Diffeomorphic Multi-Frame Non-Rigid Registration of Cell Nuclei in 2D and 3D Live Cell Images

Marco Tektonidis and Karl Rohr

***Abstract*— To gain a better understanding of cellular and molecular processes, it is important to quantitatively analyze the motion of subcellular particles in live cell microscopy image sequences. Since, generally, the subcellular particles move and cell nuclei move as well as deform, it is important to decouple the movement of particles from that of the cell nuclei using non-rigid registration methods. We have developed a diffeomorphic multi-frame approach for non-rigid registration of cell nuclei in 2D and 3D live cell fluorescence microscopy images. Our non-rigid registration approach is based on local optic flow estimation, exploits information from multiple consecutive image frames, and determines diffeomorphic transformations in the log-domain, which allows efficient computation of the inverse transformations. To register single images of an image sequence to a reference image, we use a temporally weighted mean image, which is constructed based on inverse transformations and multiple consecutive frames. Using multiple consecutive frames improves the registration accuracy compared to pairwise registration, and using a temporally weighted mean image significantly reduces the computation time compared with previous work. In addition, we use a flow boundary preserving method for regularization of computed deformation vector fields, which prevents from over-smoothing compared to standard Gaussian filtering. Our approach has been successfully applied to 2D and 3D synthetic as well as real live cell microscopy image sequences, and an experimental comparison with non-rigid pairwise, multi-frame, and temporal groupwise registration has been carried out.**

***Index Terms*— Biomedical image processing, microscopy, image sequence analysis, diffeomorphic registration.**

I. Introduction

THE motion of cellular and subcellular structures observed in temporal image sequences plays an important role in molecular biology. Biomolecular systems are dynamic, and one of the major challenges of biomedical research is to unravel the spatial and temporal relationships of these complex systems and to gain insight into the underlying biological processes. For example, the study of the movement of particles within the cell nucleus such as the promyelocytic leukemia nuclear bodies is important for understanding cell function, since these particles play a role in tumor suppression, viral defense, or DNA repair. Fluorescence microscopy allows imaging subcellular particles by using fluorescent proteins. Based on this technique time-lapse image sequences can be acquired for monitoring the position of subcellular particles over time. However, besides the particles, also the cell nucleus moves over time and undergoes deformations. The observed motion in the image data represents therefore a mixture of both the motion of the particles and the motion and deformation of the nucleus (Fig. 1). To compensate the nucleus motion and deformation, image registration can be used to align the frames of a time-lapse image sequence w.r.t. the first frame. Registration of live cell fluorescence microscopy images is challenging for a number of reasons, for example, autofluorescence causes significant background noise, and photobleaching reduces the image contrast over time. Additional challenges exist for image sequences acquired during prophase and prometaphase, when cell nuclei are going into mitosis (cell division). During these cell phases, the shape and the intensity structure of the nucleus are changing strongly due to, for example, chromatin condensation and nuclear envelope breakdown.

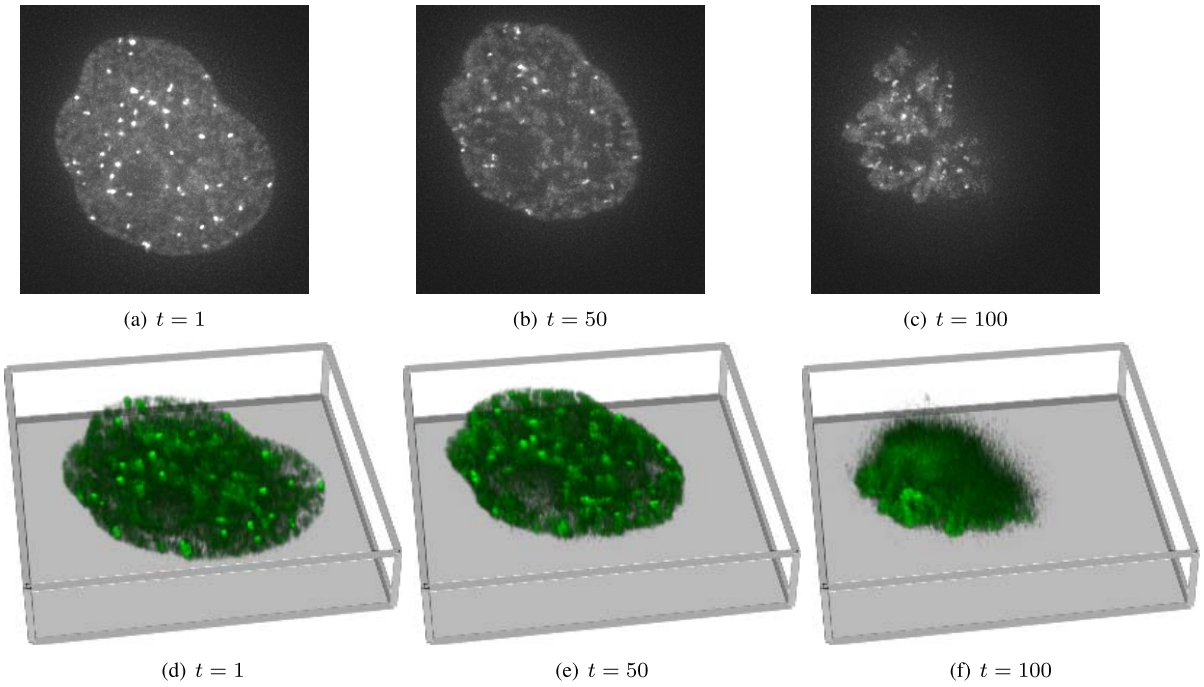
In previous work, mostly *rigid* and *affine* registration approaches have been used to register cells or cell nuclei observed in *live cell microscopy* image data (e.g., Rieger *et al.* [1], Wilson and Theriot [2], Matula *et al.* [3], de Vos *et al.* [4], Dzyubachyk *et al.* [5], Li *et al.* [6], van de Giessen *et al.* [7], Klein *et al.* [8]). However, to cope with nucleus deformations over time *non-rigid* registration approaches are required. Existing approaches for non-rigid registration of cell nuclei in live cell microscopy image data can be classified into geometry-based and intensity-based approaches. A geometry-based registration approach was described by Mattes *et al.* [9] which uses landmarks for registration of cell nuclei and a thin-plate spline deformation model. Another geometry-based registration approach was proposed by De Vylder *et al.* [10] which is based on contour matching and utilizes thin-plate splines. Sorokin *et al.* [11] described an extension of the contour-based approach in [10] and used the Navier equation as a deformation model. An intensity-based non-rigid registration approach was developed by Yang *et al.* [12] which is based on an extension of the demons algorithm with symmetric forces. Kim *et al.* [13] introduced a non-rigid intensity-based registration approach for dynamic cell nuclei images which uses an extension of an optic flow method. Tektonidis *et al.* [14] described a multi-frame registration approach for live cell nuclei which exploits information from multiple images simultaneously. Also in other biological applications, non-rigid registration of microscopy images has been used, for example, for 3D recon-

Manuscript received March 15, 2016; revised June 29, 2016 and December 23, 2016; accepted December 30, 2016. Date of publication January 15, 2017; date of current version February 3, 2017. This work was supported by the BMBF within the projects EpiSys (SysTec), ImmunoQuant (e:Bio), and CancerTelSys (e:Med), as well as by the DFG within the SFB 1129.

The authors are with the Biomedical Computer Vision Group, University of Heidelberg (BioQuant, IPMB, Dept. Bioinformatics and Functional Genomics) and German Cancer Research Center (DKFZ, Division Theoretical Bioinformatics), 69120 Heidelberg, Germany (e-mail: marco.tektonidis@bioquant.uni-heidelberg.de; k.rohr@dkfz.de).

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/TIP.2017.2653360



(a) $t = 1$ (b) $t = 50$ (c) $t = 100$

(d) $t = 1$ (e) $t = 50$ (f) $t = 100$

Fig. 1. Example images from a 3D dynamic microscopy image sequence (nucleus channel, contrast enhanced). (a)-(c): maximum intensity projection (MIP) images, and (d)-(f): rendered volumes.

struction of static fluorescence confocal microscopy images (e.g. Hogrebe *et al.* [15]), for 3D reconstruction of static electron microscopy images (e.g., Akselrod-Ballin *et al.* [16], Wang *et al.* [17]), for gene expression analysis (e.g., Peng *et al.* [18], Qu *et al.* [19]), and for cell shape modeling (e.g., Rohde *et al.* [20], Johnson *et al.* [21]). Rigid registration approaches have been used for stabilization of intravital video microscopy images of rolling leukocytes (e.g. Goobic *et al.* [22]) and for alignment of static 2D images from multi-tag fluorescence microscopy stacks (e.g. Raza *et al.* [23]).

In *medical image analysis*, *temporal groupwise* registration approaches have been proposed, which simultaneously take into account all or multiple images of a dynamic image sequence (e.g., Metz *et al.* [24], Yigitsoy *et al.* [25], De Craene *et al.* [26], Wu *et al.* [27], Durrleman *et al.* [28], Shi *et al.* [29], Zhang *et al.* [30], and Huizinga *et al.* [31]). There, MR, CT, and US images of organs are studied and typically periodic movement is considered (e.g., lung, heart). Often, *parametric* registration approaches are employed, for example, based on B-splines. In addition, *diffeomorphic* registration approaches have been proposed for medical applications (e.g., Beg *et al.* [32], Vercauteren *et al.* [33]). These approaches determine diffeomorphic transformations, which are invertible and the transformation as well as its inverse are smooth (continuously differentiable). Advantages are that the topology of objects is preserved and that foldings are prevented. This is important for dynamic image sequences where the observed objects are registered over time.

