DIGITAL IMAGE INVERSE FILTERING FOR IMPROVING VISUAL ACUITY FOR
COMPUTER USER’S WITH VISUAL ABERRATIONS

Miguel Alonso Jr.
Department of Electrical and Computer Engineering
Florida International University
Miami, Florida, U.S.A.
malons05@fiu.edu

Armando Barreto
Department of Electrical and Computer Engineering
Florida International University
Miami, Florida, U.S.A.
barretoa@fiu.edu

Malek Adjouadi
Department of Electrical and Computer Engineering
Florida International University
Miami, Florida, U.S.A.
adjouadi@fiu.edu
ABSTRACT

In this paper, we describe an approach for improving the effective visual acuity for computer users with visual aberrations. This research is based on inverse filtering of images, compensating the visual acuity loss introduced by the user’s visual aberration. These images, when displayed to the user, will counteract his/her visual aberration. The method described in this paper integrates spatial information processing mechanisms to reduce the negative impact of some shortcomings of the inverse filtering solution.

INTRODUCTION

Graphical user interfaces (GUIs) for computers have been developed to suit the needs of the general population. However, GUIs have, in some cases, disregarded the needs of a subset of computer users with visual aberrations. For blind computer users, there are programs called “screen readers” that parse the screen and provide an auditory representation of what is being displayed. There has also been significant progress in the development of 3D audio environments for visualization of computer tasks.

For individuals with low vision, there are programs that magnifying portions of the screen. However, there are other kinds of visual afflictions (e.g., Keratoconus [10]) that are not remedied by those types of assistance. Beyond magnification, other forms of enhancement are needed.

For centuries the distortion introduced in the images perceived by the human eye, due to imperfections in its refractory system, could only be determined by the
corresponding visual deficiency symptoms and specified in terms of the strength needed in their correction (e.g., “-4.5 Diopters” ). Only recently instrumentation has been developed to objectively evaluate and characterize the Wavefront Aberration in human eyes [7].

There has been previous work geared towards enhancing the computer access of individuals with severe refractive visual aberrations. These aberrations degrade the optical (as opposed to neurological) function of the user’s eye. Alonso et al. [2], have described a method to perform the pre-compensation of onscreen images according to the user’s particular visual aberration. This method relies on a priori knowledge of the visual aberration, characterized by wavefront analyzers. Through a series of transformations, these analyzers provide a way to know how the optical portion of a user’s eye will alter the images displayed on an LCD panel.

Although this method provides improvement of the visual acuity, the inherent lack of dynamic range of most display devices limits the effectiveness of the approach. Further processing is needed to improve the contrast in the final image, as perceived by the subject. This paper describes the sequence of incremental improvements that we have applied to our initial application of mean square error (Wiener) filtering to this problem [2].

BACKGROUND

The algorithm presented here relies on the linear systems approach to modeling the human visual system, known as Fourier Optics [5]. In it, the human visual system is modeled as a linear system having an impulse response $H$. The output of this system is
therefore the convolution of the input with the impulse response of the system. The impulse response of an ideal optical system, including the human eye, is a delta, or a point with zero spread.

Thus, if a user’s eye is free from any visual aberrations, the impulse response of his/her eye or Point Spread Function (PSF) will be a two-dimensional delta function. This allows a clear, undistorted projection of objects onto the retina.

If however, the user has a visual aberration, the PSF will not be a delta, and the retinal projection of the object will be distorted. This, in turn, will not allow the user to interact efficiently with the personal computer (PC) via the graphical display.

The PSF for a human eye can be measured using a wavefront analyzer, which is becoming a standard tool in ophthalmology. Thus, in the near future, the PSF will be readily available for users that stand to benefit from the approach proposed in this paper. Figure 1 shows an abstract representation of the use of this Linear Shift Invariant (LSI) model to describe the optical process that takes place in an aberrated human eye.

![Figure 1 - Block diagram of Visual System](image)
An object $O(x,y)$ (for example, a picture on a graphical display) is degraded by convolution with the PSF of the user’s visual system, $H(x,y)$, resulting in a distorted projection of the object on the user’s retina, $I(x,y)$. This is described by

$$I(x,y) = H(x,y) \ast O(x,y)$$

(1)

where $\ast$ denotes convolution.

