Reconstructing 3D skin surface motion for the DIET breast cancer screening system

Tom Botterill, Thomas Lotz, Amer Kashif, J. Geoffrey Chase

***Abstract*—Digital Image-based Elasto-Tomography (DIET) is a prototype system for breast cancer screening. A breast is imaged while being vibrated, and the observed surface motion is used to infer the internal stiffness of the breast, hence identifying tumours. This paper describes a computer vision system for accurately measuring three-dimensional surface motion. A model-based segmentation is used to identify the profile of the breast in each image, and the 3D surface is reconstructed by fitting a model to the profiles. The surface motion is measured using a modern optical flow implementation customised to the application, then trajectories of points on the 3D surface are given by fusing the optical flow with the reconstructed surfaces. On data from human trials, the system is shown to exceed the performance of an earlier marker-based system at tracking skin surface motion. We demonstrate that the system can detect a 10mm tumour in a silicone phantom breast.**

***Index Terms*—Cancer detection, Image motion analysis, 3D reconstruction, Object segmentation**

I. Introduction

Digital Image-based Elasto-Tomography (DIET) is a prototype system for breast cancer screening. It works by inducing steady-state oscillation in a breast, then imaging its surface with cameras. The observed surface motion is used to infer the internal stiffness of the breast, hence identifying tumours [1]. The system is intended to provide a safer, cheaper, and pain-free alternative to mammograms.

The first DIET system measured the skin surface motion by tracking hundreds of coloured markers [2]. Detected marker positions are matched between consecutive images, and between neighbouring cameras, enabling the 3D motion of each marker to be reconstructed. The time and inconvenience needed to apply markers to the breast is undesirable in a low-cost screening system, hence we have developed a system that reconstructs the surface motion without markers. The proposed system first reconstructs the breast's surface in 3D by fitting a model to the profiles of the breast observed in the images, then measures the surface motion of the skin by computing the optical flow between frames. The 3D trajectories of points on the breast's surface are given by combining the optical flow with the 3D surface model. This paper describes this new computer vision system, and demonstrates that its accuracy exceeds the accuracy of the previous marker-based system on images from human trials.

A. The DIET imaging system

The DIET system uses elastography to detect tumours. The breast hangs freely in the machine while being mechanically actuated from below at a frequency f (between 20Hz and 80Hz), with an amplitude of approximately 1mm. The breast enters a steady-state motion with period $1/f$. The breast is imaged by five cameras arranged in a ring (Figure 1). At N equally spaced points in the cycle (usually $N = 10$), the five cameras each image the breast, giving N images per camera showing the breast at N stages throughout the motion cycle. The cameras and actuator are synchronised by strobing a set of LEDs at specific phases in the cycle.

The current marker-based approach to 3D surface motion estimation tracks and reconstructs the positions of about two thousand coloured markers (paper dots) attached to the breast. Markers are detected by thresholding by pixel colour, and each marker's centroid is computed. Frame-to-frame motion is small compared with the size of the markers, enabling markers to be tracked between frames. To reconstruct the breast surface motion, all possible correspondences between pairs of markers (along epipolar lines) are computed, and the correctly reconstructed marker positions are selected using a local

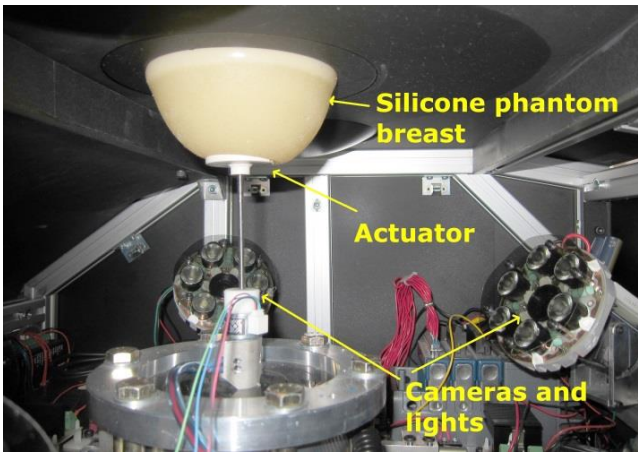


Figure 1: The DIET machine, showing two of the five cameras, the actuator, and a silicon phantom breast. The cameras are 25cm from the breast, and the machine measures 80cm by 71cm by 38cm.

The authors would like to thank the nurses and volunteers at Canterbury Breastcare, Christchurch, NZ, for their contributions in clinical trials of the DIET system. Every subject gave informed consent. Approval for this study and the use of this data was granted by the NZ National Ethics Committee, South Island Regional Ethics Committee A.

T. Botterill is with the Department of Computer Science, University of Canterbury, Christchurch, New Zealand (e-mail: tom.botterill @canterbury.ac.nz); T. Lotz, A, Kashif and J. G. Chase are with Department of Mechanical Engineering, University of Canterbury.

Random Sample Consensus (RANSAC)-based approach [2].

To detect tumours we examine the phase of the motion with respect to the actuator across the breast; this approach can detect tumours in silicone 'phantom' breasts, as the region around the tumour has a smaller phase difference than other parts of the breast [3]. Preliminary human trials have shown differences in surface motion between breasts with and without tumours [4]. An alternative approach we are currently investigating is to build a finite element model of the breast, and to optimise the internal stiffness so that the modelled surface motion matches the observed motion [5].

Alternative methods for measure breast stiffness include ultrasound elastography and Magnetic Resonance (MR) elastography. Ultrasound elastography identifies stiffer regions by measuring the relative speed and attenuation of ultrasound through tumours and healthy tissue, and is widely used for assessing breast tumour malignancy [6, 7], however requires skilled operators. MR elastography systems induce vibrations in the breast while imaging the tissue motion using an MR scanner [8], however MR imaging is expensive. The limitations of these approaches motivate the development of alternatives such as DIET.

The focus of this paper is our new system for measuring the 3D surface motion of breasts imaged in the DIET machine without requiring markers. The main contribution is to describe the computer vision system which measures this motion, and to demonstrate that it matches the performance of the old marker based system at accurately recovering the surface motion, and at detecting tumours in silicone phantom breasts.

II. Background

3D surface reconstruction is a common task in computer vision and medical imaging, and many techniques are available. Similarly, many techniques have been developed for computing optical flow. This section first reviews techniques for 3D surface reconstruction from images; and second, evaluates the optical flow methods best suited to the application.

A. Image-based 3D surface reconstruction

Many applications require a parameterised 3D model of an object to be estimated from a set of images, e.g. [9, 10, 11, 12]. There are two basic approaches to model fitting from 2D images: one approach is to reconstruct some 3D data directly from the images, then to refine the model parameters so that the model matches the 3D data. The second approach is to use the object's properties to locate the model in each 2D image, and then refine model parameters so the object's projection back into the images matches its observed location [13, 14].

