Detecting and Removing Shadows

Zvi Figov
Program of Brain Science
Bar-Ilan University,
Ramat-Gan,Israel
zvi.figov@ness.com

Yoram Tal
32 shimkin st.
Haifa 34750,
Israel
yoramtal@netvision.net.il

Moshe Koppel
Dept. of Computer Science
Bar-Ilan University
Ramat-Gan, Israel
koppel@netvision.net.il

Abstract

This paper describes a method for the detection and removal of shadows in RGB images. The shadows are with hard borders. The proposed method begins with a segmentation of the color image. It is then decided if a segment is a shadow by examination of its neighboring segments. We use the method introduced in Finlayson et. al. [1] to remove the shadows by zeroing the shadow's borders in an edge representation of the image, and then re-integrating the edge using the method introduced by Weiss [2]. This is done for all of the color channels thus leaving a shadow-free color image. Unlike previous methods, the present method requires neither a calibrated camera nor multiple images. This method is complementary of current illumination correction algorithms. Examination of a number of examples indicates that this method yields a significant improvement over previous methods.

Key Words
Imaging and image processing, computer vision ,image segmentation, shadow removal, image enhancement

1. Introduction

When people look at a picture they can judge if they are looking at a shadow of a 3-dimensional object or at a painted dark area. Deciding which areas of a picture is shape and which is reflectance is a prototypical underdetermined vision problem, which biological vision systems routinely solve. Insights into this problem may apply to other vision problems as well (such as segmentation, enhancement, etc.). [3]

A gray surface, for example, might have higher luminance in sunlight than in shade, but it still looks gray. As already observed by Helmholz in 1856 [see 4], in order to achieve "lightness constancy", the visual system must discount illumination as well as other viewing conditions and estimate the reflectance. Lightness perception seems to involve all three levels of visual system hierarchy: low level vision, mid-level vision and high level vision [4], which makes it a particularly difficult problem.

The first computational approach to solving this problem focused on low-level vision. Land and McCann [5] proposed to differentiate between changes due to illuminant and changes due to reflection as follows: changes due to the illuminant are on the whole gradual, typically appearing as smooth illumination gradients, whereas those due to changes in reflectance tend to be sharp. This so-called "Retinex (Retina + Cortex) algorithm" claimed to simulate the biological mechanism.

Adelson [4] demonstrated the limitations of the low-level processing approach by showing that low-level methods based only on a local-edge detector or filters, fail in cases of images possessing the same edges but different geometries. Another argument along this line is that according to the Retinex theory the color channels are processed independently, whereas neural processing seems to be based on an opponent-color approach [5]).

Therefore the direction that must be taken to solve this problem should involve all levels of vision, not just low-level vision. In this article some high level cues will be used to achieve a better shadow detection.

1.2. Previous Work

Recovering the illumination and reflectance of a luminance image remains a difficult problem. Denoting by I(x,y) the luminance image and R(x,y) the reflectance image and L(x,y) the illumination image, the three images are related by:

$$I(x,y) = L(x,y) \bullet R(x,y) \quad (1)$$

The number of unknowns is twice the number of equations. Therefore, in order to solve this equation, assumptions must be made about the nature of either the reflectance or the illumination (or both).

Algorithms for splitting an image into shading and reflectance images can be classified as either generative or discriminative [6]. Generative algorithms use a global approach to achieve a decomposition of the image into shading and reflectance images [7,8,9]. Discriminative algorithms attempt to locally classify the image pixels as belonging to shade, reflectance or both [5,10,11,12,13,14,15,1,6].

One approach worthy of special mention is that of Finlayson et al [1], which is based on the assumption that illumination changes simply shift the log-chromaticity color vector linearly.

The method is based on 3 stages:

1. Find a single scalar function of an RGB image that is invariant to changes in light, color and intensity.
2. Find shadow edges. This is based on the reasoning that material edges should occur in both RGB *and* invariant images whereas shadow edges appear in the RGB image but *not* in the invariant image.
3. Thicken the shadow edges and remove them in the gradient log response for each channel using the method of Weiss [2]. The gradient image is then integrated in order to recover a log response, which does not have shadows.
   A major limitation of this method is its dependence on a calibrated camera.

2. The New Method

In this paper a new method is proposed for removing shadows from images, which in contrast to previous methods uses a single image and does not need a calibrated camera. Implicit in the technique is the assumption that a gradient must be either a shadow boundary or a material boundary but not both (see [1]). Therefore the algorithm works best for cast shadows where most of the shadow gradients fall on the background.

