

Single Image Shadow Detection using Multiple Cues in a Supermodular MRF

Tomás F. Yago Vicente
tyagovicente@cs.stonybrook.edu

Chen-Ping Yu
cheyu@cs.stonybrook.edu

Dimitris Samaras
samaras@cs.stonybrook.edu

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

Abstract

In this paper we propose a complete methodology for single image shadow detection based on the learned appearance of shadows. The basis of our method is a novel single region SVM classifier with a multi-kernel model specifically tailored for shadow region classification. This classifier, which already outperforms much more complex methods, provides the unary potentials for an MRF optimization that also includes pairwise potentials encoding the relationships between neighboring regions in the image. We introduce a novel boundary classifier for shadow boundaries cast over surfaces with the same material, and two improved paired regions classifiers; one for adjacent regions of the same material taken under the same illumination, and one for regions of same material taken under different illumination. The strength of the unary classifier means that our MRF requires only relatively sparse pairwise potentials, resulting in a more efficient and accurate optimization as can be seen in our experimental results. We reduce the balanced error rate by 53% compared to the state of the art on the latest shadow detection image dataset.

1 Introduction

Shadows are commonly found in natural scenes. They result from occlusions of the scene's illumination source(s). As such, shadows are an outcome of the complex interactions between the illuminants, the geometry, and the reflectance properties of the materials present in a scene. Shadows provide useful visual cues to infer the illumination of the scene, or the size and shape of objects. However, shadows are also well known to confound many computer vision tasks such as image segmentation, scene understanding, or object recognition.

Single image shadow detection is a well studied problem. Earlier methods based on comparing gradients in the original image and its illumination invariant representation (where the effects of illumination changes are suppressed from the image), such as [2] [8], showed remarkable results in high quality images. However, their performance degrades significantly with consumer grade photographs or web quality pictures [18]. More recent methods ([15], [18], [13], [10]) use image datasets with labelled shadow ground truth to learn the appearance of shadows in images. For instance, Zhu *et al.* [25] propose a set of shadow variant

and invariant features in monochromatic images to learn a shadow region classifier, and refine the results with a CRF. Lalonde *et al.* [18] focused on shadow boundaries on the ground. They train a shadow boundary classifier (using Adaboost[9]) based on color and texture features, and combine it with scene layout cues from [12] using a CRF to encourage boundary continuity. Huang *et al.* [13] use a set of physically inspired features to train a shadow boundary pixel classifier using a SVM. They join pixels confidently predicted as shadow boundaries with weakly predicted adjacent pixels in a Canny-like manner. Shadow boundary methods ([13],[18]) achieve acceptable results but they fail to segment closed shadow contours consistently. Guo *et al.* [10] train two pairwise classifiers to find pairs of regions in an image that share the same material and are viewed under the same illumination conditions (both in shadow or both not in shadow), and same material but illuminated differently (only one region in shadow). They minimize an energy functional that combines the predictions of a single region classifier and the positive predictions of their pairwise classifiers. However, their single region classifier is not very accurate (especially for shadow regions). They only allow submodular pairwise potentials since they use graph cuts[17] to minimize the energy functional.

Our approach is also a learning-based method for shadow detection. We address the aforementioned challenges with the following contributions:

- A novel single region classifier with a multikernel model specifically tailored for shadow region classification with SVM that outperforms the state of the art.
- A novel boundary classifier for shadow boundaries cast over surfaces with the same material, and two improved paired regions classifiers.
- We pose shadow detection as an MRF binary labelling problem that combines our accurate single region classifier with both boundary cues and illumination relationships between pairs of regions of the same material obtained from precise classifiers.

We propose a single region shadow classifier based on a multikernel SVM. Our multikernel model is a linear combination of χ^2 and Earth Mover’s Distance(EMD)[22] kernels that operate on texture and color histograms disjointly. Multikernel SVMs have been employed successfully to integrate different feature sets, for instance Li *et al.* [19] use a linear combination of χ^2 kernels applied on different histogram features to regress a scoring function for figure-ground hypothesis in the context of object recognition.

