Image Reconstruction Invariant to Relighting

T. Georgiev†

Adobe Systems, 345 Park Ave, W10-124, San Jose, CA 95110, USA

Abstract

This paper describes an improvement to the Poisson image editing method for seamless cloning. Our approach is based on minimizing an energy expression invariant to relighting. The improved method reconstructs seamlessly the selected region, matching both pixel values and texture contrast of the surrounding area, while previous algorithms matched pixel values only.

Our algorithm solves a deeper problem: It performs reconstruction in terms of the internal working mechanisms of human visual system. Retinex-type effects of adaptation are built into the structure of the mathematical model, producing results that change covariantly with lighting.

1. Introduction

During the last five years there has been significant progress in the area of removing scratches, wires and other objects from images and video. As a result today there are a number of highly effective approaches to achieve this effect. Among them are Projections Onto Convex Sets [HT96], texture synthesis [EL99, WL00] inpainting [BSCB00, BBS, BVSO03], Poisson editing [Ado02, PGB03, Geo04].

This paper focuses on improving the aesthetic quality of Poisson editing. In Poisson editing the defective pixels are replaced with new pixels described by a finction $f(x,y)$ which is a solution of the Poisson equation

$$\triangle f(x,y) = \triangle g(x,y) \tag{1}$$

with Dirichlet boundary condition constraining the new $f(x,y)$ to match the original image at the boundary. In (1) $g(x,y)$ is the texture that is "seamlessly-cloned" into the reconstructed area and

$$\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \tag{2}$$

The effect is as if cloning texture, while changing its color/shading so that it seamlessly matches surrounding pixels at every boundary point.

† e-mail: tgeorgie@Adobe.com

While this method works very well in many cases, sometimes we see problems. For example consider Figure 2. It is the result of Poisson cloning pebbles from the illuminated area into the shadow area. The method correctly matches pixel values at the boundary of the patch, but the cloned pebbles are still easy to spot. There is too much variation, too high contrast, in the reconstructed area of the image. This problem is inherent in the nature of the Poisson equation (1), which transfers *variations* of g directly, without modification. It would be desirable to modify the right hand side of (1) so that variations fit surroundings more seamlessly.

We can ask the question, what would be the best expression for the right hand side of the Poisson equation (1), which would produce the highest quality seamless cloning? The type of problems described above is related to lighting conditions. This suggests the idea that the solution might be best described by a differential equation minimizing some energy expression that is invariant to illumination.

2. Approach

Our approach is to solve the problem of reconstruction based on the internal working mechanism of (a model of) the human visual system.

There is a well-known (approximate) invariance of the internal image we see with respect to change of lighting conditions. This invariance, or "color constancy", is due to the adaptation of the visual system, and it has been extensively discussed in relation to Retinex theory [Lan77, Hor74].

In this paper we observe that it is possible to cast the adaptation problem into the mathematical framework of the vector-covector formalism. In our approach luminance is a vector, retina sensitivity is a covector, and the formalism automatically handles relighting invatriance.

At each pixel (x,y) the image is a vector $f(x,y)$. For example, f could be a vector describing the three color channels, $f = (R,G,B)$. Relighting transforms this vector f multiplicatively. We say that f changes *covariantly* with illumination.

In the case of grayscale images, f is a 1-D vector. Note that a 1-D vector is different from a scalar. A scalar is an invariant quantity, while a vector changes covariantly with lighting. Unlike scalars, a 1-D vector does not have a numerical value until a basis, or a scale, is chosen. In our case this basis, or measure of scale at each point, is the *retina sensitivity*. The same vector can be seen as different *lightness* (perceived brightness) depending on adaptation to bright or dark environment.

Due to adaptation, at each pixel the retina sensitivity $\varphi(x,y)$ transforms in opposition to, or *contravariantly* with, illumination. Adapted sensitivity is a covector in the sense that a contraction, or "dot product", with the vector $f(x,y)$ must produce the *invariant* lightness that we see in the visual system. It is as a scalar invariant $\varphi \cdot f$.

The true luminance vector, $f(x,y)$ may be captured by the camera or other physical device with known sensitivity. But luminance is not observable by the human visual system. What we actually see is the scalar (i.e. invariant) quantity $\varphi \cdot f$, which is the lumninance vector acting on the sensitivity covector. Sometimes we will denote this as $< \varphi f >$ in order to make the contraction of vector and covector explicit.

To avoid confusion, we note that in the case of grayscale images the only vector and covector properties that we are going to use are that f transforms covariantly with lighting (i.e. as a vector), and φ transforms contravariantly with lighting (as a covector), due to adaptation. Lightness is a scalar $< \varphi f >$, and as such it does not change with relighting. We observe only lightness; luminance and sensitivity are not observable independently.

