Speckle reduction by I-divergence regularization in optical coherence tomography

Daniel L. Marks, Tyler S. Ralston, and Stephen A. Boppart

Beckman Institute of Advanced Science and Technology, 405 N. Mathews, Urbana, Illinois 61801

Received January 3, 2005; accepted March 31, 2005

For optical coherence tomography (OCT), ultrasound, synthetic-aperture radar, and other coherent ranging methods, speckle can cause spurious detail that detracts from the utility of the image. It is a problem inherent to imaging densely scattering objects with limited bandwidth. Using a method of regularization by minimizing Csiszar's I-divergence measure, we derive a method of speckle minimization that produces an image that both is consistent with the known data and extrapolates additional detail based on constraints on the magnitude of the image. This method is demonstrated on a test image and on an OCT image of a *Xenopus laevis* tadpole.
© 2005 Optical Society of America
OCIS codes: 100.3010, 110.4500.

1. INTRODUCTION

The problem of speckle[1] is endemic to imaging methods utilizing coherent ranging such as optical coherence tomography[2] (OCT), ultrasound, and synthetic-aperture radar. Natural objects usually contain detail at all scales, but most coherent ranging instruments can probe these details in only a narrow bandwidth of frequencies. The result is speckle, which is caused by the seemingly random interference of scatterers within a resolution cell of the instrument. Speckle randomizes the amplitudes of points in the image, which obscures fine features but generally allows coarse features to be resolved. We propose and demonstrate a method of regularization that removes these small-scale amplitude changes without smoothing the image. This is achieved by constraining the resulting reconstructed image to be consistent in a least-squares sense with the known data, while simultaneously utilizing Csiszar's I-divergence measure[3,4] to constrain and regularize the amplitudes of the reconstructed data. By ensuring that the resulting image is consistent with the known data, the detail is retained, while additional bandwidth is extrapolated to make the amplitudes of the image consistent with the regularization.

Most natural objects, especially those biological in origin, have detail at length scales from the macroscopic to the atomic. Unfortunately, because of instrument limitations and the attenutation of radiation by the sample, only a finite signal bandwidth can be utilized for useful resolution from the sample. The phenomenon of speckle in these imaging modalities is caused by the interference between scatterers that are too small and too close together to be individually resolved. What is observed is a seemingly random modulation of the amplitude of the image as a result of the interference of scattered waves from these unresolved scatterers.

Unfortunately, it is difficult in practice to differentiate between the features that are caused by the speckle modulation and the features of interest. Because scatterers are randomly placed in most natural objects, the coarse detail tends to be present in almost any frequency bandwidth that is much wider than the reciprocal of the minimum feature size one wishes to resolve. This is what enables most coherent ranging techniques to be used in practice. However, as one considers distinguishing features on the minimally resolvable scale of the instrument, it is less certain whether a certain feature is due to the random modulation of speckle or to scatterers with larger sizes.

The inspiration for our solution to this problem came from studying an expectation-maximization (EM) algorithm[5] for despeckling from synthetic-aperture radar in the presence of white Gaussian noise. In this algorithm, it is postulated that each pixel i of the object contains scatterers that have a scattering amplitude described by a complex Gaussian random variable with variance σ_i^2, so that the phase of each pixel is uniformly distributed between $-\pi$ and π. Due to speckle, the scattering from each resolution cell will be randomly modulated, because the complex amplitudes of the pixels within the resolution element of the instrument will coherently interfere. The phases of the pixels are not themselves of interest but are estimated to better determine the pixel variances σ_i^2. For a given linear relationship between the object properties and the measured data, the EM algorithm specifies an iteration that successively estimates the σ_i^2 that maximizes the likelihood of observing the data. Unfortunately, we found this method unsuitable for our purposes. First, because it is maximum likelihood,[6] it does not specify priors to the image; therefore it is not regularized and so tends to produce estimates consisting of pointlike scatterers. Second, it tends to be difficult to incorporate all but the simplest priors in the algorithm. Finally, it is computationally intractable for all but the smallest image sizes because of the need to invert the empirical covariance matrix at each iteration.

