Digital Doctoring: How to tell the real from the fake

We are living in a world where seeing is no longer believing – the technology that allows for digital media to be manipulated and distorted is developing at break-neck speeds. And at the same time our understanding of the technological, ethical, and legal implications is lagging behind. How is this technology affecting our society and how do we contend with the implications? **Hany Farid** describes the impact of digital tampering and the development of mathematical and computational algorithms to expose digital fakes.

Wars have produced some of the most memorable and powerful photographs. From a mere snapshot in space and time, these photographs seem to capture the very essence of the suffering of thousands, Figure 1. For this reason, these images hold a unique place in documenting our collective history.

For the past decade, Adnan Hajj, a renowned photographer, has produced striking war photographs from the on-going struggle in the Middle East. On August 7th of this year, the Reuters news agency published one of Hajj's photographs showing the remnants of an Israeli bombing of a Lebanese town, Figure 2. In the week that followed, hundreds of bloggers and nearly every major news organization reported that the photograph had been doctored with the addition of more smoke. The general consensus was one of outrage and anger – Hajj was accused of doctoring the image to exaggerate the impact of the Israeli shelling. The feeling was, given the sanctity of war photographs, that this manipulation was simply inexcusable. An embarrassed Reuters quickly retracted the photograph and removed from its archives nearly 1,000 photographs contributed by Hajj. A period of soul searching followed.

Photography lost its innocence many years ago. The nearly iconic portrait of the U.S. President Abraham Lincoln (circa 1860) was a fake, having been created by splicing together the head of Lincoln with the body of Southern politician John Calhoun, Figure 3. This fake was only the beginning of a long history of photographic trickery. In the early part of the 1900s Stalin famously had his enemies air-brushed out of photographs. Between 1917 and 1920 two young girls in Cottingley, Yorkshire created an international sensation when they released photographs purportedly showing tiny winged fairy creatures. It wasn't until 1984 that it was discovered that some of the most spectacular photographs of World War I aerial combat published in 1933 were fakes. The Brown Lady of Raynham, perhaps one of the most famous ghost images published in 1936, was created by superimposing two pictures on top of each other. And the list goes on and on – history is riddled with photographic tampering.

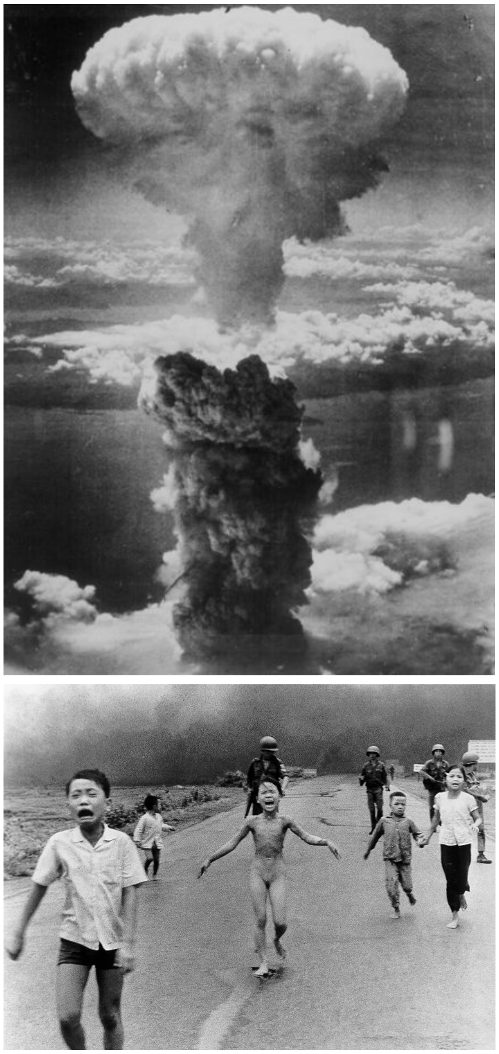

Figure 1: The mushroom cloud from the nuclear explosion over Nagasaki, August 9, 1945. A young girl flees from her village after being burned by napalm, June 8, 1972.

The case of Hajj is, of course, by no means unique. In 2003 Brian Walski, a veteran photographer of numerous wars doctored a photograph that appeared on the cover of the Los Angeles Times, Figure 2. After discovering the fake, the outraged editors of the LA Times fired Walski. The news magazines Time and Newsweek have each been rocked by scandal after it was revealed that photographs appearing on their covers had been doctored. And, in the past few years, countless news organizations around the world have been shaken by similar experiences [1].