In this work, we propose a new intensity-based *diffeomorphic* approach for non-rigid *multi-frame* registration of cell nuclei in dynamic 2D and 3D microscopy images. In our approach, each image of an image sequence is registered to a temporally weighted mean image which is determined based on multiple consecutive images and previously computed inverse transformations. The proposed *non-parametric* registration approach is diffeomorphic which guarantees the invertibility and continuity of the computed transformations. This is important for our approach, since the inverse of the computed transformations are required for multi-frame registration. The transformations are determined based on velocity vector fields which are computed in the *log-domain* and allow efficient computation of the inverse transformations. This is important since the inverse transformations are required at each time point of an image sequence. Diffeomorphic multi-frame non-rigid registration approaches have not yet been employed for dynamic live cell image data. In Rohde *et al.* [20], Johnson *et al.* [21] a diffeomorphic registration approach (Beg *et al.* [32]) was used for shape modeling of cell nuclei. However, pairwise registration was performed (using two consecutive images) and static images or 2D time-lapse data were considered. Compared to the pairwise log-domain diffeomorphic registration approach of Vercauteren *et al.* [34], which was used for medical images (MR images of the brain) and which is a *global* optimization approach, our approach exploits information from multiple frames and is based on *local* optic flow. A local approach is advantageous in our application since many local structural changes occur in our image data (e.g., due to nuclear envelope breakdown or disappearance of nucleoli). Compared to the 2D multi-frame registration approach in [14] we introduce a diffeomorphic approach, and we also perform registration w.r.t. a temporally weighted mean image which significantly reduces the computation time. This allowed us to cope with 2D as well as 3D dynamic image

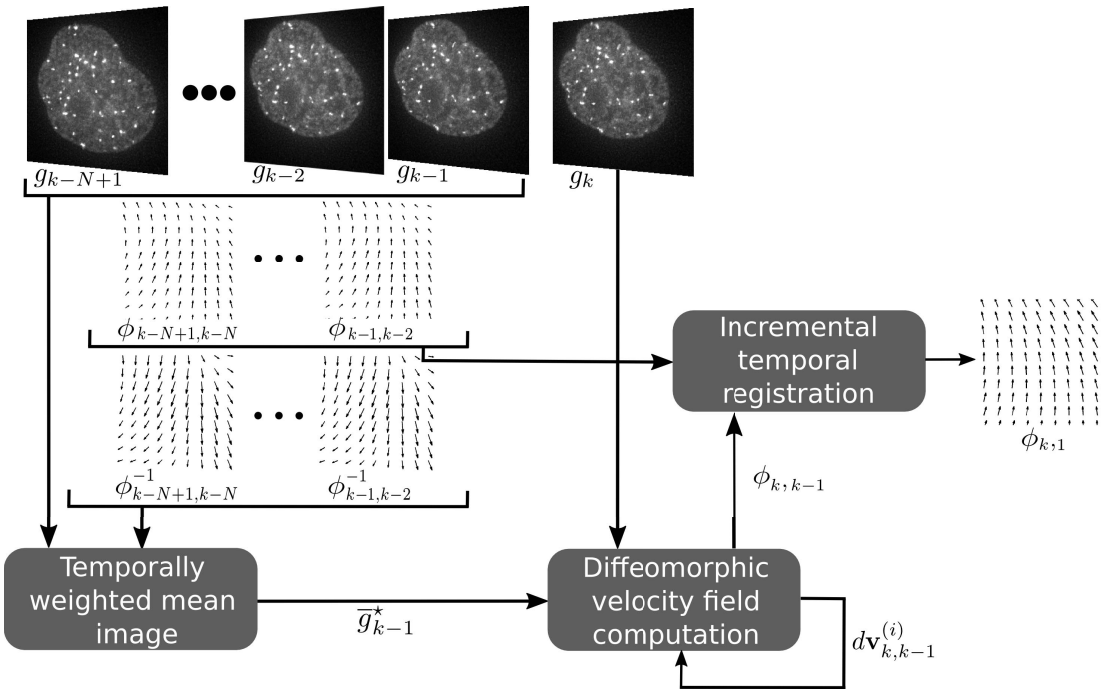


Fig. 2. Diagram of the diffeomorphic multi-frame non-rigid registration approach. First, for registration of an image g_k at time point k the temporally weighted mean image $\overline{g}^{\star}_{k-1}$ is determined based on multiple consecutive images g_l and the previously computed inverse transformations $\phi^{-1}_{l+1,l}$. Then, the temporal mean image and the image g_k are used for computing the transformation $\phi_{k,k-1}$. The computation is performed iteratively, using a log-domain diffeomorphic scheme combined with an intensity-based approach, where $d\mathbf{v}^{(i)}_{k,k-1}$ is an update vector field. Finally, the computed transformation $\phi_{k,k-1}$ is used to determine the transformation $\phi_{k,1}$ to the first image.

data. Another difference to [14] is that, for regularization of the computed deformation vector fields, we use a flow boundary preserving method instead of standard Gaussian smoothing to avoid over-smoothing. Our approach is applied to register the nucleus channel of two-channel microscopy image sequences. The computed transformations can be used to compensate the nucleus motion and deformation in the particle channel. We consider image sequences with a relatively large number of image frames (100-200 time points) compared to other work on temporal registration. The image data shows cells that go into mitosis and thus strong changes in the shape and the intensity structure of the nucleus occur which also increase over time. We have quantitatively evaluated our approach based on 2D and 3D synthetic data as well as 2D and 3D live cell microscopy images. We also performed an experimental comparison with previous pairwise [34], [13], multi-frame [14], and temporal groupwise [24] registration approaches.

This paper is organized as follows. In the following section, we introduce our diffeomorphic multi-frame approach for non-rigid registration of cell nuclei in dynamic microscopy image sequences. Then, we present experimental results for both synthetic and real 2D as well as 3D microscopy image data. Finally, we conclude the paper with a discussion and a summary.

II. Methods

In this section, we first give an overview of our diffeomorphic multi-frame non-rigid registration approach. Then, we introduce the multi-frame registration scheme based on a temporally weighted mean image. We describe the employed iterative log-domain diffeomorphic registration scheme and the intensity-based weighting optimization approach based on local optic flow. Finally, we present the flow boundary preserving method for regularization of the deformation vector fields.

A. Overview of Our Diffeomorphic Multi-Frame Non-Rigid Registration Approach

An overview of our diffeomorphic multi-frame non-rigid registration approach is given in Fig. 2. In our approach each image g_k of an image sequence at time point $k \geq 2$ is registered to the reference image g_1 at the first time point using the *diffeomorphic* transformation $\phi_{k,1}$ which is determined based on an incremental temporal registration scheme:

$$\phi_{k,1} = \phi_{k,k-1} \circ \phi_{k-1,1}, \tag{1}$$

where the transformation $\phi_{k-1,1}$ is computed analogously (recursively). The diffeomorphic transformation $\phi_{k,k-1}$ is determined by exploiting the information from $N \geq 3$ consecutive frames simultaneously, and can be described by a deformation vector field $\mathbf{u}_{k,k-1}$ which is added to the identity transformation, i.e., $\phi_{k,k-1} : \mathbf{x} \mapsto \mathbf{x} + \mathbf{u}_{k,k-1}(\mathbf{x})$. We determine $\phi_{k,k-1}$ using a temporally weighted mean image $\overline{g}^{\star}_{k-1}$ (see Sec. II-B below) based on the previous $N-1$ consecutive images (g_{k-N+1}, …, g_{k-1}) and a log-domain diffeomorphic method [34] which allows efficient computation of the inverse transformation $\phi^{-1}_{k,k-1}$. This method uses a Lie group structure [35] which defines an exponential mapping from the vector space of smooth velocity fields to diffeomorphisms.

Thus, the diffeomorphic transformation $\phi_{k,k-1}$ can be defined as an exponential of a smooth *stationary velocity vector field* $\mathbf{v}_{k,k-1}$ in the log-domain, i.e., $\phi_{k,k-1} = \exp\left(\mathbf{v}_{k,k-1}\right)$, and the inverse transformation is given by $\phi_{k,k-1}^{-1} = \exp(-\mathbf{v}_{k,k-1})$. The velocity field $\mathbf{v}_{k,k-1}$ is determined iteratively based on an update velocity field $d\mathbf{v}_{k,k-1}^{(i)}$ (see Sec. II-C below). Finally, the determined $\phi_{k,1}$ is regularized using the corresponding deformation vector field $\mathbf{u}_{k,1}$:

$$\mathbf{u}_{k,1} \leftarrow F * \mathbf{u}_{k,1}, \tag{2}$$

where F is the flow boundary preserving operator (see Sec. II-E below).

B. Multi-Frame Non-Rigid Registration Using a Temporally Weighted Mean Image

To compute the diffeomorphic transformation $\phi_{k,k-1}$ in (1) the image g_k is registered to the temporally weighted mean image determined from the $N-1$ previous images which have been transformed using the computed inverse transformations from previous time steps. Note, that for $k < N$ the temporally weighted mean image is determined based on the $k-1$ previous images. For a time point k, the temporal range of available previous images g_l is given by $\max(k-N+1, 1) \leq l \leq k-1$. The required transformations can be computed recursively by composing the inverse transformations from previous time points:

$$\phi_{l,k-1} = \phi_{l+1,l}^{-1} \circ \phi_{l+1,k-1}, \tag{3}$$

where $\phi_{l+1,l}^{-1}$ denotes the inverse transformation for the image at time point $l+1$, and $\phi_{l+1,k-1}$ is computed analogously (recursively). Note, that $\phi_{l,k-1}$ is diffeomorphic since it is a composition of diffeomorphic transformations. The computed transformation is regularized using the corresponding deformation vector field $\mathbf{u}_{l,k-1}$:

$$\mathbf{u}_{l,k-1} \leftarrow F * \mathbf{u}_{l,k-1}, \tag{4}$$

where F is the flow boundary preserving operator (see Sec. II-E below). The temporally weighted mean image is given by:

$$\overline{g}_{k-1}^{\star} = \sum_{l} \overline{W}_l \left(g_l \circ \phi_{l,k-1}\right), \tag{5}$$

where $g_l \circ \phi_{l,k-1}$ denotes the image at time point l transformed with $\phi_{l,k-1}$ and $\overline{W}_l$ is a normalized weight factor which controls the influence of the previous images on the registration result. We determine the weights $\overline{W}_l$ using a Gaussian function and based on the temporal distance between the time points l and k.