The problem of speckle is sufficiently pervasive that many approaches to minimize it have been put forth: signal processing methods such as using maximum

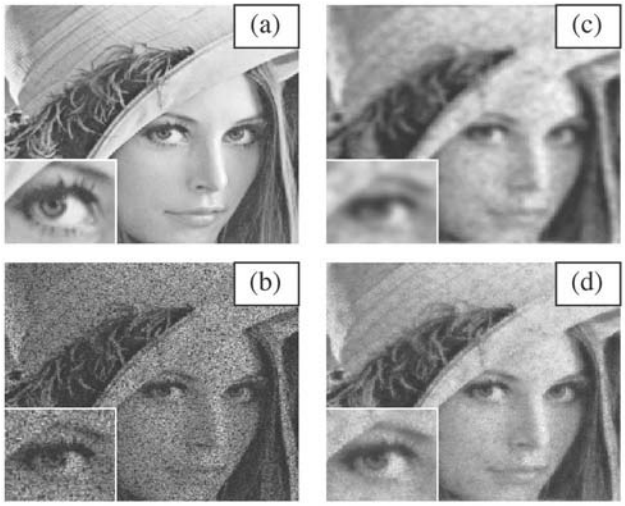

Fig. 1. Images of Lena to demonstrate a naïve Gaussian deblurring algorithm. (a) Original Lena image. (b) Speckled Lena image reconstructed with a Tikhonov-regularized least-squares algorithm. (c) Amplitude of the image of (b) blurred with a Gaussian window as in Eq. (5). (d) Image corrected with the I-divergence algorithm. Inset in each part, magnified image of Lena's right eye.

likelihood,[5,7] nonquadratic regularization,[8] deconvolution,[9] spectral estimation,[10] adaptive speckle suppression filters,[11] wavelet denoising,[12,13] the CLEAN[14] algorithm, and many others.[15,16] Other approaches such as angular or frequency compounding[17,18] have also been employed. Some of these methods suffer from the problem that increased speckle reduction entails a loss of resolution. Our research was motivated by the desire for a method that would produce a result consistent with the data, minimize resolution loss, and have a reasonable computation time.

To understand the rationale behind this method, we consider a primitive despeckling algorithm based on smoothing the amplitude of an image. We have created a speckled version of the ubiquitous Lena image, which is shown in Fig. 1, by multiplying the pixel magnitudes of the original image by a random complex phase, and then bandpassing the image with a Gaussian filter. The effect closely mimics the speckle seen in natural objects. By blurring the magnitude of the complex image with a Gaussian filter, we can produce an image with the speckle minimized, but with much fine detail also removed. In addition to being low resolution, there is no guarantee that this blurred version is consistent with the data. Consistency means that there exists a true reconstructed image of complex scattering amplitudes such that the the magnitudes of the complex data match the blurred version while the complex amplitudes match the measured data to within measurement error (perhaps using a least-squares metric). In general this blurring method will not produce a consistently despeckled image. It is because of this inconsistency that the method fails to incorporate the detail of the original data. Our method essentially produces a synthesized complex amplitude image that constrains the complex data to be consistent with the measured data in a least-squares sense, but also manipulates the magnitudes of the complex data to the degree allowed by measurement error. By doing so, this method can despeckle but retain the detail of the original data.

2. ALGORITHM

Scattering problems in coherent ranging using the single-scattering assumption can be modeled as linear systems. In general, we assume we have an object that is described by vector $\mathbf{x}_i$, and a data vector described by $\mathbf{y}_j$. The forward problem is then simply modeled as $\mathbf{y}=\mathbf{A}\mathbf{x}+\mathbf{n}$ or $\mathbf{y}_j=\Sigma_i\mathbf{A}_{ij}\mathbf{x}_i+\mathbf{n}_j$, where $\mathbf{A}$ describes the linear transformation between object and data vectors, and $\mathbf{n}$ is a noise vector. In the absence of regularization, the least-squares solution for $\mathbf{x}$ minimizes $\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2$ and has the pseudoinverse solution $\mathbf{x}=(\mathbf{A}^\dagger\mathbf{A})^{-1}\mathbf{A}^\dagger\mathbf{y}$.