3. Main Equations

The simplest way to reconstruct or inpaint the area of a scratch in an image is to replace defective pixels with a solution of the Laplace equation

$$\triangle f = 0 \tag{3}$$

with Dirichlet boundary conditions. This is equivalent to minimizing the energy expression

$$\int (\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2 dxdy \tag{4}$$

in the selected area of the scratch, where we calculate pixel values representing smallest sum of the gradients squared. In other words, we are requiring that sum of the lengths squared of the gradients be minimal.

The important observation here is that this energy is not invariant under relighting. As such, it is not appropriate as a model of the invariant (after adaptation) image that humans see.

Let's use a related invariant expression, written in terms of lightness, which is a scalar (invariant) relative to relighting

$$\int (\frac{\partial}{\partial x} < \varphi f >)^2 + (\frac{\partial}{\partial y} < \varphi f >)^2 dxdy. \tag{5}$$

Minimizing this energy means sum of the lengths squared of the perceived gradients is minimal. Writing the Euler-Lagrange equation (varying f) we get:

$$\triangle f + \frac{2}{\varphi} grad\, f \cdot grad \varphi + \frac{f}{\varphi} \triangle \varphi = 0, \tag{6}$$

where f is pixel value (luminance) and φ is retina sensitivity. This can be written in a simpler form:

$$\triangle(\varphi f) = 0. \tag{7}$$

To find one possible φ, assume adaptation to a "neutral" area of pure texture $g(x,y)$ in the image. Adaptation means $\varphi \cdot g = const$. Since in (6), (7) φ is defined up to a multiplicative constant, we can safely assume it to be 1, and then $\varphi = 1/g$.

Substituting in (7), we get our main equation

$$\triangle \frac{f}{g} = 0. \tag{8}$$

For theoretical reasons it is good to know that this equation is equivalent to

$$\triangle f - 2 grad\, f \cdot \frac{grad g}{g} - f \frac{\triangle g}{g} + 2f \frac{(grad g) \cdot (grad g)}{g^2} = 0. \tag{9}$$

In other words, we have found the correct right hand side for the relighting-covariant version of Poisson equation (1)

$$2 grad\, f \cdot \frac{grad g}{g} + f \frac{\triangle g}{g} - 2f \frac{(grad g) \cdot (grad g)}{g^2}, \tag{10}$$

that replaces $\triangle g$.

Also, note that equation (9) can be written as

$$(\frac{\partial}{\partial x}+A_x)(\frac{\partial}{\partial x}+A_x)f+(\frac{\partial}{\partial y}+A_y)(\frac{\partial}{\partial y}+A_y)f=0, \quad (11)$$

with a "guidance field" similar to [PGB03]:

$$\mathbf{A}(x,y)=-\frac{grad g}{g}. \quad (12)$$

In mathematics this is called *the connection form* and it defines a modified, *covariant derivative*, which replaces conventional derivatives as in (11). These describe perceived gradients as opposed to true gradients in images. See [Sau89, Geo05].

4. Results

Figure 1: *Original image of pebbles and a scratch.*

Figure 2: *Scratch removed by Poisson cloning from the illuminated area.*

Traditional Poisson cloning between areas of different illumination can be a problem. To provide a clean example, we have tried to remove the scratch from the shadow area in Figure 1 using only source material from the illuminated area.

In Figure 2, we see the result of Poisson cloning from illuminated area into the shadow area. It correctly matches pixel values at the boundary of the patch, but the cloned pebbles are still easy to spot. There is too high contrast in the reconstructed area of the image.

Figure 3: *Scratch removed by covariant cloning from the same illuminated area as in Figure 2, based on equation (8).*

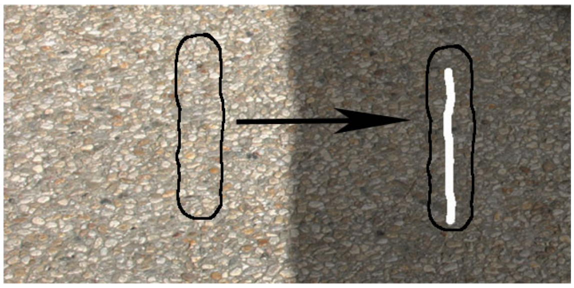

Figure 4: *Areas used for Poisson cloning in Figure 2 and covariant reconstruction, Figure 3.*

Figure 3 shows the result of our proposed illumination invariant method, equation (8). We see that not only the lighting is correct, but the method was able to clone texture in a really seamless way. Our experiments show that in a wide range of images (8) performs much better than the Poisson equation in terms of producing seamless cloning.

