



HAL
open science

Performing accurate joint kinematics from 3-d in vivo image sequences through consensus-driven simultaneous registration.

Jean-José Jacq, Thierry Cresson, Valérie Burdin, Christian Roux

► To cite this version:

Jean-José Jacq, Thierry Cresson, Valérie Burdin, Christian Roux. Performing accurate joint kinematics from 3-d in vivo image sequences through consensus-driven simultaneous registration.. IEEE Transactions on Biomedical Engineering, 2008, 55 (5), pp.1620-33. 10.1109/TBME.2008.918580 . inserm-00232732

HAL Id: inserm-00232732

<https://inserm.hal.science/inserm-00232732>

Submitted on 4 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Performing Accurate Joint Kinematics from 3D *in vivo* Image Sequences through Consensus-Driven Simultaneous Registration

J.J. Jacq, T. Cresson, V. Burdin, *Member, IEEE*, and C. Roux, *Fellow, IEEE*

Abstract—This paper addresses the problem of the robust registration of multiple observations of a same object. Such a problem typically arises whenever it becomes necessary to recover the trajectory of an evolving object observed through standard 3D medical imaging techniques. The instances of the tracked object are assumed to be variously truncated, locally subject to morphological evolutions throughout the sequence, and imprinted with significant segmentation errors as well as significant noise perturbations. The algorithm operates through the robust and simultaneous registration of all surface instances of a given object through median consensus. This operation consists of two interwoven processes set up to work in close collaboration. The first one progressively generates a median and implicit shape computed with respect to current estimations of the registration transformations, while the other refines these transformations with respect to the current estimation of their median shape. When compared with standard robust techniques, tests reveal significant improvements, both in robustness and precision. The algorithm is based on widely-used techniques, and proves highly effective while offering great flexibility of utilization.

Index Terms—Simultaneous registration, 4D medical imaging, joint kinematics.

I. INTRODUCTION

UNDERSTANDING the internal dynamics of complex joint systems – such as the tarsus or the carpus – remains a key challenge that aims at characterizing articular pathologies (e.g., arthritis) as well as designing prostheses. Working *in vivo* and non-invasively to study the precise function of such articulations has so far remained beyond the scope of the usual movement analysis techniques. Conversely, some 3D medical imaging techniques enable *in vivo* samplings of osteo-articular movements inside the most complex articulations. However, as the required sequences of 3D images turn out to appear with poor resolution, noise and time-varying truncations, these

techniques show many limitations when applied to complex articulations such as the tarsus or the carpus. Fig. 1 depicts typical data samples both for tarsus-MRI and carpus-CT.

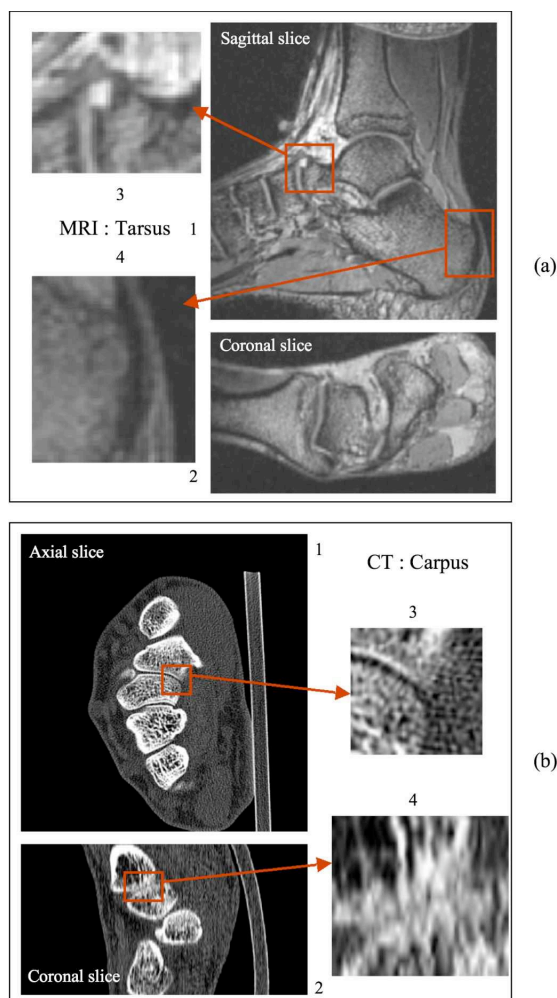


Fig. 1. Data examples obtained *in vivo*: MRI (0.5x0.5x1.5mm) of the tarsus [1] (a) and CT (0.3x0.3x1mm) of the carpus (b). Segmentation and tracking of the bone envelope involves four major difficulties. (i) The device modality may not be well adapted to bone delineation (a). Therefore, the result may depend on the operator's expertise. (ii) Due to poor voxel size, joints may appear welded (b4) and will require interactive delineation; partial volume effects may also increase. (iii) The field of view is confined. As a result, the visible part of some articular components can vary considerably as a function of the current articular configuration. The registration process will have to deal with this uneven clipping. (iv) A noticeable anisotropy may result in a dynamic evolution of biases and artifacts as a function of the current articular posture.

