

# Fluid Registration of Ultrasound using Multi-scale Phase Estimates

Matthew Mellor and Michael Brady  
Medical Vision Laboratory  
Oxford University  
matt@robots.ox.ac.uk

## Abstract

We consider the registration of successive images over a cardiac ultrasound sequence. A key challenge is speckle. We suggest that, for practical purposes, speckle should be regarded as the sum of two components: the useful, temporally correlated speckle pattern; and speckle noise, which is the unpredictable change in the speckle pattern. We also propose that an increase in robustness can be obtained by modelling the spatial correlation of speckle noise. A model is introduced that captures the correlation of speckle noise, by separating the noise into sub-bands and measuring variance within each sub-band and covariance between bands. In order to gain contrast invariance, we express similarity between sub-bands in terms of local phase only, and describe a phase-based extension of the Demons algorithm. The parameters estimated from the noise model significantly enhance the already good registration performance of the phase-demons algorithm.

## 1 Introduction

Motion estimation in echocardiography is a potential source of detailed information about cardiac function. The problem has received much attention, however robustness is still a problem for automatic measurement. A major reason for this is that ultrasound images are dominated by speckle, an interference pattern caused by sub-resolution structures. Since speckle is caused by an interaction of the ultrasound pulse with the tissue, it is often helpful to consider it as useful information. Of particular interest for motion estimation in echocardiography, is the property of temporal correlation; when only small motions are present, it is possible to recognize speckle patterns after a small movement and thereby quantify motion in otherwise featureless tissue regions. However, it is unlikely that the speckle pattern remains identical, particularly if any rotation or out of plane translation occur [1]. We therefore propose that speckle be considered as the sum of two components: a component which is predictable, which we refer to as speckle pattern, and a component which is unpredictable, which we refer to as speckle noise. This work aims to find a compromise between these two components, making use of speckle pattern, while acknowledging the possibility of speckle noise.

The first order statistical properties of speckle [2] and speckle noise [3] have been extensively modelled. Some of the simpler first order models have been incorporated into motion estimation algorithms [4]. However, the second order properties are generally

ignored, even though physical models for the process are well understood [5]. Despite this, spatial correlation is the most striking feature of speckle noise.

In this paper, we describe an algorithm to estimate tissue displacement in ultrasound based on the local image phase, which is estimated with a bank of bandpass filters. In the frequency domain, the power spectrum of speckle noise is a smoothly varying function of frequency. Since each filter responds to a limited range of frequencies, they each capture a different part of the noise spectrum. The noise spectrum can therefore be approximated by the power of the noise responses of the various filters. The more filters that are used, the more accurate the approximation. The filter responses are converted into a contrast invariant local descriptor by estimating the local phase, using the monogenic signal [6]. A similarity measure over local phases is described in section 3, which estimates the reliability of phase information on the basis of a noise model developed in section 2, and the local energy. This measure is then used to drive the demons algorithm [7, 8].

Bandpass filters have previously been used for ultrasound registration with some success. Rousseau *et al* [9] use the responses of oriented band-pass filters to represent image locations. The similarity measure is based on the Bhattacharyya distance between filter response distributions. This method is therefore texture driven, rather than feature driven, and is perhaps better suited to registering significantly different images, e.g. scans taken at different times. Yeung *et al* [10] also represent speckle with the responses of bandpass filters. However, the feature vectors to be matched consist only of the energies of the filter responses. Furthermore, the energy bands are summed together to give a local spectral power estimate which is used as a single dimensional feature. The disadvantages of this method include: the low dimension of the local descriptor, susceptibility to contrast change and the implicit assumption of equal noise in each band. Local phase techniques have been applied by Mulet-Parada to feature detection and tracking [11]. Although the results of these algorithms are promising, there are two drawbacks: the number of filters that need to be applied, and the choice of filter. The significance of both of these problems are reduced in the present work: we make use of improvements in local phase analysis to reduce the number of filters used, and we employ a multi-scale noise model provides a probabilistic description of the local phase across scales, which reduces the need for the filters to be well suited to the image structure.