Our shadow region classifier already outperforms the more complex state of art methods, without performing MRF/CRF optimization. However, the local appearance of a single region is often ambiguous. Even for a human observer it can be hard to discern if a region is in shadow or not without considering its context. Hence, it is sensible to look beyond the boundaries of a single region to decide its shadow label [10] [25]. In contrast to previous work we strive to use such contextual information sparingly. For MRF optimization reasons we prefer that most of the work is handled by the single region classifier (unary MRF potentials), with sparse pairwise connections that smooth the label changes across regions. Following this reasoning we build on the work of [10] to propose our own improved pairwise classifiers but constrained to adjacent regions: for pairs of regions sharing the same material and same illumination condition, and for same material pairs viewed under different illumination (first lit, second in shadow). We also propose a shadow boundary classifier. Since shadow boundaries often overlap with reflectance changes confounding the effects of

the illumination change, our classifier focuses on boundaries of shadows cast over surfaces with the same underlying material.

We integrate our single region classifier, our pairwise classifiers and our boundary classifier using an MRF. We pose single image shadow detection as a binary labelling problem. The MRF is a graph where the nodes represent the regions in the image. The edges in the graph are defined by the positive predictions of the two pairwise classifiers and the boundary classifier. That is, given two regions r_i and r_j , we add an edge between their corresponding nodes if: a) r_i and r_j are predicted to have the same material and same illumination, or b) r_i and r_j are predicted to have the same material but different illumination, or c) a boundary between r_i and r_j is predicted to be a shadow boundary cast over the same material surface. The key idea is to combine a very strong region classifier with a set of precise boundary and paired region classifiers that define the smoothing in the MRF. In order to achieve sparsity and increase reliability we favor precision over recall in the outcome of the context based classifiers.

The unary potentials are based on the confidences of the single region classifier. The pairwise potentials are based on the confidences of positive predictions from the two pairwise and the boundary classifiers. The same material/same illumination relationships induce submodular potentials. Potentials due to the boundary classifier are supermodular. The directionality of the same material different/illumination classifier introduces asymmetric potentials. We minimize the MRF energy using QPBO [17] [21], the asymmetric potentials led to a minor modification of the initial reparameterization stage of QPBO.

Experimental results on the dataset collected by [10] show that our method clearly outperforms the state of the art. Our single region classifier reduces the balance error rate (BER) by a 25% with respect to [10], correctly detecting 11% more shadow pixels. Moreover, with our MRF optimization we further reduce the BER to a 7.8%. That is a 53% decrease compared to the state of the art. Our total pixel accuracy is 94.1% (88% on shadow pixels).

The rest of the paper is organized as follows: Section 2 describes the pre-processing stage of our method, in Section 3 we define our single region, pairwise and boundary classifiers. In Section 4 we describe the final optimization step. In Section 5, we present our experimental results. And Section 6 concludes the paper.

2 Initial segmentation

In this section we describe how we segment the image into regions and how to extract region boundaries and pairs of adjacent regions.

2.1 Region Segmentation

The quality of the initial superpixel segmentation will affect the performance of our shadow detector. Ideally, we want to segment the image into superpixels that correspond to consistently illuminated regions. That is, either *all* pixels in a superpixel are in shadow or all are not in shadow. Furthermore, we would like to obtain a small number of superpixels to reduce the computational cost of later steps of our method. We propose a two step process that is fast and robust to the choice of thresholds. First, we apply SLIC [18] superpixel segmentation to oversegment the image, obtaining an initial set of superpixels. Then, we apply mean-shift clustering [9] over the superpixels' mean color in CIELAB color space. We merge adjacent

