Single-View Relighting with Normal Map Painting

Makoto Okabe$^1$
Takeo Igarashi$^{1,4}$
Gang Zeng$^2$
Long Quan$^2$
Yasuyuki Matsushita$^3$
Heung-Yeung Shum$^3$

$^1$The University of Tokyo, Tokyo, Japan
$^2$HKUST, Hong-Kong, China
$^3$Microsoft Research Asia, Beijing, China
$^4$PRESTO, JST

makoto21,takeo@ui.is.s.u-tokyo.ac.jp
zenggang,quan@cse.ust.hk
yasumat,hshum@microsoft.com

Abstract

This paper presents an interactive technique for changing the illumination of objects in a real photograph. Because lighting effects are very sensitive to surface normal perturbations but hard to automatically analyze, we rely on a pen-based interface to quickly draw an approximate normal map over a photo. The photo and the approximate normal map are then given as input to a novel algorithm that refines the map and assigns reflectance to every pixel. We present relighting results for a variety of scenes, and use our technique to match the illumination of multiple photographs.

1. Introduction

Illumination has often been used by painters and photographers as a tool for eliciting striking visual impressions. While physical control of illumination is sometimes possible, changing the illumination of a photo after it has been taken can prove a challenging task. One possible solution is to photograph the same scene under many different illumination conditions and then combine these photos to create new images of the same scenes [16, 14, 13, 6]. Unfortunately, these types of datasets are hard to obtain because they require a large number of photos, usually under controlled illumination conditions. The other approach is to obtain a three-dimensional (3D) model of the scene using a range scanner and apply inverse rendering techniques [15]. However, such data is not always available, especially it is difficult for the scenario of using existing photographs. When just a single photo of a scene is available, with no prior information about its shape or illumination, the relighting problem becomes especially difficult.

This work was done while the first and second authors were visiting Microsoft Research Asia.

Figure 1. A photograph(left) is relit as a novel image(right) by a user specified lighting condition.

The colors and intensities within a photo are created from the interaction of light with surfaces in the scene. This interaction depends on the reflectance properties of the surfaces, as well as their shape. As a result, accurate relighting can be done only by analyzing how surface reflectance, surface shape, and the light source position(s) affect a given photo. This is a highly under-constrained problem because every pixel in the image is a function of incident lighting, surface orientation and reflectance, none of which are known.

Our approach is to efficiently extract a representation that is specifically designed for re-lighting applications. Since we use a single photograph as input, our work is closely related to prior work on single-view modeling [17, 19, 1, 4, 3, 9]. These methods emphasized the detailed
acquisition of full 3D scene models using a variety of interactive tools. Unlike these approaches, we rely on a reduced scene representation that is based on surface normals and is much easier to interactively specify. We propose a novel user interface (UI) to assign surface normals to a photograph based on the fact that a human is capable of perceiving surface orientation more precisely than depth of the scene [10]. The surface normal representation has been used in the scenario of illumination control for cel animation [8]. In our scenario of relighting a single photograph, this representation is also useful in manipulating photographs where the dominant pixel intensity variations are caused by local lighting effects. We show that our technique works well on photographs of many complex scenes.

Our work is inspired by a simple and efficient UI of photo-editing applications [17] that allows a user to alter the appearance of an image through simple painting operations, such as brushing and flood-fill. We adopt a pen-tablet input device because it allows a quick and intuitive input method as used in the context of free surface modeling by [7]. We take the advantage of the feature of a detectable pen angle in recent products, e.g., Wacom’s Intuos2. Pen angle is directly associated with the surface normal of a scene to achieve an intuitive input of surface normals. We provide a set of input methods to efficiently assign surface normals to the scene. The assigned surface normal is then used to estimate its surface reflectance and a more precise surface normal. Once the scene is represented by our data structure where each pixel encodes surface reflectance and surface normal, the scene can be re-rendered under different illumination conditions (Figure 1).

2. Our Approach

2.1. Overview

Our method consists of two steps: Interactive surface normal assignment and simultaneous estimation of reflectance and detailed surface normal.

