Single Image Shadow Removal via Neighbor-Based Region Relighting

Tomás F. Yago Vicente, Dimitris Samaras

Image Analysis Lab, Computer Science Dept., Stony Brook University, NY, USA

Abstract. In this paper we present a novel method for shadow removal in single images. For each shadow region we use a trained classifier to identify a neighboring lit region of the same material. Given a pair of lit-shadow regions we perform a region relighting transformation based on histogram matching of luminance values between the shadow region and the lit region. Then, we adjust the CIELAB a and b channels of the shadow region by adding constant offsets based on the difference of the median shadow and lit pixel values. We demonstrate that our approach produces results that outperform the state of the art by evaluating our method using a publicly available benchmark dataset.

Keywords: Shadow Removal Illumination SVM Histogram Matching Texture Recovery Image Processing

1 Introduction

Shadows are a common phenomenon in natural scenes. Shadows appear whenever an object occludes the scene's illuminant(s). Hence, shadows are an outcome of the complex interactions between geometry, albedo and illumination sources present in a scene.

Humans can derive useful visual cues from shadows to help perceive shapes, occlussions, or objects' points of contact with surfaces. However, automatically extracting these cues from images remains a challenging task. Moreover, shadows are well known to wreak havoc in a plethora of computer vision tasks such as segmentation, object detection, tracking, scene understanding or shape-from-X. Therefore, shadow-free images would help improve the performance of all these tasks. Also, shadow removal in images may be of interest for aesthetic reasons, as well as for image editing or computational photography.

There has been a growing interest in shadow detection in the past few years. Recent works using datasets of training images with labelled shadows and learning techniques have provided great advances in the state of the art [1], [2],[3] and most recently [4].

In this paper we focus on the problem of shadow removal from a single image. In earlier work, Finlayson *et al.*[5][6] remove shadows by zeroing shadow edges in the gradient domain and then integrating to obtain a shadow free image. They achieve good results with high quality images, however the integration often introduces changes in color balance, global smoothness and loss of textural

properties, specially in the penumbra or boundary areas. In [7], Liu *et al.* propose an integration based algorithm that attempts to improve the loss of texture that commonly accompanies integration methods. They construct a gradient field for the penumbra area to cancel out the effects of the illumination change. Their results improve in terms of texture consistency but they cannot handle non uniform shadows or complex textures. Integration based methods are highly sensitive to accurate segmentation of the shadow edges.

Shor *et al.*[8] present an affine shadow formation model with a multi scale scheme to remove shadows. They require minimal user assistance to identify shadow and lit areas of the same surface material. Based on those pairings, they obtain the constant parameters of the shadow model. Due to the assumed constant coefficient their method has problems with non uniform shadows, it also presents issues with rich textures.

Wu *et al.*[9] perform shadow matting to remove shadows. They estimate shadow intensities based on intensity ratios in the umbra region and use a Bayesian framework to regularize the shadow scale factor in the shadow regions. The umbra regions of the shadows are assumed to be roughly uniform. Guo *et al.*[1] also remove shadows based on shadow matting, they generate a soft shadow mask from the ground truth and randomly sample patches from both sides of the shadow boundary to compute the illumination ratios. Guo *et al.*[1] extensively evaluate their results on a shadow dataset that is publicly available. It is the first work to present qualitative and especially quantitative evaluation results on a somewhat large dataset as opposed to a few selected images.

