A Large Image Database for Color Constancy Research

Florian Ciurea and Brian Funt
Simon Fraser University
Burnaby, British Columbia, Canada

Abstract

We present a study on various statistics relevant to research on color constancy. Many of these analyses could not have been done before simply because a large database for color constancy was not available. Our image database consists of approximately 11,000 images in which the RGB color of the ambient illuminant in each scene is measured. To build such a large database we used a novel set-up consisting of a digital video camera with a neutral gray sphere attached to the camera so that the sphere always appears in the field of view. Using a gray sphere instead of the standard gray card facilitates measurement of the variation in illumination as a function of incident angle. The study focuses on the analysis of the distribution of various illuminants in the natural scenes and the correlation between the rg-chromaticity of colors recorded by the camera and the rg-chromaticity of the ambient illuminant. We also investigate the possibility of improving the performance of the naïve Gray World algorithm by considering a sequence of consecutive frames instead of a single image. The set of images is publicly available and can also be used as a database for testing color constancy algorithms.

Introduction

Several image databases for color constancy research exist.[1-4,7,9-11] While the complex calibration procedures involved in building such databases in which the ambient illumination is properly measured and controlled represents an asset, we are often faced by the immediate limitation of this set-up: the number of images in such a database is typically small. In some cases we would prefer to have a much larger image dataset at the expense of having a less rigorously controlled illumination. For practical color constancy it is often sufficient to measure the camera (R, G, B) of the dominant illuminants present in a scene.

We construct the database using a digital video camera with a neutral gray sphere attached to the camera so that the sphere is always maintained in the field of view. Using a gray sphere instead of the standard gray card facilitates measurement of the variation in illumination as a function of incident angle. The simplicity of this set-up also facilitates recording images in locations where a conventional spectrometer would be impractical. Similarly, we can record a lot of images with relative ease. This large database with the illumination measured separately for every image allows us to study the statistics of illuminants and colors arising in a wide range of common scenes.

The Setup

The database of approximately 11,000 images are of a variety of indoor and outdoor scenes, including many with people in them, shot using a Sony VX-2000 digital video camera. The outdoor scenes were taken in two locations that differ significantly in geography and climate: Vancouver, British Columbia and Scottsdale, Arizona. The images are still frames extracted from video clips captured in progressive scan mode to avoid video interlacing. In progressive scan mode, the camera generates 15 unique frames per second. When extracting the still images from the acquired video clips, we used at most 3 frames from any second of video in order to keep from having almost identical images in the database.

All the camera settings were fixed with the exception of automatic focus and automatic exposure. Although it would be preferable to hold the aperture and shutter speed constant, this is not practical because of the large range of average scene brightness relative to the camera's limited dynamic range. The camera's white balance setting was set to "outdoors" and held that way whether or not the scene was an outdoor scene or an indoor one. With this setup, the color balance of the images taken outdoors is pretty good; those taken indoors generally have a yellow-orange color cast.

The scene illumination was measured using a smooth, small sphere connected to the video camera by means of a monopod leg as shown in Figure 1. The sphere is 4.8 centimeters in diameter and spray painted with CIL Dulux 00NN20/000 matte paint. The paint is a spectrally neutral gray of Munsell Coordinates N4.75/ and has 18% reflectance (Figure 2). The sphere is positioned so that it appears in the video image at all times.

For each image, the scene illuminant is measured in terms of the RGB values of the pixels on the sphere. This means that illuminant is specified in units that depend upon the camera. Although it might be preferable to map these coordinates to a standard coordinate system such as CIE tristimulus values, this is not easily accomplished. Firstly, it would require complete calibration of the camera, which especially for digital cameras cannot be done reliably without knowledge of the camera's internal processing. Secondly, the R, G and B sensitivity functions are probably not perfectly colorimetric. In other words, they likely are not within a linear transformation of the human cone sensitivities, which means that there would be no precise mapping from camera RGB to XYZ in any case. Given these difficulties, we decided to settle for using the camera's native coordinates for all measurements.

Figure 1. Camera with gray sphere attached

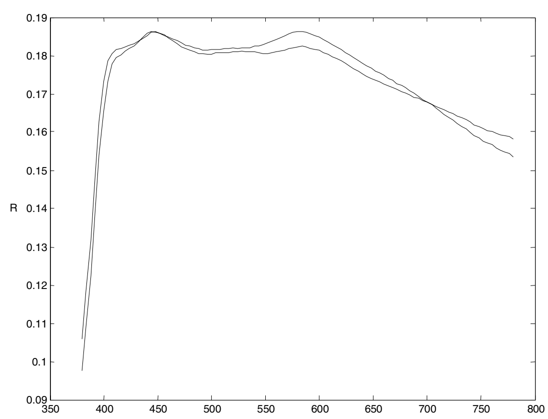

Figure 2. Reflectance spectra of gray sphere (solid line) and Munsell chip N 4.75/ (dotted line)

In order to use this camera setup effectively, we need to make sure that during video capture the illuminant cast on the gray sphere is indeed representative of the illuminant that is lighting the scene generally. In addition to being aware of the illumination conditions during video capture, we developed a post-capture validation process to discard the frames in which this condition is not met. Some video frames we discard are ones in which the dominant illuminant was coming towards the camera lens as opposed to coming from behind. Others, for example, are like those of Figure 3: an outdoor scene in which the gray sphere is temporarily in shade but the main scene is in direct sunlight.

