Rigid Image Registration by General Adaptive Neighborhood Matching

Johan Debayle, Benoît Presles

▶ To cite this version:

Rigid Image Registration by General Adaptive Neighborhood Matching

Johan Debayle[a,1,*], Benoît Presles[b,*]

[a]*Ecole Nationale Supérieure des Mines de Saint-Etienne, SPIN/LGF UMR CNRS 5307*
158 cours Fauriel, CS 62362 42023 Saint-Etienne Cedex 2, FRANCE
[b]*CREATIS, CNRS UMR5220, Inserm U1044, INSA-Lyon, Université Lyon 1, Villeurbanne; and Léon*
Bérard Cancer Center, University of Lyon, France

Abstract

This paper aims to propose a new feature and intensity-based image registration method. The proposed approach is based on the block matching algorithm [1]: a displacement field is locally computed by matching spatially-invariant intensity sub-blocks of the images before performing an optimization algorithm from this vector field to estimate the transformation. Our approach proposes a new way to calculate the displacement field by matching spatially-variant sub-blocks of the images, called General Adaptive Neighborhoods (GANs) [2]. These neighborhoods are adaptive with respect to both the intensities and the spatial structures of the image. They represent the patterns within the grayscale images. This paper also presents a consistent shape metric used to match the GANs. The performed qualitative and quantitative experiments show that the proposed GAN matching method provides accurate displacement fields enabling to perform image rigid registration, even for data from different modalities, that outperforms the classical block matching algorithm with respect to robustness and accuracy criteria.

Keywords: Block matching, Displacement field, General adaptive neighborhoods, Image registration, Shape metric

[*]In alphabetical order. Both authors contributed equally to this work.
Email addresses: debayle@emse.fr (Johan Debayle),
benoit.presles@creatis.insa-lyon.fr (Benoît Presles)
[1]Corresponding author. Tel: +33477420219. Fax: +33477499694

1. Introduction

Image registration consists in bringing two images, acquired with the same or different sensors, into spatial alignment. More formally, given two input images, registering the floating (i.e., moving) image to the reference (i.e., fixed) image entails finding the spatial transformation that minimizes the dissimilarity between the transformed floating and reference images. This process is mainly composed of three elements:

- a transformation space, which describes the set of admissible transformations from which one is chosen to apply to the floating image,

- a similarity criterion, which measures the discrepancy between the images, and

- an optimization algorithm, which traverses the transformation space, in search of the transformation that minimizes the similarity criterion.

Many registration methods have been developed and/or used in the literature using a large variety of:

- transformation spaces (linear [3], polyaffine [4, 5], elastic [6, 7], fluid [8], ...);

- similarity criteria (sum of squared differences [9], correlation coefficient [10], correlation ratio [11], mutual information [12], multiscale integral invariants [13]...); and

- optimization algorithms (Powell method [14], Levenberg-Marquardt method [15], stochastic search [16], ...).

However, registration methods can be classifed into two main categories. The first category, called *geometric methods*, is based on feature matching, where transformations are calculated using correspondences between points [17], contours [18], ... While it could be argued that these techniques enable a better control over the registration process, the feature extraction can be a difficult task [19]. The second category of methods, called *iconic methods* or *intensity-based methods*, rely on the intensities associated to pixels/voxels in the input images. Assuming a global relationship between the intensities of the images to register (affine, functional, statistical, ...), the approach consists in

2

maximizing (or minimizing) a specified similarity measure between intensities of the corresponding pixels [6, 12]. These approaches have several issues [1]. First, all the similarity measures are known to be highly non-convex with respect to the transformation parameters. Thus their global maximization is seldom straightforward. Secondly, the assumption of a global relationship between the image intensities may be violated by the presence of various image artefacts.

1.1. Presentation of the problem

The problems mentioned above have been addressed by hybrid registration methods. Such methods combine feature and intensity information [1, 20, 21, 22, 23, 24]. However, these methods still require a pre-segmentation of features or take spatial regions of interest independently of the image context. For example, the block-matching algorithm [1], a local iconic method, determines a displacement field based on intensity similarity on small sub-blocks of the image, before determining a global transformation. The blocks have a fixed size and shape and are determined independently of the local structure of the image. Local displacements cannot be accurately computed in homogeneous areas or for large transformations [25]. It can therefore lead to some outliers in the displacement field estimation. Some specific strategies including interpolation, regularization or robust optimization methods are then required to try to solve these problems [1].

Alternatively, approaches to get a more robust raw displacement field could be investigated. Combining the two strategies (estimating a robust raw displacement field and using a robust optimization method) should lead more easily to the expected transformation.

1.2. Aims and outline of the paper

This paper aims to propose a new hybrid (feature and intensity-based) registration method. It is based on the block matching algorithm [1]: but contrary to the latter, it computes and matches spatially-variant sub-blocks of the images, called General Adaptive Neighborhoods (GANs) [2], before determining a global transformation. These neighborhoods represent the patterns within the grayscale images: they are adaptive

3

with respect to the image intensities and the image spatial structures. The GANs are determined by the image itself and should provide a more robust and accurate estimation of the displacement field.

The paper is organized in the following way. Section 2 gives the concepts and definitions of the General Adaptive Neighborhood (GAN) framework. Section 3 describes the GAN-based registration method including the computation of the displacement field and the optimization algorithm to estimate the transformation. Section 4 introduces a shape metric based on [26] that is used to measure the similarity between GANs and therefore to match the GANs. The performance of this metric is evaluated on a dataset of binary images. Finally, some qualitative and quantitatives results are exposed in Section 5 highlighting the accuracy and robustness of the proposed registration method. The last section is devoted to the conclusion. In this study, the method is described in two dimensions (2D) and the geometric transformation is limited to rotations and translations.

2. GAN image representation

This paper deals with 2D intensity images, that is to say image mappings defined on a spatial support D in the Euclidean space $\mathbb{R}^2$ and valued into a gray tone range, which is a real number interval. The General Adaptive Neighborhood paradigm has been introduced [27] in order to propose an original image representation for adaptive processing and analysis. The central idea is based on the key notion of adaptivity which is simultaneously associated to the analyzing scales, the spatial structures and the intensity values of the image class to be addressed (see Subsection 2.2). This section aims to recall the concepts and definitions of the GAN framework. The interested reader can look at the references [2, 28] for more details.

2.1. GANs sets

In the so-called General Adaptive Neighborhood Image Processing (GANIP) approach [2, 28], a set of General Adaptive Neighborhoods (GANs set) is identified around each point in the image to be analyzed. A GAN is a subset of the spatial

support constituted by connected points whose measurement values, in relation to a selected criterion (such as luminance, contrast, thickness, …), fit within a specified homogeneity tolerance. In this way, the computation of a GAN can be done by using a region growing process from the current point. These GANs are used as adaptive windows for image transformations or quantitative image analysis.

Several GANs sets have been defined and each collection satisfies specific properties [2]. This paper only presents the most elementary kind of these ones, denoted $V_m^h(x)$. For each point $x \in D$ and for an image f, the GANs $V_m^h(x)$ are subsets of D. They are built upon a *criterion mapping* h (based on a local measurement such as luminance, contrast, thickness, …related to f), in relation with an *homogeneity tolerance* $m \in \mathbb{R}^+$. More precisely, $V_m^h(x)$ is a subset of D fulfilling two conditions :

1. its constituting points have a measurement value close to that of the point x :
$$\forall y \in V_m^h(x) \quad |h(y) - h(x)| \leq m$$

2. the set is path-connected (with the usual Euclidean topology on $D \subseteq \mathbb{R}^2$).