Manuscript received January 28, 2007, accepted October 10, 2007.

J. J. Jacq (corresponding author), V. Burdin, and C. Roux are both with GET-ENST Bretagne, Brest-cedex 3 CS 83818-29238, France and with INSERM, U650, LaTIM, Brest 29609, France; (e-mails: {JJ.Jacq, Valerie.Burdin, Christian.Roux}@enst-bretagne.fr).

T. Cresson is with ETS/LIO, Montréal H3C 1K3, Canada; (e-mail: Thierry.Cresson@etsmtl.ca).

Copyright © 2006 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purpose must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org

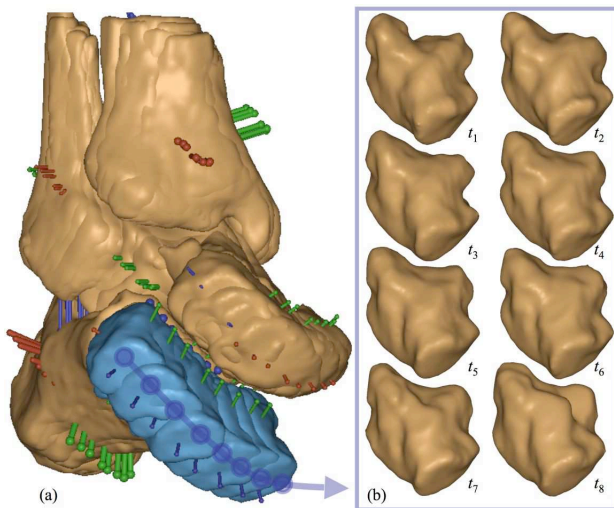


Fig. 2. Difficulties encountered while addressing the kinematics of tarsus bone components from MRI segmentation – positions 10° and 20° in pronation (t_1 , t_2), neutral position (t_3), positions 10° , 20° , 30° , 40° , 50° in supination (t_4, \dots, t_8). The raw data of one particular instance is shown Fig. 1.a. Six structures are tracked (tibia, fibula, talus, calcaneus, navicular, cuboid bone). This sequence sweep (a), whose box (b) focuses on a geometric reconstruction of the cuboid (blue shapes) from eight instances, shows that, due to time-varying segmentation errors, the mobile structure can no longer be considered as a perfectly rigid object (b). A usual tracking procedure aims at providing each instance of a bone with a kinematically equivalent trihedron – whose orthogonal axes are depicted in (a) with red, green and blue stems.

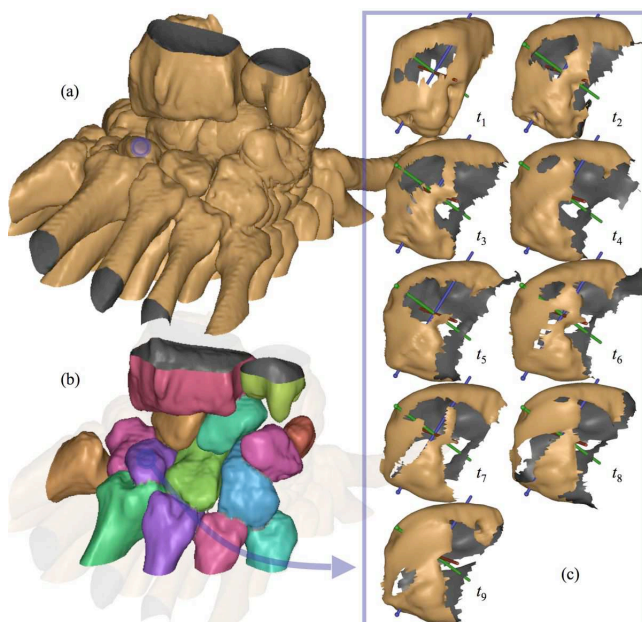


Fig. 3. Difficulties encountered while addressing the kinematics of carpus bone components from *in vivo* CT segmentations. Figure (a) depicts the superposition of nine instances of the left carpus – flexion positions 20° , 40° , 60° and Max (t_1, \dots, t_4), and extension positions 20° , 40° , 60° and Max (t_6, \dots, t_9). Fifteen structures are tracked (distal radius, distal ulna, scaphoid, semilunar, pyramidal, pisiform, hamate, capitate, trapezoid, trapezium, proximal M1–M5). A conventional protocol would require a tedious per-bone semi-interactive segmentation whereas a simple isosurface would extract outer bone surfaces instantaneously without unreliable extrapolations of subsampled congruent interfaces. As bones appear then welded, a semi-interactive dissection, through geodesic morphometry driven by torsion energy, then makes it possible to easily cut these shapes w.r.t. bone components [2]. Figure (b) depicts the resulting label map concerning the third instance. Similar results might have also been obtained through a more common use of a watershed transform on a 3D gradient image [3], [4] and [5]. Box (c) shows the corresponding observable parts of the nine trapezoid instances. Hidden parts of its intrinsic shape heavily depend on positions.

