Shadow Optimization from Structured Deep Edge Detection

Li Shen, Teck Wee Chua, Karianto Leman
Institute for Infocomm Research

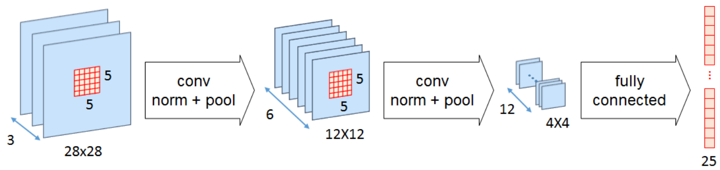

Figure 1: CNN architecture used for learning the structure of shadow edges.

In this paper we propose a novel learning-based framework for shadow detection from a single image. It's notably the first shadow detection work using structured. We exploit local structures of shadow edges by using a structured Convolutional Neural Networks (CNN) framework. A CNN learning framework is designed to capture local structure information of shadow edges and automatically learn the most relevant features. We formulate the problem of shadow edge detection as predicting local shadow edge structures given input image patches. In contrast to unary classification, we take structured labelling information of the label neighbourhood into account. Fig. 1 shows our network architecture. Our learning approach predicts a structured $y^{25\times 1}$ label which corresponds to the 5×5 shadow probability map of the central patch from a larger 28×28 image patch. The network consists of two alternating convolutional and max-pooling layers, followed by a fully connected layer and finally a logistic regression output layer with softmax nonlinear function. Fig. 2 illustrates the advantage of shadow edge detection with structured output CNN. As can be seen, the proposed structured CNN can improve local consistency over pixel labels and avoid spurious labelling.

Moreover, we introduce novel shadow and bright measures to model non-local interactions based on the spatial layout of image regions. For each image patch, shadow and bright measures are computed according to the connectivities of the patch to all the shadow and bright boundaries in the image, respectively. The shadow/bright boundaries are extracted from the shadow edges detected by the proposed CNN. The local shadow/bright measure of a patch, p, depends on its connectivities to the shadow/bright boundaries in the neighbouring area as:

$$\gamma_{shd,lit}(p) = exp(-\frac{con^2_{shd,lit}(p)}{2\sigma^2_{con}}), \tag{1}$$

The global shadow/bright measure is defined as:

$$\Gamma_{shd,lit}(p) = \frac{\sum_{i=1}^{N} w_{app}(p,p_i) w_{spa}(p,p_i) \gamma_{shd,lit}(p)}{\sum_{i=1}^{N} w_{app}(p,p_i) w_{spa}(p,p_i)} \tag{2}$$

where $w_{app} = exp(-\frac{d^2_{app}}{2\sigma^2_{app}})$, and $w_{spa} = exp(-\frac{d^2_{spa}}{2\sigma^2_{spa}})$. d_{app} and d_{spa} are the Euclidean distance between the average colors and locations, respectively. This is based on the observation that if two patches in an image are of the same colour and near to each other, they usually are both in shadow or bright regions. Global shadow/bright probability at p can be computed as $Pr^{shd,lit}_{glb}(p) = 1 - \Gamma_{shd,lit}(p)$.

Using these shadow and bright measures, we formulate a least-square optimization problem for shadow recovery to solve for the shadow map (locations). Our optimization framework combines non-local cues of region interactions in a straightforward and efficient manner where all constraints are in linear form.

$$E = \sum_{i=1}^{N} w_i^{shd} s_i^2 + \sum_{i=1}^{N} w_i^{brt}(1-s_i)^2 + \sum_{i,j} w_{ij}(s_i - s_j)^2 + \lambda \sum_{i \in \mathbf{shd},\mathbf{lit}} (s_i - \tilde{s}_i)^2 \tag{3}$$

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

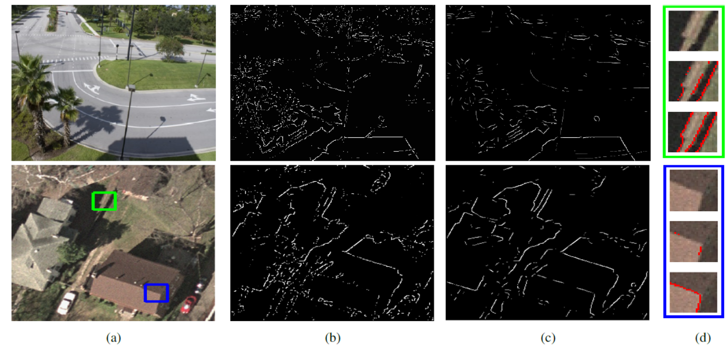

Figure 2: Structured shadow edge detection results. (a) input image. We compare the detection results using (b) 1x1 labelling CNN, (c) 5x5 structured labelling CNN. (d) Zoom-in of green and blue patches. Top to bottom:original, 1x1 and 5x5 outputs.

Methods	UCF dataset	UIUC dataset	CMU dataset
BDT-BCRF [4]	88.7%	-	-
BDT-CRF-Scene [3]	-	-	84.8%
Unary-Pairwise [1]	90.2%	89.1%	-
CNN-CRF [2]	90.7%	93.2%	88.8%
SCNN-LinearOpt	**93.1 %**	**93.4%**	**94.0%**

Table 1: Performance Comparisons of Shadow Detection Methods

where $w_i^{shd,lit} = \Gamma_{shd,lit}(p_i)$ defined in Eq. 2. $\tilde{s}_i$ is the initial values for $p \in \{\mathbf{lit},\mathbf{shd}\}$, where $\tilde{s}_i = 1$ for $p_i \in \mathbf{shd}$ and 0 for $p_i \in \mathbf{lit}$. Our shadow recovery method achieves state-of-the-art results on major shadow benchmark databases collected under various conditions.

[1] Ruiqi Guo, Qieyun Dai, and D. Hoiem. Single-image shadow detection and removal using paired regions. In *CVPR*, 2011.

[2] S.H. Khan, M. Bennamoun, F. Sohel, and R. Togneri. Automatic feature learning for robust shadow detection. In *CVPR*, 2014.

[3] Jean-Francois Lalonde, AlexeiA. Efros, and SrinivasaG. Narasimhan. Detecting ground shadows in outdoor consumer photographs. In *ECCV*. 2010.

[4] Jiejie Zhu, K.G.G. Samuel, S.Z. Masood, and M.F. Tappen. Learning to recognize shadows in monochromatic natural images. In *CVPR*, 2010.

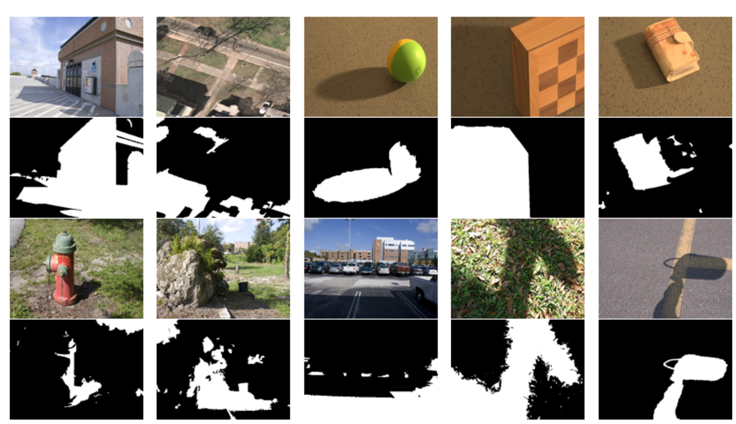

Figure 3: Shadow optimization results from detected shadow edges. Top: input shadow images. Bottom: recovered shadow regions.