

Shadow Detection based on Colour Segmentation and Estimated Illumination

Xiaoyue Jiang¹

x.y.jiang@bham.ac.uk

Andrew J. Schofield¹

a.j.schofield@bham.ac.uk

Jeremy L. Wyatt²

jlw@cs.bham.ac.uk

¹ School of Psychology

University of Birmingham

Birmingham, UK

² School of Computer Science

University of Birmingham

Birmingham, UK

Abstract

In this paper we show how to improve the detection of shadows in natural scenes using a novel combination of colour and illumination features. Detecting shadows is useful because they provide information about both light sources and the shapes of objects thereby illuminated. Recent shadow detection methods use supervised machine learning techniques with input from colour and texture features extracted directly from the original images (*e.g.* Lalonde *et al.* ECCV 2010, Zhu *et al.* CVPR 2010). It seems sensible to augment these with estimates of scene illumination, as can be obtained with an intrinsic image extraction algorithm. Intrinsic image extraction separates the illumination and reflectance components in a scene, and the resulting illumination maps contain robust intensity change features at shadow boundaries. In this paper, we make two main contributions. First we improve upon existing methods for extracting illumination maps. Second we show how to use these illumination maps together with colour segmentation to extend the Lalonde's approach to shadow detection. Illumination maps are extracted using a steerable filter framework based on global and local correlations in low and high frequency bands respectively. The illumination and colour features so extracted are then input to a decision tree trained to detect shadow edges using AdaBoost. We tested variations of our proposed approach on two public databases of natural scenes. This study showed that our approach improves on that of Lalonde both in terms of sensitivity to shadow edges and rejection of false positives. Following Lalonde we show that our detection results are further improved by imposing an edge continuity constraint via a conditional random field (CRF) model.

1 Introduction

Shadows are unavoidable in natural images. They are useful because they can provide information about both the light source [10, 11] and the shape of objects [12, 13], but they also degrade the performance of algorithms for tasks such as segmentation, object detection, tracking, and shape reconstruction. In all of these cases it is beneficial to detect shadows in images. Shadow detection is still an open problem. In this paper we improve on existing supervised learning methods for shadow detection by incorporating an illumination map –

extracted using a correlation-based approach we developed recently [10] – alongside colour features. These features are given as input to a decision tree trained using AdaBoost [9]. Our method outperforms the approach of Lalonde *et al.* [11] for both shadow edge detection and ground shadow edge detection on a database of images from Zhu *et al.* [25].

1.1 Background

A shadow is caused by an obstruction between the light source and the target surface. For a point light source, the intensity of a shadow region will be (relatively) constant and determined by the amount of ambient light. However, real light sources are seldom well modelled as infinitely small point sources. In fact, most have finite area. Consequently, shadows can be divided into umbra, penumbra and antumbra regions according to the degree, and nature of the obstruction. In common with the literature we use the term shadow to refer to umbra regions: those wholly occluded from the light source.

Recently, shadow detection methods based on multiple images [6, 8, 12, 14] have been developed for tracking systems. Weiss [24] averaged the gradient profiles of external images for the same scene taken at different times to remove shadows (external shadows move with the sun). Bousseau [5] estimated illumination in single images through interactive user input. Others have used colour features to estimate shadows in single images without user input. For example, Vazquez *et al.* [23] estimated the colour distributions inside and outside shadow regions to remove shadows. Lalonde *et al.* [11] used colour ratios across edges in three different colour spaces at different scales to discriminate shadow edges from reflectance edges. Finlayson [9] used intensity entropy as a criterion to find the optimal projection angle in a normalized RGB colour space to remove shadows. Others have used local texture/intensity features to detect shadows (for example Zhu *et al.* [25]).

