Position-Dependent Defocus Processing for Acoustic Holography Images

Ruming Yin,1 Patrick J. Flynn,2 Shira L. Broschat1

1 School of Electrical Engineering & Computer Science, Washington State University, P.O. Box 642752, Pullman, WA 99164-2752
2 Department of Computer Science and Engineering, 384 Fitzpatrick Hall, University of Notre Dame, Notre Dame, IN 46556

Received 15 October 2001; accepted 22 March 2002

ABSTRACT: Acoustic holography is a transmission-based ultrasound imaging method that uses optical image reconstruction and provides a larger field of view than pulse-echo ultrasound imaging. A focus parameter controls the position of the focal plane along the optical axis, and the images obtained contain defocused content from objects not near the focal plane. Moreover, it is not always possible to bring all objects of interest into simultaneous focus. In this article, digital image processing techniques are presented to (1) identify a “best focused” image from a sequence of images taken with different focus settings and (2) simultaneously focus every pixel in the image through fusion of pixels from different frames in the sequence. Experiments show that the three-dimensional image information provided by acoustic holography requires position-dependent filtering for the enhancement step. It is found that filtering in the spatial domain is more computationally efficient than in the frequency domain. In addition, spatial domain processing gives the best performance. © 2002 Wiley Periodicals, Inc. Int J Imaging Syst Technol, 12, 101–111, 2002; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ima.10017

I. INTRODUCTION

Ultrasound imaging techniques can be divided into two modalities: reflection mode (or pulse-echo) and transmission mode. Pulse-echo ultrasound images see widespread use in medicine for diagnosis because of their safe nature and relatively low cost. Despite its popularity, pulse-echo ultrasound has several limitations, including the typical difficulty of image interpretation, a small field of view, and the presence of significant image artifacts.

Acoustic holography is a transmission mode imaging technique that was explored intensively in the 1970s (Ermert and Karg, 1979; Mueller, 1986). It involves a two-step procedure: first, coherent interference of the image signal with a corresponding reference wave is obtained; second, the interference pattern is digitized to produce an image. The resulting images are similar to those obtained with X-rays and thus are more readily interpretable and provide a larger field of view than pulse-echo ultrasound images. A novel acoustic imaging system, employing the acoustic holography approach and called optical sonography (Advanced Diagnostics, Inc., Richland, WA), was developed to overcome many of the limitations of pulse-echo ultrasound (Garlick, 1993). The system utilizes an acoustic plane wave that is transmitted through an object to produce an acoustic hologram. The hologram is translated from acoustical to optical wavelengths and then digitized. The resulting image allows direct visualization of soft-tissue characteristics. A more detailed description appears in Fecht et al. (1998). In this work, we focus on the postprocessing of images obtained from the acoustic holography imaging system described above. Our goal is to improve the quality of images produced by the system and, thereby, to facilitate diagnosis and other applications.

While providing a larger field of view, acoustic holography also preserves the advantages of real-time imaging and low cost. As for all optical systems, however, acoustic holography introduces a defocusing problem because image reconstruction is performed optically. We approach this problem using techniques developed for typical optical images.

In an acoustic holography imaging system, an image of the object is produced from acoustic waves passing through it. Ideally, this image is only focused on tissues at a given depth along the optical axis (i.e., in one planar slice of the object). In practice, the image is focused over a small range of depths, and objects at other depths, while blurred, are still visible. This characteristic behavior motivated two research goals. First, we wish to obtain an image in which the object of interest is best focused. The best focused image is chosen from a sequence of images obtained at different focal settings. Therefore, we need a mechanism to determine whether a given object is in focus. We can realize this focus recognition using a focus measure technique, as discussed in Section III. Second, for diagnosis, we expect the image of the object to be focused. Also, techniques such as edge detection and image segmentation, which may be applied in the future, are more easily realized on focused than on defocused images. Therefore, we wish to increase the “in focus” interval. The problem of increasing the focusing range is cast as an
II. ACOUSTIC HOLOGRAPHY IMAGE FORMATION

Unlike conventional pulse-echo ultrasound, which uses reflected acoustic energy to construct an image, a holography imaging system uses a transmitted ultrasound wave to produce a fluoroscope-like image of the object. The ultrasound wave passes through the object, is mixed with a reference wave, and is then received by a holography detector, which converts the acoustic hologram into an image. The acoustic holography image formation process is analogous to typical optical imaging.