The present method follows the general strategy of [1]: shadow borders are identified and then used to remove shadow areas. However, our procedure for finding the shadow borders is different. Finding shadow borders is the result of a two stage process:

1. A color segmentation of a *reduced-* size image (in the CIE L*u*v* color space).
2. Classification of the segments into shadow areas based on a few simple rules.

The process involves a multi-resolution approach. The original image is segmented using the color values of the lower resolution image. Segments of the original image overlapping shadow segments in the lower resolution are considered shadow segments.

The rest of the paper is organized as follows: In section 2.1 we set out our segmentation method. In section 2.2 we show how to decide which segments are shadows and in section 2.3 we describe the removal of the shadows. In section 3 we show some results and state conclusions.

2.1 Segmentation

The algorithm uses two scales of the image. I_1 is the original scale image, and I_2 is the reduced scale image (we used 25% of the original). The purpose of size-reduction is to minimize the effect of color changes within the shadow, which may cause shadow segments to split even in cases of sharp borders. The image is then converted to CIE L*u*v* color representation.

The purpose of the segmentation is to divide the image into nearly homogeneous regions. What is especially important is that shadow areas form unique segments even if such segments contain different colors [16]. In Fig. 2, for example, the shadow area encompasses the grass and the pathway as well as their mutual border. Our basic assumption is that the image contains shadows with a sharp transition. This is contrary to the assumption of separating illumination and reflectance as in the Retinex. We are interested in the cases where the Retinex fails as in the cases that violate the assumption concerning illumination. Since the Retinex works well when the transition is slow we don’t expect our algorithm to work well then. All our images have failed to be corrected when running a Retinex type algorithm.

The CIE L*u*v* color representation was used since it provides a perceptually uniform color representation where differences in human perception are correctly represented. This means that a simple Euclidean distance measure can be used:

$$\| (L^*,u^*,v^*) \| = \sqrt{(L^*)^2 + (u^*)^2 + (v^*)^2} \qquad \textbf{(2)}$$

u* and v* represent the chromaticity and L* the luminance. When comparing two areas of the same material where one is lit and the other in the shade, the shadow areas are regions of similar chrominance but different luminance.

All the pixels in the new color space are then clustered using the fuzzy c-means algorithm which is the most popular derivative of the classical k-means algorithm [16]. (We used MATLAB’s version from the fuzzy logic toolbox).

The fuzzy c-means (FCM) algorithm partitions a set of data points, a color image in this case (in the L*u*v* representation where each data point is a vector of three components), into a collection of k fuzzy clusters, while minimizing an objective function. The basic limitation of this algorithm is that the number of classes is not automatically determined (although there may be methods that can predict it). Another limitation of this algorithm is that a global minimum is not guaranteed. Moreover, since the initial partition is randomly selected, different runs may lead to different segmentations (although, typically, most runs converge to the same partition).

Now that all the pixels of the lower resolution image have been classified and represented by the corresponding cluster-mean color values, the original image can be segmented using those same values. The segmented images (lower resolution and original image) are smoothed using a majority filter.

2.2 Segment Shadow Detection

For I_2, the reduced image, a segmentation L_2 is produced. A sparse graph is then formed, connecting each segment with all its significant neighboring segments. The arcs of this graph represent neighborhood relationship between segments, and their weights are controlled by two parameters:

1. The area of a significant segment relative to the area of maximum size segment. This assumes that most of the shadow area is contiguous and is not fractured into many small pieces. It also ensures that only significant shadows (we used a 10% threshold) are considered.
2. The border length ratio of a significant segment, which is the fraction of the length of the common border relative to the total segment perimeter. The inclusion of this feature is based on the idea that if an object casts a shadow onto some background, a high percentage of the border is expected to be found between the object and the background. A shadow must have a significant border in order to enable its classification (we used 20%).

Each node of the graph represents a segment, which may have one or more neighbors. The arcs were found through the corresponding significant borders. Initially, all segments that passed the first test were chosen and their borders were thickened. Then the pixels on a thickened border were examined as to their membership, and those that passed the second test were added to the graph.

We are interested in finding a *segment pair* of two neighboring segments, which forms a *shadow pair*, i.e., one of its segments being a shadow and the other being a background segment. Note that even though a node in the graph may have many connections, a qualification as a shadow pair involving one of its connections is sufficient for this purpose.