To despeckle the image in a consistent manner, we posit that we have created a despeckled image that is in general inconsistent with the data (for example, with the Gaussian blurring scheme already mentioned). For each object pixel $\mathbf{x}_i$ we have an inconsistently despeckled reference image with the magnitudes of each corresponding pixel in the reference image given by $|\mathbf{u}_i|$. The algorithm finds a consistently despeckled image $\mathbf{x}_i$ that minimizes the functional

$$L=\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2+\lambda\sum_i\left[|\mathbf{x}_i|^2\log\left|\frac{\mathbf{x}_i}{\mathbf{u}_i}\right|^2-|\mathbf{x}_i|^2+|\mathbf{u}_i|^2\right]. \quad (1)$$

The term $\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2$ ensures the image is consistent with the data in a least-squares sense. The second term is a metric called the Csiszar I-divergence[3] and is weighted by a regularization constant λ. The I-divergence is a convex measure that is related to the Kullback[19] distance, which is itself a relative entropy measure related to the maximum-entropy method.[20–24] The I-divergence is analogous to the Euclidean metric, but it is more suited to positively constrained data. The I-divergence is preferred over the Kullback distance because while both are zero when $|\mathbf{x}_i|=|\mathbf{u}_i|$, the I-divergence is also minimized, unlike the Kullback distance which is minimized at $|\mathbf{x}_i|=|\mathbf{u}_i/\sqrt{e}|$. Therefore, the I-divergence tends to enforce the condition that the magnitude of the consistently despeckled image matches the magnitude of the reference despeckled image.

To see how to optimize this function, we find the partial derivative $\partial L/\partial\mathbf{x}_i$ with respect to both the real and imaginary parts of $\mathbf{x}_i$ (called the compact gradient in Ref. 8), which is given by

$$\begin{aligned}\frac{\partial L}{\partial\mathbf{x}_i}&=\frac{\partial L}{\partial\,\mathrm{Re}\{\mathbf{x}_i\}}\frac{\partial\,\mathrm{Re}\{\mathbf{x}_i\}}{\partial\mathbf{x}_i}+\frac{\partial L}{\partial\,\mathrm{Im}\{\mathbf{x}_i\}}\frac{\partial\,\mathrm{Im}\{\mathbf{x}_i\}}{\partial\mathbf{x}_i}\\&=\frac{\partial L}{\partial\,\mathrm{Re}\{\mathbf{x}_i\}}-i\frac{\partial L}{\partial\,\mathrm{Im}\{\mathbf{x}_i\}}\\&=2\{\mathbf{A}^\dagger\mathbf{A}\mathbf{x}-\mathbf{A}^\dagger\mathbf{y}\}_i+4\lambda\mathbf{x}_i\log\left|\frac{\mathbf{x}_i}{\mathbf{u}_i}\right|, \quad (2)\end{aligned}$$

where $\{\mathbf{A}^\dagger\mathbf{A}\mathbf{x}-\mathbf{A}^\dagger\mathbf{y}\}_i$ refers to the ith element of the vector. Using this derivative we can form a simple relaxation method iteration:

$$\mathbf{x}_i^{(n+1)} = \mathbf{x}_i^{(n)} - 2\epsilon\left[\{\mathbf{A}^\dagger\mathbf{A}\mathbf{x}^{(n)} - \mathbf{A}^\dagger\mathbf{y}\}_i + 2\lambda\mathbf{x}_i^{(n)}\log\left|\frac{\mathbf{x}_i^{(n)}}{\mathbf{u}_i}\right|\right]. \tag{3}$$

Note that to make this a stable iteration, the magnitude of ϵ must be less than both $1/2\lambda$ and the reciprocal of the largest singular value of $\mathbf{A}$ squared. While a relaxation algorithm is used here, a multiplicative type algorithm[25] could also be derived.

To see what values of λ will work well, we examine the related optimization problem

$$L = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda\sum_i |\mathbf{x}_i|^2 = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda\|\mathbf{x}\|^2, \tag{4}$$

which is the least-squares solution to $\mathbf{y}=\mathbf{A}\mathbf{x}$ under Tikhonov[26] regularization constraining the magnitude of the solution $\mathbf{x}$. This simple regularization is done in practice to ensure that the vector space of $\mathbf{A}$ corresponding to small singular values does not give rise to large changes in $\mathbf{x}$. The solution is the Tikhonov-regularized pseudoinverse given by $\mathbf{x}_{\text{tik}}=(\mathbf{A}^\dagger\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^\dagger\mathbf{y}$. In practice, we have found the criterion for the selection of λ in this Tikhonov regularization case and the I-divergence regularization case to be similar, i.e., the same value of λ tends to produce good noise performance for both. A useful value of λ can usually be determined by examining the singular value spectrum of $\mathbf{A}$ and determining where it decreases sharply, setting λ equal to the square of the transition singular value.[26] More sophisticated methods assign λ based on the magnitude of the noise relative to the acquired signal. These methods show that the same order of magnitude of λ suffices for both regularizers because $|\mathbf{x}_i|^2\log|\mathbf{x}_i/\mathbf{u}_i|^2 - |\mathbf{x}_i|^2 = |\mathbf{x}_i|^2\log|\mathbf{x}_i/\mathbf{u}_i\sqrt{e}|^2$. Since the logarithm tends to be a small number, on the order of unity, the magnitude of the the regularization terms in both methods tends to be similar.

