

Discrete Colour-based Euclidean-Invariant Signatures for Feature Tracking in a DIET Breast Cancer Screening System

R. G. Brown^a, C. E. Hann^a, J. Geoffrey Chase^a, L. E. Ray^b

^aDepartment of Mechanical Engineering, University of Canterbury, P.O. Box 4800, Christchurch, New Zealand;

^bEastman Kodak Company, Kodak Research Laboratories, Rochester, New York, USA

ABSTRACT

A Digital Image-based Elasto-Tomography (DIET) system for breast cancer screening has been proposed in which the elastic properties of breast tissue are recovered by solving an inverse problem on the surface motion of a breast under low frequency (50-100 Hz) mechanical actuation. The proposed means for capturing the surface motion of the breast in 3D is to use a stroboscope to capture images from multiple digital cameras at preselected phase angles. Photogrammetric techniques are then used to reconstruct matched point features in 3D.

Since human skin lacks high contrast visual features, it is necessary to introduce artificial fiducials which can be easily extracted from digital images. The chosen fiducials are points in three different colours in differing proportions randomly applied to the skin surface. A three-dimensional signature which is invariant to locally Euclidean transformations between images is defined on the points of the lowest proportion colour. The approximate local Euclidean invariance between adjacent frames enables these points to be matched using this signature. The remaining points are matched by interpolating the transformation of the matched points. This algorithm has significant performance gains over conventional gradient-based tracking algorithms because it utilises the intrinsic problem geometry.

Successful results are presented for simulated image sequences and for images of a mechanically actuated viscoelastic gel phantom with tracking errors within 3 pixels. The errors in the phantom sequence correspond to less than 0.3 mm error in space, which is more than sufficient accuracy for the DIET system.

1. INTRODUCTION

It is well known that breast cancer is one of the most prevalent and harmful diseases among women in modern society. For example, in New Zealand it was the leading form of cancer death in women in 1999.¹ For a high chance of successful treatment, it is critical that malignant tumours are detected as early as possible. If a carcinoma is detected sufficiently early, the chances of five-year survival increase to approximately 90%.² The current gold standard breast cancer screening method is mammography, which is currently implemented in wide-scale screening programs in most developed nations. Mammography measures variation in tissue density, however the contrast between healthy and diseased tissue is relatively low (5% - 10%) which creates difficulties with patients with higher density breast tissue,³ including most younger women. The search for effective scanning methods remains an active area of research, and elastographic techniques are beginning to play a significant role. Instead of imaging tissue density, elastography measures tissue stiffness. For breast tissue this is a high-contrast feature with an order of magnitude variation between normal and cancerous tissue,⁴ and hence potentially a much better feature to image. Emerging elastographic breast cancer screening techniques include mechanical sensors on the skin surface,⁵ ultrasound elastography,⁶ and MR elastography.⁷

A Digital Image-based ElastoTomography (DIET) system is being developed at the University of Canterbury, for measuring breast elasticity by elastotomography, in which breast surface motion of a sinusoidally mechanically actuated breast is measured in 3D and used as the input to an inverse problem which solves for breast elasticity. Note that this method requires no x-rays or any other scanning agent, as it only relies upon measurement of motion at the breast surface. The method proposed to measure the motion of the breast surface is to use computer vision algorithms to track and reconstruct the breast surface in 3D using multiple digital cameras. The use of widely available digital cameras as the measurement sensors creates the potential for a cheap, portable system suitable for use in a large-scale breast screening programme. It also allows for the possibility of tracking

the surface motion at very high resolution, allowing very small variations in surface motion to be detected. This paper presents a part of the 3D surface motion capture process, the problem of tracking a dense feature set between frames from a single camera.

2. BACKGROUND

2.1. DIET Project Requirements

Preliminary research with 2D simulations has suggested that detecting the change in breast surface motion due to small carcinomae may require tracking the breast surface with a resolution of the order of hundreds of points per square cm. This means that the 3D motion of each point in a dense set of points with interpoint spacing less than 1mm will be needed to interpolate the surface motion with the required degree of accuracy. Actuation amplitudes are likely to be at least 1mm, which means that individual points could move significantly more than that in one cycle due to resonance effects. Therefore, in one cycle, it is likely that point motion will be significantly higher than the average distance between neighbouring points.

