

Shadow Identification using Colour Ratios

Kobus Barnard[1] and Graham Finlayson[2]
[1]Computer Science Division, U. C. Berkeley, Berkeley, CA, 94720
Email: kobus@cs.berkeley.edu
[2]Department of Computer Science, University of East Anglia
Email: graham@sys.uea.ac.uk

Abstract

In this paper we present a comprehensive method for identifying probable shadow regions in an image. Doing so is relevant to computer vision, colour constancy, and image reproduction, specifically dynamic range compression. Our method begins with a segmentation of the image into regions of the same colour. Then the edges between the regions are analyzed with respect to the possibility that each is due to an illumination change as opposed to a material boundary. We then integrate the edge information to produce an estimate of the illumination field.

Introduction

In this paper we present a comprehensive method for identifying probable shadow regions in an image. Shadow identification is an important problem which receives relatively little attention. Shadow identification is relevant to computer vision, colour constancy, and image reproduction. In the case of computer vision, a shadow boundary often has the same effect as a segmentation error. Such ambiguities can increase the amount of searching among the possible collections of segments used to form objects. In the case of colour constancy, illumination which varies spatially over a scene presents major problems for most colour constancy algorithms. Furthermore, it has been shown that illumination change can be very useful for colour constancy.[1,2] but the algorithms so far require the illumination change to be identified. This is done in [2] in the case of smooth illumination, but fails if there is a hard shadow boundary. The work in this paper is a first step in the integration of abrupt illumination change and colour constancy.

Moving onto image reproduction, an important application of this work is dynamic range compression. As is well known, the dynamic range of natural scenes is much larger than that which can be reproduced. Thus, there is much interest in compressing image dynamic range for more appropriate reproduction,[3-6] especially as the ability to record the range of such scenes is improving. The major contributor to the large dynamic range of scenes is illumination variation. Material changes rarely have a ratio larger than 30 to 1, whereas the differences between bright sunlight and a nearby shadow can easily exceed this.[3] Despite this observation, current methods do not model shadows. Rather, pixels are modified based on how they compare to simple statistics of surrounding neighbourhoods, often combining the results of a number of scales or instances. Such methods cannot distinguish between a dark shadow and a dark region with the same camera response, even if it can be argued that there is ample evidence suggesting one case or the other. Thus we posit that for dynamic range compression it is helpful, and perhaps necessary, to take a physics based approach. We will now outline such a method.

We begin with a conservative segmentation of the image. As in Barnard et al.[2] since we are mostly concerned with analyzing the illumination, it does not matter if areas due to surfaces with very similar colour (such as a white piece of paper on a white wall) are combined. These considerations mean that we can normally obtain a sufficiently accurate segmentation for the task at hand.

Given a segmentation, we then apply a number of tests to pairs of neighboring segments. Some of these tests (in fact, the most interesting ones) include consideration of other segment pairs. Each test leads to a different degree of belief about the boundary between the two segments being a shadow, and each test therefore has an associated score. The final score for a pair of neighboring segments is simply the maximal score found among all tests. From these edge scores we then compute the relative illumination field for the image, which can be used for colour constancy, computer vision applications, and dynamic range compression. We now provide additional details.

Determining Crisp Shadow Edges

Our method deals both with crisp shadows and soft shading. We first develop the theory for crisp shadows, and then introduce modifications to deal with, and take advantage of, the observation that shadow boundaries are often gradual. To estimate the plausibility that an edge is a shadow edge we use a number of tests, each of which has a score associated with it, which roughly corresponds to the amount of evidence that passing the test represents. The actual amounts are set somewhat arbitrarily. We are currently working towards a more principled scoring, but we note that preliminary results indicate that the exact numbers are not that important. We remind the reader that the final score for the boundary is the maximal score found among all tests. The tests will be described in rough order of strength.

