Fast and Robust Symmetric Image Registration Based on Intensity and Spatial Information

Johan Öfverstedt, Joakim Lindblad, Member, IEEE, and Nataša Sladoje, Member, IEEE

Abstract—Intensity-based image registration approaches rely on similarity measures to guide the search for geometric correspondences with high affinity between images. The properties of the used measure are vital for the robustness and accuracy of the registration. In this study a symmetric, intensity interpolation-free, affine registration framework based on a combination of intensity and spatial information is proposed. The excellent performance of the framework is demonstrated on a combination of synthetic tests, recovering known transformations in the presence of noise, and real applications in biomedical and medical image registration, for both 2D and 3D images. The method exhibits greater robustness and higher accuracy than similarity measures in common use, when inserted into a standard gradient-based registration framework available as part of the open source Insight Segmentation and Registration Toolkit (ITK). The method is also empirically shown to have a low computational cost, making it practical for real applications. Source code is available.

Index Terms—Image registration, set distance, gradient methods, optimization, cost function, iterative algorithms, fuzzy sets, magnetic resonance imaging, transmission electron microscopy.

I. INTRODUCTION

Image registration [1]–[3] is the process of establishing correspondences between images and a reference space, such that the contents of the images have a high degree of affinity in the reference space. An example of such correspondence is a mapping of an image (often referred to as floating image) of a brain to a reference space of another image (often referred to as reference image) of a brain where their important structures are well co-localized. There are two main categories of approaches for image registration: feature-based methods extract a set of feature points between which a correspondence is found, whereas intensity-based methods use the voxel-values directly, and evaluate candidate mappings based on a similarity measure (affinity). There are also two main categories of transformation models: linear (which include, as special cases, rigid, similarity, and affine transformations), and non-linear (deformable). The combination of differentiable transformation models and differentiable similarity measures enables the use of gradient-based local optimization methods.

Medical and biomedical image registration, [4], is an important branch of general image registration and a lot of effort has been invested over the last decades to refine the tools and techniques, [5]. Although a majority of the recent research has been devoted to non-linear registration techniques, the most prevalent registration method used in the clinic is still linear registration. In a number of situations, the deformations allowed by non-linear registration can be difficult to evaluate and may affect reliability of diagnosis, [6]; hence, physicians may prefer a more constrained rigid or affine alignment. Considering their wide usage as fundamental tools, improvement of rigid and affine registration in terms of performance and reliability is highly relevant in practice.

Feature-based image registration is dependent on the ability of the feature extraction method to locate distinct points of interest appearing in both (all) images. Feature-extractors (e.g. SIFT [5]) typically presuppose the existence and relevance of specific local characteristics such as edges, corners and other salient features; if no, or too few, such distinct points are found, the registration will fail. This is often the case in medical and biomedical applications, [6], [7], where intensity-based registration, therefore, tends to be the method of choice. Figure 1 shows an illustrative example of a biomedical application where feature-based method fails, whereas intensity-based method can be successful.

Intensity-based registration is, in general, formulated as a non-convex optimization problem. The similarity measures commonly used as optimization criteria typically exhibit a high number of local optima [8], [9]; a count which tends to rapidly increase under noisy conditions. A small region of attraction of a global optimum imposes that the starting position has to be set very close to the optimal solution for it to be found by an optimizer. This leads to reliability challenges for automated solutions.

In this study we develop a registration framework based on a family of symmetric distance measures, proposed in [9], which combine intensity and spatial information in a single measure. These measures have been shown to be characterized by smooth distance surfaces with significantly fewer local minima than the commonly used intensity-based measures, when studied in the context of template matching and object recognition. In this work we demonstrate that these distance measures can, with a slight modification, be successfully used for fast and robust affine image registration. By differentiating the distance measure we are able to use efficient gradient-based optimization. The proposed method outperforms the commonly used similarity measures in both synthetic and real scenarios of medical and biomedical registration tasks, which we confirm by (i) landmark-based evaluation on transmission electron microscopy (TEM) images of cilia [10], with the aim of improving multi-image super-resolution reconstruction, as well as (ii) evaluation on the task of atlas-based segmentation of magnetic resonance (MR) images of brain, on the LPBA40-dataset [11].
Intensity interpolation is typically a required tool in the context of intensity-based registration performed with commonly used similarity measures since the sought transformation (and intermediate candidates) is likely to map points to regions outside of the regular grid. Treating the reference and floating images differently in terms of the interpolation introduces a significant source of asymmetry [12] and can cause a registration task to succeed or fail depending on which image is selected as reference and which is floating. Our proposed approach requires no off-grid intensity values, and is interpolation-free in terms of intensities; empirical tests confirm that it is highly symmetric in practice.

Noting that intensity-based image registration can be computationally demanding, we also demonstrate that the proposed measure is fast to compute, in comparison with the implementations of the measures existing in the ITK-framework [12]. The registration framework, with the proposed measure, is implemented in C++/ITK, and its source code is available.

II. PRELIMINARIES AND PREVIOUS WORK

A. Images as Fuzzy Sets

First we recall a few basic concepts related to fuzzy sets [13], a theoretical framework where gray-scale images are conveniently represented.

A fuzzy set \( S \) on a reference set \( X_S \) is a set of ordered pairs, \( S = \{(x, \mu_S(x)) : x \in X_S\} \), where \( \mu_S : X_S \rightarrow [0, 1] \) is the membership function of \( S \). Where there is no risk for confusion, we equate the set and its membership function and let \( S(x) \) be equivalent to \( \mu_S(x) \).

A gray-scale image can directly be interpreted as a spatial fuzzy set by rescaling the valid intensity range to \([0, 1]\). We assume, w.l.o.g., that the images to be registered have an intensity range \([0, 1]\) and we directly interpret them as fuzzy sets defined on a reference set which is the image domain, and is in most cases a subset of \( \mathbb{Z}^n \). We use the terms image and fuzzy set interchangeably in this text.

A crisp set \( C \subseteq X_C \) (a binary image) is a special case of a fuzzy set, with its characteristic function as membership function

\[
\mu_C(x) = \begin{cases} 
1, & \text{for } x \in C \\
0, & \text{for } x \notin C .
\end{cases}
\]

The height of a fuzzy set \( S \subseteq X_S \) is \( h(S) = \max_{x \in X_S} \mu_S(x) \). The complement \( \overline{S} \) of a fuzzy set \( S \) is \( \overline{S} = \{(x, 1 - \mu_S(x)) : x \in X_S\} \). An \( \alpha \)-cut of a fuzzy set \( S \) is a crisp set defined as \( \alpha S = \{x \in X_S : \mu_S(x) \geq \alpha\} \), i.e., a thresholded image.

