Motion-induced error correction in ultrasound imaging.

David Lloret, Joan Serrat,
Antonio M. López and Juan J. Villanueva

Computer Vision Centre and Dept. d'Informàtica
Universitat Autònoma de Barcelona
08193 Cerdanyola, Spain

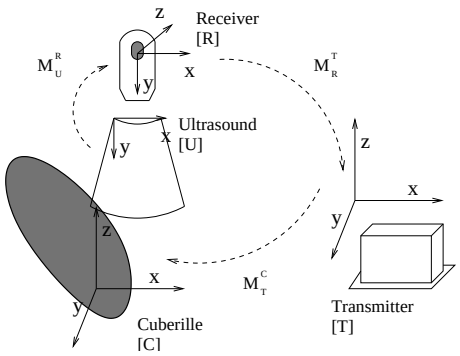

Figure 1. Coordinates systems involved in the free-hand ultrasounding

Abstract

Ultrasound imaging is a helpful aid for diagnosis in many medical specialities. In order to enhance the information available to the examiner, an upcoming approach is to compound the sequence of video images into a single 3D volume, which needs a real-time tracking of the transducer. However, errors in the positioning occur, and the compounded image becomes degraded. We propose a method which automatically corrects the positioning error, by means of a registration step between the live ecography and the compounded volume. The algorithm has been tested on sequences taken on an in-vitro human brain.

1. Introduction

Ecographic images are very useful because it is a real-time modality: the clinician is able to see immediately the region to be explored. But, because of the way images are acquired, their quality is poor compared to other modalities. 3–D ultrasound represents an improvement to the traditional viewing procedure. In this modality, all the information is located accurately in a cubical voxel array, so accurate volume or distance measurements are possible, and also many of the image processing techniques already available for conventional radiologic images.

The usual approach is called the free-hand paradigm: the transducer is moved freely along the surface to be scanned. It requires the live follow-up of the transducer, which is achieved by means of a special device called tracker. These capabilities may be of medical use in many areas. For instance, it may be useful to measure the brain shift during a neurosurgery operation, as ultrasound devices are already employed for visualisation purposes. We have investigated this matter in a number of papers [2, 5].

The goal of this paper is to fully present a method for compounding accurate 3D ultrasound images: a novel method to correct the errors in the position which cause the blurring. The method, fast, reliable and fully automatic, is based on the registration of the individual 2D frames to the compounded image. We present for visual inspection the image compounded with the corrected position.

2. Materials

A free hand system must be able to track the position of the transducer, grab the image acquired at that particular moment and then combine all the information. Ours consists of the following elements (trademarks actually employed in brackets): a) an ultrasound ecographer with video output (Siemens SONOLINE Versa Pro, with three interchangeable probes: 10, 3.5 and 6.5 Mhz). b) a 3-D tracker to track the position of the ultrasound transducer as accurately as possible (Minibird 800, Ascension Technology, Vermont). c) a computer, with two input signals: the ultrasound video image and the tracker position. It must be able to store in real time all the incoming information. The slowest of the these elements determines the final rate in frames per second of the system.

The experiments were performed on an in-vitro adult human brain, in order that the object to be scanned would resemble as close as possible to that of an actual surgery intervention. This approach had the advantage to elude the pressure and restrictions of the operating room. We chose this organ because another set of experiments, with the goal of measuring the shift of the brain during surgery, were carried out simultaneously.

3. Methods

The operation of placing each pixel from the ecography to the final world coordinates is described by means of transformation matrices. A pixel with coordinates $P_{\overline{\mathbf{x}}} = (x, y)$ in **U** will be transformed to some other coordinates $C_{\overline{\mathbf{x}}} = (x', y', z')$:

$$C_{\overline{\mathbf{x}}} = P_{\overline{\mathbf{x}}} \cdot M_U^R \cdot M_R^T \cdot M_T^C \quad (1)$$

M_U^R , so-called calibration matrix, includes the scaling parameters and the position of the top left corner of the image with regard to the centre of the receiver. Its values are constant provided the receiver's position on the transducer is kept stable and the ecography settings are not changed.

