

Probabilistic Speckle Decorrelation for 3D Ultrasound

Catherine Laporte and Tal Arbel

Centre for Intelligent Machines, McGill University, Montréal, Canada
{cathy,arbel}@cim.mcgill.ca

Abstract. Recent developments in freehand 3D ultrasound (US) have shown how image registration and speckle decorrelation methods can be used for 3D reconstruction instead of relying on a tracking device. Estimating elevational separation between untracked US images using speckle decorrelation is error prone due to the uncertainty that plagues the correlation measurements. In this paper, using maximum entropy estimation methods, the uncertainty is directly modeled from the calibration data normally used to estimate an average decorrelation curve. Multiple correlation measurements can then be fused within a maximum likelihood estimation framework in order to reduce the drift in elevational pose estimation over large image sequences. The approach is shown to be effective through empirical results on simulated and phantom US data.

1 Introduction

Freehand 3D ultrasound (US) involves integrating the information contained in 2D US images into a 3D model. For this difficult task, the positions of the 2D images relative to each other must be known. A typical implementation involves attaching a tracking device to the US probe. Unfortunately, the position tracking device is often cumbersome to the clinician. Furthermore, the accuracy of the pose measurements depends on temporal and spatial calibration procedures [9] which are technically non trivial and often time consuming. This has led to the development of a position tracking methodology based entirely on image content. In-plane probe motion can be estimated through standard image registration techniques. This paper focuses on the estimation of out-of-plane motion from the elevational decorrelation of US speckle between image frames [4,13,10,3,6].

A typical framework for estimating the elevational separation between US images involves a calibration phase where a speckle phantom is scanned at regular intervals, allowing the construction of a decorrelation curve, depicting the measured average correlation coefficient between pairs of images separated by known elevational distances. The shape of this curve is more or less Gaussian, with width dependent on the elevational beam width of the transducer and the axial depth at which correlations were measured [13]. To estimate the distance between two US image patches, their correlation coefficient is subsequently measured and the corresponding elevational distance estimate is read off the curve. This process is subject to error, especially in the nonlinear portions of the curve

(very high and very low correlations) [12]. This results in significant large scale drift error when attempting to estimate the relative positions of a large number of frames. Large scale accuracy is of paramount importance when measuring the dimensions of organs from 3D reconstructions of the US data, for instance.

Each pair of frames (not necessarily subsequent) in an US scan provides correlation measurements which can be used for estimating their positions. Recent work [6] has attempted to reduce the impact of measurement errors by averaging independent interleaved reconstructions using only correlation measurements lying on the “well-behaved” (i.e. linear) portion of the decorrelation curve. This paper presents an alternative approach which allows for more measurements to be exploited simultaneously, with the goal of further reducing drift error. Measurement uncertainty is represented by a probabilistic speckle decorrelation (PSD) model which captures statistical information available at calibration time but not represented by the average decorrelation curve. The model is used within a maximum likelihood estimation framework to fuse multiple correlation measurements of arbitrary quality in order to reduce residual uncertainty for the benefit of additional processing steps such as 3D interpolation. Preliminary results obtained with parallel frames of simulated and phantom US data show the approach to be effective at limiting drift error over long image sequences.

The remainder of this paper is structured as follows. Section 2 describes the PSD model. The maximum likelihood data fusion method is presented in section 3. Finally, the experimental results on simulated and phantom US data are discussed in section 4.

2 Probabilistic Speckle Decorrelation Model

Inspired by the approach taken in [10], the approach proposed here begins with the subdivision of the US image into a number of small patches, each having its own speckle decorrelation model (see figure 1). This accounts for the variation of correlation length with axial depth. Moreover, such a subdivision yields as many correlation measurements per frame pair as there are patches: enough to estimate yaw and tilt as well as elevational translation.

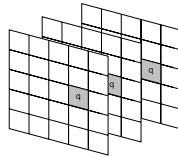


Fig. 1. The images are divided into non-overlapping patches. Corresponding patches (e.g. the highlighted patches labeled q) in different images are used to build local decorrelation models and are treated as mini-frames when estimating elevational positions.

