A Non-Linear Image Registration Scheme for Real-Time Liver Ultrasound Tracking using Normalized Gradient Fields

Lars König, Till Kipshagen and Jan Rühaak

Fraunhofer MEVIS Project Group Image Registration, Lübeck, Germany
lars.koenig@mevis.fraunhofer.de

Abstract. We propose a novel scheme for annotation tracking in long liver ultrasound sequences. It is based on a variational non-linear image registration method using Normalized Gradient Fields, extended by a moving window strategy based on registrations to the provided annotation on the first frame. By this we achieve robustness against error accumulation, while handling large deformations at the same time. The method is evaluated on 21 datasets with up to five annotations as contribution to the MICCAI CLUST14 challenge. We achieved a mean tracking error of 1.31 mm with a standard deviation of 1.63 mm, while running at close to real-time speed, exceeding acquisition rate in ten cases with up to 44 frames per second on standard hardware.

Keywords: tracking, non-linear image registration, normalized gradient fields, liver ultrasound, real-time, CLUST14

1 Introduction

Ultrasound imaging provides unbeaten acquisition speed while having low requirements in component setup. This makes ultrasound a preferable choice where real-time information about patient condition is needed, e.g. for fusion of intra-operative ultrasound images to pre-operative CT images [9] or motion compensation in image guided radiation therapy [5].

To enable fusion of real-time image sequences to planning data, often tracking of relevant features in ultrasound images is needed. Especially in long time series, due to noise and breathing motion, this can be a challenging task [2]. Many different approaches to ultrasound tracking exist, ranging from optical flow methods and speckle tracking up to different forms of image registration [1]. In image registration, especially deformable methods are of interest, as they provide deformation models that are able to represent non-linear deformations in soft tissue. As ultrasound images are typically acquired at high frame-rates, common non-linear image registration schemes are not capable of achieving real-time performance. However, due to recent developments of highly efficient computation schemes [6], even sophisticated variational methods have become an attractive choice for real-time tracking.

In this paper, we present a new tracking scheme based on a fast non-linear image registration algorithm that allows real-time ultrasound tracking. The algorithm does not rely on image segmentations, makes no assumptions about the expected motion and does not require a training phase. By computing registrations on moving image windows, which are related to the given annotation of the time-series, we achieve robustness against error accumulation, while handling large deformations at the same time. We evaluated this new scheme participating in the MICCAI CLUST14 liver ultrasound tracking challenge.

2 Method

The proposed tracking scheme is based on a variational image registration approach [7]. It is embedded in a specialized framework allowing for processing of image sequences and efficient compensation of breathing motion. In Section 2.1, we first describe the non-linear registration algorithm, that is then used as a basis for the tracking algorithm described in Section 2.2.

2.1 Image Registration

Let $\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}$ denote the fixed reference image and $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}$ the moving template image with compact support in domain $\Omega \subseteq \mathbb{R}^2$. The goal of image registration is to find a *transformation* $y : \Omega \to \mathbb{R}^2$ that encodes the spatial correspondence between the two images $\mathcal{R}$ and $\mathcal{T}$. In variational approaches, this is modeled by an objective function $\mathcal{J}$ called *joint energy function* which typically consists of a distance term $\mathcal{D}$ describing image similarity and a regularizer $\mathcal{S}$ which penalizes implausible deformations [7]. Image registration then translates to minimizing the functional

$$\mathcal{J}(y) = \mathcal{D}(\mathcal{R}, \mathcal{T}(y)) + \alpha \mathcal{S}(y). \tag{1}$$

Here, the regularization parameter α enables a balance between data fit and deformation regularity.

