EVALUATION OF COLOR SPACES FOR EDGE CLASSIFICATION IN OUTDOOR SCENES

Erum A. Khan and Erik Reinhard

School of Computer Science, University of Central Florida, Orlando, FL32816-2362
{ekhan,reinhard}@cs.ucf.edu

ABSTRACT

Knowing which edges in an image denote shadow edges and which are due to object boundaries or changes in surface reflectance has important applications in both computer vision and mixed reality. We show that the choice of color space has a significant effect on our ability to differentiate shadow edges from reflectance edges, particularly in sunlit scenes. We have evaluated the performance of 11 color spaces on an input data of more than a hundred colors imaged in a variety of illumination conditions. We quantify the performance of these color spaces using Receiver Operating Characteristic curves, and use the z statistic to find if the difference in performances is statistically significant.

1. INTRODUCTION

Image or video understanding is the process of deducing information about a scene that was captured by an image or a video. This process may use many low or mid-level tasks like segmentation and object tracking, and has important applications such as surveillance, object recognition, and action recognition. Current segmentation and object tracking algorithms do not always produce accurate results due to complicating factors such as the presence of shadows [1]. If shadows were to be removed from the image or video, then these algorithms may perform more reliably.

To successfully remove shadows, they must first be identified. Shadow identification may proceed by considering edges of shadows and distinguishing them from all other types of edges. Edges may be classified as either reflectance edges or luminance edges [2]. Reflectance edges are changes in reflected light caused by changes in the reflectance of two adjacent surfaces. Luminance edges are changes in reflected light caused by different amounts of light falling on a single surface of homogenous reflectance. These edges may be caused by cast shadows, reflected highlights on glossy surfaces, or changes in surface orientation. A subset of luminance edges is formed by cast shadows. We call these edges "shadow edges". If the objects in the scene are predominantly Lambertian, shadow edges will account for most of the luminance edges in the scene. We are therefore interested in distinguishing shadow edges from all other types of edges.

Efforts to classify edges so far rely on simplified scenes to make the problem tractable. Camera calibration may aid edge classification [3]. Other simplifying assumptions may be that objects do not occlude each other [4] or that shadows are cast on flat surfaces [5]. Moving shadows may be detected along with moving objects using the difference between the current frame and a reference frame [6].

Color has also been used to detect shadow edges. The most obvious characteristic of a shadow is that it appears darker than the same surface not in shadow, although shadowed regions still receive light from their surroundings. This non-direction specific light is called "ambient light". The similarity of the spectral composition of ambient light to direct light depends in part on the albedos of objects in the scene, as well as the presence of participating media. A simplifying assumption may be that the average albedo of all surfaces in the scene is a spectrally flat gray [7]. This is called the gray world assumption. With only one type of light source lighting the environment, and no scattering events taking place in participating media such as water and sky, a gray world allows ambient light to have the same spectral composition as the dominant light source in the scene. Thus, the difference between a shadow and non-shadow region will only be apparent in luminance, but not in chromatic content and this feature may be exploited for the purpose of shadow detection [1, 5, 8].

The choice of color space in which images are analyzed may make the task of edge classification easier. We hypothesize that color opponent spaces qualify as the most suitable representations of color for the purpose of edge classification in sunlit scenes. We provide support for our conjecture by means of experimental results which rank a set of 11 color spaces with respect to suitability for shadow edge classification. The remainder of this paper is organized as follows: Section 2 includes our theoretical analysis of the color spaces for the purpose of shadow detection, and this analysis is corroborated with experimental evidence in Section 3. We conclude with a discussion of our results in Section 4.