## 2 Speckle Noise Model

Speckle is an interference pattern caused by sub-resolution tissue structures. We have stressed the difference between a speckle pattern and speckle *noise*. When registering consecutive frames of ultrasound, this distinction is important, since there will be some correlation of the speckle pattern over time and these correlated speckle patterns are valuable features. Here, we aim to estimate speckle noise, rather than speckle patterns. Speckle has a certain amount of spatial correlation, depending on the shape of the Point Spread Function (PSF) of the imaging system and the spatial correlation of the scatterers. For cardiac ultrasound, the PSF tends to dominate. In this case, we expect the correlation of the speckle pattern, and therefore of the speckle noise, to be determined by the auto-correlation of the PSF [5]. In general, we expect the PSF to be a smooth, symmetrical function, with a slowly varying, lowpass frequency spectrum. Therefore, the noise power spectrum should also be a slowly varying function with greater power at low frequencies.

When the noise spectrum is relatively simple, it is possible to approximate it by the sum of just a few functions. In this paper, we approximate the noise spectrum with a sum of the spectra of Difference of Gaussian (DoG) filters; in principal, any bandpass filter family could be used, but DoG filters are convenient to work with and cover the frequency band of interest economically. The filters are chosen so that the positive Gaussian of one filter is the negative Gaussian of the next filter, enabling the frequency band between the highest and lowest frequency filters to be reconstructed by summing the filters:

$$I'(x, y) = \sum_{i=1}^N d_i(x, y) \quad (1)$$

where  $I'(x, y)$  is the reconstructed image from which high and very low frequencies are missing,  $N$  is the number of filters used to represent the image, and  $d_i(x, y)$  is the response of the  $i$ th difference of Gaussians filter at point  $[x, y]$ . The noise is therefore also represented with sum of filter responses. It is assumed that the noise in each filter band is Gaussian distributed and is an additive component. While this does not perfectly agree with the theoretical distribution of ultrasound, it is often a good approximation (see section 5). The aim of this work is to model the correlation of speckle noise. The classical tool for this is the Power Spectral Density (PSD). Because each filter responds to a different range of frequencies, it is possible to approximate the variation of noise power with frequency (i.e. the PSD), by the mean noise in each filter band. Since the expected PSD of speckle is smoothly varying [5], it should be possible to approximate it with a small number of such samples. Let  $n(x, y)$  denote an instance of the noise signal with (discrete) Fourier transform  $N(u, v)$  and let  $d_{ni}(x, y)$  be the response of filter  $f_i$  to the noise, with Fourier transform  $D_{ni}(u, v)$ . Ignoring the noise that falls outside the frequency response of any of the filters, we have:

$$\begin{aligned} n(x, y) &= \sum_i d_{ni}(x, y) \\ N(u, v) &= \sum_i D_{ni}(u, v). \end{aligned} \quad (2)$$

An estimate for the power spectral density may then be found in terms of the response of the filters to a sample of noise:

$$\begin{aligned} W(u, v) &= \langle N(u, v)^2 \rangle \\ &= \left\langle \left( \sum_i D_{ni}(u, v) \right)^2 \right\rangle \\ &= \sum_i \sum_j \langle D_{ni}(u, v) D_{nj}(u, v) \rangle \end{aligned} \quad (3)$$

where  $\langle x \rangle$  denotes the expectation of  $x$ . The PSD may now be approximated by introducing the assumption that, for each each pair of filters  $f_i(x, y)$  and  $f_j(x, y)$ , the noise has constant power  $\sigma_{ij}^2$ . The expectations in equation 3 may then becomes:

$$\langle D_{ni}(u, v) D_{nj}(u, v) \rangle \approx F_i(u, v) F_j(u, v) \sigma_{ij}^2. \quad (4)$$

