Computer Vision and Image Understanding 115 (2011) 1340–1354

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

A comparative study of state-of-the-art evolutionary image registration methods for 3D modeling ☆

J. Santamaría [a,*], O. Cordón [b,c], S. Damas [b]

[a] Dept. of Computer Science, EPS Linares, University of Jaén, 23700 Linares, Jaén, Spain
[b] European Centre for Soft Computing, Edif. Investigación, Mieres, Asturias, Spain
[c] Dept. of Computer Science and Artificial Intelligence, ETSIIT, University of Granada, Granada, Spain

ARTICLE INFO

Article history:
Received 19 February 2010
Accepted 20 May 2011
Available online 30 May 2011

Keywords:
Image registration
Evolutionary computation
Metaheuristics
Computer vision
Range images
3D modeling
ICP

ABSTRACT

Image registration (IR) aims to find a transformation between two or more images acquired under different conditions. This problem has been established as a very active research field in computer vision during the last few decades. IR has been applied to a high number of real-world problems ranging from remote sensing to medical imaging, artificial vision, and computer-aided design. Recently, there is an increasing interest on the application of the evolutionary computation paradigm to this field in order to solve the ever recurrent drawbacks of traditional image registration methods as the iterated closest point algorithm. Specially, evolutionary image registration methods have demonstrated their ability as robust approaches to the problem. Unlike classical IR methods, they show the advantage of not requiring a good initial estimation of the image alignment to proceed. In this contribution, we aim to review the state-of-the-art image registration methods that lay their foundations on evolutionary computation. Moreover, we aim to analyze the performance of some of the latter approaches when tackle a challenging real-world application in forensic anthropology, the 3D modeling of forensic objects.

1. Introduction

Image registration (IR) [1–3] is a fundamental task in computer vision (CV) used to finding either a spatial *transformation* (e.g, rotation, translation, etc.) or a correspondence (matching of similar image features) among two or more images acquired under different conditions: at different times, using different sensors, from different viewpoints, or a combination of them. IR aims to achieve the best possible overlapping transforming those independent images into a common one. Over the years, IR has been applied to tackle many real-world problems ranging from remote sensing to medical imaging, artificial vision, and computer-aided design (CAD). Likewise, different techniques facing the IR problem have been studied resulting in a large body of research. Several recent contributions reviewing the state of the art on IR methods can be found in [1–5].

In a nutshell, IR involves finding the optimal *transformation* achieving the best fitting between typically two *images*, usually called scene and model. They both are related by the said transformation and the degree of resemblance between them is measured by a *similarity metric*. Such transformation estimation is usually formulated as an *optimization* problem solved by an iterative procedure in order to properly explore the search space of candidate solutions to the problem. The optimization process applied by traditional IR methods is highly influenced by image noise, image discretization, and orders of magnitude in the scale of the IR transformation parameters, among other phenomena. Specially, that is the case of the approaches based on the classical iterative closest point (ICP) algorithm [6,7], which are likely to provide incorrect registration transformation estimations. This is due to the fact that those methods are usually prone to be trapped in local minima [8–11] since they assume a rough prealignment of the images typically provided by the user.

After a couple of decades, evolutionary computation (EC) [12] has demonstrated its ability to deal with complex real-world problems in CV and image processing. As an example, several special issues on the topic have been published in international journals in the last few years [13–15]. In particular, evolutionary algorithms (EAs) [12,16] have been successfully applied to tackle IR problems without requiring a good initial estimation of the image alignment. That advantage is mainly motivated by the global optimization nature of evolutionary approaches, which allows them to perform a robust search in complex and ill-defined search spaces.

☆ This work is partially supported by both the Spanish Ministerio de Educación y Ciencia (Ref. TIN2009-07727) including EDRF fundings and the University of Jaén (Ref. R1/12/2010/61) including fundings from *Caja Rural de Jaén*.
* Corresponding author. Fax: +34 953648559.
E-mail addresses: jslopez@ujaen.es (J. Santamaría), oscar.cordon@softcomputing.es (O. Cordón), sergio.damas@softcomputing.es (S. Damas).

1077-3142/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2011.05.006

The first attempts to face the IR problem using EC can be found in the eighties [17]. Since then, evolutionary IR (EIR) has become a very active area and several well-known EAs have been considered to tackle the IR optimization process, causing an outstanding interest [18–28]. Nevertheless, those EIR methods have not been covered by any of the IR surveys existing in the specialized literature. The aim of the current contribution is to bridge that gap in a two-fold manner. On the one hand, by reviewing the extensive literature in EIR. On the other hand, by developing an experimental study on the performance of 12 EIR methods when tackling the 3D modeling of some real-world forensic objects digitized by a laser range scanner.

The structure of this contribution is as follows. Section 2 describes the IR problem. Next, Section 3 describes the key concepts of he EC paradigm, it presents the first EIR methods and it reviews the state-of-the-art EIR methods. Section 4 is devoted to a deep experimental study developed on the said real-world IR application. Finally, some conclusions are drawn in Section 5.

2. Image registration

There is not a universal design for a hypothetical IR method that could be applicable to every real-world application [3]. However, IR methods consist of the following four components:

- Two input **Images** named as *scene* $I_s = \{\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n\}$ and *model* $I_m = \{\vec{p}'_1, \vec{p}'_2, \ldots, \vec{p}'_m\}$, with $\vec{p}_i$ and $\vec{p}'_j$ being image points.
- A **Registration transformation** *f*, relating the two images. Typically, it is a parametric function.
- A **Similarity metric function** *F*. It aims to measure a qualitative value of closeness or degree of fitting between the transformed scene image, noted by $f(I_s)$, and the model image.
- An **Optimizer**. It is a method that seeks the optimal transformation *f* inside the defined solution search space.

