Kernel Bundle Diffeomorphic Image Registration Using Stationary Velocity Fields and Wendland Basis Functions

Akshay Pai*, Stefan Sommer, Lauge Sørensen, Sune Darkner, Jon Sporring, and Mads Nielsen

***Abstract*—In this paper, we propose a multi-scale, multi-kernel shape, compactly supported kernel bundle framework for stationary velocity field-based image registration (Wendland kernel bundle stationary velocity field, wKB-SVF). We exploit the possibility of directly choosing kernels to construct a reproducing kernel Hilbert space (RKHS) instead of imposing it from a differential operator. The proposed framework allows us to minimize computational cost without sacrificing the theoretical foundations of SVF-based diffeomorphic registration. In order to recover deformations occurring at different scales, we use compactly supported Wendland kernels at multiple scales and orders to parameterize the velocity fields, and the framework allows simultaneous optimization over all scales. The performance of wKB-SVF is extensively compared to the 14 non-rigid registration algorithms presented in a recent comparison paper. On both MGH10 and CUMC12 datasets, the accuracy of wKB-SVF is improved when compared to other registration algorithms. In a disease-specific application for intra-subject registration, atrophy scores estimated using the proposed registration scheme separates the diagnostic groups of Alzheimer's and normal controls better than the state-of-the-art segmentation technique. Experimental results show that wKB-SVF is a robust, flexible registration framework that allows theoretically well-founded and computationally efficient multi-scale representation of deformations and is equally well-suited for both inter- and intra-subject image registration.**

***Index Terms*—Kernel bundle framework, registration, reproducing kernel hilbert spaces, wendland kernels.**

I. INTRODUCTION

In this paper, we propose a novel parameterization of stationary velocity fields for diffeomorphic registration using a class of multi-scale, multi-shape regularizing kernels called the Wendland kernels. The proposed registration scheme is termed wKB-SVF. We propose a framework that incorporates the best characteristics of state-of-the-art registration schemes: 1) we restrict the space of velocity fields to a specific class of function spaces, reproducing kernel Hilbert spaces (RKHS) [1]–[6]; 2) we parameterize the velocity fields using compactly supported reproducing kernels inherited from the RKHS structure, and we subsequently represent the high dimensional ODE via a smaller set of control points and vectors [7], [8]; 3) we provide a multi-scale representation of the velocity fields using the kernel bundle framework [9].

In the absence of validated models for inter-/intra-subject anatomical variability, deformations characterizing anatomical changes such as, change in organ growth, are generally assumed to be smooth and invertible. Three popular choices of diffeomorphic deformation models are: a) large deformation diffeomorphic metric mapping (LDDMM), b) freeform deformations, and c) stationary velocity fields (SVFs). Among them, a) and c) naturally generate diffeomorphisms and b) requires explicit regularization terms to ensure diffeomorphic transformations. For a discussion about commonly used constraints on deformation models, see [10].

Each of these methods involve finding an optimal diffeomorphism that connects two images. SVFs are less computationally expensive compared to LDDMM due to the stationary velocity field assumption. In this study, we will mainly focus on SVFs together with some key concepts from LDDMM because SVFs satisfy the dual goal of generating diffeomorphisms while keeping computational complexity low. A key feature of LDDMM is that the velocity fields are modeled on a Hilbert space. This space can be constructed using reproducing kernels, and this approach allows optimal solutions to specific optimization problems to be found as linear combinations of the reproducing kernels. In this paper, we model the stationary velocity fields on a Hilbert space constructed using a class of reproducing kernels called Wendland kernels. A key property of Wendland kernels is that they are of compact support. The construction reduces computational complexity because both the deformation field and the regularization term evaluate to zero outside the support of the kernel [6]. Existing parametric versions of SVFs [7], [8] use kernels where the evaluation of the energy term often requires spatial discretization (bending energy for instance); Wendland kernels require no such spatial discretization. In addition, we will use these kernels in a kernel bundle framework to provide a multi-scale modeling of the deformation field. Using the multi-scale feature from the kernel

---

Manuscript received September 10, 2015; revised December 07, 2015; accepted December 12, 2015. Date of publication December 23, 2015; date of current version May 28, 2016. This study was funded by the Danish Research Foundation (Den Danske Forskningsfond), the Danish National Advanced Technology Foundation (Højteknologifonden) and Biomediq A/S. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation. *Asterisk indicates corresponding author.*

*A. Pai is with the Department of Computer Science, University of Copenhagen, 2100 Copenhagen, Denmark (e-mail: akshay@biomediq.com).

S. Sommer, S. Darkner, and J. Sporring are with the Department of Computer Science, University of Copenhagen, 2100 Copenhagen, Denmark (e-mail: sommer@di.ku.dk).

L. Sørensen and M. Nielsen are with the Department of Computer Science, University of Copenhagen, 2100 Copenhagen, Denmark, and also with Biomediq A/S, 2100 Copenhagen, Denmark.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2015.2511062

0278-0062 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

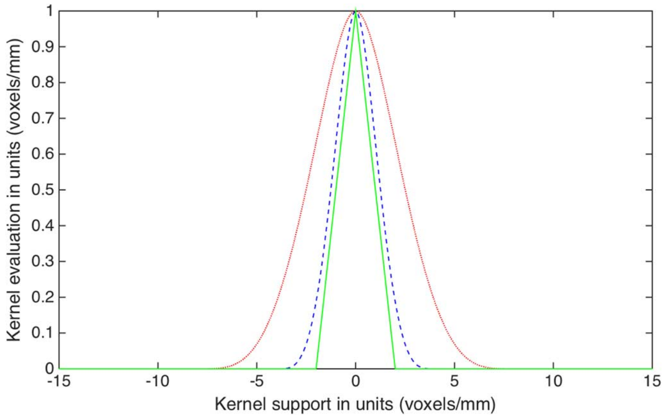

Fig. 1. A graphical representation of a 1-D Wendland kernel of different scales and orders. The bundle construction allows the use of kernels of multiple orders and scales simultaneously. We choose these particular kernels since they emulate B-Splines both in terms of the smoothness properties and the shape of the kernel.

bundle framework allows us to express the combination of various scales as a simple sum while still remaining in the Hilbert space.

The main contributions of this paper are as follows:

- We parameterize velocity fields using compactly supported norm-minimizing kernels, in particular Wendland kernels. We utilize the benefits of directly choosing reproducing kernels from the RKHS to parameterize the SVFs. Further, due to compact support, the regularization term which is a sum over kernel products, can be efficiently computed.
- We propose a kernel bundle framework where multiple kernels at different scales are combined to accurately model the deformation. Although we propose to use compactly supported kernels only, a combination of both kernels with infinite support and compactly supported kernels can also be used.

Fig. 1 illustrates the orders of kernels that, through the bundle construction, can be used *simultaneously* in wKB-SVF. In addition, the proposed framework allows the flexibility to simultaneously optimize for scales like in the LDDMM-based kernel bundle framework [9].

In terms of performance, wKB-SVF provide better overlaps than the registration methods considered in the study [11] on the publicly available MGH10 dataset. On the CUMC12 dataset, the proposed method provides significantly better overlaps than most of the registration methods, except for SPM-DARTEL where the difference was insignificant. The framework also separates diagnostic groups of Alzheimer's disease (AD) and normal controls (NC) better than the Freesurfer longitudinal segmentations when used to compute longitudinal atrophy scores. The results illustrate that the wKB-SVF is well suited for both inter- and intra-subject registration.

A. Background

In flow-based registration schemes, deformations are generated by integrating a smooth velocity field over time. Two prominent flow-based image registration frameworks are the LDDMM (time-varying ODEs) [2] and SVF (time-constant ODEs) [12], [13]. In LDDMM, the deformations can be parameterized by initial velocity fields (or their dual—momenta) and the resulting diffeomorphism paths are endpoints of the corresponding Riemannian geodesics. This particular setting is computationally expensive since it involves solving a geodesic equation on an infinite dimensional group. An alternative to LDDMM are the SVFs. Here, the diffeomorphisms are one-parameter subgroups parameterized by time-constant velocity fields through the Lie group exponential. The Lie group exponential is realized as a time-integration of the velocity field. The time-integration is usually approximated using integration schemes such as Euler's or scaling-and-squaring [12]. The generated diffeomorphism paths are geodesics with respect to the canonical Cartan connections [14]. The main drawback of SVFs is the lack of metric on space of diffeomorphisms which is important for performing statistics such as PCA [15] or regression [16]. SVFs were initially proposed by [12] and were further utilized with modifications in [7], [8], [13], [14], [17]. Among these, [13], [14], [18], [17] use the entire image space for dissimilarity minimization. In studies [7], [8], the velocity fields are instead parameterized by interpolating kernels like B-Splines. A thorough overview of existing registration schemes can be found in the study [10], [11]. In this paper, we will restrict our focus to flow-based registration schemes and specifically to the parameterization of the SVF.