2. BEHAVIOR OF COLOR ACROSS EDGES

Direct sunlight has a distinctly yellow spectrum, whereas Rayleigh scattering causes the rest of the sky to appear blue [9, 10]. This means that ambient light in outdoor scenes has a blue tinge. We may therefore expect that shadows will have a blue color cast, while lit regions will have a yellow cast. Thus a difference in luminance across a shadow edge in sunlit scenes will be accompanied by a difference in chromatic content, as shown in Figure 1. In indoor scenes however, where it may be assumed that the gray world assumption holds and that no scattering events are taking place, ambient light will have the same chromatic content as direct light and there will be no change in color across shadow edges. The information encoding in the human visual system gives us some insight into how humans treat color: the output of photoreceptors in the eye is recombined to form a new decorrelated color space where the axes consist of a luminance channel and two chromatic channels [2]. One chromatic channel encodes red versus green, and the other encodes yellow versus blue. Research on monkey foraging behavior in Kibale Forest, Uganda confirms that the red-green channel is invariant to shadows [11]. While shadows do not show in the red-green channel, this is generally not the case for the yellow-blue channel. In sunlit scenes some difference will occur in

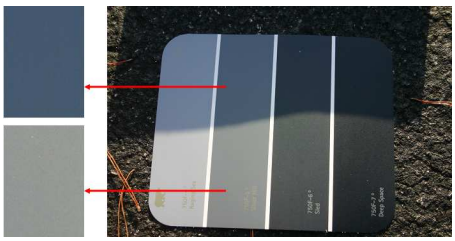


Fig. 1. Colors appear to have a bluer shade in shadows.The patches on the left have been obtained by cropping the image on the right.

the yellow-blue channel across shadow edges since the color of the scene will change from a yellow cast in lit regions to a blue cast in shadows across shadow edges. While color opponent spaces have been used in image processing tasks before [8], the two chromatic channels are invariably used indiscriminately, even though the yellow-blue channel is not invariant to shadows.

We hypothesize that opponent color spaces encode information in a way that is most suitable for edge classification for sunlit scenes due to the invariance of the red-green channel to shadows. In the following sections we test this hypothesis by measuring the suitability of 11 different color spaces in edge classification tasks for sunlit scenes.

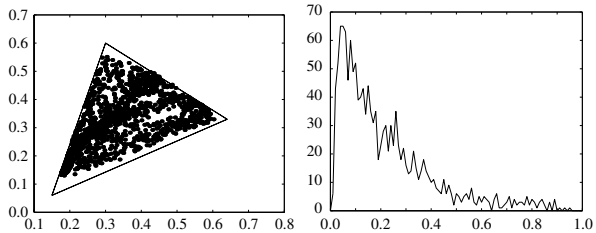
3. QUANTITATIVE EVALUATION

When an image is represented in a particular color space, one channel in the color space might encode the difference across reflectance edges only while another channel might encode the difference across both shadow and reflectance edges. In that case, classification of edges in the image may be done by first converting the image to such a color space, and then using these two channels to distinguish between the two types of edges. The accuracy of the classification depends on how well the first channel discriminates between the two kinds of edges. Hence, in our evaluation of color spaces, we base our rating on two characteristics. The color space should have a discriminating channel as well as a non-discriminating channel. In addition, the discriminating channel should be the best discriminator of edges among the color spaces compared, such that a suitable threshold over the change in values in the channel yields the most accurate classification of edges.

3.1. Data Collection

We have photographed a set of partly shadowed color samples and cropped them to obtain patches, each patch consisting of only a single color, as it appears in shadow or lit regions. Figure 1 shows two patches obtained from a single color as it appears in or out of shadow. We have photographed 143 different colors in a variety of sunlit scenes to ensure that the change across shadow edges is not always the same. Ideally, the set of colors we use should fill the range of representable colors in current display devices. We have converted the set of colors to Yxy color space and plotted the *CIE xy* chromaticity coordinates and luminance values in Figure 2. From this figure, we conclude that our set of patches indeed fills both the range of representable chromaticities (indicated by the triangle) and the range of representable luminance values.