Moreover, as each bone component must be identified in order to perform its kinematics estimation, the first processing step has to involve a segmentation task. Owing to the fact that the available unsupervised rigid registration techniques addressing accurate kinematics objectives do not cope with segmentation errors, one usually has to undergo some tedious semi-interactive preliminary work requiring expert-level anatomical skills while still obtaining unavoidable errors because of incorrect interpretations of MRI data – see Fig. 2. Moreover, save for segmentation errors, one may still have problems with time-varying large truncations – see Fig. 3. Today, these are the main factors limiting the full development of kinematics non-invasive protocols based on 3D+T imaging.

Provided that (i) there are no segmentation errors nor truncations, (ii) the dimensions of the moving object are large compared to the data resolution, available kinematic-oriented procedures assume that the inertia trihedron of a rigid structure constitutes an equivalent coordinates system in relation to any movement [6], [7], [8]. 3D CT imaging of the wrist kinematics is addressed in [9], [10], [11] and [12] whereas 3D MRI imaging of the tarsus kinematics is addressed in [13], [14], [15], [1] and [16]. However, whenever an object description becomes unreliable (segmentation errors and truncations), the registration must be refined through more flexible techniques – *e.g.*, the ICP-based matching method [17]. In the presence of a bone structure whose truncation evolves with the position, [18] suggests an intrinsic clipping technique that, provided that the shape instances meet some geometric properties, enables an equivalence between the inertia trihedron estimations to be maintained throughout the sequence. However this approach does not deal with all geometries and only focusses on the shape parts common to each instance – thus, possibly discarding a large useful section of the available data. Recently, a voxel-based registration approach aiming at carpal bones kinematics through 3D CT sequences was proposed in [19]. However, even if this approach performs high resolution reconstructions of the common underlying bones, it still requires an accurate segmentation of the first instance. Thus, as available approaches still have problems with time-varying large truncations and segmentation errors, they cannot deal both accurately and robustly with segmentation results like those involved in the kinematic applications depicted in Fig. 2 and 3.

On account of the *information redundancy* involved by any kinematics-oriented objective, we propose a new point-based 4D rigid registration framework – whose main lines were first introduced in [20] – which is shown to be robust against significant shape variations due to (i) noise, (ii) large and time-varying missing parts, (iii) large segmentation errors. Moreover, while taking into account these drawbacks, our new technique is able to deliver both accurate kinematics estimations and accurate reconstruction of the bone components. We can then consider unreliable segmentations as an input and thus rely on automatic and efficient segmentation algorithms. Currently, our technique has to operate on explicit shape descriptions – *i.e.*, tessellations. This implies that this new methodology does not depend directly on

the image modality – typ. CT or MRI. The main novel idea behind our new algorithm is to look *robustly* at the 3D+T dataset as a whole 4D network with full connectivity w.r.t. time axis – *i.e.*, each instance shares a distinct undirected link with each of the other instances. In this framework, it becomes possible to robustly perform both Simultaneous Matching and Fusion (SMAF) of the relevant data within the same algorithmic process. This work can also be seen as a first experimental attempt to generalize the well-known mean shape notion [21] while reconstructing a specific type of root shape – that we call *Median Consensus Shape* (MCS) – through robust statistics. Thus, this general working scheme will, in the near future, also consider applications that go beyond the bounds of kinematics-oriented applications.

The paper is organized as follows. Section II introduces the technical background of our approach w.r.t. computer vision research addressing simultaneous registration. Section III describes our new algorithm. Section IV summarizes key aspects in the validation of the algorithm using both synthetic sequences and true data sequences. Section V discusses some methodological choices, and section VI concludes with some perspectives for future work.

II. SIMULTANEOUS REGISTRATION

A. Introduction

A simultaneous registration approach is expected to optimally merge redundant information so that we may accurately set up a relevant trihedron marker within each instance while dealing with segmentation problems such as those underlined in Fig. 2 and 3. This first objective appears upstream within the scope of a more general framework that addresses inter-relationships between articular surfaces and their kinematics. Within this field, the methodology outlined below focuses first on rigid movements. A second and complementary objective – not explicitly addressed here – is to produce an accurate shape description of the bone components. This will make it possible to study the geometry of the joint surfaces.

Bearing in mind these objectives, we naturally set out to measure bone kinematics using the movements of their external cortical interfaces. Thus, as a working hypothesis, rigid shapes resulting from the segmentation step are assumed to be available in a polyhedral form – *i.e.*, a list of facets linked to a cloud of vertices. From a typology aspect, registration techniques can be mainly divided into two categories: iconic (*i.e.*, voxel-based) and geometric methods. The particular case of registration of two surfaces is one of the main problems belonging to the second category. A bibliography of surface registration techniques used in the domain of medical imaging is available in [22]. As the registration procedure formulated below derives from the *Point Matching* approach the rest of this paper makes use of its terminology. In this framework, terms like object, cloud, shape and structure are equivalent, and correspond to different levels of abstraction in a same entity: the surface shape. More precisely, the term “cloud” only refers to the knowledge of the

vertex set. Matching a pair of features will imply the definition of a vector linking the source point to the target point – these corresponding points are also referred to as markers. By definition the source point is linked to the source object to which corrective positioning movements will be applied, while the target point is linked to the reference object.