In order to create an accurate reconstruction of the surface in 3D, there are a few important steps. Firstly, the cameras need to be precisely calibrated to allow for accurate reconstruction in world space. Then, to compute 3D coordinates of a point, the point needs to be unambiguously identified in at least two images from different cameras, and its image coordinates triangulated. Additionally, the points needed to tracked from frame to frame from a single camera in order to measure the motion of the surface. At this stage, tracking using the natural detail of human skin as the feature points looks difficult due to lack of sufficient contrast, so it will be necessary to apply some fiducial points artificially. These fiducials will need to be chosen to be suitable to both the correspondence problem of matching points between cameras, and the tracking problem of matching points between frames.

2.2. Existing Feature Tracking Techniques

Techniques for tracking feature points between frames fall into a two broad categories. There are those that work directly on the image data by matching intensity regions in images, and there are those which are based on geometric properties of 2D points, lines, etc. The methods which match the image intensities generally work on grayscale images by finding *interest* features in the image, such as image gradients or regions of low self-similarity, e.g. the Moravec Corner Operator, Harris Corner Detector, SUSAN, SIFT and many others. See Wikipedia⁸ for a summary of many techniques and comprehensive reference list. In these methods, points are then matched by comparing properties based on these features between sets of potential matches. The use of this type of method is problematic in this project for a number of reasons. Because of the large number of feature points required, it is difficult to introduce potentially thousands of points to the breast surface without introducing a high degree of self-similarity between the feature points. Therefore large numbers of feature points detected with the interest point detectors will be very similar, making unambiguous matches difficult. Another significant reason that a number of these methods aren't practical is that they require a large amount of computational time, as they work primarily on the raw images, which, given the accuracy requirements of the DIET project, contain millions of pixels. Therefore it is important to limit the amount of image processing on the image data as much as possible. The second type of tracking technique, known in the literature as point pattern matching, relies on geometric properties of a set of image points, e.g. a set of easily extracted dots on a surface. The advantage of this is that the features, or fiducials can be designed such that the point locations can be extracted from the raw images at low computational cost. The matching process can then be based around the motion of the coordinates of the points, rather than image intensity regions and thus falls more into the realm of computational geometry rather than image processing. Existing methods include relaxation and, more recently, thin plate splines for non-rigid point matching (TPS-RPM⁹). However, the existing methods are more suited to applications with a reasonably small set of points, and tend to require the convergence of some reasonably expensive iterative process. Because of the potential magnitude of data acquisition in a breast screening application, it is desirable to keep the surface acquisition step as computationally cheap as possible.

2.3. Tracking using invariant signatures

The approach considered in this paper is to construct a mapping of image points sets based on properties of neighbouring points, to some abstract space which is invariant under the transformations in question. The matching procedure is then reduced to matching signatures of points from two images in the signature space. The concept of a signature for registration was presented in Calabi, Olver et. al.¹⁰ where a signature curve which is invariant under Euclidean transformations was constructed for any smooth curve in \mathbb{R}^2 and used for applications in object recognition.

In this paper, the types of transformations considered are those which can be well-approximated locally by members of a transformation group. The signature is defined in terms of the joint invariants of the transformation group, for example for the Euclidean group it will be dependent on interpoint distances of a set of points, for the affine group it would be dependent on area ratios, for the Projective group, the cross ratio of five points, etc. As a proof of concept, a locally Euclidean-invariant signature is constructed which is suitable for application to the tracking problem, where transformations are small, and affine and projective distortion is low. This signature is used to successfully track points both in computer simulation and in images taken of a sinusoidally actuated gel phantom. It is determined empirically that the transformations between frames can be reasonably well-approximated locally by Euclidean transformations, and hence the relative motion of the points in signature space is less than in the image space. The approximation is not perfect however, meaning that the points in the signature space do move; they are not entirely invariant. The simple Euclidean-invariant signature used alone is thus not suitable for correctly matching all points, due to this non-Euclidean motion and the signature space being crowded with many points. However if the signature is augmented with some known information about the transformation, for example an upper bound on the magnitude of the transformation, this is sufficient to rule out almost all false matches, to achieve a proportion of correct matches of around 99%.