To compute the response of a color space to a shadow or reflectance edge, we convert the two patches representing the edge



(a) (b)

Fig. 2. Subset of colors used in the experiment. 2(a) is the plot of x,y coordinates of the colors and 2(b) is the histogram of their Y values.

into that color space, average the values of each patch to obtain a single color value, and assess how different these values are for the two patches. For each of the shadow and reflectance edges, we compute the difference across the edges for all channels in the following color spaces: *RGB*, *LMS*, *XYZ*, *Yxy*, *Luv*, *HSV*, *normalized RGB*, AC_1C_2, *CIELAB*, and $L\alpha\beta$, of which *CIELAB*, $L\alpha\beta$, *Linear* $L\alpha\beta$, and AC_1C_2 are opponent color spaces [12]. Most of these are standard color spaces, with the exception of $L\alpha\beta$, which is the result of running a Principle Components Analysis on an ensemble of spectral images depicting natural scenes [13]. The result is a color space akin to opponent color spaces. This color space is defined in logarithmic space. For our comparisons we include a version of $L\alpha\beta$ in linear space to allow assessment of the importance of this non-linearity.

We have generated another set of patches from the same color samples to represent shadow and reflectance edges in indoor scenes. We have obtained 349 shadow edges and 349 reflectance edges in both lit and shadow regions separately, all under tungsten light. We have measured the discriminating ability of channels separately for indoor and sunlit scenes and compare the results.

3.2. Analysis

To measure the discriminating ability of all channels in the 11 color spaces, we use Receiver Operating Characteristic (ROC) curves. The area under an ROC curve has been empirically shown to be the equivalent to the Wilcoxon statistic, and computes the probability of correct classification, θ [14]. The standard error σ for this index may be computed as follows [14]:

$$SE(\theta) = \sqrt{\frac{\theta(1-\theta)+(n_F-1)(Q_1-\theta^2)+(n_T-1)(Q_2-\theta^2)}{n_F n_T}}$$

where θ is the probability of correct classification, n_T is the number of true data items, n_F is the number of false data items, and Q_1 and Q_2 may be computed as follows:

$$Q_1 = \theta/(2-\theta)$$
$$Q_2 = 2\theta^2/(1+\theta)$$

Quantitative comparison of 2 ROC curves is achieved by computing the *z* statistic. The value of this statistic indicates whether or not the difference in the discriminating abilities of the two classifiers as shown by the Area under the ROC curves (AUCs) is statistically significant, and it may be computed as follows:

$$z = (\theta_1-\theta_2)/\sqrt{SE^2(\theta_1)+SE^2(\theta_2)}$$

Channel	Area1	$\sigma 1$	Area2	$\sigma 2$
RGB - R	0.435	0.013	0.598	0.018
RGB - G	0.477	0.013	0.573	0.018
RGB - B	0.649	0.012	0.567	0.018
XYZ - X	0.362	0.012	0.553	0.019
XYZ - Y	0.397	0.013	0.547	0.019
XYZ - Z	0.590	0.013	0.521	0.019
LMS - L	0.384	0.013	0.554	0.019
LMS - M	0.412	0.013	0.545	0.019
LMS - S	0.590	0.013	0.521	0.019
Yxy - Y	0.397	0.013	0.547	0.019
Yxy - x	0.795	0.009	0.902	0.009
Yxy - y	0.798	0.009	0.881	0.010
Luv - L	0.397	0.013	0.547	0.019
Luv - u	0.588	0.012	0.643	0.017
Luv - v	0.647	0.012	0.612	0.018
HSV - H	0.798	0.009	0.847	0.011
HSV - S	0.707	0.011	0.760	0.015
HSV - V	0.372	0.013	0.549	0.019
normalized - R	0.796	0.009	0.903	0.009
normalized - G	0.806	0.009	0.891	0.010
normalized - B	0.782	0.010	0.790	0.014
CIELAB - L	0.394	0.013	0.569	0.018
CIELAB - a	0.826	0.008	0.848	0.012
CIELAB - b	0.762	0.010	0.778	0.015
$L\alpha\beta$ *- L*	0.422	0.013	0.532	0.019
$L\alpha\beta$ *-* α	0.798	0.009	0.910	0.009
$L\alpha\beta$ *-* β	0.818	0.009	0.895	0.009
linear $L\alpha\beta$ *- L*	0.408	0.013	0.530	0.019
linear $L\alpha\beta$ *-* α	0.682	0.012	0.626	0.018
linear $L\alpha\beta$ *-* β	0.635	0.012	0.613	0.018
AC_1C_2 *- A*	0.388	0.013	0.549	0.019
AC_1C_2 *-* C_1	0.627	0.012	0.653	0.017
AC_1C_2 *-* C_2	0.679	0.012	0.601	0.018