In order to account for the dependencies of the main algorithmic building-blocks involved in the simultaneous registration objectives, Fig. 4 depicts the progressive nesting of the sub-problems that are to be solved. First of all, two types of problems have to be distinguished: *pose estimation* on one hand, and *matching* on the other. Pose estimation of two instances of an object assumes that exact point correspondences are known beforehand; this classical problem involves minimizing a constrained error norm applied to vectors defining the point correspondences. An important feature of this sub-problem is its degree of *robustness* w.r.t. to false correspondences. The second type of problem, matching, addresses both dynamic building of correspondence vectors and robust pose estimation. The up-do-date algorithm described in section II.B, hereafter designed by the acronym ICPr, addresses robust matching – *i.e.*, the left column of Fig. 4. Another dimension of the nesting of sub-problems is connected with the number of instances to be taken simultaneously into account. Section II.C presents a short survey of the algorithms handling simultaneous pose estimation or simultaneous matching – *i.e.*, the bottom row of Fig. 4. The robust and simultaneous matching of multiple objects – *i.e.*, the full domain of Fig. 4 – is still a relatively open problem and will be addressed in section III.

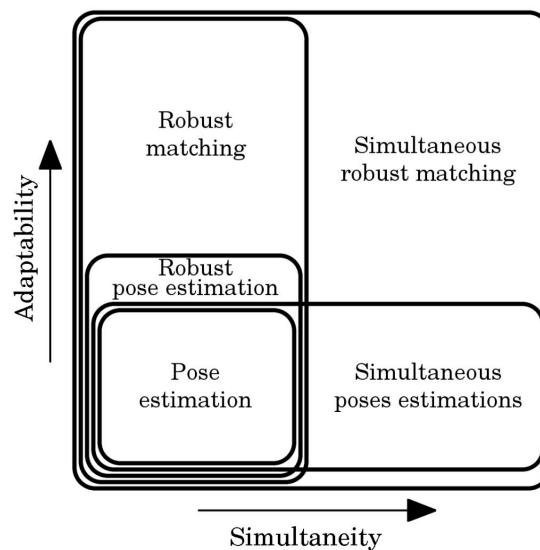


Fig. 4. Dependencies of the main sub-problems involved in robust simultaneous registrations.

B. Robust matching of two pointsets

This classical problem will remain at the core of almost simultaneous matching approaches. As stated above, it involves an iterative cascading of two sub-problems: pose estimation and matching. Pose estimation of two instances of an object assumes that point correspondences are known and do not depend on the current location of the object instances.

Let $Q = \{q_j, j = 1, \dots, L\}$ be a reference cloud of points and $P = \{p_i, i = 1, \dots, N\}$ a source cloud we want to move over the reference cloud through the rigid pose estimation transformation $p_i \rightarrow T(p_i) = R p_i + t$. While addressing the pose estimation sub-problem, the homologous target point $q_i \in Q$ is thus assumed to be known for each of the source points $p_i \in P$. Let w_i denote a predefined weight – defined on $[0, 1]$ – associated with the affinity level of the correspondence (p_i, q_i) . The optimal rigid transformation \hat{T} will first be assumed to be the one minimizing the generalized least squares $\hat{T} = \arg \min_T \sum_i w_i e_i^2$ where $e_i^2 = |q_i - T(p_i)|^2$. This first choice implies that the error distribution should be an isotropic Gaussian one. The standardized evaluation of Chi-2, defined as $\chi^2 = \sum_i w_i e_i^2 / \sum_i w_i$, will be used as the indicator of registration quality. This is a well-known sub-problem for which a comparative test of the main closed-form expressions of \hat{T} is summarized in [23]–[24]. One of the most appropriate techniques makes use of singular values decomposition (SVD) and was proposed in [25]. An important feature of this sub-problem is its degree of robustness relating to false correspondences – *i.e.*, outliers. Thus, searching for the optimal transform \hat{T} requires resorting to a robust norm $\sum_i \rho(e_i)$ instead of the standard quadratic norm $\sum_i e_i^2$. The function ρ usually refers to an M-estimator [26]–[27]. The main objective is to lessen the influence of correspondence errors whose distribution does not fit a Gaussian model. Let the derivative $\rho'(e)$ denote the influence function on the M-estimator. The norm permissiveness implies that $\rho'(e)$ is bounded when $|e| \rightarrow \infty$. Many M-estimators have a redescending influence function. This characteristic may be parameterized by writing that $\rho'(e) \rightarrow 0$ when $|e| \rightarrow e_r$, where e_r is the rejection point. As, the minimization of $\sum_i \rho(e_i)$ usually involves an Iterative Reweighted Least Squares (IRLS) method [27], [28], an M-estimator is sometime named W-estimator. Let $w(e_i) = \rho'(e_i) / e_i$ denote an auxiliary weighting function defined on $[0, 1]$ – which can be seen as the Gaussian likelihood of a residual e_i . Each IRLS step can then proceed through the update of the weights followed by a weighted quadratic optimization of the transform parameters. As the second iteration step may then refer to a closed-form solution, this results in a very efficient algorithm. While addressing the robust pose estimation of two point sets, this type of iterative, robust, efficient approach was first proposed in [29].