We present a novel method for single image shadow removal based on region relighting. We leverage the use of a dataset with annotated shadows to train a classifier that identifies non-shadow regions that neighbor shadow regions of the same material. We propose to use a neighboring lit region to relight a shadow region. To do so, we first match the luminance values of the shadow pixels to the luminance histogram of the lit region. Then, we adjust the shadow region chromaticities by adding the difference between the median CIELAB a and b values of the lit region and the shadow region. However, the image segmentation often outputs inaccurate boundaries such that shadow(lit) pixels leak into a lit(shadow) region. Hence, we perform the relighting process only on the core pixels of the regions. That is, we ignore the outer perimeter pixels of each region. We iteratively find pairs of shadow and lit neighbors and relight the shadow regions. Finally, we process the shadow boundaries. Our results outperform the state of the art in the benchmark dataset[1]. For shadow pixels we obtain a shadow removal error, measured as Root Mean Square Error (RMSE), of 9.24, a 21% reduction compared to [10]. This article contains the following main contributions:

- A novel technique for shadow region relighting based on a neighboring lit region.
- A new classifier to identify pairs of shadow and lit regions of the same material.
- Extensive evaluation on the only published benchmark dataset.

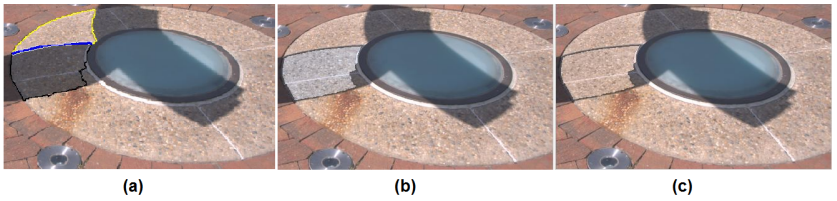


Fig. 1. a) Shadow region and lit neighbor: shadow region depicted with black boundaries, lit neighboring region depicted with yellow boundaries, common boundary drawn in blue. b) RGB reconstuction showing the result of histogram matching on L channel for the shadow region. c) Shadow region relit, results after the adjustements in a and b channels.

The rest of the paper is organized as follows: Section 2 contains a detailed description of our method to relight a shadow region. Section 3 describes the pre-processing stage of our work. Section 4 describes the lit neigbor classifier. In Section 5 we describe the full pipeline for shadow removal. Experimental results are presented in Section 6. Finally, Section 7 concludes the paper.

2 Region Relighting

Given a shadow region R_s and a neighbor non-shadow region of the same material R_l, we look for a transformation T that relights R_s. Since the two regions are close to each other and have the same material, a transformed version of R_s should closely resemble the appearance of the lit region R_l. The relighting transformation T depends on the appearance of the lit region. We have:

$$T(R_s, R_n) = \widehat{R_s}, \text{ such } \widehat{R_s} \approx R_n \tag{1}$$

We perform the relighting transformation in CIELab color space. First, we compute the 50 bin histogram of the luminance values, L channel, of R_l ($H_{R_l(L)}$). Then, we carry out histogram matching so that the shadow region L values match the lit region histogram [1]. Figure 2 contains an example of this step.

The resulting luminance histogram, $H_{\widehat{R_s}(L)}$, resembles that of the lit region $H_{R_l(L)}$ while still preserving a similar shape to the original shadow values $H_{R_s(L)}$. Figure 1(b) depicts the results of this step if we convert back to RGB with the adjusted luminance values for the shadow region. Figure 1(a) shows the original input image with R_l boundaries drawn in yellow and R_s boundaries drawn in black. The image segmentation often produces small inaccuracies around the regions' boundaries. That is, few shadow pixels leaking into a lit region (or vice versa) or small chunks of different material(s) are getting added to an otherwise homogeneous region. These spurious pixels modify the range of luminance values of a given region, which can severely affect the histogram matching results.

[1] We use Matlab's histeq function.

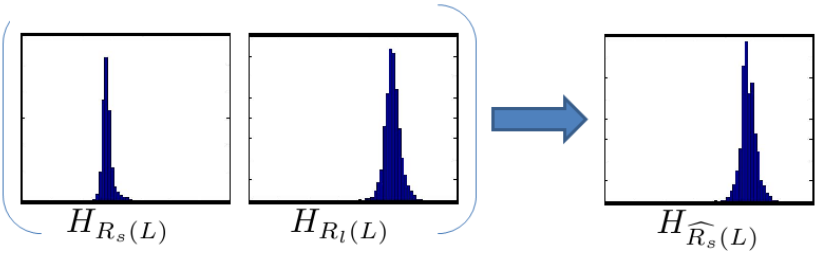


Fig. 2. Histogram matching for region relighting. L channel histograms of the shadow and lit regions depicted in figure 1(a), and the resulting matched histogram corresponding to relit region in 1(c).