Figure 3. A discarded image: the sphere is in the shadow, while the scene is in direct sunlight

The Image Database

We have recorded approximately 2 hours of digital video from which we have validated more than 11,000 images. The following images are examples of validated frames.

Figure 4. Database image (Vancouver, BC)

Figure 5. Database image (Apache Trail, Arizona)

Figure 6. Database image (Scottsdale, Arizona)

Statistics of Illuminants

In general, the gray ball can be seen to have two main illuminants falling on it. For example, in Figure 5, there is a bright sunlit region and a shadow region. Hence, for each image we computed the rg-chromaticity of the two illuminants that appear on the gray sphere. The first, the dominant illuminant, is the one having the biggest impact on the scene. We determine this illuminant from the brightest region of the gray sphere. The secondary illuminant is present in the form of shading on the gray sphere. For outdoor scenes, the first illuminant is sunlight, while the second illuminant is skylight. For indoor scenes, the coordinates of the first and second illuminants are usually very similar. The following two images show the distribution of the first and second illuminants, respectively, in rg-chromaticity coordinates.

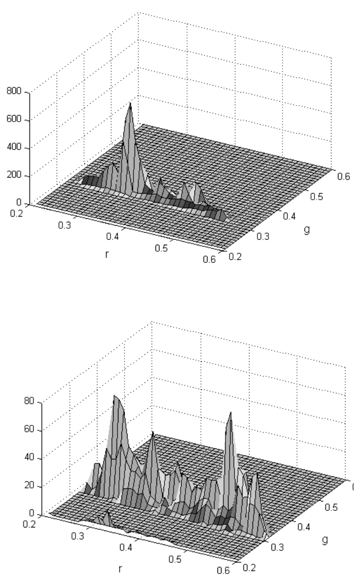

Figure 7. Distribution of dominant illuminant (top) and secondary illuminant (bottom)

Most Discriminatory Colors

We conducted an analysis to find the colors whose presence in a scene reveals the most information about the illuminant. We call these the most discriminatory colors. For this purpose we built a correlation matrix[6] relating the rg-chromaticity of image colors to the dominant scene illuminant. The rg-chromaticity space was divided into 30x30 bins. In this discrete chromaticity space we represent both the colors measured by the camera and the color of the illuminants.

Using Bayes' rule we compute the probability that the chromaticity of the dominant illuminant of a scene was illum given that we observe chromaticity color:

$$P(illum \mid color) = \frac{P(color \mid illum) \cdot P(illum)}{P(color)}$$

All the terms in the right-hand side are computed from the image database. P(color|illum) is the probability of observing the given color under the particular illuminant illum. P(color) is the probability of observing the particular color in any scene and P(illum) is the probability that the scene was observed under illuminant illum. For a given color we then compute the degree of discriminability as the maximum probability of any given illuminant relative to the average probability of all illuminants that correlate with the color:

Discriminability(color) = max(P(illum|color))/avg(P(illum|color)) for all illuminants illum

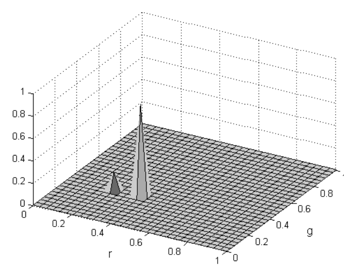

Figure 8. Probability distribution of the illuminants for a discriminatory color

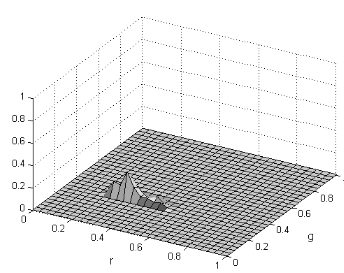

Figure 9. Probability distribution of the illuminants for a less discriminatory color

Figure 8 and Figure 9 show the probability distribution of the illuminants for a color with a high degree of discriminability and for a color with a low degree of discriminability respectively.

The following figure shows the discriminatory colors and the less discriminatory colors in the rg-chromaticity space. It is interesting to see that the colors that are far from white are more discriminatory than the neutrals. This means that when trying to infer the color of the scene illuminant, the colors farther from white are the ones that contain more information about the illuminant.