In both SVF and LDDMM, the vector fields belong to a subspace V of square integrable functions, $\mathcal{L}^2$. The subspace V is generally completed using a Hilbert norm induced by a differential operator [19]. With sufficient conditions on the operator, the space is a RKHS [20]: the Riesz representation theorem states that every linear form arises as an inner product with the representer. The representer is the reproducing kernel. A linear form is an evaluational functional that provides a mapping of a vector to $\mathbb{R}$. In case of finite dimensional optimization problems, RKHSs allow evaluation of the optimal space in terms of the reproducing kernel itself: For instance, for an interpolation problem defined as to find $\mathbf{v} \in V$ of minimum norm that satisfies $v(\mathbf{x_i}) \approx \beta_i$, $\beta_i \in \mathbb{R}$, the solution can be expressed as the regularized optimization problem (Ref. theorem 9.7 [20]), $\|\mathbf{v}\|_V^2 + C\sum_1^D |\mathbf{v}(\mathbf{x}_i) - \beta_i|^2$. The solution then attains the form, $\hat{\mathbf{v}} = \sum_1^D K(\mathbf{x}, \mathbf{x}_i)\alpha_i$. Here, K is the reproducing kernel. The whole problem is thus reformulated to a finite dimensional optimization problem involving only the vectors α_i.

In contrast to the most common approach, we will take advantage of the fact that there is flexibility in choosing reproducing kernels directly as opposed to being imposed by an operator. This approach allows us to minimize computation through the use of compactly supported kernels. Alternate options to parameterizing velocity fields are by either using B-Splines or truncated Gaussians. The latter is no longer continuous and the evaluation of the energy term of the former (like bending energy) is an approximation since it depends on spatial discretization. Both B-Splines and truncated Gaussian kernels have negative Fourier transforms, which mean they are not semi-positive definite kernels and hence are not reproducing kernels. Due to the aforementioned reasons, they are not considered in this work. In [22], we have compared the application of Wendland kernels versus B-Splines to parameterize velocity field in the context of brain atrophy estimation.

Wendland kernels [21] (the choice of reproducing kernels in this paper) emulate B-Splines in both computational complexity and smoothness. In addition, they also provide the necessary mathematical properties (smooth, C^k for some k, norm-minimizing) to realize a diffeomorphic transformation model. The role of reproducing kernels and the corresponding regularization in the context of LDDMM has been explored in [3], [6], [9].

1) Multi-Scale Kernel-Based Image Registration: Inter-subject registration often requires smooth yet large deformations, whereas intra-subject registration requires deformations at much smaller scales. For example, anatomical changes in the hippocampus are often minute and changes in ventricles (cerebrospinal fluid) may require large deformations. It is desired that such deformations are recovered using the same transformation model. The type of the resulting deformation is restricted by the scale of the parameterizing kernel. Scale, in this case, can be interpreted as either the support of the kernel or the spacing between the control points. If the scale of the kernel is large, then matching of the larger structure may be good and the transformations smooth. However, the matching of smaller structures like the hippocampus may not be satisfactory. On the other hand, if the scale is small, the matching may be good but the resulting transformation is spiky and may lead to undesirably large Jacobians [3]. One way to handle such variability in deformation scales (also to avoid local minima's in optimization) is via a pyramidal approach, i.e., by changing the scales of image smoothing or the resolution of the control points. This approach, however, is still limited by the range of deformations achievable by the shape and size of the kernel. The kernel bundle framework handles this by providing a scale-space representation of the kernels. A very attractive feature of the kernel bundle framework is that the representation of the multi-scale kernel is a simple linear combination of kernels of different support or resolution. Standard parameterizations of velocity fields like ones using B-Splines require additional routines such as, knot splitting, to combine various scales of the velocity field. The idea of exploiting RKHS kernels to build multi-scale kernel-based diffeomorphic image registration is not new in the context of LDDMM [3], [9].

However, the application of kernel bundle framework in the context of SVF-based image registration is novel. Furthermore, a combination of compactly supported reproducing kernels and the kernel bundle framework has not been explored in diffeomorphic image registration.

B. Outline

We start by describing SVF-based image registration and presenting the application of RKHSs in the context of SVFs. We then discuss how computational complexity can be minimized by representing velocity fields with compactly supported reproducing kernels. Following this, we discuss the adaptation of the kernel bundle framework to SVFs together with compactly supported Wendland kernels. Next follows an evaluation of the registration performance on the MGH10 and CUMC12 datasets and an evaluation on the Alzheimer's disease neuroimaging (ADNI) dataset for atrophy scoring. We end the paper with a discussion of the results and concluding remarks. This paper is an extension of our previous work [22]. The key extension is the multi-scale representation using the kernel bundle framework and an extended validation of both intra- and inter- subject registration.

II. REGISTRATION

Given a floating image I_1 and a reference image I_2 with a spatial domain $\Omega \in \mathbb{R}^d$, image registration involves finding a transformation $\varphi : \Omega \times \mathbb{R} \to \Omega$ that aligns the images. The transformation is found by minimizing a dissimilarity measure between the images under certain constraints encoded in a regularization term. A general cost function is of the form:

$$\begin{aligned}&\arg\min_{\varphi}\left(E(I_1, I_2)\right)\\&\quad = \arg\min_{\varphi} E_D\left(I_1(\varphi), I_2\right) + \lambda E_R(\varphi) + \lambda_{\mathrm{ICC}} E_{\mathrm{ICC}}(\varphi) \qquad (1)\end{aligned}$$

where $\lambda, \lambda_{\mathrm{ICC}}$ are user-specified constants controlling the degree of regularization, E_D is a dissimilarity measure that allows comparison of the floating image to the reference image, E_R is a regularization term that encodes the desired properties of φ, and E_{ICC} can be included as an additional penalty term to enforce inverse consistency, see Section II-B. The regularization term can either be explicitly minimized as, in the parametric approach, or can be implicitly restricted by convolving with a low pass filter [13]. The transformation now is restricted to the group of diffeomorphisms $\mathrm{Diff}(\Omega)$. In flow-based schemes, a time-dependent velocity field $\mathbf{v}(\mathbf{x}, t) : \Omega \times \mathbb{R} \mapsto \mathbb{R}^d$ is integrated to obtain a displacement. The governing differential equation is of the form

$$\frac{\partial \varphi(\mathbf{x}, t)}{\partial t} = \mathbf{v}\left(\varphi(\mathbf{x}, t), t\right), \varphi(\mathbf{x}, 0) = \mathbf{x}$$

where φ is the displacement and $\int_0^T \|\mathbf{v}(\mathbf{x}, t)\|_{1,\infty} dt < \infty$ where $[0, T]$ is the time interval. The path of diffeomorphisms $\varphi(\cdot, t)$ is in practice obtained by numerical integration. Solving the non-time stationary differential equation is generally computationally expensive.

A. Stationary Velocity Fields

With stationary velocity fields (SVF) [14], the velocity field $\mathbf{v}(\mathbf{x}, t)$ is constant in time. The paths parameterized by SVFs are exactly one parameter subgroup of $\mathrm{Diff}(\Omega)$. These paths are quite different from the Riemannian geodesics in the sense that the paths are metric-free [14]. Let Ω be the spatial domain of I_1 with $\mathrm{x} \in \Omega$ as a spatial location. Let $G \subset \mathrm{Diff}(\Omega)$ be a subspace containing the diffeomorphic transformations parameterized by SVFs, and let V be the tangent space of G at identity Id containing the velocity fields $\mathbf{v}$. A path of diffeomorphisms is generated by the stationary flow equation,

$$\frac{\partial \varphi(\mathrm{x}, t)}{\partial t} = \mathbf{v}\left(\varphi(\mathrm{x}, t)\right) \qquad (2)$$

with initial condition $\varphi(x, 0) = \mathrm{x}$. The final transformation $\varphi(x) = \varphi(\mathrm{x}, 1)$ is the Lie group exponential map $\mathrm{Exp}(\mathbf{v})$. This Lie group exponential can be approximated by Euler integration

[7]. For example, given p steps and $\varphi^t = \varphi(\mathrm{x}, t)$, the Euler integration is

$$\varphi^0 = \mathrm{x}, \tag{3}$$

$$\varphi^{\frac{1}{p}} = \mathrm{x} + \frac{\mathbf{v}(\mathrm{x})}{p}, \tag{4}$$

$$\varphi^{t+\frac{1}{p}} = \varphi^t \circ \varphi^{\frac{1}{p}}. \tag{5}$$

In the study [12], the scaling-and-squaring method to exponentiate velocity fields was proposed. Here, the final deformation was estimated by composing successive exponentials. However, a major drawback to this method is that at every squaring step the velocity fields need to be re-interpolated at integer positions. This may lead to undesired smoothing (interpolation) in the velocity field over which there is no apparent control. Bossa *et al.*, [23] point out the instability in the convergence properties of scaling-and-squaring. Therefore, we choose the relatively stable forward Euler's scheme for integrating the velocity fields.