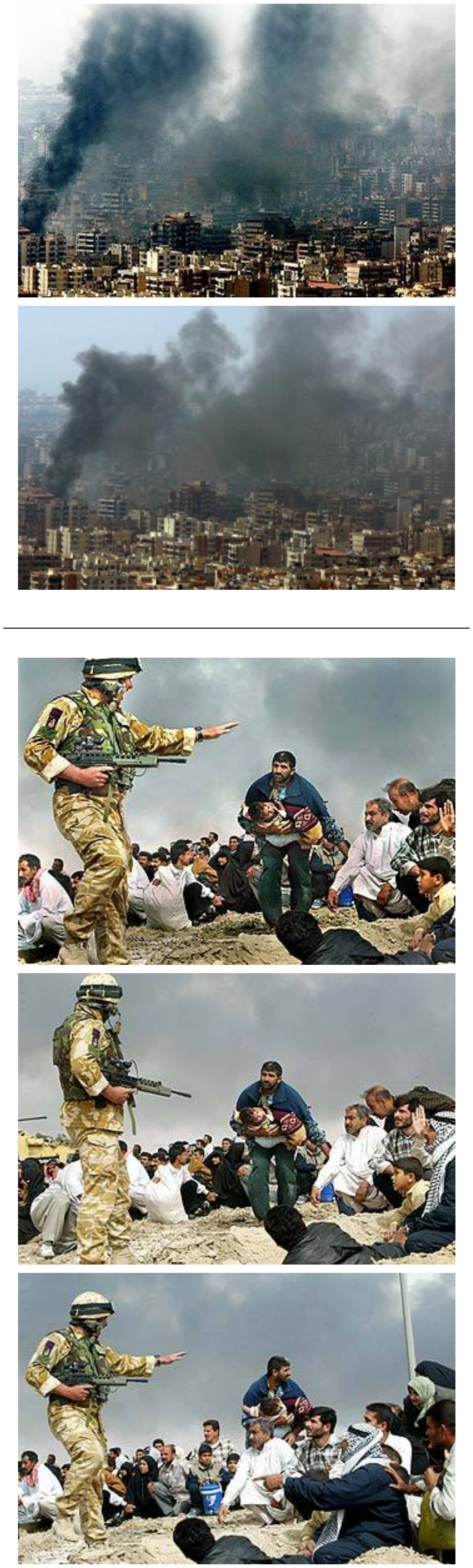

Figure 2: Shown from top to bottom are the published and original photos by Adnan Hajj, and the published photo by Brian Walski created from a composite of the two images shown below.

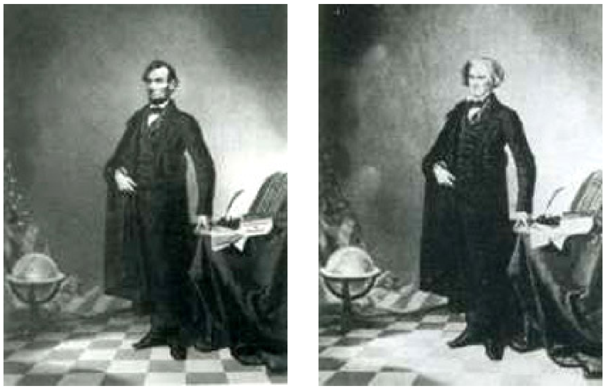

Figure 3: The 1860 portrait of President Abraham Lincoln and Southern politician John Calhoun.

The reality is that photo-journalists everywhere are altering, manipulating and distorting the images that we see every day.

Detecting Tampering

Cumbersome and time-consuming darkroom techniques were required to alter history on behalf of Stalin. Today, powerful and low-cost digital technology has made it far easier for nearly anyone to alter digital images. And, the resulting fakes are often very difficult to detect.

Over the past seven years my students and colleagues (Kimo Johnson, Siwei Lyu, Alin Popescu, Weihong Wang and Jeffrey Woodward), and I have been developing a suite of computational and mathematical techniques for detecting tampering in digital images. Our approach in developing each forensic tool is to first understand how a specific form of tampering disturbs certain statistical properties of an image, and then to develop a mathematical algorithm to detect this perturbation. I briefly describe three of these techniques, see [2] for more information on these and related work.

