

A Total Variation Approach for Customizing Imagery to Improve Visual Acuity

CARLOS MONTALTO*, IGNACIO GARCIA-DORADO*, and DANIEL ALIAGA

Purdue University

and

MANUEL M. OLIVEIRA

Instituto de Informática - UFRGS

and

FENG MENG

Purdue University

We describe a technique to generate imagery with improved sharpness for individuals having refractive vision problems. Our method can reduce their dependence on corrective eyewear. It also benefits individuals with normal vision by improving visual acuity at a distance and of small details. Our approach does not require custom hardware. Instead, the calculated images can be shown on a standard computer display, on printed paper, or superimposed on a physical scene using a projector. Our technique uses a constrained total-variation method to produce a deconvolution result which upon observation appears sharp at the edges. We introduce a novel relative total variation term that enables controlling ringing reduction, contrast gain and sharpness. The end result is the ability to generate sharper appearing images, even for individuals with refractive vision problems including myopia, hyperopia, presbyopia, and astigmatism. Our approach has been validated in simulation, in camera-screen experiments, and in a study with human observers.

Categories and Subject Descriptors: I.3.3 [Computer Graphics]: Picture/Image Generation—*Display algorithms*; I.4.4 [Image Processing and Computer Vision]: Restoration—*Inverse filtering*

General Terms: Displays, Deconvolution, Image correction

Additional Key Words and Phrases: High contrast, sharpness, total variation

ACM Reference Format:

Montalvo, C., Garcia-Dorado, I., Aliaga, D., Oliveira, M. M., and Meng, F. 20XX. Customized Imagery for Improving Visual Acuity. ACM Trans. Graph. 00, 0, Article 000 (XXX XXXX), 15 pages.

DOI = 10.0000/0000000.0000000

<http://doi.acm.org/10.0000/0000000.0000000>

* The first two authors contributed equally to this work.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© XXXX ACM 0730-0301/XXXX/15-ART000 \$10.00

DOI 10.0000/0000000.0000000

<http://doi.acm.org/10.0000/0000000.0000000>

1. INTRODUCTION

We present a personalized computational method that modifies images to improve sharpness for observers with unaided eyes (e.g., no glasses, contacts, nor head-mounted displays) (Figure 1). Our approach is useful for observers with refractive vision problems or for observers with normal/corrected vision (e.g., improving sharpness and contrast for distant reading of signs or small nearby details). Moreover, our computed image can be viewed on a computer screen, with a projector, or on printed media so that upon direct observation it appears sharper than viewing the original unprocessed image.

Recently, several hardware-based display frameworks have been proposed for providing improved sharpness and focus for observers. On the one hand, adaptive optics has emerged as a field that attempts to measure, in real-time, high-order continually changing optical aberrations using custom refractive and reflective hardware (e.g., [Bass et al. 2009; Hampson 2008]). On the other hand, light-field and multi-layer displays have recently been proposed in computer graphics to improve static long-term optical aberrations, such as defocus and astigmatism in human vision systems. For example, Pamplona et al. [2012] show a lightfield display that dynamically adapts its content to the observer’s specific condition, causing it to appear focused despite of his/her refractive vision problems. Their prototype is only able to show a very small region of a person’s field-of-view (FOV) – a more typical field-of-view is shown only in simulation by assembling many different photographs a posteriori. Huang et al. [2012] propose custom multi-layer display hardware and a deconvolution-based method that yields imagery appearing sharper for viewers with refractive vision errors. Nevertheless, their solution yields results with significant contrast loss. Further, it currently does not produce color imagery visible to the naked eye – their color images can only be obtained by long-exposure camera photography. Because of the need of precisely calibrated custom hardware and strong restrictions of observer movement, none of these methods can be used in practical applications to improve focus and contrast upon direct observation. In comparison to these and other previous works, our approach

- (i) provides naked-eye color viewing that yields improved sharpness despite refractive vision problems;
- (ii) offers high contrast;
- (iii) does not require glasses, contacts, or custom hardware (i.e., works with the unaided viewing of current displays and printed content); and

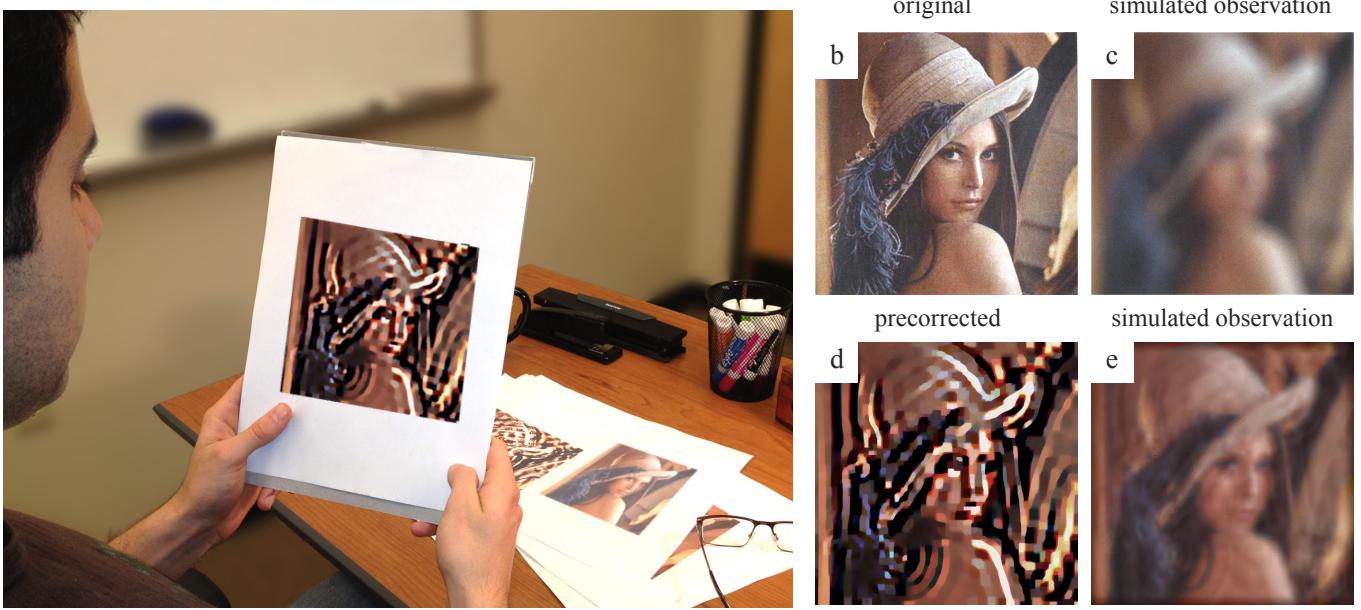


Fig. 1. PreCorrection on Paper. a) We depict the scenario where an observer sees our precorrected image, without his corrective eyewear, and perceives a sharper picture than viewing the original image; b) original image (*kindly provided by the USC SIPI Image Database*); In c-e, we show images from a scenario where the human observer is at a distance of 2.5m from the image and suffers from -2.5D of refractive error; c) simulated result of the observer looking at ‘b’; d) precorrected image; e) simulated result of the observer looking at ‘d’.

(iv) does not severely restrict observer movement.

Our methodology is based on a constrained *total variation* (TV) deconvolution method which carefully controls the process of creating a *precorrected image* that when observed appears sharper than directly viewing the *original image*. The refractive ability of the observer, for example as estimated by NETRA [Pamplona et al. 2010], can be modeled by a point-spread-function (PSF). Our method supports, but is not restricted to, PSFs created using Zernike polynomials [Thibos et al. 2000] that are widely-used in ophthalmology to model optical aberrations. Since one can only have positive and bounded pixel values on a display and the PSF causes some spatial frequencies to be severely attenuated, deconvolution-based methods are known to produce ringing artifacts (i.e., unwanted intensity oscillations) in the computed images. Ringing is typically reduced at the cost of a severe reduction in contrast. Our constrained TV-based method ensures pixel values are optimized in the $[0, 1]$ range and introduces a new term, called *relative total variation*, which enables controlling the tradeoff of ringing-reduction vs. contrast gain. In practice, our method can produce images with relatively high contrast and little ringing.

We demonstrate using our approach to yield improved sharp imagery in: (i) synthetic simulations, (ii) physical setups of premeasured lenses inserted between a camera and either a standard computer display or a printed sheet of paper, (iii) captured images of physical scenes where objects are altered to have their precorrected appearances by the use of projectors, and (iv) under human observation. Our results show that our system is able to produce considerably higher contrast than Huang et al.’s [2012] (e.g., seven times more contrast using the same image content), Alonso and Barreto’s [2003], Mohammadpour et al.’s [2012], and Lakshminarayanan et al.’s [2011]. It also yields sharpness for high blur levels (i.e., we show results up to six diopters of myopia or hyperopia).

The images produced by our method are the first full FOV high-contrast color images to be tested by human observers (and preferred by most of observers as per our 22-person user study – see Section 7.7). The impact of providing a computational-only method to yield improved vision clarity is significant. Addressing refractive vision problems is a component of the World Health Organization’s (WHO) VISION 2020 program [VisionURL]. It is estimated that between 800 million to 2.3 billion people worldwide suffer from refractive visual deficiencies, including 5 to 15 percent of children most of them being untreated [Dunaway and Berger 2006]. While our method does not avoid the need for glasses altogether, it does reduce the dependency on them. Further, our approach provides a passive way to obtain improved visual acuity and better visibility at a distance even for people with normal/corrected vision.

In summary, our main **contributions** include:

- A constrained total-variation-based formulation for generating personalized precorrected images, which upon observation by an individual exhibit sharper edges compared to observing the original unprocessed images and higher contrast in comparison to previous methodologies; improvements are obtained both for individuals with refractive errors as well as for those with normal/corrected vision;
- A novel relative TV term for controlling the tradeoff of ringing reduction versus contrast gain in the precorrected images; and
- A closed-loop automatic algorithm for effective PSF estimation and precorrected image calculation.

2. PREVIOUS WORK

2.1 Perceptual and Rendering Approaches

A large body of perceptual and rendering experiments has been described to improve the appearance of subtle details and/or to improve perceived sharpness. Rendering techniques, such as shading exaggeration [Rusinkiewicz et al. 2006], attempt to produce compelling and “clearer” content. Within perception [Cornsweet 1970], phenomena such as edge enhancement, opponent-color theory, psychophysical reading aspects (e.g., [Legge et al. 1985]), and numerous so-called visual illusions are exploited (e.g., [Wandell 1995; Pizlo et al. 2007; Pizlo 2010]). Some approaches combine perceptually-based and ophthalmologically-based techniques for improving image viewing (e.g., [Peli and Woods 2009]) and for improving resolution in videos and animations (e.g., [Didyk et al. 2010; Berthouzoz and Fattal 2012a; Berthouzoz and Fattal 2012b; Templin et al. 2013; Masia et al. 2013; Stengel et al. 2013]). For instance, the Cornsweet illusion has been exploited to provide enhanced local contrast [Ritschel et al. 2008] and to provide a perceptually-based metric for achieving compelling and efficient stereo disparity [Didyk et al. 2011]. Metrics have also been designed to quantify human sensitivity to blur (e.g., [Karunasekera and Kingsbury 1994; Marziliano et al. 2002; Marziliano et al. 2004]). However, none of these methods can compensate for the significant amount of blur perceived by an observer with several positive or negative diopters of refractive error, or by an observer with normal vision seeing distant or small content.

2.2 Custom Hardware

We refer the reader to a recent survey by Masia et al. [2013] on custom display technologies. The most relevant hardware solutions are those of Pamplona et al. [2012] and Huang et al. [2012]. However, the prototype shown by Pamplona et al. [2012] is limited to a very small FOV (e.g., would require a gigapixel display for standard desktop viewing) and requires careful positioning (or tracking) of the viewer in front of the display. Huang et al. [2012] use a multi-layer display that helps to recover some frequencies lost by the eye’s refractive problems. Their solution improves upon a single-layer display solution but still produces very low-contrast images. Their current prototype cannot show color images and does not operate fast enough to exceed the human flicker fusion threshold required by their method; thus, it is currently not usable for humans or user studies. Zhang and Nayar [2006] address a related problem where digital projectors are improved so that the projected image appears more in focus to the naked eye.