Common approaches for estimating 3D data directly from 2D images include dense stereo algorithms [15], feature matching algorithms [14, 16], and Shape from Shading algorithms [17]. These algorithms make assumptions about the scene in order for the reconstructions to be unambiguous, for example dense stereo algorithms make the assumption that the surface is either smooth or piecewise planar [18, 19]; feature-matching approaches generally discard ambiguous matches [16], and can only reconstruct objects where matchable features can be identified; and Shape from Shading algorithms assume the surface has uniform colour [17]. These assumptions are not appropriate for breast surfaces, where matchable features aren't present, and where there is significant variation in colour from lighting and reflectance variation across the breast. Other active sensors suitable for 3D imaging in clinical settings include structured light systems or Time-of-Flight cameras, however these sensors add to the cost of a system.

A 3D reconstruction method that is well suited to breast surface reconstruction is Shape from Silhouette [20]. Shape from Silhouette methods find an object's outline in each image, then optimise model parameters so that the model backprojected into each image matches the silhouette [21, 22]. The error term which is minimised is analogous to the reprojection error in bundle adjustment [13, 14], and as with bundle adjustment, additional knowledge of the object's structure can be incorporated into the optimisation. Shape-from-silhouette is ideal for breast surface reconstruction, as the outline of breasts can easily be extracted from images, and knowledge that the breast surface is smooth can be used to constrain the reconstruction.

For fitting smooth 3D solid objects to a set of silhouettes, Jones and Oakley proposed the radial intersection set representation [23], where objects are represented as a parametric function of the angles (θ, ϕ) in spherical polar coordinates. The surface is discretised into a grid of (θ, ϕ) points. For each (θ, ϕ), the set of points where that ray intersects the surface is estimated, enabling complex non-convex 3D shapes to be represented. A similar representation is proposed by Staib and Duncan [24], who represent smooth 3D surfaces by a function of two parameters. The Fourier coefficients are optimised to match the surface to image edges. Prior knowledge is incorporated by adding constraints on the Fourier coefficients to the optimisation. The method is used to reconstruct the surface of a brain from MRI images. A similar approach is used by Zheng et al. [12] to fit parametric models of femurs to X-ray images. Hybrid silhouette and feature-based systems are the current state-of-the-art methods for markerless hand tracking and human motion capture [11, 25].

B. Breast model reconstruction

Systems for estimating 3D models of breasts have been developed for monitoring the effects of breast surgery. These use a range of 3D imaging sensors, including NMR and laser scanners [26], and structured light [27]. Carter [28] uses a combination of stereo images, detected markers, and structured light for modelling the breast position during surgery, so that the position of a tumour can be tracked as the breast deforms. Any of these sensors could be used for breast model reconstruction in the DIET system, however the current colour camera-based system is inexpensive, provides sufficiently accurate 3D models, and is also well suited for surface motion estimation.

1. **Segment breast images**
   For each camera:
      Localise breast and actuator
      For each timestep:
         Refine contour location to accurately match breast edge
2. **Reconstruct surfaces**
   For each timestep:
      Reconstruct breast surface
3. **Compute optical flow**
   For each camera:
      For each consecutive pair of frames:
         Compute optical flow
4. **Reconstruct surface motion**
   Project optical flow onto surfaces to recover surface motion.

Figure 2: Overview of image processing pipeline.

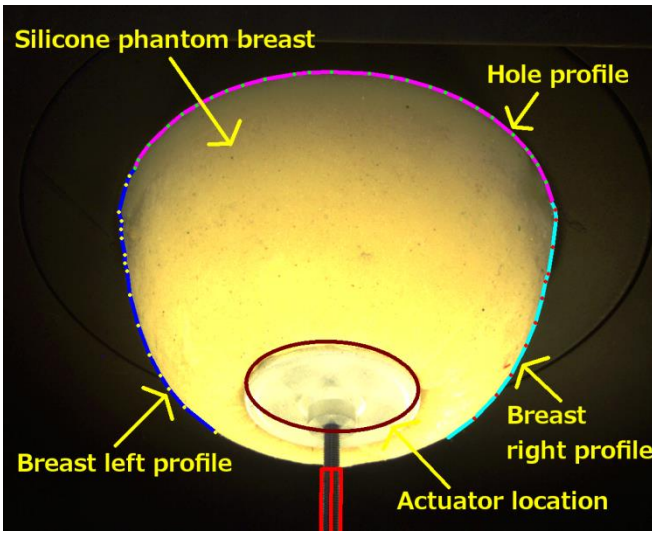


Figure 3: A silicone phantom breast without markers, showing the segmentation of the breast's profile and actuator location.

C. *Surface tracking and optical flow estimation*

When two images of a moving scene are captured in quick succession, the positions of objects in the scene move a small amount between views. Optical flow algorithms find the mapping between the positions of objects in two successive views. There are two basic types of optical flow algorithm: sparse optical flow algorithms select a set of points in the image which make good features to track (by applying a corner detector), then identify the location of the same points in the second image, e.g. by using the Lucas Kanade tracking algorithm [29]. By contrast, dense optical flow algorithms find a continuous, or piecewise continuous, vector field mapping each pixel in one image to its corresponding location in the other [30]. Modern dense optical flow algorithms [31, 32] are generally based on the classic formulation by Horn and Schunck [33]. These algorithms optimise an optical flow field to minimise both the difference in appearance between corresponding pixels and its smoothness. Both dense and sparse optical flow methods are subject to outliers (incorrect motion estimates) when the motion is ambiguous, however dense methods have the advantage that smoothness constraints enforce consistent flow estimates across untextured surfaces. A further advantage for our application is that small variations in texture can be tracked, without having to identify trackable features.

Dense optical flow was used by Bradley et al. for tracking skin surface motion for modelling facial expressions [34]. The skin texture is sufficient that very small (subpixel) motions can be tracked, enabling optical flow-based systems to replace earlier marker-based systems. Dense optical flow is often used for deformable image alignment, for example for aligning Magnetic Resonance (MR), Computed Tomography (CT), or ultrasound images of organs as they deform [35]. These applications are susceptible to lighting or image intensity variation, so lighting should be controlled in hardware or corrected in software to ensure accurate results.