3. DESPECKLING AN IMAGE

To address the issue of despeckling images, we restrict ourselves to the case where the matrix $\mathbf{A}$ implements a space-invariant convolution. In this case, we can describe $\mathbf{A}$ in the Fourier domain by a transfer function $\tilde{A}(\boldsymbol{\nu})$ with a finite bandpass, where $\boldsymbol{\nu}$ is spatial frequency complementary to spatial coordinate $\mathbf{r}$. We can then interpret the algorithm, construing $\tilde{A}(\boldsymbol{\nu})$ to be the spatial bandpass of coherent ranging instruments such as OCT. In the functional of Eq. (1), the least-squares term fits the data only inside the bandwidth of the instrument. In space-invariant systems, the least-squares solution does not infer any frequencies outside of the bandpass of the instrument. The I-divergence term of Eq. (1), however, constrains the magnitudes of the image points, and thus can infer frequencies outside the bandpass. To simultaneously minimize both terms, the I-divergence term would affect mainly frequencies outside the bandpass, while the image inside the bandpass would be determined by the least-squares minimization term. In this way, this despeckle technique can be seen as a way of bandwidth "extrapolation" that uses additional bandwidth to minimize the speckle, while keeping the image consistent with the data. Without additional bandwidth, the algorithm would have little freedom to optimize the image and keep it consistent with the data.

An important question then is the amount of additional bandwidth needed to extrapolate detail. This issue is common to such algorithms, especially maximum-likelihood algorithms. While there are technically no priors on the image in this case, the practical choice of the sampling rate of the reconstructed image is itself a prior. If too much bandwidth is added, the data will underspecify the image, and the image reconstruction commonly will become overly sharpened and sensitive to noise. We have found empirically that the amount of bandwidth required to allow for speckle reduction is 1.4 to 2 times the original instrument bandwidth, but this is a question that will benefit from further exploration.

The most important question with this method is how to choose the reference image $|u(\mathbf{r})|$. This is the most subjective aspect of the algorithm, because the algorithm only attempts to guarantee that the image $x(\mathbf{r})$ is consistent with the data while keeping $|x(\mathbf{r})| \approx |u(\mathbf{r})|$, but this does not necessarily result in an aesthetically despeckled image unless $|u(\mathbf{r})|$ is chosen properly. Because the algorithm will use consistency to keep the detail in the image, the reference image $|u(\mathbf{r})|$ need not contain all of the detail in the original image. Choosing the reference image to have smoothed magnitudes relative to the original image tends to enforce a similar condition in the despeckled image.

There are a few properties that a reference image should have to produce good results. Of course, the $|u(\mathbf{r})|$ must be positive. In addition, the values of $|u(\mathbf{r})|$ should be similar to the magnitudes of a nondespeckled image. In practice, this can be achieved by basing the reference despeckled image on the speckled Tikhonov-regularized inverse solution $\tilde{x}_{\text{Tik}}(\boldsymbol{\nu}) = \tilde{y}(\boldsymbol{\nu})\tilde{A}(\boldsymbol{\nu})^*/(|\tilde{A}(\boldsymbol{\nu})|^2 + \lambda)$. To create a simple reference image $|u(\mathbf{r})|$, we can convolve the magnitude of the Tikhonov-regularized solution with a Gaussian kernel:

$$|u(\mathbf{r})|^\gamma = (\pi w^2)^{-N/2}\int_{-\infty}^{\infty} |\eta x_{\text{Tik}}(\mathbf{r}')|^\gamma \exp\left(-\frac{|\mathbf{r}-\mathbf{r}'|^2}{w^2}\right) \mathrm{d}^N r'. \tag{5}$$