In a convex lens optical system, the object image formed on the receiver plane is focused if the object is located at the position predicted by the lens equation. The relationship between the object position and the lens parameters follows the well-known lens formula

\[
\frac{1}{f} = \frac{1}{u} + \frac{1}{v},
\]

where \(f\) is the focal length of the lens, \(u\) is the distance between the object and the lens, and \(v\) is the distance between the image plane and the lens. If one variable of the lens formula is changed, the image formed on the receiver plane is blurred. Figure 1 shows the focused image \(f(x, y)\) and the blurred image \(g(x, y)\) for image plane positions of \(v\) and \(s\), respectively. The degree of blurring increases as the difference between \(v\) and \(s\) increases. The radius \(R\) of the blur circle is given by

\[
R = \frac{D}{2v} \delta,
\]

where \(D\) is the diameter of the lens, and \(\delta = |x - v|\). Figure 2 shows the acoustic holography imaging system, which has a similar image formation mechanism and consists of an optical system with two lenses, \(L_1\) and \(L_2\). Its effective focal length \(f\) can be adjusted by moving \(L_1\) with respect to \(L_2\). \(L_2\) is used to maintain the image magnification while \(L_1\) is moving. In the schematic, a point source \(s\) is imaged as the out-of-focus image \(g(x, y)\) on the receiver plane, whereas \(f(x, y)\) is an ideal focused image.

A. Image Defocusing Problem. When considered as a two-dimensional (2-D) signal, an image is the output signal produced by the imaging system from the source signal. Ideally, the system is linear and shift-invariant. In fact, Horn (1986) claims that most complicated incoherent optical image processing systems actually are linear and shift-invariant. Thus, since a linear and shift-invariant system performs a convolution, we can think of defocusing as a convolution problem. We note that this is not strictly true for our situation due to the complexity of ultrasonic transduction and acoustic lens nonlinearities. Nonetheless, we employ an image restoration framework and present results to demonstrate its utility.

Let \(f(x, y)\) be the focused image when the receiver plane is at the focus position; the out-of-focus image \(g(x, y)\) is the convolution of the focused image \(f(x, y)\) with the point spread function \(h(x, y)\) of the defocusing transform. This relationship can be expressed as

\[
g(x, y) = f(x, y) \ast h(x, y).
\]

Given an ideal image \(f(x, y)\), the degree of defocusing depends on the point spread function (PSF) \(h(x, y)\). From Figure 2, we see that blurring increases as the receiver plane shifts away from the correct position along the optical axis \(z\). Therefore, the PSF \(h(x, y)\) is a position-dependent function. We introduce one more parameter \(\tau(z)\), called the focus spread parameter, into \(h(x, y)\) to indicate this dependency

\[
h(x, y) = h(x, y, \tau),
\]

where \(\tau = \tau(z)\) is a function of the location on the optical axis \(z\).

With the convolution model, if the source \(s\) is an ideal point source, the blurred image \(g(x, y)\) is the PSF \(h(x, y)\) of the imaging system. The blurred image \(g(x, y)\) is symmetric, because the optical system is circularly symmetric around the optical axis \(z\). Therefore, the PSF should be a circularly symmetric function.

B. PSF Models. The PSF of the acoustic holography imaging system is not easily obtained. Two commonly used PSF models for optical systems are the uniform function and the Gaussian function. In a spatially incoherent optical imaging system, diffraction is limited, and the blurred image of a point source is circular. Blurring due to defocusing can be modeled as a convolution with a circular pulse (Horn, 1986). If the system is assumed to be lossless, the intensity is constant within the blur circle and zero outside the blur circle. The defocusing PSF is then given by

\[
h(x, y) = \begin{cases} 
\frac{1}{\pi R^2} & \text{when } x^2 + y^2 \leq R^2 \\
0 & \text{otherwise}
\end{cases},
\]

where \(R\) is the radius of the blur circle.
After the defocusing convolution has been performed, the degree of focusing depends on the parameter $R$ of the uniform PSF. In a general PSF $h(x, y, \tau)$, the influence of focus depth on the defocusing operation is represented by the focus spread parameter $\tau$. The parameter $R$ of the uniform PSF model is proportional to $\tau$. In this work, the radius of the blur circle $R$ is assumed to be equal to the focus spread parameter $\tau$.

In practice, the defocused image of a point object is roughly circular, with the intensity falling off gradually to zero. Therefore, a 2-D Gaussian function is a more faithful approximation to the physical defocusing than a uniform spot (Pentland, 1987). The defocused image can thus be described as the result of convolving the focused image with a Gaussian PSF

$$h(x, y) = \frac{1}{2\pi\sigma} e^{-(x^2+y^2)/2\sigma^2},$$

where $\sigma$ is the standard deviation of the Gaussian distribution and is assumed to be proportional to the radius of the blur circle. Hence, we treat $\tau$ and $\sigma$ as identical:

$$\tau = \sigma.$$

To find $g(x, y)$, we need to determine only the focus spread parameter $\tau$. As explained earlier, the true relationship governing the defocusing transformation is probably more complicated than this simple model. However, we are aware of no other work in the literature that addresses the modeling of defocus that is inherent in acoustic holography imaging path.