III. Proposed Motion Reconstruction Algorithm

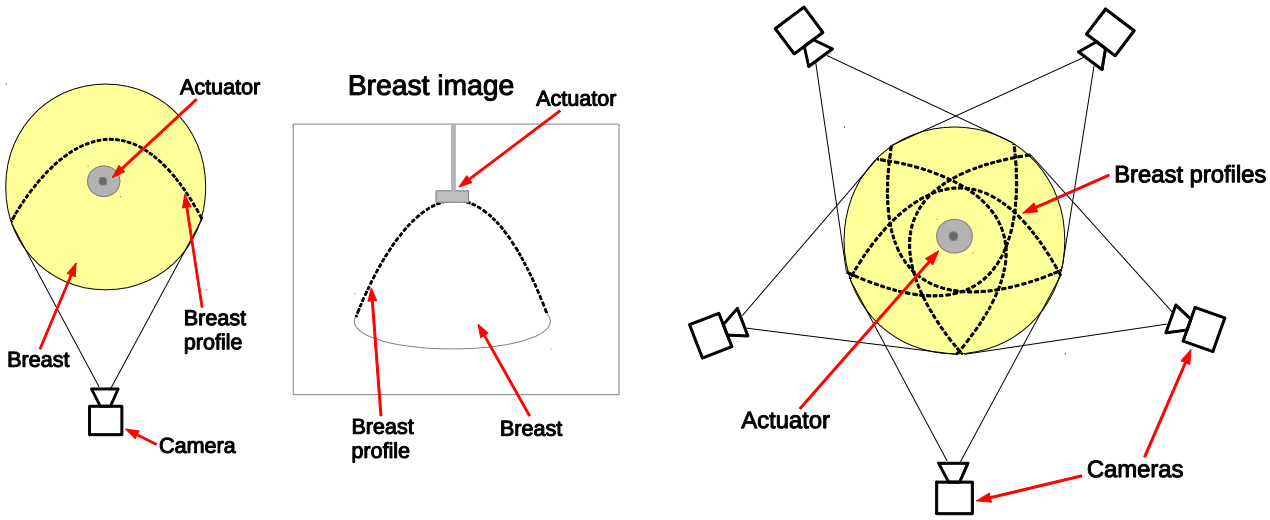
This section describes the combination of algorithms used to compute the 3D motion of points on the breast surface, given a set of $5 \times N$ images captured by the five cameras. This process consists of four stages: first, each image is segmented, to accurately localise the profile of the breast. Second, for each timestep, a 3D surface model of the breast is estimated by fitting a model to the breast profiles. Third, the skin surface motion between images from each pair of consecutive timesteps is computed with a dense optical flow algorithm. Fourth, the 3D surface motion is estimated by combining the optical flow fields with the reconstructed surfaces. An overview of this processing pipeline is given in Figure 2.

Extensive use is made of standard methods in multi-view geometry, in particular the Levenberg-Marquardt algorithm for nonlinear optimisation, and the projective geometry of calibrated cameras. Details of these methods are in Hartley and Zisserman [13].

A. *Model-based segmentation of breast*

The DIET machine is designed to image a range of breast shapes and sizes, so the position of the breast and actuator in each image varies substantially between trials. The positions of the boundaries of these components are prerequisites for reconstructing the breast surface so are computed first. To segment these components, a sequence of standard computer vision methods is used. First, images are normalised to reduce the effects of variation in colour and lighting between datasets and at different timesteps. Second, a simple colour-based segmentation of the different components of the scene is performed. Third, a model-based approach is used to find different scene components in the colour-segmented images.

For the colour-based segmentation, the distributions of the intensity and saturation on the breast, background, and actuator are each modelled as a 2D Gaussian. These models



(a) Each camera views the breast from one side. The 2D profile of the breast (dashed) gives a constraint on the 3D surface.

(b) View showing the five cameras and five profiles from below the breast. The profiles, together with a smoothness constraint, enable the 3D surface to be recovered.

Figure 4: The DIET machine contains five cameras arranged in a ring around the breast. The 3D surface is found by fitting a model to the five profiles.

give the likelihood of each pixel being background, breast or actuator. This segmentation produces an 'actuator image' where the actuator components are clearly visible, and a 'skin image', from which the breast location and profile can be extracted.

The model-based segmentation extracts the actuator from the actuator image, and the breast from the skin image. The approximate positions of both the actuator and the breast are initialised by applying Kadane's two-dimensional Maximal Subarray Algorithm [36] to these images. The actuator position is refined using Levenberg-Marquardt's method to optimise the parameters of the model (a rod connected to an ellipse) to match each image.

To locate the breast in the skin image, an Active Contour Model (or 'snake') is used [37]. The Active Contour Model fits a polyline to the breast profile by iteratively moving the polyline control points to maximise the fit of the line to the image (minimise 'external energy'), while minimising a measure of the line's curvature ('internal energy'). External energy is computed from the response of an edge detector applied to the skin image. Once refinement is complete, knowledge of the approximate hole location is used to divide the polyline into three parts: the two breast profiles (left and right), and the profile of the circular hole around the breast. A segmented image is shown in Figure 3.

To ensure that the segmentation is efficient, and to aid convergence, an image pyramid is used: the segmentation is first performed on a downsampled image, then the results of this segmentation is used to initialise the segmentation at a larger scale. For each camera, the segmentation algorithm is applied first to the average of the N images, then this segmentation is used to initialise the segmentation for the other images. The breast profile and actuator position change from frame-to-frame, so are refined for each image.

Once the segmentation is computed, the 3D positions of the actuator, and the circular hole around the breast, are found by fitting a 3D circle to each using Levenberg-Marquardt optimisation. On each iteration, points are sampled from the 3D circles and projected into the images. The distance between these points and the measured hole or actuator positions (the reprojection errors) are minimised.

B. 3D surface reconstruction from profile

The next stage of the reconstruction is to estimate a parametric 3D model of the breast surface from the segmented images, for each of the N frames. The surface model is estimated so that the profile of the breast model projected back into the images matches the contours detected in the images (Figure 4).

We use a parametric representation of the breast surface in spherical polar coordinates (r, θ, ϕ). The origin is fixed at the centre of the machine's circular hole. We estimate the radial distance r for each of a grid of (θ, ϕ) points by using Levenberg-Marquardt optimisation to minimise both the fit between the model and the breast contours and a measure of surface smoothness. The reprojection error, which measures the fit between the model and the breast contours, is computed by projecting model points into the images, then comparing these points with the measured profile positions. To impose the smoothness constraint we minimise the difference between values of r for neighbouring points. A grid of reconstructed 3D surface points backprojected onto a breast image is shown

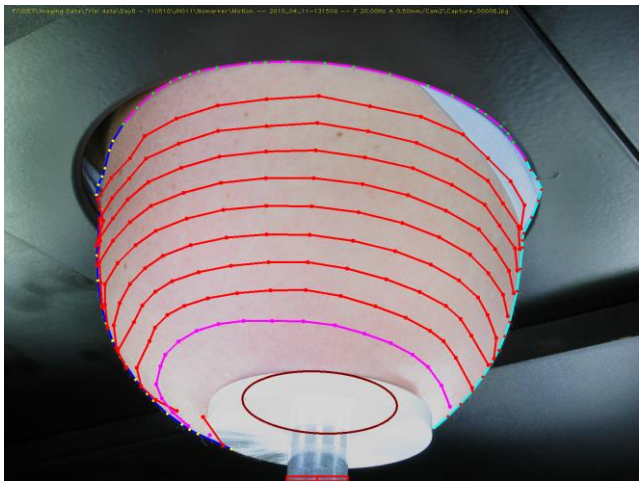


Figure 5: 3D surface backprojected onto a breast image, showing rings of 3D surface points and the detected positions of the breast profile and actuator.

in Figure 5.

C. Optical flow computation

We estimate skin surface motion by applying a modern dense optical flow algorithm to pairs of successive frames. The aim is to recover this motion by tracking the skin's texture. There is sufficient variation in skin texture on a small scale that the surface motion is obvious when images are displayed successively; however a larger source of variation between frames is the illumination change caused by the skin angle changing with respect to the lights, and by spatial variation in light intensity [38]. To ensure accurate optical flow estimation, we have adapted the method of Sun et al. [31], to correct for variations in lighting, and to enforce smooth and continuous optical flow.

