

# Three-Dimensional Ultrasound Mosaicing

Christian Wachinger<sup>1,2</sup>, Wolfgang Wein<sup>1,2</sup>, and Nassir Navab<sup>1</sup>

<sup>1</sup> Computer Aided Medical Procedures (CAMP), TUM, Munich, Germany  
{wachinge, wein, navab}@cs.tum.edu

<sup>2</sup> Siemens Corporate Research, Princeton, NJ, USA

**Abstract.** The creation of 2D ultrasound mosaics is becoming a common clinical practice with a high clinical value. The next step coming along with the increasing availability of 2D array transducers is the creation of 3D mosaics. In the literature of ultrasound registration, the alignment of multiple images has not yet been addressed. Therefore, we propose registration strategies, which are able to cope with problems arising by multiple image alignment. Among others, we use *simultaneous registration* which urges the usage of *multivariate similarity measures*. In this paper, we propose alternative multivariate extensions based on a maximum likelihood framework. Experimental results show the good performance of the proposed registration strategies and similarity measures.

## 1 Introduction

At the moment, a paradigm shift takes place in ultrasound (US) imaging, moving from 2D to 3D image acquisition. The next generation of 2D array US transducers with CMUT<sup>1</sup> technology could accelerate this shift by offering superior and efficient volumetric imaging at a lower cost. From a current perspective, the only drawbacks that remain are the limited field-of-view (FOV) of the acquired images and the reflectance of the beam from structures with high acoustical impedance causing occlusion. The idea of mosaicing is to address these issues by combining the information of several images taken from different poses. The focus can rest on quality improvement by imaging the same scene from different directions, or the extension of the FOV by stitching together consecutively taken images. Whatever we are interested in, the first step is to calculate the correct global alignment for which we propose solutions in this report.

The rigid intensity-based registration that we use for the alignment is not trivial to compute because of the limited amount of overlap between the images. For mosaicing the registration scenario changes since the perfect alignment does not correspond to a maximal overlap, like it is in most cases, putting a special interest on the overlap invariance of the measures. An additional difficulty lies in the interface enhancing nature of ultrasound images, making acquisitions of the same object but from varying viewing angles not necessarily look the same. Feature-based registration methods like in [1] were discarded due to the problems of automatic salient feature point identification.

<sup>1</sup> Capacitive Micromachined Ultrasound Transducer.

### 1.1 Clinical Value of Ultrasound Mosaicing

The usage of ultrasound mosaicing provides the sonographers not just with a compounded volume of higher quality; recent studies also state a couple of other clinical advantages that come along with the extended FOV. First, the spatial relationship among structures that are too large for a single volume is easier to understand [2]. Second, sonographers have the flexibility to visualize anatomical structures from a variety of different angles [3]. Third, size and distance measurements of large organs are possible [2]. Fourth, individual structures within a broader context can be identified by having an image of the whole examination area [4]. And last, because of the increased features in the compounded view, specialists that are used to other modalities than ultrasound can better understand the spatial relationships of anatomical structures [5]; helping to bridge the gap between the modalities and making it easier to convey sonographic findings to other experts.

But it is not just the improvement of already existing workflows, the creation of high quality mosaics may also create new medical applications for ultrasound that do not yet exist at all or are reserved for other modalities.

### 1.2 Problems Statement

In the literature of ultrasound mosaicing, the global alignment of multiple images is deduced from a sequence of pairwise ones. Gee et al. [6] reduce the 3D-3D registration problem to a 2D-2D one by registering the dividing planes to each other. Poon et al. [7] use a block-based rigid and block-based warping approach for the registration. The disadvantages that come along with the usage of pairwise registrations for ultrasound mosaicing are twofold. First, by stitching together pairwise aligned images, registration errors can be accumulated leading to a non-consistent global alignment, see figure 1. Second, during the pairwise registrations only a fraction of the available information is taken into account making it prone to misregistrations. The registration is further complicated by the viewing angle dependent US images and the high demands on the overlap invariance by mosaicing.

