Image-Based Surface Detail Transfer

Zicheng Liu and Zhengyou Zhang
Microsoft Research

Ying Shan
Sarnoff Corporation

A novel technique transfers geometric details from one surface to another with simple 2D image operations.

Changing an object's appearance by adding geometric details is desirable in many real-world applications. We might want to know, for example, how a wall would look after adding some geometrical bumps, or how a person would look after adding wrinkles on his or her face.

Direct methods for adding geometric details to an object require modeling the object and the surface details. It's usually not a trivial task to build a 3D model for a real object. It's also tedious and labor intensive to model and create surface details with existing geometric modeling tools.

Bump mapping has acted as an alternative to adding geometrical details to an otherwise smooth object. But constructing visually interesting bump maps requires practice and artistic skills.

Computer vision techniques have helped for modeling real-world objects and their surface details. Techniques include the use of laser scanner, stereo algorithms, shape from lighting variation,[1] shape from shading,[2] and so on. However, some of these techniques require specialized equipment. Many other techniques require at least two images of each object, and capturing high-resolution geometrical details robustly can be difficult.

Although the shape-from-shading technique only requires a single image, this approach usually requires knowledge of the lighting condition and reflectance functions of the surface material. Three-dimensional recognition and reconstruction from a single image have used linear classes of objects.[3,4] But this reconstruction technique doesn't generate accurate geometric details.

In our approach, for cases where we only want to transfer geometrical details from one object to another, we might not need to explicitly compute 3D structures. In particular, our novel technique captures the geometrical details of an object from a single image in a way that is independent of the object's reflectance property. We can then transfer geometrical details to another surface, producing the appearance of a new surface with added geometrical details while preserving the object's reflectance property. Our method's advantages are that it's simple to implement, requires only a single image for each object, and produces effective results.

An earlier version of this article appeared in the *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.[5] In this revised version, we show that this technique is suitable as a fast and effective touch-up tool for photo editing. In particular, it can remove wrinkles and color spots on face images.

Image-based surface detail transfer

Our technique—image-based surface detail transfer (IBSDT)—transfers geometric details between the images of two surfaces without knowing their 3D information.

Notations and problem statement

For any point P on a surface S, let $\mathbf{n}(P)$ denote its normal. Assume there are m point light sources. Let $\mathbf{l}_i(P)$, $1 \le i \le m$ denote the light's direction from P to the ith light source and l_i its intensity. Suppose the surface is diffused, and let $\rho(\mathrm{P})$ be its reflectance coefficient at $\mathbf{P}$. Under a Lambertian model, the recorded intensity of point P in image I is

$$I(\mathbf{p}) = \rho(P)\sum_{i=1}^{m} l_i \mathbf{n}(P)\cdot\mathbf{l}_i(P) \qquad (1)$$

where $\mathbf{p} = \mathbf{C}(\mathrm{P})$ is the 2D projection of P onto the image, and $\mathbf{C}(\cdot)$ is the camera projection function.

Two surfaces S_1 and S_2 are aligned if there exists a one-to-one mapping F such that for all $P_1 \in S_1$ there exists $P_2 = F(P_1) \in S_2$ such that

$$||P_1 - P_2|| \le \varepsilon \qquad (2)$$

where ε is a small positive. Furthermore, there exist neighborhoods $\Theta(P_1)$ of P_1 and $\Theta(P_2)$ of P_2 such that

$$||\bar{\mathbf{n}}(P_1) - \bar{\mathbf{n}}(P_2)|| \le$$

Published by the IEEE Computer Society

where δ is a small positive, and $\bar{\mathbf{n}}(P_1)$ and $\bar{\mathbf{n}}(P_2)$ are the mean normal defined in the neighborhoods of $\Theta(P_1)$ and $\Theta(P_2)$, respectively.

We can then state the problem as the following. Given images I_1 and I_2 of two aligned surfaces S_1 and S_2, respectively, what is the new image I_2' of S_2 if we modify its surface normal such that

$$\mathbf{n}_2'(P_2) = \mathbf{n}_1(P_1) \tag{3}$$

where P_1 and P_2 are the corresponding points defined by the mapping function F.

