Indirect Image Registration:
Geodesic Shooting for LDDMM and Deep Learning Approaches

Rosa Kowalewski

February 2018

Internship Project at

CAMBRIDGE IMAGE ANALYSIS
DAMTP - UNIVERSITY OF CAMBRIDGE

supervised by Dr. Carola B. Schönlieb

as a Part of the Master's Programme at

UNIVERSITÄT ZU LÜBECK

Contents

1 Large Deformation Diffeomorphic Metric Mapping for Indirect Image Registration

1.1 The LDDMM Model

The image registration problem is to find a deformation ϕ, that maps a given template image T onto a reference image R. Mathematically a gray scale image is modelled as a function from $\mathbb{R}^d$ to $\mathbb{R}$ that is zero outside a domain $\Omega \subset \mathbb{R}^d$ and satisfying certain smoothness assumptions. We will denote the space of images by X. The deformation ϕ is a mapping $\phi \colon \mathbb{R}^d \to \mathbb{R}^d$, assumed to be the identity mapping outside Ω.

One approach to solve this problem is by minimization of a cost functional

$$\mathcal{E}(R,T;\phi) = \mathcal{R}(\phi) + \lambda \mathcal{D}(R,T;\phi), \tag{1.1}$$

where $\mathcal{D}$ is a similarity measure for the deformed reference image and target image, and $\mathcal{R}$ is a regularization term that forces the deformation to satisfy certain smoothness conditions. The regularization parameter λ specifies the influence of the regularization term compared to the similarity measure.

For many applications it is desirable to have a deformation ϕ that is invertible, and both ϕ and ϕ^{-1} should be sufficiently smooth. This prevents the deformation from creating holes or foldings when applied to an image. Under this assumption the set of deformations forms a group with the identity mapping as neutral element.

We will consider diffeomorphisms $\phi \in \text{Diff}(\Omega)$ on a domain $\Omega \subset \mathbb{R}^d$ as deformations for the image registration framework. $\text{Diff}(\Omega)$ contains continuously differentiable maps that have a continuously differentiable inverse and therefore satisfies these assumptions.

To create a diffeomorphism ϕ one can consider small perturbations of the identity mapping

$$\phi = \text{Id} + v,$$

where v is a small enough vector field. For image registration it will generally not be enough to consider small deformations, hence we need to find a way to create larger deformations. One can consider the composition of small deformations

$$\phi_i = (\text{Id} + \epsilon v_i) \circ \phi_{i-1} = \phi_{i-1} + \epsilon v_i \circ \phi_{i-1}, \tag{1.2}$$

where $\phi_0 = \text{Id}$ and for some $\epsilon_0 > 0$, $\text{Id} + \epsilon v_i$ with $\epsilon < \epsilon_0$ is a diffeomorphism. Writing (1.2) as a difference equation $(\phi_i - \phi_{i-1})/\epsilon = v_i \circ \phi_{i-1}$ motivates the expression of a differential equation as a continuous form of it, depending on the time variable t.

Definition 1 (The Flow Equation). Let $v : \Omega \times [0,1] \to \Omega$, $v_t(x) := v(x,t)$ be a time-dependent vector-field. The flow equation reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_t^v = v_t \circ \phi_t^v,$$

and v is called the flow of ϕ^v.

2

To make sure that the evolved deformation ϕ_t^v is a diffeomorphism, it is necessary for v_t to satisfy strong enough smoothness constraints [11].

Definition 2 (Admissible Space)**.** V is an admissible space if it is continuously embedded in $C_0^1(\Omega, \mathbb{R}^d)$, or equivalently if there exists a constant $c > 0$, such that $\|u\|_{1,\infty} \leq c\|u\|_V$.

As a sufficient condition of the well-posedness of the flow equation, one assumes v_t to be in an admissible space. Section 1.1.1 will discuss the admissible space in more detail.

Note that the diffeomorphism generated by a given flow field is unique, while one diffeomorphism can generally be evolved from more than one flow field.

Considering again the energy functional (1.1), we now specify the functionals $\mathcal{R}(\phi)$ and $\mathcal{D}(R \circ \phi^{-1}, T)$. For the regularization term $\mathcal{R}$ one needs to define a metric on the space of deformations.

Theorem 1 (A Metric on G)**.** *For G a group of diffeomorphisms, $\psi, \bar{\psi} \in G$ and $\phi_t^v \in G$ the solution of the flow equation for a given vector field v_t,*

$$\mathrm{d}_G(\psi, \bar{\psi}) = \inf_{v_t \in V} \left\{ \int_0^1 \|v_t\|_V \mathrm{d}t, \psi = \bar{\psi} \circ \phi_{t=1}^v \right\}$$

is a metric on G and (G, d_G) is a complete metric space.

Proof. A proof can be found in [17, chapter 8.2] $\square$

As the similarity metric $\mathcal{D}$ in the image space we take the squared L^2-norm $\|\cdot\|_{L^2}$. Plugging in these terms into the energy functional (1.1), we obtain the the optimization problem as

Problem 1 (Optimization over Velocity Fields)**.**

$$\arg\min_{v_t} \ \mathcal{E}(v_t) = \frac{1}{2} \int_0^1 \|v_t\|_V \mathrm{d}t + \frac{\lambda}{2} \|R \circ (\phi_1^v)^{-1} - T\|_{L^2}^2 \tag{1.3}$$

$$s.t. \qquad\qquad \begin{aligned} \phi_t &= v_t \circ \phi_t \\ \phi_0 &= \mathrm{id} \end{aligned} \tag{1.4}$$

1.1.1 Admissible Spaces: Reproducing Kernel Hilbert Spaces

Admissible spaces fall into the class of reproducing kernel Hilbert spaces [2], which are introduced in the following.