Let \( p \) be an element of the reference set \( X_S \). A fuzzy point \( p \) (also called a fuzzy singleton) defined at \( p \in X_S \) with height \( h(p) \), is defined by a membership function

\[
\mu_p(x) = \begin{cases} 
h(p), & \text{for } x = p \\
0, & \text{for } x \neq p .
\end{cases}
\]

B. Intensity-Based Registration and Point-Wise Distances

Intensity-based registration is a general approach to image registration defined as a minimization process, where a distance measure between the intensities of overlapping points (or regions) is used as optimization criterion. Given a distance measure \( d \) and a set of valid transformations \( \Omega \), intensity-based registration of two images \( A \) (floating) and \( B \) (reference) can be formulated as the optimization problem,

\[
\hat{T} = \arg \min_{T \in \Omega} d(T(A), B),
\]

where \( T(A) \) denotes a valid transform of image \( A \) into the reference space of image \( B \).

Intensity-based similarity/distance measures which are most commonly used for image registration are Sum of Squared Differences (SSD) [14], Pearson Correlation Coefficient (PCC) and Mutual Information (MI) [15]. These measures are point-based, i.e. they are functions of the intensities of points belonging to the overlapping regions of the two compared sets. Their evaluation, therefore, typically requires interpolation of image intensities.

For two images \( P \) and \( Q \) defined on a common reference set \( X_{P,Q} \) of overlapping points, these measures are defined as

\[
\text{SSD}(P, Q) = \sum_{v \in X_{P,Q}} (P(v) - Q(v))^2,
\]

\[
\text{PCC}(P, Q) = \frac{\sum_{v \in X_{P,Q}} (P(v) - \text{avg}(P))(Q(v) - \text{avg}(Q))}{\sqrt{\sum_{v \in X_{P,Q}} (P(v) - \text{avg}(P))^2} \sqrt{\sum_{v \in X_{P,Q}} (Q(v) - \text{avg}(Q))^2}}
\]

and

\[
\text{MI}(P, Q) = H_P + H_Q - H_{P,Q}.
\]
The fuzzy point-to-set bidirectional distance \( d^\alpha \) is defined as
\[
d^\alpha (p, S) = d^\alpha (p, S) + \overline{d}^\alpha (p, S).
\]
which consider a random subset of the points in each optimization iteration, incurring a two-fold benefit: utilizing randomness to escape shallow local optima in the implicit distance surface, while also decreasing the computational work required per iteration. The size of the random subset is usually given as a fraction of the total number of points, and denoted as the sampling fraction. Approximation of the cost function by random subset sampling (where a new random subset of points is chosen in every iteration) has been, in previous studies, shown to perform well for intensity-based registration.

III. PROPOSED IMAGE REGISTRATION FRAMEWORK

A. Distances

To extend the family of distance measures given by \cite{15, 23}, to be suitable for registration, optionally with random subsampling optimization methods \cite{23}, we here define a new related family of distance measures.

Definition 1 (Asymmetric average minimal distance). Given fuzzy set \( \mathcal{A} \) on a reference set \( X_A \subset \mathbb{R}^n \), fuzzy set \( \mathcal{B} \) on reference set \( X_B \subset \mathbb{R}^n \), and a weight function \( w_A: X_A \rightarrow \mathbb{R}_{\geq 0} \), the asymmetric average minimal distance from \( \mathcal{A} \) to \( \mathcal{B} \), is

\[
\frac{1}{\mu_A(x)} \sum_{x \in X_A} w_A(x)d(\mathcal{A}(x), \mathcal{B})
\]

We consider point-to-set distances defined by \cite{11} or \cite{13}.

Building on the asymmetric distance, we formulate a symmetric distance as follows:

Definition 2 (Average minimal distance). Given fuzzy set \( \mathcal{A} \) on reference set \( X_A \subset \mathbb{R}^n \), fuzzy set \( \mathcal{B} \) on reference set \( X_B \subset \mathbb{R}^n \), weight functions \( w_A: X_A \rightarrow \mathbb{R}_{\geq 0} \) and \( w_B: X_B \rightarrow \mathbb{R}_{\geq 0} \), the average minimal distance between \( \mathcal{A} \) and \( \mathcal{B} \), is

\[
\frac{1}{2} \left( \frac{1}{\mu_A(x)} \sum_{x \in X_A} w_A(x)d(\mathcal{A}(x), \mathcal{B}) + \frac{1}{\mu_B(x)} \sum_{x \in X_B} w_B(x)d(\mathcal{B}(x), \mathcal{A}) \right)
\]

In the context of image registration, we utilize \( d_{AMD} \) to express a (weighted) distance between transformed fuzzy points \( T(\mathcal{A}(x)) \), and the image \( \mathcal{B} \), where the transformation of a fuzzy point \( \mathcal{A}(x) = \{ (x, \mu_A(x)) \} \) is given by the transformation of the reference point \( x \):

\[
T(\mathcal{A}(x)) = \{ (T(x), \mu_A(x)) \}
\]

To reflect the bounded image domain, only the transformed points falling on a predefined region \( M_B \subset \mathbb{R}^n \) are considered. Note that, when \( \mathcal{A} \) and \( \mathcal{B} \) are digital images, \( X_A \) and \( X_B \) are typically subsets of \( \mathbb{Z}^n \) and the transformed points \( T(x) \mid x \in X_A \) do not necessarily coincide with the points of the reference set \( X_B \); an illustrative example is given in Fig. 2.

We, therefore, provide the following definitions suited for the task of image registration:

Definition 3 (Asymmetric average minimal distance for image registration). Given fuzzy set \( \mathcal{A} \) on reference set \( X_A \subset \mathbb{R}^n \), fuzzy set \( \mathcal{B} \) on \( X_B \subset \mathbb{R}^n \), a weight function \( w_A: X_A \rightarrow \mathbb{R}_{\geq 0} \), and a crisp subset (mask) \( M_B \subset \mathbb{R}^n \), the asymmetric average minimal distance for image registration from \( \mathcal{A} \) to \( \mathcal{B} \), parameterized by a transformation \( T: X_A \rightarrow \mathbb{R}^n \), is

\[
\frac{1}{\mu_A(x)} \sum_{x \in \hat{X}} w_A(x)d(T(\mathcal{A}(x)), \mathcal{B})
\]

where \( \hat{X} = \{ x : x \in X_A \land T(x) \in M_B \} \).

