USER-AIDED SINGLE IMAGE SHADOW REMOVAL

Han Gong, Darren Cosker, Chuan Li, and Matthew Brown

University of Bath, Department of Computer Science
Bath, BA2 7AY, United Kingdom
{h.gong, d.p.cosker, c.li, m.brown}@bath.ac.uk

ABSTRACT
This paper presents a novel user-aided method for texture-preserving shadow removal from single images which only requires simple user input. Compared with the state-of-the-art, our algorithm addresses limitations in uneven shadow boundary processing and umbra recovery. We first detect an initial shadow boundary by growing a user specified shadow outline on an illumination-sensitive image. Interval-variable intensity sampling is introduced to avoid artefacts raised from uneven boundaries. We extract the initial scale field by applying local group intensity spline fittings around the shadow boundary. Bad intensity samples are replaced by their nearest alternatives based on a log-normal probability distribution of fitting errors. Finally, we use a gradual colour transfer to correct post-processing artefacts such as gamma correction and lossy compression. Compared with state-of-the-art methods, we offer highly user-friendly interaction, produce improved umbra recovery and improved processing given uneven shadow boundaries.

Index Terms— shadow removal, single image, user-aided

1. INTRODUCTION

Shadows are ubiquitous in natural scenes, and their removal is an interesting and important area of research. As well as attempting to solve this problem for image editing by e.g. artists, shadows affects many computer vision algorithms, e.g., unwanted shadow boundary’s causing artefacts in image segmentation and contributing to drift issues when tracking given moving scenes.

There are several automatic methods for shadow detection and removal. For example, intrinsic image based methods [1, 2] and feature (e.g. intensity and texture) learning based methods [3]. However, automatic shadow detection is ill-posed and often difficult even for humans [4].

This paper focuses on user-aided single image shadow removal. User-aided methods generally achieve better shadow detection and removal at the cost of user input. Many of them [4, 5, 6, 7] are also texture-preserving. Most past work requires precise input defining the shadow boundary and can have issues when properly relighting the umbra.

According to past work [1], shadow effects can be represented as an additive scale field $S_c$ in the log domain. In the log domain, an image $\tilde{I}_c$ with shadow effects added to an original image $I_c$ can be represented as follows:

$$\tilde{I}_c(x, y) = I_c(x, y) + S_c(x, y)$$

where $c \in \{R, G, B\}$ indicates each RGB colour space channel, $x$ and $y$ are the coordinates of pixels. The scales of lit area pixels are 0 and the other pixel scales are negative numbers.

The umbra is the darkest part of the shadow while the penumbra is the wider outer boundary with a nonlinear intensity change between the umbra and lit area [8]. The penumbra scale is non-uniform and shadowed surface textures generally become weaker [6]. Contrast artefacts can also appear in shadow areas due to image post-processing [7, 6, 8]. Arbel et al. [7, 8] developed curved surface shadow removal by formulating scale field estimation as thin-plate surface fitting requiring prior shadow feature points. Wu et al. [4] proposed a shadow removal method requiring a supplied quadmap and applied a Bayesian optimisation framework. Mohan et al. [5] apply a curve-fitting model and remove shadows in the gradient domain. Given precise shadow boundaries, Liu et al. [6] apply spline-fitting and texture transfer to remove shadow. Shor et al. [9] proposed a shadow removal method that requires one shadow pixel but does not preserve texture shadows. Although some past work preserves penumbra texture, umbra removal and uneven boundary processing are still problematic. Assuming a uniform umbra, Liu et al. [6] often introduces over-saturation artefacts in some cases. Uneven shadow boundaries can affect penumbra detection and scale estimation. Most user-aided methods to assist boundary detection [4, 6, 8] require careful highlighting of the boundary. We propose a method that requires one rough stroke to mark an umbra sample. It can process uneven shadow boundaries and achieves better umbra removal compared with past work. Our major contributions are as follows:

Easy user input. Past work, e.g. [4, 6, 7, 8], requires precise user-inputs defining the shadow boundary. Our
method only requires an umbrella segment highlighted by one rough stroke and grows it on an illumination-sensitive image to obtain initial shadow boundaries.

Interval-variable sampling. Past work, e.g., [6, 8], applies interval-fixed sampling around penumbra that causes artefacts near uneven shadow boundaries. To address this, we develop interval-variable sampling according to shadow boundary curvature.