We compute the transformation $\phi_{k,k-1}$ by minimizing for each time point k and for each image point $\mathbf{x} = (x, y)$ or $\mathbf{x} = (x, y, z)$ the sum of squared intensity differences between the image g_k and the temporally weighted mean image $\overline{g}_{k-1}^{\star}$ in (5) within a neighborhood region Ω:

$$\sum_{\mathbf{x}\in\Omega} \left(\left(g_k \circ \phi_{k,k-1}\right)(\mathbf{x}) - \overline{g}_{k-1}^{\star}(\mathbf{x})\right)^2, \tag{6}$$

where $g_k \circ \phi_{k,k-1}$ denotes the image at time point k transformed with $\phi_{k,k-1}$.

Note that our multi-frame approach has several main differences to [14]. First, according to (6) we compute the sum of squared intensity differences only between the image g_k at time point k and the temporal mean image $\overline{g}_{k-1}^{\star}$, whereas in [14] the sum of squared intensity differences are computed between $N-1$ pairs of consecutive images (g_k and each $g_l \circ \phi_{l,k-1}$). This reduces the computation time significantly and allows application to 3D image data. Second, our multi-frame approach employs diffeomorphic transformations in contrast to [14] which guarantees that the transformations as well as the compositions of transformations in (3) are smooth and invertible.

C. Computation of the Diffeomorphic Velocity Vector Field

The velocity vector field $\mathbf{v}_{k,k-1}$ used for determining the diffeomorphic transformation $\phi_{k,k-1}$ in (1) is computed iteratively in the log-domain. At each iteration i the current $\mathbf{v}_{k,k-1}^{(i)}$ is updated by [34]:

$$\mathbf{v}_{k,k-1}^{(i)} = F * \left(Z\left(\mathbf{v}_{k,k-1}^{(i-1)}, G * d\mathbf{v}_{k,k-1}^{(i)}\right)\right), \tag{7}$$

where $d\mathbf{v}_{k,k-1}^{(i)}$ is an *update vector field* for the transformation between the image $g_k^* = g_k \circ \exp\left(\mathbf{v}_{k,k-1}^{(i-1)}\right)$ and the temporal mean image $\overline{g}_{k-1}^{\star}$, and G is a Gaussian kernel for regularization of the update field. The employed update scheme ensures that the computed transformation is diffeomorphic in contrast to the additive [36] or the compositive [37] update scheme. $Z\left(\mathbf{v}_{k,k-1}^{(i-1)}, G * d\mathbf{v}_{k,k-1}^{(i)}\right)$ denotes a velocity vector field such that:

$$\exp\left(Z\left(\mathbf{v}, d\mathbf{v}\right)\right) \approx \exp\left(\mathbf{v}\right) \circ \exp\left(d\mathbf{v}\right), \tag{8}$$

where for simplicity $\mathbf{v} = \mathbf{v}_{k,k-1}^{(i-1)}$ and $d\mathbf{v} = G * d\mathbf{v}_{k,k-1}^{(i)}$. This velocity vector field is determined using the Baker-Campbell-Hausdorff formula [38]:

$$Z\left(\mathbf{v}, d\mathbf{v}\right) = \mathbf{v} + d\mathbf{v} + \frac{1}{2}\left[\mathbf{v}, d\mathbf{v}\right] + \frac{1}{12}\left[\mathbf{v}, \left[\mathbf{v}, d\mathbf{v}\right]\right], \tag{9}$$

where the Lie bracket $[\mathbf{v}, d\mathbf{v}]$ represents a vector field at each position $\mathbf{x}$:

$$[\mathbf{v}, d\mathbf{v}](\mathbf{x}) = \mathrm{Jac}(\mathbf{v})(\mathbf{x}) d\mathbf{v}(\mathbf{x}) - \mathrm{Jac}(d\mathbf{v})(\mathbf{x})\mathbf{v}(\mathbf{x}), \tag{10}$$

where Jac denotes the Jacobian of the respective vector field. The iteration terminates either when a certain number of iterations is reached or when the sum of squared intensity differences in (6) is below a threshold value.

D. Intensity-Based Non-Rigid Diffeomorphic Multi-Frame Registration Based on Local Optic Flow

For minimizing the sum of squared intensity differences in (6), we use a local *weighting* optimization scheme which was previously used in [14]. In our *diffeomorphic multi-frame weighting* approach, we minimize for each image point $\mathbf{x}_c$ of an image the sum of squared intensity differences between the image $g_k^* = g_k \circ \exp\left(\mathbf{v}_{k,k-1}^{(i-1)}\right)$ and the temporally

weighted mean image $\overline{g}^{\star}_{k-1}$ within the neighborhood region Ω. The approach includes a weighting matrix and performs an optimization similar to the method of Levenberg-Marquardt (e.g. [39]). The optimization is performed for each time point k over the neighborhood region Ω around each voxel $\mathbf{x}_c$ of an image, where a vector $\mathbf{V}$ at the position $\mathbf{x}_c$ of the update vector field $d\mathbf{v}^{(i)}_{k,k-1}$ (7) in the log-domain is computed by:

$$\sum_{\mathbf{x}\in\Omega}\left(\nabla g^{*}_{k}\nabla g^{*T}_{k}+\mathbf{D}\right)\mathbf{V}=\sum_{\mathbf{x}\in\Omega}\nabla g^{*}_{k}\left(\overline{g}^{\star}_{k-1}-g^{*}_{k}\right), \quad (11)$$

where ∇g^{*}_{k} denotes the gradient of the image g^{*}_{k} at voxel $\mathbf{x}$, and the weighting matrix $\mathbf{D}$ is a diagonal matrix which indicates whether the current estimate of $\mathbf{V}$ is close to the solution or not. The elements of $\mathbf{D}$ are given analogously as in [14] except that g_{k-1} is substituted by the temporally weighted mean image $\overline{g}^{\star}_{k-1}$.

E. Flow Boundary Preserving Smoothing of Vector Fields

In previous non-parametric registration approaches typically Gaussian filters were used for regularization of deformation vector fields (e.g., [14], [34]). In our approach, we use a method which prevents over-smoothing across flow boundaries. Flow boundaries represent discontinuities in a vector field, i.e., neighboring vectors have strongly different directions. In our non-rigid registration approach we employ a flow boundary preserving method [40] which was previously used for optic flow computation and determines the smoothed vector $\mathbf{U}^{*}_{\mathbf{x}_c}$ at each position of the deformation vector field $\mathbf{u}_{k,k-1}$ as the weighted average of the vectors $\mathbf{U}_{\mathbf{x}_i}$ of $\mathbf{u}_{k,k-1}$ within a neighborhood Ω_{FB}:

$$\mathbf{U}^{*}_{\mathbf{x}_c}=\sum_{\mathbf{x}_i\in\Omega_{FB}} w_{\mathbf{x}_i}\mathbf{U}_{\mathbf{x}_i}, \quad (12)$$

where $\mathbf{x}_c$ denotes the center position and $\mathbf{x}_i$ are positions in the neighborhood Ω_{FB}. We compute the weight $w_{\mathbf{x}_i}$ for each vector based on the Euclidean distance between the positions $\mathbf{x}_c$ and $\mathbf{x}_i$ as well as the intensity difference between the corresponding positions in the image g_k:

$$w_{\mathbf{x}_i}=\exp\left(-\frac{\|\mathbf{x}_c-\mathbf{x}_i\|^2}{\sigma^2_{dist}}-\frac{|g_k(\mathbf{x}_c)-g_k(\mathbf{x}_i)|^2}{\sigma^2_{int}}\right), \quad (13)$$

where σ_{dist} and σ_{int} control the influence of the Euclidean distance between the positions and the intensity difference, respectively. The weight $w_{\mathbf{x}_i}$ describes how likely the voxels at positions $\mathbf{x}_i$ and $\mathbf{x}_c$ belong to the same image region. In Fig. 3, an example for smoothing a deformation vector field using the flow boundary preserving method in comparison to using a Gaussian kernel with standard deviation σ_G is shown for a cropped region of a 2D vector field for cell nuclei registration. For the Gaussian kernel we used $\sigma_G = 1$ pixel (we used a kernel size of 5×5 pixel), and for the flow boundary preserving method we used $\sigma_{dist} = 2$ pixel, $\sigma_{int} = 4$, and a 5×5 pixel neighborhood for Ω_{FB}. The region in the center of the figures is located on a flow boundary (the vectors in this region have very different directions). It can be seen, that using a Gaussian kernel results in over-smoothing, yielding close to zero vectors or vectors with significantly reduced magnitude in the center due to the averaging of the original vectors with very different directions (Fig. 3, left). For example, see the vector at position $\mathbf{x} = (211, 239)$. In comparison, using the flow boundary preserving method the vector is well preserved (Fig. 3, right). This example demonstrates that the flow boundary preserving method is superior to Gaussian smoothing at flow boundaries, however, the computation time is also higher. To reduce the computation time of our non-rigid registration approach we use the flow boundary preserving method only for vectors located at flow boundaries, while vectors in homogeneous regions of the vector field are smoothed using a Gaussian kernel [40]. We detect flow boundaries based on the gradient magnitude of the deformation vector field:

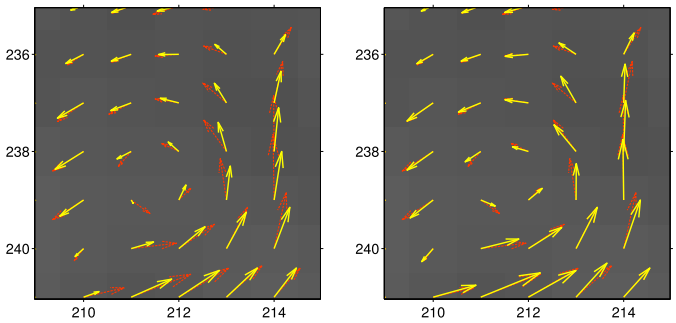


Fig. 3. Examples demonstrating the smoothing of a deformation vector field using a Gaussian kernel (left), and the flow boundary preserving method (right). Dashed red arrows depict the original vectors and yellow arrows depict the smoothed vectors.

$$\mathbf{U}^{FB}_{\mathbf{x}}=\{\mathbf{U}_{\mathbf{x}} \mid \|\nabla\mathbf{U}_{\mathbf{x}}\| \geq \overline{\|\nabla\mathbf{u}\|}+c\,\sigma_{\|\nabla\mathbf{u}\|}\}, \quad (14)$$

where $\overline{\|\nabla\mathbf{u}\|}$ and $\sigma_{\|\nabla\mathbf{u}\|}$ denote the mean and standard deviation of the Euclidean norm of the gradients of the vector field, respectively, and c is a weighting factor (we used $c = 1$). In our non-rigid registration approach we use this smoothing scheme for regularization of the vector fields in (4), (7), and (2).