B. Inverse Consistency

In a continuous setting, diffeomorphisms generated by SVFs are invertible transformations with differentiable inverses. However, due to the numerical integration of the velocity field, inverse consistency is not achieved in practice and needs to be explicitly enforced, typically through a regularization term. In [24], inverse consistency was enforced by penalizing the displacement error generated after composing a transformation with its inverse. However, in this method, the computation of the inverse is a computationally expensive approximation [25]. Forward transformations are first computed and then the inverse transformations are approximated. In this study, we will maintain a single parameterization of the velocity field. Both the forward and backward registrations are performed simultaneously. The inverse consistency term is computed as,

$$E_{\mathrm{ICC}} = \left\|\varphi'\left(\varphi(\mathbf{x})\right) - \mathbf{x}\right\|^2. \tag{6}$$

where $\varphi = \varphi(\mathbf{x}, 1) = \mathrm{Exp}(\mathbf{v})$ and $\varphi' = \mathrm{Exp}(-\mathbf{v})$ are the backward and forward registration transformations.

III. REPRODUCING KERNELS AND THE KERNEL BUNDLE FRAMEWORK

In this section, we present the main contribution of the paper—multi-scale parametrization of velocity fields using compactly supported reproducing kernels. Unlike the parametrization of velocity fields using compactly supported kernels like B-Splines, Wendland kernels have the reproducing property that make the norm on these kernels minimizing.

A. Reproducing Kernels From Operators

In SVF-based image registration, the velocity fields $\mathbf{v}$, are chosen to belong to a subspace of absolutely integrable functions in $\mathcal{L}^2$. To complete this subspace, the norm associated with an appropriate differential operator L, $\langle u, v\rangle_V = \langle Lu, v\rangle_{\mathcal{L}^2}$, $u, v \in V$ is utilized. Usually, L is chosen to be a diffusive model of the form $L = \mathrm{Id} - \alpha\nabla^2$ [19] where ∇^2 is a Laplacian operator. Other choices for the operator exist and discussion of them can be found in [26].

The operator L provides a mapping of the velocity field $\mathbf{v}$ from V to its dual space V^*. When V is admissible [20], the dual space contains linear evaluational functionals $\delta_{\mathrm{x}} : \mathbf{v} \to \mathbf{v}(\mathbf{x})$. The evaluation functionals, that for each $\mathrm{x} \in \Omega$ provide a mapping of the vector space to $\mathbb{R}$, can be written as $(\delta_{\mathbf{x}}|\mathbf{v}) = \mathbf{v}(\mathbf{x})$. According to the Riesz representation theorem, there exists spatially dependent kernels $K_{\mathbf{x}} = K(\cdot, \mathbf{x}) : \Omega \times \Omega \to \mathbb{R}^{d\times d}$ such that $\mathbf{v}(\mathbf{x}) = \langle K_{\mathbf{x}}, \mathbf{v}\rangle_V = (\delta_{\mathbf{x}}|\mathbf{v})$. This implies that $\langle K_{\mathbf{x}}, \mathbf{v}\rangle_V = (LK_{\mathbf{x}}|\mathbf{v})$ and $LK_{\mathbf{x}} = \delta_{\mathbf{x}}$ [20]. Note that $\langle\cdot,\cdot\rangle$ denotes the inner product and $(\cdot|\cdot)$ denotes evaluation of a functional, i.e., $(\delta|v) = \delta(v)$ where $\delta \in V^*$ and $v \in V$. We can therefore view K as an inverse of L. In fact, the kernel is also a Green's function with respect to the differential operator L. If the operator is differential, then K is positive definite. As a consequence, if K is constructed from a differential operator, then K is always of infinite support [27]. It may be computationally intensive to evaluate the deformation field and the norm if velocity fields are parameterized using kernels of infinite support.

B. Compactly Supported Reproducing Kernels

The approach in the previous section essentially involves first finding a mapping from V to V^*, and then constructing kernels that provide a mapping back to V. In this paper, we will use the significant benefits in taking the reverse approach: instead of constructing kernels from differential operators which force the support of the kernels to be infinite, we choose the kernels directly. This particular arrangement allows the use of kernels to intentionally *minimize* computation via the compact support.

Following Moore-Aronszajn theorem [28]: for every symmetric positive (semi-) definite kernel K exists a unique RKHS that has K as its reproducing kernel. The corresponding RKHS is the completion of the linear space spanned by the functions of the form,

$$f(\mathbf{x}) = \sum_i K_{\mathbf{x_i}}\mathbf{a}_i,$$

for all choices of $\mathbf{a}_i \in \mathbb{R}$ which is the parameter attached to each kernel centered at the points $\mathbf{x}_i \in \Omega$. The inner product on this space provides the reproducing property such that,

$$\langle K(\cdot, \mathbf{x}), f(\cdot)\rangle_V = f(\cdot). \tag{7}$$

This essentially implies that we can choose an appropriate symmetric semi-positive definite kernel with compact support and this kernel has a unique RKHS associated with it. The Gaussian kernels are an example of reproducing kernels. However, due its infinite support, parameterizing velocity fields with Gaussians in dense image matching may be expensive. Using reproducing kernels to generate transformations is not new in LDDMM. Studies usually [3], [6], [9] utilize reproducing kernels such as Gaussian kernels to parameterize velocity fields.

C. Regularization

A regularization term is required to ensure sufficient smoothness in the solution of the ODE. In flow-based registration schemes, this term is usually formulated as the squared norm on the velocity field. Given a reproducing kernel, the evaluation of the squared norm is simply the kernel product and does not depend on any spatial discretization like other regularization terms such as bending energy of B-Splines. The regularization term may be evaluated as,

$$E_R(\varphi) = \|\mathbf{v}\|_V^2. \tag{8}$$

and by linearity of the inner product and the reproducing property associated with the corresponding RKHS (7), the norm on linear combinations of the kernel can be evaluated by

$$\begin{aligned}\left\|\sum_i K(\cdot,\mathbf{x}_i)\mathbf{a}_i\right\|^2 &= \left\langle \sum_i K(\cdot,\mathbf{x}_i)\mathbf{a}_i, \sum_j K(\cdot,\mathbf{x}_j)\mathbf{a}_j \right\rangle_V \\ &= \sum_{i,j} \langle K(\cdot,\mathbf{x}_i)\mathbf{a}_i, K(\cdot,\mathbf{x}_j)\mathbf{a}_j\rangle_V \\ &= \sum_{i,j} \mathbf{a}_i^T K(\mathbf{x}_i,\mathbf{x}_j)\mathbf{a}_j. \end{aligned} \tag{9}$$

Because of this reproducing property, it is often useful to parameterize the optimal function directly using the these kernels since optimal solutions are linear combinations of the reproducing kernels of the norm. In the case of infinitely supported kernels like the Gaussian, the entire double sum needs to be evaluated, which can be computationally expensive depending the number of kernels. In contrast, with finite support, the kernel product is zero outside the support making the evaluation of (9) efficient.

D. *Kernel Bundle Framework*

In this section, we will outline the second contribution of this paper. The contribution involves extending the kernel bundle concept [9] to compactly supported reproducing kernels and using it in the SVF framework.

The reasoning behind the need for a multi-scale representation of a deformation has been well discussed in previous works such as [3], [9]. In brief, image deformations often occur at different scales. For instance, in inter-subject registration, large-scale transformations may be required and in intra-subject registration relatively small-scaled deformations are required. For example, the deformations around the hippocampus can be small, while in regions like cerebrospinal fluid, the deformation may be larger. The key is to obtain computationally efficient representations of transformations without limiting the range and capacity of the deformation. This can be achieved by combining multiple kernels at multiple scales in the same registration framework. Typically, in a kernel-based image registration scheme, the support (or scale) of the kernel is fixed. For instance, in cubic-spline, the support is fixed to four by design.

The kernel bundle framework in LDDMM [9] incorporates multiple scales of kernels as a sum in the same optimization function. We utilize the fact that the sum of multiple RKHS spaces is still RKHS. We set out to achieve a similar construction with SVFs. We extend the concept of the space of velocity fields V to a family of spaces of velocity fields $\hat{V}$. We consider r spaces where each $\hat{V}_m$ is equipped with a norm $\|\mathbf{v}_m\|_{\hat{V}_m}$, $m = 1\ldots r$. The velocity fields are linear sums of individual kernels at r levels. It is represented as,

$$\mathbf{v}(\mathrm{x}) = \sum_{m=1}^{r} \mathbf{v}_m = \sum_{m=1}^{r}\sum_{i_m}^{N_m} K_m\left(\mathbf{x}_{i_m},\mathrm{x}\right)\mathbf{a}_{i_m}^m. \tag{10}$$

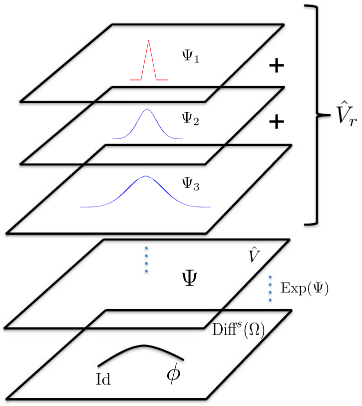

Fig. 2. Kernel bundle framework for SVFs. Different colors represent different classes of kernels.