In the first step, a surface normal is assigned via a set of input methods designed for this purpose. Three different types of brushes and a flood-fill tool are provided to facilitate the assignment process. Once the surface normal is assigned by a user, the second step uses the normal map for automatically and simultaneously estimating the surface reflectance property and a finer surface normal map. Two techniques are proposed to achieve single-view relighting. One is a smooth flood-fill tool for surface normal assignment, and the other is an automatic refinement of surface normals and reflectance.

Our contributions are as follows. We provide a set of tools to assign surface normals to a single image. The pen angle obtained from a pen-input device is used for intuitive assignment of surface normals. The system provides three types of brushes to allow users to intuitively paint surface normals. To facilitate surface normal assignment, a flood-fill tool is developed that fills in the specified area by smoothly propagating surface normals. We have also developed an algorithm to simultaneously estimate surface reflectance and a finer surface normal from a coarse hand-drawn surface normal. Our system permits relighting of a single photograph under different illumination conditions.

2.2. Interactive Normal Editing

Labeling Multiple Photo-object Our method begins with labeling photo-objects that are of interest for relighting. Once the labels are assigned to photo-objects, the boundaries are used for the following surface normal editing process.

To achieve multi-object interactive segmentation, we have added an interactive control to the segmentation method used in [20]. The input image is first smoothed while preserving edges by a variant of the anisotropic diffusion and bilateral filtering algorithms. After smoothing, each pixel is assigned its own segment label. These labels are merged interactively given a user defined threshold defined by a RGB Euclidean distance between average colors of the segments. A user can interactively merge/split the segments by changing the threshold. Once a desired segment is obtained at a certain threshold, a user can mark the edge as an object boundary and continue the labeling process by changing the threshold. In this way, a user can
efficiently label photo-objects with changing the segment boundaries (Figure 2).

We have also developed a tool to explicitly specify important boundaries by stroking. Given a stroke, the system automatically creates an area by dilating the stroke and calculating graphcut in the area using the method proposed by [12]. In the case the color gradient in the area is reasonably small, the algorithm tries to set a boundary along the user input stroke. In this way, important boundaries can be explicitly specified. Figure 3 illustrates the usefulness of boundary stroke.

Surface normal modeling Three different types of brushes and a flood-fill tool are designed for the interactive surface normal assignment. The basis of our UI is a painting UI as used in [17] for depth assignment. Our painting UI has three different brushes to facilitate the task of surface normal assignment. The first one is a normal brush with which a user can directly input a surface normal. As a key feature of the normal brush, we use a pen-tablet input device taking advantage of the ability of commercial pen-tablets providing information about angle of pen-input. Pen angle is directly associated with the surface normal in the image (Figure 4). This allows a user to intuitively assign surface normals in an image. The second brush is a copy brush that allows copying a surface normal of the specified point and paints areas with the copied surface normal regardless of the pen angle. The copy brush is useful when the same surface orientation exists at different locations, as often is the case with architecture. A user can also blur the assigned normal using a blur brush, which is useful to create smooth variations of the surface normal. The blur brush works as a Gaussian diffusion over the specified area of the normal map. For all brushes, the brush size can be interactively varied.

The three brushes provide an intuitive and efficient normal-input mechanism, however, it is still a time-consuming task when filling-in a large image area. To accelerate normal assignment, our system provides a smooth flood-fill tool that can fill up the selected image area by propagating sparse user input.

**Smooth Flood-fill Tool** To facilitate the surface normal assignment process, we developed a flood-fill tool that propagates sparse surface normal input to smoothly fill in the specified image area. We use intensity similarity measure to propagate the coarsely specified surface normal \( \vec{N} = (N_x, N_y, N_z) \) as Levin et al. used in the context of color propagation [11]. The propagation is bounded by the pre-defined label and user-specified area where intensity change is due to the variation in surface orientation. It is achieved by minimizing the energy functional \( J \):

\[
J(U) = \sum_{r} \left( U(\vec{r}) - \sum_{s \in A(\vec{r})} w_{ij} U(\vec{s}) \right)^2,
\]