The magnitude of the Gaussian is chosen such that the total integral of the Gaussian is 1, with N being the number of dimensions in $\mathbf{r}$. This is done to ensure that $\int|u(\mathbf{r})|^\gamma \mathrm{d}^N r = \int|x_{\text{Tik}}(\mathbf{r})|^\gamma \mathrm{d}^N r$ when $\eta=1$, so that the magnitude of the image $|u(\mathbf{r})|$ roughly matches that of $|x_{\text{tik}}(\mathbf{r})|$. The width of the Gaussian kernel w is chosen to be a small multiple of the resolution of the instrument, so that in practice the smoothing window covers several speckles. The exponent γ controls the relative weighting of low and high magnitude values. We have chosen $\gamma=1$ for simplicity and performance here, but others may be chosen, e.g., $\gamma=2$ has the benefit of preserving image energy. The additional constant factor η is intended to compensate for the additional energy that bandwidth extrapolation will add to the image, so that the higher the amount of bandwidth that is to be extrapolated, the higher this number

will need to be. Sometimes if this factor is not present, the I-divergence regularization will attempt to constrain the values of $|x(\mathbf{r})|$ to be too small to be consistent with the additional bandwidth added. Usually $1.0<\eta<1.5$ produces good results. Other methods of producing an inconsistently despeckled image, such as wavelet speckle denoising[27] on $|x_{\mathrm{Tik}}(\mathbf{r})|$, may result in a reference image for $|u(\mathbf{r})|$ that provides better results than the simple Gaussian blurring performed here.

Figure 1(d) shows how this algorithm performs on the speckled Lena data. The parameters here were $\lambda=0.03$, $\eta=1.2$, $\gamma=1$, and $w=3$ pixels, with the largest singular value of the bandpass normalized to 1. As can be seen, the speckle is substantially reduced. The despeckled image of Lena keeps the detail that is seen in the speckled version of part (b), which corresponds to $|x_{\mathrm{tik}}(\mathbf{r})|$, while incorporating the smoother magnitude of part (c) which corresponds to $|u(\mathbf{r})|$ calculated by Eq. (5). The despeckled version is a compromise between these two versions.

4. EXPERIMENTAL DEMONSTRATION

As a demonstration of this technique on a biological sample, we imaged a *Xenopus laevis* tadpole using OCT. Our OCT system is a time-domain, fiber-optic-based Michelson interferometer with the source being a mode-locked Ti:sapphire laser with 810 nm center wavelength and 100-nm bandwidth to achieve approximately 3-μm axial resolution. A sample power of 10 mW was measured. The delay mechanism consists of a galvanometer-driven retroreflector. The laser beam is scanned by galvanometer-driven mirrors and focused with a 20-mm-focal-length near-infrared achromatic lens, achieving a focused spot size of 15 μm and a depth of field of 400 μm. The scanned area was 2 mm in the transverse direction by 1.5 mm in the depth direction. For improved imaging fidelity, the tadpole was placed in water under a glass coverslip to index-match the medium.

To apply the algorithm, we directly digitized the cross-correlation interferogram between the sample and reference beams as the delay was scanned. One thousand axial scans of 80,000 samples per scan were acquired. To reduce the quantity of data, the portion of each axial scan corresponding to delays that contained the tadpole was extracted, and then each axial scan was bandpass filtered and decimated using the one-dimensional fast Fourier transform (FFT) so there were only 1000 decimated samples per axial scan. During decimation, the negative frequency components were zeroed to form the complex analytic signal[28] (CAS) version of the interferogram, because the original signal was real valued. The bandwidth of the extracted spectrum was approximately 1.5 times the original source spectral width to allow for additional extrapolated spectrum to be added for despeckling regularization. The power spectral density of the laser source including the additive noise was estimated by averaging together the magnitude of the Fourier transform of all of the axial scans. The additive noise floor was subtracted from this estimate (with zero as a minimum power spectral density) to produce the spectrum $\tilde{A}(\boldsymbol{\nu})$ used in the algorithm. The highest magnitude frequency, corresponding to the highest singular value, was normalized to 1, therefore scaling only the overall magnitude of the reconstructed image.

