

Measurement Selection in Untracked Freehand 3D Ultrasound

Catherine Laporte^{1,2} and Tal Arbel²

¹ Dept. of Electrical Engineering, École de Technologie Supérieure, Montréal, Canada

² Centre for Intelligent Machines, McGill University, Montréal, Canada

catherine.laporte@etsmtl.ca, arbel@cim.mcgill.ca

Abstract. In freehand 3D ultrasound, out-of-plane transducer motion can be estimated via speckle decorrelation instead of using a position tracking device. This approach was recently adapted to arbitrary media by predicting elevational decorrelation curves from local image statistics. However, such adaptive models tend to yield biased measurements in the presence of spatially persistent structures. To account for such failures, this paper introduces a new iterative algorithm for probabilistic fusion and selection of correlation measurements. In experiments with imagery of animal tissue, the approach yields significant accuracy improvements over alternatives which do not apply principled measurement selection.

1 Introduction

Freehand 3D ultrasound (US) involves the reconstruction of a volume from 2D US images [15]. The relative positions of the 2D images are generally obtained from an external position tracking device, at the expense of clinical convenience. Alternatively, in-plane motion can be recovered using image registration techniques, and out-of-plane motion can be estimated from elevational speckle decorrelation [2,16,11,4]. By scanning a speckle phantom at regular elevational intervals and pooling the correlation coefficients computed between pairs of images, a transducer specific decorrelation curve can be constructed so as to estimate out-of-plane motion between pairs of correlated images in new data sets [2,16].

The above approach only works under Rayleigh scattering conditions, which seldom occur in real tissue. Recent solutions to this problem predict the shape of the local decorrelation curve from signal statistics within the image plane [3,8]. While this dramatically reduces error, experiments suggest that accuracy deteriorates for larger displacements [6], as shown in Fig. 1. The problem is that these adaptive models predict complete image decorrelation as displacement increases. Yet, in the presence of spatially persistent structures like the bright layer of fat in the pork flank sample of Fig. 1(b), correlation remains significant even for large displacements. When computing transducer trajectory, distance estimates arising from inaccurate parts of the adapted model should be discarded. Moreover, the displacement range over which the adaptive model is accurate and the extent of the bias outside this range are subject dependent. Therefore, the measurement selection strategy must dynamically adapt to the data.

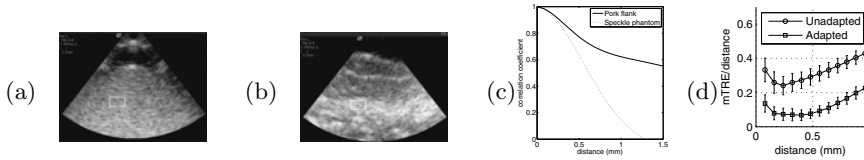


Fig. 1. (a) Speckle phantom image; (b) Pork flank image; (c) Decorrelation curves associated with the data in the white rectangles; (d) Relative motion estimation error obtained for the pork flank data set with and without medium specific adaptation [8]

Measurement selection for image-based transducer tracking is discussed in [4], which proposes to compute a robust average of reconstructions obtained from interleaved subsets of the image frames. Optimal assignment of frames to independent reconstructions requires consecutive frames of a given subset to be correlated, but not so much that they lie on the flat, error-prone part of the decorrelation curve [13]. However, this does not explicitly account for the kind of significant but uninformative image correlations shown in Fig. 1.

This paper proposes a new iterative measurement fusion and selection algorithm based on statistical tests to account for persistent structures in the medium. The method extends the probabilistic approach of [7] for use with imagery of real tissue using the technique in [8] and prevents errors arising from situations like Fig. 1. The structure of this paper is as follows. Section 2 reviews the probabilistic speckle decorrelation model and its extension to arbitrary tissue. Section 3 describes the new iterative measurement selection and probabilistic fusion scheme based on this model. Section 4 discusses experiments on animal tissue data, showing that the new approach offers improved accuracy over approaches that do not use principled measurement selection.