C. Determining the Focus Spread Parameter $\tau$. The parameter $\tau(z)$ describes the amount of defocusing in the out-of-focus image $g(x, y)$. For the geometric optics model, the amount of blurring can be described by the radius of the blur circle $R$ given by Eq. (2), and we can approximate $\tau$ by $R$ to get

$$\tau = \frac{D}{2v} \delta.$$

In our imaging system, a sequence of images is obtained by changing the focal plane position within a given interval. $M$ frames of images are recorded as $I_N (N = 1, 2, \ldots, M)$, where $N$ is called the index number of the image. An object will be most in focus in one frame $I_N$ and will be blurred in all other frames $I_N (N \neq N_f)$ of the image sequence. The focal plane of a defocused image $I_N$ is displaced by $\delta_N$ from the focal plane of the focused image $I_{N_f}$. The corresponding index shift $\Delta_N$ is defined as the absolute value of the difference between the index numbers $N$ and $N_f$:

$$\Delta_N = |N - N_f|.$$

Figure 3 depicts the focal displacement $\delta_N$ and the index numbers of the image sequence. The focal displacement $\delta_N$ increases as the index shift $\Delta_N$ increases. We assume that a simple linear relationship exists between $\delta_N$ and $\Delta_N$:

$$\delta_N = a\Delta_N + a_0.$$

III. FOCUS MEASURE

Focus measure is a technique used in conjunction with automatic focusing in many optical systems (Nayar, 1992). For our imaging system, we use focus measure to find the best focused image frame from a sequence of images. Previously, many different approaches have been explored, and several have yielded good results for different situations. All of them have examined the effect of focus deviation of the object in depth (the $z$ axis) from the focus position.

A. Gradient Magnitude. A higher contrast edge tends to have a larger gradient magnitude across the edge. Tenenbaum (1970) proposed the gradient magnitude method to find the best focus using the magnitude of the gradient in the region of interest. The best focus occurs when the gradient magnitude, defined by

$$|\nabla g(x, y)| = \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2},$$

Figure 3. The focus plane in the acoustic holography imaging system.
is at a maximum. The focus measure at a point \((x, y)\) is obtained by accumulating gradient magnitude estimates over a small region \(W_{x,y}\) around \((x, y)\):

\[
F(x, y) = \sum_{(x', y') \in W_{x,y}} |\nabla g(x', y')|,
\]

where \((x, y)\) is the center pixel of the region \(W_{x,y}\). The Sobel masks (Rosenfeld and Kak, 1976), depicted in Figure 4, are used to estimate the gradient magnitude.

A similar approach, the modulus method, was proposed by Jarvis (1976). The modulus difference is defined as the summation of differences between each pixel and its neighboring pixels in two orthogonal directions:

\[
F(x, y) = \sum_{(x', y') \in W_{x,y}} \left[ |g(x, y) - g(x, y - 1)| + |g(x, y)
- g(x - 1, y)| \right].
\]

Hence, the modulus method approximates the gradient magnitude.

B. Laplacian. The Laplacian operator is a high pass filter and, thus, can be used to describe the high frequency content of an image. Because a focused image contains more high frequency information, Muller and Buffington (1974) proposed the use of the Laplacian as a focus measure. The Laplacian at the pixel \((x, y)\) is given by

\[
\nabla^2 g(x, y) = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}. \tag{16}
\]

In a digital image, many methods can be used to yield an approximate value for the Laplacian operation. The mask most frequently used to compute the Laplacian is shown in Figure 5.

The focus measure at a point \((x, y)\) is obtained by summing the Laplacian over a small region \(W_{x,y}\) around \((x, y)\):

\[
F(x, y) = \sum_{(x', y') \in W_{x,y}} \nabla^2 g(x', y'). \tag{17}
\]

From Eq. (16), we see that the second derivatives in orthogonal directions tend to cancel each other when they have opposite signs. This may cause the Laplacian’s value to be unstable. Nayar (1992) proposed a modified Laplacian, which is defined by

\[
\nabla^2 g(x, y) = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}. \tag{18}
\]

We found that the typical \(3 \times 3\) Laplacian operator in Figure 5 is not sensitive to the change in focusing. To accommodate the variations in size of the image contents, our method uses a variable spacing between pixels to compute the derivatives. This modified Laplacian is given by

\[
\nabla^2 g(x, y) = [2g(x, y) - g(x - k, y) - g(x + k, y)] + [2g(x, y) - g(x, y - k) - g(x, y + k)], \tag{19}
\]

where \(k\) is the variable spacing, and \(k = 2\) is used in our experiments.