III. Experimental Results

We first describe the used 2D and 3D live cell microscopy image data as well as the parameter setting. Then, we present experimental results based on synthetic and real image data.

A. Image Data and Parameter Setting

We have used four 2D and two 3D real live cell microscopy image sequences (denoted as A, B, C, D, for the 2D, and E, F, for the 3D sequences). The 2D data consist of 100 up to 200 images with a size of 384×384 up to 512×512 pixel, and the 3D data consist of 100 images with a size of $512\times512\times10$ and $512 \times 512 \times 15$ voxel (see Table I). The data has kindly been provided by Y.-C. Chen and D.L. Spector, Cold Spring Harbor Laboratory. For the acquisition of the images, a DeltaVision RT widefield microscope (Applied Precision) at a resolution of $0.216\mu m \times 0.216\mu m$ for the 2D sequences, and $0.216\mu m\times0.216\mu m\times0.5\mu m$ as well as $0.216\mu m\times0.216\mu m\times 1.5\mu m$ for the 3D sequences was used. The image sequences consist of two channels: the first channel displays nuclei of

TABLE I
REAL LIVE CELL MICROSCOPY IMAGE DATA USED FOR THE EXPERIMENTAL EVALUATION

	Real image sequence	Image size [pixel/voxel]	Number of time points
2D	A	512×512	150
	B	384×384	200
	C	512×512	100
	D	512×512	149
3D	E	$512 \times 512 \times 15$	100
	F	$512 \times 512 \times 10$	100

human live cells (U2OS cell line) with different chromatin stainings (H2A-mCherry, YFP-SP100 with Hoechst), and the second channel displays subcellular particles (CFP stained PML bodies). Our approach is applied for registration of the nucleus channel. The real image data was used for quantitative evaluation of our approach and comparison with previous approaches (Sec. III-C). We also used the real image data for generating synthetic image sequences (see Sec. III-B). In the experiments based on synthetic data we investigated the influence of image noise as well as intensity scaling.

For our diffeomorphic multi-frame non-rigid registration approach the same parameter setting was used for all 2D and 3D image data. We have determined the parameter setting based on the real and synthetic image sequences (see Sec. III-B and III-C). We first used an extensive range of parameter values for one real image sequence. Then, the result for the other image sequences was checked and a subset of the parameters was refined. Finally, this parameter setting was applied for all real and synthetic data in our experiments. The evaluation was based on the minimization of the registration errors and also the smoothness of the computed transformations was taken into account. Generally, a certain amount of smoothness is necessary, however, too strong smoothing results in over-smoothing. Below, we provide the parameters for the 3D case, the parameters for the 2D case are obtained by omitting the z-dimension. For computing the velocity vector field $\mathbf{v}_{k,k-1}$ in (7) we have used a maximum number of 10 iterations and a threshold of 0.0001 for the sum of squared intensity differences in (6) (in most cases the maximum number of iterations was used since the threshold is relatively small), and for regularization of the computed deformation vector field we have used the flow boundary preserving operator ($\sigma_{dist} = 5$ voxel, $\sigma_{int} = 7$, $5 \times 5 \times 3$ voxel neighborhood for Ω_{FB}) at flow boundaries combined with Gaussian smoothing with standard deviation $\sigma_G = (2, 2, 1)$ voxel (for the x-, y-, and z-dimension, respectively, we used a kernel size of $7 \times 7 \times 5$ voxel) in other image parts (see Sec. II-E above). For regularization of the deformation vector fields $\mathbf{u}_{l,k-1}$ in (4) and $\mathbf{u}_{k,1}$ in (2) we have used the flow boundary preserving operator ($\sigma_{dist} = 2$ voxel, $\sigma_{int} = 4$, $5 \times 5 \times 3$ voxel neighborhood for Ω_{FB}) combined with a Gaussian kernel of $\sigma_G = (1, 1, 0.5)$ voxel. We found, that less strong smoothing of these deformation vector fields compared to smoothing of the computed vector field in (7) generally improves the result. This is expected, since the deformation vector fields in (4) and (2) are based on the composition of transformations which are represented by already smoothed deformation vector fields. The update vector field in (7) was regularized using a Gaussian kernel with $\sigma_G = (2, 2, 1)$ voxel since for this vector field the flow boundary preserving method did not yield a significant improvement but requires more computation time. For the 3D data we reduced the smoothing strength (size of the operators) w.r.t. the z-dimension because of the lower resolution compared to the x- and y-dimensions. We found, that this generally improves the result. In (9) we used the two first terms of the Baker-Campbell-Hausdorff formula (as suggested in [34]) since adding more terms did not significantly change the result. For computing the update vector $\mathbf{V}$ in (11) a $5 \times 5 \times 3$ voxel neighborhood was used.

B. Experimental Results Using Synthetic Image Data

In this section, we describe experimental results of our diffeomorphic multi-frame approach for synthetic 2D and 3D image data. To simulate the motion and deformation of cell nuclei over time, we have generated four 2D (denoted by S1, S2, S3, S4) and two 3D (denoted by S5, S6) synthetic image sequences using the first frame of the different real microscopy image sequences (sequences A-D for 2D and sequences E, F for 3D, see Table I) and computed registration results. An advantage of the generated image sequences is that we have ground truth for the evaluation and that the image quality is similar to real data. We used the first image g_1 from the nucleus channel of a real image sequence as well as the transformations $\phi_{1,k} = \phi_{k,k-1}^{-1} \circ \phi_{1,k-1}$ which are obtained based on previously computed inverse transformations. To generate the synthetic data, we used the smoothed transformation $\phi_{1,k}$ to transform g_1 yielding an image $g_1 \circ \phi_{1,k}$ for each time point k of an image sequence.

We have applied our approach to the synthetic data and compared for each time point the computed vector field with the ground truth vector field. As performance measure we used the endpoint error (EE) which is defined as the Euclidean distance between the vectors of the computed vector field and the ground truth vector field. EE is often used for evaluation of optic flow approaches (e.g. [41]). For each time point we determined the mean endpoint error EE_{mean} over all computed vectors of a cell nucleus. For our approach we used different numbers of multiple frames ($N = 3, 5, 10, 20$) and compared the results with pairwise registration ($N = 2$).

The registration error for our approach averaged over all time points of each image sequence ($\overline{EE}_{mean}$) and over the four 2D and the two 3D synthetic image sequences for different values of N is shown in Table II (columns "Original synthetic sequences"). It can be seen, that the error using multiple frames ($N \geq 3$) is lower for all N compared to the pairwise registration ($N = 2$), both for the 2D and 3D case. In both cases, the result improves with increasing N, and the lowest error is obtained for $N = 20$. The improvement varies from 19% to 30% for the 2D data, and from 8% to 17% for the 3D data, compared to pairwise registration.

We also studied synthetic image data with added Gaussian noise (standard deviation $\sigma_n = 3$ voxel). The results are provided in Table II (columns "Sequences with Gaussian noise").

TABLE II

MEAN ENDPOINT ERROR $\overline{EE}_{mean}$ AVERAGED OVER FOUR 2D AND TWO 3D SYNTHETIC IMAGE SEQUENCES FOR THE NEW DIFFEOMORPHIC MULTI-FRAME REGISTRATION APPROACH AND ITS PAIRWISE VARIANT. PERCENTAGES INDICATE THE CHANGE COMPARED TO PAIRWISE REGISTRATION AND BOLD VALUES INDICATE THE LOWEST ERROR

[pixel] / [voxel]		Original synthetic sequences		Sequences with Gaussian noise		Sequences with linear intensity decrease and Gaussian noise	
		2D	3D	2D	3D	2D	3D
pairwise		2.05	2.41	2.87	2.55	4.28	2.84
multi-frame	$N = 3$	$1.65_{-19\%}$	$2.21_{-8\%}$	$2.26_{-21\%}$	$2.33_{-9\%}$	$3.19_{-25\%}$	$2.49_{-12\%}$
	$N = 5$	$1.51_{-26\%}$	$2.09_{-13\%}$	$1.94_{-32\%}$	$2.18_{-15\%}$	$2.78_{-35\%}$	$2.31_{-19\%}$
	$N = 10$	$1.47_{-28\%}$	$2.02_{-16\%}$	$1.73_{-40\%}$	$2.08_{-18\%}$	$2.66_{-38\%}$	$2.24_{-21\%}$
	$N = 20$	$\mathbf{1.44_{-30\%}}$	$\mathbf{2.01_{-17\%}}$	$\mathbf{1.66_{-42\%}}$	$\mathbf{2.05_{-20\%}}$	$\mathbf{2.65_{-38\%}}$	$\mathbf{2.23_{-22\%}}$

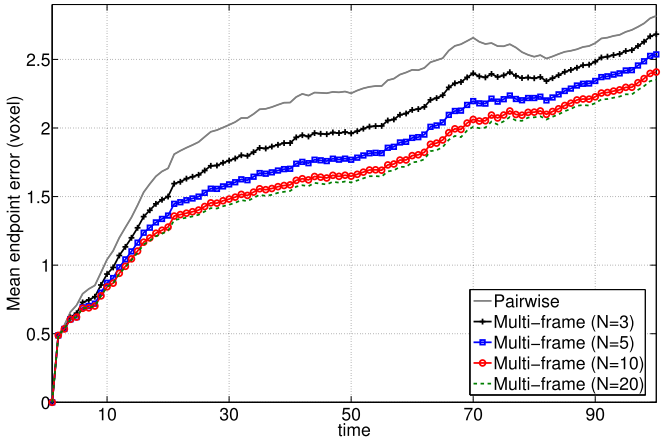


Fig. 4. Mean endpoint error EE_{mean} for the 3D synthetic sequence S6 with Gaussian noise ($\sigma_n = 3$ voxel) as a function of time using the new diffeomorphic multi-frame registration approach ($N = 3, 5, 10, 20$) and its pairwise variant ($N = 2$).