Geometric viewpoint

The following discussion assumes a single-point light source to simplify the derivation. Extension to multiple light sources is straightforward. Because the distance between P_1 and P_2 is small according to Equation 2, it's reasonable to assume that the light is always sitting far away enough such that $\varepsilon << d_l$, where d_l is the average distance from light to the points. This leads to the approximation $\mathbf{l}(P_1) \approx \mathbf{l}(P_2)$. From Equations 1 and 3, we then have

$$\begin{aligned}\frac{I_2'(\mathbf{p}_2)}{I_2(\mathbf{p}_2)} &\equiv \frac{\rho(P_2)l\mathbf{n}_2'(P_2)\cdot\mathbf{l}(P_2)}{\rho(P_2)l\mathbf{n}_2(P_2)\cdot\mathbf{l}(P_2)} \\ &\approx \frac{\rho(P_1)l\mathbf{n}_1(P_1)\cdot\mathbf{l}(P_1)}{\rho(P_2)l\mathbf{n}_2(P_2)\cdot\mathbf{l}(P_2)}\frac{\rho(P_2)}{\rho(P_1)} \\ &\equiv \frac{I_1(\mathbf{p}_1)\rho(P_2)}{I_2(\mathbf{p}_2)\rho(P_1)}\end{aligned} \tag{4}$$

where ρ has the same meaning as in Equation 1; $\mathbf{p}_1 = \mathbf{C}_1(P_1)$; $\mathbf{p}_2 = \mathbf{C}_2(P_2)$; and I_1, I_2, and I_2' have the same meaning as in the problem statement. Notice that the $\mathbf{C}(\cdot)$ functions are different for the two surfaces. This is because two different cameras could have taken surface images I_1 and I_2. This leads to

$$I_2'(\mathbf{p}_2) \approx \frac{I_1(\mathbf{p}_1)\rho(P_2)}{\rho(P_1)} \tag{5}$$

To compute the ratio of $\rho(P_1)$ and $\rho(P_2)$, let us define the smoothed image of I as

$$\bar{I}(\mathbf{p}) \equiv \sum_{\mathbf{q}\in\Omega(\mathbf{p})} w(\mathbf{q})I(\mathbf{q}) \tag{6}$$

where $\Omega(\mathbf{p}) = \mathbf{C}(\Theta(P))$ is the neighborhood of $\mathbf{p}$, and $w(\cdot)$ is the kernel function of a smooth filter such as a Gaussian or average filter. Assuming that the size of $\Theta(P)$ is relatively small compared to its distance to the light source, we have $\mathbf{l}(P) \approx \mathbf{l}(Q)$, $\forall Q \in \Theta(P)$. Also assuming that $\rho(P) \approx \rho(Q)$, $\forall Q \in \Theta(P)$, from Equations 1 and 6, it's then obvious that

$$\bar{I}(\mathbf{p}) \approx \rho(P)l\left(\sum_{Q\in\Theta(P)} w(\mathbf{C}(Q))\mathbf{n}(Q)\right)\cdot\mathbf{l}(P) \tag{7}$$

It turns out that

$$\sum_{Q\in\Theta(P)} w(\mathbf{C}(Q))\mathbf{n}(Q) = \bar{\mathbf{n}}(P)$$

where $\bar{\mathbf{n}}(P)$ is the mean normal as mentioned in the problem statement. Equation 7 becomes $\bar{I}(\mathbf{p}) \approx \rho(P)l\,\bar{\mathbf{n}}(P)\cdot\mathbf{l}(P)$, which means that smoothing in the image space is equivalent to smoothing the 3D surface if the reflectance is constant in the neighborhood. For surfaces S_1 and S_2, we then have

$$\frac{\bar{I}_2(\mathbf{p}_2)}{\bar{I}_1(\mathbf{p}_1)} \approx \frac{\rho(P_2)l\bar{\mathbf{n}}(P_2)\cdot\mathbf{l}(P_2)}{\rho(P_1)l\bar{\mathbf{n}}(P_1)\cdot\mathbf{l}(P_1)} \tag{8}$$