Figure 5 is another comparison between covariant cloning and Poisson cloning.

Implementation can be done in the following 3 steps:

(1) Divide the image by the sampling (texture) image, in which pixel value zero is replaced with a small number. This produces the first intermediate image $I_1(x,y)$.

$$I_1(x,y)=\frac{f(x,y)}{g(x,y)} \quad (13)$$

(2) Solve the Laplace equation for the second intermediate image

$$\triangle I_2(x,y)=0, \quad (14)$$

with Dirichlet boundary conditions defined by $I_1(x,y)$ at the boundary of the reconstruction area.

(3) Multiply the result by the texture image $g(x,y)$

Figure 5: *Left: Poisson cloning from selected area, and right: covariant cloning from the same area.*

$$h(x,y) = I_2(x,y)g(x,y), \tag{15}$$

and substitute the original defective image $f(x,y)$ with the new image $h(x,y)$ in the area of reconstruction.

A multigrid approach to solving (14) with good performance is described in [PTVF92]. In practical terms, the tool works sufficiently fast for using it in interactive mode. For example, on a laptop running Windows XP with a 2 GHz Pentium 4 processor, applying a brush of radius 100 pixels takes less than 0.25 seconds to converge.

5. Conclusion and future work

We have been able to find energy expression for reconstruction, covariant with illumination. Our results are much better than Poisson cloning. The proposed invariant energy approach potentially applies to any image processing algorithm that can be defined in terms of energy and corresponding differential equations. We modify the energy to produce invariant energy and covariant equations. Results are better because now perception and adaptation are taken into account. In this way we have embedded adaptation of the visual system into the mathematical formalism of the problem.

References

[Ado02] ADOBE: *Photoshop 7.0 User Guide*. Adobe Systems, San Jose, 2002.

[BBS] BERTALMIO M., BERTOZZI A. L., SAPIRO G.: Navier-stokes, fluid dynamics, and image and video inpainting. *Proceedings of the International Conference on Computer Vision and Pattern Recognition , IEEE, Dec. 2001, Kauai, HI, volume I, pp. I-355-I362.*

[BSCB00] BERTALMIO M., SAPIRO G., CASELLES C., BALLESTER C.: Image inpainting. In *Siggraph 2000, Computer Graphics Proceedings* (2000), Akeley K., (Ed.), Annual Conference Series, ACM Press / ACM SIGGRAPH / Addison Wesley Longman, pp. 417–424.

[BVSO03] BERTALMIO M., VESE L., SAPIRO G., OSHER S.: Simultaneous structure and texture image inpainting. In *IEEE Transactions on Image Processing, 12(8)* (2003), pp. 882–889.

[EL99] EFROS A. A., LEUNG T. K.: Texture synthesis by non-parametric sampling. In *IEEE International Conference on Computer Vision* (Corfu, Greece, September 1999), pp. 1033–1038.

[Geo04] GEORGIEV T.: Photoshop healing brush: a tool for seamless cloning. In *Workshop on Applications of Computer Vision in conjunction with ECCV 2004* (Prague, May 2004), pp. 1–8.

[Geo05] GEORGIEV T.: Vision, healing brush, and fiber bundles. In *Proceedings of SPIE Vol. 5666* (San Jose, January 2005).

[Hor74] HORN B. K. P.: Determining lightness from an image. *Computer Graphics and Image Processing 3*, 1 (Dec. 1974), 277–299.

[HT96] HIRANI A. N., TOTSUKA T.: Combining frequency and spatial domain information for fast interactive image noise removal. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1996), ACM Press, pp. 269–276.

[Lan77] LAND E. H.: The retinex theory of color vision. *Scientific American 237*, 6 (Dec. 1977), 108–128.

[PGB03] PÉREZ P., GANGNET M., BLAKE A.: Poisson image editing. *ACM Transactions on Graphics 22*, 3 (July 2003), 313–318.

[PTVF92] PRESS W. H., TEUKOLSKY S. A., VETTERLING V., FLANNERY B. P.: *Numerical Recipes in C: The Art of Scientific Computing (2nd ed.).* Cambridge University Press, Cambridge, 1992. ISBN 0-521-43108-5.

[Sau89] SAUNDERS D. J.: *The geometry of jet bundles*. Cambridge Univ. Press, Cambridge, 1989.

[WL00] WEI L.-Y., LEVOY M.: Fast texture synthesis using tree-structured vector quantization. In *Siggraph 2000, Computer Graphics Proceedings* (2000), Akeley K., (Ed.), Annual Conference Series, ACM Press / ACM SIGGRAPH / Addison Wesley Longman, pp. 479–488.