A Multigrid Approach for Fourth-Order Equations in Image Registration

Saskia Neuber[1(✉)] and Jan Modersitzki[1,2]

[1] Institute of Mathematics and Image Computing, Universität zu Lübeck, 23562 Lübeck, Germany
{s.neuber,jan.modersitzki}@uni-luebeck.de
[2] Fraunhofer Institute for Digital Medicine MEVIS, 23562 Lübeck, Germany

Abstract. Image registration is a demanding task that is required in many different areas of application, in particular in medical imaging. Due to the ill-posedness of image registration problems, regularization is unavoidable. This paper focuses on a family of so-called vector-field (VF) regularizers which consist of second-order energies based on a convex combination of gradients of divergence and rotation. Following a discretize-then-optimize approach, this paper proposes a staggered-grid discretization of the VF regularizers and applies a quasi-Newton type minimization to the image registration problem. Here the most costly part is solving linear systems, which can be regarded as a discretization of a linearization of a partial differential equation of fourth order. This paper proposes a highly efficient multigrid (MG) solver. In particular, the paper presents a local Fourier analysis to show that the suggested discretization is well suited for MG. More specifically, the paper provides an explicit number for the h-ellipticity measure and the local smoothing factor for a collective ω-relaxed Jacobi-type iteration. Our numerical results, including 3D image registration tasks, underline that the MG solver has in fact a complexity of $\mathcal{O}(n)$, where n is the number of unknowns.

Keywords: Image Registration · Efficient Solvers · Multigrid · Local Fourier Analysis

1 Introduction

Image registration is a challenging problem, especially in medical imaging. The goal of image registration is to establish geometric correspondences between two or more images; see, e.g., [16]. Various approaches have been proposed, based e.g. on Deep Learning [11], Optical Flow [7] or variational [25] methods.

This paper follows a variational approach [16,24,25], i.e. a joint energy is minimized, which consists of a data fitting term and a regularizer. As described for example in [16], a regularizer is typically based on either first- (e.g. diffusion or elastic) or second-order (e.g. curvature) derivatives. The elastic regularizer couples the components of the transformation, whereas diffusion and curvature

T. A. Bubba et al. (Eds.): SSVM 2025, LNCS 15668, pp. 391–403, 2025.
https://doi.org/10.1007/978-3-031-92369-2_30

are uncoupled. A coupled approach may be more useful in a biomedical setting, while uncoupled approaches allow for more efficient numerical schemes. Compared to first-order regularization, second-order regularization leads to smoother transformations and hence allows the inclusion of constraints such as corresponding landmarks [6] or the so-called 2D/3D registration [13]. The drawback of second-order terms is that the numerical treatment of the registration problem is significantly more complex, in particular for coupled approaches.

This paper focuses on a family of coupled second-order so-called vector-field (VF) regularizers [1,18,21,23]. The considered family is based on a convex combination of derivatives of divergence and rotation of the transformation and includes the curvature [6] energy, also known as thin-plate-spline energy.

In order to minimize the joint energy this paper follows a discretize-then-optimize approach with a quasi-Newton optimization scheme [19]. Here the computational demanding part is the solution of the potentially large scale Newton-system $Hs_h = \nabla J$, where H denotes the Hessian and ∇J the gradient of the discretized objective function and s_h an update direction. The system can be considered as a discretization of a linearization of a fourth-order partial differential equation and is solved with an efficient multigrid solver (MG) [3,22].

MG is established for many first-order [8,10,14,20] and second-order [4,12,15] regularizers already, but to the best of our knowledge a MG analysis for the considered VF regularizers is lacking. This paper presents a staggered grid discretization of the VF regularizers, which is ensured to be well suited for MG with local Fourier analysis. More exactly, the h-ellipticity of the differential operator of fourth order as well as the local smoothing factor for a collective ω-relaxed Jacobi-type iteration are determined. Numerical results underline the efficiency of the proposed MG solver and demonstrate the application to image registration problems.

The paper is organized as follows. The VF regularizers and their staggered-grid discretization are presented in Sect. 2. The MG components and a local Fourier analysis of the differential operator of fourth order are provided in Sect. 3. Numerical examples demonstrating the efficient use of MG solvers in 3D image registration can be found in Sect. 4.

2 Vector-Field Regularization and Its Discretization

A general discussion about image registration can be found in [16]. We assume that two images are given, i.e. $\mathcal{T}, \mathcal{R} : \mathbb{R}^d \to \mathbb{R}$, continuously differentiable and compactly supported on a set $\Omega \subset \mathbb{R}^d$, where $d \in \mathbb{N}$ denotes the spatial dimension. To simplify the presentation we restrict ourselves to $d = 2$. The aim is to find a transformation $y : \mathbb{R}^d \to \mathbb{R}^d$ such that $\mathcal{T} \circ y$ corresponds to $\mathcal{R}$. Since the specific data fitting is irrelevant for this work and to simplify the presentation, we stick to the L_2 norm. With a regularizer $\mathcal{S}$, the variational approach is to determine a minimizer y^* of the joint energy

$$\mathcal{J} : \mathcal{H}^k(\Omega, \mathbb{R}^d) \to \mathbb{R}, \quad \mathcal{J}(y) = \|\mathcal{T} \circ y - \mathcal{R}\|_{L_2}^2 + \alpha \mathcal{S}(y - y_{\text{ref}}), \tag{1}$$