Since the two surfaces are aligned, we have $\mathbf{l}(P_1) \approx \mathbf{l}(P_2)$, and $\bar{\mathbf{n}}(P_2) \approx \bar{\mathbf{n}}(P_1)$. We can then rewrite Equation 8 as

$$\frac{\rho(P_2)}{\rho(P_1)} \approx \frac{\bar{I}_2(\mathbf{p})}{\bar{I}_1(\mathbf{p})} \tag{9}$$

Substituting Equation 9 into Equation 5 leads to

$$I_2'(\mathbf{p}_2) \approx \frac{I_1(\mathbf{p}_1)}{\bar{I}_1(\mathbf{p}_1)}\bar{I}_2(\mathbf{p}_2) \tag{10}$$

Equation 10 shows that we can approximate the transfer of the surface normal by some simple operations on the surfaces' images.

Intuitive signal-processing viewpoint

We now rewrite Equation 10 as

$$I_2'(\mathbf{p}) \approx \frac{I_1(\mathbf{p})}{\bar{I}_1(\mathbf{p})}\bar{I}_2(\mathbf{p}) \equiv \left(1 + \frac{I_1(\mathbf{p}) - \bar{I}_1(\mathbf{p})}{\bar{I}_1(\mathbf{p})}\right) \tag{11}$$

From a signal-processing viewpoint, Equation 11 substitutes the high-frequency components of I_2 with those from I_1. The high-frequency components $I_1 - \bar{I}_1$ in I_1 are normalized by $\bar{I}_1$ to cancel the intensity scale difference between the low-frequency components of I_2 and I_1. Generally, I_1 could be any image, regardless of the conditions we gave in the previous "Geometric viewpoint" section. But the resultant image could be meaningless because of the inconsistency between the transferred detailed components from I_1 and native low-frequency components on I_2. This could occur when I_1 and I_2 are the images of two surfaces that are not aligned.

Implementation

Given images I_1 and I_2 of similar shapes, to perform surface detail transfer, we first align the two images. For simple geometrical shapes such as rectangles and spheres, we usually only need to perform global transformations including rotation, translation, and scaling. For more complicated shapes such as human faces, we first manually put markers on the boundaries and the feature points, and then obtain pixel alignment through image warping.[6] In our implementation, we use a simple triangulation-based image-warping method. Once we align the images, we can run a Gaussian filter with a user-specified σ on I_1 and I_2 to obtain $\bar{I}_1$ and $\bar{I}_2$. Finally, we apply Equation 10 to obtain I_2'.

Intuitively, the σ of the Gaussian filter controls how

(a) (b) (c)

1 (a) Synthetic sphere with many small bumps generated by using a bump-mapping technique. (b) Photograph of a nectarine. (c) Result obtained by transferring the geometrical details of the sphere to the nectarine.

(a) (b) (c)

2 (a) Photograph of an orange. (b) Photograph of a nectarine. (c) Synthesized image obtained by transferring the geometrical details of the orange to the nectarine.

(a) (b) (c)

3 (a) Photograph of a tissue. (b) Synthesized image of a rectangle. (c) Image obtained by transferring the geometrical details of the tissue to the rectangle.

(a) (b) (c)

4 (a) Photograph of a tissue. (b) Image of a piece of wood. (c) Synthesized image obtained by transferring the geometrical details of the tissue to the wood.

much geometrical smoothing we perform on the surface of I_1. So the σ determines the transferred surface details' scale. A small σ allows transfer of fine geometrical details; a large σ allows only large-scale geometrical deformations.

Results

Figure 1 shows the results of transferring the geometrical details of a synthetic sphere to a real nectarine. We generated the bumps on the synthetic sphere using a bump-mapping technique, and set the sphere's surface reflectance property as uniform. We put a point light source on top of the sphere so that its lighting condition is somewhat close to the nectarine's lighting condition. The bumps on the synthetic sphere transferred nicely to the nectarine except at the bottom where the synthetic sphere is dark. The image sizes are 614×614 pixels, and the σ is 8.