Here, $\mathbf{v}$ is the final velocity field, K_m is the kernel at each level and $\mathbf{a}^m$ is the parameter associated with it. The variable i_m is the kernel centers at each level and N_m is the number of kernels at each level. Note that $K_m(\mathbf{x}_{i_m},\mathbf{x}) = K(\|\mathbf{x}_{\mathbf{i_m}} - \mathrm{x}\|_2/q_m)$, where q_m is the support of the kernel at each level. The expression of the cost function (1) in a kernel bundle framework can be written as illustrated in the case statement (11).

$$\begin{aligned}&\arg\min_{\mathbf{v}} E\left(I_1, I_2\left(\mathrm{Exp}(\mathbf{v}_1)\right) + E\left(I_2, I_1\left(\mathrm{Exp} - (\mathbf{v}_1)\right)\right),\right. \\ &\arg\min_{\mathbf{v}} E\left(I_1, I_2\left(\mathrm{Exp}(\mathbf{v}_1+\mathbf{v}_2)\right)\right) \\ &\quad + E\left(I_2, I_1\left(\mathrm{Exp} - (\mathbf{v}_1+\mathbf{v}_2)\right)\right), \\ &\vdots \\ &\arg\min_{\mathbf{v}} E\left(I_1, I_2\left(\mathrm{Exp}\left(\sum_{m=1}^{L}\mathbf{v}_m\right)\right)\right) \\ &\qquad + E\left(I_2, I_1\left(\mathrm{Exp}\left(\sum_{m=1}^{L} -\mathbf{v}_m\right)\right)\right)\end{aligned} \tag{11}$$

The kernels at each level can be of any support. For instance, one can have infinitely supported Gaussian kernels in a coarser registration scale and have compactly supported kernels handle finer resolutions in the registration. Fig. 2 illustrates the kernel bundle framework. Different colors represent different classes of kernels.

E. *Wendland Kernels*

In this section, we will describe the compactly supported Wendland kernel [21] used in parameterizing the velocity fields. Note that compactly supported reproducing kernels can also be constructed instead of choosing them directly. We direct the readers to [6] for one such example.

Wendland kernels were originally developed for multi-dimensional, scattered grid interpolation. They are positive definite functions with positive Fourier transforms and minimal degree polynomials on [0,1]. They yield C^{2s} (s is the desired degree of smoothness) smooth radial basis functions on $\mathbb{R}^d$. Application of Wendland kernels in landmark registration can be found in [29]. They are defined as follows,

$$K_{d,s}(r) = \mathcal{I}^s\left\{(1-r)_+^g\right\}, \tag{12}$$

where s is the smoothness of the kernel, r is the euclidean distance to the center of the kernel scaled by the support, g is the dimension given by $\lfloor d/2 \rfloor + s + 1$ and $\mathcal{I}^s$ is the integral operator applied s times given as,

$$\mathcal{I}\left\{K(r)\right\} = \int_{r}^{\infty} tK(t)dt. \tag{13}$$

We will utilize two classes of Wendland kernels in the kernel bundle framework. They are,

$$\begin{aligned} K_{1,0}(r) &= (1-r)_{+} \\ K_{3,1}(r) &= (1-r)_{+}^{4}(4r+1), \end{aligned} \tag{14}$$

and a plot is shown in Fig. 1. The plot shows both the linear (green) and C^2 smooth Wendland kernels (red and blue). They all have unit coefficient. Note that (12) refers to the general family of Wendland kernels. We will however choose only particular kernels ($d = 3$, $s = 1$) since they emulate B-Splines both in terms of the smoothness properties and the shape of the kernel.

IV. REGISTRATION COST FUNCTION AND MINIMIZATION

In this paper, we will optimize the kernel bundle framework in a hierarchical parallel fashion. For instance, on level one, only $\mathbf{v_1}$ is optimized, on level two both $\mathbf{v_1}, \mathbf{v_2}$ is optimized, so on and so forth. A limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) [30] scheme is used for optimization. The optimization was done using the min-Func package (see http://www.di.ens.fr/ mschmidt/Software/minFunc.html, version 2012). For optimization, we manually supply the derivatives. Normalized mutual information (NMI) [31] is used as a similarity measure for both the linear transformation and non-linear transformation. We initialize the non-linear transformation with a linear transformation with 9 degrees of freedom (DOF). We refer the reader to [32] for formulations of the linear transformations that we use. We can write the cost function as (15).

$$\begin{aligned} &\arg\min_{\mathbf{v}} \left(E(I_1, I_2)\right) \\ &\quad = \arg\min_{\mathbf{v}} E_D\left(I_1\left(\mathrm{Exp}(\mathbf{v})\right), I_2\right) + \lambda \sum_{1}^{m} \|\mathbf{v}_m\|_{\hat{V}_m}^{2} \\ &\qquad + \lambda_{\mathrm{ICC}} \left\|\varphi'\left(\varphi(\mathbf{x})\right) - \mathbf{x}\right\|^2, \\ &\quad = \arg\min_{\mathbf{v}} E_D\left(I_1\left(\mathrm{Exp}(\mathbf{v})\right), I_2\right) \\ &\qquad + \lambda \sum_{1}^{m} \sum_{i_m, j_m} {\mathbf{a}_{i_m}^{m}}^{T} K_m\left(\mathbf{x}_{i_m}, \mathbf{x}_{j_m}\right) \mathbf{a}_{j_m}^{m} \\ &\qquad + \lambda_{\mathrm{ICC}} \left\|\varphi'\left(\varphi(\mathbf{x})\right) - \mathbf{x}\right\|^2, \end{aligned} \tag{15}$$

Note that the forward and backward transformations are represented as $\varphi(\mathrm{x}, 1) = \mathrm{Exp}(\mathbf{v})$ and $\varphi'(\mathrm{x}, 1) = \mathrm{Exp}(-\mathbf{v})$. The derivative of the cost function at each level can be derived as (16).

$$\begin{aligned} \nabla_{\mathbf{a^m}} E = {}& \frac{\partial E_D}{\partial u} \frac{\partial u}{\partial \mathbf{a}^m} + \frac{\partial}{\partial \mathbf{a}^m} \lambda \sum_{1}^{m} \sum_{i_m, j_m} {\mathbf{a}_{\mathbf{i_m}}^{\mathbf{m}}}^{T} K_m \\ & \times \left(\mathbf{x}_{i_m}, \mathbf{x}_{j_m}\right) \mathbf{a}_{\mathbf{j_m}}^{\mathbf{m}} + \frac{\partial}{\partial \mathbf{a}^m} \lambda_{\mathrm{ICC}} \left\|\varphi'\left(\varphi(\mathbf{x})\right) - \mathbf{x}\right\|^2, \end{aligned} \tag{16}$$

$$\frac{\partial u_e}{\partial \mathbf{a}_n^m} = \left(Id + D\varphi_{e-1}\delta c\right) \frac{\partial \varphi_{e-1}}{\partial \mathbf{a}_n^m} + \frac{\partial \mathbf{v}_{e-1}}{\partial \mathbf{a}_n^m} \delta c. \tag{17}$$

In both (15) and (16), the backward registration is computed in a similar way however, by replacing $\mathbf{v}$ by $-\mathbf{v}$. In (16), u is the intermediate steps of the Euler integral as explained in (3), $\partial u_e / \partial \mathbf{a}_n^m$ is the derivative of the eth composition with respect to the nth parameter at mth level, $e = t + (1/p)$ is the composition number (p, total number of Euler compositions), $D\varphi_{e-1}$ is the spatial Jacobian of the previous composition, δc is $1/p$, and $\partial \mathbf{v}_{e-1} / \partial \mathbf{a}_n^m$ is nothing but the kernel evaluation since $\mathbf{v}$ is linear in $\mathbf{a}$. Furthermore, (16) represents the gradient of the cost function with respect to the parameters on the mth level. Note that E_D is chosen to be NMI. The flow chart of the registration algorithm can be found in Algorithm 1. The parameter used for the registration can be found in Table I. For computational reasons, only every second voxel was used to evaluate the similarity measure. Note that for both inter-subject and intra-subject registration, the same set of parameters are used. Note that all the r levels are optimized simultaneously, depending on the level.

Algorithm 1 wKB-SVF Registration Algorithm

```
1: Affine registration with 12 degree's of freedom.
2: Non-rigid registration
       Initialization, $\mathbf{a}^m = 0, m = 1 \ldots r$
3: loop over the number of levels r
       Smooth both floating and fixed image with a
       Gaussian of standard deviation $\sigma_f$.
       Compute velocity field $\mathbf{v}$ (10).
       Compute displacement field $\varphi = \mathrm{Exp}(\mathbf{v})$.
       Compute similarity measure NMI.
       Compute the update $\nabla_{\mathbf{a^m}} E$ (16).
```

V. EXPERIMENTS

In this section, we will present the experiments conducted to evaluate the performance of the presented registration method on the paradigms of inter-subject registration and intra-subject registration. In the first experiment, we will evaluate the performance of inter-subject registration on a publicly available MGH10 and CUMC12 evaluation dataset [11]. In the second experiment, we will evaluate the performance of intra-subject registration on the publicly available Alzheimer's disease neuroimaging initiative (ADNI) dataset [33] by measuring the ability to separate the diagnostic groups of Alzheimer's disease (AD) and normal controls (NC) based on atrophy scores.