Hence the noise power spectrum is approximated by the various covariances,  $\sigma_{ij}$ . The resulting PSD estimate may be written concisely in the form of the quadratic product of a filter vector,  $\underline{F}(u, v)$ , with a covariance matrix:

$$\underline{F} = [F_1(u, v), \dots, F_n(u, v)] \quad (5)$$

$$C_{ij} = \sigma_{ij}^2 \quad (6)$$

$$W(u, v) = \underline{F}^T C \underline{F} \quad (7)$$

In order to complete the noise model, it is necessary to define a way to estimate the model parameters from the data. Assuming the availability of a noise sample,  $n(x, y)$ , it is possible to derive estimates for each  $\sigma_{ij}$  directly, by summing equation 4 over all frequencies:

$$\begin{aligned} \sum_{u,v} \langle D_{ni}(u, v) D_{nj}(u, v) \rangle &= \sum_{u,v} F_i(u, v) F_j(u, v) \sigma_{ij}^2 \\ \left\langle \sum_{u,v} D_{ni}(u, v) D_{nj}(u, v) \right\rangle &= \sigma_{ij}^2 \sum_{u,v} F_i(u, v) F_j(u, v) \end{aligned} \quad (8)$$

In the case that the noise response  $d_{ni}$  is real and the applied filter is symmetric, it can be shown that:

$$\left\langle \sum_{u,v} D_{ni}(u, v) D_{nj}(u, v) \right\rangle = \left\langle \sum_{x,y} d_{ni}(x, y) d_{nj}(x, y) \right\rangle \quad (9)$$

which identifies the r.h.s of 8 as the covariance of the filter noise-responses,  $C_{ij}^d$ . An equivalent expression holds for the filters. It follows that:

$$\sigma_{ij}^2 = C_{ij} = \frac{\langle \sum_{x,y} d_{ni}(x, y) d_{nj}(x, y) \rangle}{\sum_{x,y} f_i(x, y) f_j(x, y)} = \frac{C_{ij}^d}{C_{ij}^f} \quad (10)$$

where  $C_{ij}^f$  is the covariance of the filter impulse responses. Figure 1 illustrates the performance of an algorithm based on this technique on a 1 dimensional signal. The blue curve is an estimate of the PSD of a synthetic signal designed to have properties comparable to those of a line from a speckle pattern, after wide-band filtering to remove frequencies outside the range of the chosen bandpass filters. The five black curves show the shape of the power spectra of the five bandpass filters and the red curve shows the sum of these and the cross power terms, weighted by C. In practice, because the noise is always observed through the filters, it is possible to ignore the power spectral density and work with the filter noise-response covariance  $C^d$  directly, rather than the PSD model parameter C. Nevertheless, for a given bank of filters one specifies the other. note also that the procedure outlined above can be reversed, so that a known PSD (e.g. from calibration) may be used to define the similarity measure directly, without the need to extract a noise signal.

### 3 Similarity Measure

The problem of analysing one image with an inconvenient noise model has been split into the simultaneous analysis of several images with simple noise models. The next step

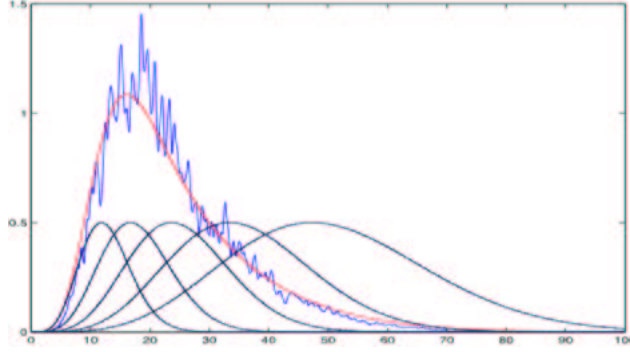


Figure 1: Approximating a PSD with variances of bandpass filters. Blue: the PSD estimated from data sample. Black: five of the fifteen cross power terms of eq. 4. Red: the approximated PSD is a good fit.

