

A Scale-Space Approach to Landmark Constrained Image Registration

Eldad Haber¹, Stefan Heldmann², and Jan Modersitzki³

¹ Dept. of Math. and Computer Science, Emory University, Atlanta, USA,
`haber@mathcs.emory.edu`

² Inst. of Mathematics, University of Lübeck, Lübeck, Germany
`heldmann@math.uni-luebeck.de`

³ Dept. of Computing and Software, McMaster University, Hamilton, Canada
`modersit@cas.mcmaster.ca`

Abstract. Adding external knowledge improves the results for ill-posed problems. In this paper we present a new multi-level optimization framework for image registration when adding landmark constraints on the transformation. Previous approaches are based on a fixed discretization and lack of allowing for continuous landmark positions that are not on grid points. Our novel approach overcomes these problems such that we can apply multi-level methods which have been proven being crucial to avoid local minima in the course of optimization. Furthermore, for our numerical method we are able to use constraint elimination such that we trace back the landmark constrained problem to a unconstrained optimization leading to an efficient algorithm.

1 Introduction

Image registration is a challenging problem in digital imaging. Roughly speaking, the problem can be described as follows. Given a reference image \mathcal{R} and a template image \mathcal{T} , find a *reasonable* spatial transformation y such that the transformed image $\mathcal{T}[y]$ is *similar* to the reference. Image registration is required whenever images resulting from different times, devices, and/or perspectives need to be compared or integrated. Alone in the area of medical applications, registration is used in radiation therapy, surgery planing, treatment evaluation, motion correction and estimation and many more, see, e.g. [1–6] and references therein. See also [7–9] for related work.

However, although the registration problem is easily stated it is hard to be solved. A key difficulty is the ill-posedness of the problem: For a particular point x , scalar intensity values $\mathcal{R}(x)$ and $\mathcal{T}(x)$ are given but a transformation vector $y(x)$ vector is to be computed. A common approach is to phrase image registration as an optimization problem involving a distance measure \mathcal{D} reflecting *similarity* of images and a regularization term \mathcal{S} reflecting *reasonability* of the transformation. Though appropriate regularization results in a well-posed problem in the sense of Hadamard [10] (see, e.g. [11–13]), it is sometimes difficult or even impossible to find an application conform regularization.

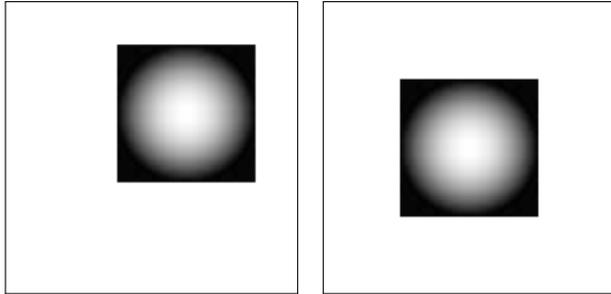


Fig. 1. Reference (left) and template (right) images.

A simple example is shown in Fig. 1, where the reference and template image cover the full intensity range and share some obvious symmetries. Considering only rigid transformations, there are four different solutions for any reasonable distance measure. Regularization can be used to privilege one of these (for example by penalizing rotations). However, any regularization is somehow artificial and may favor a meaningless solution.

One way to obtain better results and to guide the model towards a more realistic solution is by using landmarks. In the above example, just adding the information that the top-left corner of the square in the reference image corresponds to the bottom-right corner of the square in the template image eliminates three of the above solutions. Adding landmark information to image registration is far from being new, see e.g. [14–18, 4] and references therein.

Although landmarks have been used extensively in the past, the effective numerical implementation of image registration with landmark is unsatisfactory. For example, no landmark registration scheme known to us allow for the incorporation of scale space or multi-level techniques which are frequently used to avoid local minima. Typically, landmark constraints are described in a discrete sense, where the i th pixel in a fixed discretization is constrained. This causes troubles if the discretization is variable, that is, if discretization on different scales is used.

The goal of this paper is to develop a multilevel technique for the incorporation of landmarks in a registration process. We stress that ignoring issues such as local minima, different algorithms than the one proposed here (for example [17]) should give similar results. Thus, the focus of this work is on numerical implementation of multilevel algorithms with landmark constraints. Starting with a variational formulation of the landmark constrained registration problem, this paper provides a consistent numerical approach. The new approach is based on discretize-then-optimize approach and takes advantage of a multi-level discretization. The new approach automatically resolves the problem resulting from a fixed number of constraints versus a varying number of unknowns and related inconsistency of the constraints. A numerical stable and computational feasible basis of the constrained manifold is derived. Using a reduced formulation gives

a handle to an elegant algorithm, where indefinite Karush-Kuhn-Tucker systems [19] can be avoided.

