Surface Reconstruction for a DIET breast cancer screening system

Richard G. Brown*, Christopher E. Hann*,
J. Geoffrey Chase*, Larry E. Ray**

* Centre of Bioengineering, University of Canterbury, Christchurch,
New Zealand.
** Eastman Kodak Company, Kodak Research Laboratory, Rochester,
New York, USA.

Abstract: This paper outlines a method for accurately tracking surface motion on an actuated breast as part of a Digital Image-based Elasto-Tomography (DIET) breast cancer screening project. The tracking method is validated in both simulation and experiment on a silicon breast phantom which has similar elasticity characteristics to human breast tissue. Given mean reprojection error ranges between 0.3 and 0.5 pixels for 2M pixel cameras the end result is sub-millimetre surface tracking of a silicon phantom actuated at 100 Hz. This highly accurate, fast and low cost method of tracking is thus well suited to a DIET system.

1. INTRODUCTION

Breast cancer is a significant health problem in both developed and developing countries. It is estimated that each year the disease is diagnosed in over one million women worldwide and is the cause of death in over 400,000 women Fentiman [2002]. In New Zealand, breast cancer accounts for the highest mortality rate of all cancers among women and it has the sixth highest death rate out of 173 developed countries.

There are many treatment options available, including surgery, chemotherapy, radiation therapy, and hormonal therapy. These treatments are significantly more effective in reducing the mortality of the disease with early detection through breast cancer screening programmes. The most common method for early detection of breast cancer is mammography. However, mammography can cause significant discomfort to the patient and requires radiation exposure, a further health concern.

Digital Image-based Elasto-Tomography (DIET) is an emerging technology for non-invasive breast cancer screening. The DIET system uses digital imaging of an actuated breast surface to determine tissue surface motion from a specified input. It then reconstructs the 3D internal tissue stiffness distribution from that motion. Regions of high stiffness suggest cancer since cancerous tissue is between 3 and 10 times stiffer than healthy tissue in the breast Samani et al. [2003], Krouskop et al. [1998], Wellman and Howe [2000].

This approach eliminates the need for X-Rays and excessive, potentially painful compression of the breast Peters et al. [2005] as required in a mammogram. Hence, screening could start much younger and might enjoy greater compliance Robertson [2005].

Presently, there are other Elasto-Tomographic methods based on magnetic resonance Oida et al. [2004] and ultrasound Maurice et al. [2004] modalities. Both methods are capable of measuring the tissue elasticity and are undergoing rapid development. However, they are also costly in terms of equipment and take significant time to use. They are therefore limited for practical screening applications.

The DIET system, in contrast, is computer-based and is thus potentially low cost and portable, so the technology could be used in any medical centre, particularly in remote areas. In addition, the use of silicon technology ensures that as it improves and scales upward in capability so will the DIET system performance. This scalability of performance is not true for X-Ray or ultrasound based approaches.

This paper is focussed on the imaging side of the DIET system and tests the ability to accurately capture actuated tissue motion on realistically shaped breast silicon phantoms.

2. BACKGROUND

2.1 DIET project

The DIET (Digital Image-based ElastoTomography) concept is a new methodology for breast cancer screening. The breast is mechanically actuated sinusoidally at a low frequency (around 50-100Hz). This motion is visible as a wave pattern on the breast surface. With a linear elasticity model, knowledge of the breast surface motion in 3D is theoretically sufficient information to reconstruct the stiffness distribution of the interior of the breast, and thus, because cancerous tumours are around an order of magnitude stiffer than healthy tissue, could be used to diagnose breast cancer. In practice the elasticity reconstruction will be performed by a parametrised inverse problem.

The surface motion is captured by using a combination of strobed lighting and a number of calibrated digital cameras. Motion in 3D is computed by matching moving points between different cameras, and using computer

vision techniques to construct the corresponding points in 3D.