3. TECHNICAL DETAILS

3.1. Fiducial Design

As the signatures are based on matching small sets of neighbouring points to corresponding small sets of points, it is necessary to design a method of reliably selecting those sets of points so that the same set will be chosen from both images, and will hence have the correct signature computed. Depending on the invariant used, the sets of points will be small. For example, for the Euclidean group the sets can be as small as two points as the fundamental Euclidean joint invariant, the distance between points, can be computed from just two points. Note that in practice, this would be too few, because the resulting one-dimensional signature space consisting of two-point distances would be too crowded to allow unique signature matches. The signature constructed in this paper is a function of only three points, so all that is required is a method of choosing sets of three points.

As discussed earlier, in a clinical context it will be necessary to apply points to the breast surface to achieve a sufficient density of consistently recognisable points. These points, to be suitable for use in practice would need to be easy to apply and remove. It was therefore decided to use small coloured dots in three colours (red, green, and blue) that could be randomly applied to the breast surface. In practice this could be achieved by using small plastic coloured particles, or droplets of coloured fluid sprayed from an atomiser. In the practical experiments undertaken, finely cut pieces of coloured paper were used. The colours can be chosen to stand out clearly from the surface, and can easily be extracted by comparing the different colour channels in an RGB image without using any sophisticated colour space transformations. Ordered point triples can be identified with the points of one colour, e.g. red, and chosen by selecting the closest blue and green points to it. The assumption made is that nearest neighbours won't change under the types of transformation under consideration. If the nearest neighbours do occasionally change, this still should not create a false positive match, as the signatures will simply fail to correspond.

3.2. Proof of concept 3-point joint Euclidean-invariant signature

For this paper, an example signature is presented as a proof of concept, and tested both in computer simulation and in some laboratory experiments with gel phantoms. The signature is based on a 3-point joint invariant of the Euclidean group, which in turn is a function of the three 2-point joint invariants, the distances between

each pair of points. The reason why the Euclidean group was chosen over a less restrictive group such as the affine group or the projective group was that the invariants, because they only involve distances, not ratios, are more resilient to noise. It will be shown that the signature space does distort under the types of transformation encountered, but that the signature space is sufficiently spread that the 3-point clusters can still be successfully matched, with the augmentation of some additional information about the transformation, such as a magnitude upper bound.

3.2.1. Signature definition

Consider image points in two-dimensional Euclidean space $\mathbf{x} \in \mathbb{E}^2$, and denote the Euclidean distance metric by $d(\cdot, \cdot)$. A signature function $f : \mathbb{E}^2 \times \mathbb{E}^2 \times \mathbb{E}^2 \rightarrow \mathbb{R}^3$ can be defined on triples of these points as follows:

$$f : (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) \rightarrow (d(\mathbf{x}^{(2)}, \mathbf{x}^{(1)}), d(\mathbf{x}^{(3)}, \mathbf{x}^{(1)}), d(\mathbf{x}^{(3)}, \mathbf{x}^{(2)})) \quad (1)$$

Note that because Euclidean transformations have three degrees of freedom (two for translation, and one for rotation), and a three point configuration of 2D points has six degrees of freedom, there are only three functionally independent Euclidean joint invariants. Any other joint Euclidean invariant, e.g. angles, can be derived from these three distances.

For the fiducials, red, green and blue points are randomly applied to the surface. To identify the coloured points with the ordered point triples, the sets of three points are defined to be each red point, along with its nearest blue and green neighbours. To facilitate matching high densities of points the red points are applied at a significantly lower density than the green and blue points, which means that a smaller number of points are matched using the signatures, allowing the rest to be matched by interpolating the motion of the signature-matched points.