2.3 Constrained Deconvolution

Our approach is based on a constrained deconvolution process, which unlike standard image deconvolution, is a less studied topic (e.g., [Krishnan et al. 2009; Beck and Teboulle 2009]). In particular, defocus phenomena caused by refractive errors or by observing far away (or very small) objects can be modeled using convolution. The inverse operation of convolution is known as deconvolution. Many variations of image deconvolution have been proposed in computer graphics and computer vision (e.g., [Joshi et al. 2009; Krishnan and Fergus 2009; Yuan et al. 2008; Fortunato and Oliveira 2014]) using one of several methodologies (e.g., Wiener filtering, Richardson-Lucy).

Thus, at first sight the problem of generating a precorrected image may appear to be a standard application of image deconvolution to the original image. In practice, however, they are significantly

different. The optimal solution of standard deblurring has pixel values similar to those of the blurred image. However, the optimal solution for a precorrected image, that upon blurring becomes sharp, is usually very dissimilar to the original image. In fact, it often contains pixels with negative and very large intensity values which are not suitable for a computer display or printed media. Succinctly, the deconvolution challenges are the following:

- **Stability:** Inverting a convolution operator is well known to be highly sensitive to small amounts of noise in the image or in the PSF. The modulation transfer function (MTF) may contain zeros or very small values, leading to unstable deconvolution and noisy results. Because of this instability, deconvolution is an ill-conditioned operation and may introduce strong visual artifacts in the resulting images.
- **Non-Negativity and Bounded Values:** Since we ultimately seek to display an image, its pixel values must be constrained to be non-negative and are limited by the maximum intensity of the display or by the maximum amount of light that can be reflected off a printed precorrected image. Enforcing this constraint tends to severely reduce contrast.
- **Post-Deconvolution Sensitivity:** A precorrected image will be seen by a human eye, which is very sensitive to ringing artifacts. Thus, we must control the level of ringing in the precorrected image. This control, however, cannot be done by only inspecting the sharpness of the synthetically convolved precorrected images (Figure 2).

The simplest approach to constrained deconvolution is to first obtain a solution by regular deconvolution and then to restrict the solution to a fixed range in a second step. For example, Alonso and Barreto [2003], Brown et al. [2006], Marchand et al. [2011], Oyamada et al. [2007], and Huang et al. [2012] use or extend Wiener filtering to compute either a deblurred image or a precorrected image. All of these bring the solution into a valid pixel range (e.g., $[0, 1]$) by either clamping values, re-scaling the full range, or reducing enough the contrast of the original image. These methods suffer from ringing artifacts and/or very low contrast images. Although Huang et al. [2012] does use multi-layer displays to increase the obtained contrast, there is still a significant loss as compared to the original.

Total Variation. In our work, we explicitly address the constrained deconvolution problem using a total variation method. The TV approach was first introduced by Rudin et al. [1992] as a regularization technique capable of handling edges for the denoising problem. Since then there have been many applications of TV regularization to different problems including deblurring. We refer the reader to Chan and Shen [2005], Vogel [2002] and to the references within for a comprehensive study. While most TV algorithms have been proposed to deal with unconstrained minimization, we focus our treatment to the objective of precorrected image generation which necessitates non-negativity and bounded value constraints. We adapt and extend the dual-approach method of Beck and Teboulle [2009] to find an optimal solution. By considering the problem as a constrained deconvolution from the beginning, we find an optimal precorrected solution for constrained pixel values. Our result exhibits higher contrast than previous works and does not require multi-layer displays to improve contrast or to reduce ringing artifacts.

2.4 Deconvolution Comparison

To illustrate differences from previous approaches, Figure 3 shows several representative solutions under the same amount of -2.5D

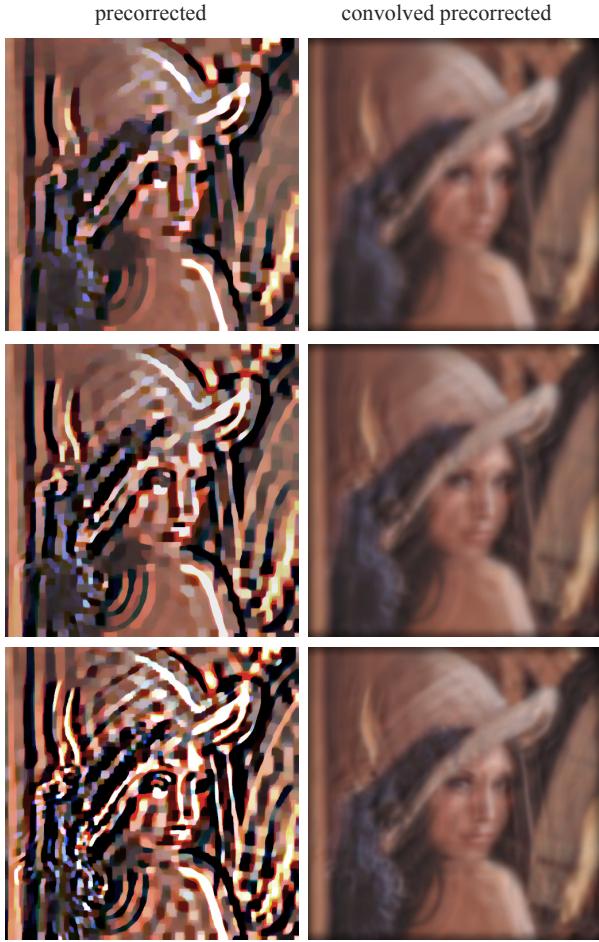


Fig. 2. Post-Deconvolution Sensitivity. We illustrate how different levels of ringing in the precorrected images (left column) for a fixed contrast range viewed under -2.5D result in similar synthetically convolved precorrected images (right column). Thus, controlling ringing is an important deconvolution issue to address. (*These simulations were created using our technique applied to the Lena picture, kindly provided by the USC SIPI Image Database*).

of defocus blur at 3 meters. Previous methods use either inverse Fourier transform (e.g., [Alonso Jr. and Barreto 2003]) or a Wiener filtering based extension (e.g., [Alonso Jr. et al. 2006; Mohammadpour et al. 2012; Lakshminarayanan 2012; Huang et al. 2012]) with pixel values clamped or rescaled to the valid range (e.g., $[0, 1]$). Some methods perform additional denoising schemes (e.g., [Alonso Jr. et al. 2008]) or edge tapering (e.g., [Mohammadpour et al. 2012]) as a post-process to reduce ringing - the results are slightly improved. The first and second columns of Figure 3 show original and blurred images. The third column contains precorrected images. The fourth and fifth columns are the synthetically convolved precorrected image and a camera-captured picture of the precorrected image (under -2.5D of blur).

The first four rows are representative ideal results from the aforementioned methods using Wiener filtering (SNR = 0.0016). In the first row, the precorrected image computed by Wiener filtering has a dynamic range of $[-19, +19]$ (note: the image we see in this document is implicitly rescaled to $[0, 1]$). We also repeated this experi-

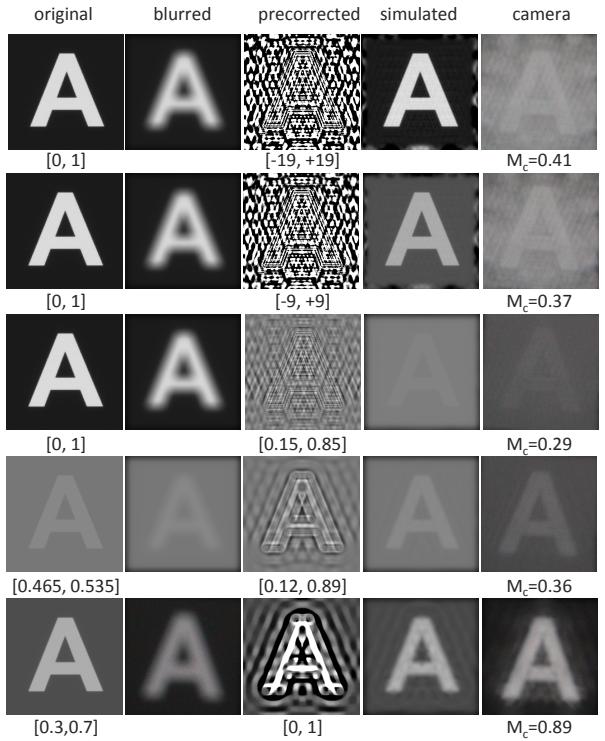


Fig. 3. Deconvolution Comparison. We show several results for deconvolution. Top four rows use Wiener filtering. Bottom row is our approach. 1st / 2nd row: full/half precorrected image range (which is not physically possible, so captured observation is unsatisfactory). 3rd row: precorrected image rescaled results in very low contrast. 4th row: input contrast range is reduced so that precorrected image fits the available range - still low contrast. 5th row: our approach yields clearly superior contrast.

ment using a standard TV-based (unconstrained) deconvolution and the result is very similar. In the second and third rows, we rescale the dynamic range of the precorrected image to $[-9, +9]$ and to $[0.15, 0.85]$, respectively. In the fourth row, we reduce contrast at input sufficiently so that the precorrected image has values within the $[0, 1]$ range. In all rows, the convolved precorrected image is computed synthetically and thus pixel values outside $[0, 1]$ can be used. The results are good though at a lower contrast as the pixel range of the precorrected image is reduced. However, the occurrence of negative values in the precorrected images demonstrate how such pixel values cannot be used by a physical system. It is worth noting that the third row is effectively the single layer solution implemented by Alonso et al. [2008].

The fifth (bottom) row of Figure 3 gives a prelude to our technique. In particular, our precorrected image is optimized within the $[0, 1]$ range for a reduced-contrast input image in the $[0.3, 0.7]$ range. The synthetically-convolved precorrected image and the captured precorrected image both show notably improved output contrast as opposed to the previous rows; in fact, a 2x improvement or more in this example as per the Michelson contrast ratio (i.e., the ratio between the difference and the sum of the highest and lowest luminance values) [Michelson 1995].

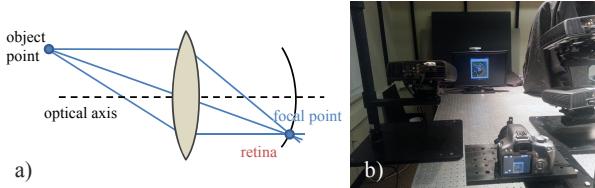


Fig. 4. **Physical and Optical Systems.** a) Depiction of human optical system and b) setup for our experiments with a camera-screen system.

3. VISION MODELING AND PRECORRECTION

3.1 Image Formation

The image formation process within the human eye is conceptually well understood (e.g., Trevor et al. [1984]). Light from the exterior world goes through the cornea, aqueous humour, lens, and interior of the eye, until reaching the retina. Figure 4a shows a conceptual simplification of this process using a thin-lens. If an object is too far (or too near) for a person to visually resolve its details, it appears blurred because the focal point is not on the retina. For example, a myopic person (i.e., nearsighted) focuses distant objects at a point in front of the retina. The opposite occurs with hyperopia (i.e., farsightedness). For astigmatism, the eye has different focal points for different meridians. Other aberrations, such Keratoconus and cataracts, occur due to a combination of additional factors. In the case of a person with normal vision, light coming from afar focus on a small point on the retina [Saladin 2007].