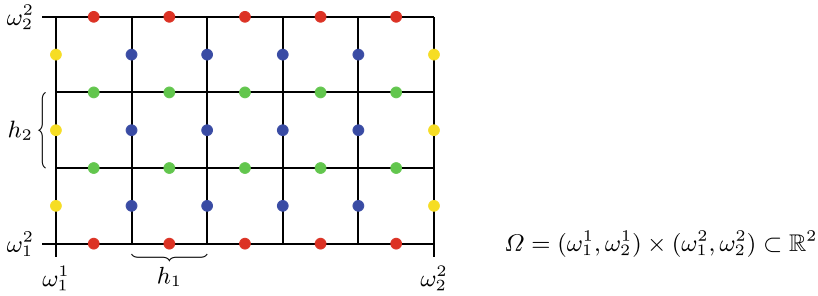

Fig. 1. Inner-staggered grid for $d = 2$: first component in blue and second component in green, Dirichlet-0 boundary points are in yellow and red. (Color figure online)

where $\alpha > 0$ is a regularization parameter, $\mathcal{H}^k(\Omega, \mathbb{R}^d)$ is a Sobolev-space where $k \in \mathbb{N}$ depends on the order of the regularizer and $y_{\text{ref}} : \mathbb{R}^d \to \mathbb{R}^d$ denotes a reference transformation, e.g. $y_{\text{ref}}(x) = x$ for all x.

2.1 Vector-Field Regularization

For the regularization, we focus on the family of coupled second-order vector-field (VF) regularizer [1,18,21,23], $\mathcal{S} : \mathcal{H}^2(\Omega, \mathbb{R}^d) \to \mathbb{R}$,

$$\begin{aligned}\mathcal{S}(y;\gamma) &:= \gamma \int_\Omega (\mathcal{B}_{\text{div}}\, y : \mathcal{B}_{\text{div}}\, y) + (1-\gamma) \int_\Omega (\mathcal{B}_{\text{curl}}\, y : \mathcal{B}_{\text{curl}}\, y)\; dx \\ &= \int_\Omega \gamma \|\nabla \text{div}\; y\|^2 + (1-\gamma)\|\nabla \text{curl}\; y\|^2\; dx\end{aligned}$$

with a weighting factor $\gamma \in (0,1)$, $(A : B) = \sum_{j,k} A_{j,k} B_{j,k}$, and, for $d = 2$,

$$\mathcal{B}_{\text{div}} = \begin{pmatrix} \partial_{11} & \partial_{12} \\ \partial_{12} & \partial_{22} \end{pmatrix} \quad \text{and} \quad \mathcal{B}_{\text{curl}} = \begin{pmatrix} -\partial_{12} & \partial_{11} \\ -\partial_{22} & \partial_{12} \end{pmatrix}.$$

Note that for $\gamma = 0.5$ the VF regularizer simplifies, leading to a decoupling of the components and thus the curvature regularization. In the following, we use an arbitrary γ and omit the dependency on γ.

For the minimization of the joint energy (1), the fourth-order differential operator $\mathcal{A}$ enters with the Euler-Lagrange equation into play,

$$\begin{aligned}\mathcal{A} &= \gamma \mathcal{B}_{\text{div}}^\top \mathcal{B}_{\text{div}} + (1-\gamma)\mathcal{B}_{\text{curl}}^\top \mathcal{B}_{\text{curl}} \\ &= \begin{pmatrix} \gamma\partial_{1111} + \partial_{1122} + (1-\gamma)\partial_{2222} & (2\gamma-1)(\partial_{1112} + \partial_{1222}) \\ (2\gamma-1)(\partial_{1112} + \partial_{1222}) & (1-\gamma)\partial_{1111} + \partial_{1122} + \gamma\partial_{2222} \end{pmatrix}.\end{aligned} \tag{2}$$

2.2 Discretization of the Differential Operators

We follow a discretize-then-optimize framework for the minimization of the joint energy. For the discretization of the integrals we use a midpoint quadrature rule [5] and focus on the discretization of the VF regularizer and in particular the fourth-order differential operator $\mathcal{A}$ in (2) with finite differences. Assuming that global affine parts are captured by a preregistration y_{ref}, thus Dirichlet-0 boundary conditions are added for the displacement $y_{\text{ref}} - y$. Hence, we discretize the displacement on an inner-staggered grid as indicated in Fig. 1 for $d = 2$.

We start with the discretization of $\mathcal{B}_{\text{div}}$ and $\mathcal{B}_{\text{curl}}$. For the discretization of each operator two distinct versions, see Table 1, are needed, since the components of the displacement are discretized cell-centered in one and inner-nodal in the other direction. For example, for $d = 2$ the discretized differential operators of second-order are

$$B_{\text{div}} = \begin{pmatrix} I_2^c \otimes D_{11}^{in} & D_2^{in} \otimes D_1^c \\ D_2^c \otimes D_1^{in} & D_{22}^{in} \otimes I_1^c \end{pmatrix} \quad \text{and} \quad B_{\text{curl}} = \begin{pmatrix} -D_2^c \otimes D_1^{in} & I_2^{in} \otimes D_{11}^c \\ -D_{22}^c \otimes I_1^{in} & D_2^{in} \otimes D_1^c \end{pmatrix},$$

where $\otimes$ denotes the Kronecker-product.