Definition 3 (Reproducing Kernel Hilbert Space)**.** Consider a Hilbert space $V \subset L^2(\Omega, \mathbb{R}^d)$. V is a *reproducing kernel Hilbert space* (RKHS) if the point evaluation functional $\delta_x \colon V \to \mathbb{R}$, $\delta_x(v) := v(x)$ is continuous on V.

This property allows the definition of the reproducing kernel of V. The Riesz representation theorem states that, for every element μ in the dual space V^*, there exists a unique representation $\mu(y) = \mu_x(y) = \langle x, y \rangle$, $x \in V$. This implies that there exists a unique $K(x) \in V$ with the reproducing property $\mu(x) = \delta_x(\mu) = \langle \mu, K(x) \rangle \forall \mu \in V$. The kernel K has the properties to be symmetric and positive definite [17].

In [17], a method for creating such a RKHS is introduced. An operator $L \colon D \to H$ is chosen. Here H is a Hilbert space, typically $H = L^2(\Omega, \mathbb{R})$ and D is densely included in H. With certain

monotonicity contraints on L, the induced inner product $\langle\cdot,\cdot\rangle_L = \langle\cdot, L\cdot\rangle_{L^2}$ can be extended to a dense subspace $V \subset H$, which is a RKHS. L can be extended to an operator $\hat{L}\colon V \to H$.

On the space V there exist two different scalar products, one being the L^2 scalar product defined by the Hilbert space H, which we will denote $\langle\cdot,\cdot\rangle_{L^2}$ or $\langle\cdot,\cdot\rangle$, and the scalar product induced by the reproducing kernel denoted by $\langle\cdot,\cdot\rangle_V$.

It can be shown that the reproducing kernel K for a space designed like this is well-defined on V [17]. This leads to the property $\langle K(x), v\rangle_V = \langle\hat{L}K(x), v\rangle = \delta_x(v)$, or $K = \hat{L}^{-1}$. K can be seen as the inverse Operator of $\hat{L}$ and is also called its Green's function.

In the LDDMM setting, it is common to regularize by minimization over the L^2-norm of a self-adjoint differential operator F applied to v^e. The operator F is defining the reproducing kernel $K^{-1} := L := F^\top F$ of the space, as

$$\int_0^1 \| v_t \|_V \, \mathrm{d}t = \int_0^1 \| Fv_t \|_{L_2} \, \mathrm{d}t \quad = \langle Fv_t, Fv_t\rangle_{L^2} = \langle Lv, v\rangle_{L^2} \quad = \langle m_t, Km_t\rangle_{L^2} = \langle m_t, v_t\rangle$$

$$(1.5)$$

A common choice for K in the LDDMM setting is the Gaussian kernel, which corresponds to the differential operator $F = -\alpha \bigtriangledown^2 + \gamma\mathrm{Id}$, where $\bigtriangledown^2$ is the Laplacian operator, see [8] and [3]. Other kinds of kernels have been used and will have a different smoothing property.

1.1.2 Lagrangian and Eulerian Reference Frames and Adjoint Action

The previously defined velocity field is given in Eulerian coordinates. This means it is seen as observed in the current configuration at each time point, therefore depending on the deformation that happened before. Another possibility of describing the velocity is to see it from the initial configuration, which is called the velocity field in the Lagrangian frame or Lagrangian coordinates. In the following chapters we will write v^l for the velocity in the Lagrangian frame and v^e for the velocity field given in Eulerian coordinates.

Figure 1.1 illustrates the correspondence of the two reference frames. On the left, the velocity is shown as in the initial configuration. The velocity vector is given relative to the neighbouring control points. By the deformation ϕ the initial configuration is transformed into the Eulerian reference frame. The position of the control points is changed aswell as the direction and length of the velocity vector in the spatial grid, while the relative position to the neighbouring control points remains the same.
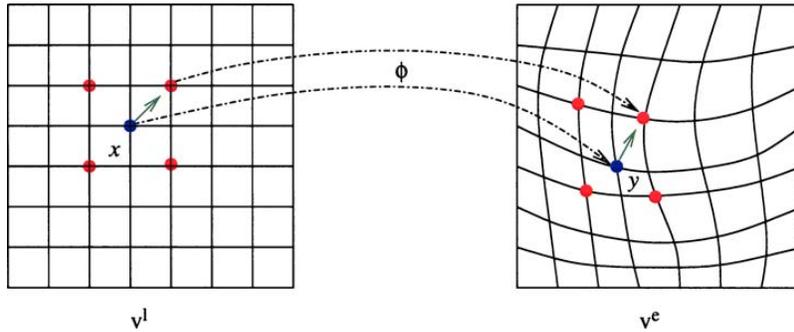
Figure 1.1: Velocity fields in Lagrangian and Eulerian reference frame

4

It will be necessary to change between the two frames, which is done by the adjoint action.

Definition 4 (Adjoint action of the group G on V)**.** The adjoint action of $\phi \in G$ on the tangent space V is

$$\mathrm{Ad}_\phi \colon V \to V$$
$$\mathrm{Ad}_\phi v \coloneqq \mathrm{d}\phi \cdot v \circ \phi^{-1}.$$

The Langrangian velocity can be defined via

$$v^l = \mathrm{Ad}_{\phi^{v^{-1}}} v^e.$$

The term adjoint action comes from Lie-group theory, where it is interpreted as the push-forward of v by ϕ. In the literature, LDDMM theory is often described in the terminology of Lie groups, interpreting the function space of velocity fields V as the tangent space (or Lie algebra) of the diffeomorphism group G. However, even though this terminology brings in useful notation and intuition, one can not speak of a Lie group structure on G in a rigorous sense since the infinite-dimensional setting is not compatible with the regularity assumptions of classical Riemannian geometry.