This paper is organized as follows. Sect. 2 introduces the basic notation and states the problem in a variational framework. A discretized then optimize approach is used to numerically solve the constrained registration problem. Details are outlined in Sect. 3, where the discretization, the construction of a basis for the constraint manifold, the numerical optimization, and a multi-level strategy are described. Sect. 4 presents some numerical results. Conclusions are given in Sect. 5.

2 Variational Formulation

In this section we formulate the constrained registration problem. Let $d \in \mathbb{N}$ denote the spatial dimension (typically $d = 2, 3$) and $\Omega \subset \mathbb{R}^d$ the region of interest and let $\mathcal{T}, \mathcal{R} \in L_2(\mathbb{R}^d, \mathbb{R})$ denote the template and reference image, respectively. The objective is to find a transformation $y : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that the transformed image $\mathcal{T}[y]$ is similar to \mathcal{R} and the transformation y is regular, where similarity and regularity are measured by \mathcal{D} and \mathcal{S} , respectively. More precisely,

$$\begin{aligned} \mathcal{T}[y](x) &:= \mathcal{T}(y(x)) \quad \text{for all } x \in \Omega, \\ \mathcal{D}[\mathcal{T}, \mathcal{R}] &:= \frac{1}{2} \int_{\Omega} (\mathcal{T} - \mathcal{R})^2 dx, \\ \mathcal{S}[y] &:= \frac{\alpha}{2} \int_{\Omega} |\mathcal{B}y|^2 dx, \quad \mathcal{B} := I_d \otimes \Delta. \end{aligned}$$

Here, for ease of presentation, it is assumed that similarity is quantified by the energy in the difference image. However, other distance measure like mutual information [20, 21] or normalized gradient fields [22, 23] can be handled similarly. Regularity is measured using the curvature regularizer [24, 25] where the partial differential operator \mathcal{B} is the vector valued Laplacian, $|\cdot|$ denotes the Euclidian norm in \mathbb{R}^n , and α is a regularization parameter. Note that the order of the regularizer has to be sufficiently high to cover the landmark constraints [26, 4].

It is assumed that a number L of landmarks $r_1, \dots, r_L \in \mathbb{R}^d$ in the reference and corresponding landmarks $t_1, \dots, t_L \in \mathbb{R}^d$ in the template image are given. The automatic detection of landmarks is beyond the scope of this paper; see [16] for an overview. The point evaluation functional is denoted by δ_x . With

$$(I_d \otimes \delta_{r_\ell})[y] = (y^1(r_\ell), \dots, y^d(r_\ell)) = y(r_\ell) \in \mathbb{R}^d$$

the landmark constraints can be phrased as

$$\mathcal{C}[y] = t := (t_1, \dots, t_L)^\top \in \mathbb{R}^{L,d},$$

where $\mathcal{C}[y] = ((I \otimes \delta_{r_1})[y], \dots, (I \otimes \delta_{r_L})[y])^\top \in \mathbb{R}^{L,d}$, and the landmark constrained registration problem reads:

$$\text{minimize } \mathcal{J}[y] = \mathcal{D}[\mathcal{T}[y], \mathcal{R}] + \mathcal{S}[y - y_{\text{ref}}] \quad \text{subject to } \mathcal{C}[y] = t, \quad (1)$$

where y_{ref} allows for a bias towards a particular solution.

The above problem is strongly related to plain landmark based registration, where $\mathcal{D} = 0$ and $\mathcal{S} = \mathcal{S}^{\text{TPS}}$ is the bending energy of a thin-plate-spline; see, e.g. [26, 4] for an extended discussion. The solution y_{TPS} is explicitly known and a linear combination of shifts of a radial basis function ρ associate to \mathcal{S} and a polynomial correction. Following [4], the k th component of y_{TPS} reads

$$y_{\text{TPS}}^k(x) = \sum_{\ell=1}^L \theta_{\ell}^k \rho(|x - r_{\ell}|) + (1, x^1, \dots, x^d)(\theta_{L+1}^k, \dots, \theta_{L+d+1}^k)^{\top}, \quad (2)$$

where the coefficients are given by $A\theta^k = (t_1^k, \dots, t_L^k, 0, \dots, 0)^{\top}$ with

$$A = \begin{pmatrix} [\rho(|r^i - r^j|)]_{i,j=1}^L & P^{\top} \\ P & 0 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & \cdots & 1 \\ r_1 & \cdots & r_L \end{pmatrix} \in \mathbb{R}^{d+1, L},$$