Definition 4 (Average minimal distance for image registration). Given fuzzy set \( \mathcal{A} \) on reference set \( X_A \), fuzzy set \( \mathcal{B} \) on \( X_B \), weight functions \( w_A: X_A \rightarrow \mathbb{R}_{\geq 0} \) and \( w_B: X_B \rightarrow \mathbb{R}_{\geq 0} \), and crisp subsets (masks) \( M_A, M_B \subset \mathbb{R}^n \), the average minimal distance for image registration between \( \mathcal{A} \) and \( \mathcal{B} \), parameterized by an invertible transformation \( T: \mathbb{R}^n \rightarrow \mathbb{R}^n \),
with inverse $T^{-1}$, is defined as

$$d_{\text{AMD}}^{\alpha}(\mathcal{A},\mathcal{B};T,w_A,w_B,M_A,M_B) = \frac{1}{2} (d_{\text{AMD}}^{\alpha}(\mathcal{A},\mathcal{B};T,w_A,w_B,M_A) + d_{\text{AMD}}^{\alpha}(\mathcal{B},\mathcal{A};T^{-1},w_B,M_B)).$$

(21)

The distance $d_{\text{AMD}}^{\alpha}$ is based on full sampling, taking into account all points in the two sets which have non-zero weights, as long as they are transformed to points inside the mask associated with the other set. To reduce the computational cost of the distance and, in addition, to enable random iterative sampling, we propose an approximate version of $d_{\text{AMD}}^{\alpha}$:

**Definition 5** (Subsampled average minimal distance for image registration). Given fuzzy set $\mathcal{A}$ on reference set $X_A$, fuzzy set $\mathcal{B}$ on $X_B$, weight functions $w_A : X_A \rightarrow R_{\geq 0}$ and $w_B : R \rightarrow R_{\geq 0}$, and crisp subsets (masks) $M_A,M_B \subset R^n$, the subsampled average minimal distance for image registration between $\mathcal{A}$ and $\mathcal{B}$, parameterized by an invertible transformation $T : R^n \rightarrow R^n$, with inverse $T^{-1}$, and crisp sets $P_A \subseteq X_A$ and $P_B \subseteq X_B$, is defined as

$$\tilde{d}_{\text{AMD}}^{\alpha}(\mathcal{A},\mathcal{B};P_A,P_B,T,w_A,w_B,M_A,M_B) = \frac{1}{2} \left( d_{\text{AMD}}^{\alpha}(\mathcal{A} \cap P_A,\mathcal{B};T,w_A,w_B,M_A) + d_{\text{AMD}}^{\alpha}(\mathcal{B} \cap P_B,\mathcal{A};T^{-1},w_B,M_B) \right).$$

(22)

Inserting (11) or (13) as point-to-set distance in (20), and hence indirectly in (21) and (22), provides extensions of this family of distances to the $\alpha$-cut-based distances, which we denote $d_{\text{AMD}}^{\alpha},d_{\text{AMD}}^{\alpha \cdot} ,d_{\text{AMD}}^{\alpha \cdot},d_{\text{AMD}}^{\alpha \cdot \cdot},d_{\text{AMD}}^{\alpha \cdot}$ and $d_{\text{AMD}}^{\alpha \cdot \cdot \cdot}$.

Normalizing of the weights of the sampled points, introduced through Def. 1 renders the magnitude of the distance (and subsequently its derivatives) invariant to the size and aggregated weight of the sets or of the sampled subsets. Since the normalization is done separately from $\mathcal{A}$ to $\mathcal{B}$ and from $\mathcal{B}$ to $\mathcal{A}$, both directions are weighted equally even if the total weights of the point subsets from the two sets are different. This normalization can simplify the process of choosing e.g. step-length of various optimization methods, and makes it more likely that default hyper-parameter values can be found and reused across different applications.

**B. Registration**

We propose to utilize symmetric distances $d_{\text{AMD}}^{\alpha}$ and $\tilde{d}_{\text{AMD}}^{\alpha}$ as cost functions in (3) to define concrete image registration methods. Inserting $d_{\text{AMD}}^{\alpha}$ into (3) we obtain

$$\hat{T} = \arg \min_{T \in \Omega} d_{\text{AMD}}^{\alpha}(\mathcal{A},\mathcal{B};T,w_A,w_B,M_A,M_B).$$

(23)

For the case of subset sampling with sets $P_A$ and $P_B$, registration is defined as

$$\hat{T} = \arg \min_{T \in \Omega} \tilde{d}_{\text{AMD}}^{\alpha}(\mathcal{A},\mathcal{B};P_A,P_B,T,w_A,w_B,M_A,M_B).$$

(24)

By selection of new random subsets $P_A$ and $P_B$ in each iteration, various stochastic gradient descent optimization methods can be realized.

To solve the optimization problems stated in (23) and (24) with efficient gradient-based optimization methods, the partial derivatives of the distance measures with respect to the transformation parameters of $T$ are required.

**C. Gradients**

The derivative of (9), the crisp point-to-set distance measure $d(T(x),S)$ (in $n$-dimensional space), with respect to parameters $T_i$ of the transformation $T$ applied to a point $x \in X$, yielding $y = T(x)$, can be written as

$$\frac{\partial d}{\partial T_i} = \sum_{j=1}^{n} \frac{\partial d_j}{\partial T_i} \frac{\partial y_j}{\partial T_i}.$$ 

(25)

The values $\frac{\partial d}{\partial T_i}$ are the components (partial derivatives) of the gradient $\nabla d(y,S)$ of the point-to-set distance in point $y \in Y \subset R^n$, and are not dependent on the transformation model.

The gradient of the fuzzy point-to-set distance measure (11) is given by the integral over $\alpha$-cuts, of gradients of the (crisp) point-to-set distances:

$$\nabla d(x,S) = \int_{0}^{\infty} \nabla d(x,\alpha S) d\alpha.$$ 

(26)

**D. Algorithms for Digital Images on Rectangular Grids**

The distances and gradients can be computed efficiently for the special case of digital images on rectangular grids. For images quantized to $\ell \in N_{>0}$ non-zero discrete $\alpha$-levels the integrals in (11) and (26) are suitably replaced by sums. The number of quantization levels is typically taken to be a small constant; a choice of $\ell = 7$ non-zero equally spaced $\alpha$-levels has previously shown to provide a good trade-off between performance, speed and noise-sensitivity [9], and we keep it for all experiments.

We need a discrete operator to approximate the gradient of $d(x,S)$ for a set $S$ defined on a rectangular grid with spacing $s \in R_{>0}^n$. We propose to use the following difference operator providing a discrete approximation of $\nabla d(x,S)$:

$$\Delta d(x) = \gamma_x(\delta_1[d(x)],\ldots,\delta_n[d(x)]),$$

(27)

where

$$\delta_i[d(x)] = \frac{1}{2\alpha}(d(x + s_iu_i,S) - d(x - s_iu_i,S)),$$

(28)

$\gamma_x$ is an indicator function,

$$\gamma_x = \begin{cases} 1 & \text{for } d(x,S) \neq 0, \\ 0 & \text{for } d(x,S) = 0. \end{cases}$$

(29)

and $u_i$ is the unit-vector along the $i$-th dimension.

The indicator function $\gamma_x$ causes the operator $\Delta[S](x)$ to be zero-valued for points included in $S$ (i.e., where the distance transform is zero-valued). This prevents discretization issues along set boundaries, where the standard central difference operator yields non-zero gradients, which can cause the measure to overshoot a potential voxel-perfect overlap.