Table 1. Channels, their ROC curve areas and Standard Error for Sunlit Scenes (Area1 and $\sigma 1$), and indoor scenes (Area2 and $\sigma 2$)

If the probability of falsely rejecting the null hypothesis is 0.05 and that of falsely accepting it is 0.1, the value of *z* above 1.96 indicates that the discriminating abilities are indeed significant.

For our experiment, ROC curves are created by thresholding each channel's response to sample reflectance and shadow edges, and then plotting the true positive fraction against the false positive fraction, (positive/negative are the classifications and true/false specify whether the classification is correct). The AUC for each channel measures the probability of correct edge classification by that channel. A suitable color space should have one channel that has an AUC close to 0.5, and another channel with an AUC close to 1. The most suitable color space is one for which the discriminating channel's AUC is closest to 1. The ROC curves computed for each channel for sunlit scenes are shown in Figure 3 and the AUCs for each channel for both sunlit and indoor scenes are shown in Table 1 along with their standard error. We note that all color spaces except *normalized RGB* color space have at least one discriminating and one non discriminating channel, and thus provide some discrimination between the two types of edges. For all these color spaces, we select the channel with the highest AUC. We select the four color spaces with the highest AUCs and compare them to all

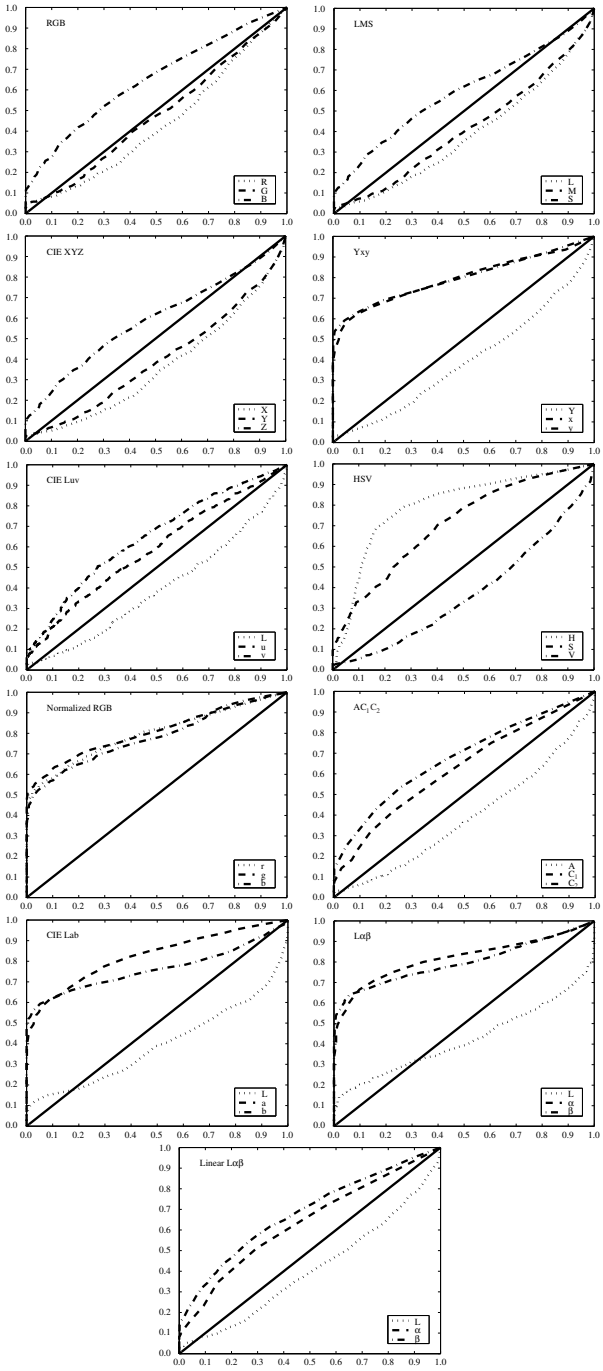


Fig. 3. ROC curves for the 11 color spaces provide qualitative assessment of the color spaces.

other color spaces using z values for sunlit scenes in Table 2 and for indoor scenes in Table 3.

