

Liver Feature Motion Estimation in Long High Frame Rate 2D Ultrasound Sequences

Tuathan O'Shea¹, Jeff Bamber¹ and Emma Harris¹

¹ Joint Department of Physics, Institute of Cancer Research & Royal Marsden NHS Foundation Trust, London and Sutton, UK
tuathan.oshea@icr.ac.uk

Abstract. This study investigates the use of a 2D normalized cross-correlation (NCC)-based algorithm to estimate *in vivo* motion of liver features in 2D B-mode ultrasound (US) images. Datasets included 23 volunteer imaging sequences, each containing first frame annotated points of interest (POI). Images had a range of spatial (0.28 – 0.71 mm) and temporal (11 – 25 Hz) resolution. Image quality was also highly variable. A 2D block-matching algorithm was developed to track POI motion throughout the imaging sequence. A correlation and displacement thresholding tracking approach, which used knowledge of previous displacement and (1) linear extrapolation, (2) a regularizing sinusoidal breathing model or (3) hybrid fixed-reference / incremental tracking was used to account for potential tracking errors. The overall mean error in vessel tracking was 2.15 ± 2.7 mm. This approach to motion estimation shows promise for applications such as radiation therapy tumor tracking.

1 Introduction

This study investigates the estimation of liver feature motion in variable quality volunteer 2D ultrasound amplitude demodulated data. In conformal radiation therapy, some form of (intra-fraction) motion management is often required [1]. If motion cannot be minimized using a method such as respiratory gating [2], then this motion should be tracked in as close to real-time as possible. Tracking cardiac and respiratory induced motion requires an imaging method which samples the target position with an adequate temporal resolution. In radiation therapy, most current tracking systems are based on kV x-ray imaging [3], [4]. Ultrasound has two major advantages over these methods: (i) it does not impart ionizing radiation (imaging dose) and (ii) it allows the visualization of soft tissue. Correlation-based techniques have been used to investigate ultrasound speckle and feature-based motion tracking. Ultrasound speckle tracking of respiratory induced phantom motion and *in vivo*

feature-based tracking has been studied [5]. Good agreement (mean absolute difference < 2 mm) was found between tracked and manually annotated displacements using a mechanically swept 3D probe limited to a 0.5 Hz imaging rate. Lediju Bell et al. [6] used a 2D matrix array transducer to acquire *in vivo* liver motion data from three volunteers at imaging rates of up to 48 Hz. In the study, volumetric data was acquired at high imaging rates without the restriction of a mechanically swept ultrasound transducer. It was found that volume rates of 8 to 12 Hz were required to track cardiac and respiratory induced liver motion. In many instances out-of-plane motion is small and 2D imaging is a valid approach to tissue motion estimation. De Luca et al. [7] presented a scale adaptive block-matching approach to liver vessel tracking in long 2D ultrasound sequences. The method achieved a mean tracking accuracy of < 1 mm.

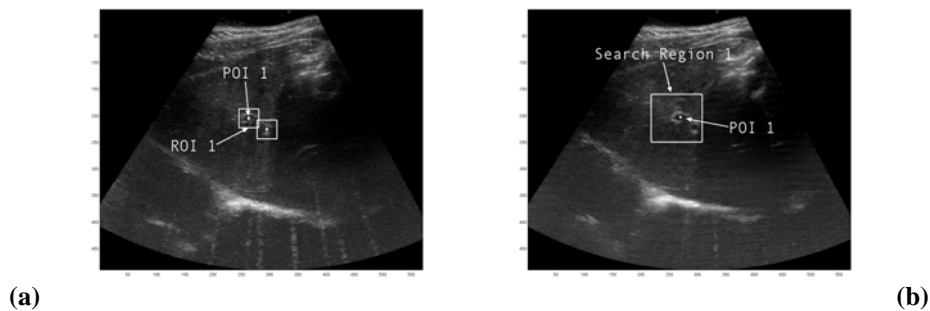


Fig. 1. Ultrasound B-mode data for one of the volunteers to illustrate image quality and method employed to track liver features. The first frame and annotated points of interest (POI) are shown (a). A region of interest (ROI) is defined around each POI. A correlation-based block matching algorithm was used to locate this same POI, within a larger search region, in a subsequent ultrasound frame (b).