Hence, we apply the relighting transformation using only the core pixels of each region. That is, we exclude the outer perimeter pixels (resulting of eroding each region with a 3 by 3 identity matrix as neighborhood structure).

As a second step, we adjust the a channel of the shadow region by adding the difference between the median a values of R_l and the median a values of R_s. Finally, the same operation is carried out for the b channel to complete the relighting process $T(R_s, R_n)$ yielding $\widehat{R_s}$. In figure 1(c), we can see the reconstructed RGB image showing the final relighting results.

3 Preprocessing

In this section we describe the initial preprocessing stages of our method. Our algorithm takes as input images containing shadows and their respective shadow masks. We begin this section by describing the segmentation of the input image into regions or superpixels. Then, we introduce the processing of the ground truth.

3.1 Region Segmentation

The quality of the region segmentation will affect the performance of our shadow removal as we operate at the region level. Ideally, we want to segment the image into superpixels that correspond to consistently illuminated regions. That is, either all pixels in a region are in shadow or all are not in shadow. Furthermore, we would like to obtain homogeneous regions in terms of material.

To segment the images into regions we use the segmentation method proposed by Yago [4], where they segment images into regions for shadow detection. The method consists of a two step process that is fast and robust to the choice of thresholds. First, SLIC [11] superpixel segmentation is applied to oversegment the image, obtaining an initial set of superpixels. Then, mean-shift clustering[12] over the superpixels' mean color in CIELAB color space is performed. Lastly,

Fig. 3. a) SLIC superpixel segmentation. b) Resulting merged superpixels from a) after mean-shift clustering over mean CIELAB colors. c) Input ground truth shadow mask. d) Overlay of processed ground truth onto segmented regions.

adjacent superpixels that belong to the same mean-shift clusters are merged into a larger region. After this second step, the final regions are considerably less than the number of superpixels in the previous step. Most of the segmented regions are consistently illuminated. In figure 3 (a) and (b) we can see example results of the first and second step of the segmentation algorithm. Once an image is segmented, we compute which regions are adjacent to which (i.e. they share a common boundary), hence defining pairs of neighboring regions.

3.2 Ground Truth Processing

The data set presented in [1] contains binary shadow masks as shadow ground truth. To generate region level shadow labels we overlay the regions segmented as described in the previous section 3.1. We label regions as shadows if they contain more than 50% shadow pixels. We implemented a Graphical User Interface in Matlab to manually annotate pairs of shadow and lit regions of the same material. With this GUI we can also refine the shadow labels.

4 Classifier for Lit Neighbors

We propose a classifier that takes as input a shadow region and a neighboring lit region. For each shadow region R_s we need to identify which of its lit neighbors R_i shares the same material with R_s. If a lit neighbor shares the same material then it can be used to relight R_s by applying the transformation T previously described in Section 2. Hence, we select features that describe: i) the similarity between R_s and R_i, ii) the transformation defined by the pair of regions $T(R_s, R_i)$, and iii) the results of applying that transformation. If R_r and R_i, have the same material, the relit shadow region $\widehat{R_s}$ and the lit region should be similar in color and texture. We compute the following features:

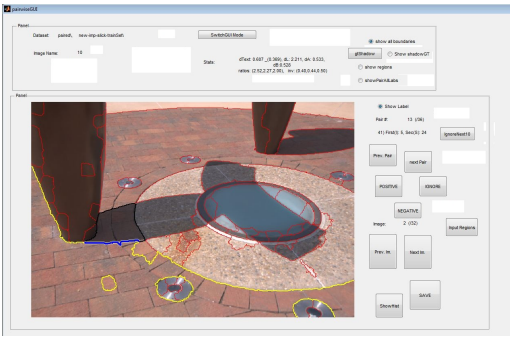


Fig. 4. GUI to annotate training data for shadow and lit regions of the same material, and to refine shadow labels.