3.2 Convolution and Precorrection

Although human focusing is complex, spatially-variant, and distance-dependent [Villegas et al. 2008], the perceived blur can be well modeled by convolving a 2D PSF with the original image. Our experimental setup uses a digital camera with a chosen amount of defocus (via added lenses) observing an LCD screen (Figure 4b). Table I summarizes the main terminology defined in this section and in our approach. We denote by $L^1(\mathbb{R}^2)$ the set of functions defined in \mathbb{R}^2 such that the integral of their absolute value is finite. The convolution of the in-focus (single-channel) *original image* $t \in L^1(\mathbb{R}^2)$ with the *kernel* $k \in L^1(\mathbb{R}^2)$, that models the eye's PSF, simulates defocus

$$b = k * t. \quad (1)$$

The image b is the *blurred version* of the original image. Moreover, if we consider the function $K(t) = k * t$, then by Young's inequality $K(t) : L^1(\mathbb{R}^2) \rightarrow L^1(\mathbb{R}^2)$. For RGB color images, Equation (1) is evaluated once for each of the three color channels. We use the term *convolved image* when b is computed with a nu-

merical simulation and *captured image* for pictures acquired by our digital camera. The PSF is defined by the observer's aperture (i.e., pupil size) and refractive behavior.

An individual's PSF can be modeled by using Zernike polynomials (e.g., [Lakshminarayanan and Fleck 2011]). In our implementation, we use second-order Zernike polynomials (i.e., Z_2^{-2} , Z_2^0 , and Z_2^2) which are able to capture defocus and astigmatism aberrations. Thibos et al. [2000] lists the standard equations and notation for 2nd order Zernike's (Equation 1 and Table 2 in that reference) and example PSFs are shown in our video. Although we use Zernike polynomials, our method is not dependent on any particular PSF formulation or equation set – thus it can be used with other PSF representations.

The objective of precorrection is to generate a *precorrected image* p that when observed appears sharper than directly observing the original image t . The problem becomes to invert the functional $K(t) : L^1(\mathbb{R}^2) \rightarrow L^1(\mathbb{R}^2)$. To invert K , it is natural to consider the Fourier Transform of a function $f \in L^1(\mathbb{R}^2)$ given by

$$\hat{f}(w) = \frac{1}{2\pi} \int e^{-iw \cdot z} f(z) dz, \quad (2)$$

where $z = (x, y)$. The absolute value of the Fourier Transform is known as the *modulation transfer function* (MTF) of f . If we assume that the MTF of k is non zero in the support of t , the *optimal precorrected image* p^* is given by

$$p^*(z) = \frac{1}{2\pi} \int e^{iw \cdot z} \frac{\hat{t}(w)}{\hat{k}(w)} dw. \quad (3)$$

Our method seeks to compute a precorrected image p that is as close as possible to the aforementioned p^* and thus leads to a convolved precorrected image $k * p$ similar to $k * (p^*) = t$.

4. IMAGE-BASED PRECORRECTION

4.1 Formulation Overview

We use a total-variation-based deconvolution method to calculate a precorrected image that upon observation appears sharper than looking at the original image. The solution also has no more than a user-specified amount of ringing τ^* and the maximum contrast range possible $[c_{low}, c_{high}]$. A precorrected image calculated by our approach is denoted by $p = p(\theta, t_c)$ and defined as

$$p(\theta, t_c) = \arg \min_{0 \leq \|p\|_\infty \leq 1} (\|k * p - t_c\|_{L^2} + \theta \|\nabla p\|_{L^1}), \quad (4)$$

where $\theta > 0$ is a regularization weight and $t_c = t \cdot (c_{high} - c_{low}) + c_{low}$ is a rescaled version of the original image t to the contrast range $[c_{low}, c_{high}]$. The term $\|k * p - t_c\|_{L^2}$ in Equation (4), ensures the convolved precorrected image is visually similar to the original image in an L^2 -sense. The term $\|\nabla p\|_{L^1}$ is the TV-norm of p , i.e.,

$$\|\nabla p\|_{L^1} = \int \int |\nabla p| dx dy, \quad (5)$$

where $|\cdot|$ denotes the absolute value in \mathbb{R}^2 [Chan and Shen 2005].

To quantify ringing, we introduce a term called *relative total variation* $\tau = \tau(p, t_c)$ that measures the additional amount of ringing in the precorrected image p as compared to the rescaled contrast image t_c . We define τ as

$$\tau(p, t_c) = \frac{\|\nabla p\|_{L^1} - \|\nabla t_c\|_{L^1}}{\|\nabla t_c\|_{L^1}}, \text{ if } t_c \neq \text{constant}. \quad (6)$$

In Equation (6), note that when t_c is constant (i.e., an image of constant color), $\nabla t_c = 0$ and hence the division is not well-defined. In

Table I. Terminology.

Name and Notation	Description
Original image t	In focus image
Point spread function k	Models individual vision
Precorrected image p	Image such that $k * p \approx t$
Convolved precorrected $k * p$	Synthetically convolved image
Captured image	Captured by the camera
Observed image	Observed by an individual

We summarize the main terminology used in this article and in our system.

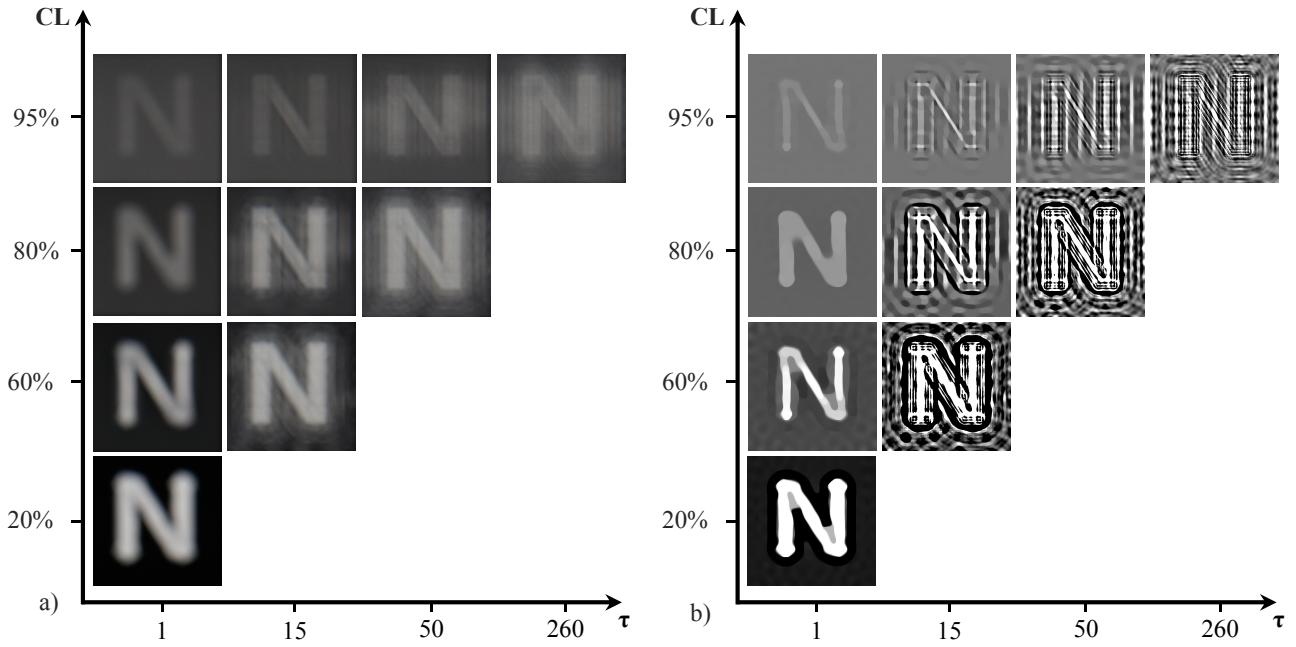


Fig. 5. **Relative TV vs. Contrast.** a) Array of captured precorrected images of varying relative TV and contrast ranges under -4.75D of refractive error at 1.5m. The changing contrast range is obtained by increasing the contrast loss CL along the vertical axis. b) The corresponding precorrected images for a) are shown. These two arrays illustrate the tradeoff between contrast and ringing.

practice, this situation does not occur. Further, note that τ depends on t_c and p which in turn depends on θ . Hence in the following, we succinctly write this as $\tau(\theta, t_c)$. Moreover, given t_c Equation (4) tell us that τ is decreasing w.r.t. $\theta > 0$ and thus

$$-1 = \lim_{\theta \rightarrow \infty} \tau(\theta, t_c) \leq \tau(\theta, t_c) \leq \tau(0^+, t_c) = \max_{\theta > 0} \tau(\theta, t_c). \quad (7)$$

Given a user-specified amount of ringing τ^* (Section 4.2), our method seeks the minimum reduction in contrast such that

$$|\tau^* - \tau(\theta, t_c)| \leq \epsilon, \quad (8)$$

for some $\theta > 0$, where ϵ is a small tolerance error. If at a given contrast the solution cannot be found, i.e., $\tau^* - \max_{\theta > 0} \tau(\theta, t_c) > \epsilon$, it implies that the desired amount of ringing does not occur at the specified contrast range. Thus, we reduce the contrast range (by either increasing c_{low} , decreasing c_{high} , or both) until inequality (8) is met. Since the contrast range can ultimately be reduced until $c_{low} = c_{high}$ (e.g., a constant image), this process converges for reasonable values of τ^* . To quantify the contrast loss we define the percentage contrast loss by $CL = (c_{low} + (1 - c_{high})) \cdot 100$.

Intuitively, as we decrease the contrast range of the original image, larger amount of ringing (i.e., larger τ) can be achieved in the precorrected image. This occurs because the bounded pixel value constraint becomes less stringent, which in turns provides more freedom to form a sharper $k * p$. This agrees with the effect seen in Huang et al. [2012] and Mohammadpour et al. [2012], where lower-contrast original images are able to be seen with improved sharpness. Further, small θ values essentially ignore the regularization and large values cause the image to be “washed out”. As per Equation (4), the pixel values of p are constrained to the $[0, 1]$ range so as to ensure that a precorrected image can be correctly displayed or printed. Algorithm 1 presents pseudo code to explain our technique.

4.2 Relative Total Variation

In the previous section we introduced the relative TV value τ that measures the relative gain in total variation of the precorrected image as compared to the original image. We show the effectiveness of τ as a parameter to measure the ringing in the precorrected images. Moreover, we illustrate the existence of an optimal value τ^* that balances contrast reduction in the original image t , ringing artifacts in the precorrected image p , and sharpness in the convolved precorrected $k * p$.

Figure 5 shows two arrays of images depicting various precorrected and convolved precorrected images at different contrast levels. All images have as the original image a sharp “N” placed in the middle of the image. The right image array shows several versions of the precorrected image p . The left image array shows the corresponding convolved precorrected images, all using the same Zernike-based PSF. Within each array, the contrast range decreases along the vertical axis and ringing increases along the horizontal axis. The labels on the horizontal axis show the approximate values of τ . As expected, the amount of perceived ringing is about the same along each column. For any contrast range, the details of the convolved precorrect image increases with relative TV as well as ringing and additional artifacts in the image (Figure 5). We observed that there is a value τ^* for the relative TV that corresponds to an optimal balance of ringing and sharpness. This optimal value depends on the inherent characteristics of the image. Once a desired τ^* has been selected, we find the maximal contrast that allows this relative TV value.