Table 1. Identity, first- and second-order finite difference operators for cell-centered and inner nodal grids of size $m = m_1 \times \cdots \times m_d$ with grid-width $h_j = (\omega_2^j - \omega_1^j)/m_j$ and Dirichlet-0 boundary condition;identity: $I_j^c \in \mathbb{R}^{m_j, m_j}$ and $I_j^{\text{in}} \in \mathbb{R}^{m_j+1, m_j-1}$, first-order: $D_j^c \in \mathbb{R}^{m_j+1, m_j}$ and $D_j^{\text{in}} \in \mathbb{R}^{m_j, m_j-1}$ and second-order: $D_{jj}^c \in \mathbb{R}^{m_j, m_j}$ and $D_{jj}^{\text{in}} \in \mathbb{R}^{m_j+1, m_j-1}$.

$$I_j^c := \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}, \qquad I_j^{\text{in}} := \begin{pmatrix} 0 & & & \\ 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix},$$

$$D_j^c := \frac{1}{h_j} \begin{pmatrix} 2 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -2 & \end{pmatrix}, \qquad D_j^{\text{in}} := \frac{1}{h_j} \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & \end{pmatrix},$$

$$D_{jj}^c := \frac{1}{h_j^2} \begin{pmatrix} -3 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & 1 & -2 & 1 & \\ & & & 1 & -3 & \end{pmatrix}, \quad D_{jj}^{\text{in}} := \frac{1}{h_j^2} \begin{pmatrix} 0 & & & & & \\ -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & 1 & -2 & 1 & \\ & & & 1 & -2 & \\ & & & & 0 & \end{pmatrix}.$$

The discretization of the fourth-order differential operator $\mathcal{A}$ from Eq. (2) thus is

$$A = \gamma B_{\text{div}}^\top B_{\text{div}} + (1 - \gamma) B_{\text{curl}}^\top B_{\text{curl}}.$$

We demonstrate that the proposed discretization A is consistent in $\mathcal{O}(h^2)$.

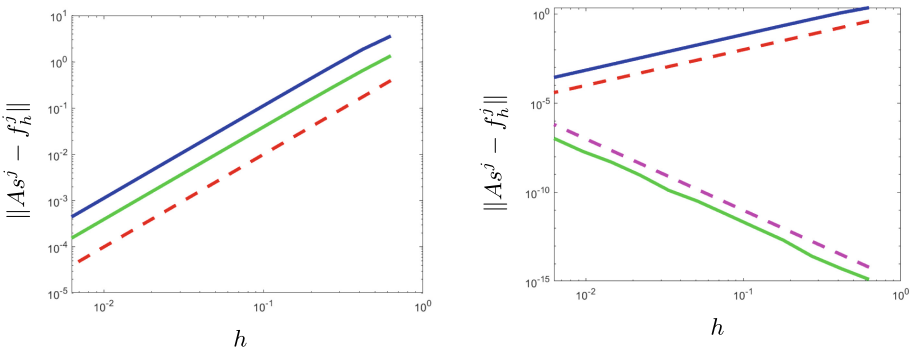

Fig. 2. Consistency check for Ex. 1 with $\gamma = 0.1$ (left) and $\gamma = 0.5$ (right). A log-log-plot of grid-width h against the consistency error $\|As^j - f_h^j\|$ for the first ($j = 1$, blue) and second ($j = 2$, green) component. Additionally the slopes 2(red) and -4 (magenta) are visualized in dashed. Catastrophic cancellation is visible for the second component for $\gamma = 0.5$, as no consistency error occurs. The discretization A of the fourth-order differential operator is consistent in $\mathcal{O}(h^2)$.

Example 1. On $\Omega = [0, 2\pi]^2$ we consider the linear system $\mathcal{A}\, s = f$ with

$$s(x) = \begin{pmatrix} \sin(x_1)\cos(2x_2) \\ 0 \end{pmatrix} \quad \text{and} \quad f(x) = \begin{pmatrix} (20 - 15\gamma)\sin(x_1)\sin(2x_2) \\ 10(1 - 2\gamma)\cos(x_1)\cos(2x_1) \end{pmatrix}$$

and their inner-staggered discretizations s_h, f_h. The consistency error $\|As_h - f_h\|$ for the proposed discretization A is plotted against the grid-width h in a log-log plot in Fig. 2. For $\gamma = 0.1$ both components have the desirable slope 2. For $\gamma = 0.5$ the slope of the second component is -4. The reason for this observation is twofold. The discretization of A never has a (block) diagonal form, due to the staggered grids. Nevertheless if the discretization is correct, we would expect an $\mathcal{O}(h^2)$ approximation of zero. Since the fourth-order differential operator approximates a constant (here zero) to all relevant digits, catastrophic cancellation [9] occurs, that is weighted with the factor $\frac{1}{h^4}$. Catastrophic cancellation is a well known phenomenon for finite differences, but it is only visible here as no consistency error occurs. Hence the discretization A is consistent in $\mathcal{O}(h^2)$.