In the current study we employ a 2D correlation-based block-matching algorithm to track features (blood vessels center-of-mass) in 2D B-mode ultrasound image sequences (from 23 volunteers). The tracking code was applied to ultrasound data from three different scanners / transducers with a range of image resolutions. We investigated non-incremental (fixed reference) tracking. For non-incremental tracking, the mean inter-frame displacement is greater than for incremental tracking and there is a higher probability of tissue deformation and rotation [4], however, incremental tracking can be prone to drift error accumulation. In speckle tracking, it is known that tissue deformation and rotation corrupt the speckle pattern [8]. Low imaging rates have limited speckle-tracking accuracy to the extent that adequate *in vivo* motion estimation has been obtained by tracking features (blood vessels) only [5]. In the current study, data was acquired at high frame rates (11 – 25 Hz) such that inter-frame rotation and deformation is expected to be small and, additionally, we

tracked tissue features which were generally highly visible throughout the imaging sequence. Nevertheless, in 2D images out-of-plane motion can be an issue. The tracking performance of our 2D correlation-based automated tracking code was quantified by comparison with manual annotations of the tissue features throughout the ultrasound sequence.

2 Materials and method

2.1 Ultrasound data

B-mode ultrasound data was provided by the CLUST 2014 (“*MICCAI Challenge on Liver Ultrasound Tracking*”) [9]. 2D volunteer liver image data from 23 patients (MED and ETH datasets) was acquired by one of three ultrasound systems (Siemens Antares, DiPhAs Fraunhofer and Zonare z.one). Data had varying spatial (0.28 – 0.71 mm) and temporal resolution (11 – 25 Hz) and sequences lasted from 121.2 – 580.64 s. Examples of the volunteer image data from the Siemens Antares are shown (Fig. 1). Some of the B-mode data contained what appeared to be electronic interference (Fig. 1 (a)) and large shadowing artifacts (Fig. 1 (a) and (b), from ETH-05 dataset). Two of the 23 volunteer datasets (MED-04 and ETH-05) were provided with ground truth annotations (of liver blood vessels) throughout the acquisition sequence which were used to assess tracking code performance and enable code development. Annotations were provided in the following form: frame number, x-pixel (lateral position) and y-pixel (axial position). For the remaining datasets, liver features (blood vessels centers) were annotated in the first frame only. The number of annotations per image sequence ranged from one to five liver features (cf. Table. 1).

2.2 Tracking code

To detect the motion of liver blood vessels center-of-mass an automated (serial) tracking code was developed in MATLAB R2011b (MathWorks, Inc. MA, USA). The code was based on the use of normalized cross-correlation (NCC) as a similarity metric between the current and a reference ultrasound frame. A reference region of interest (*ROI*) was defined around each annotation or point-of-interest (*POI*) in the first ultrasound frame. In subsequent ultrasound frames a larger search region was defined. The position of maximum correlation from the NCC code was used to identify the new position of the the *ROI*. To improve the precision of coarse pixel displacement estimates, the sub-pixel (fine) displacement was calculated by fitting the maximum correlation and two surrounding values in the correlation matrix with a

second order function and finding the peak (i.e. when the slope, $m == 0$). The tracking code output tracking results in the format: frame number, x-subpixel (lateral position) and y-subpixel (axial position).

To track annotated POI motion three tracking methods were developed and each was used to track features in a subset of the US sequences: (1) a simple correlation and displacement thresholding (fixed-reference) tracking approach, which used knowledge of previous displacement and linear extrapolation [10], (2) a regularized model-based tracking code using a sinusoidal breathing model (which was applied to two of the US sequences to investigate improvements in tracking results) and (3) a hybrid fixed reference / incremental (updated ROI) tracking approach (further details below). We visually assessed which US sequences were best suited to which method by plotting (overlying) the raw displacement tracking code output (vessel center-of-mass) on the current ultrasound frame in “real-time”. In this way, tracking errors due to, for example false matches within the search area, became obvious. The tracking method which gave the best (visually assessed) results for a particular US sequence was then selected.