By differentiating Ad at the identity of the group, we get the adjoint action $\mathrm{ad}_u : V \to V$ on the tangent space V:

Definition 5 (Adjoint Action of V on itself and Coadjoint Action)**.** For u, v in an admissible space V, the adjoint action of V on itself is

$$\mathrm{ad}_u v \coloneqq \partial_\phi(\mathrm{d}\phi \cdot v \circ \phi^{-1})|_{\phi=\mathrm{Id}} u = \mathrm{d}u \cdot v - \mathrm{d}v \cdot u$$

The coadjoint actions $\mathrm{Ad}^* \colon G \times V^* \to V^*$ on the group G and $\mathrm{ad}^* \colon V \times V^* \to V^*$ on V are defined via

$$\langle \mathrm{Ad}_\phi^* \mu, u \rangle = \langle \mu, \mathrm{Ad}_\phi u \rangle,$$

and

$$\langle \mathrm{ad}_v^* \mu, u \rangle = \langle \mu, \mathrm{ad}_v u \rangle.$$

The coadjoint actions act as the transpose of the operators Ad and ad in the sense of linear algebra, acting on the dual space of V.

1.2   Geodesic Shooting

The optimization problem (1.3) is expressed as a minimization over the time-dependent velocity fields v_t. The flow of a diffeomorphism is not unique, so the idea of the shooting formulation is to not optimize over all velocity fields v_t, but instead only over those with the least norm for a given ϕ. In Lie-group terminology we want to find the geodesics on the manifold G.

Considering a minimizer v^* of the optimization problem (1.3), v^* has to be a minimizer of the regularization term $\int_0^1 \| v_t \|_V \, \mathrm{d}t$ with the constraint (1.4) for a fixed deformation ϕ. The shooting equations are equivalent to this statement.

Similar to the physical setting of a dynamical system, one can introduce the concept of momentum for the image deformation case.

Definition 6 (Momenta in the Eulerian and Lagrangian Reference Frames)**.** The Eulerian momentum m^e is defined as the dual of the velocity in the RKHS,

$$m^e := Lv^e,$$

and the Lagrangian momentum is defined by

$$m^l := \mathrm{Ad}_{\phi^{-1}}^* m^e.$$

The Lagrangian momentum m^l is defined in such a way that

$$\langle Lv^e, v^e \rangle = \langle m^e, v^e \rangle = \langle m^l, v^l \rangle.$$

Using the identity $\|v_t\|_V^2 = \langle m_t, Km_t \rangle$ from equation (1.5), the energy functional can be expressed as a function of the momentum m_t^e, giving the minimization problem

Problem 2 (Optimization over the Momentum)**.**

$$
\begin{aligned}
\underset{m_t^e}{\arg\min} \quad & \mathcal{E}(m_t^e) = \frac{1}{2} \int_0^1 \langle m_t^e, Km_t^e \rangle \mathrm{d}t + \frac{\lambda}{2} \parallel R \circ \phi_1^{-1} - T \parallel^2 \\
s.t. \quad & \phi_t = v_t^e \circ \phi_t \\
& \phi_0 = \mathrm{id} \\
& m_t^e - Lv_t^e = 0.
\end{aligned}
$$

Here the constraints (1.4) are extended by the constraints that define the correspondence of the momentum m^e to the velocity v^e.

1.2.1 Geodesic Equation for the Lagrangian Reference Frame - Conservation of momentum

Considering the path of least energy for a particle in the Lagrangian reference frame leads to a very intuitive result known from mechanics. An important theorem in mechanics is Noether's theorem, which states that if a mechanical system is invariant under the action of a group G, the generalized momentum is conserved. For the Lagrangian reference frame the velocity v_t^l is always given as in the initial configuration. Hence, it doesn't depend on the deformation ϕ_t and is invariant under the group action. Due to Noether's theorem, the Lagrangian momentum m_t^l is a constant of motion, i.e.

Theorem 2 (Conservation of momentum)**.**

$$\frac{\mathrm{d}}{\mathrm{d}t} m_t^l = 0. \tag{1.6}$$

Proof. In [2], this is proven by directly differentiating the energy functional (1.3). □

Knowing the initial momentum m_0^l, it follows from Theorem 2, that the momentum and therefore the velocity field over the whole time interval $[0, 1]$ is known. Adding this as a constraint to the problem 2 allows optimization over the initial momenta instead of time-dependent momenta, reducing the parameter space. In the following the equivalent formulation for the Eulerian reference frame is examined.

1.2.2 Geodesic Equation for the Eulerian Reference Frame - The Euler-Poincaré Equation

To obtain the geodesic equation for the Eulerian velocity v^e, one can apply the Euler-Lagrange equation to the regularization term. We define $\phi \colon [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$ as the deformation over time and $\phi_t \coloneqq \phi(t, \cdot)$ as the deformation for a given timepoint t. We consider the energy

$$\mathcal{L}(g, w) = \frac{1}{2} \langle L(w \circ g^{-1}), w \circ g^{-1} \rangle$$

as Lagrange-function. Evaluating $\mathcal{L}(\phi_t, \partial_t \phi_t)$ gives $\mathcal{L}(\phi_t, \partial_t \phi_t) = \frac{1}{2} \langle Lv^e, v^e \rangle$, the kinetic energy for the LDDMM regularization term. Now we are interested in finding the minimizer of

$$\mathcal{S}(\phi) = \int_0^1 \mathcal{L}(\phi_t, \partial_t \phi_t) \; \mathrm{d}t.$$

The Euler-Lagrange equation states that for stationary points of the functional $\mathcal{S}(\phi)$, ϕ must satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{\partial \mathcal{L}}{\partial w} (\phi_t, \partial_t \phi_t) \right] = \frac{\partial \mathcal{L}}{\partial g} (\phi_t, \partial_t \phi_t). \tag{1.7}$$

Calculating the derivatives of $\mathcal{L}$ and inserting into (1.7) as in [7] will give the Euler-Poincaré equation:

Theorem 3 (The Euler-Poincaré Equation). *Along a geodesic path, the Eulerian momentum m_t^e satisfies the Euler-Poincaré equation*

$$\frac{\mathrm{d}}{\mathrm{d}t} m_t^e + \mathrm{ad}_{v_t^e}^* m_t^e = 0,$$

which can be explicitly written as the EPDiff equation

$$\frac{\mathrm{d}}{\mathrm{d}t} m^e + \mathrm{d}m^e v^e + \mathrm{d}v^e m^e + m^e \mathrm{div} v^e = 0. \tag{1.8}$$

The Euler-Poincaré equation holds for any Lie group with right invariant Lagrangian, while EPDiff is its special case for a diffeomorphism group. The Euler-Poincaré equation can also be derived as an equivalent formulation of the geodesic equation for the Lagrangian momentum. In [2] this is done by inserting $m_t^l = \mathrm{Ad}_{\phi_t^{-1}}^* m_t^e$ in (1.6) and differentiating the operator Ad^* with respect to time.

The shooting formulation of problem 2 is

Problem 3 (Shooting Formulation for the Eulerian Momentum).

$$\underset{m_0^e}{\arg\min} \quad \mathcal{E}(m_0^e) = \frac{1}{2} \langle m_0^e, K m_0^e \rangle + \frac{\lambda}{2} \parallel R \circ \phi^{-1}(1) - T \parallel^2$$

$$s.t. \quad \phi_t = v_t^e \circ \phi_t$$
$$\phi_0 = \mathrm{id}$$
$$m_t^e - L v_t^e = 0.$$
$$\frac{\mathrm{d}}{\mathrm{d}t} m_t^e + \mathrm{ad}_{v_t^e}^* m_t^e = 0$$

The shooting step to obtain the deformed image from the initial momentum is done by integrating (1.8), then convolving m_t for each timestep with the kernel K to obtain the velocity fields. By integration of the velocity fields one obtains the deformation that can be applied to the image.

1.2.3 Geodesic Equations for the Hamiltonian Frame

Another approach in the literature is to consider the Hamiltonian system. The reduced Hamiltonian function $\mathcal{H}(g,p)$ is derived from the Lagrangian $\mathcal{L}(g,w)$ by a change of variables $(g,w) \longmapsto (g,p)$ through the Legendre-transformation. This leads to a system of first order differential equations, as opposed to the Euler-Lagrange equation as a second order differential equation. The introduced variable p is called the conjugate momentum or generalized momentum. The Hamiltonian function is defined as

$$\mathcal{H}(g,p) = p \cdot w - \mathcal{L}(g,w),$$

where w depends on g and p and is the solution of the equation

$$p = \frac{\partial \mathcal{L}}{\partial w}(g,w). \tag{1.9}$$

Hamilton's equations state that if there are p,g that minimize the energy, then they must satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}p = -\frac{\partial \mathcal{H}}{\partial g} \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t}g = \frac{\partial \mathcal{H}}{\partial p}.$$

The first equation follows from the Euler-Lagrange equation. The second equation follows directly from the definition of $\mathcal{H}(g,p)$.

Evaluating (1.9) for the LDDMM setting with $g = \phi$ and $w = \frac{\mathrm{d}}{\mathrm{d}t}\phi$ will lead to the conjugate momentum p that corresponds to the momenta m^l and m^e as $p = \mathrm{d}\phi^{-\top}m^l = |\mathrm{d}\phi|\, m^e \circ \phi$. Similar to the Lagrangian frame the momentum is given in the initial configuration.

One can check that p and w are defined in such a way that $\langle p,w \rangle = \frac{1}{2}\langle L(w \circ \phi^{-1}), w \circ \phi^{-1} \rangle$, which gives the Hamiltonian

$$\mathcal{H}(\phi,p) = \langle p,w \rangle - \mathcal{L}(\phi,w)$$
$$= \langle p,w \rangle - \frac{1}{2}\langle L(w \circ \phi^{-1}), w \circ \phi^{-1} \rangle$$
$$= \frac{1}{2}\langle p,w \rangle$$

When solving (1.9) for w, one will obtain the expression

$$w(g,p) = \int_{\mathrm{R}^d} K(g(x), g(y))p(y)\mathrm{d}y.$$

Defining K^g as the kernel evaluated at g, that gives $w = K^g p$, we have $\mathcal{H}(g,p) = \frac{1}{2}\langle p, K^g p \rangle$. Evaluating the partial derivatives of $\mathcal{H}(\phi,p)$ leads to the geodesic equations

$$\frac{\mathrm{d}}{\mathrm{d}t}p = -(\nabla K^g \circ \phi)p$$
$$\frac{\mathrm{d}}{\mathrm{d}t}g = (K^g \circ \phi)p, \tag{1.10}$$

see [13], [12] for a more detailed derivation of this. The solutions of the Hamiltonian system are then obtained solving (1.10) for g forward in time and for p backward in time, see Algorithm 1 in [12].

1.2.4 Comparison of Optimization Strategies

An important advantage of the shooting method is the reduction of the parameter space, which makes the algorithm more efficient. In [17] it is stated that convergence of the shooting method is fast, but not guaranteed. Especially for large complex deformations the sensibility to small changes in the control variables can be large, which leads to an unstable optimization procedure.

The geodesic equations result in slightly more complex computations, while optimization over time-dependent quantities requires a larger computation memory. However, those algorithms are more less likely to get stuck in local minima for large deformations. The solutions of such an optimization algorithm are not guaranteed to be a geodesic on G, but can differ significantly from the geodesic equations unless the discretization in time is chosen fine enough.