C. Gray Level Variance. In the spatial domain, gray levels can be viewed as random variables. The arithmetic mean and the variance are very commonly used for the study of image properties. The mean \(\hat{\mu}_{x,y}\) and the variance \(\hat{\sigma}^2_{x,y}\), at a point \((x, y)\) are calculated from a small region \(W_{x,y}\) around \((x, y)\):

\[
\hat{\mu}_{x,y} = \frac{1}{\|W_{x,y}\|} \sum_{(x', y') \in W_{x,y}} g(x', y') \tag{20}
\]

\[
\hat{\sigma}^2_{x,y} = \frac{1}{\|W_{x,y}\|} \sum_{(x', y') \in W_{x,y}} (g(x', y') - \hat{\mu}_{x,y})^2, \tag{21}
\]

where \(\|W_{x,y}\|\) denotes the cardinality of \(W_{x,y}\) (i.e., the number of pixels in region \(W_{x,y}\)), and the summations are limited to pixels within \(W_{x,y}\). The best focus occurs when the sample variance is at a maximum. Jarvis (1976) used the variance as a measure of focus. The focus measure at a point \((x, y)\) can be defined as the gray level variance at the point

\[
F(x, y) = \hat{\sigma}^2_{x,y}. \tag{22}
\]

D. Choice of Focus Measure. Different approaches for the measurement of focus were discussed above. A technique suitable for one kind of image may not be appropriate for another kind of image. There is no universal technique appropriate for all images, because the properties of images are different. Here, we test the different techniques for our imaging system using a sequence of images of a small bead. The focus measures are applied in a small region around the object. Figure 6 shows two sample image frames in which the object is in focus and out of focus. Focus changes can be subtle, and identification of an image as “in focus” or “out of focus” involves examining (through animation) a sequence of im-

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-1</td>
<td>-2</td>
<td>-1</td>
</tr>
</tbody>
</table>

Figure 4. Sobel convolution masks: (a) sensitive to vertical gradients; (b) sensitive to horizontal gradients.
ages several times and carefully checking the detail of interest. We expect a focus measure to be largest for focused images and to monotonically decrease for out-of-focus images with large depth disparities. The best focus measure should be superior to others in terms of monotonicity about the peak and robustness to image noise.

The plots in Figure 7 show the focus measures for the different methods discussed. To demonstrate the dependency of the measurements on the degree of focusing, the focus measure \( F \) is plotted as a function of the index number shift \( \Delta_n \) with respect to \( N_f \) (where \( N_f \) is the index number of the best focused image). Image frames in front of the focused image correspond to a negative value, and image frames behind the focused image have a positive value along the horizontal axis. Comparisons between the different focus measures show that the gray level variance and the gradient magnitude perform better than the other measures. The gray level variance has the sharpest peak and therefore is chosen as the basis for our focus measure. Because there are more details in a focused image, we know intuitively that a focused image has more variation, so the gray level variance is a reasonable focus measure.

For our imaging system, we observe that the gray level in the region of interest varies as the focus changes. In Figure 8(a), which shows the gray level mean, it is clear that better focusing results in a lower gray level intensity. In situations when the gray level is not constant over the testing region, a variance measure is no longer suitable for detecting the best focus. Thus, to ensure validity of the variance as our focus measure, we normalize it. We choose a simple method, normalizing the original variance with respect to its mean:

\[
F(x, y) = \frac{\sigma^2_{x,y}}{\mu_{x,y}},
\]

(23)

This normalization can be applied either before or after the variance calculation. Figure 8(b) shows the modified gray level variance with respect to the index number using Eq. (23). To make our focus measure more robust to noise, we can also apply a low pass filter before the focus measure procedure. Because we want to test the validity of the modified gray level variance as our focus measure, we do not include this step in our experiments.

IV. IMAGE FOCUS ENHANCEMENT

As discussed earlier, the focusing depth of an acoustic holography imaging system is limited and defocusing occurs. Using the focus measure technique described in Section III, we can find the image frame in which the object of interest is best focused, but we can only find it for a cross section of the object. When an object is not flat, only part of it is exactly focused in a particular frame, while parts not at the same depth are defocused. The degree of defocusing depends on the distance from the plane of focus. In this section, enhancement techniques modeled on restoration processes are developed to remove the focus degradation, and the corresponding inverse processes are applied to recover the original focused image at each pixel. The enhancement procedure is applied on the best focused image selected from an image sequence.