A regularization value of $\lambda=0.03$ was chosen based on examination of the spectrum and the noise floor. The Tikhonov-regularized image was then calculated in the frequency domain using $\tilde{x}_{\mathrm{Tik}}(\boldsymbol{\nu})=\tilde{y}(\boldsymbol{\nu})\tilde{A}(\boldsymbol{\nu})^*/(|\tilde{A}(\boldsymbol{\nu})|^2+\lambda)$. The magnitude of the CAS was used to find the magnitudes of $|x_{\mathrm{Tik}}(\mathbf{r})|$ and $|x(\mathbf{r})|$ because it followed the envelope of the interferogram[29] and is independent of the center frequency of the signal (this can be found by assuming all of the negative frequency components of a signal are zero). Based on this, a Gaussian-blurred reference image $|u(\mathbf{r})|$ was formed using Eq. (5) with $w=3$, $\gamma=1$, and $\eta=1.2$. The two-dimensional discrete convolution was performed by using the FFT on the entire image.

To implement the despeckling algorithm, the iteration of Eq. (3) was used, with $\epsilon=0.75$ and $\mathbf{x}^{(0)}=\mathbf{x}_{\mathrm{Tik}}$. The convolution in the operation $\mathbf{A}^\dagger\mathbf{y}$ was precalculated using the 1D FFT by applying the spectrum $\tilde{A}(\boldsymbol{\nu})^*$ to each axial scan datum $\tilde{y}(\boldsymbol{\nu})$. The convolution in the operation $\mathbf{A}^\dagger\mathbf{A}\mathbf{x}^{(n)}$ was performed by applying $|\tilde{A}(\boldsymbol{\nu})|^2$ to each axial scan $\tilde{x}(\boldsymbol{\nu})$. The iteration was repeated 200 times to ensure convergence of the norm $\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2$ to less than 0.1% of its initial value for $\mathbf{x}^{(0)}$. For values of ϵ that satisfy the convergence criterion, the value of $\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2$ monotonically decreased to near zero without stagnation, suggesting that the iteration produces a solution consistent with the data.

Figure 2 shows the result of the experimental despeckling. Part (a) shows the CAS magnitude of the interferogram acquired with OCT. Part (b) shows the Tikhonov-regularized image $|x_{\mathrm{Tik}}(\mathbf{r})|$, and part (c) shows the reference image $|u(\mathbf{r})|$. Part (d) shows the results of the despeckling algorithm. At this scale, one can tell there is a reduction of speckle from the original data and the Tikhonov regularized version, but the scale is too coarse to see the difference from the reference image. For this reason, Fig. 3 and Fig. 4 show two magnified regions in

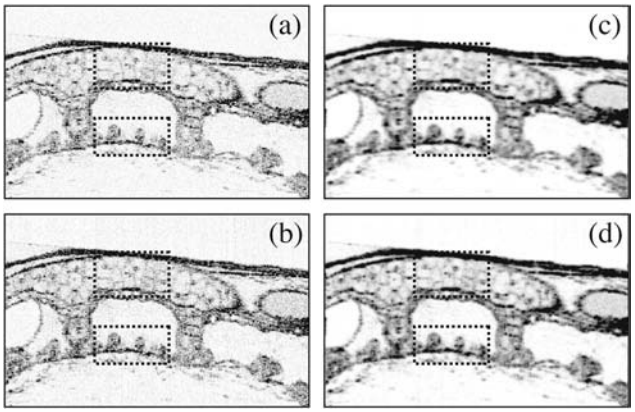

Fig. 2. OCT images of *Xenopus laevis* tadpole processed in various ways. The image size is 1000 μm transverse by 600 μm axial. (a) Original, unprocessed amplitude data. (b) Tikhonov-regularized least-squares solution for the image. (c) Reference image that is the Gaussian blurred amplitude of (b). (d) Despeckled using I-divergence minimization from the reference image. The upper dotted box in each image corresponds to the magnified area in Fig. 3; the lower dotted box in each image corresponds to the magnified area in Fig. 4.