As image edges are prominent features in ultrasound images, we choose the edge-based Normalized Gradient Fields (NGF) distance measure [4]

$$\mathcal{D}(\mathcal{R}, \mathcal{T}(y)) := \int_\Omega 1 - \left(\frac{\langle \nabla \mathcal{T}(y(x)), \nabla \mathcal{R}(x) \rangle_\eta}{\|\nabla \mathcal{T}(y(x))\|_\eta \, \|\nabla \mathcal{R}(x)\|_\eta} \right)^2 \, \mathrm{d}x \tag{2}$$

with

$$\langle f, g \rangle_\eta := \sum_{j=1}^{2} f_j g_j + \eta^2 \qquad \text{and} \qquad \|f\|_\eta := \sqrt{\langle f, f \rangle_\eta}.$$

Note that $\|\cdot\|_\eta$ does not define a norm in the mathematical sense as $\|0\|_\eta \neq 0$ for $\eta \neq 0$. Roughly speaking, NGF measures the angle between reference and template image intensity gradients at each point and aims for alignment of these image gradients. The *edge parameter* η is introduced to suppress the influence

of small edges e.g. caused by noise. The choice of the parameters η and α is discussed in Section 2.4.

As we generally expect smooth deformations between the ultrasound time frames, we select the *curvature regularizer* as proposed in [3] which is based on second order derivatives. With the decomposition $y(x) = x + u(x)$, the curvature regularizer is given by

$$\mathcal{S}(y) := \frac{1}{2} \int_{\Omega} \|\Delta u_x\|^2 + \|\Delta u_y\|^2 \, \mathrm{d}x,$$

where u_x, u_y denote the components of the displacement in x- and y-direction, respectively. The curvature regularizer penalizes the Laplacian of the displacement components, thus generating very smooth deformations.

The minimization of (1) is performed following the discretize-then-optimize paradigm [7]. In this ansatz, all components (distance measure, regularizer and transformation) are first carefully discretized, yielding a continuous, yet finite dimensional optimization problem. This enables the usage of standard algorithms from numerical optimization [8]. We employ the quasi-Newton L-BFGS optimization scheme to minimize the objective function $\mathcal{J}$ for its speed and memory efficiency. The implementation is based on the two-loop recursion formulation as presented in [8]. The occurring linear equation system in each iteration step of the Newton scheme is solved using a conjugate gradient method. Furthermore, to avoid local minima, the iteration scheme is embedded in a multi-level approach [7], where the optimization problem is solved consecutively on coarse to fine image resolution levels.

The evaluation of the objective function $\mathcal{J}$ together with its derivative is performed using the algorithm presented in [6]. This approach includes an explicit calculation rule for the derivative of $\mathcal{J}$, which does not require any storage of Jacobian matrices and allows for a full pixelwise parallel computation.

2.2 Tracking Algorithm

The non-linear registration described in Section 2.1 is embedded in a larger framework to enable efficient usage in annotation tracking on ultrasound sequences. By calculating registrations of moving windows on each image of the time series, we enable the tracking of an annotation $a_1 \in \mathbb{R}^2$, given on the first frame, over time. The proposed tracking scheme is illustrated in Figure 1.

Let $I_k \in \mathbb{R}^{M \times N},\ k = 1, \dots, T$ denote the k-th frame of the ultrasound sequence of length T. We then define $W_n(I_k) : \mathbb{R}^{M \times N} \to \mathbb{R}^{w_1 \times w_2},\ n, k = 1, \dots, T$ as a window of I_k with extent $w_1, w_2 \in \mathbb{N}, w_1 \leq M, w_2 \leq N$ and center position $c_n \in \mathbb{R}^2$. Starting with the original annotation $a_1 \in \mathbb{R}^2$ on I_1, window $W_1(I_1)$ with center $c_1 = a_1$ is chosen. The extent w_1, w_2 is kept constant throughout the algorithm and will be discussed in Section 2.4. Initially, a registration between $W_1(I_1)$ as reference and $W_1(I_2)$ as template image is performed where the continuous image representation, as defined in Section 2.1, is obtained by bilinear interpolation. Using the registration result y_1, the initial annotation a_1 is then