2.2 Camera Model

Digital CCD cameras can be accurately modeled as perspective projection pinhole cameras. The 3D world space $\mathbb{R}^3$ can be embedded in projective space $\mathbb{P}^3$ with homogeneous coordinates (X, Y, Z, W) and the image coordinate space $\mathbb{R}^2$ can be embedded in the projective plane $\mathbb{P}^2$ with homogeneous coordinates (u, v, w). The corresponding coordinates in $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively, are given by $(\frac{X}{W}, \frac{Y}{W}, \frac{Z}{W})$ and $(\frac{u}{w}, \frac{v}{w})$. Similarly, measured 3D coordinates and 2D image coordinates can be embedded in $\mathbb{P}^3$ and $\mathbb{P}^2$ respectively by the maps $(X, Y, Z) \mapsto (X, Y, Z, 1)$ and $(u, v) \mapsto (u, v, 1)$.

The camera represents a projection between $\mathbb{P}^3$ and $\mathbb{P}^2$, which can be represented by a homogeneous matrix $P \in \mathbb{R}^{3 \times 4}$, whose kernel is the projection centre of the camera. The projection from world coordinates $\mathbf{X} = (X, Y, Z, W)$ to image coordinates $\mathbf{u} = (u, v, w)$ is then described by the linear equation

$$\lambda \mathbf{u} = P\mathbf{X} \qquad (1)$$

where λ is a nonzero scalar. The calibration matrix $\mathbf{P}$ can be factored into $P = K[R\,T]$ where K is an upper triangular matrix representing the intrinsic camera parameters, $R \in \mathbb{R}^{3 \times 3}$ is a rotation matrix describing the relative orientations of the camera and world frames, and $T \in \mathbb{R}^3$ is the origin of the world coordinate system in the camera frame. A camera is said to be fully calibrated if K, R, T and hence P are known.

3. METHOD

3.1 Fiducial system

Human skin lacks high-contrast features which are easily extractable from digital images. In order to reconstruct individual points in 3D, points on the surface must be able to be extracted from images taken from different cameras, i.e. they must be viewpoint invariant. The easiest way to define such features is to introduce artificial fiducials to the surface. For reasons that will be described in the following sections, the fiducials used in this paper are randomly applied identical points in three colours, red, green, and blue.

3.2 Feature tracking

The features are tracked using the novel Euclidean-invariant algorithm described in Brown et al. [2007] which is based on a system of coloured fiducial points as mentioned in the previous section. This method of feature tracking uses the geometrically invariant properties of local configurations of the coloured point locations to match points between frames, rather than using image-based correlation techniques, the advantage being that points can be matched over a wider range of transformations, notably those involving a large translational component. As well, a significant increase in computational speed is

observed. The output of this process is the path of each fully tracked point over one phase of motion in each set of images. Note that because the breast is well-modelled with a linear stiffness model, motion of each point on the breast surface resulting from sinuosidal actuation will be an ellipse, and hence the paths in images from each camera will also be elliptical. This property can be used to decrease the number of frames necessary to identify the motion path of each point. Elliptical features can also be used as primitives for the correspondence problem, see Quan [1996], however the ellipses are too small for this to be of practical use in this problem.

3.3 3D Reconstruction

The main contribution of this paper is a method of reconstructing the coordinates of a dense set of points on the surface in 3D based on their image locations. Conventional approaches to this problem involve finding means of identifying corresponding points between images, and then triangulating these to find the resulting 3D position of the point. In this paper a somewhat alternative approach is utilised. The epipolar constraint is used to find a set of candidate matches for each point. All potential matching pairs are then reconstructed in 3D, giving a large point cloud, contained within which is the desired surface. With suitable point density and surface smoothness, the desired 3D surface can be extracted from this set of 3D points.