- RGB color ratios between R_i and R_s: t_r, t_g, t_b encoded as $\frac{t_r+t_g+t_b}{3}$, $\frac{t_r}{t_b}$, $\frac{t_g}{t_b}$ [2].
- Earth Mover's Distance(EMD) between each region's luminance histograms.
- Median based a and b offsets defined by $T(R_s, R_i)$.
- EMD between the a and b histograms of R_i and the resulting region $\widehat{R_s}$
- χ^2 distance between the texton histogram of the relit region $\widehat{R_s}$ and the texton histogram of the combined regions $\widehat{R_s}$ and R_i.

For all the feature computations we only consider the central pixels of each region. The L, a and b histograms contain 50 bins. Positive training examples are pairs of neighboring regions sharing the same material with one being shadow and the other lit. For negative training examples the lit region is of a different material than the shadow region. We train a probabilistic SVM classifier[13] with a Gaussian RBF kernel. For model selection we perform grid search with 5-fold cross validation. We use the fast version of libSVM implemented by [14].

To generate a texton codebook we ran the full MR8 filter set [15] on the whole data set and cluster the filter responses into 128 textons using K-means.

5 Shadow Removal Pipeline

Our shadow removal method takes as input an RGB image and a binary shadow mask. As a first step, we segment the image into regions and automatically label each region as shadow or lit (as described in Section 3). For each shadow region we extract its lit neighboring regions building a set of lit-shadow pairs.

Second, we compute the features for each pair of regions, as described in the previous section, and run the classifier. The positive classifications are selected as candidate relighting pairs. If for a shadow region more than one lit neighbor is classified as positive we only consider the one with the highest classification confidence.

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

Fig. 5. Shadow removal pipeline. (a) Input image with overlaid shadow mask, boundary of segmented regions depicted in red. (b) Removal results after first iteration of our method. (c) Removal results after the second interation. (d) Final removal results after boundary areas are relit.

On the next stage, region relighting is performed on the candidate relighting pairs according to the process described in section 2. After that, we label the set of relit regions as lit. Hence, new pairs of lit-shadow regions are created so we can start a new cycle of identifying candidate relighting pairs using the classifier and then relighting regions based on the positive classifications. Figures 5(b) and 5(c) depict the shadow removal results after the first and second iterations of our method, respectively. As we can observe, there are three isolated shadow regions (no lit neighbors) that are successfully relit in the second round.

As a final step, we address the so far ignored boundary pixels. To remove the shadow in the outer perimeter p_s of a relit shadow region $\widehat{R}_s$, we propose a two step operation:

1. Adjust the L, a and b values of the pixels in p_s based on the core pixels of $\widehat{R}_s$. First, we compute the mean L, the median a and the median b for the core pixels and for the boundary pixels. Then, we add the differences to the pixels in p_s.
2. Smooth the new boundary pixels' values. We convert the results from the previous step to RGB. Then, we run a Gaussian filter at the locations of p_s to obtain the final values for the boundary pixels.

6 Experiments and Results

In this section we present quantitative and qualitative results of our shadow removal method using the dataset presented in [1]. This dataset contains 32 training images for which we manually annotated ground truth for our lit neighbor classifier. The testing split contains 48 shadow images for which there is a corresponding shadow-free image, considered as ground truth for shadow removal evaluation.

6.1 Quantitative Results

In table 1, we present our quantitative results compared to the state of the art by Guo *et al.*[10]. As evaluation metric we use the Root Mean Squared Error

(RMSE) in CIELab space between the shadow-free images and the results of applying our full shadow removal pipeline. We compare to the results presented by [10] when using ground truth shadow masks as input for their shadow removal method. Note that we also take shadow masks as input. Furthermore, the performance of [10] deteriorates considerably when the shadow removal is applied on their shadow detection results.