Shadow segments obviously have a lower intensity than their neighboring (non-shadow) segments. However, for *shadow pairs* this intensity gap cannot be too large since, at least for natural scenes, shadows are never absolute (zero intensity). The intensity ratio of a *shadow pair* must therefore stay within certain limits in order to be recognized as such. We used the intensity of the CIE L*u*v* to verify this property. The ratios of the average intensities of neighboring segments (computed via the corresponding border pixels) were calculated. Only segment pairs satisfying those requirements were kept as a possible shadow pair (shadow and background).

For all the segment pairs that passed the previous tests the corresponding color ratios were calculated. Color ratios [17] or reflectance ratios [18] have been used to detect shadow changes or reflectance changes. If we look at two points next to each other where one is in the shadow and the other is not the brightness values I_1 and I_2 of the two points may be written as:

$$I_1 = L_1 \bullet R_1 \tag{3}$$

$$I_2 = L_2 \bullet R_2 \tag{4}$$

where L_1, L_2 are the illumination values and R_1, R_2 are the reflectance values for the two points. Since $R=R_1=R2$ their reflectance is the same, the ratio of their brightness values is simply:

$$\text{Ratio} = L_1 / L_2 \tag{5}$$

For a given pair of regions the ratio is calculated using mean values: Ratio = mean(L_1)/mean(L_2). This ratio is calculated for all bands. For shadow/non-shadow changes where the reflectance is the same, the ratio is the result of illumination differences, whereas for material changes the ratio is the result of reflectance change (which is therefore invariant under different illuminations).

The problem is how to decide if a given ratio represents a shadow change or a material change. One way of solving this problem is finding a shadow bordering on two different areas (different reflectances) or finding two shadows bordering on the same area (same reflectance).

Consider the color ratio in an ideal situation where areas have constant reflectance and there is one light source with constant illumination, in the case of one shadow and two different areas:

Let A_1 and A_2 denote the shadow areas, while B and C denote two areas with different reflectance values:

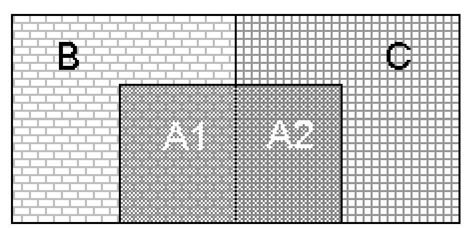

Fig. 1. B and C are two areas representing different reflectance values. A_1 and A_2 are the shadows falling respectively on B and C.

If all areas are illuminated by a single source, and Ra_1,Ra_2,Rb,Rc are the respective reflectance values and La_1,La_2,Lb,Lc are the illumination values then:

$$\text{Ratio}(A_1,B) = \frac{La_1 \bullet Ra_1}{Lb \bullet Rb} = \frac{La_1}{Lb} \tag{6}$$

This is because Ra_1=Rb. The shadow is only an illumination change.

$$\text{Ratio}(A_2,C) = \frac{La_2 \bullet Ra_2}{Lc \bullet Rc} = \frac{La_2}{Lc} \quad \textbf{(7)}$$

We will assume that in the simple case there is one light source, so that Lb=Lc (the same illumination) and $La_1=La_2$ (but $La_1 \neq Lb$ and $La_2 \neq Lc$). It follows that:

$$\text{Ratio}(A_1,B) = \text{Ratio}(A_2,C) \quad \textbf{(8)}$$

This is true for an ideal case. In cases of more light sources, a shadow segment may have neighbors, which differ in both reflectance and illumination. To cope with such problems, it is assumed that the illumination changes are locally constant (even though the illumination may globally change throughout the image). The ratio is calculated using neighboring border pixels in the RGB color space. For each segment we seek the minimum distance between the ratios of the current segment and all his possible shadows neighbors (traversing the graph) using the root mean square error criterion. A threshold for this distance value (P) is calculated based on the error of the clustering process.

$$p = \sqrt{\sum_i (\text{Ratio}_i(A_1 - B) - \text{Ratio}_i(A_2 - C))^2} \quad \textbf{(9)}$$

where i=1, 2, 3 (for Bands RGB)

Implicit in our technique is the assumption that a gradient can always be associated either with a shadow or with an object but not with both (as in [1]). Therefore the algorithm works best for cast shadows (where most of the shadow gradients fall on the background).

The method was tested on various images, seven of which are displayed below. The first 4 images (Figures 2-8) were taken from Finlayson et. al[19]. The other three images (Figure 10) were taken with a standard digital camera (Nikon CoolPix 950) using a minimal JPEG compression. The value of k=5 was chosen for all runs of the fuzzy c-means algorithm, and was used for all the images. The clustering was run 5 times with the smallest error taken. All the images were run using the same parameters.