The illumination map of an image provides information not available to most of the shadow detection methods described above. It describes all the changes in illumination within a scene, including shading caused by changes in surface orientation, and incident illumination changes such as attached and cast shadows. However, extracting illumination maps from images is itself a challenging task [1, 13, 19, 22], and thus so far such maps have not been used widely in shadow detection. Jiang *et al.* [8] proposed a steerable filter framework to derive illumination and reflectance maps of a scene from a single image based on the relationship between luminance and texture, colour, and local contrast. We extend this framework to improve upon existing methods for cast shadow detection in two ways. First Jiang *et al.*'s algorithm works well for low- but not high-frequency components of the original image. Since shadow edges are typically high frequency changes in the illumination map we extend their method to work well at high frequencies. Second, we extend the method of Lalonde *et al.* for shadow detection by incorporating the features extracted from the illumination map as additional inputs to a decision tree that has been trained using supervised learning. The paper is structured as follows. First we outline our algorithm. In Section 2 we describe the extraction of colour edge features which is the basis of Lalonde *et al.*'s approach, and which is also utilised here. In Section 3 we describe how we extend the illumination map estimation technique of Jiang *et al.* Next (Section 4) we describe how we perform shadow edge detection by combining the colour features from the first approach with illumination features from the second approach. Finally (Section 5) we present experimental analysis of a number of variations of our method and that of Lalonde *et al.*

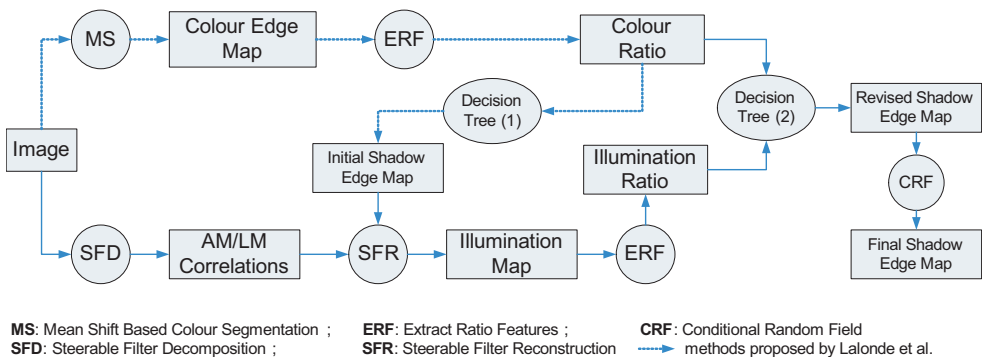


Figure 1: A flow chart giving a high level overview of our proposed algorithm.

1.2 Outline of algorithm

In this paper, intensity features from illumination maps [2] and colour features [4, 11] from original images are combined to address the problem of shadow detection. The advantage of adding illumination features is that luminance changes due to changes in albedo will have already been removed from such maps. Thus these maps provide robust features for detecting shadow regions. However, the approach of Lalonde *et al.* only extracted colour ratio and texture features from an image to use as inputs to a supervised learning method.

The structure of our approach is captured in Figs 1 and 2. The upper path corresponds broadly to the approach of Lalonde *et al.* without their texture features. When trained the decision tree (Decision Tree 1) produces an initial shadow edge map. We use this to help improve the illumination map estimation process which is itself an improved version of the algorithm of Jiang *et al.*, and is the lower path in Fig. 1.

The illumination estimation process is shown in Fig. 2. We improve on Jiang *et al.*'s intrinsic image extraction framework [2] in order to obtain more accurate illumination maps in higher frequency bands. Jiang *et al.* look for global correlations between AM (*Amplitude Modulation* or local luminance contrast) and LM (*Luminance Modulation* or average local luminance) to spot the difference between reflectance changes (where AM and LM are uncorrelated) and illumination changes (where AM and LM are correlated). The global strategy works well at low frequencies, but not for high frequency changes. We therefore calculate local correlations between LM and AM, and separately between LM and colour features. Both these help us to separate high-frequency filter responses, at a local level, for later reconstruction into either illumination or reflectance maps. For the lower-frequency bands, the global correlation between luminance and contrast in each band still determines the corresponding weights for reconstruction [2].

Having improved the separation of intrinsic images in the lower path, we then extract features from the illumination map. Using these illumination ratio features and the colour features extracted in the upper path, a second decision tree (also trained using AdaBoost) produces an improved estimate of the shadow edge probability map. This more effectively discriminates shadow edges from reflectance edges. This revised shadow edge probability map can then be improved further using a Conditional Random Field model with constraints that edges should be continuous and that neighbouring pixels should have the same state.