## 2 Mosaicing Strategies

In this section, we present registration strategies that directly address the problems that arise during the mosaic creation as mentioned in section 1.2. We denote the  $n$  images by  $\mathcal{U} = \{u_1, \dots, u_n\}$  with the global transformations  $\mathcal{T} = \{T_1, \dots, T_n\}$ , and the pairwise transformation  $T_{i,j}$  between each overlapping image pair  $u_i$  and  $u_j$ .

### 2.1 Pairwise Registration with Lie Normalization

The first strategy is based on pairwise registrations and uses a consecutive normalization to reduce the accumulated error. Supposing that we would have all

correct global transformations, we could express the pairwise registration error  $\varepsilon_{i,j}$  as

$$\varepsilon_{i,j} = T_i^{-1} \cdot T_j \cdot T_{i,j}.$$

In practice, the opposite holds since we know the pairwise registrations  $T_{i,j}$  and use them to estimate the global transformations  $T_i$ . The best estimation of the global alignment is reached when the overall error is minimized. The minimization is not trivial because rigid transformations do not belong to a vector space but rather lie on a non-linear manifold forming a Lie group [8]. We use the Lie group based normalization framework, as it was proposed by [9] for the alignment of 2D optical images, to align the 3D ultrasound images. An error function  $\mu_\varepsilon$  is introduced to assign each transformation  $\varepsilon_{i,j}$  a distance value serving as score for the optimization. Assuming  $\varepsilon_{i,j}$  being a sample of the random error  $\varepsilon$  with Fréchet matrix mean identity and covariance matrix  $\Sigma_{\varepsilon\varepsilon}$ , the Mahalanobis distance that we use as error function is

$$\mu_\varepsilon^2(\varepsilon_{i,j}) = \log_{\text{Id}}(\varepsilon_{i,j})^T \cdot \Sigma_{\varepsilon\varepsilon}^{-1} \cdot \log_{\text{Id}}(\varepsilon_{i,j}).$$

The global pose estimation is expressed by the following least-squares criterion

$$[\hat{T}_1, \dots, \hat{T}_n] = \arg \min_{[T_1, \dots, T_n]} \frac{1}{2} \sum_{(i,j)} \omega_{i,j} \cdot \mu_\varepsilon^2(\varepsilon_{i,j}).$$

with the quality weights  $\omega_{i,j}$ . These weights model the quality of each pairwise registration. Since we are interested in an automated registration we use the amount of overlap as an indicator of the registration quality. The final algorithm using the Lie normalization is stated in the following listing. The registration is accepted if the total error  $\varepsilon_t = \sum_{(i,j)} \omega_{i,j} \cdot \mu_\varepsilon^2(\varepsilon_{i,j})$  is below a scenario dependent threshold  $\delta$ .

- 
1. Start with initial global transformations  $\mathcal{T} = \{T_1, \dots, T_n\}$
  2. **Do**
    - 2.1 Deduce initial pairwise transformations  $T_{i,j}$  from  $\mathcal{T}$
    - 2.2 Compute all pairwise registrations  $T_{i,j}$
    - 2.3 Estimate new  $\mathcal{T}$  from calculated  $T_{i,j}$  with Lie normalization
  3. **While** ( $\varepsilon_t > \delta$ )
  4. Return  $\mathcal{T}$
- 

## 2.2 Simultaneous Registration

The second strategy is based on simultaneous registration which is an active field of research and has so far mainly been used for population studies [10] in medical imaging. The principle of simultaneous registration is to consider all available images at the same time during the registration process. The registration framework has to be extended to deal with multivariate similarity measures and the simultaneous optimization of  $n \cdot 6$  parameters. Up to now, only a limited number