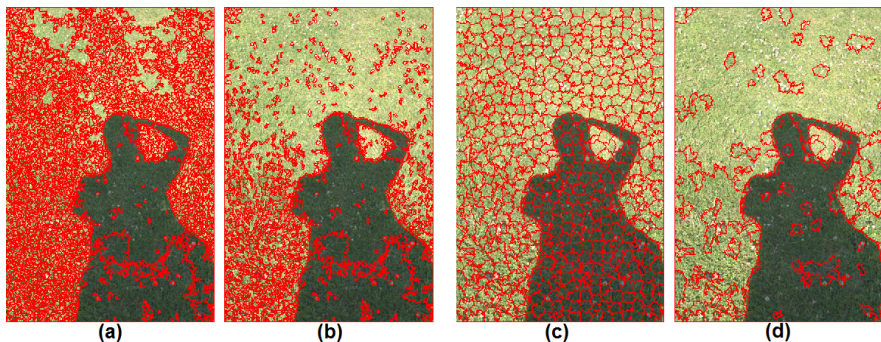


Figure 1: a) Mean-shift image segmentation with spatial bandwidth = 17 and feature bandwidth = 3. b) Merging superpixels from (a) still yield a very high amount of shadow region proposals that are small and spurious. c) SLIC superpixel segmentation. d) Merging superpixels from (c) yields a small number of shadow region proposals that are compact and accurate.

superpixels that belong to the same mean-shift clusters into a larger region. This operation can be thought of as reducing the over-segmentation of the superpixel method while preserving strong boundaries. By contrast, fine tuning is required to segment consistently illuminated superpixels with pixel level Mean-shift, which often leads to generation of many small regions. By reducing the over-segmentation of the superpixels we reduce the number of regions that need to be considered which results in less processing time in subsequent stages. We reduce the number of regions by a factor of five in our experiments. Figure 1 shows Mean-shift segmentation (a,b) compared to our pre-processing result (c,d), and that our merging approach retains all the important regions and significantly reduces the number of shadow region proposals.

2.2 Image connectivity graph

The refined superpixel segmentation defines the set of regions that will eventually be labelled as shadow or non shadow. We extract boundaries in the image as the shared boundary pixel segments between regions. The image connectivity graph is defined as follows. Each node represents a region, edges denote that two regions are adjacent. That is, there exists at least one boundary segment in common. This way, each edge connects two adjacent regions, and also uniquely identifies a set of boundary segments between those two regions.

3 Shadow Detection

In this section we describe our proposed classifiers and the features they use: a unary classifier for shadow regions, two pairwise classifiers that operate on pairs of regions, and a boundary classifier for boundaries of shadows cast over the same material.

3.1 Unary classification

In order to determine if a region is in shadow we will look at its chromatic, intensity and textural properties. For each region, we compute a histogram with 21 bins for each of the

components(L^*, a^*, b^*) of the perceptually uniform color space CIELAB. To represent texture, we compute a 128 bins *texton*[10] histogram. Shadow regions tend to be less textured and darker. The CIELAB color space has been shown to perform well for shadow edge identification in outdoor scenes [15] as well as to improve reflectance segmentation [9]. The two color opponent channels behave differently under illumination changes. Especially in outdoor environments, the b^* channel (yellow-blue) is more sensitive to shadows than the a^* channel (red-green), which is shadow invariant to a certain degree [23].

To compare textures between regions we use the χ^2 distance between their *texton* histograms. To compare the color histograms we find it more appropriate to use the Earth Mover’s Distance (EMD)[22]. Neighboring bins in the L^*, a^*, b^* histograms represent proximate values and their ground distance is uniform (property of the CIELAB space), in contrast to *texton* histograms. Furthermore, EMD is more accurate in measuring distances between histograms of continuous entities (such as L, a, b), as it is less sensitive to quantization error and it can be efficiently computed for 1D histograms. The kernel for an SVM provides a notion of similarity in the feature space. Since our features are normalized histograms (unit mass), both the EMD and the χ^2 distance are metrics. Hence, we can use them in the form of generalized Gaussian distances[24] to provide suitable kernels([24]) for SVMs:

$$K(x, y) = \exp(-\gamma D(x, y)), \quad (1)$$

where $D(x, y)$ can be $\chi^2(x, y)$ or $EMD(x, y)$. To combine the 4 histogram feature sets (*texton*, L^* , a^* , b^*) we propose a multikernel model consisting of a weighted sum of χ^2 and EMD kernels, each operating on the appropriate disjoint feature subset. That is, an EMD kernel for each L^* , a^* and b^* histogram and a χ^2 kernel for the *texton* histogram. Given two regions x , y with features encoded as $x = \{x_L, x_a, x_b, x_t\}$ and $y = \{y_L, y_a, y_b, y_t\}$, the kernel model is:

$$K(x, y) = \sum_{i \in \{L, a, b, t\}} \exp(-\gamma_i D_i(x_i, y_i)) \quad (2)$$

where $D_t = \chi^2$ and $D_L = D_a = D_b = EMD$. In order to avoid the expensive combinatorial search for the γ_i parameters we propose the following heuristic. We sample a set of pairs of similar shadow regions and a set of pairs of similar non shadow regions from the training data. Then, we compute the variance of the *EMD* and χ^2 distances between each one of the L^* , a^* , b^* and *texton* histograms of similar regions. Based on these results, we set the relative coefficients c_i for each γ_i parameter to be inversely proportional to its empirical variance. This way, given a canonical $\hat{\gamma}$ value we set $\gamma_i = c_i \hat{\gamma}$.

We implemented the stacking of multiple *EMD* and χ^2 kernels within the libSVM([8]) version of [19]. Results in Section 5 outperform the state of the art, suggesting that measuring separately *texton*, L^* , a^* , b^* histogram similarities with χ^2 and *EMD*, respectively, provide suitable similarity metrics since a good separation boundary is created in the feature space.

3.2 Looking into context

Local appearance of a region is often ambiguous. Even for human observers, it may be hard to determine if a region is in shadow without looking at its context. Hence, it is sensible to look beyond the boundaries of a single region to decide if it is in shadow or not. We propose to integrate information from illumination relationships between paired regions (as in Guo *et al.* [10]) and from a shadow boundary classifier with our single region classifier. We train two pairwise region classifiers, one for regions sharing the same material and same

illumination conditions, and another one for regions of same material but under different illumination. These classifiers are described in the following subsections. Confident positive predictions of the pairwise and boundary classifiers are used to define the pairwise potentials and the graph topology of an MRF where the unary potentials are defined based on the single region classifier. We want to minimize the following functional:

$$E(\mathbf{x}) = \sum_{i \in \mathbf{R}} \phi(x_i) + \sum_{i, j \in \Omega^s} \psi_s(x_i, x_j) + \sum_{i, j \in \Omega^d} \psi_d(x_i, x_j) + \sum_{i, j \in \beta} \psi_b(x_i, x_j)$$

where Ω^s , Ω^d denote pairs of regions of the same material confidently predicted to have same and different illumination respectively, and β is the set of confidently predicted shadow boundaries.

3.2.1 Pairwise classifiers

Inspired by Guo *et al.* [10] we define two pairwise region classifiers (same material and same illumination, same material/different illumination). The goal is to identify illumination relationships between pairs of regions sharing the same material. Motivated by the good performance of our single region classifier, we propose to enrich the feature set used in [10] with the EMD distances computed between each of the L^* , a^* , b^* histograms of the two considered regions. We restrict our work to adjacent regions. For each pair of adjacent regions, the following features are computed:

- χ^2 distance between each region's texton histogram.
- EMD between each region's L^* , a^* and b^* histograms respectively.
- Average RGB ratios. Given two regions i and j , compute the ratios of average region intensity for each R, G and B channels: $\rho_R = R_i / R_j$, $\rho_G = \rho_R / \rho_G$ and $\rho_B = B_i / B_j$.
- Chromatic alignment computed as: ρ_R / ρ_G and ρ_G / ρ_B .

Intuitively, regions of the same underlying material should portray very similar chromatic, lightness and texton distributions. Slight shifts on the lightness distribution may occur due to shading. Also, all the three color ratios should oscillate around unity. Under different illumination conditions, two regions of the same material should still retain some texture similarity, while they would differ significantly in lightness and color intensity. The chromaticities of the scene's illuminant(s) determine the degree of chromatic resemblance between the shadow and non-shadow regions of the same material. For example, as mentioned in Section 3.1, in outdoors scenes we expect the differences in b^* to be stronger than the differences in a^* . We expect the EMD's on a^* , b^* and the color ratios that are caused solely by an illumination change to be learnable with respect to the EMD's and ratios due to reflectance changes between the pair of regions. The 9-dimensional feature vectors are input to a probabilistic SVM with a gaussian RBF kernel that we tune for precision.