By creating tables for the distance and gradient sums for each image as a pre-processing step, using either of the procedures in Alg. 1 ($\Delta_0 DT$ for inwards distances and $\Delta_0 DT_{BD}$ for bidirectional distances), the distance and gradient may then be readily computed with Alg. 2. $|T|$ denotes the number
of parameters of the transformation, which is 6 for two-dimensional (2D) affine transformations, and 12 for three-dimensional (3D) affine transformations.

Algorithm 1 Distance and Gradient Maps

**Input:** Image $\mathcal{A}$. Mask $M_A$. Set of $\alpha$-levels $\{\alpha_1, \ldots, \alpha_\ell\}$.

**Output:** Stack of pre-computed distance sums $D[0 \ldots \ell]$, and corresponding discrete gradient approximations $G[0 \ldots \ell]$.

```plaintext
1: procedure $\Delta \alpha DT(\mathcal{A}, M_A, (\alpha_1, \ldots, \alpha_\ell), d_{MAX})$
2:     $\alpha_0 \leftarrow 0$, $D[0] \leftarrow 0$, $G[0] \leftarrow (0, 0, 0)$
3:     for $i = 1$ to $\ell$ do
4:         $DT_i \leftarrow DT^{\alpha_i}[\mathcal{A} \cap M_A]$, Compute DT of $\alpha$-cut.
5:         $DT_i \leftarrow \min(DT_i, d_{MAX})$
6:     for all $v \in X_A$ do
7:         $D[i](v) \leftarrow (\alpha_i - \alpha_{i-1})DT_i(v) + D_{i-1}(v)$
8:         $G[i](v) \leftarrow (\alpha_i - \alpha_{i-1})\Delta DT_i(v) + G_{i-1}(v)$
9:     end for
10: end for
11: return $D, G$
12: end procedure
```

Algorithm 2 Point-to-Set Distance and its Gradient w.r.t. $T$

**Input:** Fuzzy point $\mathcal{A}(v)$ and associated weight $w_A(v)$, fuzzy set $B$ and associated (binary) mask $M_B$, distances $D[0 \ldots \ell]$ and gradients $G[0 \ldots \ell]$; transformation $T$.

**Output:** Point-to-set distance $d$, derivatives $\frac{\partial d}{\partial T}$, weight $w$.

```plaintext
1: procedure $D_{PTS}(\mathcal{A}(v), B; T, w_A(v), M_B, D_B, G_B)$
2:     $v \leftarrow T(v)$
3:     if $v \in M_B$ then
4:         $h \leftarrow \text{QUANTIZE}(\mu_A(v))$
5:         $d, g \leftarrow \text{INTERPOLATE}(DB[h], GB[h], v)$
6:         for $i = 1$ to $|T|$ do
7:             $\frac{\partial d}{\partial T} \leftarrow \sum_{j=1}^{n} \frac{\partial g_j}{\partial T} g_j$
8:         end for
9:         return $w_A(v) d, w_A(v) \frac{\partial d}{\partial T}, w_A(v)$
10: else
11:     return $0, 0, 0$ \quad Zero dist., grad., and weight.
12: end if
13: end procedure
```

The procedures in Alg. 1 have linear run-time complexity $O((\ell + 1)|X_A|)$, achieved by using a linear-time algorithm for computing the distance transform (e.g. [24]) in line 2 of Alg. 1. The space complexity of the algorithm is $O((\ell + 1)|X_A|)$ and the $D, G$ structures must remain in memory to enable fast lookup in Alg. 2. Figure 3 shows an example of the distance and gradient of a sample $\alpha$-level. Alg. 2 computes the point-to-set distance and gradient w.r.t. the transformation using the pre-computed tables and has run-time complexity $O(|T| \alpha)$ thus being independent of $\ell$ and the size of $\mathcal{A}$ and $B$.

Algorithm 3 Symmetric Registration

**Input:** Fuzzy sets $\mathcal{A}, \mathcal{B}$, with binary masks $M_A, M_B$, and weight functions $w_A, w_B$, initial transformation guess $T$. No. of $\alpha$-levels $\ell$, distance saturation $d_{MAX}$, step-lengths $\lambda[1 \ldots N]$, and iteration count $\tau$ are hyper-parameters.

**Output:** Symmetric set distance $d$ ($d_{\alpha_{AMD}}$), estimated transformation $T$.

```plaintext
1: $D_A, G_A \leftarrow \Delta \alpha DT_{BD}(\mathcal{A}, M_A, (\alpha_1, \ldots, \alpha_\ell), d_{MAX})$
2: $D_B, G_B \leftarrow \Delta \alpha DT_{BD}(\mathcal{B}, M_B, (\alpha_1, \ldots, \alpha_\ell), d_{MAX})$
3: for $i = 1$ to $\tau$ do
4:     $d_1, \frac{\partial d_1}{\partial T} \leftarrow \sum_{v \in X_A} D_{PTS}(\mathcal{A}(v), B; T, w_A(v), M_B, D_B, G_B)$
5:     $d_2, \frac{\partial d_2}{\partial T} \leftarrow \sum_{v \in X_B} D_{PTS}(\mathcal{B}(v), A; T^{-1}, w_B(v), M_A, D_A, G_A)$
6:     $T \leftarrow T - \frac{\lambda_i}{2} \left( \frac{1}{w_1} \frac{\partial d_1}{\partial T} + \frac{1}{w_2} \frac{\partial d_2}{\partial T} \right)$
7: end for
8: $d \leftarrow \frac{1}{2} (\frac{1}{w_1} d_1 + \frac{1}{w_2} d_2)$ \quad Output final distance value.
9: $\hat{T} \leftarrow T$ \quad Output final transformation estimate.
```

Algorithm 3 performs a complete registration given two images, their binary masks, weight functions, and an initial transformation. Algorithm 3 completes $\tau$ full iterations, however other termination criteria may be beneficial (see Sec. IV). The registration consists of pre-processing, followed by a loop where the symmetric set distance and derivatives are computed and $T$ is updated. $\frac{\partial T^{-1}}{\partial T}$, in line 6 of Alg. 3 denotes a matrix $\begin{bmatrix} \frac{\partial T}{\partial T} \end{bmatrix}$ of partial derivatives of the parameters of the inverse transformation $T^{-1}$ w.r.t. the parameters of the forward transform $T$. This matrix relates the computed partial derivatives $\frac{\partial d}{\partial T}$ with the parametrization of the forward transform.

The overall run-time complexity is $O(N|T| \tau(|X_A| + |X_B|) + (\ell + 1)(|X_A| + |X_B|))$. Practical choices for $N$ tend to be in the range $[1000, 3000]$, depending on hyper-parameters (e.g. $\lambda$), and distance in parameter-space between starting position and the global optimum. The evaluation in Sec. V confirms empirically that convergence, according to (31) or (32), tends to be reached after 1000 to 3000 iterations, using an optimizer with a decaying $\lambda$.