Cloning: In order to create more smoke in his photograph, Hajj cloned (duplicated) parts of the existing smoke using a standard tool in Photoshop, a popular photo-editing software. In this case the duplication was fairly obvious because of the nearly identical repeating patterns in the smoke. When care is taken, however, it can be very difficult to visually detect this type of duplication. We have developed a computer program that can automatically detect image cloning [3]. A digital image is first partitioned into small blocks. The blocks are then re-ordered so that they are placed a distance to each other that is proportional to the differences in their pixel colors. With identical and highly similar blocks neighboring each other in the re-ordered sequence, a region growing algorithm combines any significant number of neighboring blocks that are consistent with the cloning of an image region. Since it is statistically unlikely to find identical and spatially coherent regions in an image, their presence can then be used as evidence of tampering.

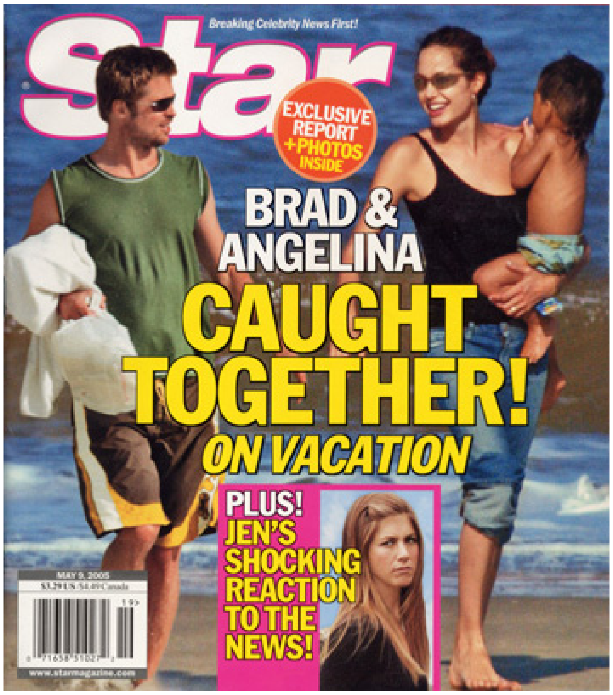
Star
Breaking Celebrity News First!
EXCLUSIVE REPORT +PHOTOS INSIDE
BRAD & ANGELINA
CAUGHT TOGETHER!
ON VACATION
PLUS! JEN'S SHOCKING REACTION TO THE NEWS!
MAY 9, 2005
www.starmagazine.com

Figure 4: The lighting of Pitt and Jolie is inconsistent in this composite.

Lighting: In April of 2005, the cover of the tabloid magazine Star featured a photograph of Brad Pitt and Angelina Jolie, at the time rumored to have a romantic relationship, Figure 4. The cover was sensational. It was also a fake – a digital composite of a picture of Pitt taken in the Caribbean in January of 2005 and a picture of Jolie taken in Virginia some time in 2004. A close examination reveals traces of tampering. The setting and shadows suggest that this photograph was taken outdoors on a sunny day. There are several clues in this photograph as to the location of the sun. Jolie's shadow cast onto the sand, the shadow under her chin, her evenly illuminated face, and the lighting gradient around her right leg, all suggest that she is facing the sun. Given this position of the sun, we would expect the right side of Pitt's face to be illuminated. It is not. It is in shadow, which is impossible. It is clear that Pitt is facing the sun, which places the sun at a location at least 90 degrees different than the position of the sun illuminating Jolie. Were the lighting differences in this image more subtle, our manual analysis would most likely have been insufficient. We have, therefore, developed a computer program that automatically estimates the direction of an illuminating light source for each object or person in an image [4]. By making some initial simplifying assumptions about the light and the surface being illuminated, we can mathematically express how much light a surface should receive as a function of its position relative to the light. A surface that is directly facing the light, for example, will be brighter than a surface that is turned away from the light. Once expressed in this form, standard techniques can be used to determine the direction to the illuminating light source for any object or person in an image. Any inconsistencies in lighting can then be used as evidence of tampering.