As we can see in table 1, our overall error is almost half a unit lower than the state of the art, 5.96 versus 6.4. For shadow region pixels our performance reduces the error by a 21%, yielding an RMSE of 9.24 units. The performance we get on non-shadow regions is slightly worse, 4.9 versus 4.7. This is due to small faults in the segmentation such that lit pixels leaked into shadow regions. With no shadow removal applied, the error in non-shadow regions is 4.6.

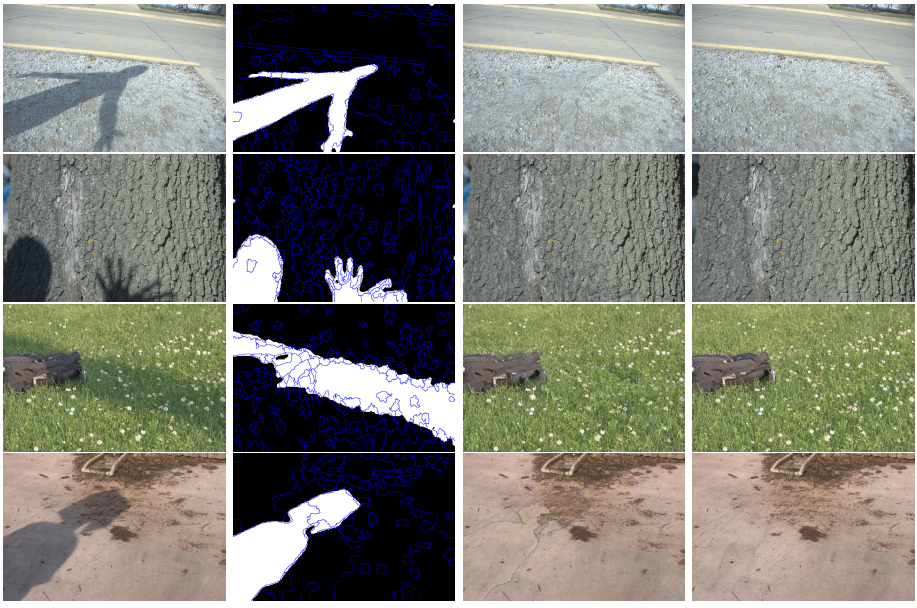
Region Type	Original	Guo *et al.*[10]	Region Relighting(Ours)
Overall	13.7	6.4	5.96
Shadow regions	42.0	11.8	9.24
Non-shadow regions	4.6	4.7	4.9

Table 1. Shadow removal evaluation on the dataset presented in Guo *et al.*[1]. First column shows the error when no shadow removal is carried out. Second column are the state of the art results by Guo *et al.*[10], their method applies matting on the ground truth shadow mask. Third column are the results of our method.

We also evaluate our performance in shadow regions for the core pixels and for the border pixels separately, see table 2. The shadow removal error for core regions is 8.81, whereas the error in the border regions is noticeably worse at 14.09. Moreover, some shadow regions cannot be relit as no suitable lit neighbor is detected by our classifier, or does not exist in the image. The third row of the table shows the error on the core pixels of the shadow regions that were actually relit by our method. The RMSE obtained drops to 8.12.

Region Type	Original	Region Relighting
Border Shadow Regions	36.92	14.09
Core Shadow Regions	42.45	8.81
Relit Shadow Regions	37.33	8.12

Table 2. Per pixel RMSE of the shadow removal. Fist column shows the results when no shadow removal is performed. Second column are the results of our method. The core shadow regions are the shadow regions excluding their outer perimeter pixels (resulting of eroding each region with a 3 by 3 identity matrix as neighborhood). The border regions are the excluded perimeter pixels. The relit shadow regions are the shadow regions for which our method actually performed shadow removal (excluding the outer perimeters).



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

Fig. 6. Shadow removal results. (a) Input image. (b) Ground truth shadow pixel mask with the region segmentation overlaid in blue. (c) Our shadow removal results. (d) Ground truth shadow free image.