The weakest test is passed if the change in camera response (RGB) over the boundary is simply consistent with a possible illuminant change. This, and many of the conditions which follow, depend somewhat on the accuracy of the diagonal model of illumination change,*[7,8] which holds fairly well for our camera.[9,10] Following Finlayson's work on colour constancy,[11] we consider the set of possible illuminants to be restricted to common indoor and outdoor illuminations. Doing so constrains the chromaticities of the expected illuminants, but we consider their magnitudes to be unconstrained. Thus the RGB of the possible illuminants form a cone in RGB space. We construct this cone from a set of roughly 100 measurements of sources, indoor illuminations, and outdoor illumination.[10] Our illuminant set is available on-line.[12]

Given the diagonal model and a set of illuminants, we can then determine if the RGB ratio over a boundary can be due to an illumination change. If it is, then one side of the boundary (the shadow side) must be illuminated by a shadow illumination (R_S, G_S, B_S), and the other side must also be illuminated by the shadow illumination (R_S, G_S, B_S) together with some additional non-shadow illumination (R_N, G_N, B_N). Thus the RGB ratio from the shadow side (the darker side) to the non-shadow side is given by:

$$\left(\frac{R_S}{R_S + R_N}, \frac{G_S}{G_S + G_N}, \frac{B_S}{B_S + B_N}\right) \tag{1}$$

where both (R_S, G_S, B_S) and (R_N, G_N, B_N) must be in the 3D cone described above. Thus we can pre-compute the possible

$$\left(\frac{R_S}{R_S + R_N}, \frac{G_S}{G_S + G_N}, \frac{B_S}{B_S + B_N}\right)$$

based on the cones, smooth the results, and enter them into a discretization of the unit cube for fast access. We note that the fact that a shadow edge must have a strict decrease in each of three channels excludes some regions which appear to the be plausible shadows due to being darker overall. Specifying the nature of the decrease, as done above, excludes additional candidates.

* Consider a white patch under two different illuminants. Suppose that under the first illuminant the color is [r,g,b] and under the second illuminant the color is [r',g',b']. It is thus possible to map the color of white under the first illuminant to the color under the second by post-multiplication by a diagonal matrix: [r',g',b'] = [r,g,b] diag(r'/r, g'/g, b'/b). If the same diagonal matrix transforms the RGB of all surfaces (not just the white ones) to a good approximation, then we say that we have a diagonal model of illumination change.

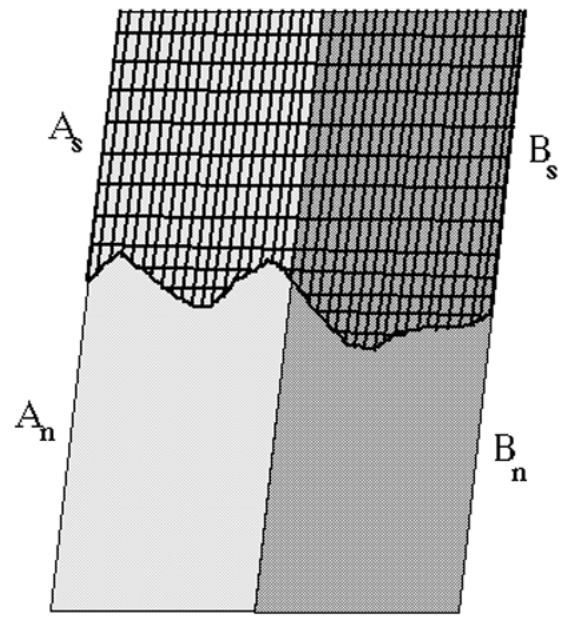


Figure 1: Strong evidence for a shadow: First, the change across the shadow boundary for both regions A and B is consistent with a valid shadow boundary. Second, the changes across the adjacent shadow boundaries are similar. Third, the reflectance A and B are quite different.