Given two images, $\boldsymbol{I}_1$, $\boldsymbol{I}_2$ with grey (or colour) values $I_1(\boldsymbol{x})$ and $I_2(\boldsymbol{x})$ at a pixel location $\boldsymbol{x}$, optical flow algorithms seek to estimate a 2D vector field $\boldsymbol{U}$ with value $\boldsymbol{U}(\boldsymbol{x})$ at each pixel, with the properties that $\boldsymbol{U}$ is (piecewise) smooth, and the appearance of $\boldsymbol{I}_1$ warped by $\boldsymbol{U}$ matches the appearance of $\boldsymbol{I}_2$. $\boldsymbol{U}$ is found by minimising a global error function which penalises differences in appearance between $\boldsymbol{I}_1$ warped by $\boldsymbol{U}$ and $\boldsymbol{I}_2$ (the 'data' term), and which enforces $\boldsymbol{U}$ to be smooth (the 'spatial' term). At each pixel $\boldsymbol{x}$, the data term is given by:

$$c_{data}(\boldsymbol{x}) = \rho_{data}(I_1(\boldsymbol{x}) - I_2(\boldsymbol{x} + \boldsymbol{U}(\boldsymbol{x}))). \quad (1)$$

c_{data} measures the similarity of corresponding points in the two images. The spatial error term is given by:

$$c_{spatial}(\boldsymbol{x}) = \sum_{\substack{Neighbours \\ \boldsymbol{x}' of\ \boldsymbol{x}}} \left(\rho_{spatial}(\boldsymbol{U}(\boldsymbol{x}) - \boldsymbol{U}(\boldsymbol{x}'))\right). \quad (2)$$

$c_{spatial}$ measures the smoothness of $\boldsymbol{U}$ at $\boldsymbol{x}$. ρ_{data} and $\rho_{spatial}$ are cost functions penalising errors in the data and spatial terms. $\boldsymbol{U}$ is then estimated by minimising the total cost:

$$\sum_{All\ pixels\ \boldsymbol{x}} \left[c_{data}(x) + \alpha\, c_{spatial}(\boldsymbol{x})\right], \quad (3)$$

where α controls the relative importance of these two constraints.

$\boldsymbol{U}$ is estimated by an iterative gradient descent procedure: on each iteration Equation 3 is linearised and an update to $\boldsymbol{U}$ is computed. $\boldsymbol{U}$ is estimated on a sequence of downsampled images (an image pyramid), the results from each of which is used to initialise the next higher resolution level. This allows large motions (tens of pixels) to be tracked.

Horn and Schunck [33] used quadratic cost functions $\rho_{spatial}$ and ρ_{data}, however a key innovation in modern optical flow algorithms is to use robust cost functions which enable optical flow discontinuities at occlusion boundaries to be estimated correctly [31]. On the breast surface, these are no occlusion boundaries, so a cost function penalising large flow discontinuities (i.e. a quadratic $\rho_{spatial}$) is appropriate. Another important innovation in optical flow estimation is to apply a median filter to smooth optical flow updates, which suppresses artefacts introduced by image noise, while preserving flow discontinuities [31]. As before, there are no flow discontinuities on the breast, so a Gaussian filter is more appropriate.

A limitation of dense optical flow techniques is their

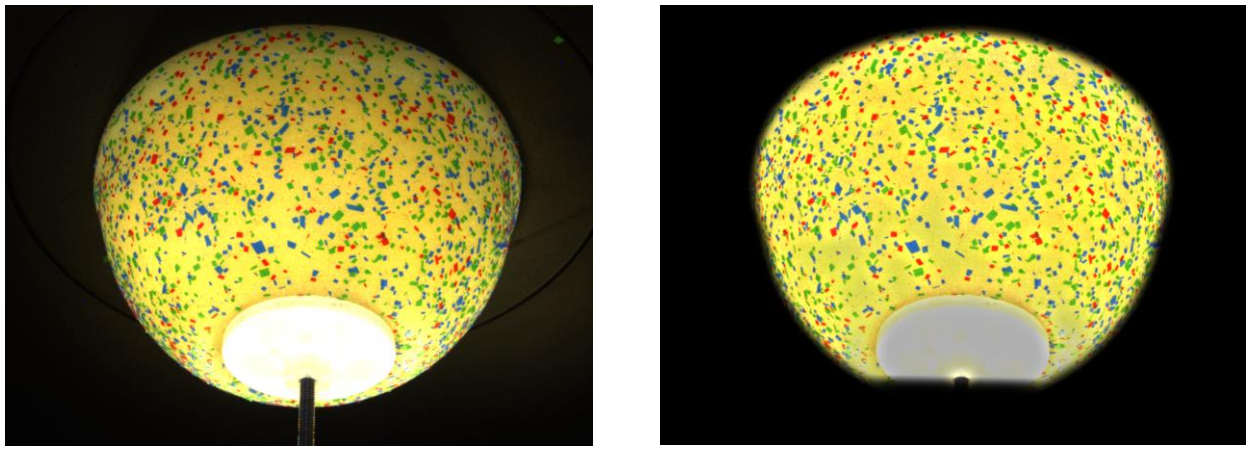


Figure 6: Silicone phantom breast with markers attached. Left: before correcting for lighting variation; right: after lighting correction.

sensitivity to lighting variation [31, 35]. Without prior knowledge of the scene, methods such as the Rudin-Osher-Fatemi method [39] are often applied to reduce the effects of lighting variation and image noise. The lighting variation is spatially smooth on the breast surface, so lighting levels across the breast are estimated by convolving the greyscaled image with a large Gaussian kernel. We then correct for lighting variation by scaling each pixel colour so that the lighting across the breast is uniform (Figure 6).

The optical flow implementation has several free parameters to set, including the filter sizes for spatially averaging updates and for removing the effect of lighting variation (equivalent to the scale on which variations in intensity are due to lighting variation rather than texture), and the parameter α controlling the relative importance of the spatial and data terms. These parameters are set using a grid-search. An image set not used elsewhere is used, and parameters for colour and greyscale are set separately.

The optical flow field is computed independently for each pair of consecutive frames, for every camera. The following section describes how these optical flow fields are combined with the 3D surfaces to estimate the 3D surface motion.

D. 3D surface motion reconstruction

The final stage of our proposed processing pipeline is to reconstruct the motion of a set of 3D points on the breast surface. We can reconstruct the motion of any 3D point on the breast surface by using another optimisation to fuse the optical flow and the 3D surface model.

Each 3D surface point has coordinates $\mathbb{X} = \{\boldsymbol{X_1}, \dots, \boldsymbol{X_N}\}$ over N timesteps. These points each lie on the corresponding surface S, the estimate of which has errors which are assumed to be independent and normally distributed with standard deviation σ_s. We seek $\mathbb{X}$ minimising:

$$c^2_{surface} = \sum_{i=1}^{N} \left(\frac{d(\boldsymbol{X_i}, S)}{\sigma_s} \right)^2 \tag{4}$$

where $d(\boldsymbol{X}_i, S)$ is the distance of point X_i from the surface S.