In our current system, the desired value τ^* is selected by the user. Empirically, we observe that for text images a larger relative TV is acceptable since the simple image content can afford it. For a busy color image, smaller values of τ are preferred, and any introduced

Algorithm 1 Pseudocode. A summary of our overall formulation.

```

Input: user desired level of ringing ( $\tau^*$ ); ringing error tolerance
to desire value ( $\epsilon$ ); contrast reduction step ( $\delta$ , typically 0.05); minimum
 $\theta$  value to explore ( $a$ , typically  $10^{-10}$ ); and maximum  $\theta$  value
to explore ( $b$ , typically 1).
Output: precorrected image  $p$ 
1: // Compute a precorrected image with the minimum CL
2:  $c_{low} := 0$ ; ▷ find minimum CL that reaches  $\tau^*$ 
3:  $c_{high} := 1$ ;
4:  $t_c := t$ ;
5: while (!DOESCONTRASTLEVELREACHTAU( $t_c, \tau^*, a$ ) do
6:    $c_{low} := c_{low} + \delta$ ;
7:    $c_{high} := c_{high} - \delta$ ;
8:    $t_c := t \cdot (c_{high} - c_{low}) + c_{low}$ ;
9: return COMPUTEPRECORRECTED( $\tau^*, t_c, a, b$ ); ▷ compute  $p$ 
10: // Auxiliary functions
11: function DOESCONTRASTLEVELREACHTAU( $t_c, \tau^*, a$ )
12:    $\theta := a$ ;
13:    $p(\theta, t_c) := \arg \min_{0 \leq \|p\|_\infty \leq 1} (\|k * p - t_c\|_{L^2} + \theta \|\nabla p\|_{L^1})$ ;
14:    $\tau(p, t_c) := \frac{\|\nabla p\|_{L^1} - \|\nabla t_c\|_{L^1}}{\|\nabla t_c\|_{L^1}}$ ;
15:   return  $\tau(p, t_c) < \tau^*$ ; ▷ return true if can be reached
16: function COMPUTEPRECORRECTED( $t_c, \tau^*, a, b$ )
17:    $\theta := \frac{a+b}{2}$ ;
18:    $p(\theta, t_c) := \arg \min_{0 \leq \|p\|_\infty \leq 1} (\|k * p - t_c\|_{L^2} + \theta \|\nabla p\|_{L^1})$ ;
19:    $\tau(p, t_c) := \frac{\|\nabla p\|_{L^1} - \|\nabla t_c\|_{L^1}}{\|\nabla t_c\|_{L^1}}$ ;
20:   if  $|\tau - \tau^*| \leq \epsilon$  then
21:     return  $p$ ;
22:   else if  $\tau < \tau^*$  then
23:     return COMPUTEPRECORRECTED( $t_c, \tau^*, \theta, b$ );
24:   else
25:     return COMPUTEPRECORRECTED( $t_c, \tau^*, a, \theta$ );

```

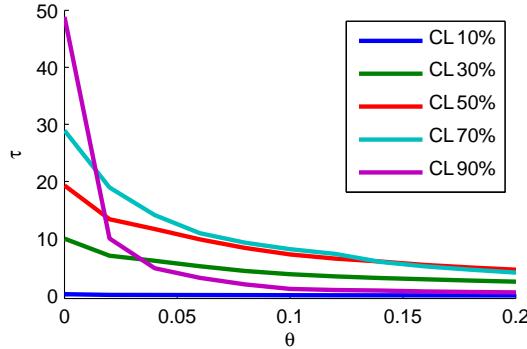


Fig. 6. Contrast Loss Graph. The vertical axis corresponds to increasing relative TV. The horizontal axis indicates the TV weight of Equation (4). Each colored curve represents a different contrast loss (CL). As we reduce contrast, the range of possible relative TV values τ is increased.

ringing is very distracting. In Section 7.7, our user study provides feedback about the visual tolerance of ringing for human observers.

4.3 Regularization

The term $\theta \|\nabla p\|_{L^1}$ of Equation (4) is the regularization term that represents the TV of the precorrected image p . As expected, this regularization term allows us to control ringing. A less obvious ob-

servation is that it also serves to enforce sharp edges in the convolved precorrected image $k * p$.

4.3.1 Ringing Control of Precorrected Image.

We illustrate with an example how the regularization term controls ringing in the precorrected image. Figure 8a depicts a portion of a 1D step signal with a sharp edge at zero with intensity difference of h . For $a < 0 < b$, we approximate the signal by the function p , that is, 0 for $x \leq a$, $h(x-a)/(b-a)$ for $a < x < b$; and h for $x > b$. As illustrated in Figure 8a the TV-norm of p is

$$\|\nabla p\|_{L^1} = \int_a^b \frac{h}{b-a} dx = h. \quad (9)$$

Note that the result is independent of a and b ; it only depends on the magnitude of the jump in intensity. To visualize what happens when there is more ringing, Figure 8b shows an original (top) and precorrected (bottom) image pair. For each image, a 1D slice corresponding to a line at the center of the image is also shown. It can be observed (Figure 8b right) that the amount of ringing in the precorrected image is roughly “three times” that in the original image. Using the graph (Figure 8b left), the total variation can be computed to be $\|\nabla p\|_{L^1} \approx 3 \|\nabla t\|_{L^1} = 6h$ (three times $2h$). Hence, this regularization term is a good indicator of the amount of ringing and can be used to control it during deconvolution.

4.3.2 Edge Sharpening of Convolved Precorrected.

One clearly desirable property is that the convolved precorrected image $k * p$ (i.e., the observation by the viewer) presents sharp edges. Upon close inspection, we observe that the ringing in $k * p$ is due to the instability of deconvolution, the bound constraints, and Gibbs phenomenon (e.g., the difficulty inherent in approximating a signal with sharp edges by a finite series of continuous waves). In particular, Gibbs phenomenon occurs as oscillations that decay away from an edge. Hence, a consequence of Gibbs phenomenon is that ringing is stronger near edges of p . Thus, the areas where ringing occurs are also precisely where edge sharpness is most desired. From this reasoning, a tentative regularization parameter for the optimization Equation (4) would be to use

$$\|\nabla(k * p)\|_{L^1}, \quad (10)$$

instead of $\|\nabla p\|_{L^1}$, to reduce ringing in $k * p$ and thus encourage edge sharpness. We will show that the effect of using the regularization term $\|\nabla(k * p)\|_{L^1}$ will in fact also be accomplished by using $\|\nabla p\|_{L^1}$.

4.3.3 Ringing Controls Sharpening.

We show that the TV-norm of the convolved precorrected $k * p$ is bounded by the product of the L^1 -norm of the PSF kernel k and the TV-norm of the precorrected image p , i.e.,

$$\|\nabla(k * p)\|_{L^1} \leq \sqrt{2} \|k\|_{L^1} \|\nabla p\|_{L^1}. \quad (11)$$

Since the L^1 -norm of the PSF k is typically normalized to 1, this inequality implies that if the ringing in p is reduced then sharpness in $k * p$ is also improved. It is worth noting that the reverse inequality of (11) is not true in general (e.g. for k with small L^1 -norm).

To prove inequality (11), we use two facts. First, from Young’s inequality for convolution it follows

$$\|\nabla(k * p)\|_{L^1} \leq \|k\|_{L^1} \|\nabla p\|_{L^1}. \quad (12)$$

Second, we use the following inequalities

$$\sqrt{|u|^2 + |v|^2} \leq |u| + |v| \leq \sqrt{2} \sqrt{|u|^2 + |v|^2}, \quad (13)$$

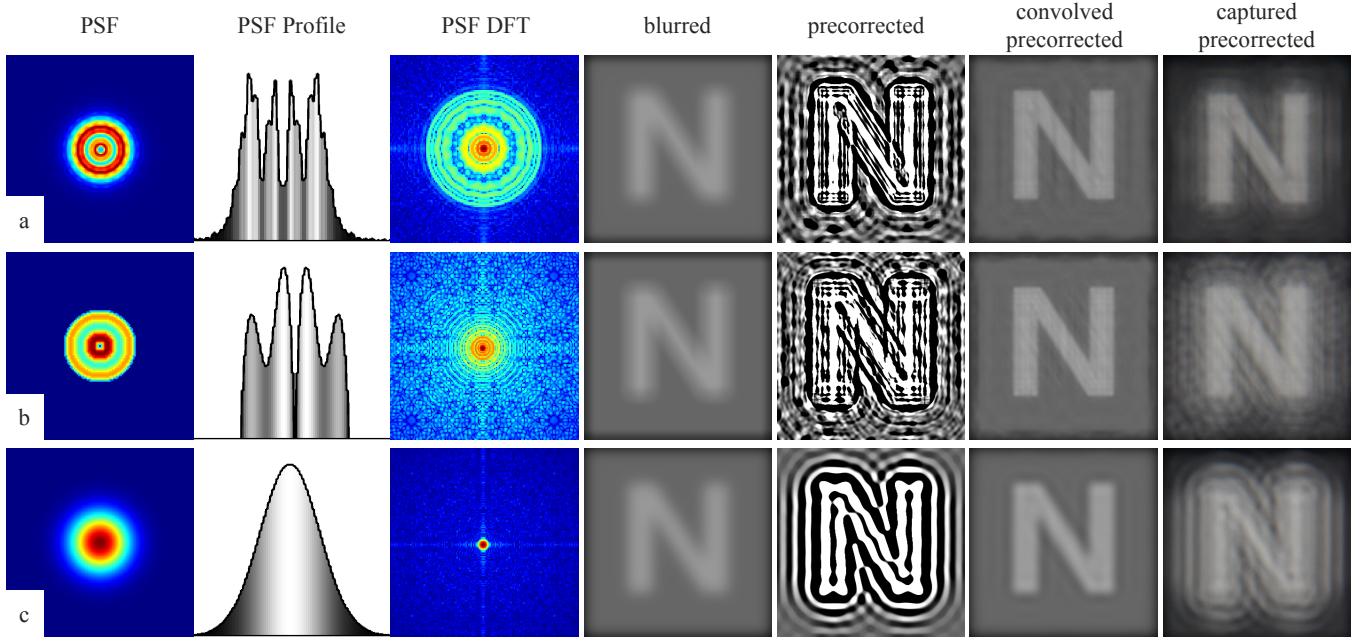


Fig. 7. **PSF Comparison.** a) A Zernike PSF, its vertical profile, its DFT, the blurred observation, a precorrected image, a convolved precorrected image, and a captured precorrected image. b) Similar sequence but with a spline curve found via optimization (see text). c) Similar sequence using a best-fitting Gaussian. Comparing the convolved precorrected and captured precorrected, Zernike PSF produces the most similar result.

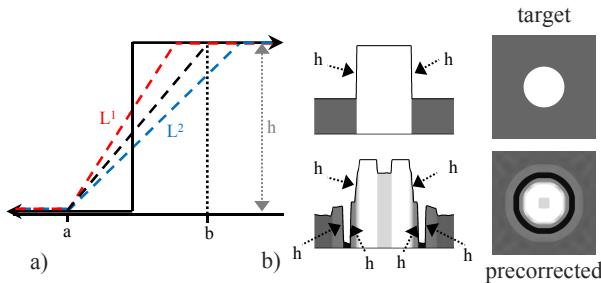


Fig. 8. **Regularization Terms.** a) Our regularization term encourages the presence of sharp edges; and b) quantifies the amount of ringing.

for u, v real numbers. These inequalities can be verified by squaring them. Using Equations (12-13), we get the following inequalities

$$\|\nabla(k*p)\|_{L^1} \leq \left\| k * \frac{\partial p}{\partial x} \right\|_{L^1} + \left\| k * \frac{\partial p}{\partial y} \right\|_{L^1} \leq \sqrt{2} \|k\|_{L^1} \|\nabla p\|_{L^1}.$$

These last inequalities prove (11), a detail proof can be found in Montalto [2014].

We saw in Figure 2 that two precorrected images with very dissimilar ringing could give a similar convolved precorrected. An advantage of using the regularization term $\|\nabla p\|_{L^1}$ instead of $\|\nabla(k*p)\|_{L^1}$ is that we can better monitor the ringing of p . By measuring the ringing of the precorrected rather than the convolved precorrected $k*p$, we can better differentiate between solutions and allow each individual to select its desired level of ringing. In addition, using $\|\nabla p\|_{L^1}$ as a regularization term compensates for model inaccuracies of the visual aberrations by the PSF.

4.4 Point Spread Functions

Analytically, Equation (2) guarantees existence and uniqueness of the precorrected image p , under the assumption that the MTF of k has no zeros in the support of t . However, when the MTF of k is zero (or close to zero) in the support of t , the problem of recovering p becomes very unstable. Hence, the shape of the PSF has an obvious impact on the precorrection ability.