It can be seen, that the error increases compared to the data without noise (as expected), but also in this case the multi-frame approach outperforms pairwise registration (for the 2D and 3D data, and for all values of N). Averaging the errors over the four 2D image sequences and the two 3D image sequences, the best result is obtained for $N = 20$. The improvements are 42% and 20% for the 2D and 3D data, respectively. In Fig. 4 we show an example of the registration error over time for a 3D synthetic sequence with Gaussian noise. It can be seen, that the multi-frame approach yields a better result than pairwise registration, and the improvement increases with increasing N (for $N = 20$ the improvement is 21.9%). For visual inspection of the results we show an example for a cropped region of the ground truth and the computed inverse vector field using the pairwise and the multi-frame approach for $N = 10$ (Fig. 5). It can be seen, that the computed vectors using the diffeomorphic multi-frame approach are more similar to the ground truth vectors compared to pairwise registration, where the vectors are very different from the ground truth.

To simulate photobleaching in fluorescence microscopy image sequences, we also generated data where the image intensities are linearly scaled over time. We used a scaling factor $\lambda_k = 1 - 0.001k$, i.e. the intensities are linearly decreased by 10% for every 100 time points (we have determined this value based on the real image data). In addition, to introduce a random component, the scaling factor λ_k was multiplied by a uniformly distributed random factor ρ_k ($0.95 \leq \rho_k \leq 1.05$) which is determined for each time point independently. We also added Gaussian noise with $\sigma_n = 3$ voxel. The results for these synthetic image sequences can be found in Table II (last two columns). We can observe, that our diffeomorphic multi-frame approach outperforms pairwise registration for all values of N, for the 2D and 3D data. The best result is obtained for $N = 20$ for the 2D and 3D data, and the improvement is 38% and 22%, respectively. Interestingly, the error for the 3D data is smaller compared to the 2D data. The reason for this is that the linear intensity decrease is a *global* change of an image, and for 3D data there are more measurements compared to the 2D data. Therefore, the registration approach can cope better with the intensity changes.

To investigate whether the improvement of the diffeomorphic multi-frame approach compared to pairwise registration in Table II is statistically significant, we performed a Wilcoxon signed-rank test (non-parametric test). Prior application of a Shapiro-Wilk test showed that the data does not follow a normal distribution in all 12 cases (3 types of image sequences, 4 different values of N). Using a significance level of 5% we obtained $p < 0.016$. Thus, the multi-frame approach yields a statistically significant improvement compared to pairwise registration.

C. Experimental Results Using Real Image Data

We have also applied our diffeomorphic multi-frame registration approach to four 2D and two 3D real microscopy image data (cf. Sec. III-A above). We performed a quantitative evaluation based on manually determined structures in the nucleus channel of the image sequences. For each sequence of the 2D image data, positions of 9 spot-like structures within cell nuclei were determined for 38 up to 125 subsequent time points, and for the 3D data positions of 6 spot-like structures were determined for 28 up to 83 subsequent time points. For each structure and each time point we computed the registration error as the Euclidean distance of the current position (center of gravity) to its position at the first time point.

For our diffeomorphic multi-frame approach we used different numbers of multiple frames ($N = 3, 5, 10, 20$) and we

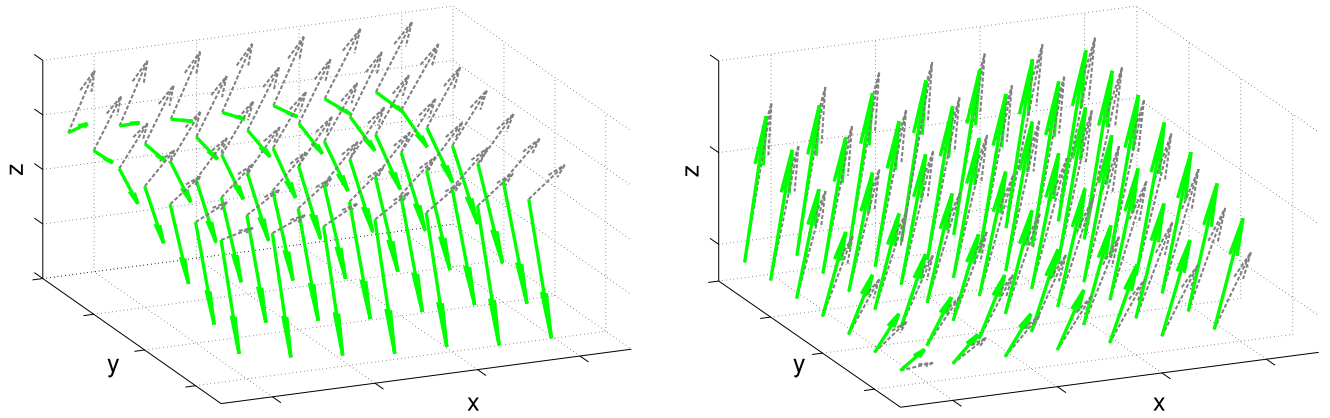


Fig. 5. Region of the ground truth inverse deformation vector field for the 3D synthetic sequence S5 with Gaussian noise for t=50 (dashed gray arrows) and computed inverse vector field (green arrows) using the diffeomorphic pairwise registration approach (left) and the diffeomorphic multi-frame registration approach (right).

TABLE III

REGISTRATION ERROR AND STANDARD DEVIATION FOR SPOT-LIKE STRUCTURES IN FOUR 2D REAL IMAGE SEQUENCES AND TWO 3D REAL IMAGE SEQUENCES. RESULTS FOR THE NEW DIFFEOMORPHIC MULTI-FRAME APPROACH ($N = 10$) AND ITS PAIRWISE VARIANT, AS WELL AS FOR OTHER APPROACHES. PERCENTAGES INDICATE THE CHANGE COMPARED TO THE CORRESPONDING PAIRWISE REGISTRATION APPROACH, AND BOLD VALUES INDICATE THE LOWEST ERROR

[pixel] / [voxel]	2D Sequences						3D Sequences			
	$\overline{e}_{mean}$					$\overline{\sigma}_{e_{mean}}$	$\overline{e}_{mean}$			$\overline{\sigma}_{e_{mean}}$
	A	B	C	D	Average	Average	E	F	Average	Average
Unregistered	6.56	13.97	26.65	27.84	18.76	9.53	24.94	21.24	23.09	11.69
Pairwise weighting [13]	3.23	2.87	4.39	7.08	4.39	2.49	10.01	7.15	8.58	5.15
Multi-frame weighting [14]	2.99	2.45	3.56	5.31	$3.58_{-19\%}$	$2.11_{-15\%}$	6.98	4.63	$5.81_{-32\%}$	$3.75_{-27\%}$
Log-Demons [34] extension	2.78	2.93	7.83	11.14	6.17	3.54	17.51	12.66	15.08	8.87
Temporal groupwise [24]	4.78	8.38	17.20	21.29	12.91	4.69	-	-	-	-
New diffeomorphic pairwise	2.20	1.94	3.27	4.99	3.10	2.06	9.03	5.02	7.03	4.55
New diffeomorphic multi-frame	**2.10**	**1.38**	**2.19**	**2.93**	$\mathbf{2.15}_{-31\%}$	$\mathbf{1.64}_{-20\%}$	**4.90**	**2.41**	$\mathbf{3.66}_{-48\%}$	$\mathbf{2.66}_{-41\%}$

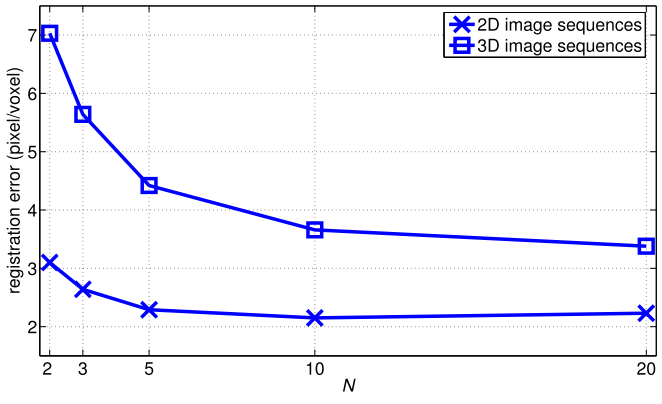
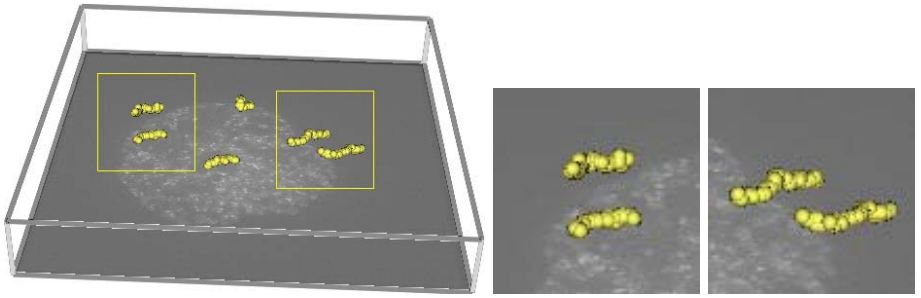


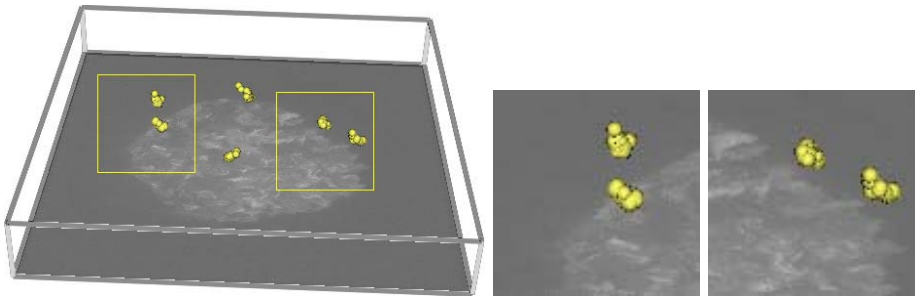
Fig. 6. Mean registration error ($\overline{e}_{mean}$) averaged over all spot-like structures of the four 2D and the two 3D real image sequences as a function of N.

computed for each spot-like structure the mean registration error e_{mean} over all time points. The results for the mean error averaged over all spot-like structures for the four 2D and the two 3D image sequences ($\overline{e}_{mean}$) are shown in Fig. 6 as a function of N. It can be seen, that for the 2D data the multi-frame approach with $N = 3$ yields a lower error than the pairwise approach, and the error is further reduced by increasing N up to $N = 10$, where the error reaches its minimum. For $N = 20$ the error is slightly higher. For the 3D data the error decreases with increasing N up to $N = 20$.