3.2.2 Boundary classifier

Shadow boundaries may overlap with reflectance changes confounding the effects of the illumination changes. Hence, we propose a classifier that focuses only on the boundaries of shadows cast over surfaces with the same underlying material. For each boundary segment

we compute a set of color ratios [2] computed at different scales on different color spaces, namely RGB, CIELAB and [4]. These features are computed in a different manner by [18].

For each boundary pixel two opposite half gaussian filters are overlaid to lay on each side of the boundary. The ratio is calculated between the two gaussian weighted sums as $\frac{f_d(p)}{f_l(p)}$. The numerator, $f_d(p)$, is the result of applying the half filter on the darker side of the boundary at pixel p . The denominator, $f_l(p)$, is the result of the filter applied on the lighter side of the boundary at p . For each color space channel, we compute the ratio at 4 different scales $\sigma^2 = \{1, 2, 3, 8\}$ and size $2\sigma^2$, obtaining 36 ratios. At the same 4 scales, we also calculate the χ^2 distance between the texton histograms of each side, and the difference in skewness of pixel intensities as suggested by [25]. Finally, we include the filter responses over the intensity channel for the dark side of the pixel. This way, we obtain a 48-dimension pixel feature vector. The feature vectors of all the pixels that belong to the same boundary are averaged to obtain that boundary's features. We determine the darker side as the one with the lowest sum of average lightness(L*) computed all along the boundary.

4 Optimization

We pose shadow detection in a single image as a binary labelling problem. Given a segmented image \mathbf{R} , each region $r_i \in \mathbf{R}$ will be labelled as shadow $x_i = 1$, or non shadow $x_i = -1$. The MRF is defined over a graph with a node x_i for each region of the image. We establish three sets of edges: 1) Ω^s connecting the nodes x_i and x_j if the corresponding pair of regions (r_i, r_j) is predicted as positive by the same material same illumination classifier, 2) Ω^d connecting the nodes x_i and x_j if the corresponding pair of regions is predicted as positive by the same material different illumination classifier (notice that this relationship is directional, r_i is the non shadow region and r_j is shadow), 3) β connecting the nodes x_i and x_j if a boundary between the regions is predicted as positive by the boundary classifier. We want to minimize the following functional:

$$E(\mathbf{x}) = \sum_{i \in \mathbf{R}} \phi(x_i) + \sum_{i,j \in \Omega^s} \psi_s(x_i, x_j) + \sum_{i,j \in \Omega^d} \psi_d(x_i, x_j) + \sum_{i,j \in \beta} \psi_b(x_i, x_j) \quad (3)$$

The unary potential $\phi(x_i)$ encourages agreement between the node's labelling and the prediction of the single region classifier: $\phi(x_i) = -\omega_i P(r_i = x_i)$. The pixel area of region i is denoted by ω_i .

$$\psi_s(x_i, x_j) = \begin{cases} \omega_{ij} P^s(r_i, r_j = 1), & x_i \neq x_j \\ 0, & x_i = x_j \end{cases}$$

$$\psi_d(x_i, x_j) = \begin{cases} 0, & x_i = -1, x_j = 1 \\ \omega_{ij} P^d(r_i, r_j = 1), & \text{otherwise} \end{cases}$$

$$\psi_b(x_i, x_j) = \begin{cases} 0, & x_i \neq x_j \\ \omega_b P^b(r_i, r_j = 1), & x_i = x_j \end{cases}$$

The pairwise potentials $\psi_s(x_i, x_j)$, $\psi_d(x_i, x_j)$, $\psi_b(x_i, x_j)$ penalize label assignments that contradict the positive predictions of the respective classifiers. $P^c(r_i, r_j = 1)$, with $c \in \{s, d, b\}$, denotes the confidence of the positive prediction by the respective classifier. ω_{ij} is the geometric mean of the areas of regions i and j . ω_b is the geometric mean of the areas of the two regions connected by the boundary b .