A. Inter-Subject Registration

1) Data Description:

- MGH10: 10 subjects were scanned at the MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging using a 3 T Siemens scanner and standard head coil. The data were inhomogeneity corrected, affine-registered to

TABLE I
VARIOUS PARAMETERS USED DURING THE REGISTRATION. THE CONTROL POINT SPACING AND THE IMAGE SMOOTHING ARE STATED AS A SINGLE NUMBER REPRESENTING ALL THREE DIMENSIONS ON EACH KERNEL BUNDLE LEVEL

Parameter	**Value**
Rigid	
Degrees of freedom	12
Number of intensity bins in joint histogram	64
Evaluation points (Rigid)	Every 2^{rd} voxel
Image smoothing (Rigid) (σ in mm)	15, 8, 2
Non-Rigid	
Number of intensity bins in joint histogram	64
Evaluation points (non-rigid)	Every 2^{nd} voxel
Control point spacing	$6, 5, 4$
Kernel support (q)	$8, 4, 2$
Regularization constant (non-rigid), λ)	0.3
Regularization constant inverse consistency (non-rigid, λ_{icc})	0.01
Number of compositions (p)	16
Image smoothing (non-rigid) (σ_f in mm)	0.2, 0.2, 0.2
Kernel orders	3, 3, 1 (Eq (14))

TABLE II
BASELINE DEMOGRAPHICS (AGE, GENDER AND MINI MENTAL STATE EXAMINATION)

	Group	
	NC (n=169)	AD (n=101)
Age, yr $\pm\sigma$	76.0 $\pm$ 5.1	75.3 $\pm$ 7.4
Male (%)	50.9	50.5
MMSE $\pm\sigma$	29.2 $\pm$ 1.0	23.2 $\pm$ 1.9

TABLE III
MEAN AND STANDARD DEVIATION OF THE OVERLAPS FOR OTHER SVF-BASED METHODS AND WKB-SVF ON THE MGH10 DATASET

Methods	Target Overlap	Union Overlap
wKB-SVF	0.5844 (0.0329)	0.4134 (0.0313)
Demons	0.5228 (0.0286)	0.3497 (0.0236)
SPM-D	0.5431 (0.0547)	0.3654 (0.0426)

the MNI152 template [34], and segmented using SPM2 software [35]. The images were manually labeled by Tourville of Boston University using Ghosh's ASAP software [36]; the labeling protocol [37] is similar to Cardviews, and in the version used for this study produces 74 labeled regions.

- CUMC12: This dataset consists of 12 subjects that were scanned at the Columbia University Medical Center and manually labeled by one technician trained according to the Cardviews labeling scheme created at the CMA. The images have 128 labeled regions.

2) Experiments: We performed an evaluation on both the evaluation dataset similar to [11]. Here, the ability to match a set of manually segmented regions of interest via pair-wise registration is evaluated. The parameters used in the registration can be found in Table I.

Once the registrations are performed, the manually labeled segmentation from the floating image are warped to the reference image using a nearest neighbor interpolation. Following this, the accuracy of the overlap for each anatomical region is assessed using the following measures [11],

$$\mathrm{TO}_l = \frac{|S(\phi)_l \cap T_l|}{|T_l|}, \tag{18}$$

$$\mathrm{UO}_l = \frac{|S(\phi)_l \cap T_l|}{|S(\phi)_l \cup T_l|}. \tag{19}$$

where S is the source segmentation, ϕ is the transformation, T is the target segmentation, TO is the target overlap, UO is the union overlap and l is the index of the anatomical region. Since these measures were used in the Klein *et al.*, we also use the same metrics in this article. Furthermore, for assessing the impact of kernel bundle framework, we perform a similar experiment but by just maintaining a single fine layer of SVF (control point resolution of 4 mm).

B. Evaluation of Intra-Subject Registration

In order to evaluate the registration on intra-subject registration, we measure atrophy (or volume change) in (Alzheimer's) disease-specific brain regions such as whole brain (WB), Hippocampus (Hip), Ventricles (Vent) and Medial Temporal Lobe (MTL).

1) Atrophy Estimation: Usually, given a deformation field and an anatomical mask, regional atrophy is estimated by summing the Jacobian determinant over the region of interest (ROI). However, for registration schemes where the analytical expression of the transformation is not available, the Jacobian determinant needs to be approximated using finite differencing schemes [14]. In this paper, we will instead utilize Cube Propagation (CP) to measure atrophy. Here, each face of a cubic voxel is triangulated and the volume under each triangle after transformation is summed to get the volume of the transformed cube. Tetrahedral meshing is also similar in terms of both numerical precision and meshing, however with CP one needs to triangulate the surface, which is simpler in terms of bookkeeping of the indices and computations. A detailed description of CP can be found in [32].

2) Data: Data used in the preparation of this article were obtained from the ADNI database (adni.loni.ucla.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging, positron emission tomography (PET), other biological markers, together with clinical data and neuropsychological assessments can be combined to measure the progression of mild cognitive impairment and early Alzheimers disease. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials, see http://www.adni-info.org/. The description of the ADNI standardized dataset used here can be found in [33]. We analyzed baseline and 12-month follow up 1.5 T1 weighted MRI volumes (169 AD and 101 NC). The demographics of the subjects can be found in Table II.

The raw DICOM images were preprocessed using Freesurfer's [38] intensity correction method, following which

TABLE IV
MEAN AND STANDARD DEVIATION OF THE OVERLAPS FOR OTHER SVF-BASED METHODS AND WKB-SVF ON THE CUMC12 DATASET

Methods	Target Overlap	Union Overlap
wKB-SVF	0.5203 (0.0341)	0.3557 (0.0284)
Demons	0.4623 (0.0286)	0.3054 (0.0233)
SPM-D	0.5216 (0.0567)	0.3571 (0.0439)

the images were resampled as isotropic voxels with a dimension of 256^3 with each voxel of size 1 mm^3. Segmentations used for the quantification of atrophy were also obtained using Freesurfer (cross-sectional). We add a routine of differential bias correction [39] to take care of intensity bias due to the scanners.

3) Metrics for Evaluation: To characterize the performance of the registration schemes, we record the mean and standard deviation of atrophy in the diagnostic groups as well as compute the diagnostic capability in terms of effect size measured as the Cohen's D and the area under the receiver operator characteristic curve (AUC). Cohen's D—disease-specific effect size is computed as,

$$\text{Cohen's D} = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2 + s_2^2}{2}}},$$

where m_1, s_1 and m_2, s_2 are the means and standard deviations of NC and AD subjects respectively. Power analysis is used to estimate the sample size required to detect a 25% reduction of mean atrophy in the AD group, using a two-sided test and a significance level of 0.05 [40]. The sample size is computed as follows,

$$\text{sample size} = \frac{2\sigma_d^2 \left(z_{1-\frac{\alpha}{2}} - z_{power}\right)^2}{(0.25\beta)^2} \tag{20}$$

where σ_d is the group standard deviation in atrophy, z is the quantile of a normal distribution at $\alpha = 0.05$ and power of 80% and β is the mean atrophy across the subjects involved in the study. The 95% confidence interval for the sample size was computed based on bootstrap resampling with 10,000 samples, with a bias corrected and accelerated percentile method [40], [41]. To compute the p-value for the pairwise method comparison, we carried out a two-tailed t-test for the null hypothesis of equal measures $N_1 - N_2 = 0$, where N_1 and N_2 are independent random measures. We compute a probability distribution for the difference between the Cohen's D for the two measures and compute p as $p(N_1 > N_2) = 1 - \text{cdf}_{N_1 - N_2}(0)$ and $p(N_2 > N_1) = \text{cdf}_{N_1 - N_2}(0)$ [42]. The p-values for comparing the AUCs were computed using the DeLong test [43]. Note that in this study we presented numbers based on the full data set.

VI. RESULTS

A. Inter-Subject Registration

The resulting overlaps on the MGH10 and CUMC12 are freely available to compare against.[1] The significance between methods was computed using a two-sample t-test. On MGH10, all the 14 registration algorithms have significantly lower overlaps (although FLIRT is only linear). The mean target overlap of wKB-SVF on the MGH10 dataset was 0.5844 when compared to 0.5431 of SPM-DARTEL, 0.5228 of Demons, 0.5683 of SyN, 0.5611 of ART and 0.5490 of IRTK. Similarly, the mean union overlap of wKB-SVF was 0.4134 when compared to 0.3654 of SPM-DARTEL, 0.3497 of Demons, 0.3946 of SyN and 0.3793 of IRTK. Figs. 5 and 6 present the target and overlaps evaluated on the MGH10 dataset. On CUMC12 dataset, wKB-SVF had a mean target overlap of 0.5203 when compared to 0.5216 of SPM-DARTEL, 0.4632 of Demons, 0.5138 of SyN, and 0.5150 of IRTK. Similarly, the mean union overlap of wKB-SVF was 0.3557 when compared to 0.3571 of SPM-DARTEL, 0.3054 of Demons, 0.3519 of SyN and 0.3542 of IRTK. Figs. 7 and 8 present the target and overlaps evaluated on the CUMC12 dataset. Figs. 3 and 4 illustrate the overlaps when the kernel bundle layer is replaced by a single layer of Wendland SVF.