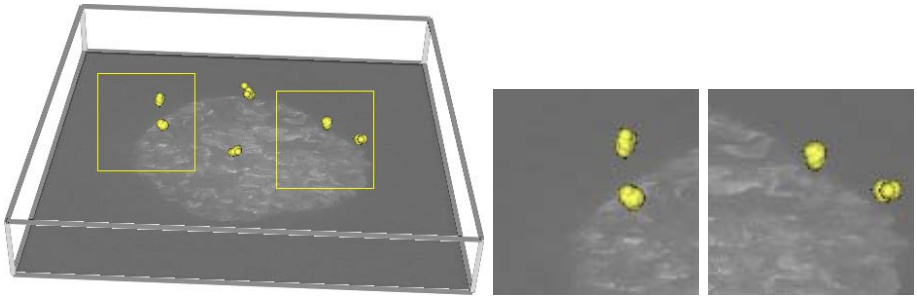
In Table III we show the results for $\overline{e}_{mean}$ for each 2D and each 3D real image sequence (image sequences A-F, see Table I and Sec. III-A) using the diffeomorphic multi-frame registration approach ($N = 10$) and the pairwise approach. We also computed the standard deviation of the mean error $\sigma_{e_{mean}}$ for each spot-like structure as well as the average over all structures and all 2D as well as all 3D image sequences ($\overline{\sigma}_{e_{mean}}$). It turned out that our multi-frame approach yields lower errors compared to pairwise registration, and the improvement for $\overline{e}_{mean}$ and $\overline{\sigma}_{e_{mean}}$ averaged over all four 2D sequences is 31% and 20%, respectively. For the 3D data, the improvement for $\overline{e}_{mean}$ and $\overline{\sigma}_{e_{mean}}$ averaged over the two 3D image sequences is 48% and 41%, respectively. In Fig. 7 we have visualized the positions of the spot-like structures (used as ground truth) over time for the 3D image sequence E for the unregistered case as well as for registration using our multi-frame approach and the pairwise approach. It can be seen, that the displacements of the positions over time w.r.t. the ground truth positions at the first frame (which represent the registration error) for the multi-frame approach



(a) Unregistered



(b) New diffeomorphic pairwise approach



(c) New diffeomorphic multi-frame approach

Fig. 7. Positions over time for 6 spot-like structures overlaid with a MIP image ($t = 50$) from the 3D image sequence E: (a) Original image and positions over time for the unregistered case, (b), (c) registered images and positions over time for the new diffeomorphic multi-frame approach and its pairwise variant. Enlarged sections for the marked regions are shown on the right.

are much smaller compared to the unregistered case and smaller compared to pairwise registration.

In addition, we have performed an experimental comparison with previous non-rigid registration approaches for the 2D and 3D data. We have used the pairwise weighting approach of Kim *et al.* [13] and the multi-frame weighting approach of Tektonidis *et al.* [14] (with $N = 10$). Both approaches are based on local optic flow estimation and have been previously used for registration of cell microscopy images. Moreover, we have used the log-domain diffeomorphic demons registration approach of Vercauteren *et al.* [34] (denoted with Log-Demons) which was previously used for pairwise registration of static medical images (MR images of the brain). We have extended this approach for application to temporal images using a pairwise consecutive scheme (analogously to the pairwise variant of our diffeomorphic approach) and have included it in our comparison. For regularization of the vector fields, we used a Gaussian kernel with $\sigma_G = (2, 2, 1)$ voxel for the update velocity fields $d\mathbf{v}_{k,k-1}^{(i)}$ as well as the velocity fields $\mathbf{v}_{k,k-1}^{(i)}$, and $\sigma_G = (1, 1, 0.5)$ voxel for the inverse deformation fields $\mathbf{u}_{l,k-1}$ as well as the deformation fields $\mathbf{u}_{k,1}$. We tested different parameters for the Gaussian kernel and used the ones which yielded the best results. For the evaluation based on 2D data we have also applied the temporal groupwise diffeomorphic registration approach of Metz *et al.* [24] which was previously used for the registration of medical images (CT, MR, and US images of the heart and the lung). With this approach, all images of an image sequence are registered simultaneously by minimizing the sum of squared intensity differences between all images and the mean image of the whole sequence. For our comparison, we used the implementation of the approach in elastix [42]. We used a grid spacing of 16×16 pixel, each time point of the sequences, and three multiresolution levels. We tested different values for the grid spacing and used the grid with the best result. An overview of the investigated non-rigid registration approaches and their main characteristics is given in Table IV. In Fig. 8 (left) we show an example of the registration error for one spot-like structure in the 2D real image sequence C using our diffeomorphic multi-frame approach and its pairwise variant

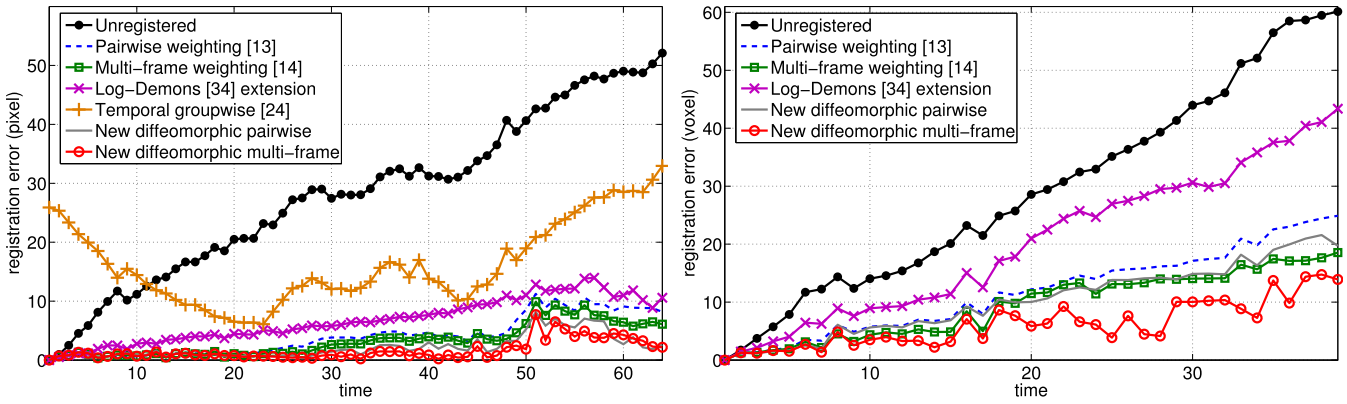


Fig. 8. Registration error for a spot-like structure in the 2D real image sequence C (left) and the 3D real image sequence E (right) as a function of time for the new diffeomorphic multi-frame approach (N=10) and its pairwise variant as well as for other registration approaches.

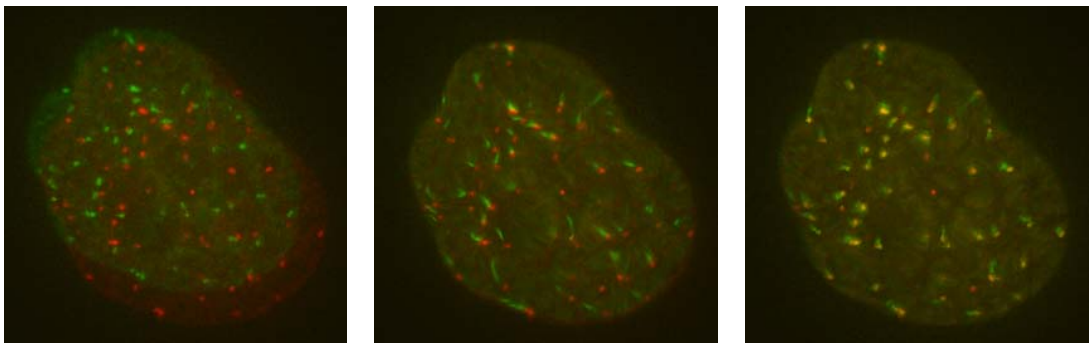
TABLE IV
NON-RIGID REGISTRATION APPROACHES INVESTIGATED IN THE EXPERIMENTS

Non-rigid registration approaches	Temporal registration	Spatial optimization	Diffeo-morphic	Flow boundary preserving regularization
Pairwise weighting [13]	pairwise	local	No	No
Multi-frame weighting [14]	multi-frame	local	No	No
Log-Demons [34] extension	pairwise	global	Yes	No
Temporal groupwise [24]	groupwise	global	Yes	No
New diffeomorphic pairwise	pairwise	local	Yes	Yes
New diffeomorphic multi-frame	multi-frame	local	Yes	Yes