The QUANTIZE procedure in Alg. 2 takes the membership of point $v$, $\mu_A(v)$, and gives the index $i$ of the minimal $\alpha$-level ($\alpha_1, \ldots, \alpha_\ell$) for which $\mu_A(v) \geq \alpha_i$. If the membership is below all $\alpha$-levels, the index is 0. For $\ell$ equally spaced $\alpha$-levels, the quantization can be expressed as

$$\text{QUANTIZE}(\mu_A(v)) = \lfloor \ell \mu_A(v) + 0.5 \rfloor. \quad (30)$$

The INTERPOLATE procedure in Alg. 2 computes the value of the discrete functions $D$ and $G$ in point $\hat{v}$ which may not be on the grid due to application of $T$. There are many interpolation schemes proposed in the literature, but...
we suggest that either nearest neighbor (for maximal speed) or linear interpolation (for higher accuracy) are used here, since the distance and gradient fields are smooth. By linearity of integration and summation, nearest neighbor and linear interpolation may be performed on the pre-processed \( D \) and \( G \) and yield the same result as if each level was interpolated before integration, allowing efficient computation. The (discretized) measure does not require intensity interpolation; the interpolation operates on distances and gradients only.

IV. IMPLEMENTATION

We implemented the proposed distance measure and registration method in the open-source Insight Segmentation and Registration Toolkit (ITK) \(^4\). We chose this particular software framework because it

- enables the use of an existing optimization framework,
- allows for a fair comparison against well written, tested, and widely used implementations of reference similarity measures, with support for resolution-pyramids,
- supports anisotropic/scaled voxels in relevant algorithms,
- facilitates reproducible evaluation,
- makes the proposed measure easily accessible for others.

The built-in ITK optimizer we have used for the registration tools and all the evaluation is RegularStepGradientDescentOptimizer\(^v4\). This is an optimizer based on gradient descent, with an initial step-length \( \lambda \), and a relaxation factor which reduces the used step-length gradually as the direction of the gradient changes, in order to enable convergence with high accuracy. In addition to a maximum number of iterations \( N \), two termination criteria are used: (i) a gradient magnitude threshold (GMT),

\[
\sqrt{\frac{\partial d}{\partial T_1}^2 + \ldots + \frac{\partial d}{\partial T_r}^2} < \text{GMT},
\]

and (ii) a minimum step-length (MSL),

\[
\lambda r < \text{MSL},
\]

where \( r \) is the current relaxation coefficient. We use default values of 0.0001 for both of these criteria. A relaxation factor of 0.99 is used for all experiments, since it performed well in preliminary tests; in this study we are willing to trade some (potential) additional iterations for better robustness. To maximize utilization of the limited number of \( \alpha \)-levels, images are normalized before registration to make sure that they are within the valid \([0,1]\) interval. We use the following robust normalization technique: Let \( P_\rho(X) \) denote the \( \rho \)-th percentile of image \( X \) with respect to image intensities,

\[
\text{NORM}_\rho(X) = \max \left[ 0, \min \left[ 1, \frac{(X - P_\rho(X))}{P_{1-\rho}(X) - P_\rho(X)} \right] \right].
\]

V. PERFORMANCE ANALYSIS

We evaluate performance of the proposed method, both for 2D and 3D images, in two different scenarios: (i) we perform a statistical study on synthetically generated images, where we seek to recover known transformations and measure the registration error by comparing the ground truth locations of known landmarks with the corresponding registered ones; (ii) we apply the proposed framework to real image analysis tasks: landmark-based evaluation of registration of TEM images in 2D, and atlas-based segmentation evaluation of 3D MR images of brain.

To compare the proposed measure and registration method against the most commonly used alternative methods and similarity measures, we select the widely used ITK implementations of optimization framework and similarity measures (SSD, PCC and MI) as the baseline of intensity-based registration accuracy. Note that the PCC measure is denoted Normalized Cross Correlation (NCC) in the ITK framework.

All experiments are performed on a workstation with a 6-core Intel i7-6800K, 3.4GHz processor with 48GB of RAM and 15MB cache. The operating system used is Ubuntu 16.04 LTS. The compiler used to build the framework is g++ version 5.4.0 (20160609). Version 4.9 of the ITK-framework is used for testing and evaluation.

A. Datasets

One biomedical 2D dataset and one medical 3D dataset are used for the evaluation.

1) TEM Images of Cilia (2D): The dataset of 20 images of cilia \(^1\) is acquired with the MiniTEM\(^4\) imaging system. Each image is isotropic of size 129 × 129 pixels, with a pixel-size of a few nm. An example is shown in Fig.\(^1\). A particular challenge is the near-rotational symmetry of the object: 9 pairs

\(^4\)MiniTEM imaging system is developed by Vironova AB.
of rings are located around a central pair of rings, which gives 9 plausible solutions for a registration problem. The alignment of the central pair can be taken into special consideration to reduce the number of solutions to two. The dataset comes with a set of 20 landmarks per image, indicating the position of each of the relevant structures to be detected and analysed – 20 rings (2 in the center and 18 in a circle around the center). The landmarks are produced by a domain expert and are only used for evaluation of the registrations.

2) LPBA40 (3D): LPBA40 [1] is a publicly available dataset of 40 3D images of brains of a diverse set of healthy individuals, acquired with MRI. The images are anisotropic, of size $256 \times 124 \times 256$ voxels with voxel-size $0.86 \times 1.5 \times 0.86 \text{mm}^3$. The dataset comes with segmentations of the brains into 56 distinct regions marked by a medical expert, which may be used for brain extraction. LPBA40 includes two atlases: first 20 out of 40 MR images of brain in the reference image space, after transformation of the floating image corner landmarks $\mathbf{R}$ of each of the relevant structures to be detected and analysed – of corresponding image corner landmarks $\mathbf{L}$ in a circle around the center). The atlases contain both a synthesized MR image and the fused label category in all the voxels, as well as a whole brain mask which may be used for brain extraction.

B. Evaluation Criteria

We evaluate accuracy and robustness of the registration methods in presence of noise, their robustness w.r.t. change of roles of reference and floating image (symmetry), and their speed. We quantify the performance of the observed frameworks in terms of the following quality measures:

1) Average Error measure (AE): The registration result is quantified as the mean Euclidean distance between the sets of corresponding image corner landmarks $L_R$ and $T(L_F)$ in the reference image space, after transformation of the floating image corner landmarks $L_F$, where $|L_R|$ is the number of landmarks (4 in 2D; 8 in 3D). The quality measure is defined as

$$AE(T; L_R, L_F) = \frac{1}{|L_R|} \sum_{i=1}^{L_R} d_E(L_R(i), T(L_F(i))).$$

(34)

A slight variation of this measure, the Average Minimal Error (AME), is used in the real task of cilia registration:

$$AME(T; L_R, L_F) = \frac{1}{|L_R|} \min_{T} \sum_{i=1}^{L_R} d_E(L_R(i), T(x)).$$

(35)

For the central pair, the error is simply $AME_{CP} = AME$, whereas for the outer rings we utilize the knowledge that an odd (even) landmark should be matched with an odd (even) landmark of the other image. The error function for the outer rings, [10], is therefore defined as:

$$AME_{Outer}(T; L_R^{Odd}, L_F^{Odd}, L_R^{Even}, L_F^{Even}) = \frac{1}{2}(AME(T; L_R^{Odd}, L_F^{Odd}) + AME(T; L_R^{Even}, L_F^{Even})).$$

(36)

2) Success Rate (SR): A registration is considered successful if its AE is below one voxel (pixel). Success rate (SR) at a given AE value corresponds to the ratio of successful registrations (w.r.t. the set of performed ones).