Given $O$ and $H$, we seek to inverse-filter $O$ with $H$ to produce an enhanced object, $RD$, shown in Equation (2). This new image should counteract the distortion introduced by $H$, such that when the user views the $RD$ on the graphic display, an undistorted version of $O$ would be projected onto the retina. Figure 2 illustrates this process.

$$RD(x,y) = H^{-1}(x,y) \ast O(x,y)$$

(2)

![Figure 2 - Pre-compensation process](image)
This model is essentially a noiseless deconvolution problem. The ill-conditioned nature of the inverse process applied to the PSF of the human eye [7] however, will require a robust method to generate images that allows a more efficient interaction between a user that has a refractive error present in his/her visual system and the PC. In our approach we make use of a priori information available about the properties of the original object, which makes our method similar to the “descriptive regularization” approach to the solution of ill-posed problems, previously studied by Morozov and colleagues [8] [9]. We refer the reader to [6][13][14] for an in-depth description of how the PSF of the human eye is obtained experimentally.

**INVERSE FILTERING PROCESS**

The PSF is used to inverse-filter the image or text to be displayed, providing compensation for the aberrations in the refractive portion of the users’ visual system. Initially, we chose to use the Weiner filter, which can be summarized by the following frequency-domain equation to obtain the pre-compensated display image [4]:

\[
RD(f_x, f_y) = \left[ \frac{1}{H(f_x, f_y)} \frac{|H(f_x, f_y)|^2}{|H(f_x, f_y)|^2 + K} \right] I(f_x, f_y)
\]

where \(RD(f_x, f_y), H(f_x, f_y), \) and \(I(f_x, f_y)\) are the Fourier transforms of \(RD(x,y)\), the display image, \(H(x,y)\) the PSF, and \(I(x,y)\), the original image, respectively. \(K\) is a specified constant that controls the accuracy of the deconvolution.
This, however, provides an approximation to the ideal inverse of H. Traditionally, this is what is desired in inverse processing. But in human beings, the visual system will never be perfect because it is diffraction-limited [13]. The diffraction-limited PSF is an Airy disk [11], not a perfect delta. We can use this to alleviate some of the problems with higher frequencies when using a straight application of the Weiner filter. In the inverse of H, we replace the numerator 1, with the frequency domain representation of the diffraction limited PSF. This provides a higher level of contrast in the initial generation of the pre-compensated image. This is summarized in Equation (4).

\[
RD(f_x, f_y) = \left[ \frac{H_{sp}(f_x, f_y)}{H(f_x, f_y)} \right] \left[ \frac{|H(f_x, f_y)|^2}{|H(f_x, f_y)|^2 + K} \right] I(f_x, f_y)
\]  

(4)

This equation already provides a pre-compensated image RD(x,y) with slightly better contrast than Equation (3). However, in addition to deconvolving the known PSF, this processing introduces artifacts in the form of “ripples” in the pre-compensated image [1]. Furthermore, the numerical values of RD(x,y) found mathematically by Equation (4) may not be constrained to the range of displayable grayscale levels [0, 255].

Figure 3 - Example of an icon pre-compensated using Equation (4)
The ripples that are produced as a result of using Equation (4) are a roadblock for further processing of the pre-compensated image. Because display devices can only represent grayscale levels between 0 (black) and 255 (white), the pre-compensated image must be scaled and shifted from a bipolar range, to a unipolar range for display. This process of scaling and shifting reduces the available contrast in the resulting retinal image [7]. Thus, post-processing of the image obtained from the deconvolution stage is necessary, but it must be implemented in a way that does not exacerbate the negative impact of residual ripples introduced by the deconvolution.

For example, the use of histogram equalization will improve the contrast in the pre-compensated image (which will also enhance the contrast in the image perceived by the user), but it will disproportionately amplify any ripples outside of the area of interest, which will then unduly compete for the viewer’s attention.

We propose a method to extract only the necessary information from the initial pre-compensated image, so that the post processing contrast enhancement may correctly enhance the final pre-compensated image.

In this paper, text images are considered binary images, composed of black letters on white background. When these images are pre-compensated, only a finite region around the boundary of the letter or symbol is necessary to yield its correct perception by the viewer. This is demonstrated in Figure 4, where a black bar has been inserted across the pre-compensated image and the blurring introduced by viewing through and imperfect PSF has been simulated. The bar removes information from the pre-compensated result; however, this information is not necessary for correct perception of the pre-compensated result.
Figure 4- Not all of the information in the Pre-compensated image (left) is necessary to provide an adequate retinal image (right)

The resulting simulated retinal image is still clear, showing that not all of the information in the pre-compensated image is necessary for a good reconstruction.