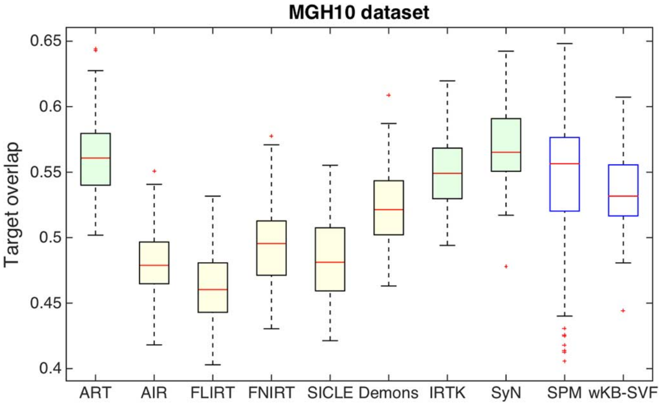

Fig. 3. Box plot of target overlaps of other registration schemes. Yellow: Significantly lower overlaps than single layer Wendland SVF. Green: Significantly higher overlaps than single layer Wendland SVF.

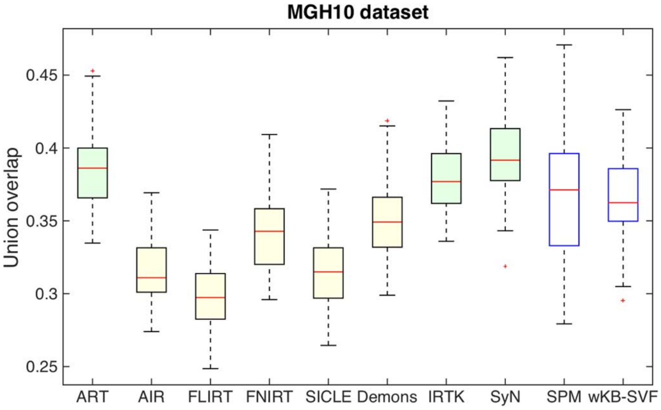

Fig. 4. Box plot of union overlaps of other registration schemes in comparison to single layer SVF and Wendland kernels. Yellow: Significantly lower overlaps than Wendland based method. Green: Significantly higher overlaps than single layer Wendland SVF.

B. Intra-Subject Registration

Table V presents the mean and standard deviation of the atrophy measured over whole brain (WB), hippocampus (Hip), ventricles (Vent) and medial temporal lobe (MTL) using the

---

[1] www.mindboggle.info

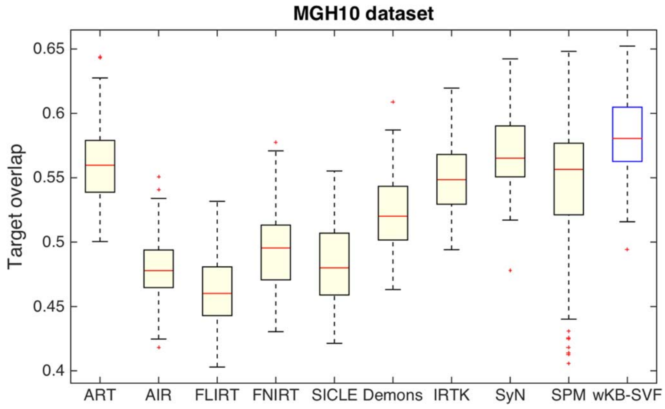

Fig. 5. Box plot of target overlaps of other registration schemes compared to wKB-SVF on the MGH10 dataset. Yellow: Significantly lower overlaps than wKB-SVF.

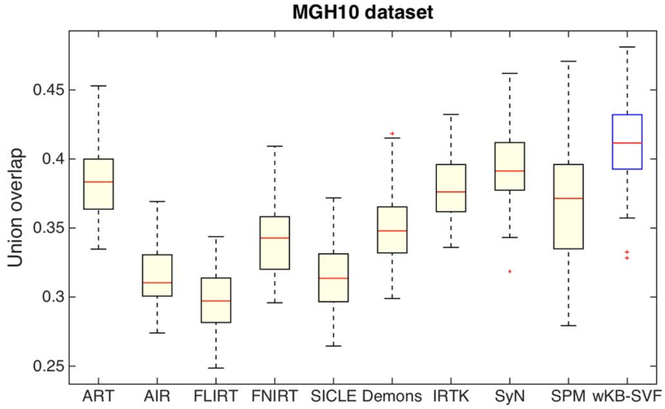

Fig. 6. Box plot of union overlaps of other registration schemes compared to wKB-SVF on the MGH10. Yellow: Significantly lower overlaps than wKB-SVF.

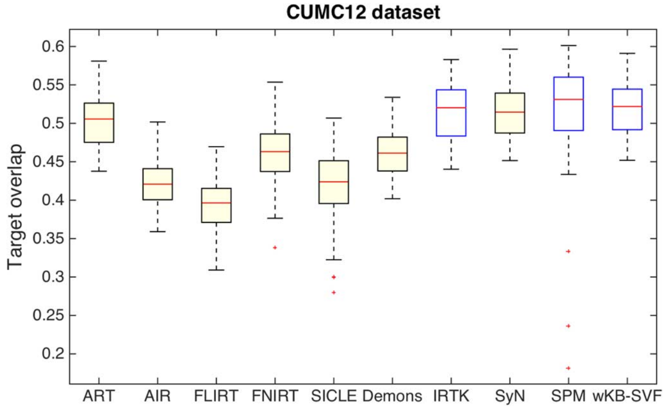

Fig. 7. Box plot of target overlaps of other registration schemes compared to wKB-SVF on the CUMC12 dataset. Yellow: Significantly lower overlaps than wKB-SVF.

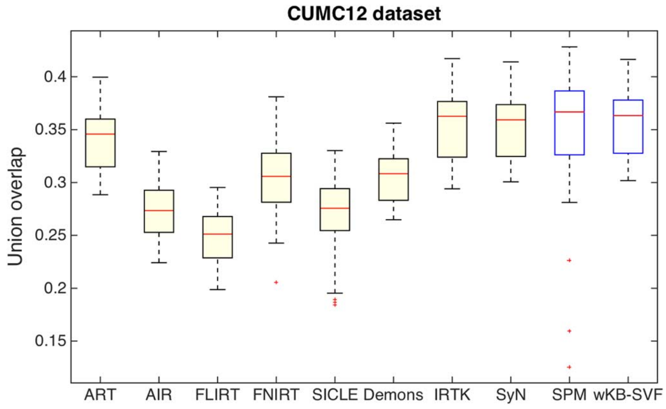

Fig. 8. Box plot of union overlaps of other registration schemes compared to wKB-SVF on the CUMC12. Yellow: Significantly lower overlaps than wKB-SVF.

volume change maps from Wendland kernel bundle registration and the longitudinal version of Freesurfer. The Cohen's D, Sample Size and AUCs are generally on the higher side for wKB-SVF measures. It is interesting to note that the mean atrophies for wKB-SVF method are on the lower side of Freesurfer. However, due to smaller standard deviations, the signal-to-noise ratio for the disease is still significant. The annual mean atrophy rate in the hippocampus is measured to be approximately 3% which is consistent with what was reported in a previous study [44]. As a consequence of atrophy in WB and other structures, we observe an expansion in the ventricles of ≈9%. Sample size is not adjusted with NC and is disease-specific only. However, the effect size after accounting for controls is reflected in the Cohen's D scores. The ventricles are clearly demarcated structures in the brain and hence both segmentation and registration of this structure is relatively easier. This is also reflected in our observation. Both Cohen's D and sample size for both the methods are quite similar. A head-to-head comparison showed that the wKB-SVF scores on the hippocampus and WB were significantly better than the respective Freesurfer longitudinal scores.