3.2.2. Matching Algorithm

The matching procedure consists of a few simple steps. First, all the red,

1. Extract all of the red, green and blue points from the image. The means of doing this is dependent on the digital images, but the colours can be chosen to make this easy to do
2. Find the nearest blue and green neighbours to each red point. These form the point triples.
3. Compute the signature (Eq. 1) for each point triple
4. Match point triples by matching their signatures in signature space, discarding matches if they violate some upper bound on the magnitude of the transformation.
5. Match the remaining unmatched points by interpolating the transformation of the matched points

4. EXPERIMENTS

Two experiments were performed to verify the viability of feature tracking using signatures for this kind of application. Real data was obtained by applying coloured points to an actuated gel phantom and taking digital images from two cameras. Additionally, the algorithm was tested in computer simulation, using a finite element model of the gel phantom used in the real experimental setup.

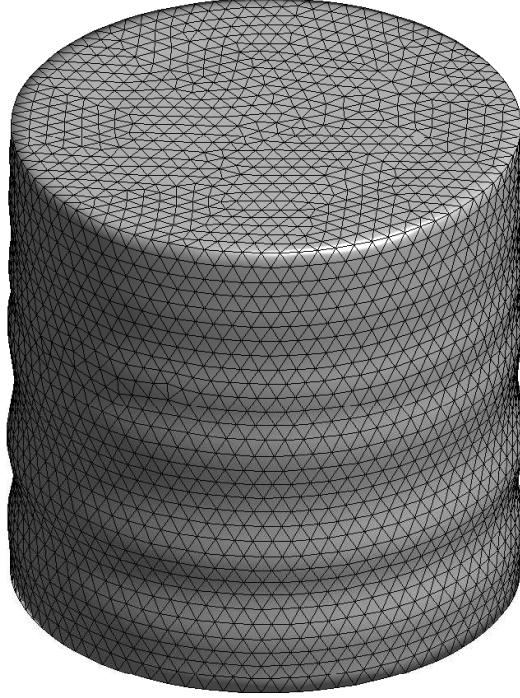


Figure 1. Finite element gel cylinder model used for simulation

4.1. Computer Simulation

4.1.1. Motion Model

To simulate the motion of the feature points, a finite element model of a gel cylinder with similar elastic properties to human breast tissue was used. The model was developed by Ashton Peters, one of the members of the DIET research group. The 17000 node finite element mesh was developed in GAMBIT, and the model simulated in Fortran 90, the matrix inversion being done with a direct sparse matrix inversion package, MUMPS. The model was simulated with a 50Hz actuation frequency, with 0.5mm peak to peak amplitude. See Fig. 1 for a picture of the cylinder. 1600 coloured points, 400 red, 800 blue, and 800 green were randomly generated on a 2x2cm patch on the surface of the cylinder, and their motion was computed by interpolating the motion of the surface nodes onto the random points using a scattered data surface fitting routine, `gridfit`¹¹ developed by John D’Errico, based on a modified ridge estimator. The image coordinates were determined by using a pinhole camera model, using the same model as computed for the real camera in the laboratory experiment.

4.1.2. Validation of Euclidean assumption

To test the validity of the assumption that the transformation between frames is locally Euclidean, small groups of 5-6 neighbouring points were randomly selected, and the motion between different pairs of frames was examined. A nonlinear least squares estimate of the best Euclidean transformation between the two groups, parameterised by a rotation angle, θ and two translation components t_x, t_y , was performed, minimising the cost function:

$$c(\theta, t_x, t_y) = \sum_i \|R(\theta)\mathbf{x}_i + \mathbf{t} - \hat{\mathbf{x}}_i\|_2^2 \quad (2)$$

where $R(\theta)$ is the standard rotation matrix parametrised by the counterclockwise rotation angle θ , the vectors \mathbf{x}_i are points in the first image with $\hat{\mathbf{x}}_i$ the corresponding points in the second image, and \mathbf{t} is the translation vector $[t_x, t_y]^T$. The points \mathbf{x}_i were transformed by the estimated Euclidean transformation, and the distance to each corresponding point, $\hat{\mathbf{x}}_i$ computed. Typical results, for any combination of two frames from the twenty,

were a mean distance error of less than 0.1 pixels, where the transformations ranged in magnitude from 1 to 15 pixels. Therefore, the transformations caused by the motion and deformation of the simulated surface can be well-approximated locally by Euclidean transformations.