Given two correlated random signals X and Y with correlation coefficient ρ_0 , the sample correlation coefficient of realisations x and y of X and Y , respectively, of length N is given by

$$\rho(x, y) = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{[N \sum x_i^2 - (\sum x_i)^2][N \sum y_i^2 - (\sum y_i)^2]}}. \quad (1)$$

For multiple realisations of X and Y , the average sample correlation coefficient should tend towards the nominal value ρ_0 . However, the individual sample correlation coefficients will exhibit some variability due to the finite length of the realisations x and y . This is true of any kind of random signal, including US image patches. In this case, the length N corresponds to the number of pixels.

The uncertainty intrinsic to correlation measurements can be modeled explicitly by a probability density function $p(\rho|\delta, q)$ relating the sample correlation coefficient ρ to elevational separation δ in patch q . Assuming that the relationship between elevational separation and the nominal correlation coefficient ρ_0 is one-to-one, as depicted by the decorrelation curves used in [4,13,6], $p(\rho|\delta, q) = p(\rho|\rho_0, q)$. The statistical variability of the sample correlation coefficient depends on ρ_0 , and on the statistical distributions of X and Y .

$p(\rho|\delta, q)$ can be estimated from data samples acquired by scanning a speckle phantom at regular distance intervals at calibration time. These are the same data required to construct decorrelation curves used in related work. While the decorrelation curve represents the *average* ρ as a function of distance, the probability density function $p(\rho|\delta, q)$ also captures higher order statistics of ρ , such as variance. Since no theoretical results are available concerning the form of $p(\rho|\delta, q)$, it is estimated using an empirical maximum entropy method by Baker [2].

From a set of correlation measurements $\{\rho_i\}, i = 1, \dots, N$, acquired between frames separated by a distance δ in patch q , Baker's method computes a number of probability densities from the exponential family given by

$$p_K(\rho|\delta, q; \mu) = p_0(\rho) \exp \left(\lambda_0(\mu) + \sum_{k=1}^K \lambda_k(\mu) \rho^k \right), \quad (2)$$

where p_0 is a uniform probability density function over the range D of the data (here, $D = [-1, 1]$) and μ is a vector of the sample moments of the data of order 1 to K . The normalisation constant $\lambda_0(\mu)$ is given by

$$\lambda_0(\mu) = -\ln \left\{ \int_D p_0(\rho) \exp \left(\sum_{j=1}^K \lambda_j(\mu) \rho^j \right) d\rho \right\}, \quad (3)$$

and the parameters $\lambda(\mu)$ are obtained from the first K sample moments by solving the system of nonlinear equations

$$\frac{\int_D \rho^k p_0(\rho) \exp \left(\sum_{j=1}^K \lambda_j(\mu) \rho^j \right) d\rho}{\int_D p_0(\rho) \exp \left(\sum_{j=1}^K \lambda_j(\mu) \rho^j \right) d\rho} = \mu_k, k = 1, \dots, K. \quad (4)$$

K is then selected such that the Akaike information criterion (AIC) [1],

$$AIC = 2K - 2 \sum_{j=1}^N \ln p_K(\rho_j|\delta, q; \mu), \quad (5)$$

is minimised. This application of the AIC encourages model goodness of fit while discouraging model complexity [2]. Using this method, conditional probability density functions $p(\rho|\delta_j, q)$ are obtained from sample correlation coefficients measured between frames of elevational separation $\delta_j = j\delta_0$ at patch q , where δ_0 is the elevational distance between consecutive frames in the calibration scan. The PSD model is illustrated by an example in figure 2.

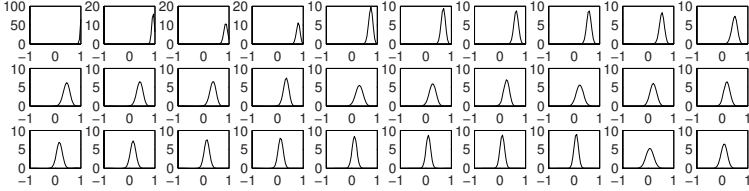


Fig. 2. The PSD model for an US image patch q obtained from speckle phantom data. Each plot is the estimated probability density function $p(\rho|\delta, q)$ (vertical axis) of the correlation coefficient ρ (horizontal axis) corresponding to elevational separation δ , with δ starting at 0.05 mm and increasing at a rate of 0.05 mm per plot from left to right and top to bottom.

Given correlation measurement ρ , the likelihood $p(\rho|\delta, q)$ for the unknown elevational separation δ can be estimated from the PSD model. The model provides samples of the likelihood at discrete values of δ corresponding to the regular intervals used for calibration. Interpolation is used to estimate the log-likelihood $L(\rho|\delta, q) = \ln p(\rho|\delta, q)$ at arbitrary δ .