For the shooting method, we examined the different shooting equations for the Lagrangian, Eulerian and Hamiltonian momentum. In the Lagrangian framework, the position of points on which the momenta are given, is changing over time. This makes the convolution and integration step computationally more demanding than for a fixed regular grid. The convolution kernel K has to be evaluated on a different grid for each timestep. For the integration, [1] say that Lagrangian integration schemes are more computationally stable and provide good numerical accuracy for a larger size of time step, compared to the Eulerian integration scheme.

For the Eulerian framework the momenta at discretized timepoints can be computed by the EPDiff-equation, and then convolved to obtain the velocity field at each timepoint for the integration. Having the vectors given on the same spatial grid at each timepoint makes the implementation easier.

In the Hamiltonian framework the geodesic equations are solved by integrating $\frac{\mathrm{d}}{\mathrm{d}t}q$ forward in time and $\frac{\mathrm{d}}{\mathrm{d}t}p$ backward in time, see [12]. Shooting algorithms for the Eulerian momentum and for the Hamiltonian momentum are widely used, whereas a shooting algorithm for the Lagrangian momentum is not common.

Note that the initial momenta are the same for each reference frame. The difference is how to look at the geodesic path on G.

At convergence, the initial momentum will be parallel to the image gradient. This is because a deformation along the level set of the image will increase the regularization term while giving the same deformed image. Thus the optimization can be done over a scalar momentum α lying in the image space and satisfying $m_0 = \alpha_0 \bigtriangledown R$. The scalar momentum α has lower dimensionality than the vectorvalued momentum m, which helps speeding the optimization algorithm. However the computations are often less stable.

1.3 Geodesic Shooting for Indirect Image Registration

In indirect image registration the target is not given in the image space X but only observed as noisy data in the data space Y. The forward operator $\mathcal{T} \colon X \to Y$, $\mathcal{T}(T) = g$ gives the relation between the image and corresponding data in the data space. The optimization problem 3 can be reformulated as

Problem 4 (Shooting Formulation of the Indirect Image Registration Problem)**.**

$$\arg\min_{m_0^e} \quad \mathcal{E}(m_0^e) = \frac{1}{2}\langle m_0^e, Km_0^e \rangle + \frac{\lambda}{2} \parallel \mathcal{T}(R \circ \phi^{-1}) - g \parallel^2$$

$$s.t. \quad \phi_t = v_t^e \circ \phi_t$$

$$\phi_0 = \mathrm{id}$$

$$m_t^e - Lv_t^e = 0.$$

$$\frac{\mathrm{d}}{\mathrm{d}t}m_t^e + \mathrm{ad}_{v_t^e}^* m_t^e = 0$$

The similarity between the transformed reference image and the target image is now measured in the data space. Depending on the kind of data the norm is chosen. Since the forward operator has no effect on the regularization term, the geodesic equations remain the same as in 3.

It is not obvious that the new minimization problem also converges to an existing solution. In [4] a proof of existence of a minimizer for the indirect registration problem is given. For the proof of convergence the forward operator $\mathcal{T}$ has to satisfy certain conditions.

1.4   Numerical Results

The indirect registration problem was implemented in Python using the Operator Discretization Library (ODL). The code is available at https://github.com/RosaKow/LDDMM.

The kernel was set to be a Gaussian kernel with standard deviation $\sigma = 5$. For the experiments the forward operator $\mathcal{T}$ was chosen to be the Ray transform, with the L^2-norm as the metric in the data space. For the optimization the gradient of the energy functional was computed using the adjoint equations as in [14]. The optimization was done by gradient descent with fixed stepsize. During the optimization process it happened that the momentum map and the velocity field exceeded the image domain, which could not be handled by the algorithm. A first attempt was to stop the iterations as soon as this problem appeared. For the examples shown here, the algorithm did not converge yet. To adress this problem one could try to use a kernel that vanishes on the image boundaries.

For the examples shown in the following, the template image is shown in the image space for better visualisation. As input for the energy functional the template was transformed to the data space. Figure 1.2 shows the result of the registration of the letters 'J' and 'V' as well as the computed initial momentum. In figure 1.3 the evolution of the template image over 10 timesteps is visualized for the same example. Figure 1.4 shows the results of the registration with noise added to the reference image in the image space. The deformed image is only very slightly less accurate. Figure 1.5 shows that the algorithm has slight difficulties remaining straight parallel edges. The straight edges of the square are blurred and rounded. This might be due to the using the Gaussian kernel. For an application where it is important to preserve this kind of features, one might consider using other kernels. Figure 1.6 shows that the algorithm is able to provide a good deformed template. The momentum is quite high in the inner region of the shape, and is quite smooth over the entire domain. Theoretically, the momentum vectors should be parallel to the image gradient. From this one can assume that the algorithm is not completely converged yet.