Each point's motion is also constrained by one camera's optical flow fields. This camera has a projection function π mapping points from 3D space to pixel coordinates. For each $\mathbb{X}$, this introduces the constraint:

$$c^2_{OF} = \sum_{i=1}^{N} \left(\frac{[\pi(X_{i+1 \bmod N}) - \pi(X_i)] - \boldsymbol{U}_i(\pi(X_i))}{\sigma_{OF}} \right)^2 \tag{5}$$

where σ_{OF} is the error in estimating the optical flow.

For each point track $\mathbb{X}$, X_1 is fixed and $\{X_2, \dots, X_N\}$ are optimised to minimise $c^2_{surface} + c^2_{OF}$. The smoothness constraints imposed earlier when estimating the surface and optical flow fields ensure that neighbouring points have similar motion. The motion of each of a set of 2500 random points is reconstructed in this way. These points provide input for the DIET diagnostic system.

E. Efficiency and implementation considerations

The system is implemented in C++, and makes extensive use of the Eigen matrix library [40] and the OpenCV computer vision library [41]. Reconstructing moving points on a breast takes 60 seconds on a Windows PC with an Intel i7 12-core processor running at 3.2GHz. This reconstruction includes segmenting 50 images sized 1600×1200, reconstructing 10 3D surface profiles, computing 50 optical flow fields, and reconstructing 2500 moving surface points. A similar time is required by each of these four stages, each of which could be further optimised to reduce the time required. Image sets captured at multiple frequencies for each patient are required; however data from different frequencies could be processed in parallel, enabling a diagnosis to be made within minutes of the images being captured.

IV. Data Collection and Accuracy Evaluation

In this paper we evaluate the system on data from human trials, on images simulated from human data, and on images of silicone phantom breasts.

Data from a preliminary human trial conducted in October 2010 is used. 16 patients receiving routine screening at Canterbury Breastcare volunteered to participate in the trial. All subjects were imaged with markers, and nine also had one breast imaged without markers. These nine datasets are used to evaluate the accuracy of the markerless system.

Silicone phantom breasts are used for the development of the DIET system. The phantom breasts are moulded from room temperature vulcanizing (RTV) silicones which are selected to match the elastic properties of human breasts. Inclusions of denser silicone are used to simulate tumours. The construction of the phantom breasts is described in detail in [42].

V. Quantifying Optical Flow Accuracy

For the human data, absolute ground truth is not available, but we can still quantify the errors in the optical flow by using the knowledge that the motion of the breast surface is cyclic. This section describes how.

The N images captured from each camera show the breast at N points throughout its cyclic motion. N optical flow fields are computed, between every pair of consecutive frames and between the Nth and first frames. If there were no errors in the optical flow, then a point tracked through the N optical flow fields should return to its start position. Because of errors in estimating the optical flow, the point's track does not return to exactly the same start point, as shown in Figure 7. We use this difference between the start and end positions of a set of tracks to estimate the error in the optical flow estimates.

A point on the breast is tracked through the N optical flow fields by accumulating N optical flow estimates $\{\hat{\mathbf{u}}^{\mathbf{1}}, \dots, \hat{\mathbf{u}}^{\mathbf{N}}\}$ along its path through the N images. We assume that each component of the optical flow estimate $\hat{\mathbf{u}}^{\boldsymbol{i}}$, $\hat{u}^i_j$, $j \in \{x, y\}$, is normally distributed about the true optical flow u^i_j with mean 0 and standard deviation $\lambda|u^i_j|$. Because the motion is cyclic,

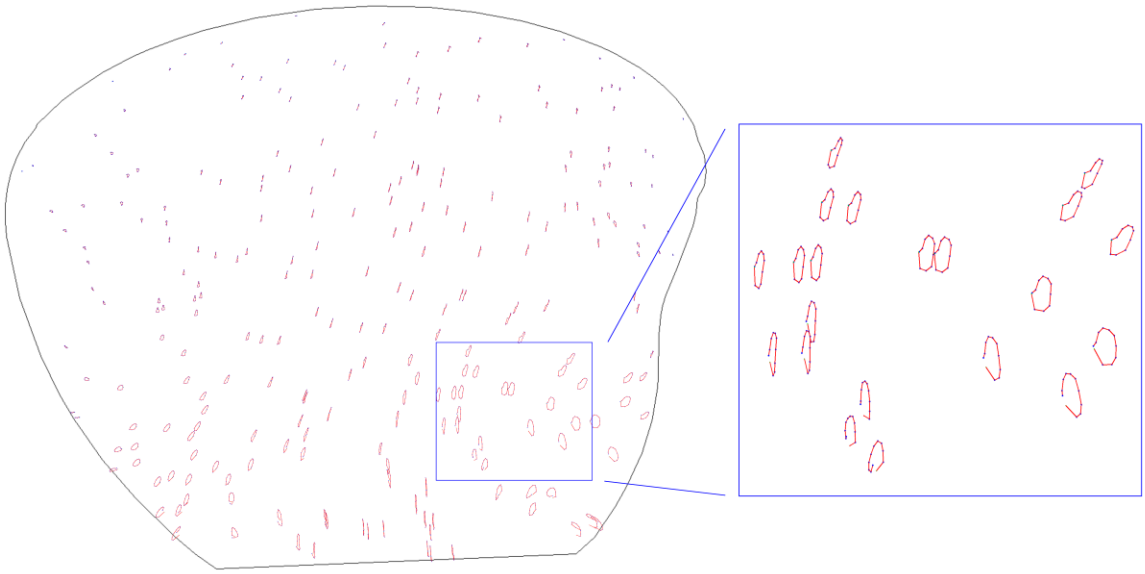


Figure 7: Points on the breast move back to the same position after *N* frames, but when we track a point through the *N* optical flow fields, the point does not return to exactly the same point because of errors in the optical flow. We use the sum of the optical flow vectors around these point tracks to quantify the error in the optical flow *(best viewed in colour)*.

each point on the breast moves back to exactly the same position after N frames, so $\sum_{i=1}^{N} \mathbf{u}^{\mathrm{i}} = 0$. We use this property to estimate the relative error λ in the optical flow as follows:

$$\hat{\mathrm{u}}_j^i \sim N\left(\mathrm{u}_j^i, \left(\lambda \mathrm{u}_{\mathrm{j}}^i\right)^2\right) \Rightarrow \sum_{i=1}^{N}\left(\hat{\mathrm{u}}_{\mathrm{j}}^i\right) \sim N\left(0, \lambda^2 \sum_{i=1}^{N}\left(\mathrm{u}_{\mathrm{j}}^i\right)^2\right) \quad (1)$$

If we assume $0 < \lambda \ll 1$, so that $\hat{\mathrm{u}}_{\mathrm{j}}^i \approx \mathrm{u}_{\mathrm{j}}^i$, then

$$s = \frac{\sum_{i=1}^{N} \hat{\mathrm{u}}_{\mathrm{j}}^i}{\sqrt{\sum_{i=1}^{N} (\hat{\mathrm{u}}_{\mathrm{j}}^i)^2}} \quad (2)$$

is a sample from $N(0, \lambda^2)$. By computing s at K points across the breast, $s_1, \ldots, s_K$, where we assume the errors to be independent, we estimate $\hat{\lambda} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(\mathrm{s_k})^2}$.

For each dataset, we compute s at $K = 250$ points across the breast, and hence estimate the relative error $\hat{\lambda}$.