The GANs are mathematically defined as follows for each point $x \in D$:

$$V_m^h(x) = C_{h^{-1}([h(x)-m, h(x)+m])}(x) \tag{1}$$

where $C_X(x)$ denotes the path-connected component [29] (with the usual Euclidean topology on $D \subseteq \mathbb{R}^2$) of $X \subseteq D$ containing $x \in D$.

Figure 1 illustrates the GANs of two points computed with the luminance criterion on a retina image. The GANs are self-determined by the local structures of the image.

Note that two distinct points x and y may lead to the same GAN. For example, if $h(x) = h(y)$ and $x \in V_m^h(y)$ then $V_m^h(x) = V_m^h(y)$.

2.2. GAN paradigm

A multiscale image representation such as wavelet decomposition [30] or isotropic scale-space [31], generally takes into account analyzing scales which are global and a priori defined, that is to say extrinsic scales. This kind of multiscale analysis presents a main drawback since a priori knowledge, related to the features of the studied image, is required. On the contrary, the GAN framework is an intrinsic multiscale representation,

5

(a) original image f with two seed points x and y (white dots)

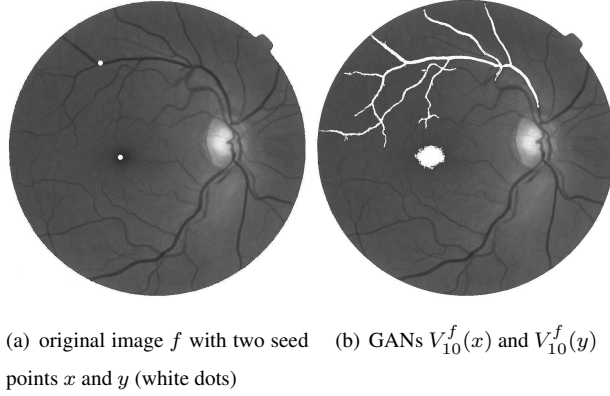(b) GANs $V_{10}^f(x)$ and $V_{10}^f(y)$

Figure 1: The GANs of the two selected points of the original image (a) are homogeneous with respect to the luminance criterion (b) using the tolerance $m = 10$.

such as anisotropic scale-space [32], where scales are self-determined by the local image structures. Such a decomposition does not need any a priori information. The GAN image representation is thus *adaptive with respect to the different scales of the image representation*.

Furthermore, the image processing techniques using spatially invariant transformations, with fixed analyzing neighborhoods, give effective and compact computing structures, in the sense where data and operators are independent. Nevertheless, they have several drawbacks such as creating artificial patterns, changing the detailed parts of large objects, damaging transitions or removing significant details [33]. Alternative approaches towards context dependent processing have been proposed [34]. The GAN image representation is one of them in the sense that it supplies *spatially adaptive* analyzing neighborhoods which are no longer spatially invariant, but vary over the whole image, taking locally into account the image context. The GAN image representation is thus *adaptive with respect to the spatial structures*.

The next section exposes the GAN-based image registration method. The proposed approach is focused on the class of rigid transformations. Note that it could be extended to other classes of transformations with appropriate similarity measures and optimization algorithms but it is out of the scope of this paper.

6

3. GAN-based image registration

The proposed method takes two images as input: a reference image I and a floating image J. The output is a transformation T bringing the two input images into spatial alignment. It is based on the block-matching strategy, where two successive tasks are performed with a multiscale and iterative scheme (Subsection 3.3):

- computing a displacement field (i.e. a set of displacement vectors) between I and J (Subsection 3.1),

- gathering these displacements to determine the global transformation T using a robust estimator (Subsection 3.2).

3.1. GAN matching: computation of the displacement field

A displacement field is determined by computing correspondences between pixels of the two input images I and J. For each pixel x of the spatial support of the image I, the GAN $V_m^I(x)$ is computed and its best correpondence (with respect to a dissimilarity measure DM) with the GANs $V_m^J(y)$ is sought in the image J for pixels y within the neighborhood of x. Note that the size m of the GANs is an important parameter to be chosen for this matching. This GAN matching is mathematically defined as:

$$\hat{y} = \arg \min_{y \in N(x)} \mathrm{DM}(V_m^I(x), x, V_m^J(y), y) \tag{2}$$

where $N(x)$ is the search neighborhood of x in the image J, and DM is the dissimilarity measure between GANs exposed in Section 4.

The best corresponding GAN $V_m^J(\hat{y})$ allows to define a displacement vector, $(x, \hat{y})$ between the pixel x of the reference image and the pixel $\hat{y}$ of the floating image.

The GAN-matching involves three parameters:

- the grid step size (Δ_x, Δ_y), underlying the spatial support of the image I, is introduced and determines the density of the displacement field. For $(\Delta_x, \Delta_y) = (1, 1)$, displacement vectors are calculated for all pixels of the spatial support of the image;

7

- the size of the neighborhood search $N(x)$ that is constant (fixed-size) over the whole image; and

- the homogeneity tolerance m for the computation of the GANs.

The main difference of this GAN matching algorithm with the block-matching one is the adaptivity of the blocks to be matched: fixed-size and fixed-shape blocks are replaced by variable-size and variable-shape blocks by using GANs.

3.2. Least trimmed squares minimization: estimation of the transformation T

The proposed GAN-matching algorithm provides a list of N corresponding 2D points, x_k and y_k. Assuming that there exists a rigid transformation T between the input images I and J, the problem is to estimate a rotation matrix R and a translation vector t that characterize the displacement. A standard approach to solve such a problem is to perform a Least Squares (LS) minimization on the residuals $r_k(R, t) = y_k - Rx_k - t$. The main advantage of the LS estimator is that the solution is unique and is quickly computed. However, LS is known to have poor robustness properties, in the sense that its solution is sensitive to outliers. M-estimators [35] appear to be the most straightforward alternative to LS to get a more robust estimation. In our experiments, the rigid transformation is estimated by Least Trimmed Squares (LTS) minimization [1]. It consists in solving the following minimization problem:

$$(\hat{R}, \hat{t}) = \arg \min_{R,t} \sum_{k=1}^{q} ||r_{(k)}(R, t)||^2 \tag{3}$$

based on the ordered residuals $||r_{(1)}|| \leq ||r_{(2)}|| \leq \cdots \leq ||r_{(N)}||$ where $||.||$ denotes the Euclidean norm. The value of the parameter q is fixed to $\lfloor 0.7 \times N \rfloor$ in this paper to achieve a 70% breakdown point.

Equation 3 is solved by means of a simple iterative LS estimation.

3.3. The iterative and multiscale scheme

For more accuracy, the method follows an iterative scheme [1]. At step n, the transformation T^n is estimated, $(T^n)^{-1}$ is then applied to the floating image J and

8

the process is again performed at step $n + 1$ between images I and $J \circ (T^n)^{-1}$ to obtain the transformation T^{n+1}. At each iteration, the current transformation is updated according to $T^{n+1} \leftarrow T^1 \circ ... \circ T^n$, the new floating image $J \circ (T^{n+1})^{-1}$ is therefore obtained by resampling only once the image J. This iterative scheme enables to enhance the accuracy of the registration method.