Epipolar constraint   Consider two images I_1 and I_2 from two fully calibrated cameras. A point p in I_1 can be represented in homogeneous coordinates as $\mathbf{p} = (p_1, p_2, p_3)$. The geometry of the two-camera configuration constrains any potential match for p in I_2 to fall upon an *epipolar line* $\mathbf{l}$ Hartley and Zisserman [2004]. This constraint can be represented algebraically by the following matrix equation

$$\mathbf{p}^T F \mathbf{q} = 0 \qquad (2)$$

where $\mathbf{p}, \mathbf{q}$ are the homogeneous coordinates of p and q, F is a rank 2 matrix known as the *Fundamental Matrix* which encapsulates the epipolar geometry and can be computed from the two camera projection matrices. The homogeneous coordinates of the epipolar line $\mathbf{l}$ are given by $\mathbf{l} = \mathbf{p}^T F$, i.e. Eq. 2 becomes $\mathbf{l}\mathbf{q} = 0$. In practice, points $\mathbf{p}$ and $\mathbf{q}$ are said to satisfy the epipolar constraint if

$$\mathbf{p}^T F \mathbf{q} < \epsilon_{\text{ep}} \qquad (3)$$

for some threshold ϵ_{ep} and appropriate scaling of $\mathbf{l}$ such that Eq. 3 represents the perpendicular distance between q and $\mathbf{l}$. ϵ_{ep} is determined by the error in calibration and image point measurement.

3D Reconstruction procedure   We assume that the points on the surface are sufficiently dense that the points within a certain radius r of each surface point can be well approximated by a plane. More concretely, let $\mathbf{X}^1 = [X^1, Y^1, Z^1, 1]^T, \ldots, \mathbf{X}^n = [X^n, Y^n, Z^n, 1]^T$ be the points within Euclidean distance r of a point $\mathbf{X}^0$, and denote this collection of points the *r-neighbourhood* of $\mathbf{X}^0$. Let $\pi = [\pi_1, \pi_2, \pi_3, \pi_4]$, with $||[\pi_1, \pi_2, \pi_3]||_2 = 1$, represent the least squares plane fit with respect to perpendicular distance of that set of points, i.e. representing the plane equation $\pi \mathbf{X} = 0$. A measure of how well π approximates $\mathbf{X}^0, \ldots, \mathbf{X}^n$ is

$$e_\pi(\mathbf{X}^0, r) = ||\pi[\mathbf{X}^0, \ldots, \mathbf{X}^n]||_\infty \qquad (4)$$

i.e. the maximum perpendicular distance from the plane to one of the data points. For a given smooth surface, r can be chosen such that

$$\max_{\mathbf{X}^0} e_\pi(\mathbf{X}^0, r) < e_{\max} \qquad (5)$$

and so that for most points on the surface the number of points in the *r-neighbourhood* $n > n_0$, where n_0 is some positive integer (around 10-20 is good).

For a given threshold, ϵ_{ep}, all pairs of points satisfying the epipolar constraint (Eq.3) are computed, and the 3D location of the point corresponding to each pair is calculated using the DLT triangulation method described in Hartley and Zisserman [2004].

The resulting cluster of 3D points contains the surface that is being sought. If the density of points on the surface is sufficiently high and ϵ_{ep} is sufficiently low, then for each point on the surface, for a suitable choice of r the *r-neighbourhood* of that point will contain mostly other points on the surface; on average a certain proportion ρ will be surface points. In this event, a plane can be robustly fitted to the *r-neighbourhood* using, for instance, RANSAC, and all of the inliers to this plane marked as surface points, where some multiple of $e_{\max}$ can be used as the threshold for being an inlier. If RANSAC detects a large number of inliers in this way, the point along with the inliers can be categorised as surface points.

For the points not on the surface, there will be a much smaller proportion than ρ of surface points in the *r-neighbourhood*. Moreover the points in its *r-neighbourhood* will not sit nicely on a plane. The failure of RANSAC to fit a plane with a large number of inliers can be recognised and allows the point to be categorised as an outlier.

The procedure is repeated until all points have been categorised.

4. RECONSTRUCTION - COMPUTER SIMULATION

A finite element of an elastic cylinder with similar elastic properties to a human breast with a high stiffness inclusion was used to generate realistic data. The model was developed by Ashton Peters, from the University of Canterbury. The 17000 node finite element mesh was produced in GAMBIT, and the model simulated in Fortran 90, with the matrix inversion being done with the direct sparse matrix inversion package, MUMPS. The model was simulated with a 50Hz actuation frequency, with a 0.5mm peak to peak amplitude. See Figure 1 for a rendering of the cylinder.