3) Symmetric Success Rate (SymSR): is defined as the ratio of performed registrations which are successful (i.e., $AE \leq 1$) in both directions, i.e., when the roles of reference and floating image are exchanged.

4) Inverse Consistency Error (ICE), [26]: Given a set of interest $X_A \subseteq A$, the transformations $T_{AB}: A \rightarrow B$, and $T_{BA}: B \rightarrow A$, the ICE of this pair of transformations is

$$ICE(T_{AB}, T_{BA}; X_A) = \frac{1}{|X_A|} \sum_{x \in X_A} d_E(T_{BA}(T_{AB}(x)), x).$$

(37)

We compute ICE considering all the points of the reference image for each of the cases where Symmetric Success is observed ($AE \leq 1$ in both directions).

5) Jaccard Index for segmentation evaluation: For two binary sets, $R_1$ and $R_2$, the Jaccard Index is defined as

$$J(R_1, R_2) = \frac{|R_1 \cap R_2|}{|R_1 \cup R_2|}.$$  

(38)

6) Execution Time: We evaluate (i) the execution times required for one iteration in the registration procedure, i.e., times needed to compute the distance (similarity) measure and its derivatives, with full sampling, and in full image resolution, between two distinct images from the same set, as well as (ii) the execution time for complete registrations.

C. Parameter Tuning

The distance measure and optimization method have a number of parameters which must be properly chosen. Synthetic tests indicated that the following values lead to good optimization performance: three pyramid levels with downsampling factors $(4,2,1)$ and Gaussian smoothing $\sigma = (5.0,3.0,0.0)$, max 3000 iterations per level and an initial step-length $\lambda = 0.5$. The number of $\alpha$-levels used is $\ell = 7$. Normalization percentile is normally 5%. This same parameter setting, if not stated differently, is used in all the tests, on both synthetic and real data.

D. Synthetic Tests

A synthetic evaluation framework is used to evaluate the performance of the proposed method, and to compare it with standard tools based on SSD, PCC, and MI, in a controlled environment. For this evaluation, we construct sets of transformed versions of a reference image and add (a new instance of) Gaussian noise to each generated image. The transformations are selected at random from a multivariate uniform distribution of rotations measured in degrees (1 angle for 2D images and 3 Euler angles for 3D images) and translations measured in fractions of the original image size.

1) 2D TEM Images of Cilia: Three sets of transformed images are built based on image Nr. 1 in the observed dataset, by applying on it the following three groups of transformations: Small, containing compositions of translations of up to 10% of image size (in any direction) and rotations by up to 10°; Medium, containing compositions of translations and rotations such that at least one of the parameters exceeds the range of Small, and falls within 10 – 20% of image size of translation (in at least one direction), or 10 – 20° of rotation; and Large,
containing compositions of translations and rotations such that at least one of the parameters exceeds the range of Medium, and falls within 20 – 30% of image size of translation (in at least one direction), or 20 – 30° of rotation. The transformed images are also corrupted by additive Gaussian noise, from \( \mathcal{N}(0, 0.1^2) \) (\( \sigma = 0.1 \), corresponding to a PSNR≈20 dB). Each group of transformations is applied 1000 times, and the resulting images are registered to image Nr. 1, each time corrupted by a new instance of Gaussian noise.

To evaluate symmetry, we performed 1000 registrations of images transformed by randomly selected translations of up to 30% of image size, and rotations by up to 30°, and corrupted by additive noise from \( \mathcal{N}(0, 0.1^2) \). Each of the registrations were performed twice, with exchanged roles of reference image and floating image.

Intensity-based registration with gradient-descent optimization can be computationally demanding, requiring the distance function and its derivative for each iteration of the optimization procedure. The time to compute the distance and derivatives is directly proportional to the number of sampled points. We, therefore, evaluate influence of the sampling fraction on registration success, observing registrations after Small transformations and added noise (with \( \sigma = 0.1 \)), over a range of sampling fractions. For each evaluated sampling fraction, 1000 registrations are performed and SR and AE are computed for successful registrations (\( AE \leq 1 \)). No resolution pyramids are used for these tests.

2) 3D MR Images of Brain: Three sets of transformed images are built based on image Nr. 1 in the observed dataset, by applying to it the following three groups of transformations: Small, containing compositions of translations of up to 10% of image size (in any direction) and rotations by up to 10° (around each of the rotation axes); Medium, containing compositions of translations and rotations such that at least one of the parameters exceeds the range of Small, and falls within 10 – 15% of image size of translation (in at least one direction), or 10 – 15° of rotation (around at least one rotation axes); and Large, containing compositions of translations and rotations such that at least one of the parameters exceeds the range of Medium, and falls within 15 – 20% of image size of translation (in at least one direction), or 15 – 20° of rotation (around at least one rotation axes). The transformed images are also corrupted by additive Gaussian noise, from \( \mathcal{N}(0, 0.1^2) \). Each group of transformations is applied 200 times, and the resulting images are registered to image Nr. 1, each time corrupted by a new instance of Gaussian noise.

E. Results of Synthetic Tests

1) 2D TEM Images of Cilia: Figure 4 shows the distributions of registration errors (AE), for the three transformation classes. Superiority of the proposed measure, and the corresponding registration framework, is particularly clear for Medium and Large transformations; it reaches a 100% success rate, with subpixel accuracy, whereas the competitors not only exhibit considerably lower accuracy, but also much lower success rate, i.e., they completely fail in a large number of cases.