Local group optimisation for picked samples. Inspired by Arbel’s [7, 8] and Liu’s [6] spline-fitting idea, we propose a local group optimisation that balances curve fitness and local group similarity. Unlike past work, we filter inferior samples that are replaced with their closest alternatives according to a log-normal probability distribution. This reduces shadow removal artefacts.

Gradual colour transfer. Post-processing effects cause inconsistent shadow removal compared with the lit areas both in tone and in contrast. We make use of statistics in thin penumbra boundaries and the shadow scale field to correct these issues.

2. PROPOSED METHOD

Fig. 1. An overview of our approach. (a): user input indicated by a purple curve, (b): initial shadow boundary detection indicated by the purple mask (§2.1), (c): interval-variable sampling lines indicated by the white lines (§2.2.1), (d): shadow removal by estimating the scale field of the shadow area (§2.2.2). Note that the area’s contrast still appears different from its surroundings. (e): image correction by our gradual colour transfer (§2.3). The input image is taken from [8].

Given an input image and a user specified umbrella segment (Fig. 1(a)), we detect the initial shadow boundaries by expanding the given umbrella segment on an illumination image using an active contour (Fig. 1(b), §2.1). To keep boundary details, we sample pixel intensities along the sampling lines perpendicular to the shadow boundary (Fig. 1(c), §2.2.1) at variable intervals. We perform a local group optimisation to estimate the illumination change which refines shadow boundary detection and provides an initial scale field. According to an adaptive sample quality threshold, sampling lines with bad samples are replaced by their nearest neighbours and a later local group optimisation is applied to them. Finally, we relight the shadow area using our scale field (Fig. 1(d), §2.2.2) and correct post-processing artefacts using our gradual colour transfer (Fig. 1(e), §2.3).

2.1. Initial Shadow Boundary Detection

Determine the initial shadow boundary is the first step of penumbra detection and is required in many previous methods including [7, 8, 6]. In this subsection, we explain how to derive an initial shadow boundary from a given rough umbrella sample segment.

Inspired by [10], we fuse four normalised candidate illumination-sensitive channels from different colour spaces into an illumination image. The chosen channels are: the V channel (C1) of HSV space, the V channel (C4) of LUV space, and the Y channel (C2) and Cb channel (C3) of YCbCr space. We measure the confidence values of each candidate channel using an exponential incentive function \( \varphi \) representing the textureness of each of their umbrella sample segments:

\[
\varphi(x) = x^{-\lambda}(\lambda > 0)
\]

where \( x \) is the pixel intensity, \( \lambda \) (default value 5) determines the steepness of the incentive function. Lower textureness is preferred as it means higher intensity uniformity of umbrella segment. The textureness is measured by standard derivation of intensities. The fused image \( F \) is computed as a weighted sum of each normalised candidate channel \( C_l \) as follows:

\[
F = \frac{\sum_{l=1}^{4} C_l \varphi(\sigma_l)}{\sum_{l=1}^{4} \varphi(\sigma_l)}
\]

where \( l \) is the channel index, \( \sigma_l \) is the standard derivation of the umbrella sample intensities of \( C_l \). To avoid texture noise, we apply a bilateral filter [11] to \( F \) first. We grow a sparse-field active contour [12] on the fused image to detect the initial shadow boundary. As shown in Fig. 2, region-growing using the fused image is more robust than relying on one channel.

2.2. Scale Field Estimation

This subsection describes our scale field estimation from initial shadow detection. Shadow affects are represented by

Fig. 2. An illumination-sensitive image \( F \) is fused from the four colour channels of the image (a) (taken from [6]) using the intensity statistics of the given umbrella region (b) inside the purple boundary. We grow the region using the image \( F \) to the entire shadow area as indicated in (g) by the purple mask. Sub-figures (c-f) are the region growing results on the corresponding four single channels.
varying (or different scale) intensity values. Using a scale field better represents the penumbra and umbra variations, and is used to relight the shadow area using Eq. 1. In § 2.2.1, we first sample the log domain pixel intensity along the sampling lines perpendicular to the shadow boundary. In § 2.2.2, we adopt a local group spline fitting optimisation through the measured sampling line pixel intensities to estimate sparse scales from the initial intensity samples. We replace bad intensity samples with their nearest alternatives and re-optimise for them. We spread the sparse scales to dense scales followed by a gradual colour transfer in § 2.3 that adjusts the colour and texture of initial shadow removal.