Ratios which are valid illumination changes, as determined above, can usually also be due to a material boundary, and thus the above is only moderate evidence for a shadow boundary. However, shadows commonly fall over more than one region in an image, and exploiting this observation is the key to developing more powerful tests. For example, if the same kind of shadow falls on two different surface reflectances (not necessarily connected), then we expect to see two similar shadow ratios with different RGB in the shadow regions and different RGB in the non-shadow regions. The occurrence of two such parallel but distinct possible shadow jumps are less likely to be due to material changes than a single shadow jump, and thus we score this test higher. Similarly, if we observe three parallel shadow jumps, all with different shadow RGB (and non-shadow RGB), the score is higher yet (in this work we stop at three). We stress that for these parallel jump tests, the proposed illumination ratios need to be similar, and the proposed surface reflectances, must all be different. For example, the same kind of proposed illumination ratio observed on both a red surface (splitting it into light red and dark red regions) and a green surface (splitting it into light green and dark green) is taken as better evidence than either alone.

We feel that when such shadows jumps are adjacent, the evidence for a shadow is even stronger. Here a shadow is proposed to cross over a surface boundary such that four segments are formed, two in shadow and two not in shadow (see Figure 1). Even more evidence, is associated when an additional similar shadow jump (but with yet a different shadow RGB) occurs elsewhere in the image. Finally, when that third shadow boundary is adjacent to the other two, the score is even higher. In this case, the shadow boundary crosses three adjacent regions, to form 6 patches.

We include one more test in our method. If the shadow jump is so large that there is no material boundary which can account for it, we also take this as very strong evidence as a shadow boundary. We currently use a ratio of 30 to 1 as the upper limit of albedo change.

Incorporating Soft Shadow Edges and Gradual Shading

So far we have been treating shadow boundaries as crisp, and thus the illumination change across the boundary is well determined. However, it is well known that shadow boundaries are often quite soft, and, in fact, this observation has been used to classify shadow and non-shadow edges.[13] Certain segmentation approaches, such as region growing based on small changes between neighboring pixels, can sometimes properly ignore the shadow boundary, and grow a region which corresponds to physical edges only. However, it is very hard to make this work in a consistent fashion when there is a mix of shadow boundary strengths. Normally any reasonable choice of thresholds leads to a segmentation which stops at some shadow boundaries, but grows into others, and, as the number of shadow boundaries which are missed increases, so do the number of other edges which are missed. Therefore, we take quite a different approach. In general, the philosophy is to find as many of the putative shadow boundaries as possible, and then use the higher level reasoning to determine which ones are shadows. It should be clear that the tests developed above will fare better when fewer shadow boundaries are absorbed. Thus we use a segmentation method which ensures that the RGB's within each region† are within a certain small range of each other, which errs on the side of having too many segments. For example, a nicely shaded ball may have a number of stripes, each one representing the next increase in illumination with respect to our allowed RGB range.

Given this sort of segmentation, applying the tests as described above may run into trouble, as the shadow boundaries may be divided into several steps at somewhat arbitrary points. The key to using such boundaries is to note that the steps have a well defined structure, discussed further below, which can be identified. Then a candidate shadow boundary is declared as two regions separated by zero or more steps. We ensure that the maximal number of steps are used, and therefore the steps themselves are not considered as shadows until later. Interestingly, the steps are characteristic of illumination change, and thus strengthen the case for the proposed shadow boundary. Therefore, in the case that the plausible shadow has these steps, the score is increased in proportion to the number of steps. Thus this method incorporates the use of the softness of shadows without relying on them. Further, it applies to quite large steps, such as a classic penumbra region, or big stripes on a lightly shaded ball.

We now provide further details on the identification of the shadow steps. The first criteria is location. In general, we want to consider regions which are between other regions in the "right way". A clean criteria that covers the desirable cases exactly is difficult to determine. However, the following heuristic seems to work well. Consider a region adjacent to two different regions A and B, and connected to them via the (possibly broken) boundaries, b_A and b_B. We test whether the center of mass of the region is in the convex hull of b_A and b_B.

We move onto the colour constraints on the step regions. We assume that the steps are due to varying illumination, specifically a blend of the illumination at the two regions surrounding them. This means, first, that the sequence of regions are each a plausible illumination change apart, and second, that the RGB of each step region is (approximately) a convex combination of the RGB's of the two regions surrounding it. So far, the combination of the location and colour tests has proven to be quite robust.