3 A Multigrid Solution for Linear Systems

In the previous section a consistent staggered-grid discretization of the VF regularizer was introduced. Following a discretize-then-optimize approach, we next focus on the optimization. For the minimization of the joint energy, an inexact Gauss Newton scheme, as proposed in [10] is used. This leads to the system $Hs_h = -\nabla J$, which we solve with multigrid (MG) methods. For a simple presentation we denote the right-hand side ∇J by f_h and omit the index h.

In a worst case scenario the Hessian consists only of the Hessian of the VF regularizer, i.e. $H = A$. Hence a local Fourier analysis of the MG properties of the operator $\mathcal{A}$ from Eq. (2) is presented.

3.1 Local Fourier Analysis

The objective of local Fourier analysis [22, §8.3] is to determine the quantitative convergence behavior and efficiency of an appropriate Multigrid algorithm. More precisely we use local Fourier analysis to demonstrate that the h-ellipticity measure $E_h(A)$ of the discretization A is bounded away from zero. Furthermore for the smoothing procedure, we propose an ω-relaxed Jacobi-type iteration, we ensure that the local smoothing factor μ_{loc} is smaller than one. The local smoothing factor can be interpreted as the worst asymptotic error reduction by one relaxation step of all high-frequency error components. Hence for an appropriate smoother $\mu_{\text{loc}} < 1$ must be fulfilled. More details about local Fourier analysis and MG in general can be found in [3,22,26].

In the following local Fourier analysis, the stencil notation [22, §8.3.1] of the operator $\mathcal{A}$ from Eq. (2) is needed. Here only the stencils corresponding to the inner grid-points are presented, as local Fourier analysis assumes infinite grids. The stencil notation of A corresponds to the finite difference operator A, introduced in Sect. 2.2. Note that staggered-grid discretizations avoid possible checkerboard instability in MG [22].

Theorem 1. *An $\mathcal{O}(h^2)$ approximation $A := \begin{pmatrix} A^{1,1} & A^{1,2} \\ A^{1,2} & A^{2,2} \end{pmatrix}$ of the operator $\mathcal{A}$ from Eq. (2) can be defined with the stencils*

$$A^{1,1} := \frac{1}{h^4} \begin{bmatrix} & \gamma & \\ 1 & -4\gamma - 2 & 1 \\ (1-\gamma) & -4(1-\gamma) - 2 \quad 10 \quad -4(1-\gamma) - 2 & (1-\gamma) \\ 1 & -4\gamma - 2 & 1 \\ & \gamma & \end{bmatrix},$$

$$A^{2,2} := \left(A^{1,1}\right)^\top, \quad A^{1,2} := \frac{(2\gamma - 1)}{h^4} \begin{bmatrix} & 1 & -1 & \\ 1 & -6 & 6 & -1 \\ -1 & 6 & -6 & 1 \\ & -1 & 1 & \end{bmatrix} = A^{2,1}.$$

The result follows directly from Taylor expansion.

First the h-ellipticity measure, see [22, §8.3.2], of A is explicitly determined.

Theorem 2. *The h-ellipticity measure of the discretized differential operator A from Thm. 1, is $\frac{1}{256}$.*

Proof. Using the stencil notation from Thm. 1 the symbol of the operator A is

$$\tilde{A}(\nu) = \frac{4}{h^4} \begin{pmatrix} \gamma\nu_1^2 + \nu_1\nu_2 + (1-\gamma)\nu_2^2 & (2\gamma - 1)\sqrt{\nu_1\nu_2}(\nu_1 + \nu_2) \\ (2\gamma - 1)\sqrt{\nu_1\nu_2}(\nu_1 + \nu_2) & (1-\gamma)\nu_1^2 + \nu_1\nu_2 + \gamma\nu_2^2 \end{pmatrix}.$$

Here the substitution $\nu_j = 1 - \cos\theta_j$ was used for the frequencies θ_j. The determinant of the symbol reads as $\det(\tilde{A})(\nu) = \frac{16}{h^8}\gamma(1-\gamma)(\nu_1 + \nu_2)^4$.

Therefore the h-ellipticity measure is

$$E(A) = \frac{\min\{|\det(\tilde{A})(\nu)| : \nu \in [0,2]^2 \backslash [0,1)^2\}}{\max\{|\det(\tilde{A})(\nu)| : \nu \in [0,2]\}} = \frac{\frac{16}{h^8}\gamma(1-\gamma)}{\frac{16}{h^8} \cdot 256\gamma(1-\gamma)} = \frac{1}{256}.$$

Note that the h-ellipticity measure is in particular independent of the constant γ and the operator A is h-elliptic for all choices of γ. Hence the suggested staggered-grid discretization of A is appropriate for a MG framework.

In a next step the smoothing properties [22, §8.3.1] of the collective ω-relaxed Jacobi-type iteration are examined, to show that this is a possible smoother.