400 Points, 80 red, 160 green, and 160 blue were projected onto one quarter of the surface of the cylinder. The motion of the finite element nodes was interpolated onto these points. Two camera models were set up 90° apart at a distance of 30cm from the cylinder. The resulting image point locations were computed over one phase of the actuation, with 20 frames. Gaussian noise was added to the measured image point locations, with standard deviation in each orthogonal direction of 0.2 pixels. This value is reasonable, as the image point locations in practice will
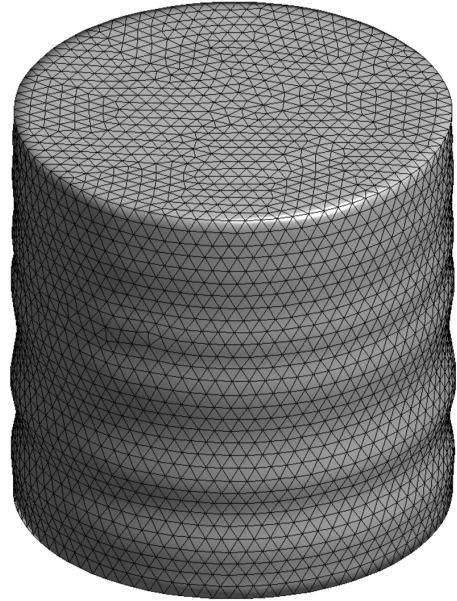

Fig. 1. Finite element cylinder model used for simulation

be determined by taking the centroid of a blob of $O(100)$ pixels. The expected measurement error should therefore be well under one pixel.

The points were then tracked using the algorithm described in Section 3 over 20 frames, including the last frame to the first frame, giving elliptical contours. Those points that were not successfully tracked through all 20 frames (i.e. the 20th point in the sequence doesn't track back to 1) were discarded. The reconstructed points are shown in Figure 2.

5. RECONSTRUCTION - LABORATORY EXPERIMENT

5.1 Experimental setup

Gel phantom   A gel phantom was created from molding silicon gel with similar elasticity characteristics to human breast tissue. The silicon was pigmented to yield a more natural flesh colour. Small red, green, and blue points were randomly applied to the surface. See Fig. 3 for a picture.

Actuator   The actuator is a voice-coil style electromechanical actuator, and provides sinusoidal actuation at frequencies between 20 and 200 Hz at amplitudes of up to 1mm. Amplitude is controlled via feedback from an LVDT position sensor from a dSpace real time control environment.

Cameras and lighting   The cameras are standard consumer Canon Powershot G5 cameras operating at 2 megapixels. Each camera is fitted with an LED ring flash which can be synchronised with an external trigger signal (see Fig. 4). These LED flashes are strobed at the same frequency as the actuation frequency via a control signal from the dSpace control system. The phase of the strobe signal can be varied with respect to the actuator signal,
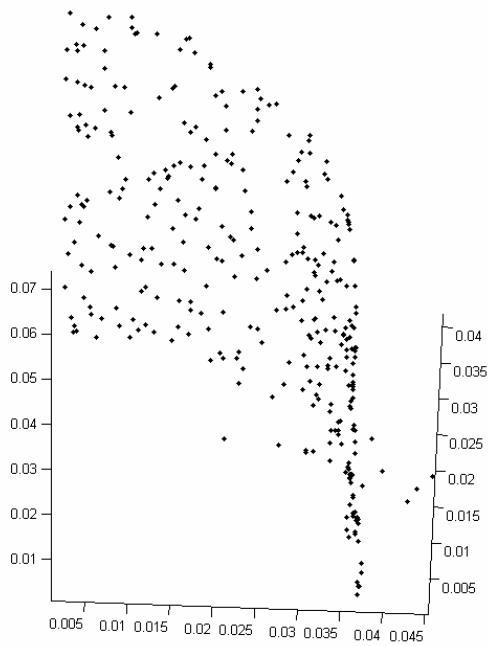
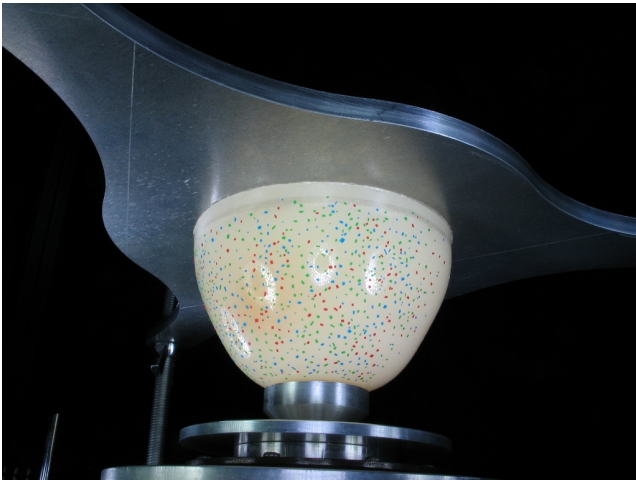Fig. 2. Reconstructed surface