Figure 2 shows the results of transferring the geometrical details of a real orange to Figure 1's nectarine. The bumps on the oranges transferred faithfully to the nectarine. The image dimensions and the σ are the same as in Figure 1. This example also reveals a limitation of our algorithm: The highlights on the orange transferred to the nectarine. This is because our algorithm treats the highlights as if they were caused by geometrical variations.

Figure 3 shows the results of transferring the geometrical details of a tissue to a synthetic rectangle. Only the geometrical bumps on the tissues transferred to the rectangle, while the material color of the rectangle is preserved.

Figure 4 shows the results of geometric detail transferring from the Figure 3's tissue to the image of a piece of wood. We took both pictures under the same lighting conditions. The small bumps on the tissue are transferred to the wood, while the wood color and global texture are preserved.

Figure 5 shows the result of transferring the same tissue's geometrical details to a table surface. This table surface has a different texture pattern than the wood in Figure 4. Comparing the result in Figure 4c with that in Figure 5c, we see that they have the same geometrical bumps but different material properties.

Aging effect synthesis

One interesting application of IBSDT is aging effect synthesis. Geometrically, older people have more bumps on their facial skin surface than a young person. If we transfer the bumps from an old person's skin surface to a young person's face, the young person's face becomes bumpy and looks older. Conversely, we can replace an old person's skin surface bumps with a young person's skin surface so that the old person's face gets smoother and looks younger. We can apply the surface details transfer technique, described previously, on human faces to generate aging effects. We align the images by first marking face boundaries and features such as eyes, noses, and mouths, and then use triangulation-based image warping to warp I_1 toward I_2. We only apply IBSDT to pixels inside of the face boundary. In addition, IBSDT doesn't modify the pixels in the regions of the eyebrows, eyeballs, nose top, and mouth.

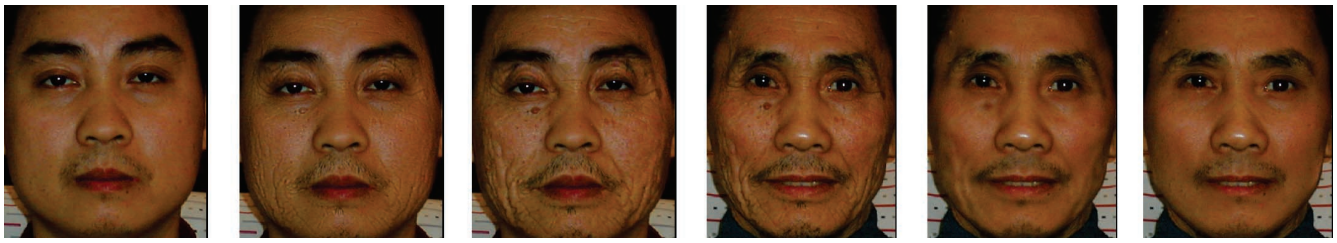
Figure 6 shows the aging effect synthesis results between the faces of a young male (Figure 6a) and an old male (Figure 6d). We took both images in a room with multiple light sources. For each face, we experimented with different σ of the Gaussian filter during the

surface detail transfer. Figures 6b and 6e show the results when $\sigma = 3$ and Figures 6c and 6f show $\sigma = 8$. Varying σ produces a reasonable, in-between aging effect such as with Figures 6b and 6e.

Obviously, surface detail transfer plays an important role when making a young person look older. However, why this technique is necessary in making an old person look younger might seem less apparent. To clarify this point, we simply smooth Figure 6d without transferring surface details from Figure 6a, while masking out the facial features as before. Figure 7a shows the results with $\sigma = 3$ and Figure 7b shows $\sigma = 8$. By comparing these results with those shown in Figures 6e and 6f with the same σs, we clearly see that the images in Figure 6 are much sharper and more convincing.

