iris image, the middle one is the color corrected image, and the right one is the iris segmentation.

We also tested our color correction method and segmentation algorithm on 50 iris images other than the Standards. Fig. 5 gives one example showing the color correction and iris segmentation results. Since there are no ground truth iris color categories for those images, we are collecting retinal pigment concentration data for further experiments.

5. Concluding Remarks

A system has been developed that automatically extracts the iris region from photographs, computes iris color in CIE *u´v´* color space, and corrects the color based on a standard calibration target. Using the uniform CIE *u´v´* color space has advantages over the color spaces used in other systems for evaluating iris color because it (3) has the property that distance between two colors is proportional to the amount of pigmentation, (2) separates luminance from color

components, and (3) allows the 2D color components to be conveniently displayed to the user.

Color correction based on a calibration target image was shown to significantly improve iris classification for a set of 18 color-skewed images of the three Wisconsin Iris Standards. Because the color intensity range of the Kodak calibration target is quite different from the range of iris colors, extrapolation of the color correction function from the calibration target values is necessary to predict color-corrected iris values. If a different calibration wedge was designed specifically for this application so that the wedge colors were more similar to real iris colors, the color-correction function could be estimated even more accurately.

In future work we intend to test how well the method works for color correction of more iris images, and how color correction improves the iris color grading. Color correction of the color bias introduced by illumination and exposure differences is also worth exploring. In CIE $u'v'$ color space the three Standards are very nearly linearly related. However, it is not clear how iris color and amount of melanin are quantitatively related. This quantitative relationship is of interest because it could be used to estimate density of melanin in the iris.

References

1. Age-Related Eye Diseases Study Research Group (AREDES). Manual of Operations, Section 15E: AREDS Lens Grading Protocol. The EMMES Corporation, Rockville, MD.
2. Bishop Y, Fienberg S, Holland P. *Discrete Multivariate Analysis – Theory and Practice*, Cambridge, MIT Press, 1975.
3. Fairchild MD. A revision of CIECAM97s for practical applications, *Color Research and Application*, 2001; *26*.
4. German EJ, Hurst MA, Wood D, Gilchrist J. A novel system for the objective classification of iris colour and its correlation with response to 1% Tropicamide. *Ophthal. Physiol. Opt*., 1998; 18:103-110.
5. Hunt RWG, *Measuring Colour*, 3rd ed. Kingston-upon-Thames, Fountain Press, 1998.
6. Klein R., Klein BEK, Jensen, SC and Cruickshanks KJ, The Relationship of Ocular Factors to the Incidence and Progression of Age-Related Maculopathy, *Arch. Ophthalmol*., 1998, 116: 506-513.
7. McLachlan G, Krishnan T. *The EM Algorithm and Extension,*Chichester, John Wiley, 1996.
8. Melgosa M, Rivas MJ, Gomez L, Lita E. Towards a colorimetric characterization of the human iris. *Ophthal. Physiol. Opt*., 2000; 20(3):252-260.
9. Moss SE, Klein R, Meuer MB, Klein BEK. The association of iris color with eye disease in diabetes. *Ophthalmology*, 1987; 94:1226-1231.
10. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge, Cambridge University Press, 1992.
11. Seddon JM, Sahagian CR, Glynn RJ, Sperduto RD, Gragouda ES. Evaluation of an iris color classification system. *Investigative Ophthalmology & Visual Science*, 1990; 31:1592-1598.
12. Takamoto T, Schwartz B, Cantor LB, Hoop JS, Steffens T. Measurement of iris color using computerized image analysis. *Current Eye Research*, 2001; 22(6):412-419.
13. Wyszecki G, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, John Wiley & Sons, Inc., 1982.