Re-touching: While attending a meeting of the United Nations Security Council in September of 2005, U.S. President George W. Bush scribbled a note to Secretary of State Condoleezza Rice. The note read "I think I may need a bathroom break. Is this possible?" Because the original image was overexposed, a Reuters' processor selectively adjusted the contrast of the notepad prior to publication. This form of photo retouching is quite common and can be used to alter a photograph in trivial or profound ways. We have developed a technique for detecting this form of tampering that exploits how a digital camera sensor records an image [5]. Virtually all digital cameras record only a subset of all the pixels needed for a full-resolution color image. Instead, only a subset of the pixels are recorded by a color filter array (CFA) placed atop the digital sensor, Figure 5. The most frequently used CFA, the Bayer array, employs three color filters: red, green, and blue. Since only a single color sample is recorded at each pixel location, the other two color samples must be estimated from the neighboring samples in order to obtain a three-channel color image. The estimation of the missing color samples is referred to as CFA interpolation or demosaicking. In its simplest form, the missing pixels are filled in by spatially averaging the recorded values. Shown in Figure 5, for example, is the calculation of a red pixel from an average of its four recorded neighbors. Since the CFA is arranged in a periodic pattern, a periodic set of pixels will be precisely correlated to their neighbors according to the CFA interpolation algorithm. When an image is re-touched, it is likely that these correlations will be destroyed. As such, the presence or lack of these correlations can be used to authenticate an image, or expose it as a forgery.

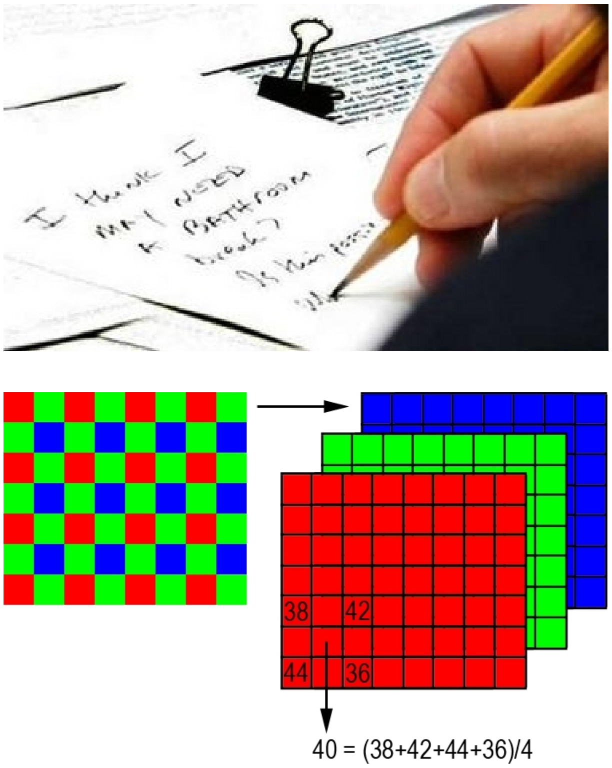

Figure 5: The note written by Bush was re-touched to improve readability, disrupting the color filter array correlations.

Science

Those in the media are not alone in succumbing to the temptation to manipulate photographs. In 2004, Professor Hwang Woo-Suk and colleagues published what appeared to be ground-breaking advances in stem cell research. This paper appeared in one of the most prestigious scientific journals, *Science*. Evidence slowly emerged that these results were manipulated and/or fabricated. After months of controversy, Hwang retracted the *Science* paper [6] and resigned his position at the University. An independent panel investigating the accusations of fraud found, in part, that at least nine of the eleven customized stem cell colonies that Hwang had claimed to have made were fakes. Much of the evidence for those nine colonies, the panel said, involved doctored photographs of two other, authentic, colonies.

While this case garnered international coverage and outrage, it is by no means unique. In an increasingly competitive field, scientists are succumbing to the temptation to exaggerate or fabricate their results. Mike Rossner, the managing editor of the *Journal of Cell Biology* estimates that as many as 20% of accepted manuscripts to his journal contain at least one figure that has to be remade because of inappropriate image manipulation [7,8].

We can better protect against this type of fraud by establishing a clear and strict editorial policy that governs the submission of scientific findings, and by incorporating a more rigorous screening process prior to publication.