VII. DISCUSSION

In this paper, we presented a novel approach to non-rigid registration using SVFs by parameterizing velocity fields using compactly supported reproducing kernels. Furthermore, we proposed a multi-scale representation of the deformations by using the kernel bundle framework. We evaluated the inter-subject registration performance of the framework on the publicly available MGH10 and CUMC12 datasets. wKB-SVF had better target and union overlaps than the 14 registration methods on the MGH10 dataset. On the CUMC12 dataset, wKB-SVF had higher overlaps compared to most of the registration methods apart from SyN, ITK and SPM-DARTEL where the difference was insignificant. The performance in intra-subject registration was evaluated by estimating the diagnostic group separation capabilities of both longitudinal Freesurfer and wKB-SVF framework. It was shown that wKB-SVF separates AD from NC better than Freesurfer. Both Cohen's D and sample size were significantly better for wKB-SVF. To assess the impact of the optimization scheme, we switched to only using sequential optimization on the MGH10 dataset and obtained lower target (0.5431) and union overlaps (0.3896) indicating the positive impact of the parallel scheme. In terms of computational time, the proposed method runs a full registration in 180 mins (single core, 2.5 Ghz Xeon processor). This is comparable and in a few cases faster than registration schemes like ANTs and

TABLE V

STATISTICS BASED ON ATROPHY ESTIMATED IN DIFFERENT BRAIN STRUCTURES USING THE PROPOSED KERNEL BUNDLE FRAMEWORK AND LONGITUDINAL FREESURFER, WB: WHOLE BRAIN, HIP: HIPPOCAMPUS, VENT: VENTRICLES, MTL: MEDIAL TEMPORAL LOBE, MEAN AND STANDARD DEVIATION ARE ATROPHY IN PERCENT

		AD	**NC**			
		mean (±std)	mean (±std)	AUC (±std)	Sample Size	Cohen's D
wKB-SVF	WB	-1.22(0.52)	-0.52(0.76)	0.76(0.03)	101 [66,181]	0.92
	Hip	-3.22(1.94)	-1.58(1.74)	0.82(0.03)	92 [66,134]	1.24
	Vent	9.27(6.15)	3.71(3.88)	0.81(0.03)	111 [82,163]	1.07
	MTL	-3.08(1.68)	-0.86(1.40)	0.86(0.03)	75 [54,113]	1.44
FSLong	WB	-1.48 (1.26)	-0.71 (1.15)	0.68 (0.03)	186 [124,308]	0.64
	Hip	-4.81 (3.75)	-1.74 (2.84)	0.75 (0.03)	154 [109,230]	0.92
	Vent	9.73 (6.49)	4.09 (4.46)	0.79 (0.03)	113 [79,182]	1.01
	MTL	-4.61 (2.94)	-1.32 (2.02)	0.82 (0.03)	103 [77,146]	1.31

IRTK. Demons runtime is lower than the presented registration scheme, however at the cost of lower segmentation overlaps. For more accurate registrations, a higher number of kernel bundle scales can be used however at a higher computational cost.

When differential operators are used to construct an RKHS, the associated reproducing kernel is of infinite support. In [45], a similar approach as in [2] was used to solve the differential equation with the space of velocity fields constructed using a Laplacian-like differential operator. Instead of parameterizing the velocity fields with K, the authors directly compute the *discretized* version Lu, u being the vector field and L the differential operator, in the Fourier domain. For Laplacian-like operators, it was shown that for certain cases the inverse of the operator equals a Gaussian kernel and, for most other cases, the Gaussian kernel is a good approximation [46]–[48]. This observation forms the basis for the Demons framework. In order to make the registration computationally efficient, a smoothing of the velocity field update is performed during optimization. There is a theoretical equivalence of the smoothing procedure to a gradient descent [49], [50].

The kernel provides a mapping from V^* to V that is equivalent to the mapping that results from the inverse of a differential operator. Classical operators like Laplacian give rise to kernels that resemble Gaussians. The native spaces of Wendland kernels are norm-equivalent to a Sobolev space of a particular order depending on the kernel chosen. The regularization effect of the norm on Wendland kernels has an equivalent effect to that of norms induced by Sobolev operators. For instance, $K_{3,1}$ (14) has a norm-equivalence to a Sobolev space $\mathcal{H}^3$. The order usually represents the differentiability of the operator. Equivalence of smoothness of the spatial transformation follows from this. The inverse consistency regularization is due to the numerics involved in the integration scheme. Taking the term into account reduces the inverse consistency error; for instance, the mean error was 1.5 mm when the inverse consistency regularization was turned off and 0.1 mm when the inverse consistency term was included.

We also extended the concept of the kernel bundle framework to SVFs and compactly supported kernels. Multi-scale representation of kernels has been previously explored in the context of LDDMM [3], [9] and freeform deformation [51]. The foundations of the sum of kernels approach is well established [3], [4]. Instead of using a single diffeomorphism group for image registration, a combination of several subgroups via a semi-direct product can be utilized. Each subgroup can parameterized by a kernel with its own characteristic shape and scale. This fact was extensively exploited by [3], [9]. In both [3], [9], a sum of Gaussian kernels with varying scales was used. Similar to the latter, it is possible to add an L^1 prior (similar to [9], [51]) to the wKB-SVF framework, adding to its flexibility. In fact, a varying combination of kernel scale and resolution of the control points is incorporated in the proposed registration framework. Also, the compactly supported kernel bundle framework can be adapted to LDDMM, thereby enabling computationally efficient integration of the LDDMM flow equations.

Another possible extension to wKB-SVF will be to utilize the full multi-grid methods [7], [52] for registration optimization. It is interesting to note that, if one uses the current standard B-Splines multi-grid approach, an additional routine like *knot splitting* needs to be included to project the coarser resolution of control points onto a finer one. Since we utilize the additive nature of the vector spaces, such an additional routine is not required. Since both the approaches (standard coarse-to-fine approach and kernel bundle) have redundant information across scales, it will be interesting to check the influence of having orthogonality between the scales. By choosing basis functions as appropriate wavelets [53], one may introduce orthogonality with respect to the inner product between scales and also avail a multi-scale feature. It will be interesting to compare the performance of such a representation with the kernel bundle framework. It has to be noted that when using wavelets, the velocity fields need to be defined on a regular grid.

A very important caveat in the comparison is the similarity measure used. The presented framework used mutual information whereas demons used sum-of-square differences and SPM-DARTEL uses congealing. In the experiments performed, we illustrated the positive impact of multi-scale representation of the velocity field. The degree of impact of the similarity measure used remains uncertain to some extent. Therefore, it is pertinent that a comparison of transformation models needs to be performed by using the same similarity measure. Another aspect of the presented image registration framework is that the parameters of the registration were not optimized. Most importantly, we use the same set of parameters for both inter- and intra-subject registration. Usually, a different set of registration parameters are used for inter-subject and intra-subject image registration. It will interesting to evaluate the robustness of the results of wKB-SVF under varying parameters. Another aspect that needs to be determined is the practical impact of Wendland kernels versus compactly supported kernels such as B-Splines or truncated Gaussian in the kernel bundle framework.

VIII. CONCLUSIONS

The nonlinear registration method, wKB-SVF, presented in the paper encapsulates the benefits from existing state-of-the-art methods. The method exploits the benefits of flow-based registration methods through the use of RKHS and SVFs. We incorporate computational efficiency through the utility of compactly supported kernels and, finally, we represent the deformation through a multi-scale kernel bundle framework. The method allows us to model precisely both inter- and intra-subject registration at several scales that are compactly represented by reproducing kernels. We demonstrated the efficiency of this method by showing that it performs better than the state-of-the-art SVF methods on a publicly available dataset on inter-subject registration. On intra-subject registration, we show that the atrophy estimated is on par or better than a state-of-the-art segmentation method—longitudinal Freesurfer.

In addition to the presented experiment, we expect that the kernel bundle framework in tandem with SVFs will be particularly powerful when applied to population analysis. For instance, the presented method can be efficiently used in estimating diffeomorphic spatio-temporal atlases for human anatomy.

ACKNOWLEDGMENT

We would like to thank Stefan Klein, Biomedical Imaging Group Rotterdam, Erasmus MC, Rotterdam, Netherlands for constructive discussions. We gratefully acknowledge the Laboratory for Computational Neuroimaging at the Athinoula A. Martinos Center for Biomedical Imaging for providing the software (FreeSurfer) used for the segmentations in this paper. We also gratefully acknowledge the Danish Research Foundation (Den Danske Forskningsfond), the Danish National Advanced Technology Foundation (Højteknologifonden) and Biomediq A/S for funding the projects "Early MRI diagnosis of Alzheimer's Disease" and "Image Based Quantification of Anatomical Change".

ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California.

REFERENCES

[1] M. Miller, "Computational anatomy: Shape, growth, atrophy comparison via diffeomorphisms," *NeuroImage*, vol. 23, no. 1, pp. 19–33, 2004.

[2] M. Beg, M. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms.," *Int. J. Comput. Vis.*, vol. 61, pp. 139–157, 2005.

[3] L. Risser *et al.*, "Simultaneous multi-scale registration using large deformation diffeomorphic metric mapping," *IEEE Trans. Med. Imag.*, vol. 30, no. 10, pp. 1746–1759, Oct. 2011.

[4] M. Bruveris, F. Gay-Balmaz, D. D. Holm, and T. S. Ratiu, "The momentum map representation of images," *J. Nonlinear Sci.*, vol. 21, no. 1, pp. 115–150, 2011.