2 Probabilistic Speckle Decorrelation Model

In this work, US images are divided into Q patches (e.g. Fig. 2, left) to account for the spatial variation of elevational beam width. Each patch has its own nominal speckle decorrelation model. During a calibration step, a speckle phantom is scanned at regular elevational intervals. Gaussians are fitted to the patch-wise axial, lateral and elevational decorrelation curves derived from this imagery; their standard deviations are termed the nominal axial, lateral and elevational *correlation lengths* of the transducer, and labelled \hat{w}_a , \hat{w}_l and \hat{w}_e . A refined nominal probabilistic elevational speckle decorrelation model is also computed for each patch by maximum entropy analysis [1,7], yielding a log-likelihood function $\hat{L}(\rho|\delta, q) = \log \hat{p}(\rho|\delta, q)$ that captures the stochasticity of correlation measurements ρ in relation to displacement δ for image location q .

In real tissue, US correlates over larger distances than in a speckle phantom. Here, this phenomenon is modelled as a scaling of the nominal elevational decorrelation model along the δ axis by a factor r_e , determined from four statistical

features measured locally within the image plane: the axial and lateral correlation lengths, w_a and w_l , and the squared signal to noise ratio $R^2 = \frac{E\{I\}^2}{VAR\{I\}}$ and skewness $S = \frac{E\{(I-E\{I\})^3\}}{VAR\{I\}^{\frac{3}{2}}}$, where I is the squared US envelope.

As described in [8], the relationship between these image statistics and the local elevational correlation length w_e is learned from a pool of synthetic US data depicting varied media, using sparse Gaussian process regression [14]. To enforce transducer independence, the regressive relationship is established using normalised features $\{r_a = \frac{w_l}{w_a}, r_l = \frac{w_l}{w_l}, r_{R^2} = \frac{R^2}{R^2}, r_S = \frac{S}{S}\}$ and $r_e = \frac{w_e}{w_e}$. In spite of its simplicity, this model was shown [8] to have accuracy comparable to that of [3]. Both models tend to fail for large elevational separations. This paper shows how to detect such failures.

In new, real imagery, r_e is estimated locally from in-plane image features using the learned regressor. The probabilistic nominal speckle decorrelation model is then scaled, yielding a locally adapted probabilistic estimator of elevational separation, $L(\rho_{ij}|\delta, q) = \hat{L}(\rho_{ij}|r_e\delta, q)$. In practice, δ is reasonably approximated by a Gaussian random variable with mean $\bar{\delta}$ and variance σ^2 given by

$$\bar{\delta} = \underset{\delta}{\operatorname{argmax}} L(\rho|\delta, q) \quad \text{and} \quad \sigma^2 = - \left(\frac{d^2}{d\delta^2} L(\rho|\bar{\delta}, q) \right)^{-1}. \quad (1)$$

3 Recovering Transducer Motion

In order to eliminate sources of error outside the control of the proposed techniques from our experiments, it is assumed throughout the paper that the transducer does not move in-plane (or else the images have been correctly re-aligned by an image registration technique) and that any out-of-plane rotations are small. Note that the measurement selection technique introduced in this paper would apply equally to more complete motion models.

Subdividing the image frames in the US data set into Q patches corresponding to those used to define the speckle decorrelation model breaks the data set into Q individual smaller US data sets consisting of “mini-frames” the size of an individual image patch. The relative positions of the full US image frames are then recovered as follows: (1) For every pair of corresponding mini-frames i and j , a correlation measurement is obtained and $r_{e_{ij}}$ is estimated using the learning-based approach of [8]; (2) For each image location $q \in Q$, estimates of elevational mini-frame positions are derived from redundant correlation-based distance estimates deemed shorter than $w_{e_{ij}} = r_{e_{ij}}\hat{w}_e$. This is done using the measurement selection and fusion algorithm of section 3.1.¹ (3) The resulting data are smoothed across each frame using a median filter (as suggested in [3]); (4) Rigid transformations relating every frame to the first are computed using a least-median-of-squares approach [12,7].