where \( \vec{r} \) and \( \vec{s} \) denote pixel labels, and \( A(\vec{r}) \) represents a set of adjacent pixels of \( \vec{r} \). The weighting factor \( w_{ij} \), which sums up to one, represents the similarity between \( \vec{r} \) and \( \vec{s} \), which is defined in [11] by

\[
w_{ij} \propto 1 + \sigma_r^{-2}(Y(\vec{r}) - \mu_r)(Y(\vec{s}) - \mu_s),
\]

where \( Y(\vec{r}) \) is the intensity at pixel \( \vec{r} \), \( \mu_r \) and \( \sigma_r^2 \) are the mean and variance of the intensities respectively in \( N(\vec{r}) \). Equation (1) is evaluated twice for \( U \leftarrow N_x/N_z \) and \( U \leftarrow N_y/N_z \), and the surface normal is re-normalized after computation.

The actual computation is performed by solving a sparse system of linear equations which is derived by taking partial derivative of Equation (1) by \( U \). For propagating surface normal smoothly, \( w_{ij} \) can be adjusted according to the image condition. For example, when \( w_{ij} \) is proportional to 1.0, all the intensity information on the image are ignored and painted normal values are simply interpolated on the image coordinate. This mode is useful for a surface with a complex texture whose intensity variations are not directly related to normal variations. Figure 5 shows the results of our flood-fill method with and without intensity similarity measure. As shown in the figure, the assigned surface normal is smoothly propagated.

**2.3. Refining Reflectance and Normals** Once the hand-drawn surface normal map is obtained, an automatic method for simultaneously estimating surface reflectance and a fine surface normal map is applied. The method takes an input image \( I(\vec{r}) \) and the user-assigned normal map \( \vec{N}(\vec{r}) \) as input. We assume that the scene is governed by a local lighting effect, more precisely, illumination can be approximated by a dominant directional lighting that is represented with a directional vector \( \vec{I} \) and magnitude \( I_p \), a small ambient lighting term \( I_a \) that approximates interreflections, and shadowed area is not dominant in the image.
We use the following local illumination model in this method:

\[ I(\vec{n}) = I_a k_a + I_p S_p(k_d(\vec{I}, \vec{n}) + k_s H_p) \]  

where \( k_a \), \( k_d \) and \( k_s \) are coefficients to describe albedo of ambient, diffuse and specular reflection respectively, \( \vec{n} \) is unit vectors of normal directions, \( S_p(\vec{n}) \in [0,1] \) represents a shadowing term, and \( H_p(\vec{n}) \) is a term that governs intensity of specular highlight.

Throughout this paper, we only consider gray images to describe our idea. For color images, the same process is applied for each red, green, and blue components.

**Light source estimation** Our second step begins with estimating the dominant lighting direction \( \vec{I} \) in the original scene. Traditionally, techniques from shape from shading [2] have been used to estimate the direction of the light source. Assuming a Lambertian image formation model [5], most of these techniques try to simultaneously recover both shape and the direction of the light source. In our case, the problem can be more constrained by the user-defined surface normal information and pre-defined label map which represents the regions of objects. Assuming that the area which shows diffuse reflection is dominant in each label i, lighting direction \( \vec{I} \) can be estimated by first minimizing \( E(\vec{I}) \):

\[ E(\vec{I}) = \sum_i \sum_{\vec{n}} (I(\vec{n}) - A_i - D_i(\vec{I}, \vec{N}(\vec{n})))^2, \]  

where \( A_i = I_a \vec{k}_i \) is a constant ambient, \( D_i = I_p \vec{k}_i \) is the diffuse component. \( \vec{k}_i \) is a constant albedo in label \( i \). \( E(\vec{I}) \) can be minimized by solving a linear system of equations which is derived by its first derivatives about \( A_i \) and \( D_i \). Once \( A_i \) and \( D_i \) are determined, \( \vec{I} \) is then estimated by searching for the minimum energy \( E(\vec{I}) \) as \( \vec{I} = \arg \min \{E(\vec{I})\} \). In this way, lighting direction \( \vec{I} \) is estimated. At the same time, \( A_i \) and \( D_i \) are assigned to each label.
Determining shadow and highlight Using the estimates $A_i$, $D_i$ and $\bar{T}$, our method detects the shadowed and specular highlight areas by evaluating deviation from the diffuse reflection model as follows:

$$S_p(\bar{T}) = \begin{cases} 1, & \text{if } I(\bar{T}) > \bar{k}_i(I_a + t_s D_i) = A_i + t_s D_i \\ \frac{I(\bar{T}) - A_i - D_i}{t_s D_i}, & \text{if } I(\bar{T}) > \bar{k}_i I_a = A_i \\ 0, & \text{otherwise} \end{cases}$$

$$I_p k_s H_p(\bar{T}) = \begin{cases} I(\bar{T}) - A_i - D_i, & \text{if } I(\bar{T}) > A_i + (1 + t_h) D_i \\ \frac{I(\bar{T}) - A_i - (1 + t_h) D_i}{4 A_i D_i}, & \text{else if } I(\bar{T}) > A_i + (1 - t_h) D_i \\ 0, & \text{otherwise} \end{cases}$$

where $t_s$ and $t_h$ are small positive thresholds that control the smoothness of the shadow and highlight. Figure 6 shows the result of shadow and highlight detection.

Refinement of Reflectance and Normals The hand-drawn surface normal is limited to representing rather coarse surface normals because it is still hard for a user to assign fine geometric details, e.g., bumps on a piece of fruit, such as an orange’s surface. This method estimates more precise surface normals and reflectance using an input image $I$ and the hand-drawn coarse surface normal map $\bar{N}$. Assuming that $\bar{T}$ is known and constant albedos $\bar{k}_i$ are obtained for labelled region $i$, we show that the surface reflectance and normal can be refined by the proposed refinement method. Due to the inaccuracy of the coarse surface normal $\bar{N}$ and constant albedo $\bar{k}_i$ in each label $i$, Eq. (3) becomes

$$e(\bar{k}_i, \bar{N}) = I(\bar{T}) - I_a \bar{k}_i - I_p S_p(\bar{k}_i, \bar{N}) + k_s H_p,$$

where $e$ is the residue caused by the error in surface normal and surface reflectance. We use the residue $e$ to refine both surface albedo $\bar{k}_i$ and normal $\bar{N}$. To determine the source
of the error, i.e., surface normal or reflectance, we propose to use a chromaticity similarity measure in the original image. In our method, the chromaticity similarity $\alpha$ between a pixel $\vec{r}$ and adjacent pixels $\vec{s} \in A(\vec{r})$ is evaluated in the specified image area as follows:

$$\alpha(\vec{r}) = \frac{1}{n} \sum_{\vec{s} \in A(\vec{r})} \text{Similarity}(\vec{C}(\vec{r}), \vec{C}(\vec{s})), \quad (8)$$

where $\vec{C}(\cdot) = (R, G, B)/(0.3R + 0.6G + 0.1B)$ represents chromaticity and $n$ represents the number of adjacent pixels $\vec{s}$. The function $\text{Similarity}(\cdot, \cdot)$ measures the distance between two chromaticities. In our case, we use the inner product of normalized vectors to measure the chromaticity similarity as $\text{Similarity}(\vec{a}, \vec{b}) = \langle \vec{a}, \vec{b} \rangle / (\|\vec{a}\|\|\vec{b}\|)$.

The residue $e$ is then distributed to surface normal $\vec{N}$ and reflectance $\vec{K}$, weighed by $\alpha$ for each pixel by plugging $e(k, \vec{N}) = (1 - \alpha)e(\vec{k}, \vec{N})$ and $e(\vec{k}, \vec{N}) = \alpha e(\vec{k}, \vec{N})$ into Equation (7) and solving respectively for each $k$ and $\vec{N}$. To obtain the refined surface normal $\vec{n}$ from $e(\vec{K}, \vec{n})$ with Equation (7), the coarse normal $\vec{N}$ is rotated on the plane spanned by $\vec{l}$ and $\vec{N}$. There are two solutions of $\vec{n}$, and the solution of $\vec{n}$ which has a smaller deviation from $\vec{N}$ is taken by evaluating $\vec{n} = \text{argmin}(\vec{n}, \vec{N})$. In this way, fine reflectance and normal map which are necessary for relighting are obtained. Figure 7 shows the result of surface normal refinement. Detailed surface normal of bumps of the orange is estimated with the proposed method.