(a) Template image

(b) Reference image

(c) Initial momentum

(d) Deformed template
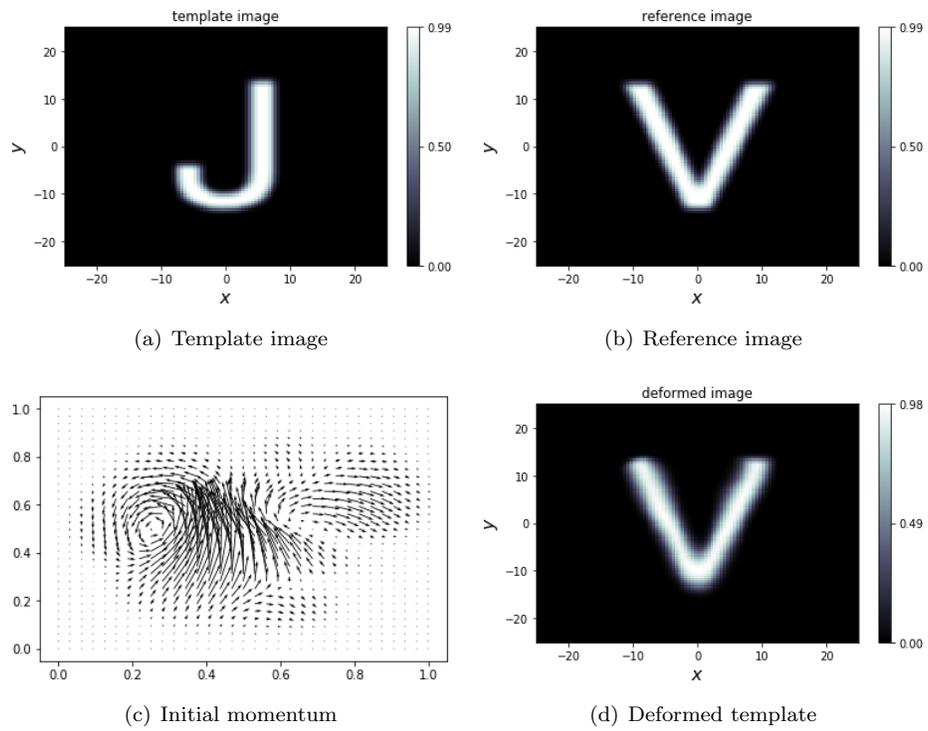
Figure 1.2: Registration results for LDDMM shooting algorithm
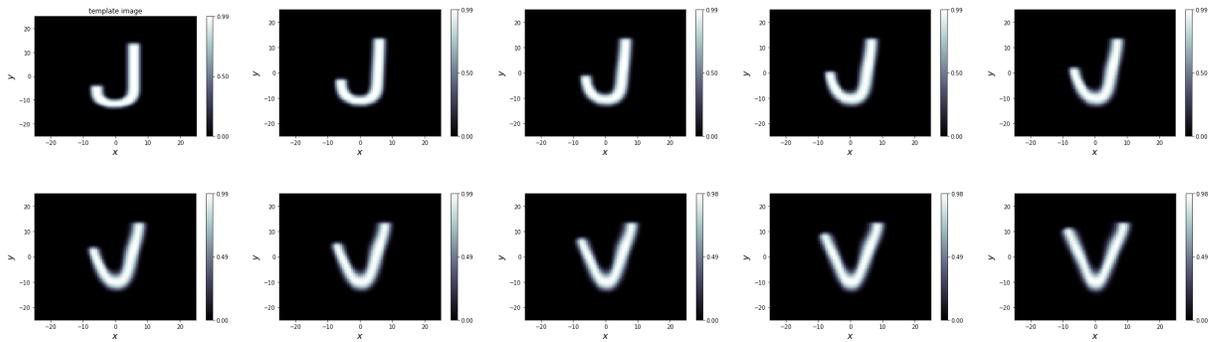
Figure 1.3: Template image deforming over time

(a) Template image   (b) Reference image

(c) Initial momentum   (d) Deformed template
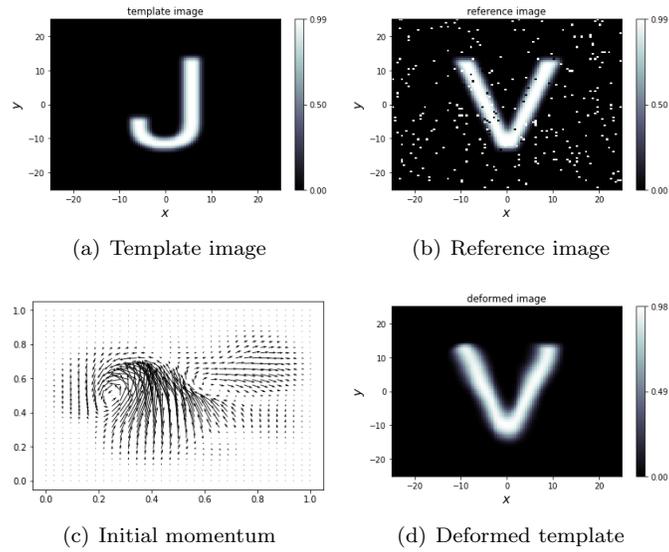
Figure 1.4: Registration results for LDDMM shooting algorithm with noisy data

(a) Template image   (b) Reference image

(c) Initial momentum   (d) Deformed template
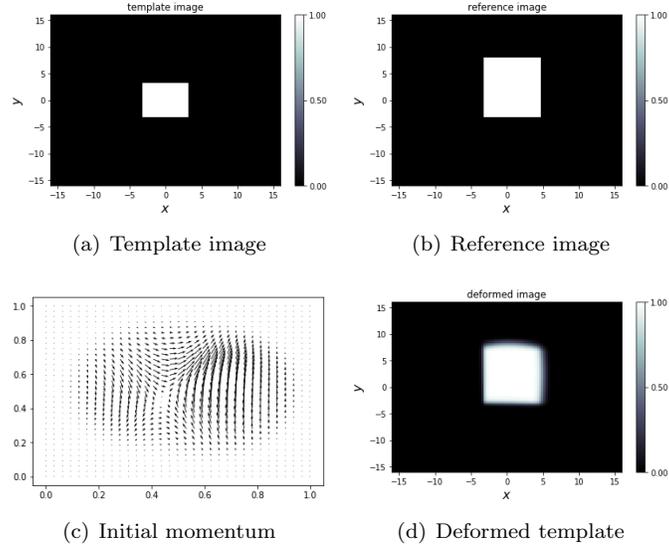
Figure 1.5: Registration results for LDDMM shooting algorithm