Fig. 3. Silicon phantom under actuation with random coloured dots applied

allowing images to be taken at arbitrary points in the actuation cycle.

Software   The dSpace control system is designed in Simulink, and has a real-time interface in Control Desk. Custom software was written to automatically coordinate the taking of the images and the adjusting of the strobe phase.

Experiment parameters   For the experiment presented in this paper, the phantom was actuated at 80Hz, with an amplitude of 0.75mm. 20 images were taken at even phase increments over the cycle from each camera.
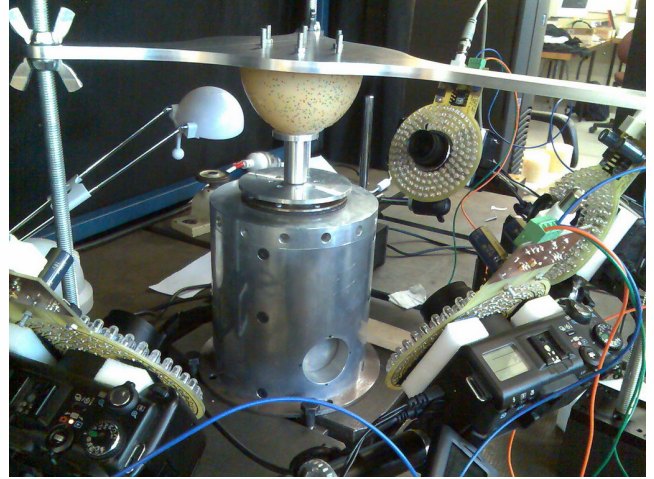

Fig. 4. Experimental setup: Actuator, gel phantom, and four cameras fitted with LED ring flashes

Calibration   In order to reconstruct the surface, it is required that the camera is fully calibrated. The projection matrix P can be computed from correspondences between known world locations of a number of points on a calibration object and their corresponding image locations. P is estimated using nonlinear least squares minimising the reprojection error. See Hartley and Zisserman [2004] for details.

The calibration object used for this experiment is an precisely machined 54mm^3 anodised aluminium die. The point locations used to compute the projection matrix are the 6 vertices on the silhouette of the cube, and the centroids of the die "dots". Note that because the dots are circular, images of these will be ellipses, and the centroid is viewpoint (projectively) invariant. These dots can also be used to uniquely identify which faces are visible, and hence calibration is fully automatic from a single image of the cube. See Fig. 5 for images of the calibration cube used.
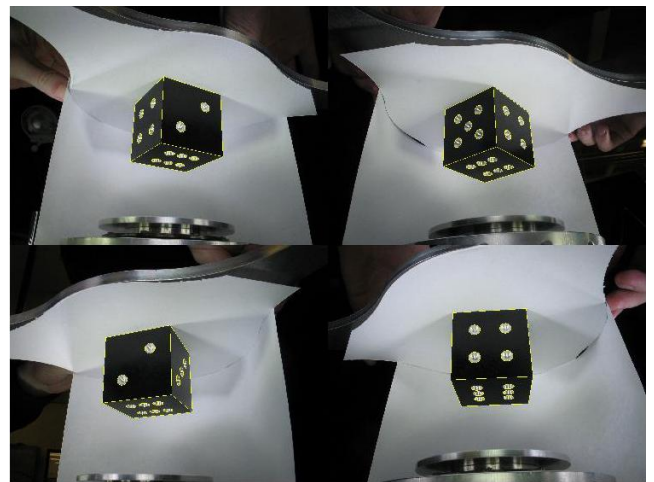