¹ Although the data sets used in the experiments involved only monotonic probe motion, the combinatorial frame ordering algorithm of [7] was also applied at this stage and may have influenced the outcome of the experiments.

3.1 Iterative Measurement Selection and Fusion

Let $z_i, i = 0, \dots, n$ be the position of the i th mini-frame corresponding to patch q along the elevational direction relative to $z_0 = 0$ and let \mathbf{z} be the vector formed by the $z_i, i = 1, \dots, n$. The correlation coefficient ρ_{ij} between mini-frames i and j provides an estimate of their elevational separation $\delta_{ij} = |z_j - z_i|$ via (1). Assuming that the distance estimates obtained for different mini-frame pairs are mutually independent given the configuration of the frames, an estimate of \mathbf{z} and its covariance matrix \mathbf{C} are obtained by solving [9,7]:

$$\mathbf{G}\mathbf{z}^* = \mathbf{b} \quad \text{and} \quad \mathbf{C} = \mathbf{G}^{-1}, \quad (2)$$

$$\text{with } G_{ij} = \begin{cases} \sum_{k=0}^n \frac{1}{\sigma_{ik}^2} & i = j \\ -\frac{1}{\sigma_{ij}^2} & \text{otherwise} \end{cases} \quad \text{and} \quad b_i = \sum_{j=0, j \neq i}^n \frac{\bar{\delta}_{ij}}{\sigma_{ij}^2}. \quad (3)$$

In principle, this approach could be used to combine all available correlation measurements into a trajectory estimate for image location q . In practice, this is highly error prone because, as illustrated in Fig. 1, the (adapted) speckle decorrelation model exhibits bias in certain parts of its range. Biased measurements can be excluded from the data fusion process through an iterative algorithm based on statistical hypothesis testing. Assuming an unbiased initial trajectory estimate $\mathbf{z}_{[0]}$, iteration t proceeds as follows:

1. A new subset of the correlation measurements are tested against the current estimate of the trajectory. For each candidate δ_{ij} derived via (1), let H_0 denote the hypothesis that it is consistent with $\mathbf{z}_{[t-1]}$, and let H_a denote the hypothesis that it is biased. H_0 is rejected in favour of H_a at the 1% significance level (and δ_{ij} is discarded) if

$$|\bar{\delta}_{ij} - |z_{[t-1]_j}^* - z_{[t-1]_i}^*|| \geq \Phi^{-1}(0.995, \sigma_{ij}^2 + C_{[t-1]_{jj}} + C_{[t-1]_{ii}} - 2C_{[t-1]_{ij}}), \quad (4)$$

where $\Phi^{-1}(\beta, \nu^2)$ is the inverse cumulative distribution function of a Gaussian random variable with variance ν^2 evaluated at probability β .

2. The measurements accepted by the above hypothesis tests are fused with all other previously accepted measurements to obtain $\mathbf{z}_{[t]}^*$ and $\mathbf{C}_{[t]}$ by solving (2).

The effectiveness of this approach depends on the initial trajectory estimate (see section 3.2) and the order in which the available measurements are processed (or *testing schedule*). This paper adopts a very simple testing schedule: the first iteration considers only correlation measurements between consecutively acquired mini-frames (unless they were already used to compute $\mathbf{z}_{[0]}$). The second iteration considers correlation measurements between frames which are two acquisition steps apart, and so on until no new measurements are accepted.

3.2 Initial Trajectory Estimate

The proposed scheme rests on the assumption of an unbiased $\mathbf{z}_{[0]}$. In the absence of other data, $\mathbf{z}_{[0]}$ is obtained from n correlation measurements. These must form

a spanning tree of the graph \mathcal{G} whose vertices and edges represent the mini-frames and the correlation measurements between them. A *minimal* spanning tree (MST) of \mathcal{G} is hence an optimal initial measurement set with respect to a cost function κ over pairs of mini-frames. One possible cost function is