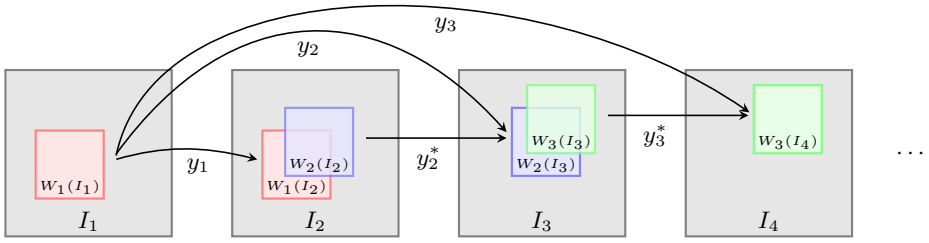


Fig. 1. Proposed tracking scheme. The registration between window $W_{n-1}(I_n)$ on current frame and window $W_1(I_1)$ on first frame is denoted by y_n. If this first registration fails, we compute y_n^* by registering corresponding windows on current and previous frame.

transformed as $a_2 = y_1(a_1)$, yielding the moved annotation a_2 tracked to frame I_2. After the transformation step, a new window $W_2(I_3)$ is computed, now with $c_2 = a_2$. Window $W_2(I_3)$ is then registered to $W_1(I_1)$, yielding transformation y_2 that is used to compute moved annotation $a_3 = y_2(a_1)$. This process is then repeated for all frames. A pseudocode of this scheme is shown in Algorithm 1.

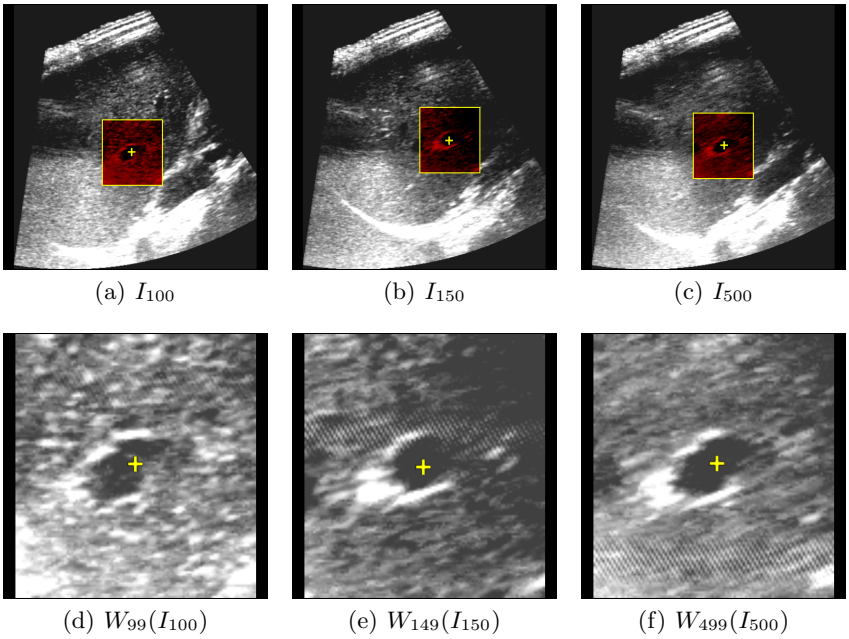
Algorithm 1 Pseudocode for the tracking algorithm of a single landmark

load I_1, a_1 ▷ Load first image and original annotation
for n in $[2, T]$ **do** ▷ Loop over all frames
  load I_n
  $W_{n-1} \leftarrow$ window around a_{n-1}
  $y_{n-1} \leftarrow$ registration$(W_1(I_1), W_{n-1}(I_n))$ ▷ Register to first window
  if registration was successful **then**
    $a_n \leftarrow y_{n-1}(a_1)$ ▷ Compute new annotation
  else ▷ change of strategy
    $y_{n-1}^* \leftarrow$ registration $(W_{n-1}(I_{n-1}), W_{n-1}(I_n))$ ▷ Register to prev. window
    $a_n \leftarrow y_{n-1}^*(a_{n-1})$ ▷ Compute new annotation
  end if
end for