6.2 Qualitative Results

In figure 6 we present some qualitative results. As we can observe, our method produces high quality shadow free images for a variety of materials and textures. In the first and forth images our shadow free image presents a noticeable boundary effect around the shadow regions. This is mostly due to inaccuracies in the region segmentation with respect to the actual shadows. However, the quality of the shadow removal in the inner areas is quite high. Table 3 contains the actual error numbers for the images depicted in figure 6. We can appreciate how the RMSE error in the core areas of the shadow regions is particularly low.

Some interesting shadow removal cases are presented in figure 7. In these cases we can observe weaker qualitative results. In the first image we can notice some boundaries between relit regions due to poor performance by our boundary processing. The image in the second row depicts a case where some regions within the person's shadow were not able to be recovered as no suitable lit region was found by the classifier. Images 3 in 4 show some strong boundary effects, very noticeable by the human eye. In these cases the segmentation does not align well with the actual shadow boundaries. For instance, in image 3 most of the outer perimeter of the shadow leaked into adjacent lit regions. However, the error for shadow regions in these images is relative low, 6.57 and 4.84 respectively; and

Image	Overall Error	Shadow Error	Shadow Core	Core Original
Fig.6 1st row	8.88	9.91	9.81	29.64
Fig.6 2nd row	13.76	12.30	12.21	37.49
Fig.6 3rd row	6.47	11.18	11.09	24.16
Fig.6 4th row	3.09	6.07	5.81	41.69

Table 3. RMSE on the images shown in figure 6. First column shows the overall error. Second column depicts the error in shadow regions. The shadow core error is the error on the core shadow pixels. Core original is the error in the shadow core pixels for the original image, with no shadow removal performed.

even lower in the core shadow pixels 6.29 and 4.58, respectively. Detailed error numbers for the images presented in figure 7 can be found in table 4.

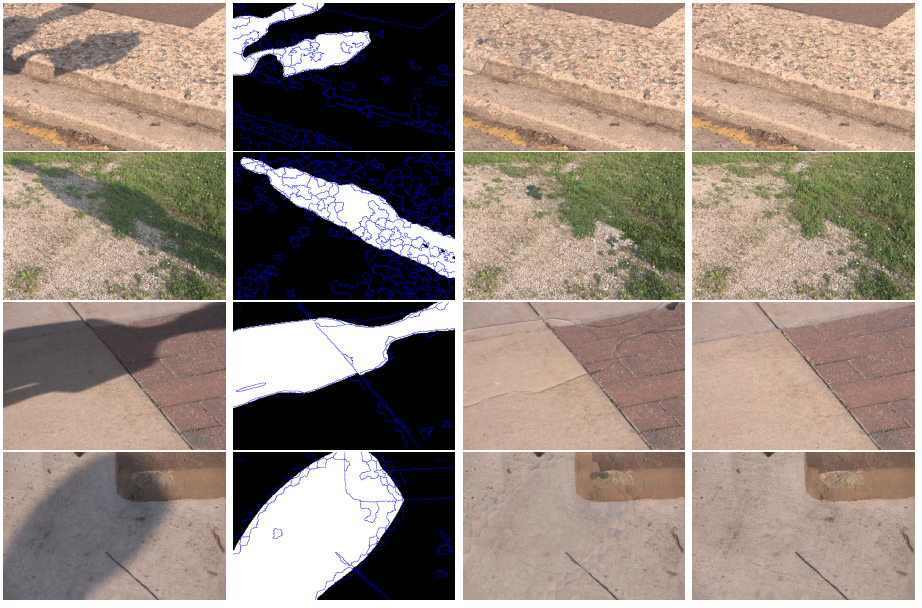


Fig. 7. Note-worthy shadow removal results. (a) Input image (b) Ground truth shadow pixel mask with the region segmentation overlaid in blue. (c) Our shadow removal results. (d) Ground truth shadow free image.