Many algorithms make use of the Tukey-Biweight M-estimator. It is expressed as $\rho(e) = (1 - e^2/e_r^2 + e^4/3e_r^4) e^2/2e_r^2$ if $|e| \leq e_r$ and $\rho(e) = 1/6$ elsewhere. The weight function involved in its IRLS formulation is $w(e) = (1 - (e/e_r)^2)^2$ if $|e| \leq e_r$ and $w(e) = 0$ otherwise. Let $\hat{\sigma}$ denote a robust estimation of the standard deviation of the residuals. In order to introduce an adaptive rejection point, it is usual to write $e_r = \lambda \hat{\sigma}$. This deviation estimation can be robustly updated through an L-estimator – *e.g.*, $\hat{\sigma} = 1.5 e_m$, where e_m denotes the median of error modulus over the current fitting. Therefore, we will also refer to this algorithm as an LW-estimator. We make use of this special acronym to underline that, unlike the standard setting of an IRLS optimization of a

robust norm, $\hat{\sigma}$ will not remain static. Indeed, its L-estimate will be updated within each IRLS iteration. On completion of each iteration, the convergence criterion tracks the evolution of the global residual χ , defined as the square root of Chi-2. Let introduce the minimal relative gain μ that enables us to define the stopping criterion $|\Delta \chi| < \mu \chi$. An upper bound iteration count I_{\max} can optionally be added to this criterion. Typical values for λ and μ are 3 and 0.1% respectively.

In fact, practical working contexts do not come down to pose estimation tasks because point correspondences remain *a priori* unknown. Thus, the registration algorithm must be able to dynamically estimate the *matching* of two clouds by identifying homologous points prior to the rigid pose estimation step. The standard matching algorithm, the Iterative Closest Point (ICP) algorithm, was proposed independently by Besl and McKay [17] and Chen and Medioni [30]. In such an algorithm, the pose estimation step is encapsulated as the second step of a two-step iterative process. The first step updates point correspondences w.r.t. the current estimation of T – an initial guess T_0 is thus required. The second step computes an updated estimation of T through a pose estimation algorithm and then moves the source cloud accordingly. These ICP algorithms mainly differ w.r.t. the metric used – a point-to-point metric or a point-to-surface metric – that is to say, the way they use or not the normal to the surface. In [17], the normals are ignored, and the minimized metric refers to the Euclidean distance between a source point and its target point. In [30], which is therefore restricted to surfaces, the metric refers to the distance between a source point and the tangent plane defined by the normal at the target point. Due to its first step, it is important to stress that an ICP algorithm performs a local search and will become unreliable if the clouds are not initially roughly registered. Alternative global search techniques are proposed in [31]. In order to reduce the search complexity (step 1) and accelerate the convergence, many ICP derivatives (*e.g.*, [32]) perform a search of matching vectors with additional heuristics constraints, such as shape invariants. However, following [33], the most interesting evolutions are those taking into account a robust norm in step 2 in the ICP algorithm.

C. Available strategies aiming at simultaneous registration

Algorithms performing simultaneous matching can be mainly divided into two categories: *incremental* algorithms and *global* algorithms. The former can be seen as greedy algorithms carrying out a sequence of pairwise registrations in order to minimize error accumulation over the network linking the overlapping instances. Their main objectives are both to cope with a large dataset and to allow for the dynamic upgrade of the network. Our application context is most concerned with the second category. In the point matching operating case, while addressing K instances, each consisting of $N_k, k = 1, \dots, K$ points, a global algorithm performs an optimization whose generic minimizer is

$$\min_{T_j, j=1, \dots, K} \sum_{k=1}^K \sum_{j \neq k, j=1}^K \sum_{i=1}^{N_k} \rho \left(|C_j(T_k(p_{k,i}); T_j) - T_k(p_{k,i})| \right),$$

where $T_l, l=1, \dots, K$ denotes the set of rigid transformations, $\rho(e)$ is a robust metric applied to the matching error e , \mathcal{C} is a point matching operator returning a prediction of the target point $q_{k,i}^j = \mathcal{C}_j(T_k(p_{k,i}); T_j)$ of the source point $T_k(p_{k,i})$ w.r.t. instance j relocated by T_j – i.e., $q_{k,i}^j$ is the closest plausible location within the framework of the distance metric under consideration. Below, $\mathcal{C}_j(p_{k,i})$ will denote $q_{k,i}^j$ whenever $T_k = T_j = I$. It should be noticed that \mathcal{C} is an asymmetric operator because of noise and outliers. Since such a problem is highly non linear, it cannot be solved analytically and it has recourse to an iterative minimization. Thus, having an initial guess $T_l^0, l=1, \dots, K$ beforehand becomes a major requirement. The minimizer does not depend on the choice of the common coordinates system. Therefore, the problem is usually regularized by assuming that the relative orientation of one of the instances – e.g., the first one – is kept aligned with the common coordinates system.