Law

The child pornography charges against Police Chief David Harrison shocked the small town of Wapakoneta, Ohio. At his trial, Harrison's lawyer argued that if the State could not prove that the seized images were real, then Harrison was within his rights in possessing the images. In 1996 the Child Pornography Prevention Act (CPPA) extended the existing federal criminal laws against child pornography to include certain types of "virtual porn". In 2002 the United States Supreme Court found that portions of the CPPA, being overly broad and restrictive, violated First Amendment rights. The Court ruled that "virtual" or "computer generated" images depicting a fictitious minor are constitutionally protected. [1] The burden of proof in the Harrison case, and countless others, shifted to the State who had to prove that the images were real and not computer generated.

By some counts, the installation of video surveillance cameras is growing at a yearly rate of fifteen to twenty per cent. The vast majority of these cameras are being used by law enforcement agencies. While their installation certainly raises complex privacy issues, they are also raising complex legal issues. In August of 2005, a magistrate in Sydney, Australia threw out a speeding case after the police said it had no evidence that an image from an automatic speed camera had not been doctored.

The courts must modernize their evidentiary rules to contend with what is unarguably a digital age. These rules can better ensure the integrity of evidence by placing strict guidelines on the handling, submission and screening of digital media.

Discussion

Today's technology allows digital media to be altered and manipulated in ways that were simply impossible twenty years ago. Tomorrow's technology will almost certainly allow for us to manipulate digital media in ways that today seem unimaginable. And as this technology continues to evolve it will become increasingly more important for the science of digital forensics to try to keep pace. Along with awareness and sensible policy and law, it is my hope that the tools that my lab is creating will help the media, the courts, and our society contend with this exciting and at times puzzling digital age.

References

1. http://www.cs.dartmouth.edu/farid/research/digitaltampering
2. http://www.cs.dartmouth.edu/farid/research/tampering.html
3. A.C. Popescu and H. Farid. Exposing Digital Forgeries by Detecting Duplicated Image Regions. Technical Report, TR2004-515, Dartmouth College, Computer Science, 2004.
4. M.K. Johnson and H. Farid. Exposing Digital Forgeries by Detecting Inconsistencies in Lighting. em ACM Multimedia and Security Workshop, New York, NY, 2005.
5. A.C. Popescu and H. Farid. Exposing Digital Forgeries in Color Filter Array Interpolated Images. *IEEE Transactions on Signal Processing*, 53(10):3948-3959, 2005.
6. D. Kennedy. Editorial retraction. *Science*, 211(5759):335, 2006.
7. H. Pearson. Image Manipulation: CSI: Cell Biology. *Nature*, 434:952-953, 2005.
8. M. Rossner and K. Yamada. What's in a picture? the temptation of image manipulation. *The Journal of Cell Biology*, 166(1):1115, 2004.
9. J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes. Computer Graphics: Principles and Practice. Addison-Wesley Publishing Company, Inc., 2nd edition, 1993.
10. P. Nillius and J.-O. Eklundh. Automatic estimation of the projected light source direction. *IEEE Conference on Computer Vision and Pattern Recognition*, 2001.

Dr. Hany Farid is an Associate Professor of Computer Science at Dartmouth College in Hanover NH, USA. He specializes in digital audio, image and video analysis and forensics. He can be contacted at farid@cs.dartmouth.edu.

[1] In the United Kingdom, under the Protection of Children Act 1978, as amended by the Criminal Justice and Public Order Act 1994, a "pseudophotograph" of a child is defined as an image, whether made by computer graphics or otherwise, which appears to be that of a child. Possession or creation of such an image is illegal.

Lighting (details): In order to estimate the light source direction, we begin by making some simplifying assumptions: (1) the surface of interest is Lambertian (the surface reflects light isotropically); (2) the surface has a constant reflectance value; and (3) the surface is illuminated by a point light source infinitely far away. Under these assumptions, the image intensity can be expressed as:

$$I(x,y) \quad = \quad R(\vec{N}(x,y)\cdot\vec{L}) + A, \tag{1}$$

where R is the constant reflectance value, $\vec{L}$ is a 3-vector pointing in the direction of the light source, $\vec{N}(x,y)$ is a 3-vector representing the surface normal at the point (x,y), and A is a constant ambient light term [9]. If we are only interested in the direction of the light source, then the reflectance term, R, can be considered to have unit-value, understanding that the estimation of $\vec{L}$ will only be within an unknown scale factor. The resulting linear equation provides a single constraint in four unknowns, the three components of $\vec{L}$ and the ambient term A.