Guo <i>et al.</i> [10]	Sha.	Non
Shadow(GT)	0.716	0.284
Non-shadow(GT)	0.048	0.952
Single region	Sha.	Non
Shadow(GT)	0.795	0.457
Non-shadow(GT)	0.043	0.957
MRF ψ_b	Sha.	Non
Shadow(GT)	0.840	0.160
Non-shadow(GT)	0.089	0.958
MRF $\psi_s\psi_d$	Sha.	Non
Shadow(GT)	0.862	0.284
Non-shadow(GT)	0.037	0.963
MRF	Sha.	Non
Shadow(GT)	0.880	0.120
Non-shadow(GT)	0.038	0.962

	Acc.	BER	Sha.	Non
Guo <i>et al.</i> [10]	0.891	0.166	0.716	0.952
single region	0.916	0.124	0.795	0.957
MRF ψ_b	0.926	0.101	0.840	0.958
MRF $\psi_s\psi_d$	0.937	0.087	0.862	0.963
MRF	0.941	0.078	0.880	0.962
single-MS	0.901	0.157	0.725	0.961
MRF-MS	0.909	0.138	0.766	0.958

Table 1: Shadow detection quantitative results.

The energy function (3) is not submodular. $\psi_b(x_i, x_j)$ introduces supermodular pairwise interactions. We optimize it using QPBO [16] [17]. Notice that $\psi_d(x_i, x_j)$ introduces asymmetric pairwise interactions. Namely, (following notation introduced in [16] [17]) for certain edges ($p \rightarrow q$) we have that: $\theta_{p,q;10} \neq \theta_{p,q;01}$. So, we introduce a new first step in the initial reparameterization stage of QPBO in order to restore the symmetry property. For each of the asymmetric edges we perform the following adjustment:

$$\begin{aligned}
 \theta_{p;1} &\leftarrow \theta_{p;1} + \theta_{p,q;10} \\
 \theta_{q;0} &\leftarrow \theta_{q;0} + \theta_{p,q;10} \\
 \theta_{p,q;10} &\leftarrow \theta_{p,q;01}
 \end{aligned} \tag{4}$$

5 Experimental results

We evaluate our shadow detection performance on the dataset collected by Guo *et al.* [10]. The training set contains 32 images for which we had to manually provide labels for shadow boundaries and material properties based on the initial segmentation described in Section 2. The testing set consists of 76 images, with pixel level shadow masks provided by [10]. In our experiments we compare to the more recent work of Guo *et al.* [10] which improves [10].

Single region results Since the training data is unbalanced, 72.4% of regions non-shadow versus 27.6% of shadow, we set different class weights for the training of the single region classifier. We compute the weight for the shadow class as the squared root of the ratio between number of non-shadow and shadow training instances. We perform grid search on the SVM’s cost parameter C and the canonical $\hat{\gamma}$ with 5-fold cross-validation to train our multikernel single region classifier. As we can see in table 5, we outperform the state of art by 2.5 percentage points in total pixel accuracy. We detect 11% more shadow pixels correctly, furthermore we obtain a 25% reduction in balanced error rate (BER) for a final BER of 12.4%.



Figure 2: Role of segmentation. a) Failure due to inadequate Mean-shift segmentation, the shadow on the concrete is missed. b) Results with our segmentation for (a). c) Problematic case for our initial segmentation. d) Detection results of our method for image (c).

For the two paired region classifiers and the boundary classifier we trained and tested SVM’s with gaussian RBF kernels on scaled data. We first compute the empirical mean and standard deviation per feature component of the training data. Then, we scale both the training and testing data by subtracting the mean and dividing by the standard deviation (z-scores). In order to properly set the γ parameter of the RBF kernel and the cost parameter C , we conducted grid search with 5-fold cross validation using the F_β -score (weighted harmonic mean between precision and recall). We set β to 0.25 to favor precision over recall.

Full method results Our full method (**MRF**) achieves an outstanding 94.1% of pixel accuracy. We detect 88.0% of the shadow pixels, that is a 22.9% increase with respect to [10]. We further reduce the BER to 7.8% (a 53% decrease). In table 5, we also show the results of running the optimization with only the pairwise potentials from boundary cues (**MRF- ψ_b**), and with just the potentials from the paired regions classifiers (**MRF- ψ_s, ψ_d**). We can observe how the boundary (ψ_b) and the paired regions potentials (ψ_s, ψ_d) present only a certain degree of redundancy, as we achieve the best results by combining both in our full model (**MRF**).