2.2.1. Interval-variable Sampling

According to Eq. 1, the logarithms of the original image are supplied for sampling. We sample pixel intensities along the lines perpendicular to the initial shadow boundary as shown in Fig. 4(c). Uneven boundaries can result in non-smooth vector normal estimations along the shadow boundary. To overcome this, we apply cubic spline smoothing to the initial boundary points before we compute their normals and curvatures. Under-sampling along the boundary neglects sharp details and causes artefacts as shown in Fig. 4 while over-sampling incurs penumbra removal noise due to texture as shown in Fig. 3. More sparse pixel scales are computed for curvy boundary parts for precise in-painting. To avoid texture artefacts, we apply a bilateral filter [11] to the input image before sampling. Unlike past work [5, 6, 8, 7], we do not adopt a fixed sampling line interval, e.g., one sampling line per boundary pixel. Our method adjusts the sampling interval according to the curvature of the smoothed boundary. The interval is the same for all RGB channels. We set a curvature accumulator for shadow boundary points and accumulate along the boundary. We place sampling marks and reset the accumulator when the curvature sum reaches a threshold $\xi$ (default value 0.05). We achieve this by adopting Eq. 4. We limit the absolute curvature of each boundary point with an upper limit $\xi$ and compute a cumulative sum array $Q$ of the saturated absolute curvature array. To determine the sampling interval, we choose boundary points where the sampling lines pass through as follows:

$$
\begin{align*}
\tilde{Q}_m &= \lfloor Q_m / \xi \rfloor & (m \leq N, m \in \mathbb{N}) \\
D_n &= \tilde{Q}_{n+1} - \tilde{Q}_n & (n \leq N-1, n \in \mathbb{N})
\end{align*}
$$

where $N$ is the number of boundary points, $m$ and $n$ specify the index of boundary points, $\tilde{Q}$ is the array of the quantized and normalized cumulative sum $Q$, $D$ is the adjacent element differences array of $\tilde{Q}$. To get the interval of sampling marks, we set the marks for the first and last boundary points and all the points in $D$ with non-zero values. If the boundary is a straight line, the initial interval is fixed up to a maxima (five boundary points). As shown in Fig. 4, our variable sampling interval avoids penumbra removal artefacts around sharp boundary parts.

2.2.2. Illumination Variance Estimation

Having obtained sparse intensity samples at different positions along the boundary, our goal is to find illumination scaling values inside the umbra, penumbra and lit area. We model the illumination scale change $S_i$ for each $i^{th}$ intensity sample of each RGB channel as follows (see also Fig. 5):

$$
S_i(x) = \begin{cases} 
K & x \leq x_1 \\
f(x) & x_1 < x \leq x_2 \\
0 & x > x_2 
\end{cases}
$$

where $x$ is a pixel location within the sampling line, $x_1$ and $x_2$ determine the start and end of the penumbra area respectively, and $K$ is a negative scale constant for sample points within the umbra area ($x < x_1$). The constant 0 is assumed for the lit area piece ($x > x_2$) as this falls inside a lit area of the image and does not require re-scaling. In order to solve $x_1$, $x_2$, $K$, and $f$, we solve a piece-wise function $G_i$ parameterised by $v_1$, $v_2$, $v_3$, and $v_4$.
where $B$ is a cubic shape function (a sinusoidal function here also produces adequate results) and $y$ is the input. Our solution is thus reduced to solving $v_2, v_3,$ and $v_4$ for each RGB channel (twelve in total). Illumination of each channel may vary differently while penumbra boundaries of three channels are usually the same. We thus assume a common penumbra width and position for each channel, determined by $v_2$ and $v_3$ respectively. We formulate solving the eight remaining parameters as an optimisation problem which balances curve fitness and local group fitting similarity. We minimise the energy function $E_i$ for the $i^{th}$ sampling line as follows:

$$E_i = \alpha_1 E_{fit}(V_i, Z_i) + \alpha_2 E_{gs}(V_i, V_{i-1}, \ldots, V_{i-r-1})$$  