An important observation is that Zernike-based PSF and TV-based regularization can produce sharp edges despite having values close to zero in the corresponding MTF. A Gaussian-based PSF has no zeros in the lower frequencies but attenuates the higher frequency components. In contrast, a Zernike-based PSF exhibits a wave-like form whose MTF has near zeros in the lower and higher frequencies, but the higher frequencies are not completely omitted. Hence, since TV-based regularization Equation (4) encourages sharper edges in the reconstruction of p , we are able to capitalize on the higher frequencies of the Zernike-based PSF despite the zeros.

We performed an informal experiment that optimizes the PSF shape (Figure 7). We captured an image t_c under -4.75D of refractive error at 1.5D, i.e., image b (synthetically, $b = k * t_c$). We then let the system optimize the shape of the PSF (k) so as to maximize the sharpness of the (synthetically) convolved precorrected image but maintain a similar amount of blur in the convolved image (b) as in the captured image. More precisely, we alter k and recompute p so as to minimize Equation (4), subject to the additional constraint $\|k * t_c - b\|_{L^2} \rightarrow 0$ for a fixed value of θ . The shape of k is altered by varying the control points of an interpolating spline that defines a circularly symmetric 2D PSF. For several images and different refractive errors, the optimization always converged to a PSF whose shape roughly mimicked the oscillating nature of Zernike-based PSFs and yielded a similar level of sharpness in the convolved precorrected image. For comparison, we also show a Gaussian-based

PSF - the width of the Gaussian is set to imitate as best as possible the observed amount of blur. It is worth noting that Raskar et al. [2006] and Veeraraghavan et al. [2007] also identified PSFs with similar non-Gaussian, irregular-like appearances for obtaining superior deblurring/deconvolution performance.

5. PRECORRECTION SYSTEM

In the following, we describe our automatic framework implementing the aforementioned precorrection computation for a given original image t and relative total variation τ^* . We designed and experimented with three deployment system types:

- *camera-screen system*: color precorrected images are displayed on a standard LCD screen or a printed sheet of paper; a pre-measured lens is placed in front of a digital camera in order to simulate a prescribed amount of positive or negative diopters of blur (Figure 4b);
- *projector-camera system*: compensation images are projected on top of a colored physical object and the object is captured by a digital camera with a premeasured lens in front of it; the compensation images are such as to alter the visual appearance of the object to that of the precorrected color image's appearance; and
- *human observation*: precorrected color images are displayed on a computer screen or on a sheet of paper for observation; this scenario was used for our user study.

For each system, our framework estimates a PSF, finds the largest contrast range for a desired relative TV, and displays/prints a pre-corrected image. *The PSFs for the three system types are not identical; a precorrected image generated for the camera-screen system is not suitable for human viewing and vice versa.*

5.1 PSF Parameter Estimation

The first step is to obtain the Zernike-based PSF k of the optical system. We obtain the Zernike parameter values (see Thibos et al. [2000]) either directly (e.g., from the refractive corrective prescription of a person) or indirectly via an optimization. The latter is the case for the camera-screen and projector-camera system whereby a photographic camera observes a standard LCD screen, a sheet of paper, or a physical scene. The camera, with an approximately known and fixed aperture, is focused on the content. Then, we induce a refractive error by placing a lens (or set of lenses) of known diopter immediately in front of the camera. We disable as much of the camera processing machinery as possible and capture an image. Afterwards, we compute an affine transformation to best align the captured original image b to image t and estimate the Zernike's pupil parameter value to best reproduce the blur in the captured image.

5.2 PreCorrection Calculation

In the next step, we seek to perform the optimization described by Equations (4-8). Given an original image t , a desired τ^* , and the estimated PSF (from Section 5.1), we solve the optimization using the following steps.

5.2.1 Contrast Level and θ Estimation.

First, we find the largest original-image contrast-range $[c_{low}, c_{high}]$ and corresponding θ that can produce the precorrected image with the desired relative TV τ^* . There are multiple ways to alter the contrast (e.g., increase/decrease one or both of c_{low} and c_{high}). By default for white text on black and for busy color images we increase

c_{low} and decrease c_{high} . However, the user can select an alternative contrast reduction strategy. Our method performs a 1D binary search to find the smallest θ value that generates a relative TV close to the desired τ^* value. The chosen value for τ^* may not be feasible for the given image (i.e., inequality of Equation (8) cannot be met). In such cases, our method reduces the contrast range by a small amount and iterates until a solution is found (see Algorithm 1).

5.2.2 PSF Tuning.

Second, once θ is calculated we compute the precorrected image p and corresponding synthetically-computed $k * p$ for a range of diopters and apertures (in the case of images captured by a camera) near the estimated PSF values. Using a *sharpness metric*, we choose from amongst these solutions the precorrected image that yields the sharpest $k * p$. By this mechanism, we can find a PSF (i.e., k) that exhibits slightly better performance during precorrection.

Our sharpness metric automatically measures the magnitude of high frequencies present near the edges of the original image. Using an edge filter (e.g., Canny), we select original image pixels near an edge. Then, we sum the magnitude of the frequencies in the Fourier transform of p that are above the median value. The computed sum is returned as the value of the sharpness metric for the entire image.

6. IMPLEMENTATION DETAILS

We implement our TV-based method by extending the algorithm Fast Iterative Shrinkage/Thresholding Algorithm (FISTA) [Beck and Teboulle 2009]. A variety of TV-based methods, including the primal-dual method by Chan et al. [1999] and variable splitting method of Wang et al. [2008], have been proposed. Beck and Teboulle employ a dual formulation of the minimization of Equation (4), as given by Chambolle [2004], which does not require an extra regularization term of the TV norm itself.

Our system is mostly implemented in MATLAB and some parts in C. It runs on a desktop PC equipped a 3.53 GHz processor. To produce RGB color images, we perform precorrection per channel. It takes on average 30-60 seconds to compute a single precorrected image. When displaying images, we must take into account the physical size of the pixels. For most of our results, we display 600x600 resolution images 1:1 on a 24 inch Flat LCD panel and take pictures with a Canon EOS Rebel T2i camera. For precorrection on paper, we print images to the same size as on the screen. For user studies, we displayed results on a 55 inch LCD. Overall, we use lenses from -6D to +6D and apertures (or pupil diameters) from 2 to 6 mm depending on the lighting conditions.

7. RESULTS AND DISCUSSION

We show results and insight into the behavior of our method. We use concrete examples to illustrate different scopes and limitations. Unless otherwise stated, all results are obtained using our experimental camera-screen system to capture actual imagery. Figure 9 demonstrates a visual summary of our pipeline. A person with myopia (i.e., nearsightedness) that observes the target image (a) perceives it similar to image (b). We generate a precorrected image (c) which upon observation looks like (d), effectively reversing the refractive aberrations and recovering image details. By additional contrast reduction, we can sharpen even more the image and restore further details as in (e). Figure 18 shows the same content but for human observation (as used by our user study). Our method is stable under lateral and some distance displacement of the viewer which makes the deployment (e.g., on paper) particularly attractive.

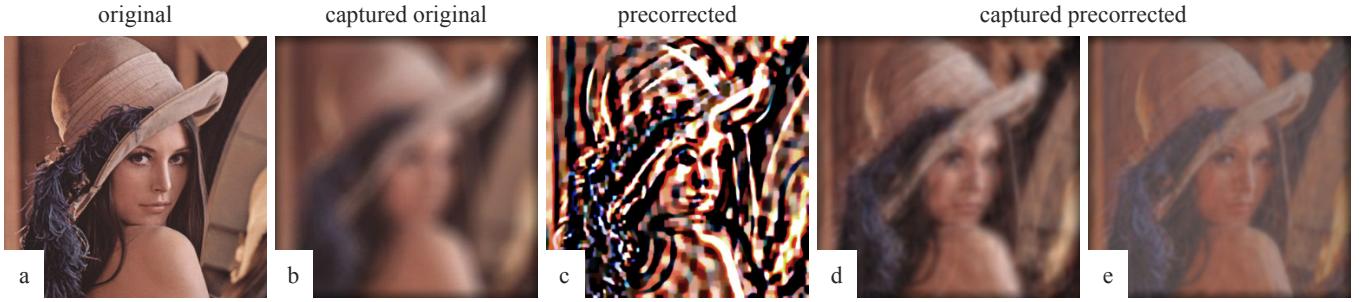


Fig. 9. **Pipeline.** Our system pipeline where (a) is the original image t (*kindly provided by the USC SIPI Image Database*), (b) the captured original image b when viewing t under -4.75D of refractive error at 1.5m , (c) the precorrected image p with $c_{low} = 0.1$ and $c_{high} = 0.9$ for our camera-screen system, (d) captured precorrected with $c_{low} = 0.1$ and $c_{high} = 0.9$ and (e) the captured precorrected with $c_{low} = 0.3$ and $c_{high} = 0.7$. Our captured precorrected images show a clear improvement over direct observation of the original image.

7.1 Different Refractive Errors

Figures 10 and 11 show captured precorrected images for several refractive aberrations. In Figure 10, we show images precorrected for -6 to $+6$ diopters; e.g., potential corrective prescriptions for myopia, hyperopia, or presbyopia. Results for this amount of refraction *have not been shown* by any of the cited previous works. Some of the observed graininess is because the images in the two leftmost columns are photographs of an LCD screen showing the original or precorrected images, respectively. Figure 11 shows an example of medium-range astigmatism and myopia. In all cases we see a clear improvement in sharpness and detail using our precorrected image. As the number of diopters increases, resolving detail is more challenging because of the wider PSF.

7.2 Robustness

Figure 12 shows the robustness of a fixed precorrected image being observed under varying refractive amounts. The figure uses a precorrected image designed for -4.5D of refractive error but is observed under -1.5D to $+1.5\text{D}$ away from its intended refractive error. The quality of the captured image degrades with higher refractive error discrepancies but does not seem overly sensitive to small variations.

In Figure 13, we show the sensitivity of our solution to observer displacement. The image in the left column of the third row is the observation for the observer/camera at intended ideal location. We show additional images corresponding to displacements of 0.25m and 0.5m to the front, back, right, and diagonal offsets as well. Observer locations maintaining the approximate same distance to the screen seem to degrade the least. Pure forward/backward motion does yield noticeable artifacts faster than other displacements. Moreover, forward displacement seems to degrade the quickest. We do not show images for displacements to the left since the visual effect is similar to that to the right.

7.3 Reading and Visual Acuity

An important application of our algorithm is that it could be used to automatically customize text for reading by people with refractive errors. We show two experiments. In Figure 14, we use a typical Snellen Chart to estimate visual acuity. In our experiment we positioned the Snellen chart at 1.5m and captured images with an induced blur of -5D . We observed a notable improvement in visual acuity from $20/70$ line and downwards, making the $20/20$ line al-

most clearly visible, in this case. The contrast loss is due to the initial reduction needed to achieve the desired total variation in the precorrected image. A similar experiment was performed by Huang et al. [2012]. Their results present images with less ringing artifacts but significantly lower contrast and require a precisely calibrated environment and customized hardware.

In Figure 15, we illustrate the appearance of a single letter at multiple sizes (or equivalently at different distances). We capture images of a letter at a distance of 1.5m with a blur of -5D . When the letter is small (e.g., 60 pixels), the captured original image is unrecognizable. In contrast, our captured precorrected image shows a letter that begins to be distinguishable. As the font size increases, the captured precorrected image becomes readable sooner than the observation of the standard text. When the letter size is large (e.g., 240 pixels), both images are readable. Nevertheless, our precorrected images show an enhanced sharpness. Special font could be generated automatically using our algorithm so as to provide a better reading experience.

7.4 Ringing and Contrast Tradeoff

Our approach is able to trade ringing for contrast as illustrated in Figure 5, where various captured images of the letter ‘N’ are shown. For each image, we alter the contrast range $[c_{low}, c_{high}]$ and relative TV. Along the horizontal axis, τ increases from 1 (i.e., little ringing) to 260 (i.e., high ringing). Each column has a similar amount of ringing, demonstrating that τ is a good variable for measuring it. As we increase τ , we increase sharpness as well as ringing. In this example, a balance of these tradeoffs occurs for a desired τ^* somewhere around $\tau = 15$. When contrast is high (i.e., images corresponding to bottom right triangle of the table are not possible), our method does not necessarily have room to increase τ to τ^* . Therefore, we decrease contrast until we find a precorrected image with relative TV equal to τ^* . Thus, we increase image sharpness at the expense of reduced contrast (Section 4.2).