In order to recover large displacements without considering large search neighborhoods $N(x)$, the method is implemented in a multiscale scheme (coarse-to-fine analysis) [1]. It enables a trade-off between performance and computational cost. In this manner, large but inaccurate displacements are determined at high levels, and small but more accurate ones are found when the scale decreases. The multiscale scheme is based on a pyramidal representation of the two input images I and J. At each level i of the pyramid, a displacement field is computed between the two subsampled images I_i and J_i and a transformation T_i is estimated. $(T_i)^{-1}$ is then applied at the next level $i - 1$ of the pyramid before calculating a new displacement field between I_{i-1} and $J_{i-1} \circ (T_i)^{-1}$. Thereafter, the transformation T_{i-1} is estimated. The process is repeated up to the lowest level of the pyramid (i.e. at the highest resolution).

This iterative and multiscale scheme involves two additional parameters to the GAN-based image registration:

- the number i of pyramid levels in the multiscale representation; and

- the number n of iterations at each level of the pyramid.

The next section presents the dissimilarity measure for GAN matching, required to compute the displacement field.

4. The GAN dissimilarity measure

The computation of the displacement field (Eq. 2) between the two input images involves a dissimilarity measure DM between GANs, which are planar shapes of the spatial support of the images. Therefore, it is necessary to use a geometric measure (shape metric) as opposed to intensity-based measures (such as correlation coefficient,

mutual information, ...). Note that the GANs are already computed using the image intensities.

Several shape metrics have been proposed in the literature. The Hausdorff distance [36] is a classical shape metric used to determine the degree of ressemblance between two objects that are superimposed. It measures the minimum distance between any point of one set from some point of the other set. This distance is particular in the sense that no descriptor is required. Nevertheless, this distance is not invariant to rigid tranformations and is sensitive to noise [37].

Other shape metrics are based on the similarity between *descriptors* of the shapes, where the descriptor should satisfies several conditions:

- invariance: the description of the shape should be invariant to the expected image deformation;

- uniqueness: two different shapes should have different descriptions; and

- continuity: the description of a shape which is slightly deformed in an unknown manner should be close to the description of the original shape.

Many authors used contour-based descriptors of the shapes. Peli [38] proposed a simple and fast shape description using radial shape vectors but the usage of this method is limited to star-shape regions only. A generalized shape description in form of a binary matrix was proposed in [39]. This shape matrix description is invariant, continuous but time consuming. The Fourier descriptors describe a shape by a fixed number of sample points on the boundary and further transformed to the frequency domain using a Discrete Fourier Transform. Nevertheless, the Fourier descriptors can describe only a single closed curve shape. So, shapes with holes are not considered.

Another group of methods uses region-based descriptors of shapes. Hu [40] derived seven orthogonal invariant moment descriptors of order 2 and 3. The first six descriptors encode a shape with invariance to translation, scale and rotation. The seventh descriptor ensures skew invariance, which enables to distinguish mirrored images. Flusser and Suk derived the affine transform invariants [41]. These decriptors are easy to implement, little time consuming but not continuous. Tang et al. [42] have recently

10

proposed a rotation and scale invariant similarity measure based on hierarchical kernels but the method needs to define adequate templates which is still an open problem.

In our particular case, the dissimilarity measure DM between GANs should fulfill another condition. Indeed, two GANs $V_m^I(x)$ and $V_m^J(z)$ with exactly the same shape but with different seed positions x and z inside the GANs have to be discriminated to avoid bad matches (see Figure 2).
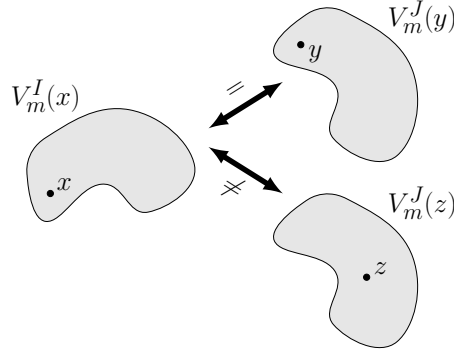

Figure 2: GAN dissimilarity measure: importance of the seed position inside the GANs. The three GANs $V_m^I(x), V_m^J(y), V_m^J(z)$ have the same geometrical shape. However, the match $(V_m^I(x), V_m^J(y))$ is correct (DM $= 0$) while $(V_m^I(x), V_m^J(z))$ is not (DM $\neq 0$). Indeed, the GANs of the two seeds y and z are equal (in terms of spatial region) but their descriptors $h_{V_m^J(y),y}$ and $h_{V_m^J(z),z}$ are different because the location of the seed point is taken into account.

Therefore, the shape descriptor has to be dependent on a reference point.

Furthermore, the computation of the dissimilarity measure between GANs will be performed a great number of times for the displacement field estimation. So, an additional requirement to the metric is needed: the shape metric should be little time consuming.

4.1. The proposed shape metric

The proposed metric for matching two shapes is based on the descriptor presented in [26] but it is here defined for binary images which represent the shapes.

In the following, the definition of the proposed shape metric is firstly given before evaluating its performance.

4.1.1. Definition

Let A and B be two shapes. For defining the descriptor of A, the distance values $d(y, x)$ between points y within A and a reference point x are considered (Fig. 3).
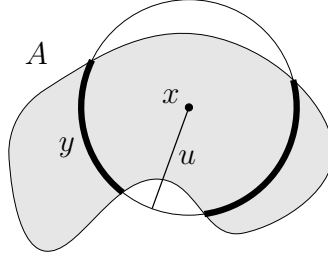
Figure 3: Points y within the shape A whose distance to the specific point x is equal to u.

The descriptor $h_{A,x}$ of the shape A is mathematically defined as the distribution (histogram) of the distance values $d(y, x)$ from the reference point x.

$$h_{A,x}(u) = \sum_{y \in A} \chi_{\{0\}}(d(y, x) - u) \tag{4}$$

where χ is the characteristic function.

In [26], the histogram of gray levels at distance u is used for defining the descriptor of a gray-level image, while here it is the number of pixels at distance u.

The proposed shape metric DM between two shapes A and B, with its two reference points x and y, is then defined as the Manhattan distance between the two descriptors:

$$\text{DM}(A, x, B, y) = \sum_d |h_{A,x}(u) - h_{B,y}(u)| \tag{5}$$

Note that these two shapes A and B will later play the role of two GANs for which the dissimilarity is measured in the displacement field computation (Eq. 2). The descriptor of each GAN will naturally be built with its seed (current) point as reference point.

DM is invariant to rigid transformations (translation and rotation): if T denotes such a transformation, $\text{DM}(A, x, T(A), T(x)) = 0$. Note that it is also invariant to mir-

12

ror transformations which can lead to some ambiguities in certain applications. Computing this distance between two shapes is few time consuming since the descriptor of each shape can be quickly calculated by using a distance map [43].

4.1.2. Performance tests

To evaluate the performance of the proposed shape metric DM, a first comparison
between pairs of similar shapes -from the Kimia dataset [44]- along a continuum is
performed (Tab. 1). Thereafter, the distance is tested (Tab. 2) with different categories
of animals (fishes, rabbits, birds and camels) from the same dataset [44] for shape
classification. The distance between shapes are computed by using the equation 5 and
the reference points used for the shapes are their barycenters. In this way, the distance
between 2 shapes A and B is given by: $DM(A, \mathrm{bar}(A), B, \mathrm{bar}(B))$ where $\mathrm{bar}(A)$ and
$\mathrm{bar}(B)$ denote the barycenter of A and B, respectively. Each bin of the histogram
counts pixels within a distance range of three pixels for computing the histograms
$h_{A,\mathrm{bar}(A)}$ and $h_{B,\mathrm{bar}(B)}$ of the shapes A and B.