A. *Simulating data to evaluate optical flow accuracy*

To verify that the errors in optical flow are being quantified correctly, we also evaluate the system on simulated image pairs, where the optical flow between the images is known. For one image from each dataset, we simulate an optical flow field and use it to warp the image to generate a second image. The warps are spatially variable, and have magnitude between 0 and 15 pixels. To simulate sensor noise, Gaussian noise with standard deviation of 5 greylevels is added to each pixel of each colour channel. To simulate lighting changes, a large specular reflection (50% brighter) is added to one image. The image noise, lighting change, and optical flow all have similar magnitude to those seen in real image pairs. We then compute the optical flow between the pairs of two images, and compute the relative error λ.

VI. Results

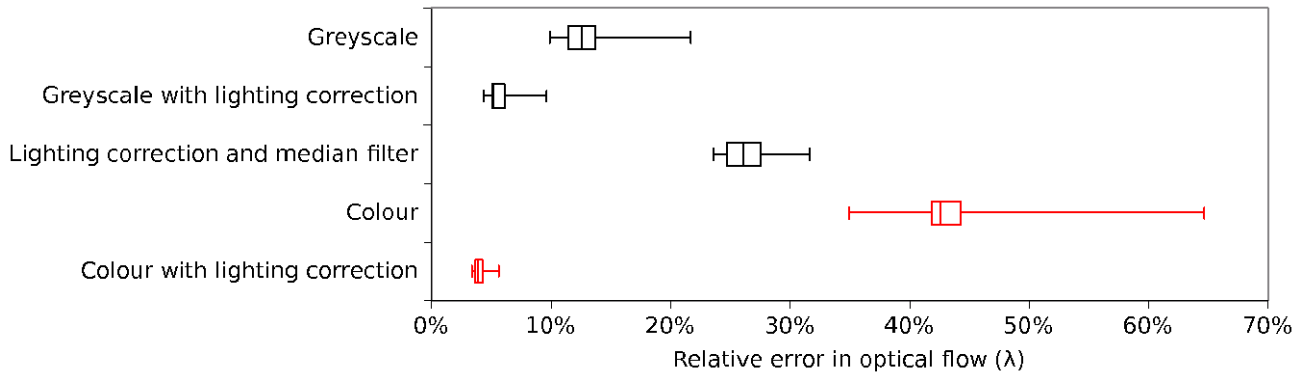
This section evaluates the accuracy of the proposed system, and compares it with the marker-based system. In Section VI.A, the accuracy of the 3D surface reconstruction is quantified. In Section VI.B, the accuracy of the optical flow estimation is estimated, and the effects of different implementation choices are evaluated. In Section VI.C, the surface motion reconstruction is compared with the surface motion computed by the marker tracking scheme. In Section VI.D, the system is used to detect tumours in silicone phantom breasts.

A. *3D surface reconstruction accuracy*

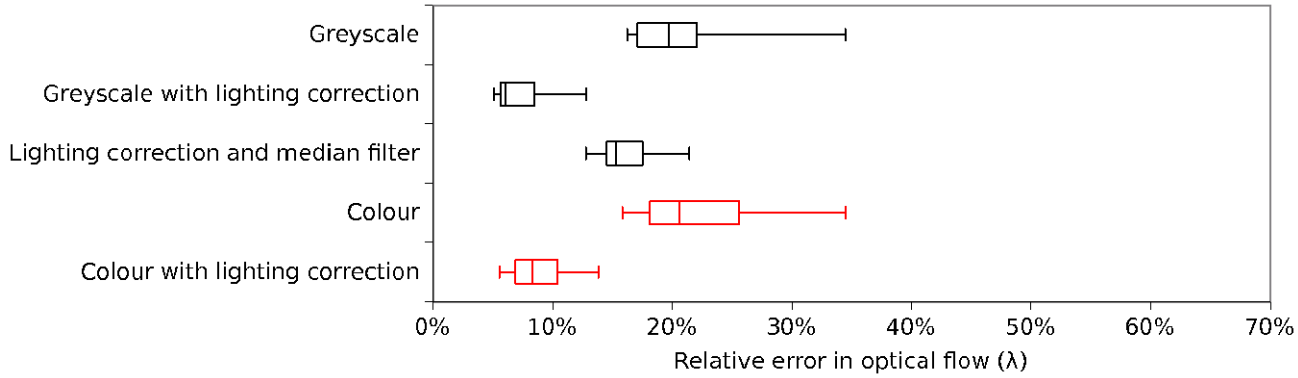
The accuracy of the 3D surface reconstruction is evaluated by refining the surface to fit only one side of each profile. The RMS error between the backprojected 3D surface and the other side of the profiles gives an estimate of the error in the surface.

The surface reconstruction accuracy is evaluated on the nine human datasets without markers. The breasts vary significantly in size and shape. The RMS error along the five half-profiles not included in the optimisation is 9.8 pixels, or 2.3mm, which is less than 2% of the breast diameter. When all ten constraints are included in the optimisation, surface height errors between profiles will be considerably lower.

A second measure of surface reconstruction accuracy is given by comparing breast volumes for the reconstructions from each of the N timesteps. We computed the standard deviation of the N breast volumes for each of the nine



(a) Effects on accuracy of different implementation choices, using simulated image pairs.



(b) Effects on accuracy of different implementation choices, inferred from cyclic motion of points in real data.

Figure 8: Accuracy of optical flow estimates (λ) on nine human breasts without markers with different optical flow implementation choices. Box-and-Whisker plots show the range, median, and interquartile range of the estimates from the nine breasts.

datasets. The average standard deviation is 1.8% of the breast volume. The breast volume is cubic in the measured surface heights, so assuming errors in computing the N breast volumes for each dataset are independent, the mean error in surface height is about 0.6%, or about 0.8mm for each breast.

B. Optical flow accuracy

This section quantifies the accuracy of the optical flow estimates, and provides justification for the different implementation choices made. Three implementation choices for the optical flow algorithm are considered: whether to use greyscale or colour images, whether to correct for lighting variations, and whether to use an edge-preserving median filter for spatial averaging of the optical flow field.

The relative errors in the optical flow are evaluated firstly on data simulated by warping real images, so that an absolute ground truth is available, and secondly on data from human trials, where the accuracy is inferred from the surface's cyclic motion. Figure 8 shows how different implementation choices affect accuracy. The choice with the largest impact is the lighting correction, which improves accuracy by an average of 69%. With lighting correction, the relative error λ averages 6%, corresponding to an average error of 0.13 pixels. On the real images, there is no advantage of using colour information—intensity variations are sufficient to capture the texture information present in the images, most likely because colour changes from Bayer interpolation artefacts and image noise are larger than variations from local differences in skin colour. On the simulated image pairs, colour information together with lighting correction gives substantially better performance than when using greyscale images; however the simulation is not modelling the Bayer imaging process, or the fact that colours measured by the camera vary with lighting levels. These factors account for the differences between real and simulated images.

For both real and simulated data the performance of a median filter for spatially averaging updates is relatively poor compared to a Gaussian filter, indicating that the spatial continuity of the optical flow field imposed by the Gaussian filter is effective at supressing the effects of noise.