Theorem 3. *For the ω-relaxed Jacobi-type iteration applied to a linear system with the discretized differential operator A from Thm. 1, then the local smoothing factor is*

$$\mu_{loc}(\omega;\gamma) = \max\left\{\left|1-\frac{16\omega}{5}(1+\rho_\gamma)\right|, \left|1-\frac{\omega}{5}(1-\rho_\gamma)\right|\right\}$$

with $\rho_\gamma := \sqrt{1-4\gamma(1-\gamma)}$.

Proof. The ω-relaxed Jacobi-type update is given via

$$y_h^{k+1} = y_h^k + \omega(z_h^{k+1} - y_h^k) \quad \text{where} \quad z_h^{k+1} = \left(I - \frac{h^4}{10}A\right) y_h^k + \frac{h^4}{10} f_h.$$

The splitting matrices read as $A^+ = \frac{10}{\omega h^4} I$ and $A^- = A - \frac{10}{\omega h^4} I$ with the symbols $\tilde{A}^+ = \frac{10}{\omega h^4} I$ and $\tilde{A}^- = \tilde{A} - \frac{10}{\omega h^4} I$.

The determinant $\det := \det(\lambda(\nu)\tilde{A}^+ + \tilde{A}^-)$ with

$$\det = \frac{\gamma(1-\gamma)16(\nu_1+\nu_2)^4}{h^8} + \frac{40(\lambda(\nu)-1)(\nu_1+\nu_2)^2}{\omega h^8} + \frac{100(\lambda(\nu)-1)^2}{\omega^2 h^8}$$

has the roots $\lambda^*(\nu) = 1 - \frac{\omega}{5}(\nu_1+\nu_2)^2\left(1 \pm \sqrt{1-4\gamma(1-\gamma)}\right)$.

Hence the smoothing factor is computed as

$$\begin{aligned}\mu_{loc} &= \sup\{|\lambda^*(\nu)| : \nu \in [0,2]^2 \backslash (0,1)^2\} \\ &= \max\left\{\left|1-\frac{16\omega}{5}(1+\rho_\gamma)\right|, \left|1-\frac{\omega}{5}(1-\rho_\gamma)\right|\right\}\end{aligned}$$

with $\rho_\gamma := \sqrt{1-4\gamma(1-\gamma)}$.

Since $\mu_{\text{loc}} < 1$ is necessary for a smoother, this leads to an upper bound for the relaxation parameter ω.

Remark 1. According to Thm. 3, the ω-relaxed Jacobi-type iteration is a smoothing procedure, if

$$\omega < \frac{10}{16(1+\sqrt{1-4\gamma(1-\gamma)})}.$$

We further provide a sanity check to emphasize the importance of choosing an appropriate relaxation parameter ω.

Example 2. For $m = [16, 16]$ and $\gamma = 0.1$ the right-hand side is set to $f = A \cdot 1$, where 1 is a vector of ones. The upper bound from Remark 1 is $\omega^* = \frac{25}{72}$. The approximated solution $\hat{s}$ of the linear system $As = f$ is obtained using the ω-damped Jacobi-type implementation in FAIR [17] with four grids, i.e. the linear system is solved directly on the coarsest grid with resolution $m = [2, 2]$. More details about the MG set-up are given in Sect. 3.2. For $\omega = 0.3 < \omega^*$ and $\omega = 0.5 > \omega^*$ the relative residuum $r = \frac{\|f - A\hat{s}\|}{\|f\|}$ is plotted against the number of iterations in Fig. 3. For the suitable relaxation parameter $\omega = 0.3$ the residuum decreases and hence the ω-damped Jacobi-type iteration is an appropriate smoother, whereas the residuum increases for the non-suitable choice $\omega = 0.5$.

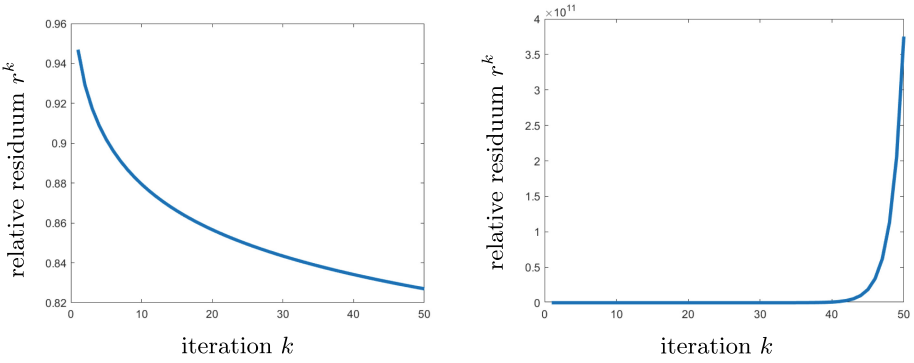

Fig. 3. The relative residuum against the number of iterations is visualized for ω-damped Jacobi-type iterations as described in Ex. 2. The relative residuum decreases for $\omega = 0.3 < \omega^*$ (left) and rises for $\omega = 0.5 > \omega^*$ (right). Hence only for $\omega = 0.3$ the ω-relaxed Jacobi-type iteration is a smoothing procedure.

Hence we identified the collective ω-relaxed Jacobi-type iteration as an appropriate smoother for the linear system $As = f$, if the relaxation parameter ω is chosen according to Remark 1.