With respect to published work addressing range image registration, the robustness treatment seems not to be the primary critical aspect. Indeed, most studies discard outlying correspondences through thresholding of the length of the matching vectors (and, when available, thresholding of the divergence angle between the source and target normals [34]) and then come down to a standard quadratic optimization, setting $\rho(e) \equiv e^2$ and thus assuming a zero mean Gaussian noise. A noticeable exception can be found in [35] where matching errors are integrated through an M-estimator and the transformations optimized through gradient descent and quaternion-based parameterization.

The main component in the iterative process is still an ICP-like loop. Thus, it alternates the update of the target point's locations and the optimization of the transformation set, so as to move the source points towards their respective target points. The local optimization sub-problem then comes down to performing simultaneous pose estimations. For this specific purpose, four types of iterative process have been proposed.

- Each of the source points is matched to the union of other instance points (i.e., getting $K-1$ independent target points per source point) [36]. This strategy is much related to the formulation of the generic minimizer and leads to iteratively solving a nonlinear least-squares problem. Neugebauer [36] operates through the point-to-plane metric whereas [35] makes use of the point-to-point metric. However, because it independently and successively matches each instance to the union of the $K-1$ other ones, [35] proposes a less optimal strategy.
- The set of matching pairs resulting from all instances are simultaneously taken into account. The transformations are optimized through a linear algebra generalization of the well-known closed-form solutions of the rigid pairwise pose estimation – their main characteristic being to enable rotation and translation to be decoupled. A generalization of Horn's quaternion-based approach is proposed in [37] whereas [38] performs a generalization of Arun's SVD-based approach. Since the latter leads to a weighted iterative process, [38] optionally proposes a seamless integration of an M-estimator managed through its usual

IRLS minimization.

- Likewise, the correspondences are again simultaneously taken into account but the optimization step is carried out through a mechanical-based analogy which simulates energy minimization over a network of spring-connected instances [39, 40].
- Each source point is matched to a virtual target point coming from the average instance [21]. The optimization step can still make use of the usual pairwise methods but needs to manage an auxiliary step in order to update the average instance. Guehring [34] confines the computation to tie points and operates through an anisotropic description of their matching noise in order to improve the convergence rate. Masuda [41] simultaneously performs registrations and explicit reconstruction of the mean shape by operating through the signed distance field of each instance as well as that of their mean instance.

Choosing an optimal strategy first requires a trade-off to be made between robustness and convergence rate. In fact, [42] compares the performance of three simultaneous pose estimation algorithms ([21], [37], [40]) and concludes that the best (resp. lower) convergence rate is performed by [37] (resp. [21]) whereas [21] seems to be the most accurate among these three methods.

III. METHODOLOGY

A. Overview

First, before returning to the real working case, let \mathcal{C} denote a hypothetical and ideally perfect operator able to successively match each point of the cloud k with its unique homologous point in any cloud k' . In this context, the pairwise rigid matching of two point clouds reverts to the classic rigid-pose-estimation problem. As stated in section II.C, Pennec [21] showed that the corresponding simultaneous rigid-pose-estimation problem can be fairly resolved through an iterative application of pairwise-like steps while introducing a simple point merging operator \mathcal{M} returning, for each source point, a unique virtual homologous point built as the *mean* of its $K-1$ homologous points and the source point. After convergence, if the noise associated with the point measurements is Gaussian, the set of virtual homologous points gives rise to an implicit *mean shape*. Unfortunately, in a real working case the available operator \mathcal{C} always becomes merely approximate. However, in the same way as one perform practical pairwise point matching, it is straightforward to insert the Pennec simultaneous pose estimation approach in an ICP-like iterative process. A recent application addressing such a mean shape computation can be found in [43].

Our present work, whose first results were published in [20], is an attempt to address medical simultaneous registration problems through some robust generalization of the Pennec mean shape approach. As we now have to cope with both time-varying large contaminations (accounting for pathologies and/or segmentation errors) and large truncations – not only Gaussian noise – this makes a world of difference. As discussed below, this context limits the achievable point

matching operator \mathcal{C} to its simplest form – *i.e.*, finding the closest point. Thus, as the returned matched points may become thoroughly unreliable, our main problem comes down to finding a robust binding of the merging operator \mathcal{M} in an optimal way w.r.t. available information. While making a syntactic parallel with the prominent work of John Tukey, who had long been dealing with *root signal* extraction through robust statistics [44], this task can be seen as performing *root shape* extraction through a dedicated operator \mathcal{M} . Since the optimization process has to rely on a robust norm managed through an M-estimator, it gives rise to an IRLS optimization process involving a per-point auxiliary weighting scheme. As a core idea, we propose an implementation of the \mathcal{M} -operator making intensive use of the valuable extra-information managed by this weighting scheme in order to make up for the numerous failures of the \mathcal{C} -operator. This is expected to introduce a *transversal* information flow and, thus, to enforce the global performance of the optimization process. Instead of juxtaposing K M-estimators, this attempts to manage a single K -dimensional M-estimator in an experimental way. As a consequence of our purpose-oriented binding of the \mathcal{M} -operator (discussed in section III.D), the resulting root shape (resp. the new algorithm) is then called the Median Consensus Shape (resp. the *Iterative Median Closest Point* (IMCP) algorithm).