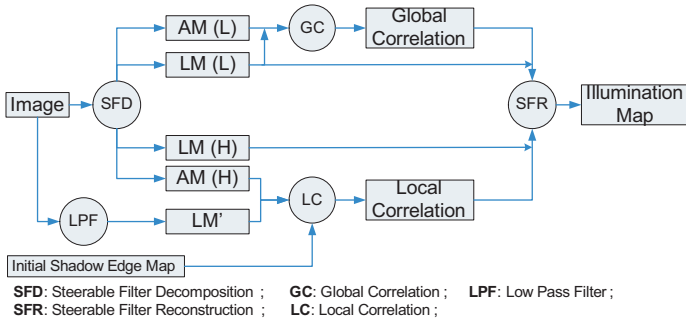


Figure 2: A flow chart giving detail on the improved approach for estimating the illumination map. This corresponds to the first two stages of the bottom row in Fig. 1. LM' is calculated directly from the undecomposed image. H and L refer to the outputs of steerable filters in High and Low frequency bands respectively. The SFD is performed in up to 8 frequency and 8 orientation bands.

2 Colour segmentation and colour feature extraction

We now detail the colour segmentation and feature extraction method. Colour features are used to aid the identification of shadows directly, and to improve the illumination maps produced by the steerable filter method. We initially produce a colour segmentation using the mean shift algorithm [10] in RGB space. The edges of the same colour region will be labelled as shadow or non-shadow edges as shown later.

Having determined the locations of colour edges we extract colour ratios (CR) across edges. The ratio is between the minimum and maximum values in different colour channels across an edge [10]. Shadow edges are due to illumination changes so the illumination ratio for regions either side of shadow edges should be the same everywhere in the image, hence hue should be roughly constant across shadow boundaries. Non-shadow edges will not be so constrained. It is unclear which colour space is the best to estimate these ratios, thus we calculate colour ratios in 3 different spaces (RGB, HSV and LAB) and at 4 different scales. Altogether, we derive a 36-dimensional vector for each edge pixel in the image. Following Lalonde *et al.*, we then use an AdaBoost-based decision tree, trained on a set of labelled images, to classify edges as shadow or non-shadow (see Section 5). This decision tree gives the probability that each edge pixel found by the mean shift algorithm is a shadow edge. The result of applying the decision tree is the initial shadow edge probability map which is used as input to the illumination map estimation process (by helping to separate reflectance and illumination changes, see Section 3). The colour features, meanwhile are also input directly into the final shadow detection process along with the illumination features.

3 Illumination estimation

Intensity values in natural images are the product of reflectance and illumination, and do not always represent the intrinsic features of the scene, such as the illumination profile, surface orientations or surface colours. It can be advantageous to separate illumination changes from reflectance changes: a type of intrinsic image extraction. Jiang *et al.* [10] proposed a steerable filter-based framework to measure the correlation between luminance and local contrast, tex-

ture, and colour in each frequency and orientation band, then used the resulting correlation coefficients to derive weights for each component of the steerable filter output in order to reconstruct the illumination and reflectance maps. Due to its global nature, this method works well on low but not on high-frequency bands because the origin and hence appropriate classification of high frequency components tends to vary across the image. Thus, we proposed to calculate local correlations in each band and apply these correlations to derive local rather than global weights. At the end of this section we show how if we calculate both local and global correlations we can automatically determine the frequency band at which to shift from one to the other, we refer to this as a hybrid correlation method.