Additionally, a safeguard procedure is included in the algorithm. In each registration step, the success of the current registration is checked by determining the final value of the distance measure (2). If this value is above a certain threshold θ, described in detail in Section 2.4, the registration is considered having failed. In this case, the registration paradigm is switched. Instead of registering $W_{n-1}(I_n)$ onto $W_1(I_1)$, we now use window $W_{n-1}(I_{n-1})$ of the previous frame as reference image and register $W_{n-1}(I_n)$ onto $W_{n-1}(I_{n-1})$ instead, yielding y_{n-1}^*. This step is based on the interpretation of a_{n-1} as the last successfully tracked annotation. It enables landmark tracking in cases where large local differences compared to the first window exist, but consecutive frames are still relatively similar. If this procedure has to be repeated multiple times, error accumulation may occur, as the tracking relies on tracked annotations from previous steps,

that may themselves contain errors. However, as soon as the differences to the first frame decrease, the original scheme takes over and possibly accumulated errors are discarded by deforming the ground truth annotation a_1 again. This mechanism enables a successful tracking also in situations where the difference to the initial frame is temporarily large.

The proposed algorithm has several benefits. It does not require a training phase, avoids error accumulation, makes no assumptions about motion periodicity and does not rely on image segmentations. By choosing $W_1(I_1)$ as a fixed reference window, we always refer to the given annotation a_1 throughout the whole tracking process. While possibly being at completely different locations, the image contents inside the windows only exhibit small movements as larger movements of the structure have already been compensated for by shifting the window according to the deformation obtained in the prior iteration, see Figure 2. This procedure generates an excellent starting point for the underlying non-linear registration scheme.



(a) I_{100} (b) I_{150} (c) I_{500}

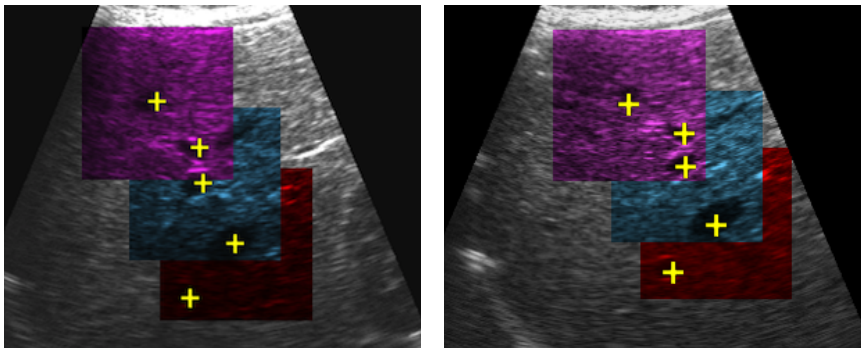
(d) $W_{99}(I_{100})$ (e) $W_{149}(I_{150})$ (f) $W_{499}(I_{500})$

Fig. 2. Selected frames of dataset ETH-07. While the colored window (tracked annotation at its center) experiences globally large movements (top row), the contents of the respective windows remain similar (bottom row).

2.3 Annotation coupling

The tracking scheme is in principle designed for tracking single landmarks. If multiple landmarks are to be tracked, the choice of window and registration

are performed separately for each landmark, yielding independent deformations and windows. In the described scheme, however, dense deformation fields are computed at all times on the area of the chosen windows. We exploit the deformation field for annotations lying close together by coupling a secondary annotation with the window of a primary annotation instead of tracking these annotations independently. The primary annotation is still tracked identically to the scheme described in Section 2.2, while the secondary annotation is deformed using the computed deformation field within the chosen window. This increases tracking performance by enabling tracking multiple landmarks with the same registration.