The skeleton of the new algorithm is described in section III.B. The operator \mathcal{C} (resp. \mathcal{M}) is then discussed in section III.C (resp. III.D). This section will end with the release of the full IMCP pseudocode.

B. Skeleton of the new algorithm

Initial instance registrations are assumed to be performed through alignment of their inertia trihedrons. Let $\rho_k, k = 1, \dots, K$ denote a set of redescending M-estimators whose indexation expresses a per-instance management of the robustness through adaptive estimation of their rejection point $\lambda\sigma_k$. According to the technical aspects discussed in section II, the *iterative* process aiming at the simultaneous matching of the K instances can be summarized through the following five main steps.

1. Update the points of each instance with their respective initial transform T_1^0, \dots, T_K^0 . Set the robust norms to the quadratic profile (*i.e.*, setting $\sigma_k \rightarrow \infty, k = 1, \dots, K$).
2. Rebuild the matching point set $\{q_{k,i}^j = \mathcal{C}_j(p_{k,i}); k = 1, \dots, K; j = 1, \dots, K, j \neq k; i = 1, \dots, N_k\}$.
3. Using merging operator \mathcal{M} , build the point set accounting for the current estimation of the implicit root instance as $\{\tilde{q}_{k,i} = \mathcal{M}_{\rho_1, \dots, \rho_K}(p_{k,i}, q_{k,i}^j, j = 1, \dots, K, j \neq k); k = 1, \dots, K; i = 1, \dots, N_k\}$.
4. Update the rejection points – *i.e.*, $\sigma_{1, \dots, K}$. Then, obtain the correction transforms $T_1^{(c)}, \dots, T_K^{(c)}$ as $\{\arg \min_{T_k^{(c)} \in \mathcal{T}_k} \sum_{i=1, \dots, N_k} \rho_k(|\tilde{q}_{k,i} - T_k^{(c)}(p_{k,i})|)\}$.
5. Update each point's location with its respective correction transform. Jump to step 2 while all of the instances have not converged.

The main part of step 4 simply comes down to a sequence of K pairwise robust pose estimations. These estimations are carried out through an IRLS optimization and, thus, each of them makes use of an auxiliary weighting map – rejection points σ_k being robustly computed through an L-estimation over $\{|\tilde{q}_{k,i} - p_{k,i}|, i = 1, \dots, N_k\}$. Since the current internal state of a robust norm ρ_k w.r.t. data is thus made explicit through $\{w_{k,i}, i = 1, \dots, N_k\}$, using an IRLS process becomes a key advantage of the algorithm. Indeed, these normalized weights can now express how well a point is likely to account for a zero mean Gaussian perturbation of the current root shape estimation. Therefore, interaction levels between global and local iterations become reinforced if the M-estimators become involved in a hidden transversal link between algorithm steps 3 and 4.

Fig. 5 provides the skeleton of the robust simultaneous registration algorithm. It involves a major evolution w.r.t. the previous conceptual five-step algorithm. This evolution assigns an iteration-level precedence to the step performing the matching update – which is thus located *within* the inner loop at line 7. This assumes that trying to obtain an optimal correction transform w.r.t. poor matching entries would not only waste computation time but would also get the whole process trapped into irrelevant minima. This strategy balances the crudeness of the matching operator and becomes a major requirement for dealing successfully with corrupted datasets.

1.	Until <i>global</i> convergence:
2.	Increment global iteration count and set transformations $T_{1, \dots, K}$ to I
3.	For current source cloud $k = 1$ to K :
4.	Reset local iteration count and copy current value of error χ_k in χ_k^0 Copy the current weights set of the cloud k in their corresponding cloud caches.
5.	Until <i>local</i> convergence:
6.	Increment local iteration count.
7.	Using kD-trees and operators (\mathcal{C}, \mathcal{M}), build the mapping set of the point's source k w.r.t. other clouds.
8.	Compute robust error statistics from mapping vectors lengths and get new estimation for σ_k .
9.	Update source point weights (step 1 of the robust norm optimization) and get new estimation for χ_k .
10.	Keeping weights constant, find the rigid transform \mathcal{T} of the current source cloud which minimizes its quadratic error norm (step 2 of the robust norm optimization – <i>i.e.</i> , weighted SVD).
11.	Update location of the cloud k with \mathcal{T} and replace T_k with $\mathcal{T} T_k$.
12.	Local convergence if gain $ \Delta\chi_k / \chi_k < \mu$ (or local count $> It_{\max}$).
13.	Update global gain of the cloud k with value $ \chi_k^0 - \chi_k / \chi_k$. Save the new weights set of cloud k in their own cloud caches and restore the previous ones. Apply $\text{inv}(T_k)$ to cloud k in order to restore its previous location.
14.	Apply $\text{inv}(T_1)$ to $T_{1, \dots, K}$ (optional step), then apply transformations $T_{1, \dots, K}$ to their respective cloud. For each of the clouds, restore from cache it new weights set and put it in its active place.
15.	Global convergence if global gain $< \mu$ for all clouds (or global iteration count $> Itg_{\max}$).