For method (1), above, the code monitored the inter-frame displacement ($mdisp$) and correlation ($mcorr$) (via user-specified thresholds) and limited the maximum displacement. In cases when the inter-frame displacement was larger (and $mcorr$ smaller) than the threshold values, the current displacement estimate was replaced with displacement predicted by linear extrapolation using the previous two displacement estimates.

For method (2), a model-based (predictive) regularization scheme was developed and used to track feature motion in two of the volunteer image sequences (5 and 14). After a user-specified period of time (number of frames), t , the tracking code fit the previous t seconds of raw motion estimation data (median filtered, $n = 3$) with a well known respiratory motion model [11] and this was used to infer the current displacement of the feature (ROI). During time periods which exhibited potential tracking errors (as monitored by $mcorr$ and $mdisp$), the model-predicted displacement could be used to infer the new position of the ROI .

For US sequences 18 – 23, out-of-plane motion, rotation and deformation changed the images to such an extent that standard fixed-reference tracking was not feasible. Instead the code monitored the correlation ($mcorr$) and displacement ($mdisp$) value and used linear extrapolation to calculate displacement and update the reference region (ROI) if the values were below the user-defined thresholds (method (3)).

2.3 Analysis

The tracking code was developed and used to track motion of first frame annotated features (POI) in twenty-three volunteer image sequences. Automated tracking

results were evaluated by comparison with manual annotations of liver feature (vessels) throughout each image sequence which were provided after automated tracking was complete. Tracking accuracy was evaluated using the Euclidean distance between tracked points and manually annotated points which was summarized by the mean and standard deviation. The run-time performance of the tracking code was also evaluated by calculating the average run-time for all cases.

Table 1. Volunteer B-mode data sequence spatial and temporal resolution, number of points-of-interest (*POI*) and error (mean \pm standard deviation) in tracking code motion estimation (as quantified relative to manual annotations of liver feature motion). Listed for *POI* with maximum (mean \pm standard deviation) error for that specific volunteer data-set only (*POI* listed in brackets). Patients with annotations available throughout the imaging sequence are highlighted in bold

Sequence number	Name	Im. Res. [mm]	Imaging rate [Hz]	No. of <i>POI</i>	Tracking method	Track. Error mean \pm SD [mm]
1	ETH-01	0.71	25	1	(i)	1.9 \pm 0.4 (1)
2	ETH-02	0.40	16	1	(i)	0.5 \pm 0.2 (1)
3	ETH-03	0.36	17	3	(i)	1.6 \pm 1.0 (1)
4	ETH-04	0.42	15	1	(i)	0.9 \pm 1.0 (1)
5	ETH-05	0.40	15	2	(ii)	1.1 \pm 1.1 (1)
6	ETH-06	0.37	17	2	(i)	0.6 \pm 0.3 (2)
7	ETH-07	0.28	14	1	(i)	0.7 \pm 0.3 (2)
8	ETH-08	0.36	17	2	(i)	0.9 \pm 0.4 (2)
9	ETH-09	0.40	16	2	(i)	0.8 \pm 0.6 (2)
10	ETH-10	0.40	15	4	(i)	1.2 \pm 1.5 (3)
11	MED-01	0.41	20	3	(i)	1.8 \pm 0.6 (3)
12	MED-02	0.41	20	3	(i)	1.8 \pm 1.8 (2)
13	MED-03	0.41	20	4	(i)	2.3 \pm 1.3 (2)
14	MED-04	0.41	20	3	(ii)	3.3 \pm 1.7 (3)
15	MED-05	0.41	20	3	(i)	2.3 \pm 1.3 (2)
16	MED-06	0.41	20	3	(i)	6.3 \pm 7.5 (3)
17	MED-07	0.41	20	3	(i)	5.3 \pm 4.2 (1)
18	MED-08	0.41	20	2	(iii)	4.9 \pm 3.5 (2)
19	MED-09	0.41	20	5	(iii)	11.7 \pm 5.6 (5)
20	MED-10	0.41	20	4	(iii)	6.6 \pm 3.9 (1)
21	MED-13	0.35	11	3	(iii)	4.4 \pm 1.4 (3)
22	MED-14	0.35	11	3	(iii)	3.4 \pm 2.0 (3)
23	MED-15	0.35	11	1	(iii)	2.4 \pm 1.4 (1)