and $\rho(t) = \begin{cases} t^2 \log t & (d = 2) \\ t & (d = 3) \end{cases}.$

In our final formulation of the continuous problem, we use this function as a reference for regularization, i.e. $y_{\text{ref}} = y_{\text{TPS}}$, and it is thus convenient, to rephrase the problem in the update $u = y - y_{\text{ref}}$:

$$\text{minimize } \mathcal{J}[u] = \mathcal{D}[\mathcal{T}[y_{\text{ref}} + u], \mathcal{R}] + \mathcal{S}[u] \quad \text{subject to } \mathcal{C}[u] = 0. \quad (3)$$

The role of the plain landmark solution as a reference is manifold. It can be seen as a good starting guess for a later implementation, minimizing the risk of being trapped by a local minimum. Moreover, it injects boundary values to region of interest. In fact, these boundary conditions make y_{TPS} linear for $x \rightarrow \infty$ and thus invertible, which is preferable for most applications. Finally, it yields homogeneous constraints. As it is pointed out later, this is a crucial point for the discretization as now the feasible set is always non-empty.

3 Numerical treatment

A discretize-then-optimize approach is used to compute a numerical solution of (3). The discretization is briefly outlined for dimension $d = 2$, see [27] for a detailed and general description. Note that the discretization is variable during the course of optimization and all quantities introduced in this section depend on the discretization with h . However, in this section a fixed discretization level is assumed and in order to keep the presentation clear, dependencies on h are neglected.

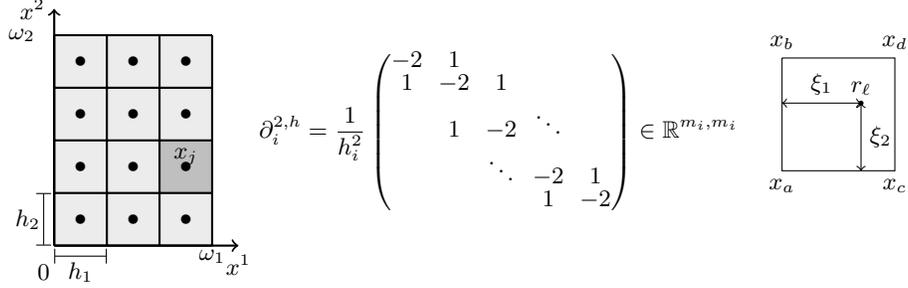


Fig. 2. Discretization of a 2D domain $\Omega = (0, \omega_1) \times (0, \omega_2) \subset \mathbb{R}^2$ (left); discrete 2nd derivative $\partial_i^{2,h}$ (middle); linear interpolation (right).

3.1 Discretization

Fig. 2.a shows the discretization of a domain Ω in $m = (3, 4)$ cells with cell-centers x_j , $j = 1, \dots, n = m_1 m_2$.

Note that all discrete quantities depend on the discretization width h , $h_i = \omega_i / m_i$ and $\hat{h} = h_1 \cdots h_d$. The next equations describes how the discrete quantities are assembled.

$$\begin{aligned} X &= (x_1^1, \dots, x_n^1, \dots, x_n^d) \in \mathbb{R}^{dn}, & R &= (\mathcal{R}(x_1), \dots, \mathcal{R}(x_n)) \in \mathbb{R}^n, \\ U &= (u_1^1, \dots, u_n^1, \dots, u_n^d) \in \mathbb{R}^{dn}, & T(U) &= (\mathcal{T}(u_1), \dots, \mathcal{T}(u_n)) \in \mathbb{R}^n. \end{aligned}$$

The discretization of the curvature operator can be expressed as Kronecker-products [25] of identity matrices $I_q \in \mathbb{R}^{q,q}$ and discrete 2nd derivatives $\partial_i^{2,h}$ (see Fig. 2.b): $\mathcal{B} \approx B = I_d \otimes (I_{m_2} \otimes \partial_1^{2,h} + \partial_1^{2,h} \otimes I_{m_1})$. Finally, the integrals are approximated using a midpoint quadrature rule. Thus

$$\mathcal{J}[u] \approx J(U) = \frac{1}{2} \hat{h} |T(y_{\text{TPS}}(X) + U) - R|^2 + \frac{1}{2} \alpha \hat{h} |BU|^2.$$