is to define a similarity criterion for registering the bandpass images. One option is to assume constant intensity and use an optical flow method, such as the Demons algorithm [7, 8]. The difficulty with this approach is that the intensity of an object in an ultrasound image can change markedly over time. To avoid this we use a similarity criterion based on constant local phase between images. Local phase is a contrast invariant, qualitative description of the local shape of an image and is accompanied by a second quantity, local energy, which is a measure of local signal activity. Local phase has previously been shown to be a powerful tool for accurate registration and motion estimation [12]. Space does not permit a detailed discussion of phase estimation techniques, which may be found in [6, 13]. We estimate phase using the monogenic signal [6]. The monogenic signal enables local phase and energy to be estimated from a bandpass image  $I_b$  by the application of two further linear filters, which can be expressed in the frequency domain as:

$$H_1(u, v) = \frac{u}{\sqrt{u^2 + v^2}}, H_2(u, v) = \frac{v}{\sqrt{u^2 + v^2}} \quad (11)$$

The responses of these filters to the bandpass image are combined, in the image domain, with the bandpass image itself to estimate local phase and energy as follows:

$$A(x, y) = \sqrt{I_b^2 + (h_1 \otimes I_b)^2 + (h_2 \otimes I_b)^2} \quad (12)$$

$$\phi(x, y) = \tan^{-1} \left( \frac{I_b}{\sqrt{(h_1 \otimes I_b)^2 + (h_2 \otimes I_b)^2}} \right) \quad (13)$$

where  $A(x, y)$  is the local energy and  $\phi(x, y)$  is the local phase. To measure phase similarity in noisy data, it is important to be able to deal with the effects of noise. Equations 12 & 13 are expressed as functions of even and odd components, where the odd part is the combined responses of the two odd filters,  $\sqrt{(h_1 \otimes I)^2 + (h_2 \otimes I)^2}$ , and the even part is simply the bandpass filtered image,  $I_b$ . The even and odd components are orthogonal to each other, so for unstructured noise (which has a flat phase distribution), the noise

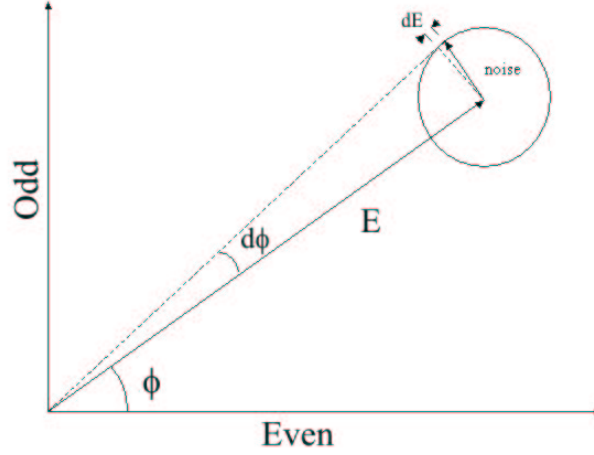


Figure 2: The effect of random noise on a phasor

responses of the even and odd components are independently distributed. It is therefore possible to model the effect of noise on phase and energy as illustrated in figure 2.

If we assume that the noise distribution is Gaussian with standard deviation  $\sigma$ , the expression for phase error at a location with local energy  $A$  is [14]:

$$P(\Delta\phi|\sigma,A) = \frac{1}{2\pi} e^{-\frac{A^2}{2\sigma^2}} + \frac{A \cos \Delta\phi}{\sigma\sqrt{2\pi}} e^{-\frac{A^2 \sin^2 \Delta\phi}{2\sigma^2}} Q\left(\frac{-A \cos \Delta\phi}{\sigma}\right) \quad (14)$$

where  $Q$  is the complementary error function. For the specific definitions of odd and even components used here, the situation in figure 2 does not hold exactly when the odd response is very small, in which case the noise, as shown, could theoretically make the odd component negative, which is impossible in practice. However, we have found that the effect is insignificant.