(a) I_1 (b) I_{1750}

Fig. 3. Tracking of five annotations in dataset MED-09, first frame (a) (part of full image shown) and intermediate result (b). The individual windows are shown as colored rectangles. The second annotation is coupled with the third and the fifth with the fourth (numbering from top to bottom).

2.4 Parametrization

Our tracking algorithm provides several parameters to adapt to varying image characteristics. While these parameters were manually determined, it generally suffices to select them once per device and not per dataset.

For the CLUST14 challenge, we focused on parameters fitting all datasets from the same scanner and probe. While better results can be obtained by choosing specialized parameters for each dataset, this would contradict the purpose of annotation tracking in a real-time setting, where future frames are unknown.

For all processed datasets, parameters for optimization, deformation resolution and multi-level scheme were kept constant. The window size was determined as one fourth of the image size in each dimension. For all registrations, we chose two levels in the multi-level scheme, with a downsampled version of half the window resolution on the finest level. The deformation resolution was determined as 17×17, thus reducing the computational costs and additionally acting as a regularizer, see [11] for details.

Important parameters that allow adapting to different noise and device characteristics are the regularization factor α and the NGF noise parameter η, see Section 2.1. The threshold θ for the safeguard strategy described in Section 2.2 is adaptively chosen as $\theta = \tau \cdot \frac{a}{4096}$, depending on the window area a. Using all given datasets, the parameters α, η and τ were manually calibrated per device and probe. For the ETH datasets, we used $\alpha = 0.1, \eta = 10$ and $\tau = 1490$, except for ETH-1, where because of the different resolution, we set $\eta = 2.5$. For the MED datasets $\alpha = 0.5, \eta = 5$ and $\tau = 1400$ were used, except for MED-15, where $\tau = 850$ was used to compensate for a single exceptional artifact. The annotations were coupled as follows. ETH-03: $3 \rightarrow 2$, ETH-10: $4 \rightarrow 3$, MED-03: $2 \rightarrow 4$, MED-05: $3 \rightarrow 2$, MED-09: $3 \rightarrow 2$, $5 \rightarrow 1$, MED-10: $2 \rightarrow 4$.

3 Results and Discussion

The non-linear image registration was implemented in C++, while the tracking framework was scripted in Python and executed in MeVisLab [10]. Our method was evaluated on all 2D annotation tracking test datasets provided by the organizers of the CLUST2014 challenge [2]. These datasets contained image sequences from 264×313 (ETH-01) to 524×591 (MED-13 – MED-15) pixels in resolution, with number of frames ranging from 2427 (MED-10) to 14516 (ETH-01) frames. Every first frame was provided with up to five annotations.

On the ETH datasets, our method achieved a mean tracking error (MTE) of 0.89 mm with a standard deviation (σ) of 1.84 mm. For the MED datasets, we achieved a MTE of 1.73 mm with $\sigma = 1.25$ mm, resulting in overall values of MTE=1.31 mm and $\sigma = 1.63$ mm. Full results are given in Table 1. It has to be noted, that the overall tracking results were negatively influenced by a tracking failure in dataset ETH-07 that affected only the last ≈ 250 of total 4588 frames.

The algorithm achieved close to real-time performance in all cases, exceeding acquisition rate in ten cases, computed on a three year old Intel i7-2600 PC with 3.40GHz running Ubuntu Linux 12.04, see Table 1 for computation speed. Thus real-time performance is easily within reach when using recent hardware.

Currently the choice of suitable parameters requires significant manual fine tuning since neither spacial image resolution information nor device characteristics such as center frequency were taken into account. Including this information is subject to future research and may enable automatic parameter calibaration.

We developed a fast and accurate ultrasound tracking algorithm capable of achieving real-time performance while relying solely on image information, without any further knowledge like image segmentation or feature recognition required. Furthermore, no prior training phase is needed and no assumptions about the type of movement are made.

The proposed scheme can easily be extended to 3D tracking. Since dense deformation fields are computed in every step, the algorithm can also be used directly for tracking of segmentations.