Since the scaling and position parameters are unknown, they must be estimated. Although some papers simply measure the distance between the receiver and the tip of the ecographer, it is much more reliable and precise to compute them statistically by scanning some object with known shape and size. This is known as calibration, and will be described in detail later in this section.

M_R^T contains the position and orientation of the receiver with respect to the transmitter. Therefore, the values will change as the clinician moves the transducer where the receiver is attached. M_R^T inevitably contains positioning errors, caused by interferences and the limited accuracy of the tracking device.

M_T^C is used for the sake of a proper location, orientation and scaling of the array of voxel system coordinates. After applying $M_U^R \cdot M_R^T$ to $P_{\overline{\mathbf{x}}}$ in the equation above, the resulting coordinates will likely be some distance away from the coordinate origin, and also the distance between pixels will be set by the tracker specifications.

But in general this will not be adequate; probably, the area to be scanned is rather small compared to the whole volume needed, so we would like to take into account only those coordinates near the scanning. Therefore, M_T^C will be chosen at user's will, in order to contain the area scanned and with a given scaling and orientation.

In order to calibrate the system, We decided to employ the well known crossing wires method, and to hold the transducer each time a frame had to be captured instead of live video framing. This approach permits to focus properly the field of view to include the point to be imaged, and presumably would reduce the error in positioning the tracker produced by the movement in the previous experiments. Numerical values of the registrations were consistent and reproducible.

Once the system has been calibrated, i.e., M_U^R is known, the information (image, position) can be used in many ways. Perhaps the most obvious is to compound a volume image, but others exist. Gee, for instance, in [1] describes a method to resample the data on a curved, not flat, surface, without actually producing the intermediate image. In the next section we review the process to compound the video sequence into the volume image.

4. Compounding the volumes

After the transducer has been calibrated, the spatial information accompanying each image can be employed in many different ways. One way is to combine each video frame with existing images of the same patient, thus providing the surgeon with the ability to complement the information displayed. This idea was explored in another paper [2].

Other possible applications demand the composition of the set of video images into a 3–D image of the scanned volume. The volume image was not compounded while acquiring the data, but the frames and the position were stored in disk. This procedure is very convenient because it permits further experimentation with the data.

We acquired three sequences of data, one for each transducer available. We preferred not to restrict the experiment to a single transducer, in order to obtain more general results.

In figure 2 (left column) we depict the compounded volumes for the three probes. Of course, the quality of the images is far from resembling other, more accurate modalities such as MR . However, the contours of the brain and also the sulci are clearly distinguishable, specially for distal slices. Proximal slices show features only at the external borders and around the longitudinal fissure, which are the zones the ultrasound signal could penetrate. No other anatomical issues are exposed, because the shape of the container did not permit any scanning of more proximal areas. Images on right column will be explained in next section.

Note the gaps in the images, in white, which are areas left non scanned. Also note that some areas are clearer than others. The reason is that, despite the correcting effects of taking the mean value, the image contents degrade as more slices are compounded in, and also with stronger positioning error. These areas appear blurred in the final image. In the next section, we will see more closely this effect and one possible way to suppress it.

5. Improving accuracy by means of 2–D to 3–D self-registration

The RMS error of the calibration was roughly 3 mm, about 30 pixels. With this numbers, we expected that little could be seen in the composed image. In reality, the positioning error had been overestimated because the composed image was fairly regular and coherent.

Still there was a chance to improve the results: to register each video 2–D image to the compounded 3–D volume.

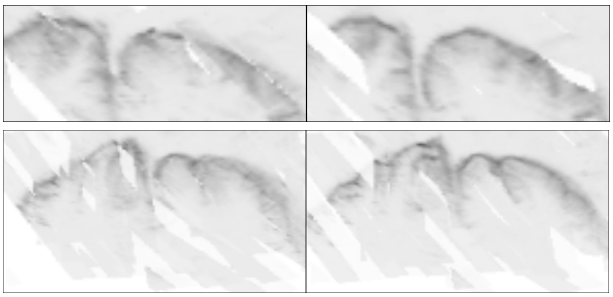

Figure 2. Rois of the compounded volume with corrected position (right) compared to the original (left).