Overall registration performance is summarized in Table I, for complete sampling (a), and for random sampling of 10% of the points (b). The proposed method has 100% success rate and also 100% symmetric success rate. The other observed measures exhibit much lower success rate and poor symmetry scores; the second best, SSD, succeeds in 54% of the cases, and succeeds symmetrically in only 31% of the cases. The registration error for successful registrations is considerably smaller for the proposed method, while the execution time is considerably lower. The reduced sampling fraction in (b) has
TABLE I
REGISTRATION OF SYNTHETIC 2D IMAGES OF CILIA. THE TABLES SHOW SUCCESS RATE (SR), AVERAGE ERROR (AE) OF SUCCESSFUL REGISTRATIONS, SYMMETRIC SUCCESS RATE (SymSR), AVERAGE INVERSE CONSISTENCY ERROR (ICE) AND AVERAGE RUNTIME FOR THE REGISTRATION WITH COMPLETE SAMPLING (A) AND WITH RANDOM SUBSAMPLING (B). SUCCESSFUL REGISTRATIONS (AE ≤ 1) OF TRANSFORMATIONS UP TO (AND INCLUDING) LARGE ARE CONSIDERED.

| Measure | SR | AE | SymSR | ICE | Time (s) |
|---------|----|----|-------|-----|---------|
| SSD     | 0.356 | 0.3086 | 0.313 | 0.2424 | 17.3 |
| PCC     | 0.363 | 0.3413 | 0.249 | 0.4227 | 20.8 |
| MI      | 0.440 | 0.3495 | 0.251 | 0.4518 | 18.3 |
| \(d_{\alpha}^{R} \) | 1.000 | 0.1295 | 1.000 | 0.0023 | 3.49 |

(a) Full sampling

| Measure | SR | AE | SymSR | ICE | Time (s) |
|---------|----|----|-------|-----|---------|
| SSD     | 0.367 | 0.6270 | 0.186 | 0.5260 | 1.761 |
| PCC     | 0.299 | 0.6364 | 0.152 | 0.5676 | 2.171 |
| MI      | 0.283 | 0.6219 | 0.068 | 0.5974 | 2.083 |
| \(d_{\alpha}^{R} \) | 1.000 | 0.1311 | 1.000 | 0.0193 | 0.834 |

(b) 0.1 sampling fraction

Successful registrations (AE ≤ 1) are observed. Here as well, the proposed method delivers 100% success rate, whereas the second best, SSD, succeeds in only 33% of the cases. The registration error for successful registrations is the smallest for the proposed method. We observe a relative increase in execution time of the proposed registration framework in 3D case, where it is slightly slower than the other measures.

3) Execution Time Analysis: The number of iterations required for convergence of the optimization (registration) typically range from 1000 to 3000. Measures SSD, PCC and MI use cubic spline interpolation. Lookups from the distance map \(d_{\alpha}^{R} \) are done using linear interpolation. Table II shows the mean (and standard deviation) execution time of one iteration, which includes computation of the measures and their derivatives, repeated 1000 times for 2D, and 50 times for 3D affine image registrations. We observe that the proposed measure is the fastest per iteration both in 2D and 3D. Note that these execution time measurements exclude pre-processing.

F. Evaluation on Real Applications

1) Registration of Cilia: Registration of multiple cilia instances detected in a single TEM sample, for enhancement of diagnostically relevant sub-structures, requires a pixel-accurate and robust method which is able to overcome the challenges posed by the near-rotational symmetry of a cilium. At most two of the possible solutions properly align the central pair, which is vital for a successful reconstruction.

We compare the performance of the proposed method with reported results of a previous study [10] which uses intensity-based registration with PCC as similarity measure. We follow the general protocol described in [10] and perform, as a first step, a multi-start rigid registration (parameterized by angle \(\theta \) in radians, and translation \(t = (tx, ty)\)) followed, in a second step, by affine registration initiated by the best (lowest final distance) registration of the 9 rigid ones.

No resolution pyramids are used since they were observed to interfere with the multi-start approach (by facilitating large movements). The registrations are performed in full resolution, without stochastic subsampling. For the rigid registration we use a small circular binary mask with radius of 24 pixels, positioned in the center, combined with a squared circular Hann window function. The affine registration is performed using a circular binary mask with radius of 52 pixels; the mask removes the outside background and the outer plasma which is not helpful in guiding the registration. No additional

FIG. 5. (Left/Blue) SR for registrations of cilia images, and (Right/Red) AE of the successful registrations, as functions of sampling fraction for the proposed method. Both measures improve (almost) monotonically with sampling fraction and flatten out after approximately 0.01.
weight-mask is used for the affine registration. Step length 0.1 was used for the rigid and 0.5 for the affine registration. We use $\ell = 7$. Normalization percentile is set to 0% for the rigid stage and 1% for the affine stage.

2) Atlas-based Segmentation (LPBA40): In [27], a protocol for evaluation of distance/similarity measures in the context of image registration was proposed. The protocol starts with affine registration, for which results are reported, and then proceeds to deformable registration. Since this study focuses on the development of an affine (linear) registration framework based on the proposed distance measure, we compare with the reported affine-only performance; an improved affine registration is of great significance since a very high correlation between the performance of the affine registration and that of the subsequent deformable registration has been established.

We start from the two atlases created utilizing the Advanced Neuroimaging Tools (ANTS) registration software suite and the open-source evaluation script provided in the reference study [27]. We utilize the atlas created using Mutual Information since that is the one found in [27] to be best performing for evaluations.

The multi-label segmentations defined by the atlas are transformed using the transformation parameters found during the registration and compared to the ground-truth segmentations for each brain. The Jaccard Index [28] is calculated per region, as well as for the entire brain mask.

For the proposed method based on $\tilde{d}_{\text{AMD}}^n$ we use $\ell = 7$, normalization percentiles 5%, $N = 3000$, 0.05 sampling fraction, and circular Hann windows as weight-masks.

G. Results of Real Applications

1) Results of Registration of Cilia: Performance of the proposed method, together with the best previously published results, are shown in Tab. III. The table shows the mean and standard deviation of registration error (AME, in pixels) of the 19 registrations, for the three considered sets of landmarks: the Central pair, the Outer rings, and All (1+9) ring pairs.

The original study includes deformable registration as a final stage, after the rigid and affine steps. Here presented framework based on $\tilde{d}_{\text{AMD}}^n$ includes linear (rigid and affine),
but not deforming registration. However, as results included in Tab. [III] confirm, the proposed method outperforms the previous state-of-the-art, even if using only rigid and affine registrations.

We note that with only rigid registration we improve the alignment of the central pair while degrading the alignment of the outer rings. After the affine registration, the alignment of the central pair is improved further, plausibly due to the less constrained transformation model of affine compared to rigid, and we observe that the alignment of the outer rings and the total alignment are improved substantially.

2) Results of Atlas-Based Segmentation of Brains: Table [IV] shows results of atlas-based brain segmentation. The Mean Jaccard Index is computed for each of the brain regions, for $d_{\text{AMD}}$ and MI, with affine registration as reported in [27]. For $d_{\text{AMD}}$, mean and std. dev. are displayed; for the comparative results (MIaff, [27]), only mean was reported.

We observe that for the whole brain mask, for the aggregated overlap, and for 43 out of the 56 distinct regions, the proposed measure outperforms the reported performance obtained with the MI metric; MI was the best performing measure out of the three evaluated in [27].

VI. DISCUSSION

Compared to the traditional similarity measures (SSD, PCC, MI), the proposed measure and associated registration method require substantial amounts of memory to store the auxiliary data-structures. A single 3D registration of two MR images of brains may require approximately 4GB of working memory with a reasonable set of parameters; contemporary machines for high-end data processing typically have a lot more memory than 4GB, but this requirement can affect how many registrations can be performed in parallel on a single machine.