**Table 1.** Summary of bi- and multivariate similarity measures in shortened notation

	Pairwise	Semi-Simultaneous	Full-Simultaneous	Voxel-Wise
SSD	$\mathbb{E}[(u - v^\perp)^2]$	$\sum_{i=2}^n \omega_{1,i} \mathbb{E}[(u_1 - u_i^\perp)^2]$	$\sum_{i < j} \omega_{i,j} \mathbb{E}[(u_i^\perp - u_j^\perp)^2]$	$\sum_{x_k \in \Omega} \omega_k \mathbb{E}_i[(\mu_k - u_i^\perp(x_k))^2]$
NCC	$\mathbb{E}[\tilde{u} \cdot \tilde{v}^\perp]$	$\sum_{i=2}^n \omega_{1,i} \mathbb{E}[\tilde{u}_1 \cdot \tilde{u}_i^\perp]$	$\sum_{i < j} \omega_{i,j} \mathbb{E}[\tilde{u}_i^\perp \cdot \tilde{u}_j^\perp]$	$\sum_{x_k \in \Omega} \omega_k \mathbb{E}[\tilde{u}_1^\perp \cdot \tilde{u}_2^\perp \cdots \tilde{u}_n^\perp]$
CR	$\frac{\text{Var}[\mathbb{E}(u v^\perp)]}{\text{Var}(u)}$	$\sum_{i=2}^n \omega_{1,i} \frac{\text{Var}[\mathbb{E}(u_1 u_i^\perp)]}{\text{Var}(u_1)}$	$\sum_{i < j} \omega_{i,j} \frac{\text{Var}[\mathbb{E}(u_i^\perp u_j^\perp)]}{\text{Var}(u_i^\perp)}$	-
MI	$\text{MI}(u, v^\perp)$	$\sum_{i=2}^n \omega_{1,i} \text{MI}(u_1, u_i^\perp)$	$\sum_{i < j} \omega_{i,j} \text{MI}(u_i^\perp, u_j^\perp)$	$\sum_{x_k \in \Omega} \omega_k \text{H}(P^k)$

of multivariate extensions for popular measures have been proposed, which we discuss together with our own extensions in section 3.

The reason for choosing a simultaneous registration approach is twofold, like the problems occurring during registration. First, the accumulated registration error that was treated in a separated normalization step by the above mentioned registration approach, is now handled intrinsically during the registration. Second, the multivariate similarity measures create more robust cost functions for the optimizer to run on because each image is put into its global context trying to get the maximum out of the depicted structures.

For our mosaicing framework we use two variants of the simultaneous approach that we refer to as *full-simultaneous* and *semi-simultaneous* registration, both using multivariate similarity measures but differing in their optimization strategy. While for the full-simultaneous registration the optimization is performed in the  $n \cdot 6$  dimensional parameter space, the semi-simultaneous registration focuses on the optimization of the 6 pose parameters of one image at a time. During one cycle each image is registered for a limited number of registration steps. Several of these cycles yield a stepwise simultaneous convergence to the best global alignment.

The reason for working with two versions lies in the increased computational complexity of simultaneous methods, which is a logical consequence of the higher dimensional parameter space and multivariate similarity metrics. The semi-simultaneous approach has lower complexity because of the reduced parameter space and the measures, which need only to be evaluated within the grid of the currently optimized image, in contrast to the whole compounding volume for the full-simultaneous one. A complete drift of the scene is avoided by normalizing the transformations so that one of them be the identity.

### 3 Multivariate Similarity Measures

Multivariate similarity measures have not yet been used for the registration of multiple ultrasound images in spite of their already mentioned advantages. We

focus our analysis on four popular measures, whose applications are not limited to ultrasound registration: sum of squared differences (SSD), normalized cross-correlation (NCC), mutual information (MI), and correlation-ratio (CR). A maximum-likelihood estimation (MLE) framework is commonly used to mathematically model the registration process. For the bivariate case the imaging process is described by  $u(x) = f(v(T(x))) + \mu$ , with the images  $u$  and  $v$ , the transformation  $T$ , the stationary white Gaussian noise  $\mu$  and the intensity mapping  $f$ . The negative log-likelihood function is

$$-\log \mathcal{L}(T, \mu, f) = -\log P(u|v, T, \mu, f) = -\log P(\mu = u(x) - f(v(T(x)))) \quad (1)$$