7.5 Displacement, Distant and Small Content

Another application for our method is improving visual acuity for individuals with normal vision. We use the definition of standard visual acuity (i.e., the ability to recognize a letter of size 5 minutes of an arc at a distance of 1 meter [Duane 1985]) to model an individual with normal vision looking at distant objects. Figure 16 shows the comparison between small images captured by a camera that was focus to simulate an individual with normal vision. The

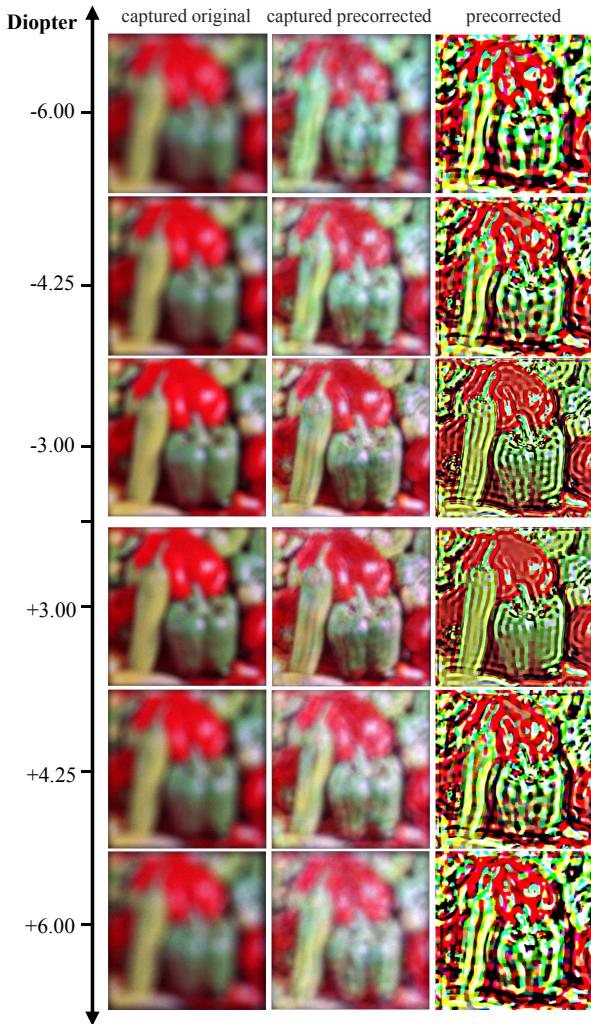


Fig. 10. Multiple Diopters. Captured original images, captured precorrected images, and precorrected images seen with -6D to +6D of refractive error at 1.5m. (These simulations were created using our technique applied to the Peppers picture, kindly provided by the USC Sipi Image Database).

ability to recognize details in the images is improved by using our method.

7.6 Contrast Comparison

Our approach works with a variety of image types and at high contrast. Figure 17 shows several original images, captured original images, and captured precorrected images all with -4.75D of refractive error at 1.5m. We also show their corresponding Michelson contrast ratios. We make the original images have maximum contrast ratio (i.e., minimum and maximum intensities are 0 and 1, respectively). Our imagery shows noticeably improved sharpness at almost the same contrast as the original images (e.g., the top row is the same image from Huang et al. [2012] but our solution obtains 7.15x more contrast as per the Michelson contrast ratios reported in their paper). While our resulting images may still suffer from some artifacts, we obtain significantly higher contrast ranges and support higher diopters than others.

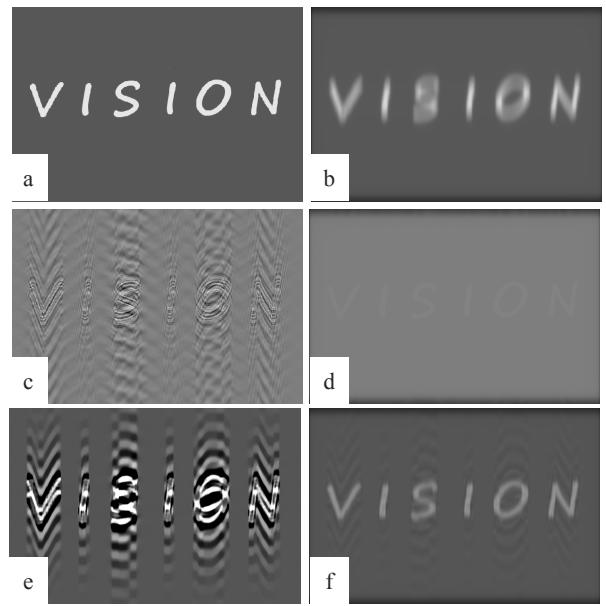


Fig. 11. Astigmatism & Myopia. Comparison for combined blur of -1D of myopia and -2.5D of astigmatism (axis: 45 degrees): a-b) original and convolved original images, c-d) precorrected and convolved precorrected images using Wiener filter, e-f) our precorrected and convolved precorrected images.

7.7 User Study

We performed an informal user study to evaluate the performance of our precorrection method. Our study is similar in spirit to that performed by Pamplona et al. [2012], but we do not require custom hardware. Huang et al. [2012] did not perform a user study. Our study consisted of 22 people (65% male, 35% female) with an average age of 28.3 years. Subjects were voluntarily recruited from a university campus. We performed five groups of tests. For each test, subjects had normal vision and we induced -2.75D of myopia by placing a lens in front of one eye (and covering the other eye). Subjects stood 2.5m from the display and viewed images of sizes ranging from 3cm (for text) to 31cm for the Lenna image. After a short explanation and calibration session (to find the optimal location for their individualized viewing), subjects were asked several questions and a response form was filled out for them. The following paragraphs provide a summary of these results.

. Study A - Precorrected vs. original: This first experiment wants to verify whether our method improves the perceived sharpness as compared to the original image. For this, we displayed side by side three pairs of images (one pair at a time): one precorrected (from Fig. 18) and the original. We ordered them randomly to prevent bias. The conclusion is that subjects preferred observing a precorrected image over the standard original image. We use a Likert-scale (where -3 means “strongly prefer original image”, -2 means “prefer original image”, -1 means “slightly prefer original image”, 0 means “similar”, +1 means “slightly prefer our method”, +2 means “prefer our method”, and +3 means “strongly prefer our method”) and the study resulted in a mean score (over all participants and images in this study) of 1.13 and confidence interval (0.48, 1.78), critical-t 2.09 and standard error of 0.31 (95% confidence interval). Moreover, 95% of

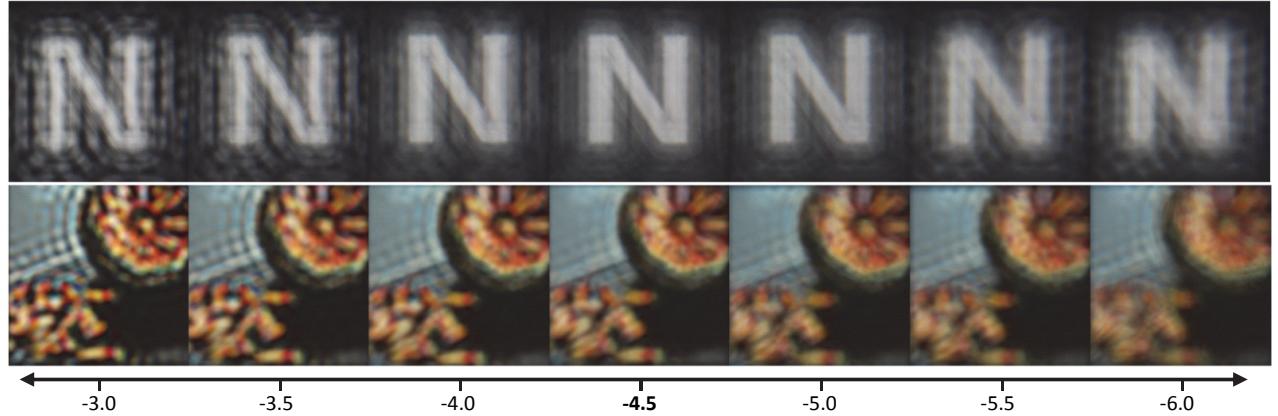


Fig. 12. **Precorrection Robustness.** Comparison of the same precorrected image (designed for -4.5D of refractive error) under different refractive errors ranging -1.5D to +1.5D from the intended refractive error. (*The 2nd row simulations were created using our technique applied to [Pamplona et al. 2012] © 2012 ACM, Inc. Reprinted by permission*).

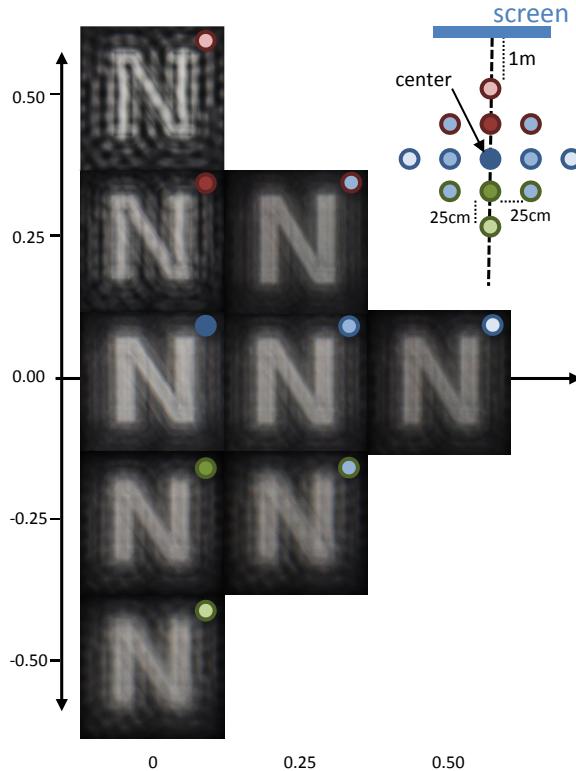


Fig. 13. **Displacement Robustness.** We present the effect of displacing the observer away from the intended observation point of a precorrected image. From the center (left column, third row), we show the captured precorrected images as the camera/observer moves forward/backward and to the right. The imagery resulting from moving to the left is similar and thus not shown.

the people selected at least one of our precorrected images over the original.

- **Study B - Displacement:** The second experiment seeks to evaluate the spatial robustness of our solution to displacement. For

this, we asked subjects to move 0.5m to the sides (left, right, front, and back) and to compare the perceived sharpness in comparison with the initial location. Moving laterally, 84.2% of the subjects claimed that this did not significantly affect image quality. Using a Likert-scale (where -2 means “much worse than center”, -1 means “worse than center”, 0 means “similar than center”, +1 means “better than center”, and +2 means “much better than center”), the mean value of the questionnaire responses for all participants was -0.052. Moving forward/backward, however, did result in increased sensitivity, with 48% of the subjects indicating moving forward or backward decreased image quality but with a mean value of just -0.11. We conclude that the displacement performed does not significantly affect solution quality.

- **Study C - Reading:** The third experiment tries to analyze whether our system improves readability. For this, we presented to the subjects in random order two different images of 3cm tall text: one precorrected by our system (similar to Fig. 15) and the corresponding original image, and ask them to read the letters. 82% of the people indicated it was easier to read our precorrected text image over the original text.

- **Study D - Contrast and Ringing Sensitivity:** This experiment analyzed whether the subjects would prefer to trade off contrast with more sharpness. For this, we displayed a set of 9 precorrected randomly placed images and asked the volunteers to rank them. The precorrected images (similar to left column of Fig. 2) presented three levels of contrast loss (20%/40%/60%) and three different τ (low/mid/high). All subjects choose as first option the least contrast loss (20%), and the largest cluster of same-respondents (45%) chose the middle τ . This implies that humans prefer high contrast and suggests that either, lower or higher levels of ringing, are counter-productive.