$$\kappa_1(i, j) = |i - j|. \quad (5)$$

The associated MST consists of the n correlation measurements between consecutive mini-frames. Because the log-likelihood function is flat for large ρ , the uncertainty of the associated $\mathbf{z}_{[0]}$ is typically high so that initially, H_0 is rarely rejected by (4). The initial estimate is thus diluted by new, more certain measurements early on in the data fusion process. As long as modelling inaccuracies occur mostly for widely separated images (e.g. Fig. 1), this can lead to good results. However, empirical evidence suggests that model failures can occur for shorter displacements when the overall US signal is low, due to the increased contribution of unmodelled noise in the image [6]. In such cases, inaccurate estimates of r_e are obtained and short displacements are over-estimated. The following two cost functions are designed to alleviate this problem:

$$\kappa_2(i, j) = \text{sgn}(r_{e_{ij}} - 1)\bar{\delta}_{ij} \quad \text{and} \quad \kappa_3(i, j) = \text{sgn}(r_{e_{ij}} - 1)\bar{\delta}_{ij}/\sigma_{ij}^2. \quad (6)$$

Where the estimated $r_e > 1$, κ_2 and κ_3 are similar to κ_1 , preferring measurements with large relative variance. This changes in areas where the estimated $r_e < 1$, which is taken to be indicative of model failure.² In these error-prone areas, κ_2 and κ_3 select measurements with smaller relative variance, making the hypothesis test of (4) more discriminative.

4 Experiments

Imagery was acquired using a video frame-grabber connected to an Acuson Cypress US scanner with a 7V3C transducer at a depth setting of 2.7 cm. Images were divided into 26 patches of 50×30 pixels and back-converted to polar coordinates. Motion was purely elevational and controlled to 0.01 mm precision using a manually operated linear stage. Fifty-one 161 frame long scans of pork tenderloin, turkey breast and beef brisket samples were acquired at a range of different regular spacings. Fig. 2 shows sample images. A nominal probabilistic speckle decorrelation model was estimated from speckle phantom imagery acquired at regular 0.05 mm spacing. The algorithm of [10] was applied to phantom images to reverse log-compression in all data sets. The mapping from US image statistics to elevational correlation length was learned from synthetic image sequences depicting varied media, as described in [8]. For each animal tissue scan, the probe trajectory was estimated as described in Section 3.

² Indeed, true $r_e < 1$ is a very rare occurrence in the noiseless training set because correlation length is theoretically minimal under Rayleigh scattering conditions ($r_e = 1$). Thus, small values of r_e , which arise in the presence of noisy real data, while often estimated as small (< 1), are unlikely to be estimated *accurately* by the learning-based algorithm.

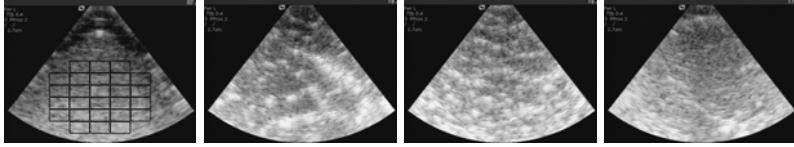


Fig. 2. From left to right: sample speckle phantom, pork, beef and turkey images. The phantom image shows the subdivision of imagery into 26 patches.

Three versions of the proposed approach, HYP-K1, HYP-K2 and HYP-K3, were considered, respectively corresponding to the initialisation cost functions of (5) and (6). Two baseline approaches were also tested: (a) a nearest-neighbour (NN) approach, which uses only deterministic distance estimates from consecutive pairs of mini-frames and (b) solving (2) once using all δ_{ij} deemed shorter than $w_{e_{ij}} = r_{e_{ij}}^* \hat{w}_e$, an approach labelled NMS (no measurement selection). The latter is a direct adaptation of [7] (shown to work well under Rayleigh scattering conditions) for use in imagery of real tissue via [8].

For statistical analysis, accuracy was summarised by the mean target registration error (mTRE) measured at the mini-frame centres, averaged over all frames and normalised by the length of the trajectory. A mixed within/between-subject factor ANOVA was carried out on the results of the 51 experiments to evaluate the effect of algorithm choice and tissue type.³ The HYP-K2 and HYP-K3 variants of the proposed approach significantly outperformed both baselines ($p < 0.001$ in both cases). HYP-K1 significantly outperformed NMS ($p < 0.001$), but was not significantly more accurate than NN ($p = 0.251$).