A few remarks with regard to the integration of the shadow step analysis with the tests explained above are in order. First, the tests are applied using illumination changes between the two outside regions. The shadow steps are not used for this. However, as mentioned above, if the two regions being considered are separated by shadow steps, then the score will increase. Finally, the shading boundaries between the steps inherit the score from the outside regions.

Determining the Illumination Field.

For some applications, such as segmentation into regions of similar surface reflectance, the shadow information computed above may be sufficient. However, for many applications, we need to integrate the information to determine the overall illumination field. Our method for doing so is as follows: We simplify matters by solving for the logarithm of the illumination at each segment. We solve for the illumination field of each channel separately. We set up a number of equations which are solved in the least squares sense. To begin, we note that our method essentially computes relative illumination, and thus we specify the unknown factor by setting the sum of the logs of the illuminations to zero. Also, since our method is based on the relative changes across adjacent regions, we need to deal with the possibility that the graph is not connected. Thus we introduce an equation setting the log of each illuminant component to zero. Since we solve the equations in the least squares sense, we arrange that these equations have little effect, except when there is ambiguity, by giving them a relatively small weight. The system of equations described so far will ensure that the

†Technically, for the results presented in this work, we segmented using (r=R/(R+G+B), g=G/(R+G+B), L=R+G+B), not RGB.

recovered relative illumination field is uniformly one in the absence of any other information.

Each edge between regions is a source of information. On the assumption that the ratio of the red pixel values for adjacent regions i and j are due to the illumination, we have:

$$\log(e_i^R) - \log(e_j^R) = \log(\frac{R_i}{R_j}) \qquad (2)$$

where e_i^R is the relative value of the red component of the illumination at region i, and e_j^R is the analogous quantity for region j. We weight this with the plausibility of the edge, p, and the length of the edge, L:

$$pL\log(e_i^R) - pL\log(e_j^R) = pL\log(\frac{R_i}{R_j}) \qquad (3)$$

Similarly, to the extent we believe the edge is not a shadow edge, the illumination should not change at that boundary:

$$(1-p)L\log(e_i^R) - (1-p)L\log(e_j^R) = 0 \qquad (4)$$

These equations are solved in the least squares sense for each channel. Then the illumination is set to the appropriate value at each region. Since there are always pixels which do not belong to any of the segments, we linearly interpolate the result for those pixels.

Results

We have tested our method on several complex real images with both distinct and subtle shadows, and have obtained very promising results. However, shadow identification on complex real images is a difficult problem, and misclassification does occur. One of the strengths of this method is that the shadow information propagation step serves to resolve ambiguity and mitigate the effect of misclassification, especially when the goal is simply an improved image for human viewing.

Figure 2 shows the results on an image with a strong shadow, and some gentle shading due to surface curvature. Although not apparent in the black and white reproduction, there is a significant colour change across the shadow boundary. The background (shadow) illuminant is quite blue, and the foreground illuminant is quite yellow. Thus this scene emulates a standard outdoor scene where the shadows are illuminated by the blue sky, and the non-shadow regions are illuminated by both the blue sky and direct sunlight. The image was taken with a Sony DXC-930 camera which supports the diagonal model well. Figure 2(a) shows the input image. Figure 2(b) shows the candidate shadow edges with line segments joining the two regions sharing the edge. The plausibility of the edges is represented by the brightness of the segments. White segments join regions where there is strong evidence that one of them is a shadow, or more specifically, that the edge between them is an illumination edge. If the process were perfect, then adjacent regions from the same surface would be joined with white segments. Figure 2(c) shows the estimated illumination field, which is very close to the actual illumination field. Since the method rarely completely commits to a specific edge being due to an illumination change (i.e. "p" in (2) is rarely 1), the least squares fitting procedure will always produce a few artifacts of the material edges. We feel this is a reasonable tradeoff for the gain in robustness. Figure 2(d) shows the spatial variation of the illumination removed by dividing by the illumination field. Applications seeking to enhance dynamic range would normally provide an image somewhere between (a) and (d), possibly with additional transforms. (Removing all shading from an image is usually detrimental). Finally we note that Figure 2(d) has artifacts at the reduced/removed shadow edges. We are currently considering several methods for dealing with such artifacts. Since our method implicitly contains quite a good understanding of the underlying processes producing the edges, we feel that these problems can be dealt with.