Note that the local energy appears in ratio with the noise standard deviation. The ratio  $A/\sigma$  is the local signal to noise ratio (SNR). Although equation 14 appears rather complex, for large values of SNR it is well approximated by a Gaussian distribution:

$$\hat{P}_\phi(\phi|\sigma,A) \approx \frac{A}{\sigma\sqrt{2\pi}} e^{-\frac{A^2 \phi^2}{2\sigma^2}} \quad (15)$$

This approximation is reasonable even for signal to noise ratios as low as two. With the phase noise expressed as a Gaussian, it is natural to express the phase distance as a Mahalanobis distance. The Mahalanobis distance,  $D_m$ , between two multiscale phase vectors is defined as:

$$D_m^2 = \Delta\Phi^T C^A \Delta\Phi \quad (16)$$

$$\Delta\Phi = [\Delta\phi_1, \dots, \Delta\phi_n] \quad (17)$$

where  $C^A$  is the covariance of the Gaussian distributed phase errors, determined from the local SNR according to (15) and  $\Delta\phi_n$  is the phase difference at scale  $n$ . It may easily be

shown that the covariance  $C^A$  can be derived at each point from the filter noise-response covariance matrix  $C^d$  and the observed energies:

$$C^{A^{-1}} = \underline{A} \odot C^{d^{-1}} \quad (18)$$

where  $\underline{A}$  is the local energy matrix, with  $\underline{A}_{ij} = A_i A_j$  and  $A_i$  is the local energy at scale  $i$ . The operator  $\odot$  denotes element-wise multiplication, i.e.  $C_{ij}^{A^{-1}} = \underline{A}_{ij} C_{ij}^{d^{-1}}$ . Note that this depends on the assumption that the noise power in the odd and even filters is the same. This is not strictly true of bandpass filtered speckle noise, but is a good approximation. In the algorithm presented in the next section  $C^d$  is assumed constant over the image (though this need not be the case). The local energy matrix  $\underline{A}$ , however, is non-constant by definition, which means that the similarity measure is itself non-constant. Where the underlying signal is strong, the energy matrix is large and so the corresponding covariance matrix is small, which results in lower toleration of phase error. Conversely, when the energy is very low, the covariance is very large and even significantly distinct phase values do not appear very different. For this reason, we refer to (16) as an ‘local energy mediated’ similarity measure.

## 4 Algorithm

Our implementation is based on the demons algorithm, a fluid registration algorithm which registers images on the assumption that intensity is constant over time [8]. We extend the same framework to the problem of minimizing the phase distance according to (16). The demons algorithm is an iterative method which updates the present estimate of the displacement between images,  $v_n$ , with a function of the phase distance and the gradient of the phase distance:

$$v_{n+1}(x,y) = G_\sigma \otimes \left( v_n + \frac{D_m(x,y)}{\|\nabla D_m(x,y)\|^2 + D_m(x,y)^2} \nabla D_m(x,y) \right) \quad (19)$$

Where  $G_\sigma$  is a regularizing Gaussian of width  $\sigma$ . The phase distance  $D_m$  is calculated at each point  $x$  between the phases of the target images and the phases of the source image deformed by the current transformation estimate  $v_n$ . The gradient of the phase distance is approximated by central difference. At each iteration, the noise is estimated by subtracting the target image from the source image deformed by the present registration estimate. Initially, this noise field contains image structure and the resulting noise model is contaminated. As the registration proceeds, the amount of signal contaminating the noise decreases. The noise field is used to generate the noise model covariance  $C_f$  by measuring the covariances of the bandpass filter responses. Since the filtered noise can be found by subtracting the filtered images, noise estimation does not require any additional filtering operations. For the results presented in this paper, five difference of Gaussian filters were used, given by the differences between six consecutive Gaussians, with standard deviations  $\sigma_n = 2^{\frac{n+2}{2}}$ ,  $n = 1 : 6$ . The regularizing Gaussian was set at  $\sigma = 4$ .