The final step is the discretization of the point evaluation functional δ_x . For an arbitrary location r_ℓ , a d -linear interpolation of discrete point evaluation functionals located at the 2^d closest grid points is exploited. For example, let $d = 2$ and let the four neighboring grid points of r_ℓ be denoted by x_a, \dots, x_d ; see Fig. 2.c. Thus,

$$\begin{aligned} \delta_{r_\ell}[u] &\approx \delta_{r_\ell}^h u = C_\ell u(X) \\ &= (1 - \xi_1)(1 - \xi_2)u(x_a) + \xi_1(1 - \xi_2)u(x_b) + (1 - \xi_1)\xi_2 u(x_c) + \xi_1 \xi_2 u(x_d), \end{aligned}$$

and C_ℓ is a sparse row vector with non-zero entries only at positions related to the locations of x_a, \dots, x_d . If for a certain discretization a landmarks r_ℓ would be located precisely on a grid point x_j , then C_ℓ has only one non-zero entry at position j .

Assembling these rows for $\ell = 1, \dots, L$ results a sparse L -by- n matrix C with at most 2^d non-zero entries per row, see Fig. 3.b. The Kronecker-products $I_d \otimes C$ enables a simultaneous treatment of all components of the discretized vector field U . Note that even for a very coarse discretization ($n < L$) there exists a feasible solution fulfilling the constraints: $U = 0$. Thus the feasible set is non-empty.

The discrete formulation of the constrained registration problem thus reads:

$$\begin{aligned} & \text{minimize} && J(U) = \frac{1}{2} \hat{h} |T(y_{\text{ref}}(X) + U) - R|^2 + \frac{1}{2} \hat{h}\alpha |BU|^2 \\ & \text{subject to} && (I_d \otimes C)U = 0, \quad U \in \mathbb{R}^{dn}. \end{aligned} \quad (4)$$

3.2 An efficient basis for the feasible set

The objective is to derive a numerical feasible basis for the nullspace of the operator C . Note the size L -by- n of C can be large (e.g. $n = 128^3$ and $L = 100$) and the rank of this matrix is generally unknown. For a coarse discretization, C has more rows than columns and a fine discretization it has more columns than rows.

The basic idea is to reorder the columns of C , such that the non-zeros columns are placed first. Let Π denote the corresponding $n \times n$ permutation matrix and C^* be a matrix consisting of the non-zeros columns of C , such that

$$C\Pi = (C^* \mid 0).$$

The size of C^* is L -by- p , where $p \leq 2^d L$ since each row of C can have at most 2^d non-zeros entries. The matrix C^* is not only relatively small but also very sparse. Assuming the number of landmarks to be less than 1.000, it is thus possible to compute a singular value decomposition (SVD) of C^* [28], i.e.

$$\begin{aligned} C^* &= W\Sigma V^\top, \quad \text{where } W^\top W = I_L, \quad V^\top V = I_p, \\ &\text{and } \Sigma = \text{diag}(\sigma_1, \dots, \sigma_{\min\{L,p\}}) \in \mathbb{R}^{L,p}, \quad \sigma_1 \geq \dots \geq \sigma_{\min\{L,p\}} \geq 0. \end{aligned}$$

The above SVD enables the computation of the numerical rank of the matrix C^* and hence C . To this end let tol be a user proscribed tolerance (e.g. $\text{tol} = 0$ or $\text{tol} = 10^{-16}$) and let k the largest integer such that $\sigma_k > \text{tol}$. The last $p - k$ columns of V are a basis of the (numerical) nullspace of the matrix C^* and thus the columns of Z form a basis for the nullspace of C , where $V(:, k+1 : p) \in \mathbb{R}^{p,p-k}$

$$Z = \Pi \begin{pmatrix} V(:, k+1 : p) & 0 \\ 0 & I_{n-p} \end{pmatrix} \in \mathbb{R}^{n,n-k},$$

and the final step undoes the permutation.

Important issues are summarized as follows. The matrix C^* is relatively small, such that the SVD becomes numerically feasible. The SVD enables a uniform treatment independent of the rank of C^* and thus handles a coarse discretization ($L > n$) as well as a fine discretization ($L < n$). Note that in the case $L > n$ the solution is the thin plate spline solution since there are 0 degrees of freedom. The columns of Z form a sparse, orthonormal, and numerically stable basis for the set of constraints. For very fine discretizations, the matrix Z is essentially the identity matrix and can be stored efficiently. Any feasible vector is given by $U = (I_d \otimes Z)w$, where $w \in \mathbb{R}^{d(n-k)}$, and there always exists a feasible point $w = 0$.

3.3 Numerical optimization

The final version of the discrete constrained registration problem is given in terms of the reduced basis and reads

$$\text{minimize } J(w) = \frac{1}{2} \hat{h} |T(y_{\text{ref}}(X) + Z_d w) - R|^2 + \frac{1}{2} \hat{h} \alpha |B Z_d w|^2. \quad (5)$$

where $Z_d = I_d \otimes Z$.