Table 1: Metric comparison between pairs of shapes along a continuum. For a better readability, the values have been divided by 20000. The first rows shows that as the quadrangle moves toward the triangle, the distance increases. An exception is observed for the fourth shape because of a best match with a π rotation angle of the shape (DM is invariant to mirror transformations).

	0.00	0.10	0.18	0.08	0.24	0.29	0.38
	0.10	0.00	0.10	0.15	0.33	0.37	0.46
	0.18	0.10	0.00	0.22	0.41	0.45	0.54
	0.08	0.15	0.22	0.00	0.21	0.25	0.34
	0.24	0.33	0.41	0.21	0.00	0.07	0.14
	0.29	0.37	0.45	0.25	0.07	0.00	0.09
	0.38	0.46	0.54	0.34	0.14	0.09	0.00

In Table 1, the first rows shows that as the quadrangle moves toward the triangle,

the distance increases. An exception is observed for the fourth shape because of a best match with a π rotation angle of the shape.

Table 2: Metric for the dataset. For a better readability, the values have been divided by 10000. The animals are well categorized: for each tested animal (in one row), the three best results are achieved for the animals of its category.

	0.00	0.07	0.07	0.16	0.13	0.20	0.24	0.17	0.18	0.46	0.51	0.50
	0.07	0.00	0.03	0.14	0.08	0.13	0.17	0.10	0.11	0.39	0.44	0.43
	0.07	0.03	0.00	0.14	0.09	0.14	0.17	0.11	0.11	0.39	0.44	0.43
	0.16	0.14	0.14	0.00	0.08	0.06	0.13	0.08	0.09	0.29	0.34	0.33
	0.13	0.08	0.09	0.08	0.00	0.06	0.14	0.09	0.10	0.32	0.37	0.36
	0.20	0.13	0.14	0.06	0.06	0.00	0.11	0.08	0.09	0.26	0.31	0.30
	0.24	0.17	0.17	0.13	0.14	0.11	0.00	0.08	0.09	0.25	0.32	0.30
	0.17	0.10	0.11	0.08	0.09	0.08	0.08	0.00	0.04	0.29	0.34	0.33
	0.18	0.11	0.11	0.09	0.10	0.09	0.09	0.04	0.00	0.27	0.33	0.32
	0.46	0.39	0.39	0.29	0.32	0.26	0.25	0.29	0.27	0.00	0.08	0.08
	0.51	0.44	0.44	0.34	0.37	0.31	0.32	0.34	0.33	0.08	0.00	0.07
	0.50	0.43	0.43	0.33	0.36	0.30	0.30	0.33	0.32	0.08	0.07	0.00

Regarding Table 2, the animals are well categorized. Indeed, for each tested animal (in one row), the three best results are achieved for the animals of its category.

In the following experiments, the noise sensitivity of the proposed shape metric is studied on the whole Kimia dataset. The performance is evaluated by considering two types of noise:

- Pepper noise: a percentage of all the pixels are reassigned as background pixels (black pixels). In the experiments, the noise amount is related to this percentage value going from 0.05 to 0.5 (with a step of 0.05).

- Boundary noise: pixels on the image border are spatially randomly reassigned from their initial location to produce the deformed shape. In the experiments,

the noise amount is related to the maximal displacement for both the x and y coordinates, going from -5 to 5 pixels (excluding 0, with a step of 1).

Figure 4 illustrates these two kinds of noise with the minimal and maximal amount of noise according to the selected parameters. In Figure 5 the average value of the dissi-

(a) original image   (b) minimal boundary noise   (c) maximal boundary noise

(d) minimal pepper noise (e) maximal pepper noise
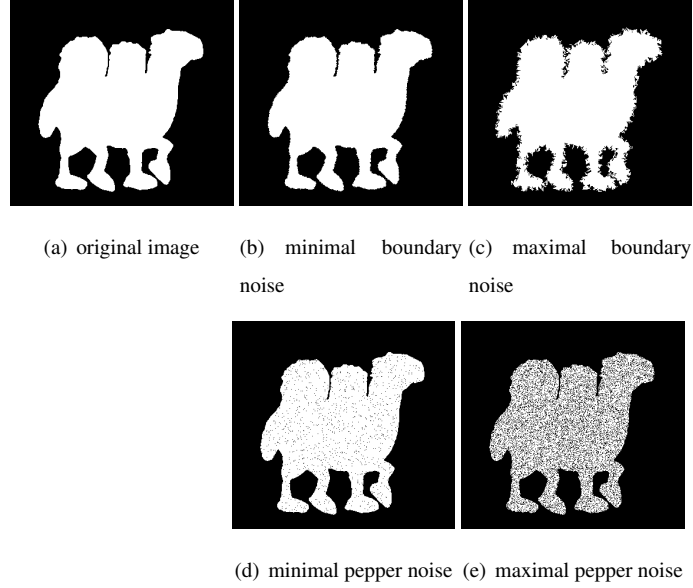
Figure 4: Damaged images with pepper and boundary noise using the minimal and maximal amount of noise according to the selected parameters.

malirity measure on the whole Kimia dataset is given as a function of the noise amount which is labeled from 1 to 10 for both boundary and pepper noise parameters. One can notice that the proposed shape metric is more sensitive to pepper noise than boundary noise. Indeed, the proposed descriptor is a region-based descriptor: each value of the histogram can be affected by several pixels for the pepper noise as opposed to the boundary noise.

Note that the Manhattan distance used in the dissimilarity measure DM could be replaced by the Earth mover's distance if one wishes to be less sensitive with respect to a small shift of the reference points of the considered shapes. Indeed, with such a shift, some pixels of a shape counted up to a specific bin of the histogram could move to the neighboring bins. In this case, the Earth mover's distance would be more robust than
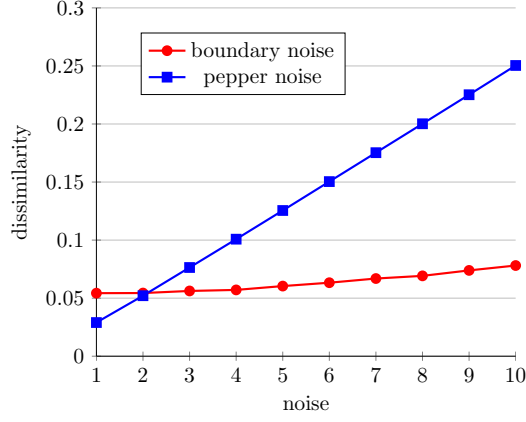
Figure 5: Noise sensitivity of the proposed dissimilarity measure. The resulting values have been normalized by the area of the original image. The noise amount is labeled from 1 to 10 for both boundary and pepper noise parameters.

the Manhattan distance.

All these experiments highlight the performance (continuity, classification accuracy, noise sensitivity) of the proposed shape metric.

4.2. Application to GANs

Using the proposed shape metric, it is possible to measure the dissimilarity between two GANs: $\mathrm{DM}(V_m^I(x), x, V_m^J(y), y)$, so as to find the best match according to Equation 2. As mentioned before, it is important to note that the descriptor of a GAN is naturally built with its seed (current) point from which it has been computed, and not with its barycenter. In this way, GANs with the same geometrical shape but with different seed positions can be discriminated (Fig. 2) to avoid mismatches.

The next section presents some qualitative results for the displacement field computation and thereafter for the registration of MR brain and retinal images, before presenting quantitative results.

5. Experimental results

This section aims to evaluate the performance of the proposed GAN-based registration method in comparison with the block matching method. Both algorithms firstly give a raw displacement vector field which is further used as input for the robust estimation of the transformation using the least trimmed squares minimization. The results of qualitative and quantitative experiments are given for standard and real images.