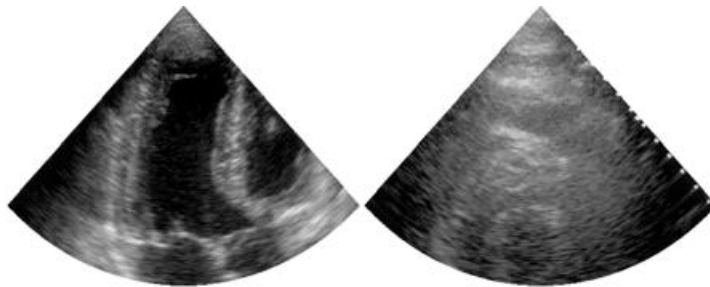


Figure 3: An example image from the high quality sequence (left) and the low quality sequence (right)

## 5 Results

We illustrate the phase-based demons algorithm on two cardiac ultrasound sequences, one judged to be of high quality, the other of low quality. An example frame from each sequence is shown in figure 3.

To evaluate the effect of the noise modelling, experiments were repeated twice, once with the noise model and local energy based similarity measure and once using the identity matrix in place of the phase noise covariance (i.e assuming white noise and ignoring local SNR). Although ground truth motion was not available for the two sequences, it was possible to compare the consistency of the two motion estimates. To do this, we compared two alternative ways of estimating the motion between next-but-one neighbors in a sequence: direct estimation where the two frames are registered directly, and indirect estimation where the motion from the first frame to the middle frame is added to the motion from the middle frame to the second frame. We make the assumption that the differences between next-but-one neighbors are purely determined by motion and noise. In this case, the difference between the direct and indirect motion estimates should be largely due to noise. Table 1 shows the mean difference between direct and indirect motion estimates for the good and the poor sequences, both with and without the proposed noise model. The consistency of the noise-model based algorithm is better for both sequences. However, even without any realistic noise model, the algorithm still performs quite consistently, suggesting that it is inherently robust. Strangely, both algorithms seem to perform better on the poorer image. The reason for this would appear to be that the high quality images actually contain large regions (mostly blood) that don't contain useful features.

To estimate the accuracy of the algorithm, we have also applied it to a synthetic data

	Without noise model		With noise model	
	High quality	Low quality	High quality	Low quality
Mean	1.38	1.31	1.16	1.12
Variance	1.15	1.14	0.5	0.5

Table 1: Mean differences, in pixels, between direct and indirect motion estimates, with and without the proposed noise model.



	Mean error	S.D.	% Err. < 0.5
With n.m.	0.80	0.048	36%
Without n.m.	0.88	0.052	26%

Table 2: Summary of results of synthetic experiment.

set. This was a set of thirty image pairs generated from the high quality ultrasound sequence. The first frame of each pair is a randomly selected frame from the sequence. The second frame is created by applying a randomly generated warp. Simulated speckle noise, created using the method described in [15], was then added to both images. The simulated speckle noise is designed to have many of the attributes of real speckle noise; it is correlated, with location dependent autocorrelation function, and was generated by taking the difference between the speckle patterns generated by two partially correlated scatterer distributions. The random warps have a mean displacement of 3 pixels, and were generated by convolving a field of zero-mean random vectors with a Gaussian and so should be fully compatible with the regularizer of the Demons algorithm. The synthetic image pairs were then registered using the phase distance derived from the true noise model.

Figure 4 shows the distribution of mean error before and after registration. In the presence of tissue the registration error is significantly reduced. We can deduce from this that the algorithm uses some of the fine scale information as well as the large features, such as the blood-tissue boundary. If this were not the case, the algorithm would suffer from the aperture problem and would not achieve such levels of accuracy. Table 2 summarizes the results of the experiment. The increase in performance introduced by the noise model is significant ( $p < 0.001$ ), but seems small ( $\approx 10\%$ ). However, since neither method can effectively register the blood pool, this performance increase is concentrated in the tissue. This is reflected in the 36% increase in the proportion of the image aligned with accuracy of better than half a pixel width.