Fig. 3 illustrates these results with examples. In most cases, HYP-K2 and HYP-K3 gave excellent results (e.g. Fig. 3(a)-(c)). Though NN and HYP-K1 may exhibit similar mTRE, they differ in qualitative behaviour. NN underestimates displacement because the deterministic decorrelation model is biased for short displacements [13]. HYP-K1 drifts because it easily accepts displacement over-estimates arising from unmodelled image noise (e.g. Fig. 3 (a), (c)), as discussed in Section 3.2. This likely explains the significantly lower accuracy obtained with turkey data (which has lower echogenicity) than with beef data ($p = 0.042$). HYP-K2 and HYP-K3 were usually able to compensate for this modelling limitation (see Fig. 3(c)).

Though they generally gave good results, HYP-K2 and HYP-K3 were not always able to overcome failures of the adaptive speckle decorrelation model. For example, in Fig. 3(d), distances were over-estimated by all approaches except, uncharacteristically, NMS. Here, the κ_2 and κ_3 cost functions probably failed to protect the trajectory estimate against distance over-estimates because there were few areas where $r_e < 1$. While $r_e < 1$ is a good indicator of model failure, $r_e > 1$ does not always indicate success.

³ The Bonferroni adjustment was used when carrying out multiple comparisons and statistical significance was evaluated at the 5% level.

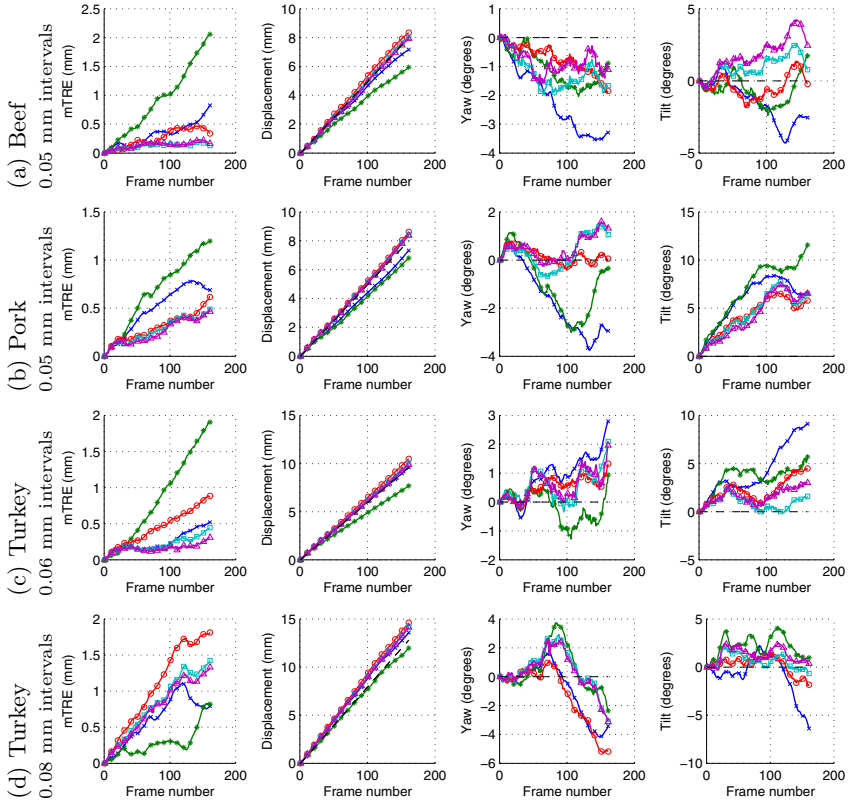


Fig. 3. Sample results, markers placed every 10 data points for greyscale legibility. From left to right: mTRE, estimated displacement, yaw and tilt evaluated at the centroid of the grid of image patches. Dashed: ground truth; blue crosses: NN; green stars: NMS; red circles: HYP-K1; cyan squares: HYP-K2; magenta triangles: HYP-K3.