[5] M. De Craene *et al.*, "Temporal diffeomorphic free-form deformation: Application to motion and strain estimation from 3D echocardiography," *Med. Image Anal.*, vol. 16, pp. 427–50, 2012.

[6] A. Jain and L. Younes, "A kernel class allowing for fast computations in shape spaces induced by diffeomorphisms," *J. Comput. Appl. Math.*, vol. 245, pp. 162–181, 2013.

[7] J. Ashburner, "A fast diffeomorphic image registration algorithm," *NeuroImage*, vol. 38, no. 1, pp. 95–113, 2007.

[8] M. Modat, P. Daga, M. J. Cardoso, and S. Ourselin, "Parametric nonrigid registration using a stationary velocity field," in *IEEE Workshop Math. Methods Biomed. Image Anal.*, 2012, pp. 145–150.

[9] S. Sommer, F. Lauze, M. Nielsen, and X. Pennec, "Sparse multi-scale diffeomorphic registration: The kernel bundle framework," *J. Math. Imag. Vis.*, vol. 46, no. 3, pp. 292–308, 2012.

[10] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1153–1190, Jul. 2013.

[11] A. Klein *et al.*, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *NeuroImage*, pp. 786–802, 2009.

[12] V. Arsigny, O. Commowick, X. Pennec, and N. Ayache, "A log-euclidean framework for statistics on diffeomorphisms," in *MICCAI*, 2006, vol. 4190, pp. 924–931.

[13] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Non-parametric diffeomorphic image registration with the demons algorithm," in *MICCAI*, 2007, vol. 4792, pp. 319–326.

[14] M. Lorenzi and X. Pennec, "Geodesics, parallel transport and one-parameter subgroups for diffeomorphic image registration," *Int. J. Comput. Vis.*, vol. 105, pp. 111–127, 2013.

[15] V. M, M. MI, Y. L, and T. A. , "Statistics on diffeomorphisms via tangent space representations," *NeuroImage*, vol. Suppl 1, no. 23, pp. S161–S169, 2004.

[16] N. Singh *et al.*, "Multivariate statistical analysis of deformation momenta relating anatomical shape to neuropsychological measures," *MICCAI*, pp. 529–537, 2010.

[17] M. Bossa, E. Zacur, and S. Olmos, "Tensor-based morphometry with stationary velocity field diffeomorphic registration: Application to ADNI," *NeuroImage*, vol. 51, no. 3, pp. 956–969, 2010.

[18] T. Mansi, X. Pennec, M. Sermesant, H. Delingette, and N. Ayache, "iLogDemons: A demons-based registration algorithm for tracking incompressible elastic biological tissues," *Int. J. Comput. Vis*, vol. 92, no. 1, pp. 92–111, 2011.

[19] M. Hernandez, M. N. Bossa, and S. Olmos, "Registration of anatomical images using paths of diffeomorphisms parameterized with stationary vector field flows," *Int. J. Comput. Vis.*, vol. 85, no. 3, pp. 291–306, 2009.

[20] L. Younes*, Shapes and Diffeomorphisms*. New York: Springer, 2010, vol. 171.

[21] H. Wendland, "Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree," *Adv. Comput. Math.*, vol. 4, no. 1, pp. 389–396, 1995.

[22] A. Pai *et al.*, "Image registration using stationary velocity fields parameterized by norm-minimizing wendland kernel," in *SPIE Med. Imag.*, 2015.

[23] M. Bossa, E. Zacur, and S. Olmos, "Algorithms for computing the group exponential of diffeomorphisms: Performance evaluation," in *Comput. Vis. Pattern Recognit. Workshops*, 2008, pp. 1–8.

[24] G. E. Christensen and H. J. Johnson, "Consistent image registration," *IEEE Trans. Med. Imag.*, vol. 20, no. 7, pp. 568–582, Jul. 2001.

[25] A. Leow *et al.*, "Inverse consistent mapping in 3d deformable image registration: Its construction and statistical properties," in *Inf. Process. Med. Imag.*, 2005, vol. 19, pp. 493–503.

[26] J. Modersitzki, *Numerical Methods for Image Registration*. Oxford, U.K.: Oxford Univ. Press, 2004.

[27] Q. Ye, Reproducing kernels of generalized Sobolev spaces via a green function approach with differential operators Illinois Inst. Technol., Tech. Rep., 2010.

[28] N. Aronszajn, "Theory of reproducing kernels," *Trans. Am. Math. Soc.*, vol. 68, no. 3, pp. 337–404, 1950.

[29] M. Fornefett, K. Rohr, and H. Stiehl, "Radial basis functions with compact support for elastic registration of medical images," *Image Vis. Comput.*, vol. 19, no. 1, pp. 87–96, 2001.

[30] D. C. Liu, J. Nocedal, D. C. Liu, and J. Nocedal, "On the limited memory BFGs method for large scale optimization," *Math. Programm.*, vol. 45, pp. 503–528, 1989.

[31] S. Darkner and J. Sporring, "Locally orderless registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1437–1450, Jun. 2013.

[32] A. Pai *et al.*, "Cube propagation for focal brain atrophy estimation," in *IEEE Symp. Biomed. Imag.*, 2013.

[33] B. Wyman *et al.*, "Standardization of analysis sets for reporting results from ADNI MRI data," *Alzheimer's Dementia*, Oct. 2012.

[34] A. C. Evans, D. L. Collins, and B. Milner, "An MRI-based stereotactic brain atlas from 300 young normal subjects," in *Proc. 22nd Symp. Soc. Neurosci.*, 1992, p. 408.

[35] K. J. Friston *et al.*, "Spatial registration and normalization of images," *Human Brain Map.*, vol. 3, no. 3, pp. 165–189, 1995.

[36] A. Nieto-Castanon, S. S. Ghosh, J. A. Tourville, and F. H. Guenther, "Region of interest based analysis of functional imaging data," *NeuroImage*, vol. 19, no. 4, pp. 1303–1316, 2003.

[37] J. Tourville and F. Guenther, A cortical and cerebellar parcellation system for speech studies Boston Univ., Tech. Rep. CAS/CNS-03-022, 2003.

[38] A. Dale, B. Fischl, and M. Sereno, "Cortical surface-based analysis. I. Segmentation and surface reconstruction," *Neuroimage*, vol. 9, pp. 179–194, 1999.

[39] B. Zhou, A. Pai, and M. Nielsen, "Simultaneous registration and bilateral differential bias correction in brain MRI," in *MICCAI Workshop IntellMR*, 2014.

[40] X. Hua *et al.*, "Unbiased tensor-based morphometry: Improved robustness and sample size estimates for alzheimer's disease clinical trials," *Neuroimage*, vol. 66, pp. 648–661, 2013.

[41] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. New York: Chapman Hall, 1993.

[42] D. Holland, L. McEvoy, A. Dale, and A. D. N. Initiative, "Unbiased comparison of sample size estimates from longitudinal structural measures in ADNI," *Human Brain Map.*, vol. 33, no. 11, pp. 2586–2602, 2012.

[43] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach," *Biometrics*, vol. 44, no. 3, pp. 837–845, Sep. 1988.

[44] C. Jack *et al.*, "Rate of medial temporal lobe atrophy in typical aging and alzheimer's disease," *Neurology*, vol. 51, no. 4, pp. 993–999, 1998.

[45] M. Hernandez, M. N. Bossa, and S. Olmos, "Registration of anatomical images using paths of diffeomorphisms parameterized with stationary vector field flows," *Int. J. Comput. Vis.*, vol. 85, pp. 291–306, 2009.

[46] C. Chefd'hotel, G. Hermosillo, and O. Faugeras, "Flows of diffeomorphisms for multimodal image registration," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2002, pp. 753–756.

[47] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," in *Int. Conf. Visualizat. Biomed. Comput.*, 1996, pp. 267–276.

[48] N. Navab, A. Kamen, and D. Zikic, "Unifying characterization of deformable registration methods based on the inherent parametrization: An attempt at an alternative analysis approach," in *Proc. 4th Int. Workshop Biomed. Image Regist.*, 2010, vol. 6204, pp. 161–172.

[49] X. Pennec, P. Cachier, and N. Ayache, "Understanding the demon's algorithm: 3d non-rigid registration by gradient descent," in *MICCAI*, 1999, vol. 1679, pp. 597–605.

[50] M. Nielsen, L. Florack, and R. Deriche, "Regularization, scale-space, and edge detection filters," in *ECCV*, 1996, vol. 1065, pp. 70–81.

[51] W. Shi *et al.*, "Temporal sparse free-form deformations," *Med. Image Anal.*, vol. 17, no. 7, pp. 779–789, 2013.

[52] E. Haber and J. Modersitzki, "A multilevel method for image registration," *SIAM J. Sci. Comput.*, vol. 27, no. 5, pp. 1594–1607, 2006.

[53] W. Cai and J. Wang, "Adaptive multiresolution collocation methods for initial boundary value problems of nonlinear PDEs," *SIAM J. Numer. Anal.*, vol. 33, no. 3, pp. 937–970, 1996.