## 6 Discussion

We have introduced a method to model the second order statistics of speckle noise by independently modelling the first order statistics of the responses of band-pass filters. We

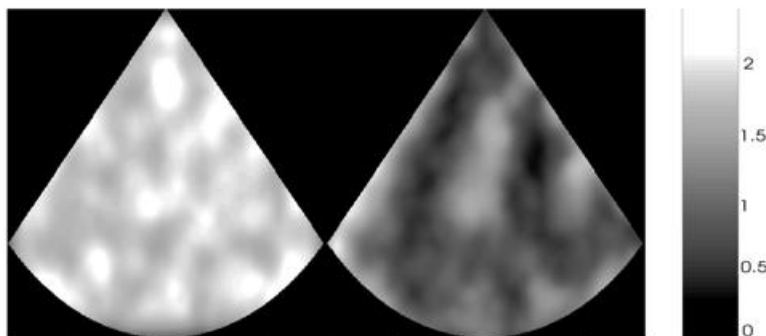


Figure 4: Average errors before and after registration for the synthetic sequence

have shown how this model can be combined with local phase estimation to produce a contrast invariant local similarity measure, with a non-white noise model. This is used to drive the demons algorithm, but could, in principal, be combined with many other cost function minimization strategies. It should also be noted that the method is not restricted to ultrasound; provided an appropriate family of bandpass filters is chosen, the method could be applied to any data with correlated noise. The registration method appears to be both accurate and intrinsically robust, but benefits significantly from a model of the noise correlation. At the present, it is not possible to say whether the benefits of modelling noise correlation outweigh those of accurate modelling of the first order properties. To this end we are currently carrying out an extensive test of the method, in competition with two other methods based on first order noise models.

## References

- [1] M. Bertrand J. Meunier. Ultrasonic texture motion analysis: Theory and simulation. *IEEE Trans. Medical Imaging*, 14(2):293–300, 1995.
- [2] V. Dutt & J. F. Greenleaf. Ultrasound echo envelope analysis using a homodyned k distribution signal model. *Ultrasonic Imaging*, 16:265–287, 1994.
- [3] I. Dinstein B. Cohen. New maximum likelihood estimation schemes for noisy ultrasound images. *Pattern Recognition*, 35:455–463, 2002.
- [4] D. Boukerroui. Velocity estimation in ultrasound images: A block matching approach. *IPMI*, pages 586–598, 2003.
- [5] J. M. Sandrik & H. Lopez R. F. Wagner, S. W. Smith. Statistics of speckle in ultrasound b-scans. *IEEE Trans. Sonics and Ultrasonics*, 30(3), 1983.
- [6] M. Felsberg; G. Sommer. A new extension of linear signal processing for estimating local properties and detecting features. *Proc. DAGM*, 2000.
- [7] J.P. Thirion. Image matching as a diffusion process: an analogy with maxwell’s demons. *Medical Image Analysis*, 2(3):243–260, 1998.
- [8] N. Ayache J. Meunier A. Guimond, A. Roche. Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. *IEEE. Trans. Medical Imaging*, 20(1):58–69, 2001.
- [9] F. Rousseau; R. Fablet; C. Barillot. Robust statistical registration of 3d ultrasound images using texture information. *IEEE Int. Conf. on Image Processing, ICIP’03*, Sept 2003.
- [10] F. Yeung; F. Levinson; D. Fu; & K.J. Parker. Feature adaptive motion tracking of ultrasound image sequences using a deformable mesh. *IEEE Trans. Medical Imaging*, 17:945–956, 1998.
- [11] J. A Noble M. Mulet-Parada. 2d+t acoustic boundary detection in echocardiography. *Medical Image Analysis*, 4:21–30, 2000.
- [12] D.J. Fleet. Disparity from local weighted phase-correlation. *IEEE International Conference on SMC*, pages 48–56, October 1994.
- [13] W.T. Freeman; W.H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.
- [14] A.B. Carlson. Communication systems: An introduction to signals and noise in electrical communication. *McGraw-Hill*, 1986.
- [15] M. Mellor. Phase methods for medical image registration. *DPhil Thesis, University of Oxford*, 2004.