3 Results and discussion

The accuracy with which the automated tracking code could track multiple liver features in the 23 volunteer B-mode imaging sequences is summarized in the final column of table 1 (the values for the POI exhibiting the largest tracking error is listed). The lowest motion estimation error (0.5 ± 0.2 mm) was for an ultrasound sequence containing a relatively large, single centrally located blood vessel. For many of the ultrasound sequences, there was relatively small out-of-plane motion or deformation of the tracked features and therefore a fixed-reference NCC-based approach was adequate. However, on occasions the tracking code detected a false match within the search region, for example the hyperechogenic blood vessel wall would disappear (out-of-plane) leaving only the hypoechogenic vessel centre (blood) and the NCC code would “find” another hyperechogenic feature (i.e. generate a false match) within the search region. When the correlation value (inter-frame displacement) for a POI decreased (increased) below a user-defined threshold (e.g. the current NCC value was < 0.8 , inter-frame displacement < 3 mm), linear extrapolation was used to account for the vessel displacement in the time interval. While linear extrapolation may not be the most accurate method [9], it appears adequate in cases when there are no times of sustained tracking errors and at the high frame rates of these data sets.

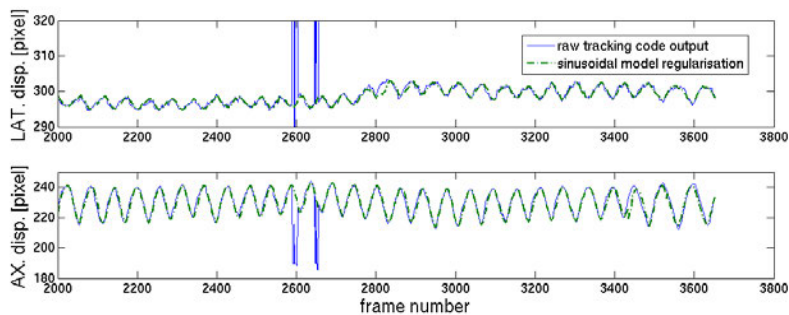


Fig. 2. Example of fixed-reference tracking code raw motion estimation (solid blue line) exhibiting some obvious tracking errors (false matches) and application of model-based regularization to improve motion estimation results (dashed green line)

For data sequences 5 and 14, a fixed reference sinusoid model-based tracking approach was adopted. The model was used to fit the last t seconds of ultrasound data to attempt to improve tracking results (Figure 2). We used a value of $t = 5$ s which is the approximate average breathing period of most patients/volunteers. The model was found to work well for relatively regular breathing motion (i.e. sequences 5 and 14) and could detect large tracking errors (“false matches”) in drifting breathing signals. However, when the algorithm was applied to US sequences which exhibited

large amounts of out-of-plane motion or deformation, it failed to model the liver feature motion.

Imaging sequences 16 – 23 exhibited relatively large out-of-plane motion and vessel deformation and this is reflected by the increased mean tracking error (Table 1). While it has previously been found that NCC-based incremental tracking of liver features in 3D US images was not as accurate as fixed-reference tracking [5], large changes over long imaging sequences meant standard fixed-reference tracking could not be used. Incremental tracking is also known to suffer from drift [5]. A hybrid fixed reference / incremental tracking method which updated the reference ROI if a low correlation value was detected was found to marginally improve results but some large errors remained (e.g. volunteer sequence 9 mean error from POI 5: 11.7 mm). In the future, motion estimation accuracy could benefit from a code which combines model-based regularization and hybrid tracking. The sinusoidal model employed in this work increased run-time by a factor of 3. Other methods, such as linear regression prediction which could be used to update the ROI location and account for displacement during times of known errors, are known to be fast and accurate [10]. Detailed analysis of the optimum use of inter-frame correlation and displacement thresholds to infer tracking errors, would help limit drift accumulation during hybrid/incremental tracking.