(7)

where $r$ is the number of members in a local sampling line group (default value 5), $\alpha_1$ and $\alpha_2$ are two balancing weights (default values 1 and 0.2 respectively), $E_{fit}$ measures the sum-of-square fitting error between the three piece-wise functions $G_i$ (defined by parameter vector $V_i$) and the original three-channel intensity sample matrix $Z_i$, $E_{gs}$ measures the parameter similarity between the neighbouring members of a local group. In practice, penumbra width affects the removal quality most significantly. We thus only compare the similarity of the shared $v_2$ in Eq. 6. $E_{gs}$ is defined as follows:

$$E_{gs}(V_i, V_{i-1}, \ldots, V_{i-r-1}) = \left( \frac{1}{n_i} - \frac{\sum_{k=1}^{n} \psi(e_{i-k})}{\sum_{j=1}^{r} \varphi(u_{i-j})b_{i-j}} \right)^2$$  

(8)

where $v_{R1}, v_{G1}$ and $v_{B1}$ are the $v_1$ for each channel, $v_{R4}, v_{G4}$ and $v_{B4}$ are the $v_4$ for each channel, $v_{RGB2}$ and $v_{RGB3}$ are the shared $v_2$ and $v_3$ for all channels, $b_i$ is the $v_{RGB2}$ of the parameter vector $V_i$, $\psi$ indicates the fitting error of the previous $u^{th}$ fittings, $\varphi$ is the function defined in Eq. 2. We solve this using a sequential quadratic programming algorithm [13].

However, interval-variable sampling can not always guarantee good sample quality. Strong surface textures introduce more significant intensity change than illumination change. Unlike past work[5, 6, 8, 7], we ignore sampling lines with higher fitting errors and pick their most suitable neighbours. Based on our empirical tests on various images, we model the initial fitting error distribution as a log-normal probability distribution. As the initial fitting error is distributed log-normally, we can convert it to its corresponding normal distribution by taking the exponential of it. According to the empirical 3-sigma rule of normal distribution, we ignore sampling lines with fitting errors higher than a threshold $\mu + \sigma$ which accounts for 15.8% of all samples where $\mu$ and $\sigma$ are the mean and standard deviation of errors. For each sampling line with bad samples, we only attempt to pick its nearest sampling lines within a short distance, i.e. no further than its neighbouring chosen sampling lines. To evaluate the replacements’ quality and compute their sparse scales, we apply the same optimisation method described previously.

After optimisation, we obtain sparsely distributed scales, defined in Eq. 1, of all sampled pixels inside and around the penumbra area. We interpolate the scales within the penumbra area using linear interpolation and extrapolate the other scales in the lit and umbra area using in-painting [14].

2.3. Gradual Colour Transfer

In practice, the theoretical shadow affect formulation does not often hold. Image acquisition devices usually apply post-processing, e.g. Gamma correction. Lossy compressions, e.g. JPEG, are also common in images such that compression artefacts (e.g. affecting contrast) in the shadow area become noticeable when removal is applied. To address this, we extend the colour transfer in [15] with scale field $S_m$. We compute the normalised scale increase $h_i$ of the $i^{th}$ sampling line according to Eq. 5 as follows:

$$h_i(x) = (\exp(f_i(x)) - \exp(K_i))/(1 - \exp(K_i))$$  

(9)

where $x$ is the location of pixel of a sampling line, $K_i$ and $f_i$ are respectively the lit constant $K$ and the cubic function piece of the $i^{th}$ sampling line. We apply the same method described in § 2.2.2 to interpolate and extrapolate sparse scale increase values computed by Eq. 9 to form a dense scale increase field $S_m$. We convert the initial shadow removal image from RGB space to LAB space. For each LAB channel, we compute the mean $\mu_u$ and deviation $\sigma_u$ of the umbra side pixel intensities near penumbra as the source and we compute the mean $\mu_t$ and deviation $\sigma_t$ of the lit side pixels near penumbra as the target. We adjust the initial removal image channel $L$ to the final image channel $\tilde{L}$ as follows:

$$\begin{cases} 
\mu_s(x,y) = \mu_u + S_m(x,y)(\mu_t - \mu_u) \\
\sigma_s(x,y) = \sigma_u + S_m(x,y)(\sigma_t - \sigma_u) \\
\tilde{L}(x,y) = \mu_t + (L(x,y) - \mu_s(x,y))\sigma_t/\sigma_s(x,y)
\end{cases}$$  

(10)

where $x$ and $y$ are the coordinates of pixels, $\mu_s$ and $\sigma_s$ are the fields of gradual source mean and deviation. We show an example of colour transfer in Fig. 6.
Fig. 7. From (a), the red light component that passes through semi-transparent object is still in (b) and is reduced by our method in (c). Ours also removes shadow residuals near the highly uneven boundary in (b). The shadow removal result of (d) shown in (e) results in over-saturation of the umbra. Also, the texture and colour variations across the shadow boundary are not smooth and consistent. We overcome this issue as shown in (f). Our removal result of (g), as shown in (i), is improved over [5] as shown in (h).