In order to find a numerical solution to (5) standard optimization techniques can be applied; see e.g. [19] for an overview. Here, we use a Gauss-Newton type algorithm with an Armijo line search as outlined in [27]. The quasi-Newton system is given by $Z_d^\top H Z_d \delta_w = -\nabla J(w)$ where δ_w is the new search direction and $H = \nabla T^\top \nabla T + \alpha B^\top B$ is an approximation to the Hessian. Note that since the regularization is quadratic the term $B^\top B$ is exactly the Hessian of the regularization part and only the data fitting term is approximated. A generalized Gauss-Newton strategy can be used to handle other distance measures as mentioned before. For a numerical solution of the Newton-systems, a preconditioned conjugate gradient solver is used with symmetric Gauss-Seidel preconditioned; see [29] for details.

3.4 The multilevel strategy

It remains to describe the multi-level framework. To this end, a multi-level representation $\{T_D^\ell, R_D^\ell, m^\ell\}$ of given discrete data is initialized, where for ease of presentation it is assumed that $m_i^\ell = 2^\ell$, $i = 1, \dots, d$, $\ell = \ell_{\min}, \dots, \ell_{\max}$. Note that $h_i^\ell = \omega_i / m_i^\ell$ depends on the level. More precisely,

$$T^{\ell_{\max}} = \text{original data}, \quad T^{\ell-1} = \text{downsample}(\text{conv}(G, T^\ell)),$$

where G is a smoothing kernel (in our numerical experiments we used the block smoother $G = (1, 1, 1)^\top (1, 1, 1) / 9$).

In general, we compute updates to the thin plate spline solution on different grids. Similar to many other multilevel algorithms, the solution on finer grid is initialized by the coarser grid solution.

To be more specific, running from coarse to fine, the continuous representation $\mathcal{T}^\ell, \mathcal{R}^\ell$ for T_D^ℓ, R_D^ℓ are computed (in our numerical experiments, spline interpolation is used). Moreover, the discretized thin-plate spline solution $Y_{\text{ref}}^\ell = y^{\text{TPS}}(X^\ell)$ (cf. (2)) for a cell-centered grid X^ℓ of size m^ℓ and the matrix Z^ℓ (cf. Sect.3.2) is initialized. A numerical solution w_{opt}^ℓ of the discretized registration problem (5) is computed and the current grid solution is given by $Y_{\text{opt}}^\ell = Y_0^\ell + Z_d^\ell w_{\text{opt}}^\ell$. On the coarse grid, the initial guess we choose $w_0^{\ell_{\min}} = 0$ as starting guess such that $Y_0^{\ell_{\min}} = Y_{\text{ref}}^{\ell_{\min}}$. The starting guess w_0^ℓ for a finer grid is chosen as the best least squares approximation of the prolonged coarser grid solution, where $P_{\ell-1}^\ell$ denotes the linear prolongation operator. Since the constraint basis Z^ℓ is orthogonal, the computation simplifies to $w_0^\ell := Z^{\ell \top} P_{\ell-1}^\ell Z^{\ell-1} w_{\text{opt}}^{\ell-1}$.

When designing a multilevel strategy we require set the number of levels. Unfortunately, setting the number of levels is non-trivial. In general, similar

to other problems, one requires that the coarsest level actually represents the problem [30].

4 Results

4.1 Artificial 2D data

We use the hand data shown in Fig. 4. In this example, a synthetic transformation y_{true} has been specified and the reference is a transformed copy of the template image $\mathcal{R} = \mathcal{T}[y_{\text{true}}]$; see Fig. 4. This construction allows a comparison with a ground truth. Here, 47 manually detected landmarks \tilde{t}_j have been chosen in the template image. Using a numerical approximation to the inverse of the transformation, the landmarks in for the reference are defined by $r_j \approx y_{\text{true}}^{-1}(\tilde{t}_j)$ and corresponding landmarks in the template image are defined by $t_j := y_{\text{true}}(r_j)$. Note that since y_{true} is explicitly known, there are no errors in the landmark pairing. The original data is 128-by-128 and the level ranges from $\ell_{\min} = 3$ to $\ell_{\max} = 7$.

Fig. 3.a shows the coarse grid representation of the data. Here, many landmarks can be found in some particular cells. The problem is over-constrained and the 47-by-64 matrix $C^{\ell_{\min}}$ is rank deficient (the rank being 27). The non-zero pattern of this matrix is shown in Fig. 3.b.