The MATLAB tracking code took approx. 190 ms per POI per frame pair (on a single Intel® Core™ 2 Duo E6750 2.66GHz CPU) using a 150 x 150 pixel search region. With model fitting disabled the run-time was approximately 60 ms per POI per frame pair with (*mdisp* and *mcorr* thresholding enabled).

4 Summary and conclusion

Long sequences of B-mode volunteer ultrasound data of variable quality have been used to develop automated tracking algorithms which were used to track annotated liver features in 23 US sequences. The tracking code used correlation and displacement thresholds to identify potential errors and linear extrapolation to account for displacement during these events. For two US sequences, previous motion estimation data were used to generate a breathing sinusoidal model which was used to account for potential future tracking errors. For six of the US sequences, out-of-plane motion, rotation and deformation meant standard fixed reference tracking was not feasible. Instead a hybrid fixed reference/incremental tracking approach was employed. Comparison of manually and automatically tracked displacement of liver features (blood vessels) for sequences (generally longer than a standard external beam radiation therapy delivery) have shown promise (overall mean error 2.15 ± 2.7 mm). Future work will investigate the optimum code parameters, including other regularization models and methods [10], and address the relatively long run-time.

The tracking code model could also be extended to allow future prediction and account for tracking algorithm latencies should this be a significant issue for radiation dose delivery.

References

1. Webb S.: Motion effects in (intensity modulated) radiation therapy: a review. *Physics in medicine and biology* 51 (2006) (13) R403
2. von Siebenthal M., Szekely G., Lomax A. and Cattin P.: Systematic errors in respiratory gating due to intrafraction deformations of the liver. *Medical Physics* 34 (9) (2007) 3620-3629
3. Hoogeman M., Prévost J. B., Nuyttens J., Pöll J., Levendag P., and Heijmen B.: Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files. *International Journal of Radiation Oncology* Biology* Physics* 74 (1) (2009) 297-303
4. Ng J., Booth J., Poulsen P., Fledelius W., Worm E., Eade T., Hegi F., Kneebone A., Kuncic Z., and Keall P.: Kilovoltage intrafraction monitoring for prostate intensity modulated arc therapy: first clinical results. *International Journal of Radiation Oncology* Biology* Physics* 84 (5) (2012) e655-e661
5. Harris E., Miller N., Bamber J., Symonds-Taylor R. and Evans P.: Speckle tracking in a phantom and feature-based tracking in liver in the presence of respiratory motion using 4D ultrasound. *Physics in medicine and biology* 55 (12) (2010) 3363
6. Bell M. A. L., Byram B., Harris E., Evans P. and Bamber J.: In vivo liver tracking with a high volume rate 4D ultrasound scanner and a 2D matrix array probe. *Physics in medicine and biology* 57 (5) (2012) 1359
7. De Luca V., Tschannen M., Székely G. and Tanner C.: A Learning-Based Approach for Fast and Robust Vessel Tracking in Long Ultrasound Sequences. *Medical Image Computing and Computer-Assisted Intervention–MICCAI* (2013) 518-525
8. Meunier J.: Tissue motion assessment from 3D echographic speckle tracking. *Physics in medicine and biology* 43 (5) (2012) 1241
9. <http://clust14.ethz.ch/>
10. Sharp G., Jiang S., Shimizu S. and Shirato H.: Prediction of respiratory tumour motion for real-time image-guided radiotherapy . *Physics in medicine and biology* 49 425-440
11. Lujan A., Larsen E., Balter J. and Ten Haken R.: A method for incorporating organ motion due to breathing into 3D dose calculations. *Medical Physics* 26 (5) (1999) 715-720