A. Region-Fusing Focus Reconstruction. A common goal of medical imaging analysis is to understand the physical relationship between objects in a focused image. Hence, for acoustic holography, with its limited focusing range, we wish to synthesize a focused image in which every pixel is “best focused.” Because the focus measure technique can be used to find the best focused image, any object can be related to an image frame where it is best focused. Thus, we can identify the focused details from different image
frames and then fuse them together into one new image. The new image contains focused details that were originally focused in different frames. The procedure computes an output image \( g_{out}(x, y) \) from the set of input images and corresponding focus measure images \([g_i(x, y), F_i(x, y)]|i = 1 \ldots N\). For each pixel \((x, y)\),

1. Calculate \( F_k(x, y) \) through \( F_N(x, y) \), the focus measures at all depth index values.
2. Obtain \( k = \arg \max F_k(x, y) \), the index of the best-focused frame at \((x, y)\).
3. Set \( g_{out}(x, y) = F_k(x, y) \).

For our region-fusing focus experiment, a sequence of images is obtained from a phantom developed to study the focal plane characteristics of the sensor. The phantom has nine monofilament strands stretched on a cylindrical frame. Monofilaments are separated by 1 cm in the \( z \) direction and are oriented at different angles in the \( xy \) plane. Figure 9 shows the geometry of the phantom. Because the monofilaments are at different depths, they are not imaged focally in one single image frame. Figure 10 shows an experiment for a sequence of images (10 frames). The relationship between strands is difficult to see, because the original images only have one monofilament best focused. The constructed image is assembled from the focused details of the 10 images. In the constructed image, more strands are focused and their relationship is clearer.

B. Spatial Domain Filtering. The goal of our image enhancement processing is to recover the image focus degraded by the defocusing transform inherent to the imaging system. We want to use a focus recovery procedure to increase the focusing range of the imaging system. In Section IV-A, we implemented a method to reconstruct a focused image by fusing image details that were originally focused in different image frames. The disadvantages of this method are the unsmoothed image region obtained and the noise introduced by the fusing operation. Another approach to the focus recovery problem is to use an image filtering technique that approximates deconvolution. Here we consider two filtering methods—the S-transform method in the spatial domain and the Wiener filter in the frequency domain—and apply them to our situation. The S-transform has the advantage of a lower computational cost. As noted above, the relationship between the untransformed signal and the digitized signal is governed by defocus but in a complex way. This makes the use of the term “restoration” inappropriate.

B.1. S-Transform. Subbarao et al. (1995) proposed a new method called the S-transform, which uses a convolution/deconvolution transform in the spatial domain. This method does not require complete knowledge of the system point spread function (PSF) (it only requires the second moment of the PSF), and it costs very little in terms of computation time. Here, we summarize the formulation of the S-transform.

As discussed in Section II, the degradation problem becomes a defocus problem, which can be modeled approximately as a convolution. We assume \( g(x, y) \) is the defocused image obtained by convolving the focused image \( f(x, y) \) with the system’s PSF \( h(x, y) \)

\[
g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - \xi, y - \eta) h(\xi, \eta) \, d\xi \, d\eta. \tag{24}
\]

Two assumptions of the S-transform are that the system’s PSF is circularly symmetric and that the image signal can be approximated as a bicubic polynomial. The focused image \( f(x, y) \) is assumed to be a bicubic polynomial defined by

\[
f(x, y) = \sum_{m=0}^{3} \sum_{n=0}^{3} a_{m,n} x^m y^n,
\]

where \( a_{m,n} \) are the polynomial coefficients. Using a Taylor series expansion at the point \((\xi, \eta)\), \( f(x, y) \) is expressed as

\[
f(x - \xi, y - \eta) = \sum_{0 \leq m + n \leq 3} \frac{(-\xi)^m (-\eta)^n}{m! n!} f^{m,n}(x, y). \tag{26}
\]

From Eq. (24) we obtain

\[
g(x, y) = \sum_{0 \leq m + n \leq 3} \frac{(-1)^{m+n}}{m! n!} f^{m,n}(x, y) h_{m,n}, \tag{27}
\]

where \( f^{m,n} \) and \( h_{m,n} \) are defined by

\[
f^{m,n}(x, y) = \frac{\partial^n}{\partial x^m \partial y^n} f(x, y), \tag{28}
\]
Thus, with these results, Eq. (27) becomes

\[ f(x, y) = g(x, y) - \frac{h_{2,0}}{2} \nabla^2 g(x, y). \]  

(34)

This equation yields a deconvolution procedure for enhancement. The focused image \( f(x, y) \) is recovered from the blurred image \( g(x, y) \), its derivatives, and the moments of the PSF \( h(x, y) \).