4.1.3. Results

The tracking procedure was performed on different pairs of images from one sinusoid, and performed well, registering over 99% of points matched, and of those, fewer than 0.5% mismatched, even for images 180 degrees out of phase, where the motion was greatest. Matching was performed with an assumed translation upper bound of 15 pixels. See Figure. 2 for an illustration of the point motion and tracking. Performance was not so good in the presence of noise. Performance was good with Gaussian noise of standard deviation of up to 0.2 pixels in each direction, however anything higher than this significantly degraded performance. This is because the signature space is a 3 dimensional space of distances, essentially in pixel units, and its scale is determined by the interpoint spacing of the points. In this case, noise any higher than about 0.2 pixels can easily cause a 15% change in the corresponding signature points. The solution to offset this problem is to consider a smaller image region by zooming. This has the effect of scaling the signature space upwards while leaving the measurement noise relatively unchanged. This, however, would mean requiring more digital cameras to comprehensively cover entire breast surface.

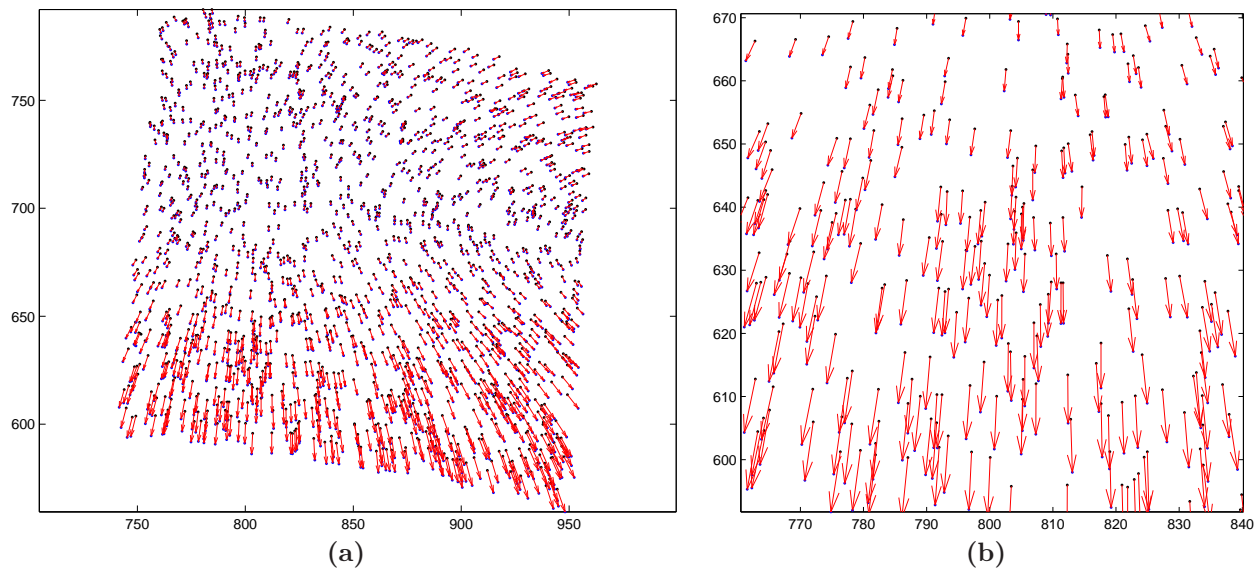


Figure 2. Tracked points from computer simulation of gel cylinder. All points are shown, with matched pairs connected with an arrow. All the points are shown in (a) while (b) is a close-up

4.2. Gel Phantom Simulation

4.2.1. Experimental setup

The computer experiment from Section 4.1 was duplicated on real laboratory equipment. A cylindrical gel phantom made from an A-431 and LSR-05 silicone gel composite was placed on the platform of a mechanical actuator. A number of red, green, and blue coloured speckles made from finely cut coloured paper were applied to the surface of the cylinder, see Fig. 3(a). The actuator was controlled by a simple control system written in Simulink and implemented in real time in dSpace to give a 50Hz, 1mm peak to peak sinusoid at the actuator plate. Still images at arbitrary phase were generated by strobing the cylinder in a dark enclosure at the actuation frequency at a user-specified phase separation from the actuator input signal, and the images were captured using a 6 megapixel Canon Powershot G5 digital CCD camera. A set of 20 images was captured in this manner to cover the entire sinusoid at 18 degree phase differences. The camera was calibrated beforehand by corresponding



Figure 3. (a)Silicone gel phantom with coloured speckles. (b) detected red, green, and blue points.

known points on a precisely machined calibration object with their image locations and solving for the projection matrix.