For outdoor scenes, *CIELAB*'s *a* channel as well as $L\alpha\beta$'s β channel provided the highest AUC. These two channels were not statistically different, since their *z*-score remains below 1.96. The *y* channel of the *Yxy* color space and *H* channel of *HSV* have the second highest AUC value after *CIELAB* and $L\alpha\beta$. They were

Channel	*CIELAB - a*	$L\alpha\beta$ - β	*HSV - H*	*Yxy - y*
RGB - B	11.868	11.274	9.725	9.757
XYZ - Z	15.328	14.729	13.177	13.209
LMS - S	15.328	14.729	13.177	13.209
Yxy - y	2.140	1.532	0.032	-
Luv - v	12.011	11.414	9.869	9.901
HSV - H	2.173	1.565	-	0.032
$L\alpha\beta$ - β	0.609	-	1.565	1.532
CIELAB - a	-	0.609	2.173	2.141
Linear $L\alpha\beta$ - α	9.889	9.290	7.744	7.776
AC_1C_2 - C_2	10.069	9.471	7.925	7.957

Table 2. *z* statistic when *CIELAB - a*, $L\alpha\beta$ - β, *HSV - H* and *Yxy - y* are compared with other discriminating channels.

Channel	$L\alpha\beta$ - β	*Yxy - x*	*HSV - H*	*CIELAB - a*
RGB - R	15.565	15.026	11.573	11.674
XYZ - X	17.417	16.876	13.406	13.510
LMS - L	17.376	16.836	13.366	13.468
Yxy - x	0.639	-	3.754	3.651
Luv - u	13.738	13.195	9.732	9.833
HSV - H	4.374	3.754	-	0.105
$L\alpha\beta$ - β	-	0.639	4.374	4.271
CIELAB - a	4.3	651.0	105.271	-
Linear $L\alpha\beta$ - β	14.454	14.421	10.968	10.558
AC_1C_2 - C_1	13.344	12.799	9.330	9.431

Table 3. *z* statistic when $L\alpha\beta$ - β, *Yxy - x*, *HSV - H*, and *CIELAB - a* are compared with other discriminating channels.

both statistically different from *CIELAB*, but there was no statistical difference between the two. For indoor scenes the two highest ranked color channels are $L\alpha\beta$'s β channel and the *x* channel of the *Yxy* color space and are not statistically different. The third and fourth color spaces are also statistically the same, but statistically different from the first two.