This should be possible also because the Minibird already provides us with an initial estimation of the position of the image in the cuberille. Otherwise, the field of view would probably be too small compared to the size of the volume image to permit any registration. Now, given the initial position, only a small transformation, correcting the positioning errors of tracker, should be expected.

This process is depicted in figure 3: the B-scan during an examination (left) must adjusted to the features already existing in the compounded volume image (right). This adjustment process is called registration: the algorithm will try a explore neighbourhood areas and will choose that with best correspondence.

Following we give the details of the registration algorithm. We have already employed a similar method to other medical image registration, CT to MR volumes [6, 3], and 2D retinographies [4], but the application to 2D–3D registration is a novelty in the subject.

Similarly to any registration algorithm, there were several issues to decide:

- **the model of the transformation to optimise**

  We modelled the position error as a rigid transformation given the matrix:

  $$M^{ERR} = Tras(T_x, T_y, T_z) * Rot(\phi_x, \phi_y, \phi_z) \quad (2)$$

  And we modified equation (1) to include this adjustment:

  $$C_{\overline{\mathbf{x}}} = P_{\overline{\mathbf{x}}} \cdot M^{ERR} \cdot M_U^R \cdot M_R^T \cdot M_T^C \qquad (3)$$

  The new matrix M^{ERR} measures the error in the position of the slice.

- **the features to be compared.**

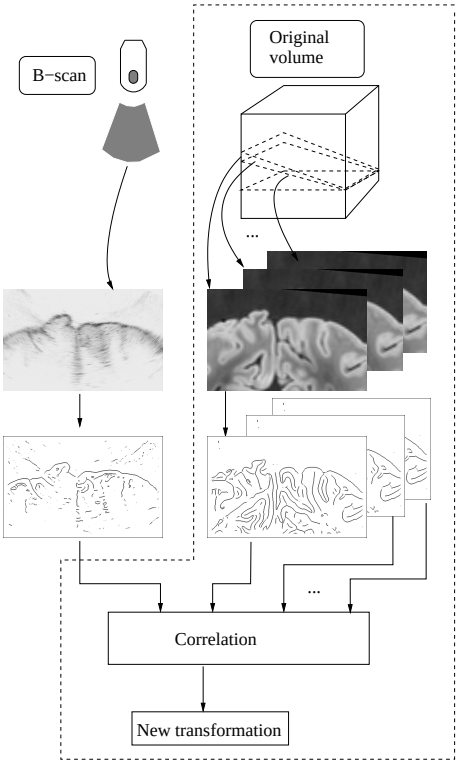

Figure 3. The iterative registration algorithm.

  The experience of our group was very valuable in this matter; similarly to our previous research, we employed a creaseness operator to extract relevant landmarks. The natural choice was to adapt the already tested algorithms, which worked for 3D–3D and 2D–2D, to work also with 2D–3D. This is, nevertheless, a very suitable choice because again the images to be compared may be very different. The creaseness image suppresses the effect of the gaps and makes them comparable.

- **the alignment estimation.**

  The correlation is suitable function, because it is fast to compute and robust to false and missing features.

- **the scheme of iterative process.**

  The iterative step would be the computation of $\mathcal{C}$, i.e. the slice to be compared in the cubic image. The iteration would modify the values of M^{ERR} , which in turn would change the contents of $\mathcal{C}$, until the desired convergence has been achieved.

The iterative scheme explained above is too simple to work for real data. In effect, false maxima occur due to artifacts, and the search is bound to get trapped. This problem

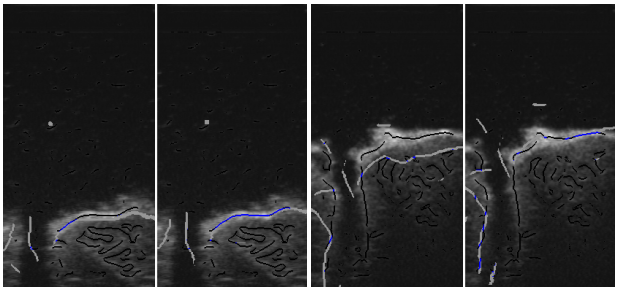

Figure 4. B-scans showing alignement before (left) and after (right) the registration for 10 MHz transducer. Creases are overdrawn in white and black.

was solved in other papers by means of an exhaustive search in the Fourier domain, which was run on the top level of a hierarchical structure. But this initial step was not suitable now, as the dimensionality of the images was different. Instead, we performed an initial 2d-2d matching.