7 Conclusion and future work

We have presented a novel method for shadow removal in single images that outperforms the state of the art. The main contribution of our work is a new region

Image	Overall Error	Shadow Error	Shadow Core	Core Original
Fig.7 1st row	4.55	11.54	11.20	43.97
Fig.7 2nd row	6.89	12.96	12.66	28.25
Fig.7 3rd row	6.31	6.57	6.29	50.91
Fig.7 4th row	3.77	4.84	4.58	38.17

Table 4. RMSE on the images shown in figure 7. First column shows the overall error. Second column depicts the error in shadow regions. The shadow core error is the error on the core shadow pixels. Core original is the error in the shadow core pixels for the original image, with no shadow removal performed.

relighting transformation based on histogram matching of luminance values between the shadow region and the neighboring lit region, plus addition of median based offsets in the a and b channels. Furthermore, we propose a new classifier to automatically identify suitable pairs of lit-shadow regions. We demonstrated that the iterative application of the proposed transformation in positively classified pairs of region is powerful enough to outperform the state of the art on the shadow removal benchmark dataset. Our results are specially accurate in the core pixels of the shadow regions.

In future work we will explore alternative ways to deal with the boundary pixels such as in-painting techniques. We are also interested in region segmentation tailored for the task of shadow removal.

Acknowledgements

This work was partially supported by NSF IIS- 1161876, IIS-1111047, NIH R21 DA034954 and the DIGITEO Institute, France.

References

1. Guo, R., Dai, Q., Hoiem, D.: Single-image shadow detection and removal using paired regions. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. (june 2011) 2033 –2040
2. Huang, X., Hua, G., Tumblin, J., Williams, L.: What characterizes a shadow boundary under the sun and sky? In: Computer Vision (ICCV), 2011 IEEE International Conference on. (nov. 2011) 898 –905
3. Lalonde, J.F., Efros, A.A., Narasimhan, S.G.: Detecting ground shadows in outdoor consumer photographs. In: Proceedings of the 11th European conference on Computer vision: Part II. ECCV'10, Berlin, Heidelberg, Springer-Verlag (2010) 322–335
4. Yago Vicente, T.F., Yu, C.P., Samaras, D.: Single image shadow detection using multiple cues in a supermodular MRF. In: Proceedings of the British Machine Vision Conference, BMVA Press (2013)
5. Finlayson, G., Hordley, S., Lu, C., Drew, M.: On the removal of shadows from images. Pattern Analysis and Machine Intelligence, IEEE Transactions on **28**(1) (jan. 2006) 59 –68

6. Finlayson, G., Drew, M., Lu, C.: Entropy minimization for shadow removal. International Journal of Computer Vision **85** (2009) 35–57 10.1007/s11263-009-0243-z.
7. Liu, F., Gleicher, M.: Texture-consistent shadow removal. In: Computer Vision–ECCV 2008. Springer Berlin Heidelberg (2008) 437–450
8. Shor, Y., Lischinski, D.: The shadow meets the mask: Pyramid-based shadow removal. Computer Graphics Forum **27**(2) (April 2008) 577–586
9. Wu, T.P., Tang, C.K., Brown, M.S., Shum, H.Y.: Natural shadow matting. ACM Trans. Graph. **26**(2) (June 2007)
10. Guo, R., Dai, Q., Hoiem, D.: Paired regions for shadow detection and removal. IEEE Transactions on Pattern Analysis and Machine Intelligence **99**(PrePrints) (2012) 1
11. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. IEEE TPAMI (2012)
12. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence **24**(5) (2002) 603–619
13. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology **2** (2011) 27:1–27:27 Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.
14. Li, F., Carreira, J., Sminchisescu, C.: Object recognition as ranking holistic figure-ground hypotheses. In: IEEE Conference on Computer Vision and Pattern Recognition. (2010)
15. Varma, M., Zisserman, A.: Classifying images of materials: Achieving viewpoint and illumination independence. In: European Conference on Computer Vision. Volume 3., Springer-Verlag (2002) 255–271