The first step in extracting the necessary information from the pre-compensated image is to find the areas of interest. This is done by first segmenting the original image, blurring the segmented image using the PSF of the user, and then thresholding the blurred, segmented original image to produce a mask that contains zeros where there is unnecessary information and ones where the information is to be extracted. The mask is then used in a point-to-point multiplication with the pre-compensated image.

Image segmentation is used to extract the relevant information in the digital image. The first step in segmenting the images is to edge-detect the image. We chose to use Robert’s cross-gradient operators [4] because their implementation is computationally efficient and provide a simple way to compute the first derivatives of an image. The kernels used are shown in Figure 5.
Because the images that we are using are noise-free, the Roberts cross-gradient operators are sufficient for providing an accurate approximation of the gradient. Thus, the kernels shown in Figure 5 are used to approximate the gradient by means of a two-dimensional filter [4]. The edge detected image is then found by taking the magnitude of the gradient vector for each pixel. Figure 6a shows an edge detected icon. We refer the reader to [4] for further details on this process.

Because gaps may exist in the edge detected image, this image is dilated using first a horizontal structuring element, and then a vertical structuring element. The structural elements are 3 pixels long. Dilation is a morphological operation and is defined as follows:
“The dilation of A by B is the set of all displacements, z, such that B and A overlap by at least one element.” [4]

This will close any gaps that may exist in the edge-detected image (Figure 6b) and is written as [4]:

\[ A \oplus B = \left\{ z \left| \left( \hat{B} \right)_z \cap A \subseteq A \right. \right\} \] (5)

The next step is to fill in the dilated image. This is done by filling ‘holes’ in the binary image. A hole is defined as a set of background pixels that cannot be reached by filling in the background from the edge of the image. The algorithm used is outlined in [12]. The result of this step is shown in Figure 6c.

The final step is to erode the image by a pyramidal structuring element using Equation (6) [4]. In words:

“The erosion of A by B, denoted by \( \ominus \), is the set of all points z, such that B, translated by z, is contained in A.” [4]

\[ A \ominus B = \left\{ z \left| \left( \hat{B} \right)_z \cap A \subseteq A \right. \right\} \] (6)
The final result of the image segmentation process is shown in Figure 6d. The segmented image can be used to create a mask that isolates the relevant regions of the raw pre-compensated image.

The procedure is summarized as follows:

1. Edge-detect the original image using an edge-detecting method, such as the one described above.

2. Blur the edge-detected image (Figure 6d) with the user’s PSF in Equation (1), where \( O(x,y) \) is the edge-detected image, \( H(x,y) \) is the user’s PSF, and \( I(x,y) \) is the blurred, edge-detected image.

3. Threshold \( I(x,y) \) at a very small value, close to zero, and replace all values above this threshold with one, and values below it with zeros, to produce the mask, \( IH(x,y) \) (e.g., Figure 7a).

4. Multiply the Pre-compensated image, \( RD(x,y) \), point-to-point with \( IH(x,y) \).
Although the PSF associated with some aberrations (e.g., Keratoconus) may not be circularly symmetric [1], this method allows for a robust generation of the appropriate mask, based on the user’s PSF.

Step four above results in an image with an artificially introduced black background and is not suitable for comfortable identification of the letters displayed. Conversely, arbitrary replacement of a completely white background gives rise to a “halo” effect that would also be distracting. In order to alleviate this, a background color closer to the un-processed pre-compensated image is necessary. This can be easily achieved by replacing the black intensity with the mean intensity located at the border of the masked pre-compensated retinal image. By edge-detecting the mask itself, we can produce a sub-mask that only extracts information from the image at the edge of the mask. Thus, if we edge-detect IH(x,y) using a simple Laplacian of the Gaussian filter [4], we obtain a mask that highlight only the edge information, show in Figure 7b.

Figure 7 - Mask ($IH(x,y)$) to be used for extracting relevant information from the initial pre-compensated image and its edge detected image
If we then multiply Figure 7b point-to-point with the blurred version of the initial deconvolution result, and then take the mean of only the non-zero values, an estimate of the appropriate background intensity can be obtained.

The integration of this mean level with the masked pre-compensated image is implemented using the following equation:

\[ P_{\text{Cm}}(x, y) = R_{D}(x, y) \circ I_{H}(x, y) + (1-I_{H}(x, y)) \cdot m_{n} \]  \hspace{1cm} (6)

where \( m_{n} \) is the estimate of the appropriate background intensity, and \( \circ \) denotes point-to-point multiplication. Figure 8 shows the result of this final step, and the corresponding simulated retinal image.