3.1 Local correlation between LM and AM

When illumination changes fall across a visual texture the luminance difference (amplitude, absolute contrast) between the light and dark parts of the texture varies with the illumination. This positive correlation between illumination and local amplitude can be used to identify shadow edges [L8]. Suppose an image has been decomposed into multiple frequency and orientation bands I_{ij} ($i = 1, \dots, N; j = 1, \dots, M$) where N is the number of frequency bands and M the number of orientation bands. Modulations of local amplitude (AM) can be represented as the envelope of the high frequency components in I_{ij} . In practice, we can extract the low frequency information from the amplitude of the high frequency components in the original image, i.e. $F_l(abs(I_{ij}))$, where F_l is a low-frequency filter. Similarly, local luminance modulation (LM) can be directly extracted by low-frequency filtering, i.e. $F_l(I)$. If we have already segmented the image into regions based on colour (e.g. Fig.4(b)) we need only examine correlations between LM and AM in local regions on either side of colour edges. Let us denote the set of colour edge pixels in the colour segmentation to be $D = \{d_1, \dots, d_k, \dots, d_K\}$, so that d_k indexes a point in any the original image I , or any of the filter outputs I_{ij} . Colour in this case comprises both luminance and hue, so the colour edge pixels will be a superset of the shadow edge pixels. Thus we need only examine the regions around these colour edges to find the shadow edges. If the correlation between LM and AM is strong enough, then we assign the associated edge components to the illumination map, otherwise they are assigned to the reflectance map. The local correlation – for a pixel d_k – between AM and LM for the i^{th} frequency band and the j^{th} orientation band is denoted by $W_{ij}^{AM}(d_k)$ and is defined as

$$W_{ij}^{AM}(d_k) = cor[F_l^{d_k}(abs(I_{ij})), F_l^{d_k}(I)] \quad (1)$$

where $W_{ij}^{AM}(p) = 0$ for any pixel $p \notin D$, $cor[s_1, s_2]$ calculates the correlation between two components s_1 and s_2 , and $F_l^{d_k}(\cdot)$ is the output of the filter around edge pixel d_k .

To calculate the local correlations we must be careful to select the appropriate region of the input image. The high frequency outputs of the steerable filter pyramid will always have large values in the region of edges in the original image, as shown in Fig.3(c). Therefore if AM is calculated in a local window including these edge pixels the correlation with LM will be compromised. In practice, the values of LM and AM in regions adjacent to each edge are the real concern. Thus an oriented rectangular window is used to calculate the local correlation. As shown in Fig.3(d), local correlation is calculated for the two regions (shaded region in Fig.3(d)) either side of the edge region, but the weights are applied to the central edge region. For component I_{ij} which is the response of steerable filter in direction $(j-1)\pi/M$, the rectangular window should be oriented at $\pi/2 - (j-1)\pi/M$; orthogonal to the filter. Fig.4(h) shows the local correlation between LM and AM in one frequency and

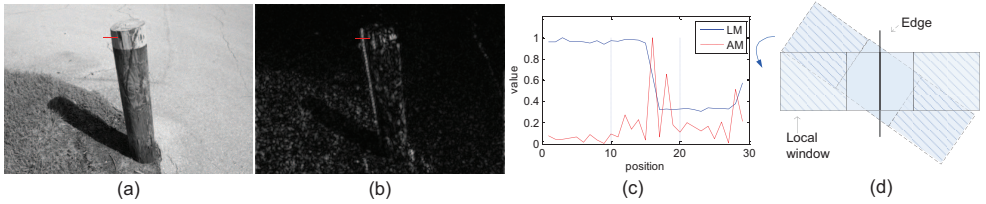


Figure 3: Local correlation between LM and AM (a) original image, red line indicates the region sampled in (c). (b) AM content at frequency level 1 and vertical orientation. (c) LM (blue line) and AM (red line) in this frequency/orientation band. The vertical blue lines represent the internal boundaries of the rectangular window (d) used to determine local correlation. The correlation between LM and AM is calculated using the two hatched regions, not the central region. The coarsely dashed box represents a window rotated to match a different orientation band.

orientation band. From the results, we can see the local correlation between LM and AM indicates changes in illumination, but may also represent reflectance changes that happen to share features with shadows. The initial shadow edge probability map is used to remove these reflectance changes from the illumination map. We now describe how this is done.

3.2 Shadow edges refine the illumination map

As described in Section 2, a decision tree based classifier is used to estimate the probability of a luminance edge being a shadow edge pixel based on the colour features. We denote this probability as $W^c(d_k)$ for the edge pixel d_k . We now want to produce a reconstructed illumination map based on the local correlation and this initial shadow edge map. The weights of each component in this reconstruction are different for each edge pixel. We denote these weights $W_{ij}^l(d_k)$, where l refers to the fact that the reconstruction is partly based on the local correlation (as opposed to the global correlation). Simply put it is the sum of $W^c(d_k)$ and the local correlation $W_{ij}^{AM}(d_k)$, for each edge pixel d_k ,

$$W_{ij}^l(d_k) = W^c(d_k) + W_{ij}^{AM}(d_k) \quad (2)$$

Intuitively, the Decision Tree 1 based on colour features indicates that, for a shadow edge, we should allocate the edge to the illumination map rather than to reflectance. $W_{ij}^l(p) = 0$ for any pixel $p \notin D$.

