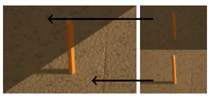
Interactive Shadow Removal and Ground Truth for Variable Scene Categories

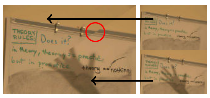
Han Gong
http://www.cs.bath.ac.uk/~hg299
Darren Cosker
http://www.cs.bath.ac.uk/~dpc

Media Technology Research Centre
Department of Computer Science
University of Bath
Bath, UK

Shadows are ubiquitous in image and video data, and their removal is of interest in both Computer Vision and Graphics. We present an interactive, robust and high quality method for fast shadow removal. To perform detection we use an on-the-fly learning approach guided by two rough user inputs for the pixels of the shadow and the lit area. From this we derive a fusion image that magnifies shadow boundary intensity change due to illumination variation. After detection, we perform shadow removal by registering the penumbra to a normalised frame which allows us to efficiently estimate non-uniform shadow illumination changes, resulting in accurate and robust removal. We also present a reliable, validated and multi-scene category ground truth for shadow removal algorithms which overcomes issues such as inconsistencies between shadow and shadow-free images and limited variations in shadows. Using our data, we perform the most thorough comparison of state of the art shadow removal methods to date. Our algorithm outperforms the state of the art, and we supply our code and evaluation data and scripts to encourage future open comparisons.

Shadow removal ground truth The first public data set was supplied in [2]. In our work, we propose a new data set that introduces multiple shadow categories, and overcomes potential environmental illumination and registration errors between the shadow and ground truth images. An example of comparison is shown in Fig. 1. Our new data set avoids these issues using a careful capture setup and a quantitative test for rejecting unavoidable capture failures due to environmental effects. Our images are also categorised according to 4 different attributes.


(a) mismatched illuminaiton


(b) unregistered pixels

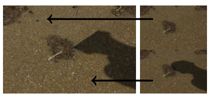

(c) our data (no artefacts)

Figure 1: For each image: top left segment – shadow-free image; bottom right segment – shadow image. (a) and (b) are taken from [2]. An example from our data without these properties is shown in (c).

Our algorithm consists of 3 steps (see Fig. 2):

1) Pre-processing We detect an initial shadow mask (Fig. 2(b)) using a KNN classifier trained from data from two rough user inputs (e.g. Fig. 2(a)). We generate a *fusion image*, which magnifies illumination discontinuities around shadow boundaries, by fusing channels of YCrCb colour space and suppressing texture (Fig. 2(c)).

2) Penumbra unwrapping Based on the detected shadow mask and fusion image, we sample the pixel intensities of sampling lines perpendicular to the shadow boundary (Fig. 2(d)), remove noisy ones and store the remaining as columns for the initial penumbra strip (Fig. 2(e)). We align the initial columns' illumination changes using its intensity conversion image (Fig. 2(f)). This results in an aligned penumbra strip (Fig. 2(g)) whose conversion image (Fig. 2(h)) exhibits a stabler profile.

3) Estimation of shadow scale and relighting Unlike previous work [1, 2], we do not assume a constrained model of illumination change. The columns of penumbra strip are first clustered into a few small groups. A unified sample can be synthesised by averaging the samples of each group (e.g. Fig. 2(i)). Our shadow scale is adaptively and quickly derived from the unified samples which cancel texture noise. The derived sparse scales for all sampled sites (Fig. 2(j)) are then propagated to form a dense scale field (Fig. 2(k)). We remove shadows by inverse scaling using this non-uniform field (Fig. 2(l)).

Evaluation Directly using the per-pixel error [2, 3] between the shadow removal result and shadow-free ground truth does not take into account the size of the shadow, or the fact that some shadows are darker than others. We therefore compute the error ratio $\mathbf{E}_r = \mathbf{E}_n/\mathbf{E}_o$ as our quality measurement where $\mathbf{E}_n$ is the RMSE between the ground truth and shadow removal result, and $\mathbf{E}_o$ is the RMSE between the ground truth and the original shadow image. This normalised measure better reflects removal improvements towards the ground truth independent of original shadow intensity and size. Our removal test is based on our data set of 186 cases, which contains shadows in variable scenarios as well as simpler shadows, plus 28 example cases from [2] – resulting in 214 test cases in total. Each case is rated according to 4 attributes, which are *texture*, *brokenness*, *colourfulness* and *softness*, in 3 perceptual degrees from weak to strong. Our method is compared with three state-of-the-art methods [1, 2, 4] and shows leading performance across all scores. Tab. 1 shows some typical visual results of shadow removal on various scenarios[1].

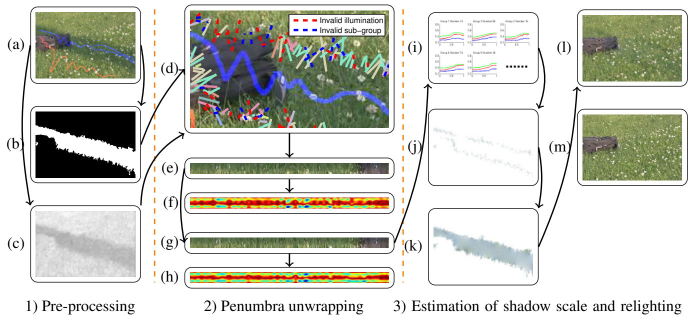


Figure 2: Our shadow removal pipeline. (a) input: a shadow image and user strokes (blue for lit pixels and red for shadowed pixels); (b) detected shadow mask; (c) fusion image; (d) initial penumbra sampling (solid lines in different colours indicate valid samples of different sub-groups. Dashed lines are invalid samples); (e) initial penumbra regularisation; (f) initial penumbra conversion image; (g) final penumbra regularisation; (h) final penumbra conversion image; (i) penumbra illumination estimation; (j) sparse shadow scale; (k) dense shadow scale; (l) output; (m) GT.

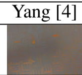
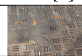
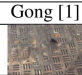
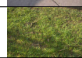
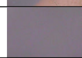
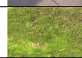
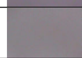
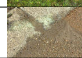
	Original	Yang [4]	Guo [2]	Gong [1]	Ours	GT
Tex.						
Sof.						
Bro.						
Col.						
Other						

Table 1: Comparisons using images in different categories.

Application Our method is exclusively suitable for real-time interactive shadow editing which offers free controls for shape, darkness and smoothness of either new or original shadows (see our supplementary material).

Conclusions We have presented an interactive method for fast shadow removal together with a state of the art ground truth. Our method balances the complexity of user input with robust shadow removal performance. Our quantitatively-verified ground truth data set overcomes issues of mismatched illumination and registration. We have evaluated our method against several state of the art methods using a thorough quantitative test and shown leading state of the art performance.

[1] H. Gong, D. Cosker, C. Li, and M. Brown. User-aided single image shadow removal. In *ICME*, pages 1–6, 2013.

[2] R. Guo, Q. Dai, and D. Hoiem. Paired regions for shadow detection and removal. *PAMI*, PP(99):1–1, 2012.

[3] Y. Shor and D. Lischinski. The shadow meets the mask: Pyramid-based shadow removal. *CGF*, 27(2):577–586, 2008.

[4] Q. Yang, K.-H. Tan, and N. Ahuja. Shadow removal using bilateral filtering. *IEEE Trans. on Image Proc.*, 21(10):4361–4368, 2012.

[1] Our supplementary material shows a wide range of other removal results with higher resolution images.