(a) (b) (c)

5 (a) Image of a tissue. (b) Image of a table surface. (c) Synthesized image obtained by transferring the geometrical details of the tissue to the table surface.

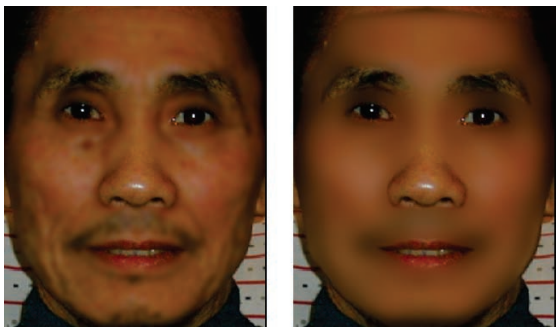


(a) (b) (c) (d) (e) (f)

6 Young adult versus a senior adult. (a) Face of a young adult. (b) Simulated old face of Figure 6a with a small σ. (c) Simulated old face of Figure 6a with a large σ. (d) Face of a senior adult. (e) Simulated young face of Figure 6d with a small σ. (f) Simulated young face of Figure 6d with a large σ.

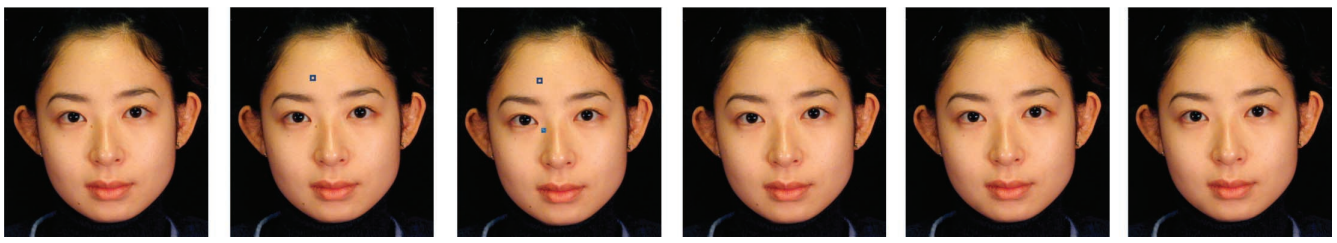
Photo touch-up

Finally, we can use this technique as a touch-up tool to remove wrinkles and color spots. The user first selects a source point on any image (which can be the same as the target image). Then he or she selects a target rectangle on the target image by clicking and dragging the mouse. Releasing the mouse applies the IBSDT operation from the source rectangle (centered at the source point with size equal to the destination rectangle) to the destination rectangle. The user can then perform a linear fade-in-fade-out blending along the boundary of the destination rectangle. Figures 8a through d show an example of this technique used to remove a color spot on the subject's forehead. Figure 8e uses the same procedure and source point to remove a different color spot (in her right chin area). Figure 8f shows removal of the



(a) (b)

7 Senior adult to young adult without using image-based surface detail transfer. (a) Result with $\sigma = 3$; (b) Result with $\sigma = 8$. The input face image is Figure 6d.

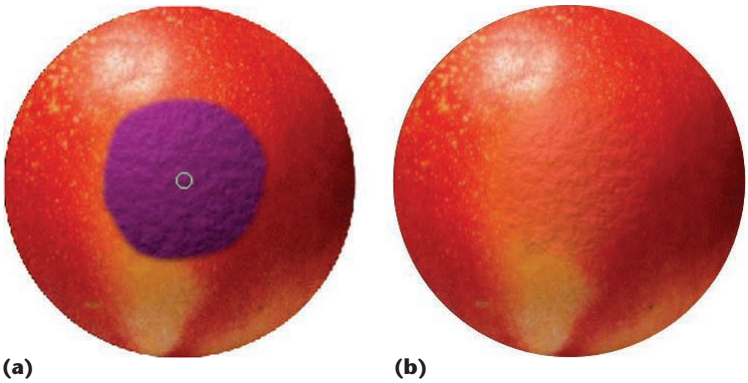


(a) (b) (c) (d) (e) (f)

8 Using IBSDT as a touch-up photo editing tool. (a) Input image. (b) Source point (the center of the small rectangle on her forehead). (c) Destination rectangle (small rectangle on top of a color spot close to her eye corner). (d) Result after applying IBSDT. (e) Removing the second spot in the chin area. (f) Removing the third spot below her right eye.