5.1. Displacement field estimation

Figures 6 and 7 compare the raw displacement fields computed with the block and GAN matching algorithms at the first iteration with the pyramidal level of highest resolution.

Regarding Figure 6, the GAN-based displacement vectors are qualitatively accurate. Moreover, contrary to the classical block matching, the proposed method is able to compute accurate displacement vectors inside homogeneous areas as highlighted in the proposed binary images.

As one can notice in Figure 7, the GAN matching method provides well-estimated displacement fields while much more outliers occur with the block matching method. Such a GAN-based displacement field is suitable for image registration.

5.2. Qualitative evaluation

5.2.1. Illustration on MR brain images

To qualitatively evaluate the GAN registration results, a rigid registration is performed on MR brain images. Proton density (PD) imaging and T1 imaging are used to discriminate anatomical structures based on their proton density or their T1 relaxation properties. Examples of intra PD/PD (Fig. 8) and inter T1/PD (Fig. 9) MR brain image registration are shown. The reference and floating images are superimposed before and after registration. After registration, the images superimpose perfectly: the rigid transformation is well estimated.

It can be noted that even with dissimilar intensities (Fig. 9), the registration succeeds: it is one of the strength of the proposed method, it is able to register data from different modalities. Indeed, the patterns of the two input images have similar shapes

17

(a) reference image I_1 and displacement field computed with block matching

(b) reference image I_1 and displacement field computed with GAN matching

(c) floating image J_1

(d) reference image I_2 and displacement field computed with block matching

(e) reference image I_2 and displacement field computed with GAN matching

(f) floating image J_2
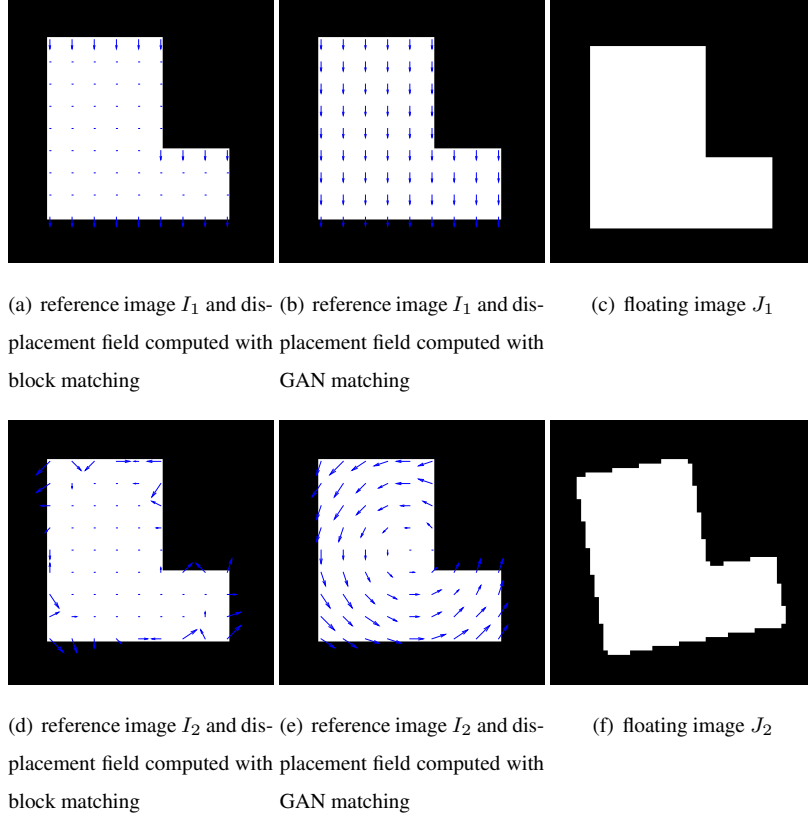
Figure 6: Displacement field computation between two binary images (I_1, J_1) and (I_2, J_2) using block and GAN matching. The parameter m of the GAN is fixed to 0 while the size of the blocks is fixed to $r = 7$ and the vectors are only computed for pixels within the object. Contrary to the classical block matching, the proposed GAN-based method is able to compute accurate displacement vectors inside homogeneous areas.
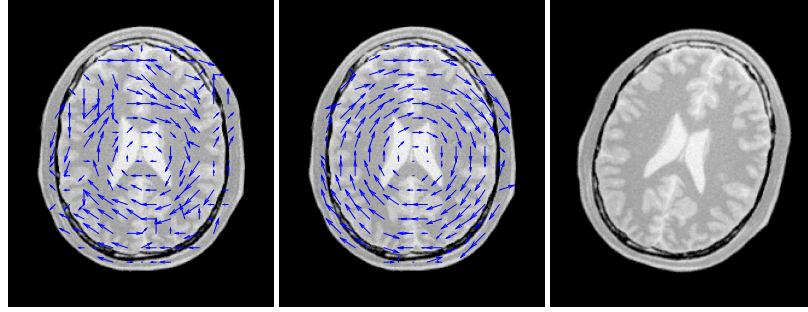
(a) reference image I and displacement field computed with block matching   (b) reference image I and displacement field computed with GAN matching   (c) floating image J

Figure 7: Displacement field computation between two MR brain images (I, J) using block and GAN matching. The GANs are computed with the luminance criterion and the tolerance is fixed to $m = 50$ while the size of the blocks is fixed to $r = 7$. The displacements vectors computed by the GAN matching method are qualitatively better estimated in comparison with the block matching method. Such a GAN-based displacement field is suitable for image registration.
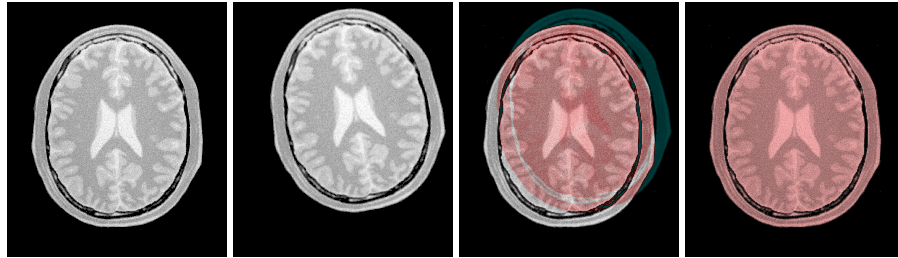
(a) PD MR brain image   (b) PD MR brain image   (c) before registration   (d) after registration using the proposed GAN matching

Figure 8: PD/PD registration of MR images. The rigid transformation is well-estimated thanks to the GAN matching.
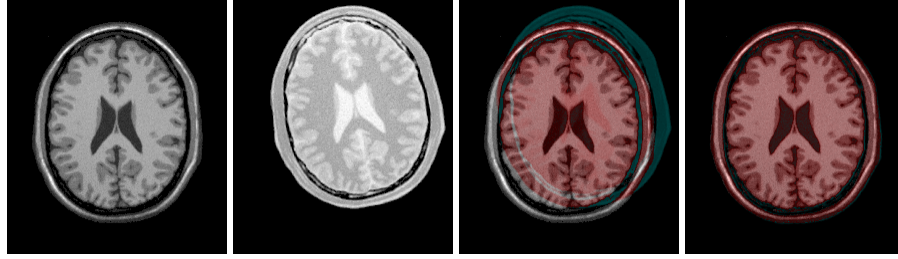
(a) T1 MR brain image   (b) PD MR brain image   (c) before registration   (d) after registration us-
ing the proposed GAN
matching

Figure 9: T1/PD registration of MR images. The structures within the two input images being similar, the
GAN matching enables a robust rigid transformation to be calculated. The dissimilarity measure compares
the spatial structures and not the local intensities.

and consequently, GANs are consistently and accurately matched, and the optimization
algorithm performed on the resulting displacement field provides the expected trans-
formation.