in comparison to the four previous approaches. It can be seen, that the temporal groupwise approach [24] yields the largest errors and a mean error of e_{mean}= 16.4 pixel. For time points in about the middle of the sequence, [24] yields the lowest errors, which is what we expect since the images at these time points are most similar to the employed mean image. The mean errors for the Log-Demons [34] extension, the pairwise weighting approach [13], and the multi-frame weighting approach [14] are e_{mean}= 6.73 pixel, e_{mean}= 4.09 pixel, and e_{mean}= 3.26 pixel, respectively. Our diffeomorphic multi-frame approach yields the lowest error e_{mean}= 1.63 pixel which is smaller than e_{mean}= 2.22 pixel for the pairwise variant. When averaging over all spot-like structures of all four 2D real image sequences our diffeomorphic multi-frame registration approach yields the lowest values for $\overline{e}_{mean}$ and $\overline{\sigma}_{e_{mean}}$ compared to the four previous approaches, and the improvement is 45% and 25%, respectively, compared to the lowest error obtained by the four previous approaches (cf. Table III). In Fig. 8 (right) we show an example of the registration error for one spot-like structure in the 3D real image sequence E using our approach in comparison to three previous approaches. Also in this example our diffeomorphic multi-frame approach yields the lowest error e_{mean}= 6.19 voxel, and the pairwise variant yields e_{mean}= 10.3 voxel. In comparison, the Log-Demons [34] extension, the pairwise weighting approach [13], and the multi-frame weighting approach [14] yield e_{mean}= 20.1 voxel, e_{mean}= 11.8 voxel, and e_{mean}= 9.53 voxel, respectively. When averaging over all spot-like structures of the two 3D real image sequences ($\overline{e}_{mean}$, $\overline{\sigma}_{e_{mean}}$) our approach yields an improvement of 37% and 31%, respectively, compared to the lowest error obtained by the three previous approaches (cf. Table III). To investigate whether the improvement of the diffeomorphic multi-frame approach compared to the previous registration approaches as well as compared to pairwise registration for the 2D and 3D real data is statistically significant, we performed Wilcoxon signed-rank tests (non-parametric test). Using a significance level of 5% we obtained $p < 0.016$ for all tests. Thus, the multi-frame approach yields a statistically significant improvement compared to the previous registration approaches as well as compared to pairwise registration.

In Fig. 9 we show a comparison of results for our diffeomorphic multi-frame approach and the Log-Demons [34] extension for the 2D real image sequence D (time points 1 and 20). It can be seen, that the alignment of the nucleus is significantly improved compared to the unregistered case for both approaches. However, the alignment of subcellular structures in the inner part of the nucleus is better for our approach. The reason is that the local optimization employed in our approach copes better with the strong local structural changes compared to the global optimization used in [34].

Finally, we determined the computation time for registration of the first 100 images of the 2D real image sequence B (384 × 384 pixel) and the first 50 images of the 3D image sequence F (512 × 512 × 10 voxel). We compared our dif-



(a) Unregistered (b) Log-Demons [34] extension (c) New diffeomorphic multi-frame

Fig. 9. Overlays of images for time points 1 (red) and 20 (green) of the 2D real image sequence D for (a) the unregistered images, and for the registered images using (b) the Log-Demons [34] extension and (c) the new diffeomorphic multi-frame approach ($N = 10$). Yellow indicates overlapping intensities.

TABLE V

COMPUTATION TIME FOR THE REGISTRATION OF THE FIRST 100 IMAGES FROM THE 2D REAL IMAGE SEQUENCE B USING THE NEW DIFFEOMORPHIC MULTI-FRAME WEIGHTING APPROACH COMPARED TO THE MULTI-FRAME WEIGHTING APPROACH [14] FOR DIFFERENT VALUES OF N. PERCENTAGES INDICATE THE INCREASE COMPARED TO PAIRWISE REGISTRATION. THE PERCENTAGES IN THE LAST COLUMN INDICATE THE DIFFERENCE OF THE COMPUTATION TIME BETWEEN THE DIFFEOMORPHIC MULTI-FRAME APPROACH AND THE APPROACH IN [14]

		Multi-frame weighting [14]	New diffeomorphic multi-frame	Difference
pairwise		0 h 40	1 h 36	$+140\%$
multi-frame	$N = 3$	1 h $16_{+90\%}$	1 h $37_{+1\%}$	$+27\%$
	$N = 5$	2 h $04_{+210\%}$	1 h $47_{+11\%}$	-14%
	$N = 10$	4 h $26_{+565\%}$	2 h $11_{+36\%}$	-51%
	$N = 20$	8 h $23_{+1158\%}$	2 h $41_{+68\%}$	-68%

feomorphic multi-frame approach with its pairwise variant as well as with the pairwise and the multi-frame weighting approaches in [14]. For the 2D data the computation times can be found in Table V. For both multi-frame approaches the computation time is increasing with increasing N, however, the increase is significantly smaller for our diffeomorphic multi-frame approach compared to [14]. For example, the increase of the computation time for $N = 10$ is 565% for [14] and 36% for our approach, compared to the respective pairwise registration. We have determined the differences of computation times in percentage between our diffeomorphic multi-frame approach and [14] for each value of N (column "Difference" in Table V). For $N = 2$ (pairwise registration) the computation time of our approach is higher compared to [14], and the difference is $+140\%$. For $N = 3$ the difference decreases to $+27\%$, and further decreases with increasing N. For $N \geq 5$ the differences are negative, i.e., the computation time is lower for our approach compared to [14]. For the 3D data, our diffeomorphic multi-frame approach with $N = 10$ yields a computation time of 27 h 50 min. The change compared to pairwise registration is $+49\%$, and the change compared to [14] is -91%. We used a workstation under Linux with an Intel Xeon E5530 CPU (2.4 GHz). Note, that our implementation is currently not parallelized and further optimization is possible, thus the computation time can be further reduced.

IV. DISCUSSION

From the experiments based on synthetic as well as real 2D and 3D microscopy image sequences, it turned out that our diffeomorphic multi-frame non-rigid registration approach yields better results than previous approaches, in particular, a previous multi-frame registration approach [14]. There are two main reasons for the improved performance. First, our approach is diffeomorphic which guarantees that the determined transformations are invertible and smooth. This is important since our approach uses inverse transformations to construct a temporal mean image which is used to register single frames of the image sequence. In contrast, the approach in [14] uses a scheme for inversion of transformations which are not guaranteed to be invertible. Second, the employed flow boundary preserving smoothing method prevents over-smoothing of deformation vector fields, leading to more accurate registration results compared to standard Gaussian smoothing. In addition, our diffeomorphic multi-frame registration approach is computationally more efficient than the approach in [14], for two main reasons. First, our new approach is based on the minimization of the sum of squared intensity differences between each frame of an image sequence and a temporal mean image, in contrast to [14], where the sum of squared intensity differences between each frame and multiple previous frames are minimized. Second, in our approach the inverse transformations are efficiently computed using vector fields computed in the log-domain. From our experimental results we found that the reduction of computation time due to these two reasons is larger for multiple frames $N \geq 5$ than the increase in computation time due to the use of the log-domain diffeomorphic update rule (compared to the additive update rule in [14]) and due to the use of the flow boundary preserving smoothing (compared to Gaussian smoothing in [14]). We have also shown that our diffeomorphic multi-frame registration approach outperforms a pairwise temporal extension of the diffeomorphic registration approach in [34]. The reason is that our approach is based on local optic flow estimation which is advantageous for our application since the image data contain many local structural changes, compared to the global optimization in [34]. In addition, employing flow boundary preserving smoothing in

our approach prevents over-smoothing of deformation vector fields, compared to Gaussian smoothing in [34]. Compared to a temporal groupwise approach [24] which was previously used for registration of medical images, our approach yields better results. The reason is, that the simultaneous registration of all frames of an image sequence with large temporal structural changes as in our case is disadvantageous since the frames of an image sequence differ significantly, and the difference generally increases with time (e.g., compared to medical images of periodic movements such as respiratory motion).

V. Summary

We have presented a diffeomorphic multi-frame approach for temporal non-rigid registration of live cell microscopy images. Our registration approach computes diffeomorphic transformations based on local optic flow estimation and exploitation of information from multiple consecutive images. The determination of diffeomorphic transformations in the log-domain allows efficient computation of the inverse transformations which are used to construct a temporally weighted mean image. The use of a temporal mean image to register single frames of an image sequence significantly reduces the computation time compared to a previous multi-frame registration approach. For regularization of vector fields, we employed a flow boundary preserving method which prevents from over-smoothing of deformation vector fields. Using synthetic as well as real 2D and 3D dynamic microscopy image sequences, we have investigated the performance of our approach and based on a quantitative comparison we found that our approach outperforms pairwise registration, a previous multi-frame registration approach, a temporal pairwise extension of a diffeomorphic registration approach, and a temporal groupwise registration approach. In future work, we plan to reduce the computation time of our approach and apply it to other biological applications.

Acknowledgments

We thank Il-Han Kim for discussions. This work benefited from the use of ITK [43], elastix [42], and the ITK implementation of the log-domain diffeomorphic demons algorithm [44].

References

[1] B. Rieger, C. Molenaar, R. W. Dirks, and L. J. van Vliet, "Alignment of the cell nucleus from labeled proteins only for 4D *in vivo* imaging," *Microscopy Res. Tech.*, vol. 64, no. 2, pp. 142–150, Jun. 2004.

[2] C. A. Wilson and J. A. Theriot, "A correlation-based approach to calculate rotation and translation of moving cells," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1939–1951, Jul. 2006.

[3] P. Matula, P. Matula, M. Kozubek, and V. Dvorak, "Fast point-based 3-D alignment of live cells," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2388–2396, Aug. 2006.

[4] W. H. De Vos, G. H. Joss, W. Haffmans, R. A. Hoebe, E. M. M. Manders, and P. Van Oostveldt, "Four-dimensional telomere analysis in recordings of living human cells acquired with controlled light exposure microscopy," *J. Microscopy*, vol. 238, no. 3, pp. 254–264, Jun. 2010.