Fig. 5. Calibration: Images of the calibration die from the four cameras

Error in calibration is measured by the mean *reprojection error*, i.e. the world locations of the visible points are projected onto the images, and the mean Euclidean dis-

tance between these projected points and their measured location is taken to be the measure of error. For these experiments, the mean reprojection error ranges between 0.3 and 0.5 pixels, where the images used are 2 megapixel images.

Feature tracking   The coloured points are extracted by applying simple colour thresholds to the individual RGB channels of the images. The point locations are taken to be the centroids of the detected coloured blobs. The feature locations are then tracked using the Euclidean Invariant algorithm described in Brown et al. [2007]. The motion of the feature points from one camera is depicted in Fig. 6.
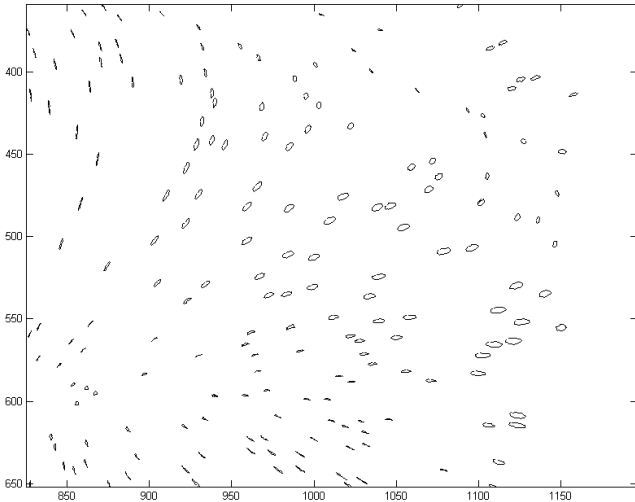

Fig. 6. Close up of the tracked motion from one camera

Another subsection

5.2 3D Reconstruction

3D reconstruction was performed on the data points from the first frame. Parameters used were $\epsilon_{ep} = 2$ pixels, $r = 10$ mm, and $e_{max} = 0.5$ mm. Plane fitting for *r-neighbourhoods* was only considered if the number of neighbours $n \geq 10$. Points were reconstructed in 3D by performing 3D reconstructions for each of the 3 adjacent camera pairs individually and collating the results.

An example of the point cloud created by reconstructing all potential matches from the epipolar constraint stage is depicted in Fig. 7. Because we have points of three different colours, we can eliminate a number of these mismatches, by only allowing potential matches to have the same colour. The same resulting point cloud is shown in Fig. 8. The effect of this is to significantly decrease the density of points which are not on the surface.

The outlier rejection algorithm was then applied to the point cloud. For the cases considered, the surface was successfully reconstructed both with and without the colour information. The case with the colour taken into account performed significantly faster, however. The resulting reconstructed surface is shown in Fig. 9.
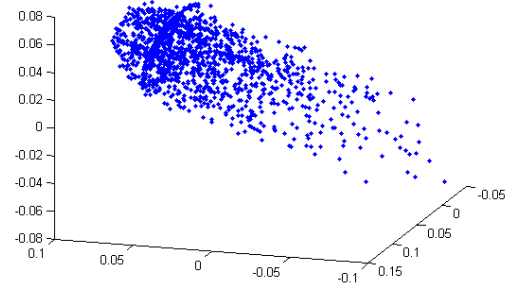