5.2.2. Illustration on retinal images

In this subsection, an application on retinal images is investigated. The global ob-
jective is to reconstruct a comprehensive single image from multiple spatial images
acquired from the same patient, in order to facilitate the disease diagnosis and treat-
ment planning. In Figure10, the spatial rigid registration of a pair of retinal images
is presented both with the classical block matching and the proposed GAN matching
method.

One can note that the rigid transformation is badly estimated with the block match-
ing method, contrary to the proposed GAN matching one. The main reason seems to be
related to the presence of several uniform regions. Indeed, local displacements cannot
be accurately estimated in homogeneous areas with fixed-size and fixed-shape blocks.
Mismatches between such blocks often occur contrary to the correspondances between
GANs which are spatially adapted to the underlying regions. For a full registration of
such a pair or retinal images, note that nonlinear registration methods are required since
the retina is not a planar surface (projective transformations are consequently needed

20
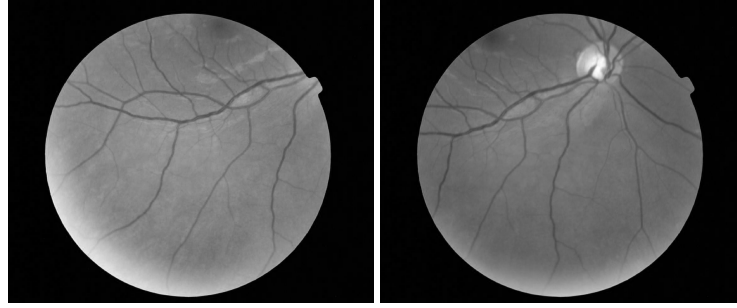
(a) first retinal image   (b) second retinal image

(c) before registration
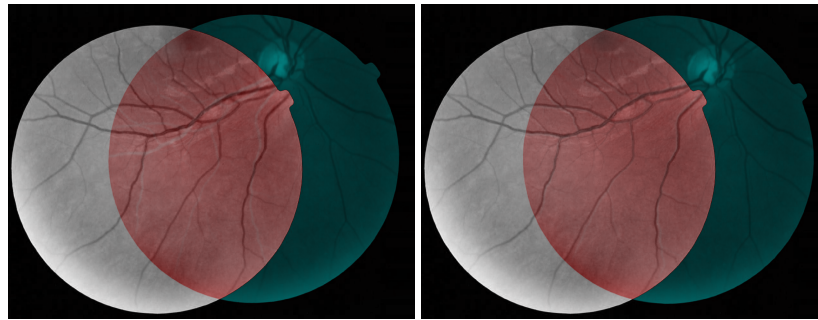
(d) after registration using the classical block matching   (e) after registration using the proposed GAN matching

Figure 10: Rigid registration of a pair or retinal images by comparing the block matching and the GAN matching. One can note that the classical method fails contrary to the proposed GAN-based method.

for such 2-D images).

5.3. Quantitative evaluation

5.3.1. GAN vs. block matching

The GAN matching is here quantitatively evaluated for rigid image registration and compared to the classical block matching on a dataset of twelve standard images of size 256×256 pixels, from Matlab (Fig. 11). For both algorithms, the following

Figure 11: Image dataset for the performance evaluation of the proposed GAN matching algorithm.

parameter values are used: $N = 3$ pixels for the neighborhood search, $\Delta_x = \Delta_y = 5$ pixels for the density of the displacement field, $i = 3$ for the number of pyramid levels and $n = 10$ for the number of iterations at each pyramid level. The resampling uses bilinear interpolation. In addition, the size of the GAN (resp. block) is controlled by the parameter m (resp. the size r of the block) and is fixed to $m = 35$ (resp. $r = 7$). The metric used for comparing blocks in the block matching algorithm is the sum of squared differences between pixel gray levels.

For each image of the dataset, 100 random rigid transformations T are generated. T is then applied to the original image I leading to the transformed image J. The

22

registration between I and J is then performed by using the block and GAN matching, giving an estimated transformation $\hat{T}$ is given.

The registration is evaluated in terms of robustness, capture range and accuracy such as proposed in [25]:

- The robustness is defined as the percentage of random experiments being a success. The registration is considered a success if the final warping index $\omega_f = \frac{1}{\#I} \sum_{v \in I} ||T(v) - \hat{T}(v)||$ is lower than 1 pixel.

- The capture range of the algorithm is defined as the maximum initial warping index $\omega_i = \frac{1}{\#I} \sum_{v \in I} ||T(v) - v||$ for which it succeeds.

- The accuracy is defined as the average ω_f for which it succeeds.

The performance of the algorithms is evaluated for small, medium and large displacements:

- T_{small}: random rotation in the range $[0, 20]$ degrees and random translation in the range $[0, 5]$ pixels

- T_{medium}: random rotation in the range $[20, 40]$ degrees and random translation in the range $[5, 10]$ pixels

- T_{large}: random rotation in the range $[40, 60]$ degrees and random translation in the range $[10, 15]$ pixels

The results are summarized in Table 3. The different values denote the average for the 100 simulations and the twelve images of the dataset.

These results demonstrate that GAN and block matching algorithms achieve perfect registrations for small displacements with a robustness of 100%. For medium and large displacements, the GAN matching provides better results than the block matching in terms of robustness, capture range and accuracy. Indeed, contrary to the GANs, searching similar spatially-invariant blocks within the original and transformed images for such transformations (rigid displacements) often leads to wrong displacement vectors which explain the difficulty of the block matching to well register the images.

23

Table 3: Quantitative evaluation of the proposed GAN matching (GM) algorithm with comparison to the block matching (BM) for 100 simulated transformations on a dataset of twelve images.

Displacement	Algorithm	Robustness (%)	Cap range (px)	Accuracy (px)
Small	BM	100.00	33.41	0.24
	GM	100.00	33.41	0.20
Medium	BM	93.58	65.62	0.27
	GM	99.92	65.95	0.19
Large	BM	64.42	91.71	0.27
	GM	81.08	95.43	0.19

5.3.2. Impact of the GAN homogeneity tolerance

In this paragraph, the impact of the GAN homogeneity tolerance m is studied. For these experiments, the 100 transformations are identical for each homogeneity tolerance value m so as to only evaluate the impact of this parameter. The GAN registration is achieved by using different values m varying from 0 to 50 with a step size of 5. The following graph (Fig. 12) shows the robustness of the proposed method as a function of the homogeneity tolerance values m of the GANs for small, medium and large displacements. The computed values denote the average for the 100 simulations and the 12 images of the dataset.

For small and medium displacements, the two curves in Figure 12 are monotonic while a maximum exists for the curve associated to large displacements. The value $m = 35$ gives good results for the small, medium and large displacements. Note that this value depends on the underlying images: it does not give necessarily the best performance for each individual image.