(a) Template image

(b) Reference image

(c) Initial momentum

(d) Deformed template
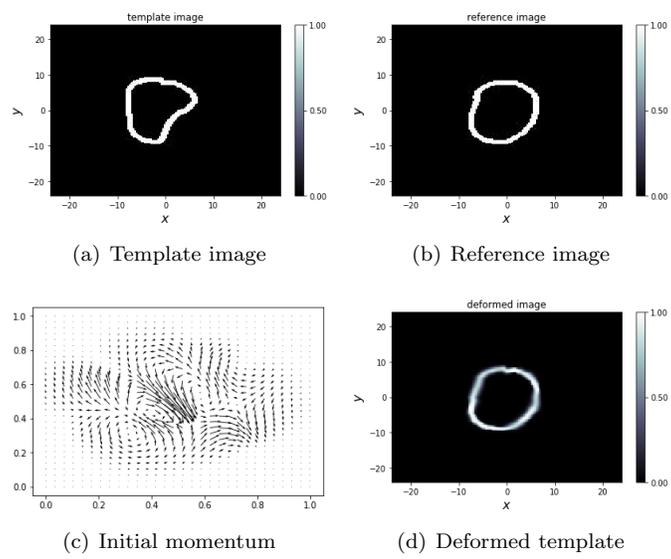
Figure 1.6: Registration results for LDDMM shooting algorithm

13

2 Deep Learning Approaches for Indirect Image Registration

2.1 Deep Learning for Image Registration

In some applications, e.g. in medical imaging, a real-time registration algorithm is required. Most state-of-the-art registration algorithms are too complex to provide good results in real time. The advantage of neural networks is that once trained, they are very fast in producing the output for new input data. For this reason, several approaches of applying learning algorithms in image registration have been made.

In deep learning one distiguishes between supervised and unsupervised learning algorithms. In supervised learning, the network is trained with a pair of input data and desired output data. In contrast, unsupervised learning algorithms don't have access to the ground truth output but work as functions to describe hidden structures from the data.

For a supervised learning approach for image registration, the training data is typically given as the set

$$\Sigma := \{R_i, T_i, \phi_i\} \subset X \times X \times G$$

of corresponding images, where $T_i \approx \mathcal{W}(R_i, \phi_i) := R_i \circ \phi_i^{-1}$ and G is the deformation group. The intention is to find the approximate inverse operator $\Lambda_\theta \colon X \times X \to G$ of the operator $\mathcal{W}$ w.r.t the deformation. This is done by minimizing a loss function

$$\mathcal{L}(\theta) = \sum_i \|\Lambda_\theta(R_i, T_i) - \phi_i\|_G \tag{2.1}$$

over θ.

For an unsupervised learning approach the loss function is similar to the energy function (1.1):

$$\mathcal{L}(\theta) = \sum_i \|\mathcal{W}(R_i, \Lambda_\theta(R_i, T_i)) - T_i\|_X \tag{2.2}$$

It is also possible to add a regularization term here. In the loss function (2.1) the accuracy is measured in the space of deformation. In contrast, the loss function (2.2) measures the accuracy of Λ_θ in the space of images. For both cases the choice of the norm is essential for the result.

Both supervised and unsupervised learning algorithms have been applied for image registration tasks, of which some examples are described in the following. In [5] the deformation is directly being learned end-to-end. In a convolutional network a displacement vector field is generated and the similarity is measured of the warped template and reference image in the image space. In [15], authors use a convolutional neural network for preprocessing, by extracting features that are important for predictions of local displacement from images and constructing a low-dimensional representation which is then fed into a registration algorithm. A supervised learning approach with given groundtruth deformation, [10] used CNNs with generated training data for rigid image registration. Rigid registration has much fewer parameters than using a deformation field, and therefore is easier to learn by a deep network. In [6] and [9], authors used a supervised CNN for

deformation prediction, that is then used as an initial guess for a registration algorithm. In [16], the initial momenta from the LDDMM setting are learned in a supervised learning framework. The following section gives a further describtion of this approach.

2.2 Deep Learning for LDDMM

The momenta in the LDDMM shooting framework have the property to be much less smooth than the resulting deformation field. This is particularly convenient for learning, as the momenta provide more characteristic features for a particular image than the smoother deformations. The prediction of the momenta is likely more robust than the prediction of the deformation field.

In [16] a deep convolutional neural network is used to learn these initial momenta. The network architecture consists of an encoder and a decoder. The encoder is fed with the corresponding reference and template image and works as a feature extraction for both images independently. The extracted features are then concatenated and fed into the decoder. The decoder consists of three independant convolutional networks that predict the momenta for the three dimensions. To recover from prediction errors, a correction networks of the same archtitecture is used to predict the prediction error. While feeding the network with whole images would require a substantially large capacity of the network, only patches of images are used as input here. This also has the advantage that from relatively few images and ground truth momenta, a large amount of training data is obtained. The patches are extracted from the reference, template and deformation at the same spatial grid locations. For this to be reasonable, it needs to be assumed that the deformation is relatively small, i.e. the deformed patch lies (at least mostly) in the same patch in the template image.

2.3 Learning Strategies for Indirect LDDMM Image Registration

In indirect image registration the target is not given in the image space X but only observed as noisy data in the data space Y. The forward operator $\mathcal{T} \colon X \to Y, \mathcal{T}(T_i) = g_i$ gives the relation between the image and corresponding data in the data space. The training data now consists of pairs of images and data

$$\Sigma := \{R_i, g_i\} \subset X \times Y,$$

where $g_i \cong \mathcal{T} \circ \mathcal{S}(R_i, m_i)$ for an unknown momentum m_i. Now one is interested in learning the operator

$$\Lambda_\theta \colon X \times Y \to V^*,$$

where V^* is the space of momenta. This can be done by minimizing one of several possible cost functions.

For a supervised learning approach a deep convolutional network can be used to minimize the cost function

$$\theta^* \in \arg\min_\theta \sum_i \|\Lambda_\theta(R_i, g_i) - m_i\|_{V^*}$$

Here it is assumed that the ground truth initial momentum is known. This can either be computed from an existing registration algorithm, like in the previously described framework. The accuracy

15

of the learned momentum is obviously restricted by the accuracy of the underlying registration algorithm. However the computation will be significantly faster.