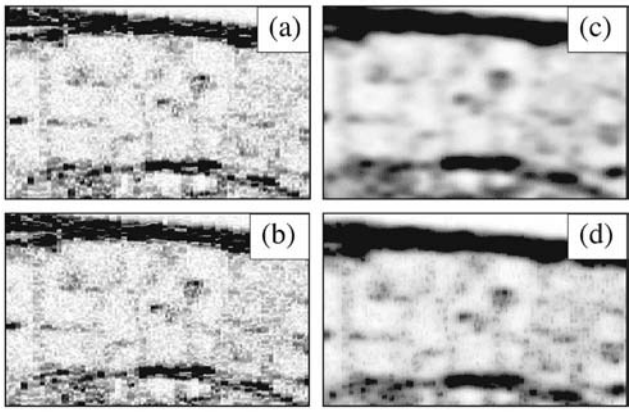

Fig. 3. Magnified OCT images of a *Xenopus laevis* tadpole that correspond to the upper dotted box in each image of Fig. 2. The image size is 240 μm transverse by 150 μm axial. (a) Original, unprocessed amplitude data. (b) Tikhonov-regularized least-squares solution for the image. (c) Reference image that is the Gaussian blurred amplitude of (b). (d) Despeckled using I-divergence minimization from the reference image. Note how the image in (d) retains the detail of the least-squares solution of (b) but includes the smoothed amplitude of the reference image of (c).

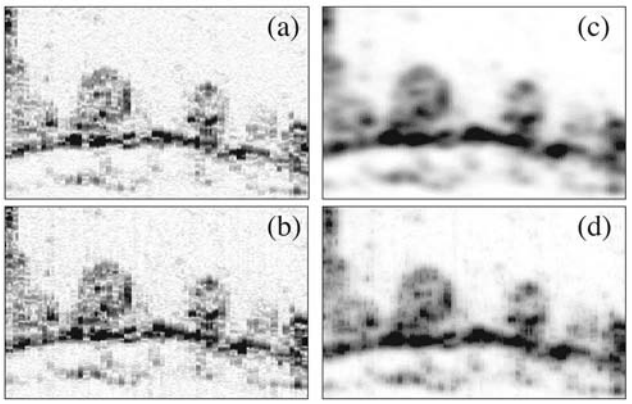

Fig. 4. Magnified OCT images of a *Xenopus laevis* tadpole that correspond to the lower dotted box in each image of Fig. 2. The image size is 240 μm transverse by 130 μm axial. (a) Original, unprocessed amplitude data. (b) Tikhonov-regularized least-squares solution for the image. (c) Reference image that is the Gaussian blurred amplitude of (b). (d) Despeckled using I-divergence minimization from the reference image. The uniform intensity areas of the reference image of (c) are "filled in" with the detail of the least-squares solution of (b) to form the despeckled image. Consistency helps to ensure that the despeckled image retains the detail provided by the data.

this image that correspond, respectively, to the upper and lower dotted boxes in each image. The four images in Figs. 3 and 4 correspond to the same respective images in Fig. 2. In Fig. 3, one can tell that the despeckled image is able to denoise the speckle in the background, but maintain the sharp boundaries between cells, unlike in the reference image. In addition, the cell nuclei are much sharper than in the reference image and look similar to those in the Tikhonov-regularized image, but have greatly reduced speckle. Fig. 4 shows another example of the algorithm eliminating speckle but preserving features. Many of the pointlike scatterers in the solid regions are blurred in the reference image, but are recovered in the despeckled image. This is maintained while the speckled background is smoothed out. The algorithm is able to keep the details consistent with the data while finding a solution with a more uniform background.

We have proposed, derived, and experimentally demonstrated a novel method of minimizing speckle based on regularizing the image with an I-divergence criterion. We believe this algorithm has the potential to enhance the ability of a human operator to recognize features. Speckle features that are consistent with a uniform intensity scattering background tend to be removed, leaving the features that are more likely to be from dense scattering. In this way, one perhaps may better differentiate strong scatterers from other points of large magnitude that are due to random interference between weak scatterers. Because of the relatively few parameters to choose (λ, η, w, and γ), it seems unlikely that the despeckled image can be overspecified from the parameters alone, and features that are seen are likely to correspond to real scatterers.

ACKNOWLEDGMENTS

We thank Ron Stack for his technical contributions to our OCT system and Wei Luo for her assistance in the care and handling of our animal model. The animals used in this study were handled and cared for under the approved protocols of the Institutional Animal Care and Use Committee (IACUC) at the University of Illinois at Urbana-Champaign. This research was supported in part by grants to S. A. Boppart from the National Institutes of Health (1 R01 EB00108-1) and the University of Illinois Intercampus Research Initiative in Biotechnology and to T. Ralston by the Beckman Institute Graduate Fellowship Program.

Stephen Boppart may be reached by e-mail at boppart@uiuc.edu.