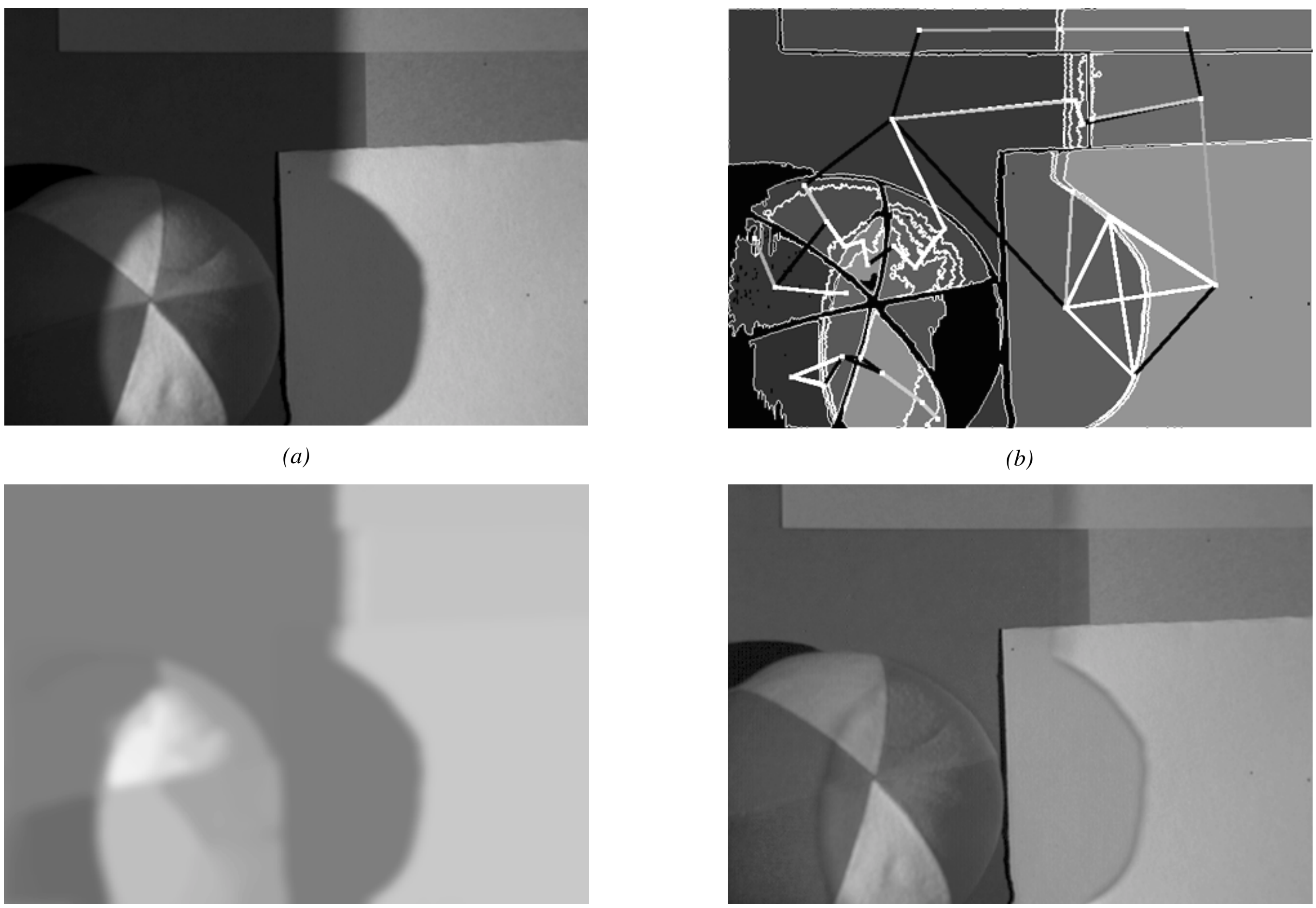
Conclusion

We have developed a method for analyzing edges with regard to whether they are likely an illumination edge or a material edge. The method includes considering whether the edge is a possible illumination edge based on a set of common illuminants, whether there are similar changes across different surfaces, whether the edge is in a geometric configuration suggestive of a shadow, and whether the edge is too strong to be a material change. We integrate these criteria with the observation that illumination changes are often gradual or somewhat gradual, leading to a specific structure with our segmentation approach, which can be exploited. Finally we show how to compute an estimate of the illumination field from the edge hypothesis. Our method is therefore ready to be used in conjunction with colour constancy algorithms, dynamic range compression, as well as a number of computer vision applications.

References

1. G. D. Finlayson, B. V. Funt, and K. Barnard, Color Constancy Under Varying Illumination, Proc. Fifth International Conference on Computer Vision, pp. 720-725 (1995).
2. K. Barnard, G. Finlayson, and B. Funt, Colour constancy for scenes with varying illumination, *Computer Vision and Image Understanding*, **65**, pp. 311-321 (1997).
3. D. J. Jobson, Z. Rahman, and G. A. Woodell, A Multi-Scale Retinex For Bridging the Gap Between Color Images and the Human Observation of Scenes, *IEEE Transactions on Image Processing: Special Issue on Color Processing*, **6**, pp. 965-976 (1997).
4. K. Barnard and B. Funt, Investigations into Multi-Scale Retinex, in *Colour Imaging in Multimedia '98*, L. W. MacDonald and M. Ronnier Luo, Eds.: John Wiley and Sons, 1999, pp. 9-17.

5. J. Frankle and J. McCann, Method and Apparatus for Lightness Imaging, United States Patent 4,384,336 (1983).