Finally, to assess the role of the initial segmentation in shadow detection, we experimented with the same Mean-shift segmentation used in [10] [10]. We trained and tested our single region classifier with the Mean-shift segmented images (single-MS). We also trained our pairwise classifiers using the labels provided by [10] and run our optimization without the boundary potentials ψ_b , (**MRF-MS**). In figure 3 (a) we can see a failure case where Mean-shift fails to segment the shadow on the concrete (smooth penumbra transition). Our segmentation presents few failure cases such as in figure 3 (c)(d).

6 Conclusion

In this paper we presented a learning based MRF optimization method for single image shadow detection. First, we proposed a powerful single region classifier that already performs better than more complex methods. Then, we incorporate contextual cues from illumination changes across region boundaries and from illumination relationships between pairs of same material adjacent regions into the MRF. The results of our full MRF model outperformed the state-of-the-art in the most recent shadow detection dataset, with a marked reduction in shadow pixel misclassification. In future work we will explore the effects of additional segmentation methods as well as shadow removal applications.

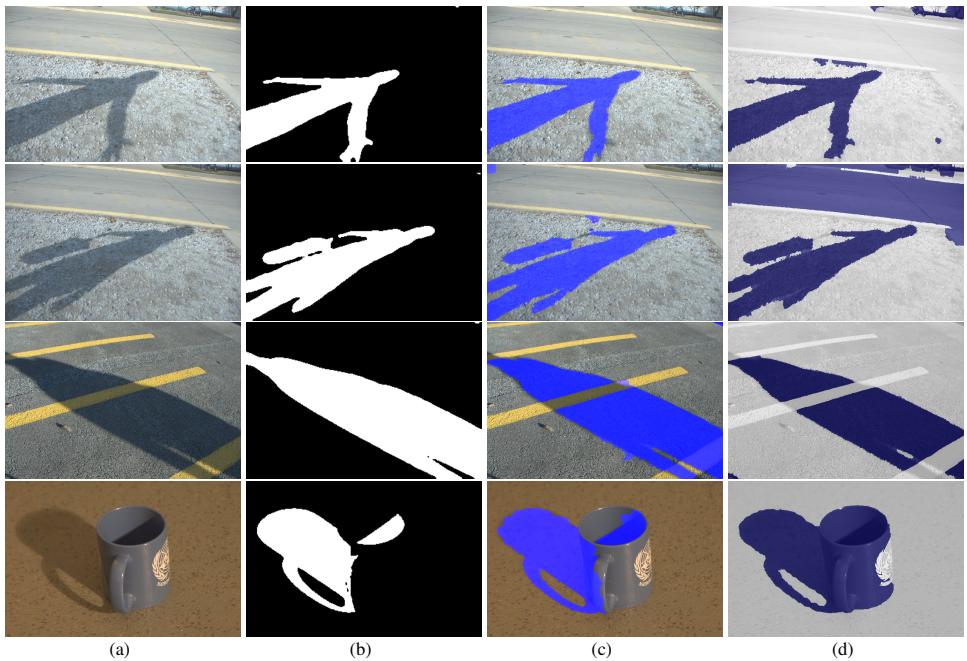


Figure 3: Comparison of shadow detection results. a) Input image. b) Ground truth shadow pixel mask. c) Our results overlaid in light blue d) Guo *et al.* [10] results in dark blue.

Acknowledgements. This work was supported by NSF IIS-1111047 and IIS-1161876 and the SUBSAMPLE Project from the DIGITEO Institute, France.