3. RESULTS

In our experiments we visually compare our shadow removal results with state-of-art methods using the same input images and some other representative images. Our method has been tested on various photos from existing work and additional representative photos. Our method is highly user-friendly as shown in Fig. 1 and produces state-of-art quality shadow removal as shown in Fig. 1, 3, 4, 7, 6, 8, and 9. For each shadow area, we require one stroke, shown as red curves in figures, marking an characteristic umbra segment. The stroke can be very rough and does not necessarily need to follow the shadow boundary. We also overcome some cases that past work fail to handle.

In our comparisons, we focus primarily on more recently studied texture-preserving methods as opposed to older texture-lossy ones such as in-painting [2]. Our previous figure (Fig. 4) has already highlighted the issues caused by uneven boundary processing. This compares our variable sampling interval – which is boundary-perpendicular – against our own test of fixed interval boundary-perpendicular sampling and the fixed interval vertical and horizontal sampling method in [6].

Compared with [6], our method reduces the red light component that passes through the semi-transparent leaf as illustrated in Fig. 7. The redness of different parts of the leaf is different and thus the amount of passed red light is not uniform. To handle this, we assume non-uniform scales in the umbra, i.e. different umbra scale constants as mentioned in § 2.2.2, and process the RGB channels separately based on a shared penumbra area. Our interval-variable sampling (see Fig. 7(c)) removes some minor residual shadow fragments near the shadow boundary (noticeable in Fig. 7(b)). In Fig. 7(d), the shadowed sandy surface is consistently recovered. As the surface in the lit area is not saturated, the same surface beneath the shadow should also not be saturated in Fig. 7(e). Fig. 7(f) shows our method avoids over-saturation artefacts and achieves more consistent texture and smoother colour variance across the shadow boundary. In Fig. 7(h), the result from [5] appears darker in both the shadow and lit areas. Our result shows a consistently coloured texture between the lit area and shadowed area. Figs. 8 and 9 demonstrate our results on images with various textures, reflectances, and shadow boundaries. Fig. 8(a) shows the removal of soft shadow cast on the curved surface and the texture consistency. Fig. 8(b) reveals our smooth and texture-consistent shadow removal applied to the earlier example illustrated in Fig. 4. The colour of the tree bottom of the hill is consistent with the trees at the hill top. The smoothness of the hill-side’s colour and texture are also recovered. In Fig. 8(d), Fig. 8(e), Fig. 8(c), and Fig. 8(f), the texture and self-shadows are kept after removal.

However, our method still shows minor limitations in highly complex cases (still unsolved in state of the art work). In Fig. 9(a), the river bed is recovered but the ripple highlights in the shadow area are missing. These highlights are mainly produced by light reflection on the wavy water surface and the light refraction from the river bed. As the directed light is blocked in the shadow area, these complex affects cannot be recovered by simple relighting. In another challenging case in Fig. 9(b) with non-white light, the book texture in the wide penumbra area is retained with minor over-saturation artefacts. This is due to the strong broken-shadow-like text texture which affects the spline fitting.

4. CONCLUSION

We have presented a user-friendly texture-preserving shadow removal method that overcomes some common limitations from the past work. Specifically, our method retains shadowed texture and preforms well on highly-uneven shadow boundaries, non-uniform umbra illumination, and non-white lighting. Our main technical contributions are: (1) highly
Fig. 8. Demonstrations. Left: original images; Right: our results. The original images in (a) and (c) are from [8] and the original images in (b) and (f) are from [6] and [5] respectively.

user-friendly input; (2) interval-variable sampling; (3) local group optimisation; and (4) gradual colour transfer. In future work, we would like to focus on more complex cases, such as highly broken shadows, shadowed surfaces with very strong shadow-like textures, and complex reflections in transparent scenes. Perceptual studies based on shadow removal are generally ignored in the literature, and our aim is to perform an extensive study in this area. We will combine this with qualitative assessment on data such as [3].

5. REFERENCES