3 Estimation of Elevational Separations

Consider a set of US images whose relative positions must be estimated. In this paper, it is assumed that: (1) there is no in-plane probe motion or the images have been correctly aligned by some registration procedure; (2) the frames do not intersect; (3) the motion of the probe in the elevational direction is monotonic; (4) the images are made of fully developed speckle. These assumptions may be relaxed through the use of complementary algorithms [5,6] as an add-on to the proposed drift-reduction scheme. They are adopted here for simplicity, and in the experimental protocol to minimise the influence of sources of error outside the control of the proposed method in the analysis of results.

Subdividing the frames in the data set into M patches corresponding to those used to define the speckle decorrelation model yields M individual US data sets consisting of “mini-frames” the size of the individual image patches (see figure 1). The first problem considered is that of estimating the maximum likelihood positions Z_i , $i = 1, \dots, n$ along the elevational direction of n mini-frames corresponding to patch q , with respect to a reference mini-frame with position $Z_0 = 0$. A correlation measurement ρ_{ij} between mini-frames i and j provides an uncertain measurement of the elevational separation $\delta_{ij} = |Z_i - Z_j|$ between

them through the log-likelihood $L(\rho_{ij}|\delta_{ij}, q)$. Enforcing the assumption of motion monotonicity, the absolute value is dropped and $\delta_{ij} = Z_i - Z_j, i > j$. Assuming mutual independence of measurements given the the configuration of the frames, the maximum likelihood position vector \mathbf{Z} is given by

$$\mathbf{Z}^* = \underset{\mathbf{Z}}{\operatorname{argmax}} \sum_{i>j} L(\rho_{ij}|Z_i - Z_j, q). \quad (6)$$

This is a difficult non-linear optimisation problem for arbitrary L . However, it was observed the log-likelihood terms in (6) are generally unimodal, suggesting that a Gaussian approximation of L is both possible and useful. This amounts to assuming that $\delta_{ij} = Z_i - Z_j$ is Gaussian with mean $\bar{\delta}_{ij} = \underset{\delta_{ij}}{\operatorname{argmax}} L(\rho_{ij}|\delta_{ij}, q)$, and variance $\sigma_{ij}^2 = -1/\left(\frac{d^2}{d\delta_{ij}^2} L(\rho_{ij}|\bar{\delta}_{ij}, q)\right)$. The approximate solution $\tilde{\mathbf{Z}}^*$ is

$$\tilde{\mathbf{Z}}^* = \underset{\mathbf{Z}}{\operatorname{argmin}} \sum_{i>j} \frac{(Z_i - Z_j - \bar{\delta}_{ij})^2}{\sigma_{ij}^2}. \quad (7)$$

This simpler optimisation problem was addressed for the context of robot localisation and can be solved analytically [8]. Re-expressing (7) in matrix form,

$$\tilde{\mathbf{Z}}^* = \underset{\mathbf{Z}}{\operatorname{argmin}} (\bar{\boldsymbol{\delta}} - \mathbf{H}\mathbf{Z})^T \mathbf{C}^{-1} (\bar{\boldsymbol{\delta}} - \mathbf{H}\mathbf{Z}), \quad (8)$$

where $\bar{\boldsymbol{\delta}}$ is the vector of all distance measurements, \mathbf{H} is a matrix consisting exclusively of 0, 1 and -1 entries expressing the linear relationships between distances and absolute positions, and \mathbf{C} is a diagonal covariance matrix made of all the σ_{ij} . The solution to the problem of (8) is [8]

$$\tilde{\mathbf{Z}}^* = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1} \bar{\boldsymbol{\delta}}, \quad (9)$$

whose computation is simplified by \mathbf{C} being diagonal due to the assumed independence of measurement errors. The approach also allows for the computation of the residual uncertainty in \mathbf{Z} , which could eventually be used to embed uncertainty in tasks such as volume interpolation or re-slicing.

Having obtained an elevational position estimate for each mini-frame, knowing that all image patches belonging to the same frame of the original data set lie on a plane provides an additional constraint. This is enforced by calculating the final position of each full frame as the least squares rigid transformation mapping the positions of the patch centers from every frame to the first.