VERIFICATION

An experiment was designed for the assessment of the enhancement in edge information preservation achieved through pre-compensation. The experiment was designed as follows: Two factors were studied: “Row”, which represents the size of the symbols used
for testing and “Pre-compensation”, which indicates whether or not the images to be blurred by simulated viewing were pre-compensated. The “Row” factor, R, consists of 9 levels, with each level consisting of the letters in rows 5 through 13 of a standard ETDRS eye chart. The “Pre-compensation” factor, M, consists of 2 levels, the perceived retinal image for a pre-compensated letter (level 1) and for the original letter (level 2).

RESULTS
In this test the objective is to estimate how well the edge information is preserved through the process of simulated viewing (convolution with the PSF chosen). As such, the measure of comparison is based in contrasting the edge-detected version of each original image with the edge-detected version of the image obtained after simulated viewing, under the two circumstances specified by levels 1 and 2 of factor M. Specifically, the dependent variable of the experiment is calculated as the sum of squares difference between the edge-detected version of the original letter and the edge-detected version of the image resulting from simulated viewing. The lower the sum of squares difference, the more similar the retinal image is (with respect to its edges) to the original image. Each treatment had 5 replications (the 5 letters in each of the rows of the eye chart).

The experiment was conducted and an analysis of variance (ANOVA) was performed on the resulting data. Table 1 shows the resulting ANOVA table.
Table 1- ANOVA: Data versus R, M

<table>
<thead>
<tr>
<th>Factor</th>
<th>Type</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>fixed</td>
<td>9</td>
<td>5, 6, 7, 8, 9, 10, 11, 12, 13</td>
</tr>
<tr>
<td>M</td>
<td>fixed</td>
<td>2</td>
<td>1, 2</td>
</tr>
</tbody>
</table>

Analysis of Variance for Data

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>8</td>
<td>2452723</td>
<td>306590</td>
<td>42.61</td>
<td>0.000</td>
</tr>
<tr>
<td>M</td>
<td>1</td>
<td>1341269</td>
<td>1341269</td>
<td>186.41</td>
<td>0.000</td>
</tr>
<tr>
<td>R * M</td>
<td>8</td>
<td>169565</td>
<td>21196</td>
<td>2.95</td>
<td>0.007</td>
</tr>
<tr>
<td>Error</td>
<td>72</td>
<td>518064</td>
<td>7195</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>89</td>
<td>4481620</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The ANOVA yields a p-value of p<=0.0009 for M, which can be interpreted with the following statement: At a significance level of alpha = 0.05, there is a statistically significant difference in the dependent variable (edge preservation) between the levels of factor M, the application of the pre-compensation (yes/no). This is shown in Figure 9. In our study, the interest is focused on the statistical difference between methods; i.e., presence of pre-compensation (M=1) vs. absence of pre-compensation (M=2).
CONCLUSIONS

In this paper, we propose an algorithm that advances the development of techniques for inverse filtering images to be displayed for computer users with significant aberrations in their visual system. The motivation for this comes from the fact that high-order aberrations are not correctable using current ophthalmic methods, making the use of computers difficult for users with these types of refractive errors. Additionally, alternative methods that exist for the correction of wavefront aberrations (e.g., adaptive optics) are costly and inaccessible to the average user.

The statistical analysis of the quantitative assessment experiment described in the paper indicates that, indeed, applying the pre-compensation described to images to be viewed through optical aberrations of the type considered may aid in the preservation of the edge information in the retinal projection.

Since the ultimate goal of this research is to facilitate human-computer interaction, the next step will be to test the method and conduct a statistical analysis based on user trials, having users attempt to identify text with and without the pre-compensation applied.

For the practical implementation of the pre-compensation method, additional features may be required in the computer system. In particular, real-time evaluation of the pupil diameter in the eye of the user would be used to adjust the PSF utilized in the pre-compensation process, as the PSF is known to vary with pupil diameter changes [3].
Additionally, future developments may lead to the establishment of relations between the information available a priori from the original images and the parameters in our process, such as the value of $K$ in Equations (3) and (4), so that the performance can be maximized for specific images.

Upon further development, software-based custom pre-compensation approaches may offer users with otherwise uncorrectable high-order refractory aberrations a way to use GUIs more efficiently in their interaction with computers.

ACKNOWLEDGMENT

This work was sponsored by NSF grants IIS-0308155, CNS-0520811, HRD-0317692 and CNS-0426125. Mr. Miguel Alonso Jr. is the recipient of an FIU Dissertation Year Fellowship.

REFERENCES