6. J. McCann, Lessons learned from Mondrians applied to real images and color gamuts, *Proc. IS&T/SID Seventh Color Imaging Conference: Color Science, Systems and Applications*, pp. 1-8 (1999).
7. G. D. Finlayson, Coefficient Color Constancy, Simon Fraser University, School of Computing, Ph.D. thesis (1995).
8. G. D. Finlayson, M. S. Drew, and B. V. Funt, Spectral Sharpening: Sensor Transformations for Improved Color Constancy, *Journal of the Optical Society of America A*, **11**, pp. 1553-1563 (1994).
9. K. Barnard and B. Funt, Experiments in Sensor Sharpening for Color Constancy, *Proc. IS&T/SID Sixth Color Imaging Conference: Color Science, Systems and Applications*, pp. 43-46 (1998).
10. K. Barnard, Practical colour constancy, Simon Fraser University, School of Computing, Ph.D. thesis (1999), available from ftp://fas.sfu.ca/pub/cs/theses/1999/KobusBarnardPhD.ps.gz.
11. G. D. Finlayson, Color in perspective, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**, pp. 1034-1038 (1996).
12. Available from http://www.cs.sfu.edu/~colour/data.
13. C. Jiang and M. O. Ward, Shadow identification, *Proc.* IEEE *Conference on Computer Vision and Pattern Recognition*, pp. 606-612 (1992).



(a) *(b)* *(c)* *(d)*

Figure 2. Illustration of the method developed in this paper applied to an image. The original image (a). The segments used showing lines joining regions separated by a plausible shadow boundary (b). The brightness of the line is proportional to the degree of belief that the edge is a shadow edge. The estimated illumination field is shown in (c), and the image with the shadow regions corrected based on the recovered illumination is shown in (d). The appropriate amount of correction is dependent on the application. Here we correct for the full illumination field effect for illustrative purposes. Other than effects at the edges due to edge localization errors, the recovery is very good. Note that some of the minor inaccuracies in the illumination field recovery are not noticeable when used to enhance the image.