Pattern Recognition 55 (2016) 45–57

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Rigid image registration by General Adaptive Neighborhood matching

Johan Debayle [a,*,1], Benoit Presles [b,c,1]

[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 UMR 5220, Inserm U1044, INSA-Lyon, Université Lyon 1, Villeurbanne, France
[c] Léon Bérard Cancer Center, University of Lyon, France

ARTICLE INFO

Article history:
Received 29 January 2015
Received in revised form
15 December 2015
Accepted 21 January 2016
Available online 4 February 2016

Keywords:
Block matching
Displacement field
General Adaptive Neighborhoods
Image registration
Shape metric

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 (Ourselin et al., 2000 [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) (Debayle and Pinoli, 2006 [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 us 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.

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], etc.);
- similarity criteria (sum of squared differences [9], correlation coefficient [10], correlation ratio [11], mutual information [12], multiscale integral invariants [13], etc.); and
- optimization algorithms (Powell method [14], Levenberg–Marquardt method [15], stochastic search [16], etc.).

However, registration methods can be classified 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], etc. 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*, relies 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, etc.), the approach consists in 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. Second, the assumption of a global relationship between the image intensities may be violated by the presence of various image artefacts.

* Corresponding author. Tel.: +33 477420219; fax: +33 477499694.
E-mail addresses: debayle@emse.fr (J. Debayle), benoit.presles@creatis.insa-lyon.fr (B. Presles).

[1] Both authors contributed equally to this work.

http://dx.doi.org/10.1016/j.patcog.2016.01.024
0031-3203/

1.1. Presentation of the problem

The problems mentioned above have been addressed by hybrid registration methods. Such methods combine feature and intensity information [1,20–24]. However, these methods still require a pre-segmentation of features or take spatial regions of interest independent 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 independent 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 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 quantitative 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 Section 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, and 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, and 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 : $\forall y \in V_m^h(x) \ |h(y) - h(x)| \le m$;
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 $x \in D$.

Fig. 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, 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