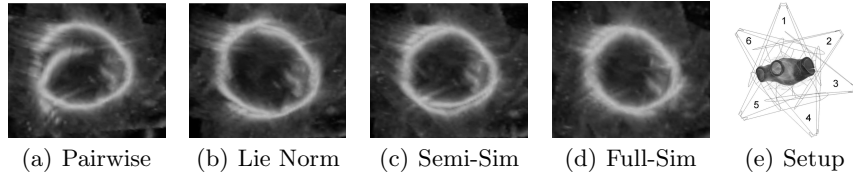
with  $P$  the probability density function (PDF). In the work of Viola [11] and Roche et al. [12] the deduction of the four measures based on this equation is shown by varying the assumptions for the intensity mapping. We are extending this approach to multiple images under the assumption of conditional independent images. The extended MLE denoting the transformed images  $u_i^\downarrow = u_i(T_i(\cdot))$  is

$$\begin{aligned} -\log \mathcal{L}(T, \boldsymbol{\mu}, \mathbf{f}) &= -\log P(u_1^\downarrow | u_2^\downarrow, \dots, u_n^\downarrow, \boldsymbol{\mu}, \mathbf{f}) \\ &= -\log P(\mu_2 = u_1^\downarrow - f_2(u_2^\downarrow), \dots, \mu_n = u_1^\downarrow - f_n(u_n^\downarrow)) \\ &= -\sum_{i=2}^n \log P(\mu_i = u_1^\downarrow - f_i(u_i^\downarrow)) \end{aligned}$$

with intensity mappings  $\mathbf{f} = (f_2, \dots, f_n)$  and Gaussian noises  $\boldsymbol{\mu} = (\mu_2, \dots, \mu_n)$ . Each summand corresponds to the bivariate formula in equation 1 and the deduction of the four similarity measures can therefore be done analogously as in [11,12]. This shows that we directly obtain multivariate extensions of that form by summing up the bivariate measures. In this type of extension we pick one reference image, in the formulae above  $u_1$ , which suits very well for the semi-simultaneous registration approach. Setting up a similarity matrix  $M$  with the entries  $M_{i,j} = SM(u_i, u_j)$ , this corresponds to summing up its first row. An adaptation of this approach to the full-simultaneous registration is obtained by summing up the whole similarity matrix, which can often be limited to the upper triangular part because of the symmetry of the measures. Additionally, the pairs are weighted by an overlap dependent factor  $\omega_{i,j}$  emphasizing pairs with high overlap. The final criteria are shown in figure 1.

A second type of extension, the *voxel-wise* one, that we are using is based on the idea of congealing [10] and puts the focus on a voxel location at a time. In the MLE framework, it is integrated by estimating PDFs for each voxel under the assumption of independent but not identical distributed coordinate samples

$$\begin{aligned} -\log \mathcal{L}(T) &= -\log P(u_1^\downarrow, u_2^\downarrow, \dots, u_n^\downarrow) \\ &= -\frac{1}{|\Omega|} \log \prod_{x_k \in \Omega} P^k(u_1^\downarrow(x_k), \dots, u_n^\downarrow(x_k)) \\ &\approx -\frac{1}{|\Omega|} \sum_{x_k \in \Omega} \log \prod_{i=1}^n P^k(u_i^\downarrow(x_k)) \end{aligned} \quad (2)$$



**Fig. 1.** Error sum up by pairwise reg. Simultaneous reg intrinsically deals with it

with the grid  $\Omega$ . By further assuming a Gaussian distribution of values at each location with mean  $\mu_k$  and variance  $\sigma_k^2$  the negative log-likelihood function is

$$\begin{aligned}
 -\log \mathcal{L}(T) &= -\frac{1}{|\Omega|} \sum_{x_k \in \Omega} \sum_{i=1}^n \log \left( \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(u_i^\downarrow(x_k) - \mu_k)^2}{\sigma_k^2}} \right) \\
 &\approx \frac{1}{|\Omega|} \sum_{x_k \in \Omega} \frac{1}{\sigma_k^2} \sum_{i=1}^n (u_i^\downarrow(x_k) - \mu_k)^2.
 \end{aligned} \tag{3}$$

We consider this criterion as a voxel-wise extension of SSD because similar assumptions as for its pairwise deduction in [11] were used. We also use a voxel-wise criterion for NCC that, in our opinion, captures the basic idea of it by multiplying the values at each voxel location of the normalized images  $\tilde{u}_i$ . We added the congealing criterion [10] as an extension of MI to the figure 1, because they are both based on the estimation of entropy  $H$ , although they have different properties. For all, we added the weighting factor  $\omega_k$  emphasizing locations with a higher number of overlapping images. The usual extensions based on higher-dimensional PDFs are not applicable to mosaicing because they are not flexible enough to allow for varying numbers of overlapping images.