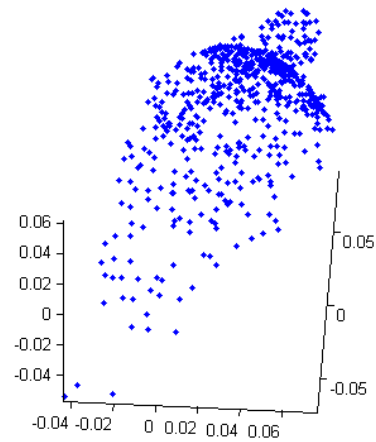Fig. 7. Point cloud from epipolar constraint matches for one camera pair, disregarding colour

Fig. 8. Point cloud from epipolar constraint matches for one camera pair, including colour information
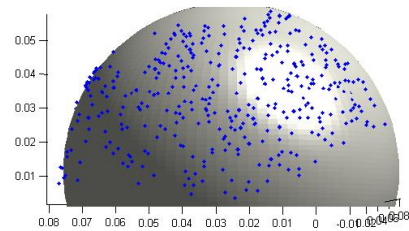

Fig. 9. Reconstructed 3D surface points, with a least squares sphere in the background to give some 3D context

6. CONCLUSION

This paper implemented an algorithm for reconstructing surface motion from digital images of an actuated gel phantom. All the key issues of Digital image acquisition, camera calibration, surface reconstruction and point tracking were addressed. The end result is highly accurate tissue surface motion tracking which would then go into a finite element based inverse problem that identifies the tissue distribution of the phantom. In the case of a breast, regions of high stiffness would suggest a tumour.

High frequencies of 100 Hz were captured using a combination of standard digital cameras and a stroboscope providing a simple cost-effective approach. Overall the results show good potential for practical implementation in a DIET system with the potential for low cost and portable breast cancer screening.

REFERENCES

R. G. Brown, C. E. Hann, J. G. Chase, and L. A. Ray. Discrete color-based Euclidean-invariant signatures for feature tracking in a DIET breast cancer screening system. In *Medical Imaging 2007: Physiology, Function, and Structure from Medical Images. Edited by Manduca, Armando; Hu, Xiaoping P.. Proceedings of the SPIE, Volume 6511, pp. 65110D (2007).*, volume 6511, March 2007.

I. S. Fentiman. Breast cancer in older women. *Breast Cancer Online*, 5, 2002.

R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, ISBN: 0521540518, second edition, 2004.

T. A. Krouskop, T. M. Wheeler, F. Kallel, B. S. Garra, and T. Hall. Elastic moduli of breast and prostate tissues under compression. *Ultrasonic Imaging*, 20:2183–2198, 1998.

R. L. Maurice, J. Ohayon, Y. Fretigny, M. Bertrand, G. Soulez, and G. Cloutier. Non-invasive vascular elastography: theoretical framework. *IEEE Transactions on Medical Imaging*, 23:164–180, 2004.

T. Oida, A. Amano, and T. Matsuda. Magnetic resonance elastography: in vivo measurements of elasticity for human tissue. *International Conference in Informatics Research for Development of Knowledge Society Infrastructure*, 2004.

A. Peters, A. Milsant, J. Rouze, L. Ray, J. G. Chase, and E. E. W. Houten. Digital image-based elasto-tomography: Proof of concept studies for surface based mechanical property reconstruction. *Japanese Society of Mechanical Engineers (JSME)*, 47:1117–1123, 2005.

L. Quan. Conic reconstruction and correspondence from two views. *IEEE Proceedings on Pattern Matching and Machine Intelligence*, 18(2):151–160, 1996.

M. Robertson. Commercialisation and ip strategies for digital image-based elasto-tomography (diet). Master's thesis, University of Canterbury, 2005.

A. Samani, J. Bishop, C. Luginbuhl, and D. B. Plewes. Measuring the elastic modulus of ex-vivo small tissue samples. *Phys. Med. Biol.*, 48:2183–2198, 2003.

P. W. Wellman and R. D. Howe. Breast tissue stiffness in compression is correlated to histological diagnosis. *Harvard BioRobotics Laboratory Technical Report*, 2000.