5 Conclusions

This paper considered a new approach for image-based recovery of out-of-plane US transducer trajectories which generalises the probabilistic data fusion of [7] to arbitrary media using the learning-based approach of [8]. In order to discard significant, but uninformative image correlation measurements arising from spatially persistent structures in the medium, a novel iterative algorithm was presented which alternates measurement fusion steps with measurement selection steps based on statistical hypothesis testing. Despite some imperfections, experiments show that, if well initialised, the algorithm is statistically significantly more accurate than a deterministic nearest-neighbour approach or probabilistic fusion without measurement selection. It would be interesting to compare the new approach to that of [4] and possibly combine them by applying the probabilistic techniques in this paper to the frame subsets selected and averaged in [4].

Future research could also exploit the measurement selection approach in a context where image-based measurements complement an external tracking device providing mostly unbiased but highly imprecise positional data [5]. The latter, along with a valid noise model, could initialise the iterative algorithm, thereby improving its robustness.

References

1. Baker, R.: Probability estimation and information principles. *Struct. Saf.* 9, 97–116 (1990)
2. Chen, J.-F., Fowlkes, J.B., et al.: Determination of scan-plane motion using speckle decorrelation: theoretical considerations and initial test. *Int. J. Imaging Syst. Technol.* 8(1), 38–44 (1997)
3. Gee, A.H., Housden, R.J., et al.: Sensorless freehand 3D ultrasound in real tissue: speckle decorrelation without fully developed speckle. *Med. Image Anal.* 10(2), 137–149 (2006)
4. Housden, R.J., Gee, A.H., et al.: Sensorless reconstruction of unconstrained freehand 3D ultrasound data. *Ultrasound Med. Biol.* 33(3), 408–419 (2007)
5. Lang, A., Mousavi, P., et al.: Fusion of electromagnetic tracking with speckle-tracked 3D freehand ultrasound using an unscented Kalman filter. In: *Proc. SPIE Med. Imaging*, pp. 72651A-1–72651A-12 (2009)
6. Laporte, C.: Statistical methods for out-of-plane ultrasound transducer motion estimation. PhD thesis, McGill University (2009)
7. Laporte, C., Arbel, T.: Combinatorial and probabilistic fusion of noisy correlation measurements for untracked freehand 3D ultrasound. *IEEE TMI* 27(7), 984–994 (2008)
8. Laporte, C., Arbel, T.: Learning a tissue invariant ultrasound speckle decorrelation model. In: *Proc. ISBI*, pp. 995–998 (2009)
9. Lu, F., Milios, E.: Globally consistent range scan alignment for environment mapping. *Auton. Robots* 4(4), 333–349 (1997)
10. Prager, R.W., Gee, A.H., et al.: Decompression and speckle detection for ultrasound images using the homodyned k-distribution. *Patt. Recognit. Lett.* 24, 705–713 (2003)
11. Prager, R.W., Gee, A.H., et al.: Sensorless freehand 3-D ultrasound using regression of the echo intensity. *Ultrasound Med. Biol.* 29(3), 437–446 (2003)
12. Rousseeuw, P.J., Leroy, A.M.: *Robust Regression & Outlier Detection*. Wiley, Chichester (1987)
13. Smith, W., Fenster, A.: Statistical analysis of decorrelation-based transducer tracking for three-dimensional ultrasound. *Med. Phys.* 30(7), 1580–1591 (2003)
14. Snelson, E., Ghahramani, Z.: Sparse Gaussian processes using pseudo-inputs. In: *NIPS*, vol. 18, pp. 1257–1264 (2006)
15. Solberg, O.V., Lindseth, F., et al.: Freehand 3D ultrasound reconstruction algorithms — a review. *Ultrasound Med. Biol.* 33(7), 991–1009 (2007)
16. Tuthill, T.A., Krücker, J.F., et al.: Automated three-dimensional US frame positioning computed from elevational speckle decorrelation. *Radiology* 209(2), 575–582 (1998)