3.2 Further Multigrid Components

Further MG components [22, §2.3] than just the smoothing procedure need to be set for an appropriate MG solver.

For the prolongation between the grids we set the bilinear interpolation and for the restriction, the full weighting operator, both implemented in FAIR [17]. Hence $o^P = 2 = o^R$, where o^P, o^R is the order of the restriction or prolongation operator. Hence the inequality $o^P + o^R \geq o$ from [3, §4.3] is fulfilled, as $o = 4$ is the order of A. Furthermore we perform at least two cycles of the MG solver as for one cycle the inequality $o^R \geq$ o needs to be fulfilled according to [3, §4.3].

The coarse grid operator of the VF regularizer is computed directly. For the computation of the coarse grid operator of the data-fitting term we follow the approach proposed in [10] and use a diagonal approximation and a Galerkin-based approach.

We apply three pre- and one post-smoothing step to the V-cycle and hence use the notation V(3,1).

Example 3. To verify the proposed MG components, we extend the sanity check from Ex. 2. We change the size to $m = (2^p, 2^p)$ for $p = 4, 5, \cdots, 12$, which leads to a linear system of equations with $2^{p+1}(2^p - 1)$ unknowns. Two cycles of the V(3,1) MG iterates are performed. The resulting runtime and the relative residual are listed in Table 2 for different values of p. As expected the runtime is linear in the number of unknowns and the relative residual converges to the asymptotic convergence factor.

4 The Multigrid Solver in Image Registration

To demonstrate the performance of the MG solver for image registration problems and further to illustrate the extension for $d = 3$, we provide 3D image registration examples from the publicly available dataset ACDC [2].

The dataset includes image pairs with three expert segmentations.

Each image pair consists of two 3D MRI scans of the human heart, one in the end-systolic, and the other in the end-diastolic phase. The 3D images are taken as a series of parallel 2D short axis slices.

For the following examples we use case 101. The end-systolic scan is used as template and the end-diastolic scan as reference image.

The image registration framework is implemented in Matlab and based on FAIR [17]. For faster computations, matrix-free implementations are used.

In all following examples the VF parameter γ is set to $\gamma = 0.3$. Additionally the following MG setting is used: The relaxation parameter is set to $\omega = 0.2$, and two V(3,1) cycles are applied.

First we have a closer look at the inner(= MG) iterations within the first outer (= inexact Gauss-Newton) iteration. The discussion about outer vs. inner iterations is beyond the scope of this contribution.

We ensure that the proposed MG set-up is appropriate for the 3D image registration framework, by measuring the relative residual in a two-grid cycle.

Table 2. Runtime and relative residual for $2^{p+1}(2^p - 1)$ unknowns, see Ex. 3. As expected from a suitable MG solver the runtime increases linear to the number of unknowns and the relative residuum converges.

p	4	5	6	7	8	9	10	11	12
runtime [s]	0.034	0.034	0.063	0.110	0.311	1.871	7.548	32.680	128.217
relative residual	0.349	0.378	0.400	0.407	0.412	0.415	0.417	0.417	0.418

Example 4. We set the resolution to $m = (16, 16, 4)$, i.e. 2688 unknowns, and perform two-grid cycles. Within the first outer iteration the number of cycles is plotted against the relative residuum of the MG solver to the left in Fig. 4. As expected the relative residuum decreases, which demonstrates that the MG set-up is suitable.

We show that for large linear systems, that potentially occur during 3D image registration, the MG solver outperforms other solver, e.g. backslash or conjugate gradient [19] solver in terms of efficiency.

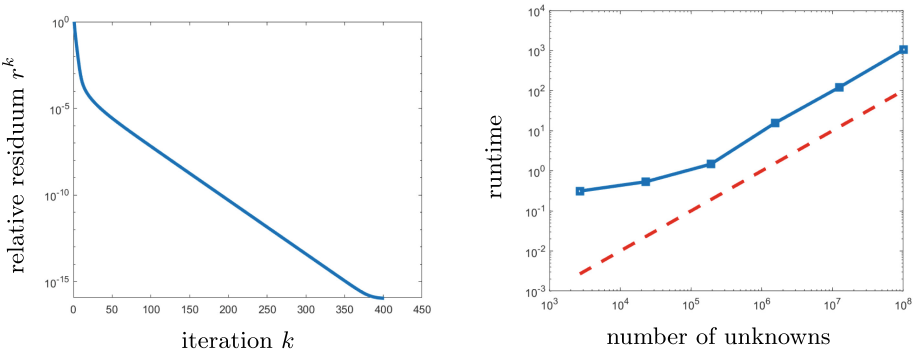

Fig. 4. Numerical results for 3D image registration of the ACDC data with a MG solver. Left: The relative residuum on a logarithmic scale against the number of MG iterations for the two-grid cycle from Ex. 4. The residuum decreases and thus validates the MG set-up. Right: The number of unknowns plotted against the runtime in seconds in a log-log-plot with the setting as proposed in Ex. 5. A linear slope is added in dashed to visualize the expected linear behaviour of the runtime for large systems.