(a) (b)

9 Removing wrinkles and color spots of a senior person. (a) Input image. (b) Result after applying the touch-up tool on some of his color spots and his forehead area.



(a) (b)

10 Applying Photoshop 7.0's healing brush from the bumpy sphere to the nectarine shown in Figure 1. (a) Applied region and (b) result.

color spot below her right eye. We can see that all three of the color spots were removed successfully.

Figure 9 shows the result of applying the touch-up tool to a senior person to remove the wrinkles on his forehead as well as some color spots.

Although some similarity exists between the functionality of our touch-up operation and that of the healing brush in Adobe Photoshop 7.0, the result is sometimes quite different. Figure 10 shows an example of using the healing brush of Photoshop 7.0 on the bumpy sphere and the nectarine as shown in Figure 1. The image in Figure 10a shows the region where the brush is applied; Figure 10b is the result. Comparing Figure 10b with Figure 1b, we can see that Photoshop 7.0 does not preserve the original texture of the nectarine while ours does.

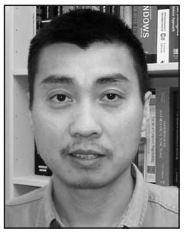
Future work

One limitation of IBSDT is that it requires similar lighting conditions between the two images. For images taken under completely different lighting conditions, we could use relighting techniques such as those reported in Marschner, Guenter, and Raghupathy,[7] Debevec et al.,[8] and Riklin-Raviv and Shashua (see "Related Work" sidebar) before applying our IBSDT technique.

Another limitation is that it assumes a smooth surface reflectance. Our algorithm regards abrupt reflectance changes—such as small color spots—in objects as geometrical details. As a result, our algorithm can transfer or remove color spots. This is both a feature and a limitation. We can use this feature to remove the color spots, as we've shown in this article. However, transferring color spots is usually not desirable. To avoid transferring color spots, we could potentially separate these color variations from geometry variations through learning approaches. We are planning on pursuing this in the future. ■

References

1. H. Rushmeier, G. Taubin, and A. Gueziec, "Applying Shape from Lighting Variation to Bump Map Capture," *Proc. Eurographics Workshop on Rendering*, Springer, 1997, pp. 35-44.
2. B. Horn and M.J. Brooks, *Shape from Shading*, MIT Press, 1989.
3. T. Vetter and T. Poggio, "Linear Object Classes and Image Synthesis from a Single Example Image," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, 1997, pp. 733-742.
4. V. Blanz and T. Vetter, "A Morphable Model for the Synthesis of 3D Faces," *Proc. Siggraph*, ACM Press, 1999, pp. 353-360.
5. Y. Shan, Z. Liu, and Z. Zhang, "Image-Based Surface Detail Transfer," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, IEEE Press, 2001, pp. 794-799.
6. T. Beier and S. Neely, "Feature-Based Image Metamorphosis," *Proc. Siggraph*, ACM Press, 1992, pp. 35-42.
7. S.R. Marschner, B. Guenter, and S. Raghupathy, "Modeling and Rendering for Realistic Facial Animation," *Rendering Techniques*, SpringerWein, 2000, pp. 231-242.
8. P.E. Debevec et al., "Acquiring the Reflectance Field of a Human Face," *Proc. Siggraph*, ACM Press, 2000, pp. 145-156.



Zicheng Liu *is a researcher at Microsoft Research. His research interests include face modeling and animation, face relighting and expression synthesis, linked figure animation, multisensory technology, and communication and collaboration. Liu received a BS from Huazhong Normal University, China, in mathematics; an MS from the Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, in operational research; and a PhD from Princeton University in computer science.*