One can see the decreasing performance from $m > 35$ for large displacements. Indeed, the GANs are truncated by the image borders caused by the transformation which is too large for the observation window. It means that for large displacements the spatial regions are not identical in the two images to be registered. It consequently leads to mismatches with the proposed shape distance. Using another distance between GANs, less sensitive to truncations or occlusions such as proposed in [13] for example,
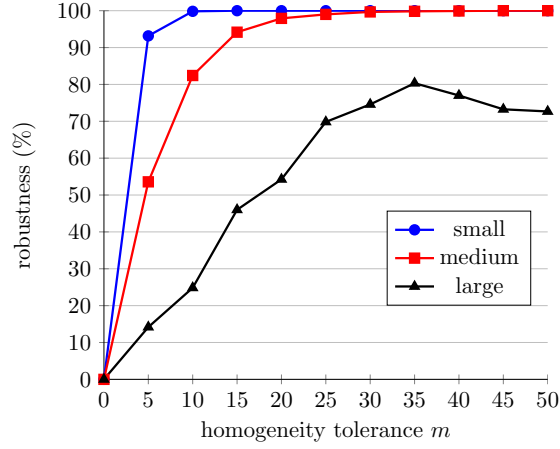
24

Figure 12: Impact of the homogeneity tolerance for GAN rigid registration with small, medium and large displacements.

would probably increase the value of m from which the robustness starts to decrease.

For small and medium displacements, the same behavior would probably appear but for larger m values ($m > 50$). Indeed, the larger the homogeneity tolerance value m is, the larger the GANs are. For small and medium displacements, mainly very large spatial regions (i.e. GANs computed with large values m) will be truncated by the observation window.

5.3.3. Noise sensitivity

In this subsection, the performance of the proposed rigid image registration is evaluated in the presence of noise. In Figure 14 the robustness of the GAN matching and block matching methods is given as a function of the noise amount for small, medium and large displacements. As in the previous experiments, the computed values denote the average for the 100 simulations and the 12 images of the dataset. Both the reference and floating images are damaged by an additive Gaussian noise. The noise amount is related to its variance going from 0 to 0.02 with a step of 0.005. Figure 13 illustrates two damaged images of 'Lena' with the minimal (0.005) and maximal (0.02) amount of noise. The results show that in noisy conditions the GAN matching method outperforms the block matching one. Moreover, the proposed method is less sensitive to

25

(a) original image   (b) minimal Gaussian noise   (c) maximal Gaussian noise
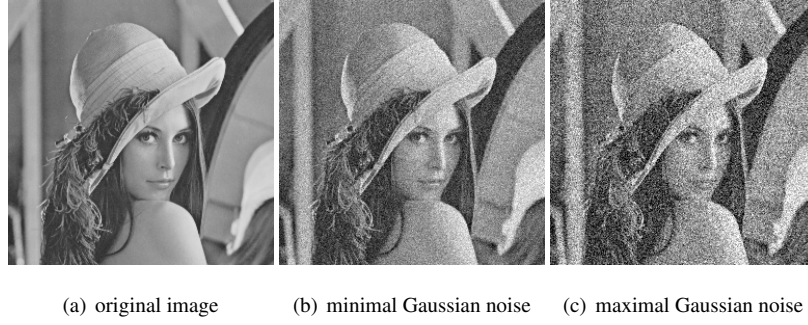
Figure 13: Damaged images with an additive Gaussian noise using the minimal (var=0.005) and maximal (var=0.02) amount of noise according to the selected parameters.
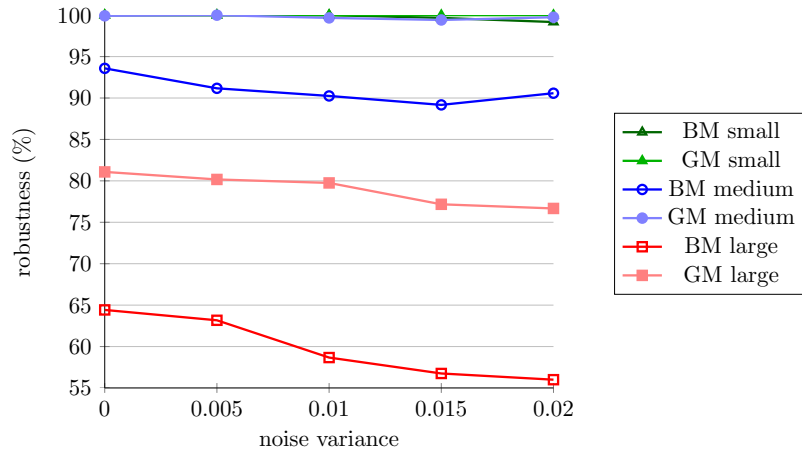

Figure 14: Noise robustness of the block and GAN matching methods as a function of the noise variance for small, medium and large displacements.

Gaussian noise than the classical method: the overall approximated slope of each GAN

455 matching curve is lower than its corresponding block matching curve.

5.3.4. Computational time

In this subsection, some experiments have been done to give the computational time for both the block matching and GAN matching methods. The algorithms have been run on a 64 bits Dual Intel Xeon X5650 Processor (2 x 6 cores at 2.66Ghz) with the

460 Matlab software running in parallel on 22 threads. In Table 4, the average CPU time per image (computed on the image database for which the registration is a success) are given in seconds. The rotations R and translations t have been fixed to ($R = 10^o, t = 2.5$px), ($R = 30^o, t = 7.5$px), ($R = 50^o, t = 12.5$px) to represent a small, medium and large displacements, respectively.

Table 4: Computational time (in average per image) of the proposed GAN matching (GM) algorithm with comparison to the block matching (BM) algorithm for small, medium and large displacements.

Displacement	Algorithm	Time (s)
Small	BM	19
	GM	182
Medium	BM	28
	GM	172
Large	BM	34
	GM	195

465 One can note that the GAN matching is less efficient with respect to computation time than the block matching but it provides better results in terms of robustness and accuracy. It is also interesting to note that the computational time for GAN matching does not increase with the size of the deformation.

6. Conclusion

470 In this paper, a new hybrid (feature and intensity-based) image registration algorithm has been presented. It is based on the block matching algorithm and proposes a

27

new way to calculate the displacement field by matching spatially-variant sub-blocks of the images, called General Adaptive Neighborhoods (GANs) [2]. These neighborhoods represent the patterns within the grayscale images: they are adaptive with respect to the intensities and the spatial structures of the image. The GANs are determined by the image itself and consequently enable a robust and accurate estimation of the displacement field by GAN matching. A consistent shape metric has also been introduced so as to match the GANs. It provides a good trade-off between performance and computational cost. Several tests have been carried out on a shape dataset to evaluate the performance of the proposed metric. Thereafter, all information of the computed displacement field have been gathered to estimate (by an optimization algorithm) the transformation between two images for image registration. According to the performed experiments, the proposed GAN-based registration method works better than the classical BM method for image rigid registration and opens large perspectives for nonlinear approaches [45]. Indeed, for nonlinear registration, the same process could be applied in order to extract a robust vector field. The main constraint concerns the dissimilarity measure used for matching GANs that should be modified so as to get new invariance and/or robustness properties [13] with respect to the expected tranformation.

[1] S. Ourselin, A. Roche, S. Prima, N. Ayache, Block matching: A general framework to improve robustness of rigid registration of medical images, in: International Conference of Medical Robotics, Imaging And Computer Assisted Surgery, Vol. 1935 of Lectures Notes in Computer Science, Springer, Pittsburg, Pennsylvanie, USA, 2000, pp. 557–566.

[2] J. Debayle, J. C. Pinoli, General Adaptive Neighborhood Image Processing - Part I: Introduction and Theoretical Aspects, Journal of Mathematical Imaging and Vision 25 (2) (2006) 245–266.

[3] M. Jenkinson, S. Smith, A global optimisation method for robust affine registration of brain images, Medical Image Analysis 5 (2) (2001) 143–156.

[4] J. Feldmar, N. Ayache, Rigid, affine and locally affine registration of free-form surfaces., International Journal of Computer Vision 18 (1996) 99–119.