References

- [1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE TPAMI*, 2012.
- [2] Kobus Barnard and Graham Finlayson. Shadow identification using colour ratios. In *in IS&T/SID 8th Colour Imaging Conference: Colour Science, Systems and Appl*, pages 97–101, 2000.
- [3] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [4] Hamilton Y. Chong, Steven J. Gortler, and Todd Zickler. A perception-based color space for illumination-invariant image processing. In *ACM SIGGRAPH 2008 papers*, SIGGRAPH '08, pages 61:1–61:7, New York, NY, USA, 2008. ACM.
- [5] Michael Collins, Robert E. Schapire, and Yoram Singer. Logistic regression, adaboost and bregman distances. *Mach. Learn.*, 48(1-3):253–285, September 2002. ISSN 0885-6125. doi: 10.1023/A:1013912006537.

- [6] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002. ISSN 0162-8828.
- [7] G.D. Finlayson, S.D. Hordley, Cheng Lu, and M.S. Drew. On the removal of shadows from images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(1): 59–68, jan. 2006. ISSN 0162-8828. doi: 10.1109/TPAMI.2006.18.
- [8] Graham Finlayson, Mark Drew, and Cheng Lu. Entropy minimization for shadow removal. *International Journal of Computer Vision*, 85:35–57, 2009. ISSN 0920-5691. 10.1007/s11263-009-0243-z.
- [9] Elena Garces, Diego Gutierrez, and Jorge Lopez-Moreno. Graph-based reflectance segmentation. In *Proceedings of SIACG 2011*, 2011.
- [10] Ruiqi Guo, Qieyun Dai, and D. Hoiem. Single-image shadow detection and removal using paired regions. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2033–2040, june 2011. doi: 10.1109/CVPR.2011.5995725.
- [11] Ruiqi Guo, Qieyun Dai, and Derek Hoiem. Paired regions for shadow detection and removal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99 (PrePrints):1, 2012. ISSN 0162-8828.
- [12] Derek Hoiem, Alexei A. Efros, and Martial Hebert. Recovering surface layout from an image. *Int. J. Comput. Vision*, 75(1):151–172, October 2007. ISSN 0920-5691.
- [13] Xiang Huang, Gang Hua, J. Tumblin, and L. Williams. What characterizes a shadow boundary under the sun and sky? In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 898–905, nov. 2011. doi: 10.1109/ICCV.2011.6126331.
- [14] Feng Jing, Mingjing Li, Hong jiang Zhang, and Bo Zhang. Support vector machines for region-based image retrieval. In *IEEE Int. Conf. on Multimedia & Expo*, pages 2–6, 2003.
- [15] E.A. Khan and E. Reinhard. Evaluation of color spaces for edge classification in outdoor scenes. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 3, pages III–952–5, 2005. doi: 10.1109/ICIP.2005.1530551.
- [16] Vladimir Kolmogorov and Carsten Rother. Minimizing non-submodular functions with graph cuts - a review. Technical report, TPAMI, 2007.
- [17] Vladimir Kolmogorov and Ramin Zabih. What energy functions can be minimized via graph cuts? *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(2):147–159, 2004.
- [18] Jean-François Lalonde, Alexei A. Efros, and Srinivasa G. Narasimhan. Detecting ground shadows in outdoor consumer photographs. In *Proceedings of the 11th European conference on Computer vision: Part II, ECCV’10*, pages 322–335, Berlin, Heidelberg, 2010. Springer-Verlag. ISBN 3-642-15551-0, 978-3-642-15551-2.
- [19] Fuxin Li, Joao Carreira, and Cristian Sminchisescu. Object recognition as ranking holistic figure-ground hypotheses. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2010.

- [20] David R. Martin, Charless C. Fowlkes, and Jitendra Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(5):530–549, May 2004. ISSN 0162-8828. doi: 10.1109/TPAMI.2004.1273918.
- [21] C. Rother, V. Kolmogorov, V. Lempitsky, and M. Szummer. Optimizing binary mrfs via extended roof duality. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on, 2007*.
- [22] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. A metric for distributions with applications to image databases. pages 59–66, 1998.
- [23] Tom Troscianko, R. Baddeley, C. Alejandro Parraga, and U. Leonards. Visual encoding of green leaves in primate vision, May 2003.
- [24] Vladimir N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995. ISBN 0-387-94559-8.
- [25] Jiejie Zhu, K.G.G. Samuel, S.Z. Masood, and M.F. Tappen. Learning to recognize shadows in monochromatic natural images. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 223–230, june 2010.