4 Experiments

In order to demonstrate the feasibility of the proposed approach with data from a clinical US scanner, experiments were carried out on both simulated US imagery

and real US data of a speckle phantom. The simulations were run on a parallel processing cluster using Field II [7], with a 3.5 MHz linear transducer scanning a moving speckle phantom at a depth of 6 cm, and the resulting RF data were scan converted and log compressed to emulate the action of a clinical US scanner. The real imagery was acquired through a video frame grabber connected to an Acuson Cypress ultrasound system using a 5 MHz sector probe at a depth setting of 2.7 cm. The probe was moved using a sub-millimeter positioning device which approximately restricted motion to the elevational direction. Calibration scans were obtained for both imaging devices, at intervals of 0.1 mm for the simulations and 0.05 mm for the phantom data. The simulated images were divided into 64 patches of 60x60 pixels and the real images were divided into 19 patches of 50x50 pixels. The log compression was reversed using the technique presented in [11].

The PSD method for computing out-of-plane probe motion was applied to 16 different parallel image sequences. A minimum correlation threshold was defined for each patch such that the lowest 20% of the range of correlation values observed during calibration was cut off. A maximum number of measurements per frame was determined as the average number of consecutive frames needed to reach this threshold. This roughly amounts to assuming that variations in probe velocity are small, and reduces the sensitivity of the method to local variations in speckle decorrelation rates.

The PSD approach was compared to two base-line methods, both of which rely only on the average decorrelation curve. The first, nearest neighbour (NN), consists in positioning each frame relative to its immediate predecessor using only the correlations between subsequent frames. The second, shifting reference (SR), consists in positioning each frame using its correlation to a reference frame until the correlation falls below the minimum value or the maximum number of frames is exceeded, at which point the last frame positioned becomes the reference. The thresholds involved are the same as for the PSD method.

The accuracy of the recovered frame positions (which include elevational translation from a global reference, frame yaw and frame tilt) was measured in terms of the average RMS error between the inferred 3D positions of the centers of the image patches and the ground truth. Results are summarised in table 1.

Predictably, the NN strategy exhibits poor performance on long sequences due to the accumulation of large relative error at every frame. The PSD and SR approaches both did much better, with PSD generally outperforming SR and relative improvement varying from -4.5% (Sequence 6) to $+62.6\%$ (Sequence 11). Detailed results for these two extreme cases are shown in figure 3. Qualitatively, the PSD method yields more stable error in elevation, yaw and tilt over time, implying that in addition to reducing large scale drift error, the proposed method improves local accuracy. In the context of 3D volume interpolation, this should lead to more accurate large scale measurements (within the limitations imposed by the sensorless framework) and better rendition of local tissue structure.

Table 1. Average RMS error (mm) for different test image sequences reconstructed with the proposed PSD approach, the NN approach and the SR approach. The best result for each sequence is in bold font. The frame positions in the random simulated sequences were sampled at exponentially distributed intervals with mean 0.15 mm.

Sequence	Type	Description	RMS error (mm)		
			PSD	NN	SR
1	Simulated	178 frame calibration scan	0.038	0.667	0.051
2	Simulated	67 frames 0.3 mm apart	0.087	0.152	0.133
3	Simulated	150 frames 0.15 mm apart	0.059	0.564	0.074
4	Simulated	150 frames 0.15 mm apart	0.066	0.560	0.074
5	Simulated	134 frames 0.05 mm apart	0.036	1.400	0.048
6	Simulated	128 frames, random	0.114	0.606	0.109
7	Simulated	130 frames, random	0.078	0.641	0.099
8	Simulated	123 frames, random	0.105	0.541	0.188
9	Simulated	149 frames, random	0.095	0.757	0.113
10	Simulated	149 frames, random	0.097	0.752	0.113
11	Simulated	134 frames, random	0.104	0.609	0.278
12	Real	101 frame calibration scan	0.036	0.124	0.071
13	Real	111 frames, 0.05 mm apart	0.171	0.196	0.168
14	Real	101 frames, 0.03 mm apart	0.086	0.145	0.100
15	Real	126 frames, 0.04 mm apart	0.165	0.253	0.180
16	Real	201 frames, 0.05 mm apart	0.221	0.515	0.352