Example 5. For resolution $m = 2^p(16, 16, 4)$, i.e. $3(2^{3p+10} - 2^{2p+7})$ unknowns, with $p = 0, 1, \cdots, 5$ we tracked the runtime of the inner iterations. As the build-in Matlab solver ("backslash") is a direct solver only one iteration is needed and the exact solution is met. For the conjugate gradient solver, we use a matrix-free implementation and ensure with the termination criterion that the solver achieves the same residual reduction, hence the same accuracy, as two cycles of the MG solver. Hence the number of iteration differs for every resolution. Both solvers are implemented in FAIR [17]. The obtained runtime for the three solvers with different resolutions are summarized in Table 3. The backslash solver is limited to a rather small resolution m. Additionally Table 3 specifies the number of iterations that the conjugate gradient solver needs until the same residual reduction as for the two MG cycles is achieved. As expected the MG solver is the fastest solver for a large linear system. Furthermore the MG solver's runtime against the number of unknowns is plotted in Fig. 4 on the right. For large linear equation systems, the linear relationship $\mathcal{O}(n)$ between runtime and number of unknowns is achieved.

Finally, we present some 3D image registration results.

Table 3. Runtime in seconds (and number of inner iterations for the conjugate gradient solver) for different solvers and for resolution $m = 2^p(16, 16, 4)$ of the ACDC scan. As expected the Multigrid solver is the fastest solver for large linear systems.

Solver \ p	0	1	2	3	4	5
Backslash	0.07	2.51	173.72	–	–	–
Multigrid	0.31	0.53	1.47	15.71	121.30	1044.70
Conjugate Gradient	0.12 (12)	0.17 (7)	0.52 (4)	8.58 (9)	380.81 (56)	22596 (425)

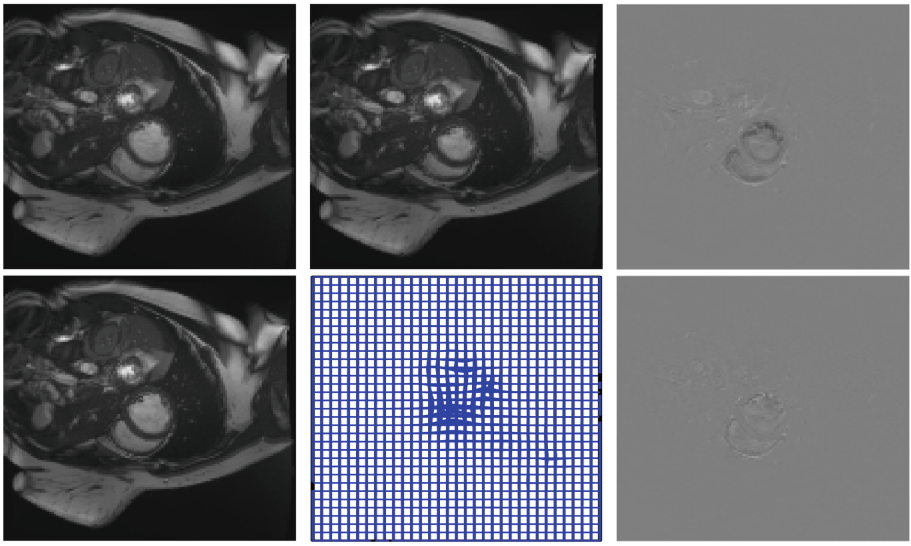

Fig. 5. Image registration results for the 3D ACDC images as described in Ex. 6. First row: a slice of the reference (left), template (middle) and their difference image (right) with no registration. Second row: corresponding slice of the deformed template image (left), a projection of the 3D deformed grid (middle) and the difference image (right) after registration. The MG solver is applicable to the 3D image registration problem.

Example 6. We follow a Multilevel framework with the finest resolution $m = (128, 128, 32)$ and set $y_{\text{ref}}(x) = x$. Here the maximal number of outer iterations is set to 30 and the proposed MG settings are used. The sum of squared differences is reduced from $1.44 \cdot 10^8$ to $6.01 \cdot 10^7$. The average Dice of the segmentations is increased from 0.69 to 0.78. The registration results are visualized in Fig. 5. Note that the registration is performed for $d = 3$ and slices of the volume

are used for the visualization. This example demonstrates the applicability of the MG solver to image registration problems.

5 Conclusion

An efficient numerical scheme, based on a MG solver for a fourth-order partial differential equation occurring during the optimization of a joint energy including the VF regularizer was proposed in this paper. Furthermore local Fourier analysis verified the suggested MG solver for the proposed discretization. Numerical examples in the form of sanity checks and 3D image registration results underline the theoretical predictions and demonstrate the effectiveness of MG methods for image registration with the VF regularizer. The family of VF regularizers is coupled and of second-order and hence suitable for medical registration frameworks with constraints such as corresponding landmarks or 2D/3D registration.