The end result of the present algorithm is that the most prominent shadows were removed in all the test images. There are some smearing effects at the shadow borders in both methods, which are due to incorrect removal of gradients at some pixels. Those pixels, which now have an erroneous zero gradient, cause an incorrect reproduction of the nearby pixels, which is visualized as smearing. This effect is even more pronounced at shadow border pixels, which are contained in textured areas due to the fact that such areas possess strong, rapidly varying gradients.

However, most of the artifacts appearing in the results of Finlayson et.al [1] do not appear in our images. For example, in Fig. 4b part of the bridge and the sky were detected by Finlayson as shadows (in the upper right part of the image), but are correctly reproduced in fig. 4c. In Fig. 6b most of the windows in the building image were incorrectly detected by Finlayson et.al as shadows whereas our algorithm handled them correctly (Fig 6c). Note that for both methods the color of the “shadowless” images are not the same as the respective original. This is due to the integration method that can recover the image color up to a multiplicative constant (same in both methods).

Now consider the images in Fig. 10. In the first image the prominent shadow was removed from the grass area, clearly revealing a red pipe (the poor grass recovery is due to the JPEG compression noise). In the second image the shadows are effectively removed except for a small shadow on the pavement. In the third image the prominent shadow is effectively removed. The other shadows at the far side of the road were correctly detected but their removal caused the far car to be smudged. This is due to the shadow falling on the car itself, a situation that violates the assumption that the shadows fall on the background, not on the object casting the shadow.

Errors in the present segment-based method may have far-reaching effects. An incorrect identification of some segment may wipe out a whole area (as in the last image in Fig. 10). In Finlayson's pixel-based method the errors are randomly distributed within the image, therefore a complete removal of non-shadow objects is virtually impossible. On the other hand, we have seen that overall the present method produces less errors than the method of Finlayson et. al. We have also seen that it is possible to obtain similar results, if not better, regardless of the camera calibration. We note that in all Finlayson’s images the shadow borders are strong, therefore the Retinex like algorithms did not manage to remove or even illuminate the shadows (we tested all our images using a similar implementation in [20]).

3. Conclusion

The method presented in this paper is successful in finding and removing shadows providing the following conditions are met:

1. The image can be segmented successfully into segments based on colors alone.
2. The shadows are significant and fall on two different background areas or there are two or more shadows next to one area.
3. The original image is not overly compressed.

There is still a good deal of work to be done on the problem of shadow detection and removal. A number of issues need further investigation including (i) finding the optimal value of k for the fuzzy c-means algorithm, (ii) dealing with shadow areas having a single border, and (iii) classifying the x,y gradients of a shadow area into shadow and reflectance using *shape* information in addition to *color* information (see [6]).