4.2.2. Results

Red, green, and blue regions were determined by directly thresholding the RGB image and comparing colour channels, and the point locations were taken to be the centroids of the coloured blobs,. This process takes a little under a second per image for a 6 megapixel image in Matlab. An example of the extracted points is depicted in Fig. 3(b). Note that with this simple approach not all the image points are recognised, however this is not a concern as only a certain density of points is needed to be matched, rather than every individual point.

The matching procedure was tested on various image pairs, and it was found that it successfully matched over 90% of the points with fewer than 1% mismatches (from visual inspection). The points that didn't get matched weren't matched for one of three simple reasons. Firstly, some points were not picked up by the simple feature detection process. Therefore some points simply did not have corresponding points. Secondly, for some points only part of the coloured region comprising the point was detected. Finally, in some regions there were not sufficient red points, and hence insufficient point triples, to allow accurate interpolation of the motion of the points matched by the signature method in order to match the remaining points. introduces an error into the centroid location which can interfere with the matching process. See Fig. 4 for an idea of the motion undergone by the matched points.

5. CONCLUSION AND FUTURE WORK

A simple Euclidean-invariant signature has been presented and tested as a proof of concept for using local signatures for feature tracking and matching. The signature used was sufficient to track features on a regularly deforming gel phantom. However, this basic signature is not particularly robust to errors because it is entirely defined by the geometry of point distances between closely spaced points. This did not have any significant effect on the laboratory experiment, as the points were not as close together as they might be in the near future, however in the simulation it caused significant problems. The solution to this, zooming in on the surface to scale up the signature space while leaving the measurement noise unchanged has the drawback of requiring more cameras to track the entire surface, meaning ultimately more expense and a less portable system. Thus, the next step is to investigate a more sophisticated signature that is more robust. This could be done by building up invariants of a larger number of points; expanding the dimension of the signature space will increase the separation of the signature points. Alternatively, signatures could be built up of less noise-sensitive quantities, for example with distributions of colour ratios in a small region. This would have the advantage of being entirely view-independent, and would be useful for the correspondence problem as well as the tracking problem.