C. Comparison with marker-based system

To compare the new system directly with the marker-based system, both were used to reconstruct points on 12 human datasets with markers. The point cloud computed by the marker-based system is compared with the surface and surface motion estimated by the new system.

The marker-based system reconstructs the 3D position of each marker at each timestep from its measured location in images from two cameras. Each marker is tracked over the N

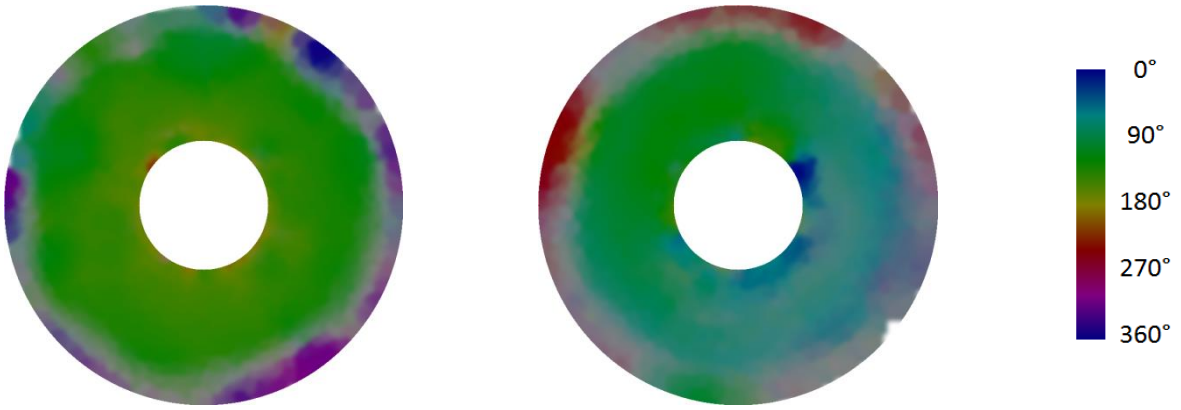


Figure 9: Breast surface motion for two silicone phantom breasts, f=39Hz. The phase difference from the actuator is indicated by hue. Left: healthy, Right: a 10mm tumour is present on the bottom-right side. The motion's magnitude is relatively small at the perimeters, and the variation in phase here is caused by boundary effects.

timesteps, giving N 3D points. Errors in localising markers lead to errors in reconstructed positions, particular in the depth direction. The marker-based system does not constrain points to fall on a smooth surface, so points lie in a cloud scattered around the true surface. We compared these points with the surfaces reconstructed by the proposed markerless system. For the 12 datasets, the mean difference between marker positions and the surface from the new system is 0.14mm, indicating that the difference between the surfaces estimated by the two systems is very small. The RMS distance from the surface is 1.30mm, indicating that the variation in point depths from the old system is considerably higher than the error in surface reconstruction in the new system.

The old marker locations are projected to the surface from the new system, and the positions of consecutive marker positions are compared to the computed optical flow vectors. The RMS difference in position is 1.9 pixels; this is large relative to the motion of points (approximately 4 pixels on average) and is large compared with typical errors in optical flow, but is biased by a small number of poorly localised or incorrectly matched markers. Our system to detect tumours based the phase of the surface motion averages motion across segments of the breast to reduce the effects of outliers and to remove the need for points to be constrained to the surface [4]. The new system explicitly constrains points to lie on a surface, and the motion of nearby points to be consistent, and hence the motion of individual points is more accurate. The new system's increased accuracy on a small scale will help with the development of finite element-based tools to detect tumours, which require a higher density of surface motion measurements [5].

D. Locating tumours in phantom breasts

Previous work on the DIET system with markers has indicated that variation in the phase of the surface motion indicate the presence or absence of a tumour [3]. Figure 9 shows plots of the phase difference with respect to the actuator in the ϕ direction for a healthy silicone phantom breast, and a phantom breast with a 10mm diameter tumour. Close to the tumour, the surface motion has a smaller phase difference, because the breast is stiffer here, and the induced wave moves more quickly through the breast. Figure 10 shows that the faster motion is present near to the tumour at a range of frequencies.

VII. Conclusion

This paper has demonstrated a computer vision system for measuring the 3D surface motion of a breast which is being mechanically vibrated in the DIET breast cancer screening machine. The measured surface motion will be used to infer the breast's internal stiffness, and hence to identify tumours. The proposed system uses a model-based shape-from-silhouette method to reconstruct a 3D model of the breast's surface, and uses a modern dense optical flow algorithm to estimates the skin surface motion between pairs of frames. The 3D surface and the optical flow are fused in order to reconstruct the skin surface motion.

Results indicate that the proposed system can reconstruct the breast surface with average errors of less than one millimetre, and the optical flow fields have an average relative error of 6%, corresponding to subpixel accuracy. A comparison with a previous marker-based system shows that the new system more accurately estimates the skin surface motion. Plots of the phase of the breast surface motion show the presence of a 10mm tumour in a phantom breast. The new system demonstrates that the time and inconvenience needed to apply markers to the breast are unnecessary for an elastography-based breast cancer screening machine.

References

[1] A. Peters, A. Milsant, J. Rouzé, L. Ray, J. Chase, and E. Van Houten, "Digital image-based elasto-tomography: proof of concept studies for surface based mechanical property reconstruction," *JSME International Journal Series C*, vol. 47, no. 4, pp. 1117–1123, 2004.

[2] R. Brown, J. Chase, and C. Hann, "A pointwise smooth surface stereo reconstruction algorithm without correspondences," *Image and Vision Computing*, 2012.

[3] T. Lotz, P. D. Simpson, D. Stocker, C. E. Hann, and J. G. Chase, "In vitro evaluation of surface based non-invasive

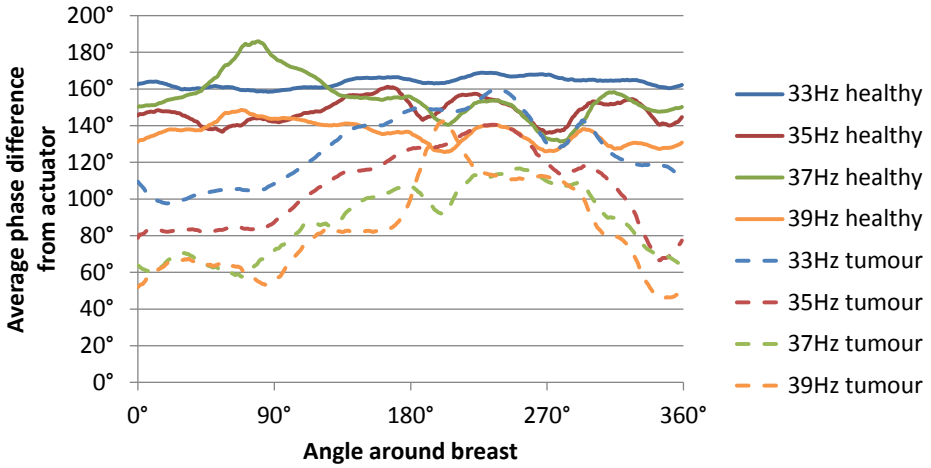


Figure 10: Average phase difference from actuator around two phantom breasts which are designed to be mechanically similar to human breasts. One breast has a 10mm tumour at 45°. The motion travels more quickly through the breast close to where a tumour is located, hence the phase difference from the actuator is smaller.

breast cancer screening with digital image based elasto tomography (DIET)," in *Proc. International Conference of the Engineering in Medicine and Biology Society (EMBC)*, 2010, pp. 3077–3080.