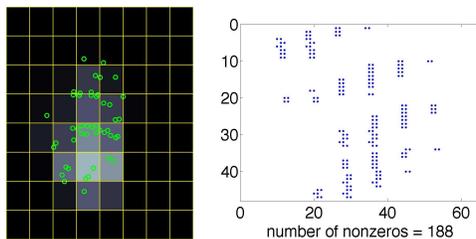


Fig. 3. Coarse grid representation of data with 47 landmarks (circles), $\ell_{\min} = 3$ (left); non-zero pattern of the matrix $C^{\ell_{\min}}$ (right).

Fig. 4 shows the original data (a,b,c) and the results based on the thin-plate-spline solution y_{TPS} (d,g), an unconstrained solution y_{un} (e,h), and the constrained solution y_{con} (f,i). The distance measure and landmark error are given by

$$\begin{aligned} \text{err}(Y) &:= 100 D(Y)/D(X)[\%], \quad D(Y) = |T(Y) - R|^2, \\ \text{LM}(Y) &:= |(I_d \otimes C)Y - t|_{\text{Frobenius}}. \end{aligned}$$

All three registration approaches (TPS, unconstrained, constrained) perform well for this example. The TPS approach gives perfect results for the landmarks

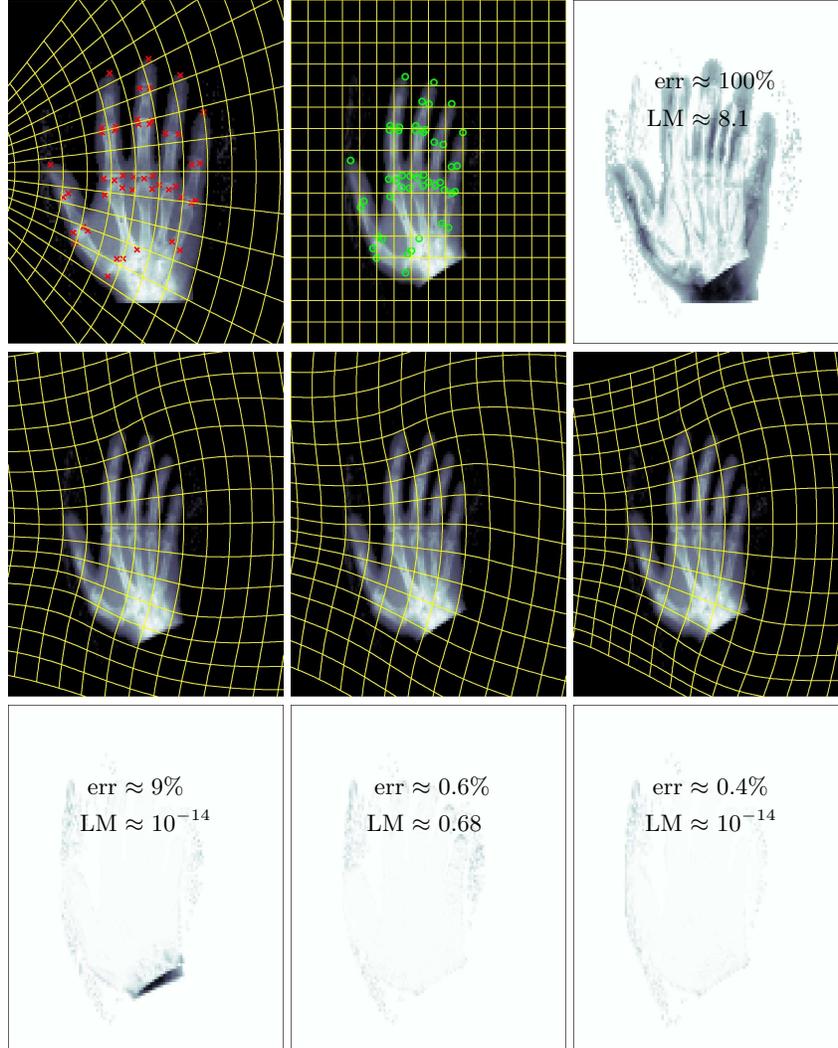


Fig. 4. Original template image with landmarks (crosses) and visualization of an artificial transformation y_{true} (top left); reference (top middle) is a transformed template $\mathcal{R}(x) = \mathcal{T}(y_{\text{true}}(x))$, with visualization of transformed landmarks and initial grid; initial difference $|\mathcal{T} - \mathcal{R}|$ (top right); transformed template $\mathcal{T}[y]$ based on thin-plate spline solution y_{TPS} (center left), unconstrained solution y_{un} (center middle), and constrained solution y_{con} (center right); differences $|\mathcal{T}[y] - \mathcal{R}|$ for $y = y_{\text{TPS}}$ (bottom left), $y = y_{\text{un}}$ (bottom middle), and $y = y_{\text{con}}$ (bottom right).

but a large difference for the trapezoid. The unconstrained approach results a very small difference but the landmark error is relatively large. Finally, the constrained approach performs perfect on the landmark and results the smallest difference. The later is due to the fact that the stopping criteria is relative to the initial guess, which is results a smaller distance in the constrained approach.