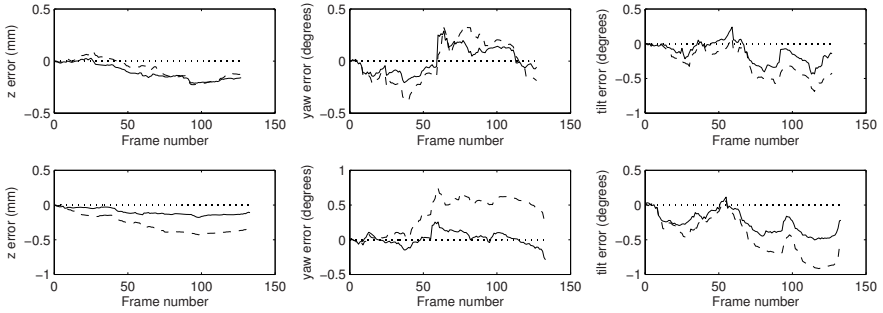


Fig. 3. Error in frame elevational translation, yaw and tilt as a function of frame number for image sequences 6 (top) and 11 (bottom). The solid line plots the results for the PSD approach, and the dashed line represents the SR approach.

5 Conclusions

This paper presented a probabilistic model for elevational speckle decorrelation in US imagery which embeds measurement uncertainty into the recovery of positional information. The model was applied to the task of fusing multiple correlation measurements obtained from long US image sequences in an attempt to recover their relative out-of-plane separations. According to the experimental

results presented in this paper, the approach is promising as it leads to a reduction of drift compared to methods which only use one correlation measurement per frame. Future work will investigate the inclusion of tissue models in the proposed probabilistic formulation in order to robustly account for local variations in scatterer organisation and density in the scanned subject, as well as global variations which occur outside the context of fully developed speckle [5].

Acknowledgments. Catherine Laporte is funded by a NSERC scholarship.

References

1. Akaike, H.: A new look at the statistical model identification. *IEEE T. Automat. Contr.* 19(6), 716–723 (1974)
2. Baker, R.: Probability estimation and information principles. *Struct. Saf.* 9, 97–116 (1990)
3. Chang, R.-F., Wu, W.-J., Chen, D.-R., Chen, W.-M., Shu, W., Lee, J.-H., Jeng, L.-B.: 3-D US frame positioning using speckle decorrelation and image registration. *Ultrasound Med. Biol.* 29(6), 801–812 (2003)
4. Chen, J.-F., Fowlkes, J.B., Carson, P.L., Rubin, J.M.: Determination of scan-plane motion using speckle decorrelation: theoretical considerations and initial test. *Int. J. Imag. Syst. Tech.* 8(1), 38–44 (1997)
5. Gee, A.H., Housden, R.J., Hassenpflug, P., Treece, G.M., Prager, R.W.: Sensorless freehand 3D ultrasound in real tissue: speckle decorrelation without fully developed speckle. *Med. Image Anal.* 10(2), 137–149 (2006)
6. Housden, R.J., Gee, A.H., Treece, G.M., Prager, R.W.: Sensorless reconstruction of unconstrained freehand 3D ultrasound data. *Ultrasound Med. Biol.* 33(3), 408–419 (2007)
7. Jensen, J.A.: Field: a program for simulating ultrasound systems. *Med. Biol. Eng. Comput.* 34, 351–353 (1996)
8. Lu, F., Milios, E.: Globally consistent range scan alignment for environment mapping. *Auton. Robots* 4(4), 333–349 (1997)
9. Mercier, L., Langø, T., Lindseth, F., Collins, D.L.: A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound Med. Biol.* 31(4), 449–471 (2005)
10. Prager, R.W., Gee, A.H., Treece, G., Cash, C.J.C., Berman, L.H.: Sensorless freehand 3-D ultrasound using regression of the echo intensity. *Ultrasound Med. Biol.* 29(3), 437–446 (2003)
11. Prager, R.W., Gee, A.H., Treece, G.M., Berman, L.H.: Decompression and speckle detection for ultrasound images using the homodyned k-distribution. *Pattern Recogn. Lett.* 24, 705–713 (2003)
12. Smith, W., Fenster, A.: Statistical analysis of decorrelation-based transducer tracking for three-dimensional ultrasound. *Med. Phys.* 30(7), 1580–1591 (2003)
13. Tuthill, T.A., Krücker, J.F., Fowlkes, J.B., Carson, P.L.: Automated three-dimensional US frame positioning computed from elevational speckle decorrelation. *Radiology* 209(2), 575–582 (1998)