## 4 Results

We tested the mosaicing strategies and multivariate similarity measures on two data sets. First, 3D images of a heart clay model in the water bath were acquired from six different angles. The imaging setup is shown in fig. 1(e). We use a cutting plane through the reconstruction volume to visualize the registration error. When using pairwise registration the summed up error leads to a large displacement between the first and sixth volume, fig. 1(a). The pairwise registration with a successive Lie normalization corrects this error, but the alignment is not perfect, fig. 1(b). The semi-simultaneous registration provides good results, fig. 1(c), but superior results are obtained with the full-simultaneous registration, fig. 1(d).

The second data set consists of four sequentially taken acquisitions from a baby phantom, see fig. 3(d) for the compounded result. For this data set we plot all the proposed similarity measures by moving the second image along the cranio-caudal axis, see fig. 2. One clearly sees the high overlap dependence of the bivariate measures, being a source for misregistrations (total overlap at -37mm

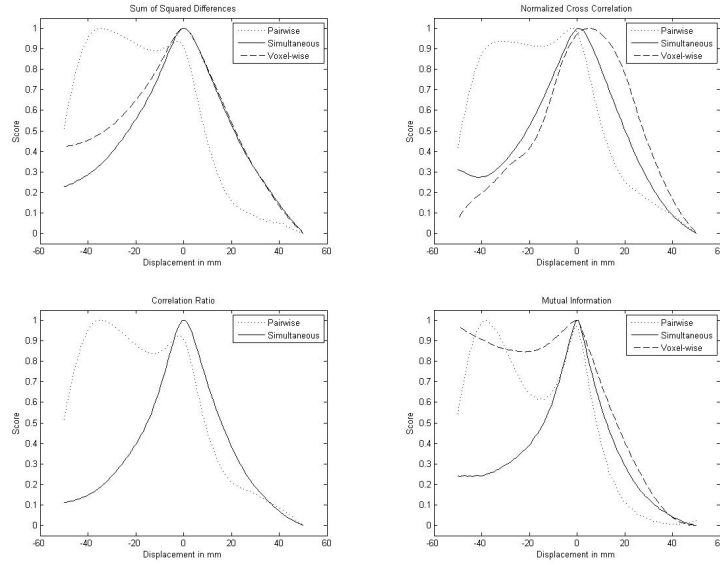


Fig. 2. Similarity plots of the measures in figure 1 on the baby phantom

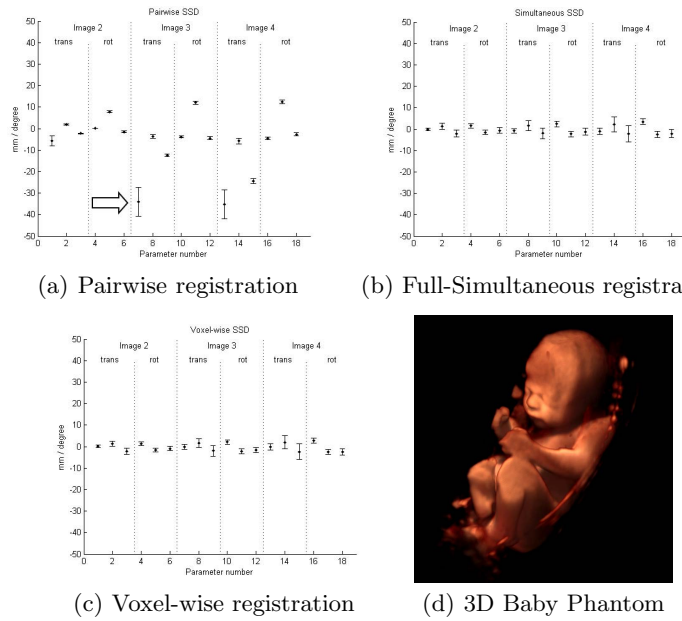


Fig. 3. Mean and standard deviation of pose parameters after 100 registrations

displacement). The multivariate measures provide a smooth cost function with a clear maximum at the correct position 0. We also ran a registration study, with an initial random deviation of maximal  $\pm 20$  mm in translation and  $\pm 20^\circ$  in rotation from the correct pose. The mean and standard deviation of each pose parameter of the three moving images after the registration are shown in fig. 3. The pairwise registration leads to a misalignment because of the total overlap of the images 2 and 3, indicated in fig. 3(a) by a mean of -34.9 mm of parameter 7. The distribution of the mean values around 0 after the simultaneous registration, together with low variances, indicates good registration results, see fig. 3(b) and 3(c). For a demonstration of the performance of the simultaneous registration on the baby phantom, see the video material.