Likewise, an iterative process is often followed (see Fig. 1). It usually finishes when convergence is achieved, i.e., when the similarity metric is bellow a given tolerance threshold. In this work, we focused our attention on the optimizer component which is of crucial importance in the success of any IR method. In particular, two search approaches for optimization have been considered in the IR literature [3]:

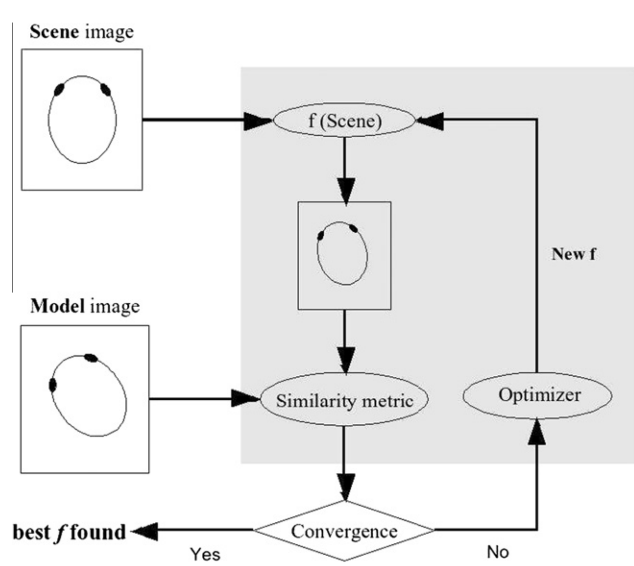

Fig. 1. The IR optimization process.

- On the one hand, we find the *matching-based* approach, where the optimization problem is intended to look for a set of correspondences of pairs of similar image features. Then, the registration transformation is derived from that set. This is the case of the well-known ICP method [6,7], whose main drawback is its sensitiveness to the initial transformation [8–11]. Thus, ICP usually gets stuck in local optima.
- On the other hand, we find the *parameter-based* IR approach which directly explores the values in the range of each transformation parameters.

A detailed description of the IR framework is out of the scope of this contribution. We refer the interested reader to [1–3]. Likewise, the formulation of the IR problem is dependent on the particular environment it is involved (remote sensing, medical imaging, CAD, etc.). Thus, in order to provide a more specific description of the problem, we focused our attention on the particular application we consider in our experiments: the IR of range images for 3D modeling [5,29–31].

Range scanners are able to capture 3D images, named range images, of the surface of the sensed object. Every range image is acquired from a particular viewpoint and it models the geometry of the scanned object partially. Thus, it is mandatory to consider a reconstruction technique to perform the accurate integration of the images in order to achieve a complete and reliable model of the physical object. This framework is usually called 3D modeling (see Fig. 2) and it is based on applying IR techniques to achieve the integration of the range images [5,29–31].

The 3D model reconstruction procedure involves several pair-wise alignments of two adjacent range images in order to obtain the final 3D model of the physical object. Therefore, every pair-wise IR method aims to find the Euclidean motion that brings the *scene* view (I_s) into the best possible alignment with the *model* view (I_m). It is usually considered an Euclidean motion based on a 3D rigid transformation (*f*) determined by seven real-coded parameters, that is: a rotation $\vec{R} = (\theta, Axis_x, Axis_y, Axis_z)$ and a translation $\vec{t} = (t_x, t_y, t_z)$, with θ and *Axis* being the angle and axis of rotation, respectively. Then, the transformed points of the *scene* view are denoted by

$$f(\vec{p}_i) = R(\vec{p}_i) + \vec{t}, \quad i = 1 \cdots N_{I_s} \tag{1}$$

Hence, the pair-wise IR task can be formulated as an optimization problem developed to search for the Euclidean transformation f^* achieving the best alignment of both images according to the considered *similarity metric F*:

$$f^* = \arg\min_f \; F(I_s, I_m; f) \;\; s.t.: \; f^*(I_s) \cong I_m \tag{2}$$

The median square error (MedSE) is usually considered the similarity metric in 3D modeling [28,30]:

$$F(I_s, I_m; f) = MedSE(d_i), \;\; \forall i \in \{1, \ldots, N_{I_s}\} \tag{3}$$

where $MedSE()$ corresponds to the median d_i value. We define $d_i = \|f(\vec{p}_i) - \vec{q}_{cl}\|^2$ as the squared Euclidean distance between the transformed scene point, $f(\vec{p}_i)$, and its corresponding closest point, $\vec{q}_{cl}$, in the *model* view I_m.

In order to speed up the computation of the closest point q_{cl} of I_m, indexing structures as kd-trees [32] or the grid closest point (GCP) transform proposed in [33] are often used. We will consider the GCP scheme in the experimental study (Section 4). In addition, we will follow a feature-based IR approach [3]. Such approaches consider a feature extraction procedure as a preprocessing step, previous to the application of the IR method. They are based on the selection of a small subset of truly representative characteristics of the images to be registered. In previous works [28,34–36], it has been demonstrated that using such IR approach offers a fast and a reliable IR result when range images are considered. In