With at least four points with the same reflectance, R, and distinct surface normals, $\vec{N}$, the light source direction and ambient term can be solved for using standard least-squares estimation. To begin, a quadratic error function, embodying the imaging model of Equation (1), is given by:

$$\begin{aligned} E(\vec{L},A) &= \left|\left| M\begin{pmatrix} L_x \\ L_y \\ L_z \\ A \end{pmatrix} - \begin{pmatrix} I(x_1,y_1) \\ I(x_2,y_2) \\ \vdots \\ I(x_p,y_p) \end{pmatrix} \right|\right|^2 \\ &= \left|\left| M\vec{v} - \vec{b} \right|\right|^2, \end{aligned} \tag{2}$$

where $||\cdot||$ denotes vector norm, L_x, L_y, and L_z denote the components of the light source direction $\vec{L}$, and

$$M = \begin{pmatrix} N_x(x_1,y_1) & N_y(x_1,y_1) & N_z(x_1,y_1) & 1 \\ N_x(x_2,y_2) & N_y(x_2,y_2) & N_z(x_2,y_2) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ N_x(x_p,y_p) & N_y(x_p,y_p) & N_z(x_p,y_p) & 1 \end{pmatrix}, \tag{3}$$

where $N_x(x_i,y_i)$, $N_y(x_i,y_i)$, and $N_z(x_i,y_i)$ denote the components of the surface normal $\vec{N}$ at image coordinate (x_i,y_i). The quadratic error function above is minimized by differentiating with respect to the unknown, $\vec{v}$, setting the result equal to zero, and solving for $\vec{v}$ to yield the least-squares estimate:

$$\vec{v} \quad = \quad (M^TM)^{-1}M^T\vec{b}. \tag{4}$$

Note that this solution requires knowledge of 3-D surface normals from at least four distinct points ($p \geq 4$) on a surface with the same reflectance. With only a single image and no objects of known geometry in the scene, it is unlikely that this will be possible.

In [10], the authors suggest a clever solution for estimating two components of the light source direction (L_x and L_y) from only a single image. While their approach clearly provides less information regarding the light source direction, it does make the problem tractable from a single image. The authors note that at the occluding boundary of a surface, the z-component of the surface normal is zero, $N_z = 0$, Figure 6. In addition, the x- and y-components of the surface normal, N_x and N_y, can be estimated directly from the image.

With this assumption, the error function of Equation (2) takes the form:

$$\begin{aligned} E(\vec{L},A) &= \left|\left| M\begin{pmatrix} L_x \\ L_y \\ A \end{pmatrix} - \begin{pmatrix} I(x_1,y_1) \\ I(x_2,y_2) \\ \vdots \\ I(x_p,y_p) \end{pmatrix} \right|\right|^2 \\ &= \left|\left| M\vec{v} - \vec{b} \right|\right|^2, \end{aligned} \tag{5}$$

where,

$$M \quad = \quad \begin{pmatrix} N_x(x_1,y_1) & N_y(x_1,y_1) & 1 \\ N_x(x_2,y_2) & N_y(x_2,y_2) & 1 \\ \vdots & \vdots & \vdots \\ N_x(x_p,y_p) & N_y(x_p,y_p) & 1 \end{pmatrix}. \tag{6}$$

This error function is minimized, as before, using standard least-squares to yield the same solution as in Equation (4), but with the matrix M taking the form given in Equation (6). In this case, the solution requires knowledge of 2-D surface normals from at least three distinct points ($p \geq 3$) on a surface with the same reflectance.

In our work [4], we extended this basic formulation in three ways. First, we estimate the two-dimensional light source direction from local patches along an object's boundary. This is done to relax the assumption that the reflectance along the entire surface is constant. Then, we introduce a regularization (smoothness) term to better condition the final estimate of light source direction. Finally, this formulation is extended to accommodate a local directional light source (e.g., a desk lamp). We are currently extending this work to estimate a low-parameter model that embodies a multitude of complex light sources.

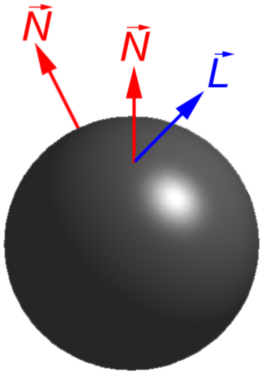

Figure 6: Schematic for estimating the light source direction, $\vec{L}$, from surface normals, $\vec{N}$.