3.3 Hybrid correlation

The higher frequency bands in the steerable filter decomposition contain fine details in the image, and are dominated by local features. Therefore a local correlation strategy is better for the high frequency bands than the global strategy which is better for lower frequency components. In our hybrid strategy we thus combine the local and global strategies prior to reconstructing the reflectance and illumination maps. The reconstruction of the illumination map proceeds as follows:

$$IM = \sum_{i=1}^{T-1} F_{inv}(W_{ij}^g \times I_{ij}) + \sum_{i=T}^N F_{inv}(W_{ij}^l \times I_{ij}) \quad (3)$$

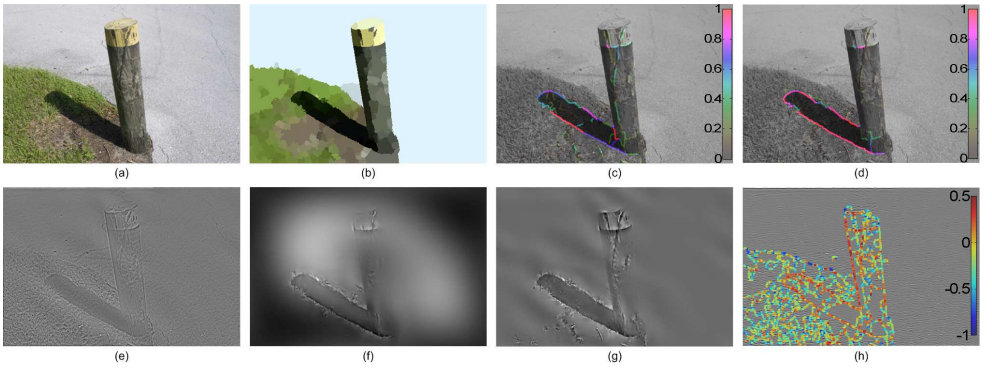


Figure 4: Intermediate results. (a) original image. (b) colour segmentation result. (c) initial shadow edge map. (d) final shadow edge map. (e) illumination map (IM) reconstructed with global correlation. (f) IM from local correlation. (g) IM from hybrid correlation. (h) local correlation between LM and AM in one component of (a), where colour represents the correlation coefficients, the background image shows the corresponding steerable filter response.

where F_{inv} is the inverse or reconstruction filter corresponding to the relevant steerable filter [20] used in the decomposition (steerable filters are their own inverse). W_{ij}^g is the global correlation, and T is the threshold frequency above which the local correlation strategy is applied (note: big i = high frequency filter). T is determined from the energy distribution of the global correlation measure W_{ij}^g as follows,

$$T = \arg \min_t \left(\sum_{i=1}^t \sum_{j=1}^M W_{ij}^g / \sum_{i=1}^N \sum_{j=1}^M W_{ij}^g > \lambda \right) \quad (4)$$

where λ is the energy threshold. We set $\lambda = 0.9$ in this paper. Fig.4(e-g) shows example reconstructed illumination maps using global, local and hybrid correlations. The hybrid method characterises the shadow best, retaining its edges and intensity.

4 Shadow detection

We now combine the colour features described in Section 2 and illumination information (Section 3) into a new shadow edge classifier trained on features from both of sources. We first calculate intensity ratios in the illumination map from regions on either side of colour edges. As with the colour ratios we calculate these intensity ratios at 4 different scales. Taken with the colour features, this provides a 40-dimensional feature vector for each edge pixel identified by the mean shift method. We trained a new decision tree (Decision tree 2 in Fig 1) to classify pixels as shadow or non-shadow. The decision tree gives the probability value p_{shd} of each pixel being on a shadow edge. To further refine the resulting classifications, we introduce a conditional random field (CRF) model [9, 21] built with the constraint that edges should be continuous. The full cost function can be written as:

$$E = \lambda \sum_i F_d(P_i, l_{P_i}) + \sum_{j(j \neq i)} F_s(l_{P_i}, l_{P_j}) \quad (5)$$

where the data cost is $F_d(P_i, l_{P_i})$, the smoothing cost is $F_s(l_{P_i}, l_{P_j})$, and λ controls the balance between these two items. For the data cost, the penalty for labelling a pixel shadow or non-shadow will be $-\log(p_{shd})$ or $-\log(1 - p_{shd})$. In order to reduce the influence of this prior, a regularisation term is added, so that $F_d(P_i, l_{P_i})$ is written as

$$F_d(P_i, l_{P_i}) = \begin{cases} -\log(p_{shd}) - \log(p_c) & \text{if } l_{P_i} = 1 \text{ (shadow)} \\ -\log(1 - p_{shd}) - \log(1 - p_c) & \text{if } l_{P_i} = 0 \text{ (non-shadow)} \end{cases} \quad (6)$$

The constraint function $F_s(l_{P_i}, l_{P_j})$ evaluates the cost of giving label l_{P_i} to point P_i , and label l_{P_j} to point P_j , where P_i and P_j are adjacent to each other. If they belong to the same colour edge, and have similar features, then they should have the same label. Otherwise, setting same labels should be penalized. Function F_s is defined as

$$F_s = \mathbf{1}(P_i \neq P_j) \exp(-\beta \|f_{P_i} - f_{P_j}\|_2^2) \quad (7)$$

where $\mathbf{1}(\cdot)$ is the indicator function; f_{P_i} is the colour ratio feature in HSV and LAB space of points P_i . The optimal label can be found through minimizing the whole cost function i.e. $\min E$ by using the graph cut algorithm [9, 10].

5 Experimental results

All algorithms tried were evaluated against the labelled dataset from [15], which comprised 355 mainly outdoor images including natural views, street views and satellite images. Some shadows are caused by natural objects, such as trees and stones, and some by man-made objects, such as buildings, cars, and lamp posts. We conducted two-fold cross validation (with odd and even numbered images making up each set). We tested various versions of our algorithm to see which combination of features was the most effective. It is important to note that Lalonde’s data set labels only shadow edges that fall on the ground plane. We therefore tested our algorithm on this criterion, and on the more difficult problem of general shadow edge detection. This involved establishing new ground truth labels for the data set adding non-ground plane shadow edges. Lalonde *et al.* designed their algorithm for the ground shadow edge task. To be fair to both methods we thus tested each on the full set of 355 with our ground truth, and also on the 135 images (taken from Zhu’s [15] and LabelMe [16]) used by Lalonde to test ground shadow detection. Lalonde *et al.* do not detail their validation methodology, so we again apply two-fold cross validation.

Fig.5 compares the performance of the various classifiers we generated. Lalonde refers to the original AdaBoost method utilising colour and texture features. We also tested Lalonde’s method switching off the texture features and using only colour ratio (CR) features. The two methods were indistinguishable based on Zhu’s data set, but produced noticeable differences on Lalonde’s images. This is because ground plane textures are more consistent than the general case. Thus texture features could work well for ground shadow, but might not for more general shadows. We considered three algorithm variants for generating the illumination maps: the original Jiang *et al.* [9] method based on global correlations only (global illumination map, GIM), one based on local correlations only (LIM), and a hybrid using global correlations in low frequency bands but local correlations in high frequency bands (HIM). None of these treatments incorporated colour features in Decision Tree 2 (to determine the revised shadow edge map), but they all employed those features in Decision Tree 1 to help them refine their illumination maps. HIM + colour refers to a Decision Tree