References

1. G.D.Finlayson, S.D.Hordley and M.S.Drew, Removing Shadows from Images, ICCV 2002
2. Y.Wiess, Deriving intrinsic images from image sequences, In ICCV 2001, pages 68-75, IEEE 2001
3. W.T.Freeman and Paul A.Viola ,Bayesian model of surface perception, Advances in Neural Information Processing Systems (NIPS) 10,1998
4. E.H.Adelson, Lightness Perception and Lightness Illusions. In the New Cognitive Neorosciences, 2nd ed. 2nd ed., M. Gazzaniga, ed. Cambridge, MA: MIT Press, pp. 339-351, 2000.
5. E.H.Land and J.J McCann, Lightness and retinex theory, J Opt.Soc Amer, 61, 1-11,1971
6. M.F.Tappen, W.T.Freeman and E.H.Adelson, Recovering Intrinsic Images from a Single Image, Advances in Neural Information Processing Systems 15, 2003
7. P. Sinha and E. H. Adelson. Recovering reflectance in a world of painted polyhedra. In Fourth International Conference on Computer Vision, pages 156-163. IEEE, 1993
8. E.H.Adelson and A.P.Pentland, The perception of shading and reflectance. In D.Knill and W.Richards, editors Perception as Baysian Inference (pp. 409-423). New York: Cambridge University Press ,1996
9. V.M. Bove and W.J.Butera, The Coding Ecology: Image Coding Via Competition Among Experts, IEEE Transactions on Circuits and Systems for Video Technology,Oct. pp.1049-1058, 2000
10. D.Marr, Vision,1982 W. H. Freeman and Co., 1982.
11. R.Gershon, A.D.Jepson and J.K.Tsotos, Ambient illumination and the determination of material changes, Journal of the Optical Society of America, Vol 3, page 1700-1707, October 1986
12. T.Gevers and H.M.G. Stokman, Classifying Color Transitions Into Shadow-Geometry, Illumination, Highlight Or Material Edges, IEEE ICIP, pp. 521-525, part I, 2000
13. B.V.Funt, M. Drew and M.Brockington, Recovering Shading from Color Images, In G. Sandini, editor, ECCV-92: Second European Conference on Computer Vision, pages 124-132. Springer-Verlag, May 1992
14. E.Salvador, A.Cavallaro and T.Ebrahimi, Shadow Identification and classification using Invariant Color Models, Proc. of IEEE Signal Processing Society ,ICASSP-2001, Salt Lake City (Utah, USA), 7-11 May,pp 1545-1548, 2001
15. Matt Bell and William T.Freeman, Learning local evidence for shading and reflectance, IEEE International Conference on Computer Vision, p.670-677, July 2001
16. L.Lucchese and S.K.Mitra, Color Image Segmentation – A State-of-the-Art Survey, Proc. of the Indian National Science Academy (INSA-A), New Delhi, India, Vol. 67, A, No. 2, March 2001, pp. 207-221.,2001
17. K.Barnard and G. Finlayson, Shadow Identification using Colour Ratios, Proceedings IS&T/SID, Eight Color Imaging Conference Color Science Systems and applications, pp 97-101, 2000
18. S.K.Nayar and R.M.Bolle, Reflectance Based Object Recognition, IJCV 17(3), 219-240 ,1996
19. Images are from http://www.cs.sfu.ca/~mark/ftp/Eccv02/ECCVShadows.ppt as of 1/07/03
20. http://www.shadowilluminator.org/ as of 1/1/04

Fig. 2. (a)Original, (b)Finlayson et al[1] results, (c)our result

Fig. 3. Shadow Borders :(a) Finlayson et al[1] result, (b)Our result

Fig. 4. Second Image (a) Original, (b) Finlayson et al removed (c) Our removed

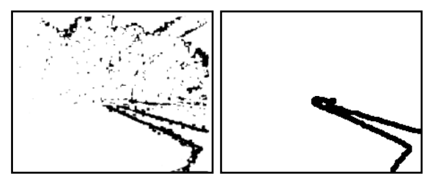

Fig. 5. Shadow Borders : (a) Finlayson et al[1] result, (b) Our result

Fig. 6. (a) Third Image Original, (b) Finlayson et al removed (c) Our removed

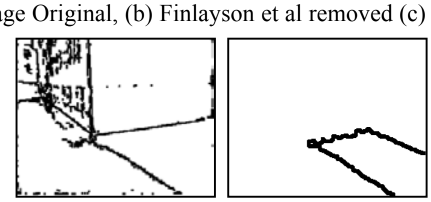

Fig. 7. Shadow Borders (a) Finlayson et al[1] result (b) Our result

Fig. 8. (a) Fourth Image Original, (b) Finlayson et al removed (c) Our removed

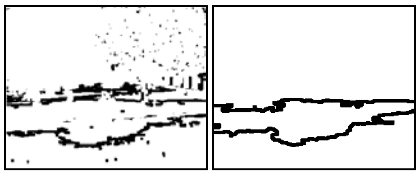

Fig. 9. Shadow Borders (a) Finlayson et al[1] result (b) Our result

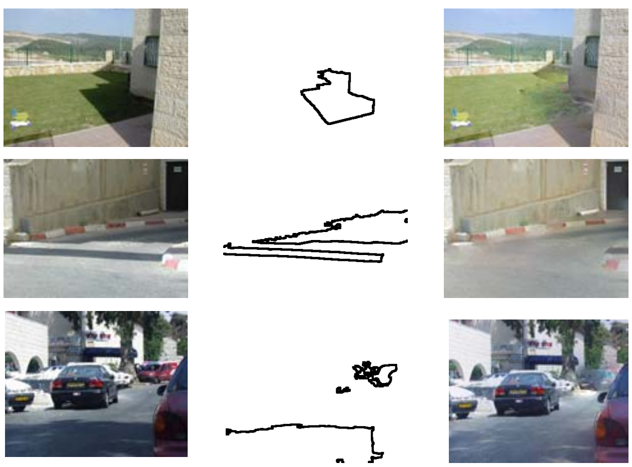

Fig. 10. More examples. The first column is the original followed by the shadow detected followed by shadow removed