- **Study E - Comparison:** The last experiment wants to compare our system with a Wiener filter alternative. For this, we display the same text with three precorrections: one Wiener-filter based (3rd row of Figure 3), one using inverse DFT, and one using our method. No subject preferred the inverse DFT. 95% of the subjects preferred our precorrected image and 5% chose the Wiener filtered image. Moreover, 86% could read our precorrected text while only 9% could read the Wiener-filtered image.

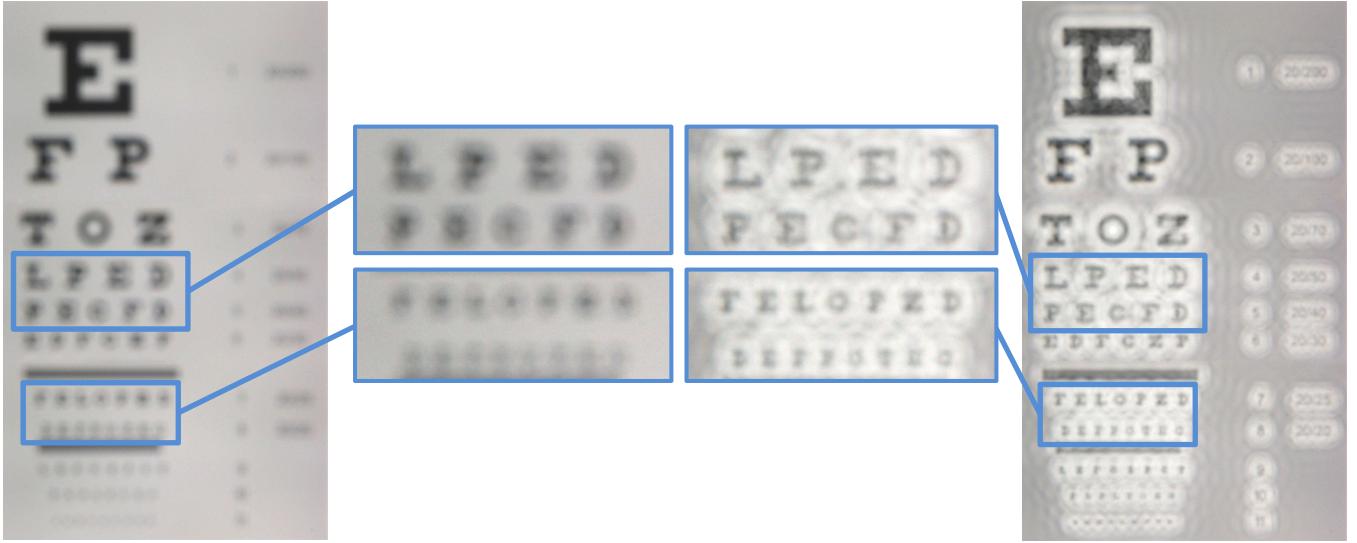


Fig. 14. **Snellen Eye Chart.** (left) A captured original image under -5D at 1.5m. (right) A captured precorrected image under the same aberration. (center) Close-ups of two regions of the eye chart. Our approach helps to improve visual acuity.

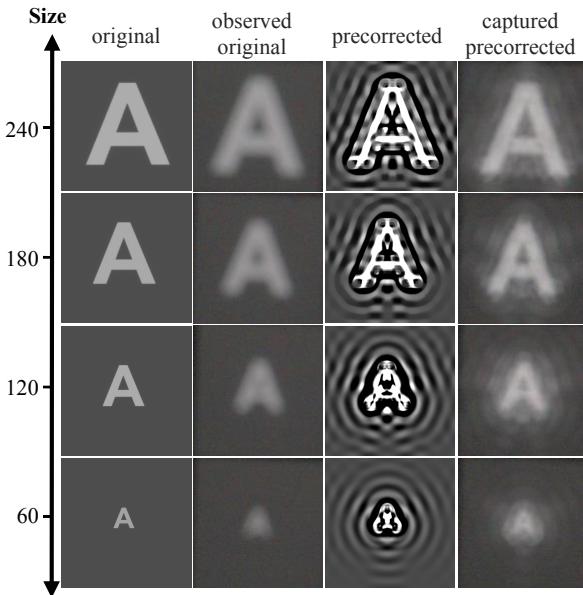


Fig. 15. **Reading Improvement.** We show an ‘A’ under -5D and at 1.5m. The letter size in pixels is shown to the left. Our captured precorrected images show a clear improvement over directly observing the original image.

7.8 3D Scene

Finally, Figure 19 shows a novel application of our method to 3D scene precorrection. We capture as the original image a 3D scene as seen from a digital camera. Then, we precorrect the original image and use a projector-camera system to alter the appearance of the physical scene (e.g., object) to that of the precorrected image. Thus, when we now look at the scene from the point of view of the camera, and with the same amount of refractive error, the scene

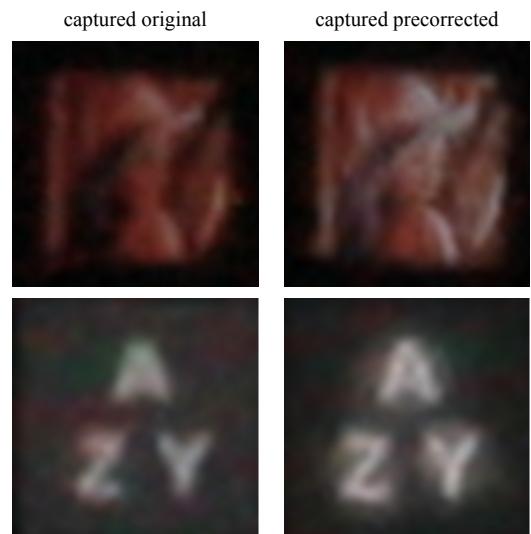


Fig. 16. **Distant and Small Content.** We focus a camera simulating a person with 20/20 vision on content of size 1.5cm at 3.1m. Even under normal vision, our precorrected image is able to produce sharper edges and details (right) as compared to viewing the original image (left). (*The 1st row simulations were created using our technique applied to the Lena picture, kindly provided by the USC SIPI Image Database*).

appears sharper. Our projector-camera system is based on a spatial-augmented reality framework, such as that of Raskar et al. [2001], but using colored objects as in Aliaga et al. [2008].

7.9 Limitations

Our method is not without limitations. To provide sharper text under a refractive aberration we require sufficient “space” between letters and “pixels per inch” in order to precorrect them properly

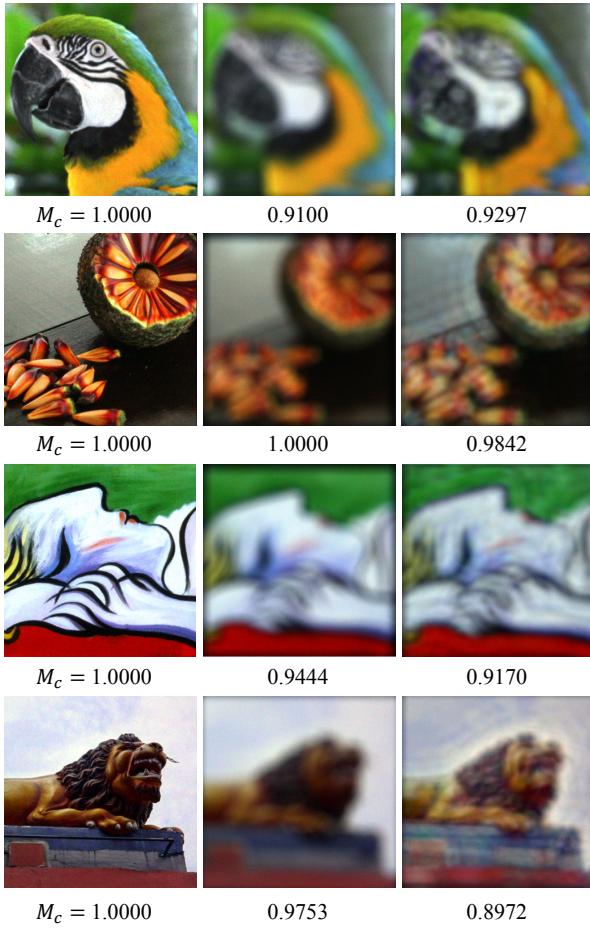


Fig. 17. **Additional Examples & Contrast Comparison.** We show several more examples as well as their Michelson contrast ratio (M_c), under -4.75D and at 1.5m. As comparison, the top right picture is shown with $M_c = 0.13$ in Huang et al. [2012]. Overall, our approach yields sharper imagery at only 10% contrast loss as compared to the original. (1st Row Left: [Huang et al. 2012] © 2012 ACM, Inc. Reprinted by permission; 2nd Row Left: [Pamplona et al. 2012] © 2012 ACM, Inc. Reprinted by permission; 3rd Row Left: © 2009-Present <http://www.pablopicasso.org/>).

(i.e., not all font styles and text sizes are suitable to our methodology; we might need to increase font size a few times in order to support mid-range corrections). Our technique needs some amount of contrast loss in order for the TV method to produce a precorrected image that can mitigate the fundamental appearance of ringing. For arbitrary images, the subtle details in between sharp edges are typically lost. Our method concentrates on improving the sharpness of strong edges at the expense of blurring the space in between.

8. CONCLUSION

We have presented a constrained total-variation-based approach to improve the visual perception of images under refractive aberrations and under normal vision. Our method can reduce ringing artifacts in the precorrected images while sharpening edges of the captured precorrected images, and by doing so we can recover sharp edges with high contrast. As compared to other methodologies, we provide higher contrast and do not need specialized hardware. Fur-



Fig. 18. **Precorrected Images for Human Observation.** We show images used during our user study for human observation. All were seen 31cm tall on a LCD with -2.5D of blur at 2.5m (assuming 5.6mm pupil). a) 20% contrast loss; b) 15% contrast loss; c) 10% contrast loss. (These simulations were created using our technique applied to the Lena picture, kindly provided by the USC Sipi Image Database).

ther, we produce imagery that is preferred by humans as per our user study. Our work makes significant strides towards printed and displayed content that can be personalized to an individual.

As future work, we are pursuing several directions. First, we initially collaborated with a perceptual psychologist and now have initiated a close collaboration with an ophthalmologist in order to further improve our methodology to include perceptual issues. Second, we would like to segment the original image in order to perform local TV analysis and potentially achieve an overall better captured precorrected image. Third, we would like to explore the option of hiding information within the deblurring. For example, similar to watermarking, we could encode one (or more) hidden original images into the precorrected image. Fourth, we are inter-

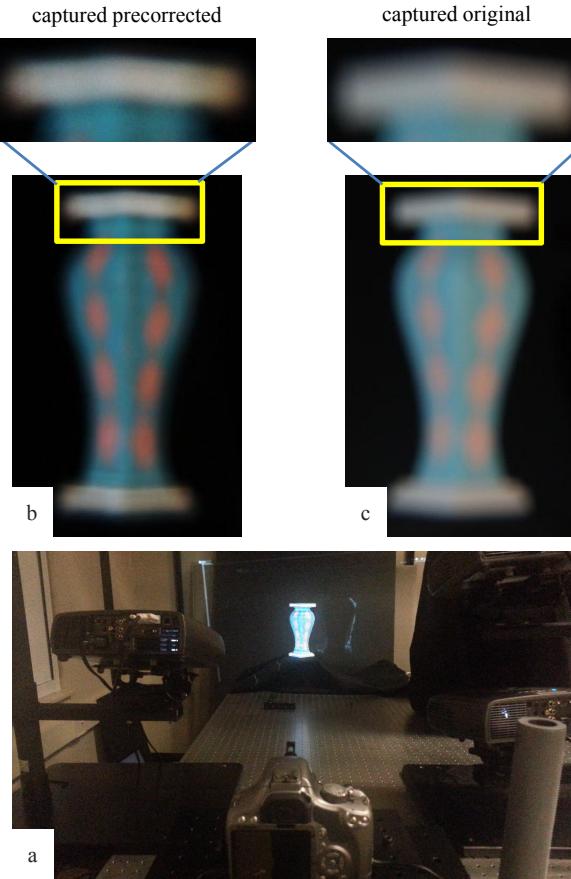


Fig. 19. **3D Scene Precorrection.** a) A projector-camera setup, b) captured precorrected object and c) captured original object. This prototype enables a glasses-free viewer to observe a 3D scene at increased sharpness (-4.75D at 2m).

ested in applying our technique directly to videos, while maintaining temporal coherence. Finally, our current solution produces the same result for each eye; thus, we are interested in exploring how to provide binocular correction.