[4] E. Van Houten, H. Kershaw, T. Lotz, and J. G. Chase, "Localization and detection of breast cancer tumors with digital image elasto-tomography," in *Proc. International Conference of the Engineering in Medicine and Biology Society (EMBC)*, 2012, pp. 2635–2638.

[5] A. Peters, J. G. Chase, and E. E. W. Van Houten, "Digital image elasto-tomography: combinatorial and hybrid optimization algorithms for shape-based elastic property reconstruction," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 11, pp. 2575–2583, 2008.

[6] B. S. Garra, E. I. Cespedes, J. Ophir, S. R. Spratt, R. A. Zuurbier, C. M. Magnant, and M. F. Pennanen, "Elastography of breast lesions: initial clinical results." *Radiology*, vol. 202, no. 1, pp. 79–86, 1997.

[7] A. Itoh, E. Ueno, E. Tohno, H. Kamma, H. Takahashi, T. Shiina, M. Yamakawa, and T. Matsumura, "Breast disease: Clinical application of us elastography for diagnosis," *Radiology*, vol. 239, no. 2, pp. 341–350, 2006.

[8] D. B. Plewes, J. Bishop, A. Samani, and J. Sciarretta, "Visualization and quantification of breast cancer biomechanical properties with magnetic resonance elastography," *Physics in Medicine and Biology*, vol. 45, no. 6, p. 1591, 2000.

[9] M. De La Gorce, D. Fleet, and N. Paragios, "Model-based 3D hand pose estimation from monocular video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 9, pp. 1793–1805, 2011.

[10] J. Rehg and T. Kanade, "Model-based tracking of self-occluding articulated objects," in *International Conference on Computer Vision*, 1995, pp. 612–617.

[11] J. Starck and A. Hilton, "Model-based multiple view reconstruction of people," in *International Conference on Computer Vision*, 2003, pp. 915–922.

[12] G. Zheng, M. Ballester, M. Styner, and L. Nolte, "Reconstruction of patient-specific 3D bone surface from 2D calibrated fluoroscopic images and point distribution model," in *Medical Image Computing and Computer-Assisted Intervention*, 2006, pp. 25–32.

[13] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, UK: Cambridge University Press, 2003.

[14] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment – a modern synthesis," in *Proc. International Workshop on Vision Algorithms: Theory and Practice*, vol. 1883/2000, Corfu, Greece, 1999, pp. 1–71.

[15] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1-3, pp. 7–42, 2002.

[16] R. Szeliski, *Computer vision: algorithms and applications*. Springer, 2010.

[17] E. Prados and O. Faugeras, "Shape from shading," *Handbook of mathematical models in computer vision*, pp. 375–388, 2006.

[18] W. Miled, J. Pesquet, and M. Parent, "Dense disparity estimation from stereo images," 2006.

[19] O. Woodford, P. Torr, I. Reid, and A. Fitzgibbon, "Global stereo reconstruction under second-order smoothness priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 12, pp. 2115–2128, 2009.

[20] A. Laurentini, "The visual hull concept for silhouette-based image understanding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 2, pp. 150–162, 1994.

[21] M. Prasad and A. Fitzgibbon, "Single view reconstruction of curved surfaces," in *Proc. Computer Vision and Pattern Recognition*, vol. 2, 2006, pp. 1345–1354.

[22] D. Terzopoulos, A. Witkin, and M. Kass, "Symmetry-seeking models and 3d object reconstruction," *International Journal of Computer Vision*, vol. 1, no. 3, pp. 211–221, 1988.

[23] M. Jones and J. Oakley, "Efficient representation of object shape for silhouette intersection," in *Vision, Image and Signal Processing*, vol. 142, no. 6, 1995, pp. 359–365.

[24] L. H. Staib and J. S. Duncan, "Model-based deformable surface finding for medical images," *IEEE Transactions on Medical Imaging*, vol. 15, no. 5, pp. 720–731, 1996.

[25] L. Ballan, A. Taneja, J. Gall, L. Van Gool, and M. Pollefeys, "Motion capture of hands in action using discriminative salient points," in *European Conference on Computer Vision (ECCV)*, 2012, pp. 640–653.
[26] L. Kovacs, M. Eder, R. Hollweck, A. Zimmermann, M. Settles, A. Schneider, M. Endlich, A. Mueller, K. Schwenzer-Zimmerer, N. Papadopulos *et al.*, "Comparison between breast volume measurement using 3D surface imaging and classical techniques," *The Breast*, vol. 16, no. 2, pp. 137–145, 2007.
[27] N. Isogai, K. Sai, H. Kamiishi, M. Watatani, H. Inui, and H. Shiozaki, "Quantitative analysis of the reconstructed breast using a 3-dimensional laser light scanner," *Annals of plastic surgery*, vol. 56, no. 3, pp. 237–242, 2006.
[28] T. Carter, "Biomechanical modelling of the breast for image-guided surgery," Ph.D. dissertation, University College London, 2009.
[29] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. International Joint Conference on Artificial Intelligence*, 1981, pp. 674–679.
[30] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *International journal of computer vision*, vol. 12, no. 1, pp. 43–77, 1994.
[31] D. Sun, S. Roth, and M. Black, "Secrets of optical flow estimation and their principles," in *Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2432–2439.
[32] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *International Journal of Computer Vision*, vol. 92, no. 1, pp. 1–31, 2011.
[33] B. Horn and B. Schunck, "Determining optical flow," *Artificial intelligence*, vol. 17, no. 1, pp. 185–203, 1981.
[34] D. Bradley, W. Heidrich, T. Popa, and A. Sheffer, "High resolution passive facial performance capture," *ACM Transactions on Graphics (TOG)*, vol. 29, no. 4, p. 41, 2010.
[35] W. R. Crum, T. Hartkens, and D. L. G. Hill, "Non-rigid image registration: theory and practice," *British journal of radiology*, vol. 77, no. suppl 2, pp. S140–S153, 2004.
[36] J. Bentley, "Programming pearls: algorithm design techniques," *Communications of the ACM*, vol. 27, no. 9, pp. 865–873, 1984.
[37] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International journal of computer vision*, vol. 1, no. 4, pp. 321–331, 1988.
[38] T. Botterill, S. Mills, R. Green, and T. Lotz, "Optimising light source positions to minimise illumination variation for 3D vision," in *Proc. 3DIMPVT*, 2012, pp. 1–8.
[39] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.
[40] G. Guennebaud and B. Jacob, "Eigen 3 matrix library," n.d., http://eigen.tuxfamily.org/.
[41] OpenCV Computer Vision Library, n.d., http://opencv.org/.
[42] A. S. Kashif, "Imaging technology for digital image based motion detection in the diet breast cancer screening system," Ph.D. dissertation, University of Canterbury, 2013.