We run successfully this algorithm for all the series we acquired. An improvement was to compute the target coordinates $C_{\overline{\mathbf{x}}}$ only for non-void pixels $P_{\overline{\mathbf{x}}}$ in the video image, which were usually only 5% of the total. The result was rewarding: the new algorithm ran within 10 seconds per frame.

6. Results

Figures 4 show the successful convergence for a few frames. In all figures, the creaseness of $\mathcal{B}$ is depicted over $\mathcal{B}$ in black, and that of the $\mathcal{C}$ in white. Left image corresponds to alignment before registration, right image after registration.

Sometimes, the algorithm fails because the compared creases are too dissimilar. Other times, as seen in last row of the previous figures, large artifacts appearing in $\mathcal{C}$ mislead the search.

We have manually classified each individual registration as valid or non valid, in order to measure the influence of the error described in the last paragraph. We have visually compared the compounded image without and with position correction. To compound the cuberille with the registered values, we have taken the choice to use the original, unregistered transformation, for those frames labelled as not properly registered. This approach permits an easier visualisation of the effects of the registration.

Figure 2 shows a selection of relevant regions of interest for experiment C. Regions shown compare favourably for the self-registered version: the contrast is clearer, there are less artifacts and contours are more continuous. This is the case for most cases, but occasionally the contrary also happens when the field of view is smaller.

7. Conclusions and future work

Our contribution in this paper has consisted on a new method to correct motion errors when compound 3D ultrasound images. The method is fast and fully automatic, and has proved acceptable accuracy. Also, we have specified the steps necessary to achieve a proper calibration of the transducer. As a future work, we plan design a filter to automatically discard wrong registration: the set of transformations obtained follows a regular fashion, thus those too large or with non-coherent directions will be safely discarded.

Acknowledgements

This paper has been partly funded by the CICYT grant contract TIC2000–1123. The authors thank Dr. Derek L.G. Hill and Dr. Donald Farr, f Guy's, King's and St Thomas' School of Medicine, King's College London, U.K., for providing help and support to make this research possible.

References

[1] A. Gee, R. Prager, and L. Berman. Non-planar reslicing for freehand 3-d ultrasound. In Springer, editor, Proceedings of the 2nd International Conference on Medical Image Computing and Computer-Assisted Intervention, volume LNCS 1679, pages 716–725, 1999.

[2] D. Lloret and D. L. Hill. System for live fusion of 2–d ultrasound scans to pre-interventional mr volumes of a patient. In M. Torres and A. Sanfeliu, editors, 8th Spanish Symposium on Pattern Recognition and Image Analysis, pages 23–24. Asociación Española de Reconocimiento de Formas y Análisis de Imágenes, 1999.

[3] D. Lloret, A. López Peña, J. Serrat, and J. Villanueva. Creaseness-based CT and MR registration: comparison with the mutual information method. Journal of Electronic Imaging, 8(3):255–262, July 1999.

[4] D. Lloret, C. Mariño, J. Serrat, A. M. López, M. G. Penedo, F. Gómez Ulla, and J. J. Villanueva. Automatic registration of full slo video sequences (submitted). IEEE Transactions On Medical Imaging, 2001.

[5] D. Lloret, J. Serrat, A. M. López, and J. J. Villanueva. Ultrasound to mr volume registration for brain sinking measurement. IEEE Transactions On Systems, Man and Cybernetics (submitted), 2001.

[6] A. López Peña, D. Lloret, J. Serrat, and J. Villanueva. Multilocal creasness based on the level set extrinsic curvature. Computer Vision and Image Understanding, 77:111–144, 2000.