[5] O. Commowick, V. Arsigny, A. Isambert, J. Costa, F. Dhermain, F. Bidault, P. Bondiau, A. Ayache, G. Malandain, An efficient locally affine framework for the smooth registration of anatomical structures, Medical Image Analysis 12 (4) (2008) 427–441.

[6] C. Davatzikos, Spatial transformation and registration of brain images using elastically deformable models, Computer Vision and Image Understanding 66 (2) (1997) 207–222.

[7] R. Gallea, E. Ardizzone, R. Pirrone, O. Gambino, Three-dimensional fuzzy kernel regression framework for registration of medical volume data, Pattern Recognition 46 (11) (2013) 3000–3016.

[8] E. D'Agostino, F. Maes, D. Vandermeulen, P. Suetens, A viscous fluid model for multimodal nonrigid image registration using mutual information, Medical Image Analysis 7 (2003) 565–575.

[9] N. Alpert, M. Berdichevsky, D. Levin, E. Morris, A. Fischman, Improved methods for image registration, NeuroImage 3 (1996) 10–18.

[10] G. Brown, A survey of image registration techniques, ACM Computing Surveys 24 (4) (1992) 325–376.

[11] A. Roche, G. Malandain, X. Pennec, N. Ayache, The correlation ratio as a new similarity measure for multimodal image registration, in: First International Conference on Medical Image Computing and Computer-Assisted Intervention, Vol. 1496 of Lecture Notes in Computer Science, Springer, Cambridge (USA), 1998.

[12] P. Viola, Alignment by maximisation of mutual information, International Journal of Comuter Vision 24 (2) (1997) 147–154.

[13] B. W. Hong, S. Soatta, Shape matching using multiscale integral invariants, IEEE Transactions on Pattern Analysis and Machine Intelligence 37 (1) (2015) 151–160.

[14] A. Collignon, F. Maes, D. Delaere, D. Vandermeuden, P. Suetens, G. Marchal, Automated multi-modality image registration based on information theory, in: International Conference on Information Processing in Medical Imaging, Kluwer Academic Publishers, 1995, pp. 263–274.

[15] G. Taubin, An improved algorithm for algebraic curve and surface fitting, in: International Conference on Computer Vision, 1993, pp. 658–665.

[16] J. Maintz, M. Viergever, A survey of medical image registration, Medical Image Analysis 2 (1) (1998) 1–36.

[17] A. Rangarajan, H. Chui, E. Mjolsness, S. Pappu, L. Davachi, P. Goldman-Rakic, J. Duncan, A robust point-matching algorithm for autoradiograph alignment, Mediacl Image Analysis 1 (4) (1997) 379–398.

[18] F. Cohen, Z. Yang, Z. Huang, J. Nissanov, Automatic matching of homologous histological sections, IEEE Transactions on Biomedical Engineering 45 (5) (1998) 642–649.

[19] B. Zitova, J. Flusser, Image registration methods: a survey, Image and Vision Computing 21 (2003) 977–1000.

[20] P. Cachier, J. Mangin, X. Pennec, D. Rivire, D. Papadopoulos-Orfanos, J. Rgis, N. Ayache, Multisubject non-rigid registration of brain mri using intensity and geometric features, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Vol. 2208 of Lecture Notes in Computer Science, Springer, Utrecht, The Netherlands, 2001, pp. 732–742.

[21] P. Hellier, C. Barillot, Cooperation between local and global approaches to register brain images, in: International Conference on Information Processing in Medical Imaging, Vol. 2082 of Lectures Notes in Computer Science, Davis, CA, USA, 2001, pp. 315–328.

[22] H. Johnson, G. Christensen, Consistent landmark and intensity-based image registration, IEEE Transactions on Medical Imaging 21 (5) (2002) 450–461.

[23] K. Rohr, P. Cathier, S. Worz, Elastic registration of electrophoresis images using intensity information and point landmarks, Pattern Recognition 37 (5) (2004) 1035–1048.

[24] D. Shen, Image registration by local matching histogram, Pattern Recognition 40 (4) (2007) 1161–1172.

[25] O. Commowick, N. Wiest-Daessle, S. Prima, Block-matching strategies for rigid registration of multimodal medical images, in: Proceedings of the IEEE International Symposium on Biomedical Imaging, Barcelona, Spain, 2012.

[26] S. Lazebnik, C. Schmid, J. Ponce, A sparse texture representation using local affine regions, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (8) (2005) 1265–1278.

[27] J. Debayle, J. C. Pinoli, Spatially Adaptive Morphological Image Filtering using Intrinsic Structuring Elements, Image Analysis and Stereology 24 (3) (2005) 145–158.

[28] J. Debayle, J. C. Pinoli, General Adaptive Neighborhood Image Processing - Part II: Practical Application Examples, Journal of Mathematical Imaging and Vision 25 (2) (2006) 267–284.

[29] G. Choquet, Topology, Academic Press, New-York, U.S.A., 1966.

[30] S. G. Mallat, A Theory for Multiresolution Decomposition : The Wavelet Representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (1989) 674–693.

[31] T. Lindeberg, Scale-Space Theory: a Basic Tool for Analysing Structures at Different Scales, Journal of Applied Statistics 21 (2) (1994) 225–270.

31

[32] P. Perona, J. Malik, Scale-Space and Edge Detection using Anisotropic Diffusion, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (7) (1990) 629–639.

[33] G. R. Arce, R. E. Foster, Detail-Preserving ranked-order based filters for image processing, IEEE Transactions on Acoustics, Speech, and Signal Processing 37 (1) (1989) 83–98.

[34] R. Gordon, R. M. Rangayyan, Feature Enhancement of Mammograms using Fixed and Adaptive Neighborhoods, Applied Optics 23 (4) (1984) 560–564.

[35] Z. Zhang, Parameter estimation techniques: a tutorial with application to conic fitting, Journal of Image and Vision Computing 15 (1) (1997) 79–98.

[36] D. Huttenlocher, G. Klanderman, W. Rucklidge, Comparing images using hausdorff distance, IEEE Transcations on Pattern Analysis and Machine Intelligence 15 (9) (1993) 850–863.

[37] R. C. Veltkamp, L. J. Latecki, Data Science and Classification, Springer, 2006, Ch. Properties and Performance of Shape Similarity Measures, pp. 47–56.

[38] T. Peli, An algorithm for recognition and localization of rotated and scaled objects, Proceedings of the IEEE 69 (1981) 483–485.

[39] A. Goshtasby, Description and discrimination of planar shapes using shape matrices, IEEE Transactions on Pattern Analysis and Machine Intelligence 7 (1985) 738–743.

[40] M. Hu, Visual pattern recognition by moment invariants, IRE Transactions on Information Theory 8 (1962) 179–187.

[41] J. Flusser, T. Suk, Pattern recognition by affine moment invariants, Pattern Recognition 26 (1993) 167–174.

[42] Y. Y. Tang, T. Xia, Y. Wei, H. Li, L. Li, Hierarchical kernel-based rotation and scale invariant similarity, Pattern Recognition 47 (7) (2014) 1674–1688.

[43] C. R. Maurer, R. Qi, V. Raghavan, A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (2) (2003) 265–270.

[44] D. Sharvit, J. Chan, H. Tek, B. Kimia, Symmetry-based indexing of image databases, Journal of Visual Comunication and Image Representation 9 (4) (1998) 366–388.

[45] J. Modersitzki, FAIR: Flexible Algorithms for Image Registration, SIAM, 2009.

610