Four of the color spaces we consider show color opponency: *CIELAB*, $L\alpha\beta$, linear $L\alpha\beta$ and AC_1C_2. While *CIELAB* is a perceptually uniform space, and $L\alpha\beta$ is perceptually uniform to a first approximation, the latter two are linear color spaces. The performance of the perceptually uniform color spaces is significantly better than their linear counterparts. To verify the performance of the red-green channels versus the yellow blue channels in the perceptually uniform color opponent spaces, we have computed *z* statistics between red-green channels and the blue-yellow channels of both these spaces. We expect a difference between these pairs of channels and find this to be the case. The *a* channel in *CIELAB* performs significantly better than the b channel, as the *z* statistic for this comparison is 4.733. In $L\alpha\beta$, the α channel performs better than the β channel, but the *z* statistic is 1.515 and so this difference is not significant.

4. DISCUSSION

For the purpose of edge classification we have evaluated several color spaces using ROC curves. For outdoor scenes, our results support our hypothesis that color opponent spaces perform well. The two perceptually uniform color opponent spaces prove to be the most suitable choices. For indoor scenes, other color spaces tend to perform as well as color opponent spaces. For applications that would process both indoor and outdoor scenes, the $L\alpha\beta$ color space would be a suitable choice. In addition, by transforming the data to log space the pixel data along each axis becomes symmetrical. This symmetry, along with the color opponency appears to aid in the separation of shadow and reflectance edges.

In conclusion, we have shown that color is an important cue in the classification of edges for sunlit scenes even when we do not make assumptions about the gray world and the effects of participating media. Through experimental results, we have validated the conjecture that opponent color spaces are a good choice for the purpose of edge classification.

5. REFERENCES

[1] Rita Cucchiara, Costantino Grana, Massimo Piccardi, Andrea Prati, and Stefano Sirotti, "Improving shadow suppression in moving object detection with hsv color information," *IEEE Intelligent Transportation Systems*, pp. 334–339, 2001.

[2] Stephen E Palmer, *Vision Science: Photons to Phenomenology*, The MIT Press, Cambridge, Massachusetts, 1999.

[3] Mark S Drew and Graham D Finlayson, "Recovery of chromaticity image free from shadows via illumination invariance," in *IEEE Workshop on Color and Photometric Methods in Computer Vision, ICCV*, 2003, pp. 32–39.

[4] Raphael Charit and Murray H Loew, "Complex shadow boundary segmentation using the entry-exit method," in *Computer Vision and Pattern Recognition*, 1988.

[5] Elena Salvador, Andrea Cavallaro, and Touradj Ebrahimi, "Shadow identification and classification using invariant color models," in *International Conference on Acoustics, Speech, and Signal Processing*, 2001.

[6] A Prati, I Mikic, R Cucchiara, and M Trivedi, "Comparative evaluation of moving shadow detection algorithms," in *IEEE CVPR workshop on Empirical Evaluation Methods in Computer Vision*, Kauai, 2001.

[7] Anya Hurlbert, "Formal connections between lightness algorithms," *Journal of the Optical Society of America A*, vol. 3, pp. 1684–1693, 1986.

[8] Theo Gevers and Harro Stokman, "Classifying color edges in video into shadow-geometry, highlight, or material transitions," *IEEE Transactions on Multimedia*, vol. 5, 2003.

[9] M Minnaert, *Light and Color in the Open Air*, Dover, 1954.

[10] D K Lynch and W Livingston, *Color and Light in Nature*, Cambridge University Press, 1995.

[11] Tom Troscianko, Roland Baddeley, Carlos A Párraga, Ute Leonards, and Jolyon Troscianko, "Visual encoding of green leaves in primate vision," *Journal of Vision*, vol. 3, 2003.

[12] G Wyszecki and W S Stiles, *Color science: Concepts and methods, quantitative data and formulae*, John Wiley and Sons, New York, 2nd edition, 1982.

[13] D L Ruderman, T W Cronin, and C-C Chiao, "Statistics of cone responses to natural images: Implications for visual coding," *J. Opt. Soc. Am. A*, vol. 15, no. 8, 1998.

[14] J A Hanley and B J McNeil, "The meaning and use of the area under the receiver operating characteristic (roc) curve," in *Radiology*, 1982, vol. 143, pp. 29–36.