3. Results

To evaluate the performance of the proposed method, we have conducted extensive experiments on various pho-
Figure 10. Adjustment of photometric consistency in superimposition application. The foreground (the girl in the top-left image) is superimposed on the different background (bottom and right image). The illumination condition of foreground (girl) is aligned to that of the background with our relighting method.

tographs. We have developed a complete system to achieve the goal and all the relighting results are rendered without any other photo-editing tools. Figure 8 displays the result of proposed relighting method. The images in the top row show the original input images, the second row shows the normal maps used for relighting, the third row shows the estimated reflectance images, and the relighting results are shown in the bottom row. To produce the relighting result, Equation (3) is used for rendering.

For surface normal assignment, like existing 3D modellers and image-based modeling methods that allow a user to assign depth values, in practice, our method is limited to rather simple scenes where image segmentation is feasible. As mentioned in [17], image segmentation usually takes many hours. The editing time of our results is also roughly a few hours on average. Ignoring the image segmentation process, we found that the normal assignment is much easier than existing methods that assign full 3D or depth by 3D modellers and depth painting method.

At the boundaries of objects where surface normals are not smooth, sometimes small artifacts due to the inaccuracy of the surface normal boundary are observed. To handle this problem, pixels that have a significant difference in chromaticity between themselves and the adjacent pixels compared to that in the original image are replaced with the smooth interpolation of adjacent pixel values weighted by the chromaticity similarity in the original image. To render highlights, we used a simple Phong model in the experiment:

$$H_s = \frac{I_p k_s (\vec{v} \cdot \vec{n})}{B_4 B_5}$$

where $\vec{v}$ and $\vec{m}$ are unit vectors of viewing and mirror reflection directions. $n_s$ is a user-defined shininess factor. In this way, $H_s$ is computed for each pixel under a new lighting condition. Figure 9 shows the close-up view of the orange surface in the fruit scene. As we can see clearly in the image, highlight and small bumps on the orange surface are naturally rendered.

Since the scene structure is represented by surface normals in our method, it is inherently difficult to generate cast shadows that are not local lighting effects. However, when the scene surface is smooth, we can compute cast shadow by taking the integral of surface normals. We have adopted a method proposed by Robertson et al. [18] to achieve cast shadow rendering. The relighting results with cast shadow
is shown on the bottom row of Figure 8.

Our method is useful for illumination alignment for photo-object superimposition (Figure 10). With our relighting method, photometric consistency between foreground and background is adjusted and visible in the illustration. In this example, we superimposed the foreground onto the background image with adjusting the illumination condition of the foreground.

4. Conclusion and Future Work

This paper presents a new approach to single-view relighting. We provide a set of input methods for interactive assignment of surface normal. An automatic method for estimating surface reflectance and detailed surface normal from the coarse surface normal input is also proposed. The proposed method uses surface normals to represent the scene geometry in an image, which is sufficient for relighting where local illumination effect is dominant. Two key techniques are developed: (1) Smooth flood-fill tool which propagates sparse normal input to smoothly fill-in the specified image area, and (2) Algorithm to refine surface reflectance and normal with the input image and the hand-drawn coarse surface normal. The applicability of the method is evaluated by relighting a variety of photographs and the application of illumination adjustment in photo superimposition.

There are many promising directions for future research. One is to investigate more efficient tools for surface normal assignment. Our current system provides normal modeling tools based on detectable pen angle of a pen-tablet device. It may be also possible to develop UIs to assign surface normals in a copy-and-paste manner by referencing well-known shapes such as cube, sphere, etc. We are also interested in other interactive tools such like Teddy system [7].

To simplify the problem, we limit our method to apply to scenes which are governed by a dominant directional lighting with an ambient lighting and are composed of surface materials that can be approximated by the Phong’s model. Extending the method to work with more complex lighting models and BRDF models is the other direction of future research, which also makes the refinement algorithm of reflectance and normal maps more complex but may produce more physically correct relighting results.

References