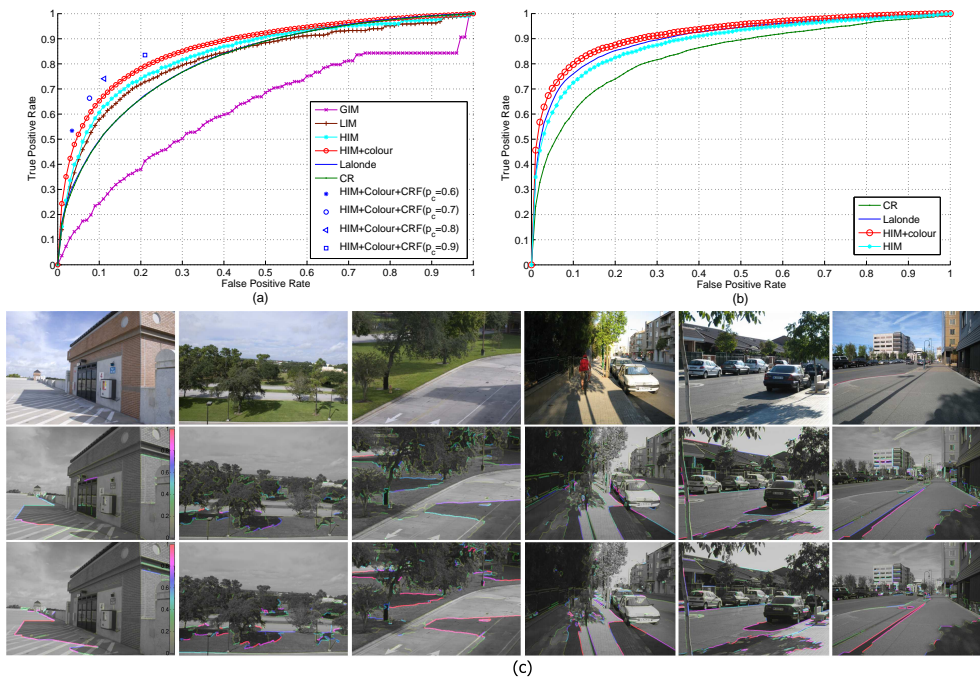


Figure 5: Shadow detection results. (a) ROC curves for the full data set from Zhu *et al.* [25] for shadow edge detection and (b) on the subset used in Lalonde *et al.* [11] for ground shadow edge detection. (c) result samples: For each column: top, original image; middle, shadow detection result using colour features; bottom, the results of proposed algorithm.

2 that was trained using both the illumination map features, and the colour ratio features. Classifications based on HIM + colour were more accurate on both ground truths than all previous methods. Colour and illumination features complement each other, and thus edge pixels are classified as shadows only when hue is relatively constant across the edge and the change in illumination is consistent with the difference between sun and ambient lighting. Finally, adding the CRF model increased classification performance to 83.5% correct rate. Here we adjusted the false alarm rate by altering the regularisation term p_c in the CRF. The other parameters were selected as $\lambda = 0.5, \beta = 4$ through cross-validation. More shadow detection samples are shown in Fig.5(c). Compared with shadow detection results based on colour features, the proposed algorithm can remove the false alarms and increase the possibility of true positives as well.

6 Conclusion

In the proposed shadow detection algorithm, features from the illumination maps and colour segmentation results are combined as inputs to a decision tree trained to detect shadow edges. A CRF model then improves the detection results by imposing an edge continuity constraint. Illumination maps are extracted using a multi-scale intrinsic feature extraction algorithm. A hybrid correlation strategy is proposed to for selecting filter responses for reconstructing

the illumination maps. The proposed algorithm is relatively accurate, outperforming other recent methods. However, there is room for improvement. First, the window size for the reconstructing the illumination map could be optimised. It may also help to make the window size for calculating correlations adaptive. For example, a single window size may be inappropriate if a scene contains a wide range of scales (near and far views) and the standard deviation within a local window could be a useful guide to for deciding the window size. Finally, we are planning to use shadow detection results to further refine intrinsic image extraction methods in an iterative fashion.