From Eqs. (32) and (34), we see that the second moment of the PSF \( h(x, y) \) determines the degree of defocusing. For this reason, we call \( h_{2,0} \) the defocus parameter. To recover an image using the S-transform technique, we need to find or estimate the value of the defocus parameter \( h_{2,0} \). Once \( h_{2,0} \) is known, the focused image \( f(x, y) \) can be recovered from the blurred image \( g(x, y) \). The defocus parameter \( h_{2,0} \) is the second moment of a circularly symmetric PSF \( h(x, y) \). As discussed in Section II, the Gaussian PSF was chosen as our approximate model for defocusing in the imaging pathway. The second moment of the Gaussian PSF is given by

\[ h_{2,0} = \frac{\sigma^2}{2}, \]  

(35)

where \( \sigma \) is the standard deviation of the 2-D Gaussian PSF. Thus, the quasi-deconvolution we employ becomes

\[ f(x, y) = g(x, y) - \frac{\sigma^2}{4} \nabla^2 g(x, y). \]  

(36)

B.2. Position-Dependent Filtering. Unlike a typical image obtained from a camera system, the image obtained via our imaging system includes 3-D information. As discussed in Section II, the defocusing of an image is different at different positions \((x, y)\). Thus, the PSF is a position-dependent function, and the parameter \( \sigma \) is a position-dependent variable. In contrast, for an optical camera system, it is a constant. This dictates the use of a position-dependent filtering operation.

From Eq. (7), we know that the spread parameter \( \tau \) in the general PSF \( h(x, y, \tau) \) is replaced by \( \sigma \) when the PSF is approximated by a Gaussian. The spread parameter \( \tau \) is related to the image index in a linear relationship described by Eq. (12). For the Gaussian PSF, this equation becomes

\[ \sigma(x, y) = b|N - N_f(x, y)| + b_0, \]  

(37)

where \( b \) and \( b_0 \) are constant parameters, \( N \) is the frame index of the processed image, and \( N_f(x, y) \) is the frame index of the best focused image at the position of interest \((x, y)\). To find \( N_f(x, y) \), the focus detection method described in Section III is used. This procedure is applied at each pixel in the image, and the results are saved in a focus index table, which stores the image index where each pixel is focused. Physically, pixels representing objects at the same depth correspond to the same index \( N_f \). Thus, the filtering operation uses a \( \sigma \) of the same value. At different depths, pixels correspond to different indices \( N_f \) so that different values of \( \sigma \) are used in the filtering operation.

B.3. Flowchart for Position-Dependent Filtering. There are three phases in the position-dependent filtering operation:
1. Construct a focus index table. For each pixel \((x, y)\):

- Choose a small region \(W\) around \((x, y)\).
- Apply the focus measure Eq. (23) to \(W\) for a sequence of images \(I_N\) \((N \in \Omega, \Omega = \{1, 2, \ldots, M\})\)

\[
F_N(x, y) = \frac{\sigma^2 \mu_N}{\sigma_N^2}
\]

where \(\mu_N\) and \(\sigma_N\) are the mean and the variance in the region \(W\) for the image frame \(I_N\).
- Find the index of the image whose focus measure is greatest

\[
N_f(x, y) = \arg \max_{N \in \Omega} \{F_N(x, y)\}.
\]
- Save the index number \(N_f(x, y)\) to a focus index table corresponding to \((x, y)\).

2. Find an initial image.

- Choose a small region \(W\) in the image area of interest.
- Apply the focus measure to \(W\) for a sequence of images.
- Find the best focused image whose focus measure is greatest.
- Use the resulting image as the initial image \(I_N\).

3. Restore the image using S-transform filtering. For each pixel \((x, y)\) of the initial image \(I_N\):

- Determine the local defocus parameter \(\sigma(x, y)\)

\[
\sigma(x, y) = b|N - N_f(x, y)| + b_0,
\]

where \(N\) and \(N_f(x, y)\) are the index of the processing image frame and the value of the focus index table, respectively; \(b\) and \(b_0\) are chosen experimentally to yield the best result.
- Apply the filter Eq. (36)

\[
f(x, y) = g(x, y) - \frac{\sigma^2(x, y)}{4} \nabla^2 g(x, y).
\]

C. Wiener Filter. In the frequency domain, a Wiener filter is commonly used for image restoration. This method requires knowledge of the system PSF, and its computational cost includes two Fourier transforms. Compared to an inverse filter in the frequency domain, the Wiener filter is more robust to noise. The Wiener filtering operation is described by

\[
F(u, v) = G(u, v)W(u, v),
\]

\[
W(u, v) = \frac{1}{H(u, v) + \lambda}\frac{H^2(u, v)}{H^2(u, v) + \lambda},
\]

where \(F(u, v)\) is the Fourier transform of the resulting image \(f(x, y)\), \(G(u, v)\) is the Fourier transform of the defocused image \(g(x, y)\), \(H(u, v)\) is the Fourier transform of the PSF \(h(x, y)\), and \(\lambda\) is the noise-to-signal power density ratio. We assume \(\lambda\) is constant in our imaging system. Its value is chosen experimentally from the range of 0.01–0.1 to yield the best result. The inverse filter is a special case for \(\lambda = 0\).