REFERENCES

1. J. Goodman, *Statistical Optics* (Wiley, 1985).
2. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," Science **254**, 1178–1181 (1991).
3. I. Csiszar, "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," Ann. Stat. **19**, 2032–2066 (1991).
4. C. L. Byrne, "Iterative image reconstruction algorithms based on cross-entropy minimization," IEEE Trans. Image Process. **2**, 96–103 (1992).
5. D. L. Snyder, J. A. O'Sullivan, and M. I. Miller, "The use of maximum likelihood estimation in forming images of diffuse radar targets from delay-doppler data," IEEE Trans. Inf. Theory **35**, 536–548 (1989).
6. M. Bertero and P. Bocacci, *Introduction to Inverse Problems in Imaging* (IOP, 1998).
7. P. Moulin, J. A. O'Sullivan, and D. L. Synder, "A method of sieves for multiresolution spectrum estimation and radar Imaging," IEEE Trans. Inf. Theory **38**, 801–813 (1992).
8. M. Cetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," IEEE Trans. Image Process. **10**, 623–631 (2001).
9. M. D. Kulkarni, C. W. Thomas, and J. A. Izatt, "Image enhancement in optical coherence tomography using deconvolution," Electron. Lett. **33**, 1365–1367 (1997).
10. S. R. DeGraaf, "SAR imaging via modern 2-D spectral

estimation methods," IEEE Trans. Image Process. **7**, 729–761 (1998).
11. J. Rogowska and M. E. Brezinski, "Evaluation of the adaptive speckle suppresion filter for coronary optical coherence tomography imaging," IEEE Trans. Med. Imaging **14**, 1261–1266 (2000).
12. X. Hao, S. Gao, and X. Gao, "A novel multiscale nonlinear thresholding method for ultrasonic speckle supressing," IEEE Trans. Med. Imaging **18**, 787–794 (1999).
13. D. C. Adler, T. H. Ko, and J. G. Fujimoto, "Speckle reduction in optical coherence tomography by use of a spatially adaptive wavelet filter," Opt. Lett. **29**, 2878–2880 (2004).
14. J. M. Schmitt, "Restoration of optical coherence images of living tissue using the CLEAN algorithm," J. Biomed. Opt. **3**, 66–75 (1998).
15. M. Bashkansky and J. Reintjes, "Statistics and reduction of speckle in optical coherence tomography," Opt. Lett. **25**, 545–547 (2000).
16. K. M. Yung, S. L. Lee, and J. M. Schmitt, "Phase-domain processing of optical coherence tomography images," J. Biomed. Opt. **4**, 125–136 (1999).
17. J. M. Schmitt, "Array detection for speckle reduction in optical coherence microscopy," Phys. Med. Biol. **42**, 1427–1439 (1997).
18. J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," J. Biomed. Opt. **4**, 95–105 (1999).
19. T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, 1991).
20. X. Zhuang, R. M. Haralick, and Y. Zhao, "Maximum entropy image restoration," IEEE Trans. Signal Process. **39**, 1478–1480 (1991).
21. S. F. Gull and T. J. Newton, "Maximum entropy tomography," Appl. Opt. **25**, 156–160 (1986).
22. M. L. Reis and N. C. Roberty, "Maximum entropy algorithms for image reconstruction from projections," Inverse Probl. **8**, 623–644 (1992).
23. B. Borden, "Maximum entropy regularization in inverse synthetic aperture radar imagery," IEEE Trans. Signal Process. **40**, 969–973 (1992).
24. B. R. Frieden and D. J. Graser, "Closed-form maximum entropy image restoration," Opt. Commun. **146**, 79–84 (1998).
25. H. Lanteri, M. Roche, and C. Aime, "Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms," Inverse Probl. **18**, 1397–1419 (2002).
26. P. C. Hansen, "Numerical tools for analysis and solution of Fredholm integral equations of the first kind," Inverse Probl. **8**, 849–872 (1992).
27. D. L. Donoho, "De-noising by soft-thresholding," IEEE Trans. Inf. Theory **41**, 613–627 (1995).
28. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge U. Press, 1995).
29. D. L. Marks, A. L. Oldenburg, J. J. Reynolds, and S. A. Boppart, "A digital algorithm for dispersion correction in optical coherence tomography for homogeneous and stratified media," Appl. Opt. **42**, 204–217 (2003).