4.2 3D example

For our 3D experiment we use real data from CT and 3D power Doppler ultrasound (US) of a human liver. The goal of this application is the alignment of vessels that have been segmented from the original data. Consequently, we have binary images allowing for a direct comparison by the SSD distance measure. The size of the data in our experiment is $171 \times 165 \times 186$ voxels. Additionally, we have 11 corresponding landmarks that were manually picked by an expert; see Fig. 5 (a,b). For the registration we used four levels starting from $22 \times 21 \times 24$ and ranging to the original resolution with $171 \times 165 \times 186$ voxels. Results for a plain landmark based registration by using only the thin-plate-spline solution y_{TPS} and the constrained solution $y_{\text{con}} = y_{\text{TPS}} + u$ are shown in Fig. 5(c,d). As it turns out, the landmark solution provides a reasonable alignment but is far from being perfect. On the other hand, using the constrained approach improved the quality of the results considerably and leads to an almost perfect alignment of large parts of the vessel system.

5 Conclusions

The paper presents a variational framework for the landmark constrained registration problem and a discretize-then-optimize approach for computing a numerical solution. A difficulty for the multi-level discretization is that the number of constraints is constant while the number of degrees of freedom varies. In particular for a coarse discretization, inconsistent constraints are to be expected.

This paper provides a technique to overcome this problem by mixing landmark and update components, which results in compatible constraints. Moreover, it is shown how to efficiently compute a stable, orthogonal, and sparse basis for the constraint manifold and thus enabling a reduced space optimization avoiding saddle point problems.

References

1. Glasbey, C.: A review of image warping methods. *Journal of Applied Statistics* **25** (1998) 155–171
2. Pluim, J., Maintz, J., Viergever, M.: Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging* **22** (1999) 986–1004
3. Hajnal, J., Hawkes, D., Hill, D.: *Medical Image Registration*. CRC Press (2001)

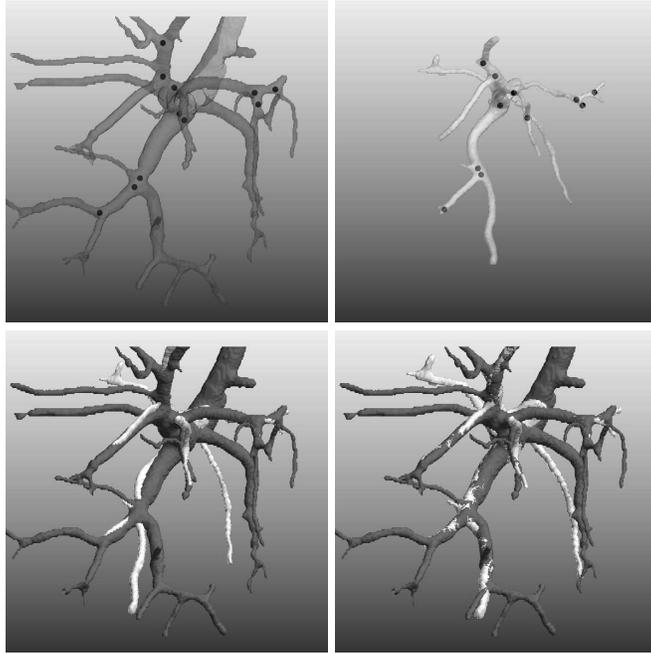


Fig. 5. 3D registration of CT and US. Reference (top left) \mathcal{R} with landmarks (black balls); (b) template \mathcal{T} with landmarks (top right); reference \mathcal{R} and deformed template $\mathcal{T}[y_{\text{TPTS}}]$ after landmark registration (bottom left); reference \mathcal{R} and deformed template $\mathcal{T}[y_{\text{con}}]$ after constrained registration (bottom right).