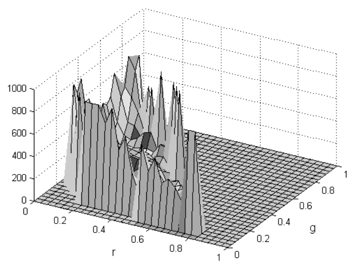

Figure 10. Distribution of discriminatory colors in the rg-chromaticity space

Gray World Over Time

Since the image database is derived from video sequences it means we can study color constancy over time. Clusters of frames that are close temporally in a video sequence may reveal more information about the ambient illumination than a single frame. One possibility is that the gray world method of estimating the illuminant may converge to a better answer if the average is computed over more frames.

From the database, we extracted clusters of sequential frames for which the scene illuminant, as measured from the gray calibration sphere, remains unchanged. The intent in clustering frames in this way is to find automatically clips in which the camera points at different parts of the same basic scene.

Gray World assumes that the average world is gray[5] and estimates the illuminant in the image by computing the departure of the current average RGB in the scene from that of the defined RGB of gray in the world. We simply extend the notion of a scene to a frame cluster and compute the average (R, G, B) over this set of images. Our defined gray in the world is the average (R, G, B) over the whole dataset. To measure how gray world performance might vary with time, we extract frame clusters based on video clips of 1.5, 3, 6 and 15 seconds. We use the following error measure:

$$Error(r_{est}, g_{est}, r_{act}, g_{act}) = ((r_{est} - r_{act})^2 + (g_{est} - g_{act})^2)^{1/2}$$

where (r_{est}, g_{est}), (r_{act}, g_{act}) represent the rg-chromaticity of the estimated and actual (measured) illuminant respectively.

The results are summarized in the following table:

Table 1. Gray World Over Time

	1 frame	*1.5 sec.*	*3 sec.*	*6 sec.*	*15 sec.*
Gray World avg. error	*0.049*	*0.048*	*0.047*	*0.045*	*0.044*
Gray World max. error	*0.331*	*0.333*	*0.309*	*0.309*	*0.309*

Our experience with this error measure indicates that an error of 0.02 - 0.03 in the illuminant estimation is hardly noticeable. The above results confirm that while Gray World performs reasonably well on average; however, it also fails badly in certain cases. The results also show that for the Gray World, little is gained from an extended exposure to the scene. This concurs with previous experiments on adaptation to the scene average.[8]

Conclusion

We have assembled a database of more than 11,000 images in which the illuminants are measured in terms of camera RGB coordinates. This large database allows several studies on the distribution of colors in everyday scenes, the distribution of illuminants colors in everyday scenes and their mutual correlation. The database is publicly available* and can also be used as a color constancy database (with the images un-corrected for different illuminations) or as a properly white-balanced database (with the images properly color corrected based on the illumination data).

Acknowledgements

The authors gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada.

References

* http://www.cs.sfu.ca/~colour/data/

1. Natural Image Database at the Tominaga Laboratory: www.osakac.ac.jp/labs/shoji/English/Database/index_e.html
2. Vision Texture database at MIT Media Lab: www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html
3. K. Barnard, L. Martin, B. Funt, and A. Coath, "A Data Set for Color Research," Color Research and Application, Vol. 27, No. 3, pp. 148-152, 2002 http://www.cs.sfu.ca/~colour/ data/
4. D. Brainard, Hyperspectral Image Data: http://color.psych. upenn.edu/hyperspectral/
5. G. Buchsbaum, "A Spatial Processor Model for Object Color Perception" J. Franklin Institute, Vol. 310, pp. 1-26, 1980
6. G. Finlayson, S. Hordley, P. M. Hubel "Color by Correlation: A Simple, Unifying Framework for Color Constancy", IEEE

Trans. on Pattern Analysis And Machine Intelligence, Vol. 23, No. 11, pp. 1209-1221, 2001

7. E. Marszalec, B. Martinkauppi, M. Soriano, M. Pietikäinen (2000), "A physics-based face database for color research", J. of Electronic Imaging, Vol. 9 No. 1 pp. 32-38, 2000
8. J. J. McCann, "Lessons Learned from Mondrians Applied to Real Images and Color Gamuts", Proc. IS&T/SID Seventh Color Imaging Conference, pp. 1-8, 1999
9. D. Ruderman, Database of hyperspectral images: ftp://ftp.sloan.salk.edu/pub/ruderman/hyperspectral/
10. J. Watlington, Image database: http://web.media.mit.edu/~wad/color/exp1/new/
11. F. Xiao, J. DiCarlo, P. Catrysse and B. Wandell, "High Dynamic Range Imaging of Natural Scenes", Proc. IS&T/SID Tenth Color Imaging Conference, Scottsdale, Arizona, pp. 338-342, 2002