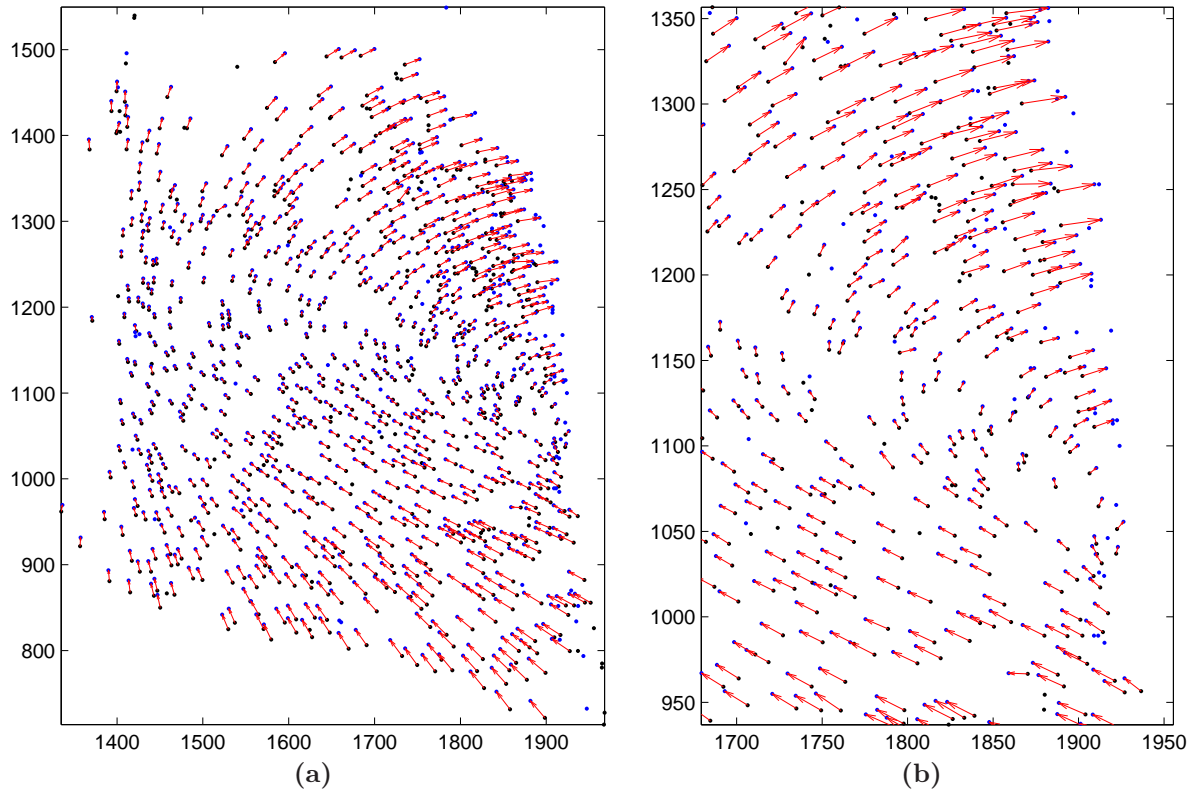


Figure 4. Tracked points from gel phantom experiment. All points are shown, with matched pairs connected with an arrow. All the points are shown in (a) while (b) is a close-up

ACKNOWLEDGMENTS

Ashton Peters, for the finite element gel phantom model, for the gel phantoms, and for helping with the image capture.

REFERENCES

1. C. S. of New Zealand, "Causes of cancer death in New Zealand," 2003.
2. C. C. Boring, Squires, and T. Tong, "Cancer statistics," *CA Cancer J. Clin* **44**, pp. 7–26, 1994.
3. M. Tilanus-Linthorst, L. Verhoog, I. M. Obdeijn, K. Bartels, M. Menke-Pluymers, A. Eggermont, J. Klijn, H. Meijers-Heijboer, T. van der Kwast, and C. Brekelmans, "A *brca1/2* mutation, high breast density and prominent pushing margins of a tumor independently contribute to a frequent false-negative mammography," *Int. J. Cancer* **102**(1), pp. 91–95, 2002.
4. T. M. Krouskup, T. Wheeler, F. Kallel, B. S. Garra, and T. Hall, "Elastic moduli of breast and prostate tissues under compression," *Ultrasonic Imaging* **20**, pp. 260–274, 1998.
5. A. Sarvazyan, "Mechanical imaging: a new technology for medical diagnostics," *Int. J. Med. Inf.* **49**(2), pp. 195–216, 1998.
6. P. Chaturvedi, M. Insana, and T. Hall, "Acoustic and elastic imaging to model disease induced changes in soft tissue structure," in *Information Processing in Medical Imaging, Proc. IMPI*, 1997.
7. J. B. Weaver, E. E. W. V. Houten, M. I. Miga, F. E. Kennedy, and K. D. Paulsen, "Magnetic resonance elastography using 3d gradient echo measurements of steady-state motion," *Medical Physics* **28**(8), pp. 1620–1628, 2001.

8. Wikipedia, "Corner detection — Wikipedia, the free encyclopedia," 2007. [Online; accessed 15-January-2007].
9. H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Computer Vision and Image Understanding* **89(2-3)**, pp. 114–141, 2003.
10. E. Calabi, P. J. Olver, C. Shakiban, A. Tannenbaum, and S. Haker, "Differential and numerically invariant signature curves applied to object recognition," *International Journal of Computer Vision* **26**, pp. 107–135, 1998.
11. J. D'Errico, "Matlab Central File Exchange — Surface Fitting using gridfit," 2006. [Online; retrieved 11-October-2006].