References

- [1] H.G. Barrow and J.M. Tanenbaum. Recovering intrinsic scene characteristics from images. *Computer Vision Systems*, pages 3–26, 1978.
- [2] Adrien Bousseau, Sylvain Paris, and Frédo Durand. User assisted intrinsic images. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2009)*, 28(5):130:1–130:10, 2009.
- [3] Michael Collins, Robert E. Schapire, and Yoram Singer. Logistic regression, adaboost and bregman distances. *Machine Learning*, 48:253–285, September 2002.
- [4] D. Comanicu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24:603–619, 2004.
- [5] G. D. Finlayson, S. D. Hordley, C. Lu, and M. Drew. On the removal of shadows from images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 28:59–68, 2006.
- [6] Ivan Huerta, Michael Holte, Thomas Moeslund, and Jordi Gonzalez. Detection and removal of chromatic moving shadows in surveillance scenarios. *IEEE Int'l Conf. on Computer Vision*, 2009.
- [7] Xiaoyue Jiang, Andrew J. Schofield, and Jeremy L. Wyatt. Correlation-based intrinsic image extraction from a single image. *European Conference on Computer Vision (ECCV 2010)*, 4:58–71, 2010.
- [8] Ajay J. Joshi and Nikos P. Papanikolopoulos. Learning to detect moving shadows in dynamic environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30:2055–2063, November 2008.
- [9] Vladimir Kolmogorov and Ramin Zabih. What energy functions can be minimized via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26: 65–81, 2004.
- [10] Jean-Francois Lalonde, Alexei A. Efros, and Srinivasa G. Narasimhan. Estimating natural illumination from a single outdoor image. *IEEE Int'l Conf. on Computer Vision*, pages 183–190, 2009.
- [11] Jean-Francois Lalonde, Alexei A. Efros, and Srinivasa G. Narasimhan. Detecting ground shadow in outdoor consumer photographs. *European Conference on Computer Vision (ECCV 2010)*, pages 322–335, 2010.

- [12] Nicolas Martel-Brisson and Andre Zaccarin. Learning and removing cast shadows through a multidistribution approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29:1133–1146, July 2007. ISSN 0162-8828.
- [13] Y. Matsushita, S. Lin, S. B. Kang, and H.-Y. Shum. Estimating intrinsic images from image sequences with biased illumination. *European Conf. on Computer Vision*, 2: 274–286, 2004.
- [14] Yasuyuki Matsushita, Ko Nishino, Katsushi Ikeuchi, and Masao Sakauchi. Illumination normalization with time-dependent intrinsic images for video surveillance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:1336 – 1347, 2004.
- [15] Takahiro Okabe, Imari Sato, and Yoichi Sato. Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting conditions. In *IEEE Int'l Conf. on Computer Vision (ICCV09)*, pages 1693–1700, 2009.
- [16] Bryan C. Russell, Antonio Torralba, Kevin P. Murphy, and William T. Freeman. Labelme: A database and web-based tool for image annotation. *International Journal of Computer Vision*, 77:157–173, 2008.
- [17] Imari Sato, Yoichi Sato, and Katsushi Ikeuchi. Illumination from shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3):290–300, 2003.
- [18] A. J. Schofield, G. Hesse, P. B. Rock, and M. A. Georgeson. Local luminance amplitude modulates the interpretation of shape-from-shading in textured surfaces. *Vision Research*, 46:3462–3482, 2006.
- [19] L. Shen, P. Tan, and S. Lin. Intrinsic image decomposition with non-local texture cues. *IEEE Computer Vision and Pattern Recognition*, pages 1–7, 2008.
- [20] E. P. Simoncelli and W. T. Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. *IEEE Second Int'l Conf on Image Processing*, pages 444–447, 1995.
- [21] Richard Szeliski, Ramin Zabih, Daniel Scharstein, Olga Veksler, Aseem Agarwala, and Carsten Rother. A comparative study of energy minimization methods for markov random fields with smoothness-based priors. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 30(6):1068–1080, 2008.
- [22] M. F. Tappen, W. T. Freeman, and E. H. Adelson. Recovering intrinsic images from a single image. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27:1459–1472, 2005.
- [23] Eduard Vazquez, Ramon Baldrich, Joost van de Weijer, and Maria Vanrell. Describing reflectances for color segmentation robust to shadows, highlights, and textures. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 33:917–930, 2011.
- [24] Yair Weiss. Deriving intrinsic images from image sequences. *IEEE Int'l Conf. on Computer Vision*, 2:68–75, 2001.
- [25] Jiejie Zhu, Kengan G. G. Samuel, Syed Z. Masood, and Marshall F. Tappen. Learning to recognize shadow in monochromatic natural images. *CVPR 10: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 223–230, 2010.