9. ACKNOWLEDGEMENTS

This work is partially funded by NSF 1250232, 0913875, and 1302172. The first author would like to thank Plamen Stefanov for his financial and academic support during some stages of this work. Manuel M. Oliveira acknowledges CNPq (grants 482271/2012-4 and 308936/2010-8). We also thank Yi-Liu Chao and Tyler Smith for their work in earlier versions of this research.

REFERENCES

- ALIAGA, D. G., LAW, A. J., AND YEUNG, Y. H. 2008. A virtual restoration stage for real-world objects. *ACM Trans. Graph.* 27, 5, 149–159.
- ALONSO JR., M. AND BARRETO, A. 2003. Pre-compensation for high-order aberrations of the human eye using on-screen image deconvolution. In *Engineering in Medicine and Biology Society*. Vol. 1. 556–559.
- ALONSO JR., M., BARRETO, A., AND ADJOUADI, M. 2008. Digital image inverse filtering for improving visual acuity for computer users with visual aberrations. *Inverse Problems in Science and Engineering* 16, 957–966.
- ALONSO JR., M., BARRETO, A., ADJOUADI, M., AND JACKO, J. A. 2006. Howard: High-order wavefront aberration regularized deconvolution for enhancing graphic displays for visually impaired computer users. In *ICCHP*. 1163–1170.
- BASS, M., DECUSATIS, C., ENOCH, J., LAKSHMINARAYANAN, V., LI, G., MACDONALD, C., MAHAJAN, V., AND VAN STRYLAND, E. 2009. *Handbook of Optics*. Vol. 3. McGraw-Hill Education.
- BECK, A. AND TEBOULLE, M. 2009. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Image Processing* 18, 11, 2419–2434.
- BERTHOUZOZ, F. AND FATTAL, R. 2012a. Apparent resolution enhancement for motion videos. In *ACM SAP*. 91–98.
- BERTHOUZOZ, F. AND FATTAL, R. 2012b. Resolution enhancement by vibrating displays. *ACM Trans. Graph.* 31, 2 (Apr.), 15:1–15:14.
- BROWN, M., SONG, P., AND CHAM, T.-J. 2006. Image pre-conditioning for out-of-focus projector blur. In *IEEE CVPR*. Vol. 2. 1956–1963.
- CHAMBOLLE, A. 2004. An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision* 20, 1-2, 89–97.
- CHAN, T. F., GOLUB, G. H., AND MULET, P. 1999. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J. Sci. Comput.* 20, 6, 1964–1977.
- CHAN, T. F. AND SHEN, J. J. 2005. *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*. SIAM.
- CORNSWEET, T. 1970. *Visual perception*. Academic Press.
- DIDYK, P., EISEMANN, E., RITSCHEL, T., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2010. Apparent display resolution enhancement for moving images. *ACM Trans. Graph.* 29, 4, 113–120.
- DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2011. A perceptual model for disparity. *ACM Trans. Graph.* 30, 4, 96–105.
- DUANE, T. D. 1985. Clinical Ophthalmology. 6, 5434.
- DUNAWAY, D. AND BERGER, I. 2006. Worldwide distribution of visual refractive errors and what to expect at a particular location. *Presentation to International Society for Geographic and Epidemiologic Ophthalmology*.
- FORTUNATO, H. AND OLIVEIRA, M. 2014. Fast high-quality non-blind deconvolution using sparse adaptive priors. *The Visual Computer* 30, 6–8, 661–671.
- HAMPSON, K. M. 2008. Adaptive optics and vision. *Journal of Modern Optics* 55, 21, 3425–3467.
- HUANG, F.-C., LANMAN, D., BARSKY, B. A., AND RASKAR, R. 2012. Correcting for optical aberrations using multilayer displays. *ACM Trans. Graph.* 31, 6, 185. ©2012 ACM, Inc. <http://dx.doi.org/10.1145/2366145.2366204>.
- JOSHI, N., ZITNICK, C., SZELISKI, R., AND KRIEGLAN, D. 2009. Image deblurring and denoising using color priors. In *IEEE CVPR*. 1550–1557.
- KARUNASEKERA, S. AND KINGSBURY, N. 1994. A distortion measure for image artifacts based on human visual sensitivity. In *IEEE ICASSP*. Vol. 5. 117–120.
- KRISHNAN, D. AND FERGUS, R. 2009. Fast image deconvolution using hyper-laplacian priors. In *NIPS*. Vol. 22. 1–9.
- KRISHNAN, D., PHAM, Q. V., AND YIP, A. M. 2009. A primal-dual active-set algorithm for bilaterally constrained total variation deblurring and piecewise constant mumford-shah segmentation problems. *Adv. Comput. Math.* 31, 1-3, 237–266.
- LAKSHMINARAYANAN, V. 2012. New results in biomedical image processing. In *PHOTONICS*. 1–3.
- LAKSHMINARAYANAN, V. AND FLECK, A. 2011. Zernike polynomials: a guide. *Journal of Modern Optics* 58, 7, 545–561.

- LEGGE, G. E., PELLI, D. G., RUBIN, G. S., AND SCHLESKE, M. M. 1985. Psychophysics of reading—I. Normal vision. *Vision Research* 25, 2, 239–252.
- MARCHAND, P., PITE, E., GUILMET, H., AND TROCME, L. 2011. Systems and methods for rendering a display to compensate for a viewer's visual impairment. WO Patent App. PCT/US2011/039,993.
- MARZILIANO, P., DUFAUX, F., WINKLER, S., AND EBRAHIMI, T. 2002. A no-reference perceptual blur metric. In *IEEE Trans. Image Processing*. Vol. 3. 57–60.
- MARZILIANO, P., DUFAUX, F., WINKLER, S., AND EBRAHIMI, T. 2004. Perceptual blur and ringing metrics: Application to JPEG2000. *Signal Process. Image Commun.*, 163–172.
- MASIA, B., WETZSTEIN, G., ALIAGA, C., RASKAR, R., AND GUTIERREZ, D. 2013. Display adaptive 3d content remapping. *Comput. & Graph.* 37, 8, 983 – 996.
- MASIA, B., WETZSTEIN, G., DIDYK, P., AND GUTIERREZ, D. 2013. Special section on advanced displays: A survey on computational displays: Pushing the boundaries of optics, computation, and perception. *Comput. & Graph.* 37, 8, 1012–1038.
- MICHELSON, A. A. 1995. *Studies in optics*. Courier Dover Publications.
- MOHAMMADPOUR, S., MEHRIDEHNAVI, A., RABBANI, H., AND LAKSHMINARAYANAN, V. 2012. A pre-compensation algorithm for different optical aberrations using an enhanced Wiener filter and edge tapering. In *IEEE ISSPA*. 935–939.
- MONTALTO, C. 2014. Stability analysis of three inverse problems. Ph.D. thesis, Purdue University, West Lafayette.
- OYAMADA, Y. AND SAITO, H. 2007. Focal pre-correction of projected image for deblurring screen image. In *IEEE CVPR*. 1–8.
- PAMPLONA, V. F., MOHAN, A., OLIVEIRA, M. M., AND RASKAR, R. 2010. Netra: Interactive display for estimating refractive errors and focal range. *ACM Trans. Graph.* 29, 4 (July), 77–84.
- PAMPLONA, V. F., OLIVEIRA, M., ALIAGA, D., AND RASKAR, R. 2012. Tailored displays to compensate for visual aberrations. *ACM Trans. Graph.* 31, 4, 81–92. ©2012 ACM, Inc. <http://dx.doi.org/10.1145/2185520.2185577>.
- PELI, E. AND WOODS, R. L. 2009. Image enhancement for impaired vision: The challenge of evaluation. *International Journal on Artificial Intelligence Tools* 18, 3, 415–438.
- PIZLO, Z. 2010. *3D shape: its unique place in visual perception*. MIT Press.
- PIZLO, Z., ALLEBACH, J., AND LEE, B.-S. 2007. Characterization of red-green and blue-yellow opponent channels. *Journal of Imaging Science* 51, 1, 23–33.
- RASKAR, R., AGRAWAL, A., AND TUMBLIN, J. 2006. Coded exposure photography: Motion deblurring using fluttered shutter. *ACM Trans. Graph.* 25, 3, 795–804.
- RASKAR, R., WELCH, G., LOW, K.-L., AND BANDYOPADHYAY, D. 2001. Shader Lamps: Animating Real Objects With Image-Based Illumination. In *Rendering Techniques*. Eurographics. 89–102.
- RITSCHEL, T., SMITH, K., IHRKE, M., GROSCH, T., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2008. 3D Unsharp masking for scene coherent enhancement. *ACM Trans. Graph.* 27, 3, 90–97.
- RUDIN, L. I., OSHER, S., AND FATEMI, E. 1992. Nonlinear total variation based noise removal algorithms. *Physica D* 60, 1-4, 259–268.
- RUSINKIEWICZ, S., BURNS, M., AND DECARLO, D. 2006. Exaggerated Shading for Depicting Shape and Detail. *ACM Trans. Graph.* 25, 3, 1199–1205.
- SALADIN, K. 2007. Anatomy & physiology, the unity of form and function, 5th.
- STENGEL, M., EISEMANN, M., WENGER, S., HELL, B., AND MAGNOR, M. 2013. Optimizing apparent display resolution enhancement for arbitrary videos. *IEEE Trans. Image Processing* 22, 9, 3604–3613.
- TEMPLIN, K., DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2013. Apparent resolution enhancement for animations. In *Proceedings of the 27th Spring Conference on Computer Graphics*. SCCG '11. ACM, New York, NY, USA, 57–64.
- THIBOS, L., APPLEGATE, R. A., SCHWIEGERLING, J. T., AND WEBB, R. 2000. Standards for reporting the optical aberrations of eyes. In *Vision Science and its Applications*. Vol. 35. Optical Society of America, 232–244.
- TREVOR-ROPER, P. D. 1984. The eye and its disorders. *Journal of Clinical Pathology*.
- VEERARAGHAVAN, A., RASKAR, R., AGRAWAL, A., MOHAN, A., AND TUMBLIN, J. 2007. Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture. *ACM Trans. Graph.* 6, 3.
- VILLEGRAS, E. A., ALCÓN, E., AND ARTAL, P. 2008. Optical quality of the eye in subjects with normal and excellent visual acuity. *Investigative Ophthalmology & Visual Science* 49, 10, 4688–4696.
- VISIONURL. 2013. www.vision2020.org, April 28.
- VOGEL, C. R. 2002. *Computational methods for inverse problems*. SIAM.
- WANDELL, B. A. 1995. *Foundations of vision*. Sinauer Associates.
- WANG, Y., YANG, J., YIN, W., AND ZHANG, Y. 2008. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Img. Sci.* 1, 3, 248–272.
- YUAN, L., SUN, J., QUAN, L., AND SHUM, H.-Y. 2008. Progressive inter-scale and intra-scale non-blind image deconvolution. *ACM Trans. Graph.* 27, 3, 74–83.
- ZHANG, L. AND NAYAR, S. 2006. Projection defocus analysis for scene capture and image display. *ACM Trans. Graph.* 25, 3, 907–915.