As discussed in Section II, a Gaussian function is the best approximation for our imaging system PSF, and the standard deviation \(\sigma\) of the Gaussian PSF is position-dependent. However, because Wiener filtering includes the use of two Fourier transforms, a position-dependent deconvolution is impractical in terms of computation time. In practice, we approximate the position-dependent \(\sigma\) by a constant that is chosen empirically. For a constant \(\sigma\) parameter, however, a Wiener filter only gives good results for part of the object. Figures 11 and 12 show results for a monofilament phantom image and a breast tissue image with Wiener filtering using \(\lambda = 0.01\) and three different values of \(\sigma\). We see that it is difficult to find a value of \(\sigma\) that results in good image restoration for all parts of the image. This is additional evidence that the true relationship between the original (unobservable) and digitized signals is complex and not easily modeled precisely.

V. EXPERIMENTS AND EVALUATION

In this section, a set of experiments is implemented to compare results for the different focus recovery methods. A quantitative measure is developed and used to evaluate each focus recovery method. The following three enhancement methods are applied to our test data: Wiener filtering with a constant defocus parameter, spatial filtering with a constant defocus parameter, and spatial filtering with a position-dependent defocus parameter.

The basic idea underlying focus evaluation is that well-focused images contain more information than poorly focused images, so most image processing techniques are evaluated using the signal-to-noise ratio (SNR). The focus measure described in Section III is based on a similar idea, that focused images have higher contrast than defocused images. Using the focus measure, we determine the
focused image of an object of interest from a sequence of images. However, because the focus measure is used to obtain the focused image, it obviously cannot be used to evaluate the result. In addition, because the focus restoration process involves a high-pass filtering operation, it increases the noise level of the image. Therefore, comparisons of the focus measure cannot be used, because a noise-degraded image also results in a larger focus measure. To evaluate the focus recovery methods, we again make use of a sequence of images. For each pixel of the image, we can calculate its focus measure frame by frame. In a manner similar to the construction of the focus index table described in Section IV-B, we then can construct a table storing the focus measure value at each pixel. These values are the focus measure of the best focused image at each pixel of the image. The objective of our focus restoration processing for one image is for the focus measures of the resulting image to be the same values as the ones in this table. Therefore, we can evaluate each focus recovery method by measuring differences between the ideal and actual results.

Let $F$ and $F_f$ correspond to the focus measure of the testing image and the value in the table, respectively; we use the mean square value of their difference to compare the two:

$$MSF = \frac{1}{\|I\|} \sum_{(x,y) \in I} \left[ F(x,y) - F_f(x,y) \right]^2,$$

where $MSF$ stands for the mean square of the focus measure over the image region tested, $\|I\|$ denotes the number of pixels in the image $I$, $F(x,y)$ is the focus measure of the resulting image at pixel $(x,y)$, and $F_f(x,y)$ is the focus measure of the best focused image frame at pixel $(x,y)$.

In the first experiment, the monofilament images are used to test the processing results. The constant parameters used in the filtering operation are chosen experimentally to yield best results. Figure 11 shows the original image and the results using a Wiener filter with $\lambda = 0.01$ and three different values of $\sigma$. Figure 13 shows the results using the S-transform with three different values of $\sigma$ and the position-dependent parameter. For the method of spatial filtering with a position-dependent defocus parameter, the focus index table is built from a sequence of images (10 frames), which includes the processing image frame. The constants $b$ and $b_0$ in Eq. (37) are 0.01 and 0.0, respectively. In the second experiment, we use a sequence of images of breast tissue, which includes one lesion. The sequence of images (10 frames) has depth intervals of approximately 1 mm. The constant parameters used in the filtering operation are chosen experimentally to yield the best results. Figure 12 shows the original image and the results using a Wiener filter with $\lambda = 0.1$ and three different values of $\sigma$. Figure 14 shows the results using the S-transform with three different values of $\sigma$ and the position-dependent parameter. In the position-dependent S-transform method, the position-dependent defocus parameter $\sigma$ is determined from Eq. (37) with $b = 0.5$ and $b_0 = 0$.

Tables I and II show the focus evaluation results using the $MSF$ defined previously. The $MSF$ only describes how close the focus measure of the resulting image is to the estimated value $F_f$. For the second experiment, the speckle-like images introduce an obvious effect on the focus measure and result in large values of the $MSF$. However, comparisons between different restoration methods are

![Figure 12. Breast tissue: (a) the original image; (b) the result of Wiener filtering with a constant $\sigma = 1$; (c) the result of Wiener filtering with a constant $\sigma = 2$; (d) the result of Wiener filtering with a constant $\sigma = 3$.](https://example.com/figure12)