4. Modersitzki, J.: Numerical Methods for Image Registration. Oxford University Press (2004)
5. Goshtasby, A.A.: 2-D and 3-D Image Registration. Wiley Press, New York (2005)
6. Joshi, A., Shattuck, D., Thompson, P.: Brain image registration using cortically constrained harmonic mappings. In Sgallari, F., Murli, A., Paragios, N., eds.: Proceedings of IPMI, LNCS (2007)
7. Grady, L.: A lattice-preserving multigrid method for solving the inhomogeneous poisson equations used in image analysis. In Forsyth, D.A., Torr, P.H.S., Zisserman, A., eds.: Scale Space and Variational Methods in Computer Vision, SSVM, ECCV (2008)
8. Koestler, H.: A Multigrid Framework for Variational Approaches in Medical Image Processing and Computer Vision. Ph.d. dissertation, University of Erlangen (2008) Netherland.
9. Keller, S., Lauze, F., Nielsen, M.: Motion compensated video super resolution. In Sgallari, F., Murli, A., Paragios, N., eds.: Scale Space and Variational Methods in Computer Vision, SSVM, LNCS (2007)
10. Hadamard, J.: Sur les problmes aux drives partielles et leur signification physique. Princeton University Bulletin (1902) 49–52
11. Weickert, J., Schnörr, C.: A theoretical framework for convex regularizers in PDE-based computation of image motion. Int. J. Computer Vision **45**(3) (2001) 245–264

12. Hinterberger, W., Scherzer, O., Schnörr, C., Weickert, J.: Analysis of optical flow models in the framework of calculus of variations. *Num. Funct. Anal. Opt.* **23** (2002) 69–82
13. Droske, M., Rumpf, M.: A variational approach to non-rigid morphological registration. *SIAM Appl. Math.* **64**(2) (2004) 668–687
14. Bookstein, F.L.: Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(6) (1989) 567–585
15. Maurer, C.R., Fitzpatrick, J.M.: A Review of Medical Image Registration. In: *Interactive Image-Guided Neurosurgery*. Park Ridge, IL, American Association of Neurological Surgeons (1993) 17–44
16. Rohr, K.: *Landmark-based Image Analysis*. Computational Imaging and Vision. Kluwer Academic Publishers, Dordrecht (2001)
17. Fischer, B., Modersitzki, J.: Combining landmark and intensity driven registrations. *PAMM* **3** (2003) 32–35
18. Ashburner, J., Friston, K.: Spatial normalization using basis functions. In Frackowiak, R., Friston, K., Frith, C., Dolan, R., Friston, K., Price, C., Zeki, S., Ashburner, J., Penny, W., eds.: *Human Brain Function*. 2nd edn. Academic Press (2003)
19. Nocedal, J., Wright, S.J.: *Numerical optimization*. Springer, New York (1999)
20. Collignon, A., Vandermeulen, A., Suetens, P., Marchal, G.: 3d multi-modality medical image registration based on information theory. *Computational Imaging and Vision* **3** (1995) 263–274
21. Viola, P.A.: *Alignment by Maximization of Mutual Information*. PhD thesis, Massachusetts Institute of Technology (1995)
22. Clarenz, U., Droske, M., Rumpf, M.: Towards fast non-rigid registration. In: *Inverse Problems, Image Analysis and Medical Imaging, AMS Special Session Interaction of Inverse Problems and Image Analysis*. Volume 313., AMS (2002) 67–84
23. Haber, E., Modersitzki, J.: Intensity gradient based registration and fusion of multi-modal images. *Methods of Information in Medicine, Schattauer Verlag, Stuttgart* **46**(3) (2007) 292–299
24. Fischer, B., Modersitzki, J.: Fast curvature based registration of MR-mammography images. In Meiler, M., et al., eds.: *Bildverarbeitung für die Medizin*, Springer (2002) 139–143
25. Fischer, B., Modersitzki, J.: A unified approach to fast image registration and a new curvature based registration technique. *Linear Algebra and its Applications* **380** (2004) 107–124
26. Light, W.A.: Variational methods for interpolation, particularly by radial basis functions. In Griffiths, D., Watson, G., eds.: *Numerical Analysis 1995*, London, Longmans (1996) 94–106
27. Haber, E., Modersitzki, J.: A multilevel method for image registration. *SIAM J. Sci. Comput.* **27**(5) (2006) 1594–1607
28. Golub, G.H., van Loan, C.F.: *Matrix Computations*. Third edn. The Johns Hopkins University Press, Baltimore (2000)
29. Barrett, R., Berry, M., Chan, T.F., Demmel, J.W., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., van der Vorst, H.: *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. second edn. SIAM (1994)
30. Trottenberg, U., Oosterlee, C., Schuller, A.: *Multigrid*. Academic Press (2001)