Another possibility is to use generated training data. The loss function for generated training data would read

$$\theta^* \in \arg\min_{\theta} \sum_i \|\Lambda_\theta(R_i, \mathcal{T} \circ \mathcal{S}(R_i, m_i)) - m_i\|_{V^*}.$$

Here $\mathcal{S} \colon X \times V^* \to X$ is the shooting operator, that maps an initial momentum and an image to the deformed image. For data generation one can pick an image and an initial momentum to then obtain the data g by the shooting algorithm and forward operator. The difficulty with this is, that the training data needs to be similarly distributed as the actual data one wants to feed to the network. Randomly chosen momenta will likely not lead to a plausible deformation field. It can be shown, that a momentum vector field will be parallel to the image gradient when minimizing the energy functional. This is because deforming constant regions of the image will not lead to a change of the data term but increase the regularization term in the energy functional (1.1). It would therefore make sense to use the scalar valued momentum α with

$$m = \alpha \cdot \nabla R$$

for training data generation. Which further conditions the momenta should satisfy is still to be investigated. It would be interesting if one could use a network that uses both precomputed and generated groundtruth momenta to have the advantages of real images as training data and at the same time not being limited by the accuracy of the used registration algorithm.

Another question in this formulation is how to chose the norm $\|\cdot\|_{V^*}$ for a sensible comparison of momenta. The standard L^2-norm would most likely not be a good choice, because it is using pointwise differences. For comparing momenta it would make sense to use a norm that takes the whole geometry of the vectorfield into account. It needs to be investigated which choice of norm has the desired properties for comparing initial momenta. To avoid this issue it is possible to formulate the loss function in the data space, as in

$$\theta^* \in \arg\min_{\theta} \sum_i \|\mathcal{T} \circ \mathcal{S}(R_i, \Lambda_\theta(R_i, g_i)) - \mathcal{T} \circ \mathcal{S}(R_i, m_i)\|_Y$$

One can also use a network that is trained withour the groundtruth momenta as a function approximation network with the loss function

$$\theta^* \in \arg\min_{\theta} \sum_i \|\mathcal{T} \circ \mathcal{S}(R_i, \Lambda_\theta(R_i, g_i)) - g_i\|_Y.$$

Here, the sought-for operator $\Lambda_{\theta,R}$ is forced to approximate the inverse of the composition of the operators $\mathcal{T} \circ \mathcal{S}_R$. The metric is in the data space, as no momentum is directly involved in the formulation.

For the network architecture, it would be possible to use a similar architecture as in [16]. For indirect image registration it would propably not work that well to extract image patches, because in the data space the information of the patch might be spread over the whole data, as for example in a sinogramm of a tomographic image. It would be interesting to see, which of the proposed loss functions work best for which kind of application. For now this is an open question for future work.

Bibliography

[1] M. Faisal Beg, Michael I. Miller, Alain Trouvé, and Laurent Younes. Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, Feb 2005.

[2] Martins Bruveris and Darryl Holm. Geometry of image registration: The diffeomorphism group and momentum maps. 73, 06 2013.

[3] Vincent Camion and Laurent Younes. *Geodesic Interpolating Splines*, pages 513–527. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[4] Chong Chen and Ozan Öktem. Indirect image registration with large diffeomorphic deformations. *arXiv preprint arXiv:1706.04048*, 2017.

[5] Bob D. de Vos, Floris F. Berendsen, Max A. Viergever, Marius Staring, and Ivana Isgum. End-to-end unsupervised deformable image registration with a convolutional neural network. *CoRR*, abs/1704.06065, 2017.

[6] Benjamin Gutierrez-Becker, Diana Mateus, Loic Peter, and Nassir Navab. Guiding multimodal registration with learned optimization updates. *Medical Image Analysis*, 2017.

[7] Darryl Holm, Tanya Schmah, and Cristina Stoica. *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*. 01 2009.

[8] Sarang C Joshi and Michael I Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370, 2000.

[9] Minjeong Kim, Guorong Wu, Pew-Thian Yap, and Dinggang Shen. A generalized learning based framework for fast brain image registration. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 306–314, 2010.

[10] Shun Miao, Z Jane Wang, and Rui Liao. A cnn regression approach for real-time 2d/3d registration. *IEEE transactions on medical imaging*, 35(5):1352–1363, 2016.

[11] Michael I. Miller, Alain Trouvé, and Laurent Younes. Geodesic Shooting for Computational Anatomy. *Journal of Mathematical Imaging and Vision*, 24(2):209–228, Mar 2006.

[12] Michael I Miller, Alain Trouvé, and Laurent Younes. Hamiltonian systems and optimal control in computational anatomy: 100 years since d'arcy thompson. *Annual review of biomedical engineering*, 17:447–509, 2015.

[13] Anna Mills, Tony Shardlow, and Stephen Marsland. Computing the geodesic interpolating spline. In *International Workshop on Biomedical Image Registration*, pages 169–177. Springer, 2006.

[14] N. Singh, J. Hinkle, S. Joshi, and P. T. Fletcher. A vector momenta formulation of diffeomorphisms for improved geodesic regression and atlas construction. In *2013 IEEE 10th International Symposium on Biomedical Imaging*, pages 1219–1222, April 2013.

[15] Guorong Wu, Minjeong Kim, Qian Wang, Brent C Munsell, and Dinggang Shen. Scalable high-performance image registration framework by unsupervised deep feature representations learning. *IEEE Transactions on Biomedical Engineering*, 63(7):1505–1516, 2016.

[16] Xiao Yang, Roland Kwitt, and Marc Niethammer. Quicksilver: Fast predictive image registration - a deep learning approach. *CoRR*, abs/1703.10908, 2017.

[17] L. Younes. *Shapes and Diffeomorphisms*. Applied Mathematical Sciences. Springer Berlin Heidelberg, 2010.