## 5 Conclusion

We have described three registration strategies for ultrasound mosaicing which are put into relationship to the standard pairwise sequential one. Our experiments clearly show that these advanced strategies are necessary to address the problems that can occur during ultrasound mosaicing. The best registration result was obtained with the full-simultaneous approach but this comes with a high computational cost. Moreover, we set up a MLE framework to deduce extensions of popular similarity measures. This allows us to derive a new class of multivariate measures by summing up the pairwise ones and also to deduce a voxel-wise extension of SSD. Our results show the good performance of these measures in contrast to the bivariate ones. Seamless extension to affine and deformable transformation models is possible especially using the proposed voxel-wise SSD.

## References

1. Brown, M., Szeliski, R., Winder, S.: Multi-image matching using multi-scale oriented patches. *Computer Vision and Pattern Recognition (CVPR)* 1 (2005)
2. Kim, S.H., Choi, B.I., Kim, K.W., Lee, K.H., Han, J.K.: Extended FOV Sonography: Advantages in Abdominal Appl. *J. Ultrasound Med.* 22(4), 385–394 (2003)
3. Peetrons, P.: Ultrasound of muscles. *European Radiology* 12(1), 35–43 (2002)
4. Dietrich, C., Ignee, A., Gebel, M., Braden, B., Schuessler, G.: Imaging of the abdomen. *Z. Gastroenterol.* 40, 965–970 (2002)
5. Henrich, W., Schmider, A., Kjos, S., Tutschek, B., Dudenhausen, J.W.: Advantages of and applications for extended field-of-view ultrasound in obstetrics. *Archives of Gynecology and Obstetrics* V268, 121–127 (2003)
6. Gee, A.H., Treece, G.M., Prager, R.W., Cash, C.J.C., Berman, L.H.: Rapid registration for wide field-of-view freehand 3d ultrasound. *IEEE Trans. Med. Imaging* 22(11), 1344–1357 (2003)
7. Poon, T., Rohling, R.: Three-dimensional extended field-of-view ultrasound. *Ultrasound in Medicine and Biology* 32(3), 357–369 (2005)
8. Pennec, X.: *Statistical Computing on Manifolds for Computational Anatomy*. Habilitation à diriger des recherches, Université Nice Sophia-Antipolis (December 2006)



9. Vercauteren, T., Perchant, A., Malandain, G., Pennec, X., Ayache, N.: Robust mosaicing with correction of motion distortions and tissue deformation for in vivo fibered microscopy. *Medical Image Analysis* 10(5), 673–692 (2006)
10. Zoellei, L., Learned-Miller, E., Grimson, E., Wells III, W.M.: Efficient population registration of 3d data. In: *ICCV* (2005)
11. Viola, P.A.: *Alignment by Maximization of Mutual Information*. Ph.d. thesis, Massachusetts Institute of Technology (1995)
12. Roche, A., Malandain, G., Ayache, N.: Unifying maximum likelihood approaches in medical image registration. *Int J. of Imaging Syst. and Techn.* 11(1), 71–80 (2000)