The application of more advanced smoothing strategies and a comparison between different smoothers is still pending and will be provided in future work as well as an examination of higher-order prolongation/interpolation operators, as second-order transfer operators are a borderline case here. Furthermore a discretization with finite elements similar to [20] could further improve the h-ellipticity, as the measure derived from the proposed discretization is rather small. Here due to the fourth order of the differential operator a non-conformal approach is necessary. Moreover, two- or three grid analysis as described in [26, §6], which is a local Fourier analysis of the whole iteration operator and not just the smoothing operator is planed. In addition, combining MG strategies with a coordinate transformation within the VF regularizer and hence to end up with a first order regularizer similar to [23] is also subject to further research.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Amodei, L., Benbourhim, M.N.: A vector spline approximation. J. Approx. Theory **67**(1), 51–79 (1991). https://doi.org/10.1016/0021-9045(91)90025-6
2. Bernard, O., et al.: Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE Trans. Med. Imaging **37**(11), 2514–2525 (2018)
3. Brandt, A.: Guide to multigrid development. In: Hackbusch, W., Trottenberg, U. (eds.) Multigrid Methods. LNM, vol. 960, pp. 220–312. Springer, Heidelberg (1982). https://doi.org/10.1007/BFb0069930
4. Chumchob, N., Chen, K., Brito-Loeza, C.: A fourth-order variational image registration model and its fast multigrid algorithm. Multiscale Model. Simul. **9**(1), 89–128 (2011). https://doi.org/10.1137/100788239
5. Davis, P.J., Rabinowitz, P.: Methods of Numerical Integration, 2nd ed. edn. Dover Publications, Mineola, NY (2007)

6. Fischer, B., Modersitzki, J.: Combining landmark and intensity driven registrations. In: PAMM: Proceedings in Applied Mathematics and Mechanics, vol. 3, pp. 32–35 (2003). https://doi.org/10.1002/pamm.200310309
7. Fortun, D., Bouthemy, P., Kervrann, C.: Optical flow modeling and computation: a survey. Comput. Vis. Image Underst. **134**, 1–21 (2015)
8. Frohn-Schauf, C., Henn, S., Witsch, K.: Multigrid based total variation image registration. Comp. Vis. Sci. **11**(2), 101–113 (2008)
9. Goldberg, D.: What every computer scientist should know about floating-point arithmetic. ACM Comput. Surv. (CSUR) **23**(1), 5–48 (1991)
10. Haber, E., Modersitzki, J.: A multilevel method for image registration. SIAM J. Sci. Comput. **27**(5), 1594–1607 (2006). https://doi.org/10.1137/040608106
11. Haskins, G., Kruger, U., Yan, P.: Deep learning in medical image registration: a survey. Mach. Vis. Appl. 1–18 (2020). https://doi.org/10.1007/s00138-020-01060-x
12. Héas, P., Mémin, E.: Three-dimensional motion estimation of atmospheric layers from image sequences. IEEE Trans. Geosci. Remote Sens. **46**(8), 2385–2396 (2008)
13. Heldmann, S., Papenberg, N.: A variational approach for volume-to-slice registration. In: Tai, X.-C., Mørken, K., Lysaker, M., Lie, K.-A. (eds.) SSVM 2009. LNCS, vol. 5567, pp. 624–635. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-02256-2_52
14. Henn, S., Witsch, K.: A multigrid approach for minimizing a nonlinear functional for digital image matching. Computing **64**(4), 339–348 (2000). https://doi.org/10.1007/s006070070029
15. Köstler, H., Ruhnau, K., Wienands, R.: Multigrid solution of the optical flow system using a combined diffusion-and curvature-based regularizer. Numer. Linear Algebra Appl. **15**(2–3), 201–218 (2008)
16. Modersitzki, J.: Numerical Methods for Image Registration. Oxford University Press (2004)
17. Modersitzki, J.: FAIR. Flexible Algorithms for Image Registration, vol. 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2009). https://doi.org/10.1137/1.9780898718843
18. Neuber, S., Schulz, P.F., Kuckertz, S., Modersitzki, J.: Segmentation-inspired image registration. In: BVM Workshop, pp. 205–210. Springer (2024)
19. Nocedal, J., Wright, S.J.: Numerical Optimization, 2nd ed. edn. Springer, New York, NY (2006)
20. Ruthotto, L., Greif, C., Modersitzki, J.: A stabilized multigrid solver for hyperelastic image registration. Numer. Linear Algebra Appl. **24**(5), e2095 (2017). https://doi.org/10.1002/nla.2095
21. Sorzano, C., Thévenaz, P., Unser, M.: Elastic registration of biological images using vector-spline regularization. IEEE Trans. Biomed. Eng. **52**(4), 652–663 (2005)
22. Trottenberg, U., Oosterlee, C.W., Schuller, A.: Multigrid. Academic Press, Orlando, FL (2001)
23. Tzitzimpasis, P., Ries, M., Raaymakers, B.W., Zachiu, C.: Generalized div-curl based regularization for physically constrained deformable image registration. Sci. Rep. **14**(1), 15002 (2024)
24. Vese, L.A., Le Guyader, C.: Variational Methods in Image Processing. CRC Press Boca Raton (2016)
25. Viergever, M.A., Maintz, J.A., Klein, S., Murphy, K., Staring, M., Pluim, J.P.: A survey of medical image registration–under review (2016)
26. Wienands, R., Joppich, W.: Practical Fourier Analysis for Multigrid Methods, vol. 4. Chapmaqn & Hall/CRC, Boca Raton, FL (2005)