[5] O. Dzyubachyk *et al.*, "Automated analysis of time-lapse fluorescence microscopy images: From live cell images to intracellular foci," *Bioinformatics*, vol. 26, no. 19, pp. 2424–2430, 2010.

[6] S. Li, J. Wakefield, and J. A. Noble, "Automated segmentation and alignment of mitotic nuclei for kymograph visualisation," in *Proc. IEEE Int. Symp. Biomed. Imag., Nano Macro (ISBI)*, Mar. 2011, pp. 622–625.

[7] M. van de Giessen *et al.*, "Fully automated attenuation measurement and motion correction in FLIP image sequences," *IEEE Trans. Med. Imag.*, vol. 31, no. 2, pp. 461–473, Feb. 2012.

[8] K. Klein *et al.*, "Label-free live-cell imaging with confocal Raman microscopy," *Biophys. J.*, vol. 102, no. 2, pp. 360–368, Jan. 2012.

[9] J. Mattes, J. Nawroth, P. Boukamp, R. Eils, and K. M. Greulich-Bode, "Analyzing motion and deformation of the cell nucleus for studying co-localizations of nuclear structures," in *Proc. 3rd IEEE Int. Symp. Biomed. Imag., Nano Macro (ISBI)*, Apr. 2006, pp. 1044–1047.

[10] J. De Vylder, W. H. De Vos, E. M. M. Manders, and W. Philips, "2D mapping of strongly deformable cell nuclei-based on contour matching," *Cytometry A*, vol. 79A, no. 7, pp. 580–588, Jul. 2011.

[11] D. V. Sorokin, M. Tektonidis, K. Rohr, and P. Matula, "Non-rigid contour-based temporal registration of 2D cell nuclei images using the Navier equation," in *Proc. IEEE 11th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2014, pp. 746–749.

[12] S. Yang *et al.*, "Nonrigid registration of 3-D multichannel microscopy images of cell nuclei," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 493–499, Apr. 2008.

[13] I.-H. Kim, Y.-C. M. Chen, D. L. Spector, R. Eils, and K. Rohr, "Nonrigid registration of 2-D and 3-D dynamic cell nuclei images for improved classification of subcellular particle motion," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 1011–1022, Apr. 2011.

[14] M. Tektonidis, I.-H. Kim, Y.-C. M. Chen, R. Eils, D. Spector, and K. Rohr, "Non-rigid multi-frame registration of cell nuclei in live cell fluorescence microscopy image data," *Med. Image Anal.*, vol. 19, no. 1, pp. 1–14, Jan. 2015.

[15] L. Hogrebe *et al.*, "Serial section registration of axonal confocal microscopy datasets for long-range neural circuit reconstruction," *J. Neurosci. Methods*, vol. 207, no. 2, pp. 200–210, Jun. 2012.

[16] A. Akselrod-Ballin, D. Bock, R. C. Reid, and S. K. Warfield, "Accelerating image registration with the Johnson–Lindenstrauss Lemma: Application to imaging 3-D neural ultrastructure with electron microscopy," *IEEE Trans. Med. Imag.*, vol. 30, no. 7, pp. 1427–1438, Jul. 2011.

[17] C.-W. Wang, E. B. Gosno, and Y.-S. Li, "Fully automatic and robust 3D registration of serial-section microscopic images," *Sci. Rep.*, vol. 5, Oct. 2015, Art. no. 15051, doi: 10.1038/srep15051.

[18] H. Peng *et al.*, "BrainAligner: 3D registration atlases of Drosophila brains," *Nature Methods*, vol. 8, no. 6, pp. 493–498, 2011.

[19] L. Qu, F. Long, and H. Peng, "3-D registration of biological images and models: Registration of microscopic images and its uses in segmentation and annotation," *IEEE Signal Process. Mag.*, vol. 32, no. 1, pp. 70–77, Jan. 2015.

[20] G. K. Rohde, A. J. S. Ribeiro, K. N. Dahl, and R. F. Murphy, "Deformation-based nuclear morphometry: Capturing nuclear shape variation in HeLa cells," *Cytometry A*, vol. 73A, no. 4, pp. 341–350, Apr. 2008.

[21] G. R. Johnson, T. E. Buck, D. P. Sullivan, G. K. Rohde, and R. F. Murphy, "Joint modeling of cell and nuclear shape variation," *Molecular Biol. Cell*, vol. 26, no. 22, pp. 4046–4056, 2015.

[22] A. P. Goobic, J. Tang, and S. T. Acton, "Image stabilization and registration for tracking cells in the microvasculature," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 2, pp. 287–299, Feb. 2005.

[23] S.-A. Raza *et al.*, "RAMTaB: Robust alignment of multi-tag bioimages," *PLoS ONE*, vol. 7, no. 2, p. e30894, 2012.

[24] C. T. Metz, S. Klein, M. Schaap, T. van Walsum, and W. J. Niessen, "Nonrigid registration of dynamic medical imaging data using nD + t B-splines and a groupwise optimization approach," *Med. Image Anal.*, vol. 15, no. 2, pp. 238–249, Apr. 2011.

[25] M. Yigitsoy, C. Wachinger, and N. Navab, "Temporal groupwise registration for motion modeling," in *Proc. Int. Conf. Inf. Process. Med. Imag. (IPMI)*, vol. 6801. 2011, pp. 648–659.

[26] M. De Craene *et al.*, "Temporal diffeomorphic free-form deformation: Application to motion and strain estimation from 3D echocardiography," *Med. Image Anal.*, vol. 16, no. 2, pp. 427–450, Feb. 2012.

[27] G. Wu, Q. Wang, J. Lian, and D. Shen, "Estimating the 4D respiratory lung motion by spatiotemporal registration and super-resolution image reconstruction," *Med. Phys.*, vol. 40, no. 3, p. 031710, 2013.

[28] S. Durrleman, X. Pennec, A. Trouvé, J. Braga, G. Gerig, and N. Ayache, "Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data," *Int. J. Comput. Vis.*, vol. 103, no. 1, pp. 22–59, May 2013.

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

[30] Z. Zhang, M. Ashraf, D. J. Sahn, and X. Song, "Temporally diffeomorphic cardiac motion estimation from three-dimensional echocardiography by minimization of intensity consistency error," *Med. Phys.*, vol. 41, no. 5, p. 0529020, 2014.
[31] W. Huizinga *et al.*, "PCA-based groupwise image registration for quantitative MRI," *Med. Image Anal.*, vol. 29, pp. 65–78, Apr. 2016.
[32] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *Int. J. Comput. Vis.*, vol. 61, no. 2, pp. 139–157, Feb. 2005.
[33] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, pp. S61–S72, Mar. 2009.
[34] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Symmetric log-domain diffeomorphic registration: A demons-based approach," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervent (MICCAI)*, 2008, pp. 754–761.
[35] V. Arsigny, O. Commowick, X. Pennec, and N. Ayache, "A Log-Euclidean framework for statistics on diffeomorphisms," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervent (MICCAI)*, 2006, pp. 924–931.
[36] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 1981, pp. 674–679.
[37] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: A unifying framework," *Int. J. Comput. Vis.*, vol. 56, no. 3, pp. 221–255, 2004.
[38] M. Bossa, M. Hernandez, and S. Olmos, "Contributions to 3D diffeomorphic atlas estimation: Application to brain images," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervent (MICCAI)*, 2007, pp. 667–674.
[39] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.
[40] D. Sun, S. Roth, and M. J. Black, "A quantitative analysis of current practices in optical flow estimation and the principles behind them," *Int. J. Comput. Vis.*, vol. 106, no. 2, pp. 115–137, 2014.
[41] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *Int. J. Comput. Vis.*, vol. 92, no. 1, pp. 1–31, 2011.
[42] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 196–205, Jan. 2010.
[43] L. Ibáñez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*, Clifton Park, NY, USA: Kitware Inc., 2005.
[44] F. Dru, P. Fillard, and T. Vercauteren, "An ITK implementation of the symmetric log-domain diffeomorphic demons algorithm," *Insight J.*, Jun. 2009.

Marco Tektonidis received the Diploma degree in computer science from the Technical University of Berlin, Germany, in 2008.

He is currently pursuing the Ph.D. degree with the Biomedical Computer Vision Group, University of Heidelberg, and the German Cancer Research Center. From 2007 to 2009, he was with the Department for Innovative Methods of Computing, Technical University of Dresden, Germany. Since 2014, he has been with the French–German Research Institute of Saint–Louis, France.

Karl Rohr received the Dipl.Ing. degree in electrical engineering from Karlsruhe Institute of Technology (KIT), Germany, and the Ph.D. and Habilitation degrees in computer science from the University of Hamburg, Germany, in 1994 and 1999, respectively. He is currently an Associate Professor and the Head of the Biomedical Computer Vision Group, University of Heidelberg, and the German Cancer Research Center. From 2000 to 2004, he was an Associate Professor with the School of Information Technology, International University in Germany. In 1999, he spent a research stay with the Surgical Planning Laboratory, Harvard Medical School, Boston, MA, USA. From 2007 to 2010, he was a Guest Professor with the International University in Germany. Since 2007, he has been a member of the Excellence Cluster CellNetworks, University of Heidelberg.

He has authored the book *Landmark-Based Image Analysis* (Kluwer Academic Publishers, 2001) covering landmark localization and elastic registration. He has published over 250 peer-reviewed scientific articles. His research interests are in biomedical image analysis with a focus on nonrigid image registration, cell and particle tracking, vessel segmentation, and landmark localization. He has been a Program and Review Committee Member of a number of international conferences and workshops. He was a Co-Organizer of the International Workshop on Microscopic Image Analysis with Applications in Biology in 2008, 2009, 2011, and he was a Program Chair of the IEEE International Symposium on Biomedical Imaging 2016. For his research work he was awarded several prizes. He was an Associate Editor of the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING from 2007 to 2012 and served the Editorial Board of the journal *Pattern Recognition* from 2000 to 2006.