![Figure 13. Monofilament phantom: (a) the S-transform with a constant $\sigma = 1.0$; (b) the S-transform with a constant $\sigma = 2.0$; (c) the S-transform with a constant $\sigma = 3.0$; (d) the S-transform with a position-dependent $\sigma$.](https://example.com/figure13)
still valid. From the results of both experiments, we see that the performance of the S-transform with a constant defocus parameter is competitive with the result of the Wiener filter, but it has the advantage of lower computational cost. For both methods, the constant defocus parameter is chosen experimentally. As discussed previously, because it is difficult to find a value that improves the entire image, it is selected on the basis of the best restoration in the region of interest. The necessity of choosing the defocus parameter reflects the limitation of constant defocus parameter restoration methods. The result of the position-dependent S-transform method shows the best performance. The out-of-focus details are improved but the focused details remain unchanged. As discussed previously, the degradation of image focusing for our imaging system is position-dependent. Hence, a variable defocus parameter is desired in the focus restoration procedure.

Table II. MSF measurements for monofilament images.

<table>
<thead>
<tr>
<th>Image</th>
<th>Algorithm</th>
<th>$\sigma$</th>
<th>MSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 11(a)</td>
<td>None</td>
<td></td>
<td>0.74</td>
</tr>
<tr>
<td>Fig. 11(b)</td>
<td>Wiener</td>
<td>1.0</td>
<td>0.78</td>
</tr>
<tr>
<td>Fig. 11(c)</td>
<td>Wiener</td>
<td>2.0</td>
<td>0.91</td>
</tr>
<tr>
<td>Fig. 11(d)</td>
<td>Wiener</td>
<td>3.0</td>
<td>2.95</td>
</tr>
<tr>
<td>Fig. 13(a)</td>
<td>S-transform</td>
<td>1.0</td>
<td>0.57</td>
</tr>
<tr>
<td>Fig. 13(b)</td>
<td>S-transform</td>
<td>2.0</td>
<td>1.45</td>
</tr>
<tr>
<td>Fig. 13(c)</td>
<td>S-transform</td>
<td>3.0</td>
<td>19.1</td>
</tr>
<tr>
<td>Fig. 13(d)</td>
<td>S-transform</td>
<td>Variable</td>
<td>0.41</td>
</tr>
<tr>
<td>Fig. 10(d)</td>
<td>Pixel-fuse</td>
<td></td>
<td>0.87</td>
</tr>
</tbody>
</table>

VI. CONCLUSIONS AND FUTURE WORK

We have presented postprocessing algorithms for an acoustic holography imaging system. It was shown that a focus measure technique can be used to find the best focused image for the object of interest. Because an acoustic holography image contains 3-D information about the object, position-dependent defocusing occurs. Thus, a position-dependent focus restoration technique is needed. Position dependence was realized by setting the focus spread parameter of the point spread function (PSF) with respect to its location at different pixels of the image, and several filtering techniques were examined for focus recovery. It was found that the position-dependent spatial filtering method yields the best results at a low computational cost. Some of the remaining questions to be addressed in future research include the following:

1. In our proposed method, a linear relationship is assumed for the focus depth and the index of the image sequence. Further calibration work is needed to configure the system operation to realize the assumed relationship.
2. In this work, we treated only one parameter of the PSF dependent on the position of a pixel. For a 2-D Gaussian PSF, the parameter $\sigma$ is assumed to be linear with respect to the image index. In the future, we will use a more complex PSF such as a generalized Gaussian model for which more parameters are dependent on the position.

ACKNOWLEDGMENTS

This work was supported by the Washington Technology Center, by Advanced Diagnostics, Inc., and by the Carl M. Hansen Foundation.

REFERENCES


Table I. MSF measurements for breast tissue images.

<table>
<thead>
<tr>
<th>Image</th>
<th>Algorithm</th>
<th>$\sigma$</th>
<th>MSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 12(a)</td>
<td>None</td>
<td></td>
<td>136</td>
</tr>
<tr>
<td>Fig. 12(b)</td>
<td>Wiener</td>
<td>1.0</td>
<td>195</td>
</tr>
<tr>
<td>Fig. 12(c)</td>
<td>Wiener</td>
<td>2.0</td>
<td>191</td>
</tr>
<tr>
<td>Fig. 12(d)</td>
<td>Wiener</td>
<td>3.0</td>
<td>198</td>
</tr>
<tr>
<td>Fig. 14(a)</td>
<td>S-transform</td>
<td>1.0</td>
<td>216</td>
</tr>
<tr>
<td>Fig. 14(b)</td>
<td>S-transform</td>
<td>2.0</td>
<td>206</td>
</tr>
<tr>
<td>Fig. 14(c)</td>
<td>S-transform</td>
<td>3.0</td>
<td>183</td>
</tr>
<tr>
<td>Fig. 14(d)</td>
<td>S-transform</td>
<td>Variable</td>
<td>128</td>
</tr>
</tbody>
</table>