Acknowledgment: This work was supported by the Fraunhofer Internal Programs under Grant No. MAVO 823 287.

Dataset	IAR	FPS	MTE_1	σ_1	MTE_2	σ_2	MTE_3	σ_3	MTE_4	σ_4	MTE_5	σ_5
ETH 01	25	43.5	0.87	0.98								
02	16	31.7	0.97	0.46								
03	17	14.6	0.37	0.21	0.64	0.36	0.47	0.24				
04	15	33.8	0.86	1.20								
06	17	14.8	0.62	0.60	1.13	0.85						
07	14	33.3	2.87	7.38								
08	17	15.0	0.59	0.32	0.68	0.45						
09	16	18.3	0.69	0.34	1.01	0.54						
10	15	11.1	1.07	0.74	0.80	0.66	0.93	1.23	0.94	1.28		
MED 01	20	15.5	1.09	0.61	0.94	0.42	1.04	0.61				
02	20	14.3	1.03	0.57	1.30	0.88	1.94	0.46				
03	20	14.9	1.23	0.62	2.72	1.84	1.20	0.75	0.91	0.49		
05	20	22.6	2.02	0.95	2.14	0.86	2.56	1.08				
06	20	14.3	1.71	0.91	1.22	0.55	1.37	0.59				
07	20	13.9	3.39	2.22	1.49	0.90	2.17	1.28				
08	20	21.9	2.03	1.06	2.52	1.52						
09	20	14.2	2.31	1.72	1.23	0.64	1.21	0.81	2.42	0.91	2.71	2.29
10	20	13.0	2.25	1.01	1.67	0.92	2.12	0.96	1.34	0.98		
13	11	11.1	1.08	0.69	2.14	1.35	1.13	0.62				
14	11	11.6	1.72	0.93	1.88	1.02	2.64	1.64				
15	11	29.9	1.32	1.28								

Table 1. Mean tracking error (MTE) and standard deviation (σ) in the CLUST14 2D annotation tracking datasets (empty cells correspond to fewer annotations). Processing speed is given as frames per second (FPS), image acquisition rate in column IAR.

Bibliography

[1] De Luca, V.: Liver motion tracking in ultrasound sequences for tumor therapy. Ph.D. thesis, ETH Zurich (2013)

[2] De Luca, V., Tschannen, M., Székely, G., Tanner, C.: A learning-based approach for fast and robust vessel tracking in long ultrasound sequences. In: MICCAI 2013, pp. 518–525. Springer (2013)

[3] Fischer, B., Modersitzki, J.: Curvature based image registration. Journal of Mathematical Imaging and Vision 18(1), 81–85 (2003)

[4] Haber, E., Modersitzki, J.: Intensity gradient based registration and fusion of multi-modal images. Methods of Information in Medicine 46, 292–9 (2007)

[5] Keall, P.J., et al.: The management of respiratory motion in radiation oncology report of AAPM Task Group 76. Medical Physics 33(10), 3874–3900 (2006)

[6] König, L., Rühaak, J.: A Fast and Accurate Parallel Algorithm for Non-Linear Image Registration using Normalized Gradient Fields. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro (2014)

[7] Modersitzki, J.: FAIR: Flexible Algorithms for Image Registration. SIAM (2009)

[8] Nocedal, J., Wright, S.: Numerical optimization. Springer (1999)

[9] Olesch, J., et al.: Fast intra-operative nonlinear registration of 3D-CT to tracked, selected 2D-ultrasound slices. In: SPIE Medical Imaging (2011)

[10] Ritter, F., Boskamp, T., Homeyer, A., Laue, H., Schwier, M., Link, F., Peitgen, H.O.: Medical image analysis. Pulse, IEEE 2(6), 60–70 (2011)

[11] Rühaak, J., Heldmann, S., Kipshagen, T., Fischer, B.: Highly accurate fast lung CT registration. In: SPIE Medical Imaging (2013)