VII. CONCLUSION

In this study we have adapted a family of distance measures to gradient descent based image registration, for 2D and 3D images. We have shown that such an extension is feasible and that the very good performance of the measures observed previously for object recognition and template matching, and their property of a large catchment basin for local optimization, also hold in the context of registration. This has been shown by evaluating the method in four main ways: (i) on synthetic tests, (ii) execution time measurement, (iii) registration of TEM-images of cilia for multi-image super-resolution reconstruction, and (iv) atlas-based segmentation with annotated MR

| Method | Transform | ΔR | R+A | R+A+D | Central Pair | Outer | All |
|--------|-----------|----|-----|-------|--------------|-------|-----|
| -      | Identity  | 4.32 (0.78) | 5.75 (3.49) | 5.61 (3.09) | - | - | - |
| PCC    | R         | 2.6 (1.5) | - | - | - | - | - |
|        | R+A       | 2.79 (1.84) | 3.27 (1.75) | 2.30 (1.30) | 2.35 (1.82) | - | - |
|        | R+A+D     | 2.65 (0.83) | 6.49 (2.64) | 6.10 (2.41) | - | - | - |

...
brain images. The performance observed in this study shows that the proposed method provides outstanding performance for intensity-based affine registration, in terms of robustness, accuracy and symmetry. It is also faster or similar in speed to the commonly used measures, which allows its practical applications.

Future work includes extending the measures to non-linear (deformable), as well as multi-modal registration.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Ida-Maria Sintorn for providing the cilia dataset and the landmarks used for the evaluation in Sec. V-F1.

REFERENCES

[1] J. Maintz and M. A. Viergever, “A survey of medical image registration,” Medical Image Analysis, vol. 2, no. 1, pp. 1 – 36, 1998.
[2] M. A. Viergever, J. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. Pluim, “A survey of medical image registration – under review,” Medical Image Analysis, vol. 33, pp. 140 – 144, 2016.
[3] B. Zitova and J. Flusser, “Image registration methods: a survey,” Image and vision computing, vol. 21, no. 11, pp. 977 –1000, 2003.
[4] A. Sotiras, C. Davatzikos, and N. Paragios, “Deformable medical image registration: A survey,” IEEE Transactions on Medical Imaging, vol. 32, no. 7, pp. 1153–1190, 2013.
[5] D. G. Lowe, “Object recognition from local scale-invariant features,” in Proceedings of the Seventh IEEE International Conference on Computer Vision, vol. 2. IEEE, 1999, pp. 1150–1157.
[6] B. Berkels, P. Binev, D. A. Blom, W. Dahmen, R. C. Sharpley, and T. Vogt, “Optimized imaging using non-rigid registration,” Ultramicroscopy, vol. 138, pp. 46 – 56, 2014.
[7] B. Fischer and J. Modersitzki, “Ill-posed medicine - an introduction to image registration,” Inverse Problems, vol. 24, no. 3, p. 034008, 2008.
[8] D. Skerl, B. Likar, and F. Pernus, “A protocol for evaluation of similarity measures for rigid registration,” IEEE Transactions on Medical Imaging, vol. 25, no. 6, pp. 779–791, June 2006.
[9] J. Lindblad and N. Sladoje, “Linear time distances between fuzzy sets with applications to pattern matching and classification,” IEEE Transactions on Image Processing, vol. 23, no. 1, pp. 126–136, 2014.
[10] A. Suveer, N. Sladoje, J. Lindblad, A. Dragomir, and I.-M. Sintorn, “Enhancement of cilia sub-structures by multiple instance registration and super-resolution reconstruction,” in Scandinavian Conference on Image Analysis. Springer, 2017, pp. 362–374.
[11] D. W. Shattuck, M. Mirza, V. Adisetiyo, C. Hojatkashani, G. Salamon, K. L. Narr, R. A. Poldrack, R. M. Bilder, and A. W. Toga, “Construction of a 3D probabilistic atlas of human cortical structures,” Neuroimage, vol. 39, no. 3, pp. 1064–1080, 2008.
[12] B. B. Avants, N. J. Tustison, M. Stauffer, G. Song, B. Wu, and J. C. Gee, “The Insight ToolKit image registration framework,” Frontiers in neuroinformatics, vol. 8, p. 44, 2014.
[13] L. A. Zadeh, “Information and control,” Fuzzy sets, vol. 8, no. 3, pp. 338–353, 1965.
[14] J. V. Hajnal, N. Saced, E. J. Soar, A. Oatridge, I. R. Young, and G. M. Bydder, “A registration and interpolation procedure for subvoxel matching of serially acquired MR images,” Journal of computer assisted tomography, vol. 19, no. 2, pp. 289–296, 1995.
[15] P. Viola and W. M. Wells III, “Alignment by maximization of mutual information,” International journal of computer vision, vol. 24, no. 2, pp. 137–154, 1997.
[16] M. Irani and S. Peleg, “Improving resolution by image registration,” CVGIP: Graphical models and image processing, vol. 53, no. 3, pp. 231–239, 1991.
[17] A. Rosenfeld, Multiresolution image processing and analysis. Springer Science & Business Media, 2013, vol. 12.
[18] J. Lindblad, V. Curic, and N. Sladoje, “On set distances and their application to image registration,” in Proceedings International Symposium on Image and Signal Processing and Analysis, Sept 2009, pp. 449–454.
[19] T. Eiter and H. Mannila, “Distance measures for point sets and their computation,” Acta Informatica, vol. 34, no. 2, pp. 109–133, 1997.
[20] P. Brass, “On the nonexistence of hausdorff-like metrics for fuzzy sets,” Pattern Recognition Letters, vol. 23, no. 1–3, pp. 39–43, 2002.

Johan Öfverstedt received the B.Sc. and M.Sc. degrees in computer science in 2012 and 2017 respectively, and is currently pursuing a Ph.D. degree in computerized image processing at Centre for Image Analysis, Uppsala University, Sweden, since 2017. His research interests include distance measures, image registration, object recognition, and optimization.

Joakim Lindblad received the M.Sc. in engineering physics and Ph.D. in computerized image analysis from Uppsala University, Sweden, in 1997 and 2003, respectively. He is currently Researcher at the Centre for Image Analysis, Uppsala University, Sweden; Senior Research Associate at Mathematical Institute of the Serbian Academy of Sciences and Arts, Serbia; and Head of Research at Topgolf Sweden AB, Stockholm, Sweden. His research interests include development of general and robust methods for image processing and analysis.

Nataša Sladoje received the B.Sc. and M.Sc. degrees in mathematics from the Faculty of Science, University of Novi Sad, Serbia, in 1992 and 1998, respectively, and the Ph.D. degree in computerized image analysis from the Centre for Image Analysis, Swedish University of Agricultural Sciences, Uppsala, Sweden, in 2005. She is a Senior Lecturer at the Centre for Image Analysis, Uppsala University, Sweden. Her research interests include theoretical development of image analysis methods with applications in biomedicine and medicine.