Fig. 5. Skeleton of the IMCP algorithm operating on K instances.

C. Matching operator

The application of the matching operator \mathcal{C} to the current source point $p_{k,i}$ has, through the point-to-point metric, to return the set $\mathcal{N}_p = \{(q_{l,j_0}, w_{l,j_0}), l \in K \setminus \{k\}\}$ of target points grouping the closest $K-1$ neighbors of $p_{k,i}$, where j_0 denotes

the index of the closest neighbor of $p_{k,i}$ assigned in cloud l , and where w_{l,j_0} denotes the current weight assigned to location q_{l,j_0} by the LW-estimator. As a way to avoid combinatorial explosion, this operator requires the management of K KD-trees. However, as the current applications of the simultaneous registration algorithm address rigid transformations, these tree-like structures can remain static – *i.e.*, they are built once and for all at the initialization step and are then accessed through their own coordinate frames.

While processing in point matching mode, not taking into account an available polyhedral description is equivalent to working with additional matching noise. In an alternative polyhedral mode, target points are thus located on the nearest surface location – the matching weight associated with the actual target point is computed as the barycentric interpolation of the three vertex weights of the triangular facet on which the target point is located. At the expense of a slight additional cost, the polyhedral mode enables us to better discriminate the contamination noise from the Gaussian noise. The examples discussed in the application section are performed through this polyhedral mode. Such a polyhedral mode should become mandatory whenever the vertex density is poor or very different from one mesh to the other. Optionally, although not used hereafter, this mode enables target points located on surface boundaries to be filtered out.

D. Merging operator

As introduced in section III.A, the median consensus approach is based on the ability to compute target points through the auxiliary weighting schemes managed by the LW-estimator. Let us first recall what, within the IMCP framework, the sub-problem to be resolved at the \mathcal{M} -operator level is. For a given source point $p_{k,i}$ of rank i in a source cloud k , on the basis of the set \mathcal{N}_p returned by the matching operator, it simply requires being able to predict the localization of the virtual target point $\tilde{q}_{k,i}$ that should be associated with it.

Let point $r_{k,i}$ denote some robust centroid, discussed below, computed over the set \mathcal{N}_p . A first algebraic rule has to be defined in order to guarantee that the IMCP algorithm provides a mean shape whenever the data set becomes ideally Gaussian – *i.e.*, free of contaminations and perfectly matched. The virtual target point is defined as $\tilde{q}_{k,i} = K^{-1}p_{k,i} + (K-1)K^{-1}r_{k,i}$. Thus, $\tilde{q}_{k,i} \rightarrow r_{k,i}$ as the number of clouds increases. Conversely, when $K=2$, this formulation clearly shows that this usual working case is also handled as a *symmetric* problem, where each of the two objects is taken in turn as the source object and matched to its intermediate mean shape if there are no outliers.

The computation of the robust centroid $r_{k,i}$ over \mathcal{N}_p first requires the definition of a relevant weighting scheme. In our robust estimation framework, the normalized target weight w_{l,j_0} returned within \mathcal{N}_p expresses how well point j_0 of cloud l is likely to account for a Gaussian perturbation of the current evaluation of the implicit root shape. Thus, it does not convey any information about the relevancy of its putative

matching with the source point. In fact, a target point may be far away from its ideal location w.r.t. $p_{k,i}$, while still remaining pertinent against the root shape. Meanwhile, these target weights still carry out a rank of merit between elements of \mathcal{N}_p .

Hopefully, the global statistic measurements managed by the IMCP provide, on a per-cloud basis, a useful robust estimation of the standard deviation of the Gaussian noise over the set of inlier points. Thus, an auxiliary normalized weighting $w_{l,j_0}^{k,i}$ can be introduced to enable *a priori* scoring of the matching of the source $p_{k,i}$ with the target q_{l,j_0} . Let us recall that, on completion of each local iteration, the LW-estimator updates the estimation of the noise deviation χ_k associated with cloud k . In order to express the probability of a source point being located in the *noise envelope* associated with the source cloud k , this auxiliary weight is defined as $w_{l,j_0}^{k,i} = (1 - (|p_{k,i} - q_{l,j_0}| / \lambda' \chi_k)^2)^2$ if $|p_{k,i} - q_{l,j_0}| \leq \lambda' \chi_k$ and $w_{l,j_0}^{k,i} = 0$ otherwise. This expression may become irrelevant whenever the source point belongs to a contaminated location of cloud k . To robustify this new weighting, we start from the assumption that $w_{l,j_0}^{k,i}$ will remain combined with the weighting attached to the target point, which filters out target points without intrinsic significance. Thus, the expression of the weighted centroid is $r_{k,i} = \sum_l w_{l,j_0}^{k,i} w_{l,j_0} q_{l,j_0} / \sum_l w_{l,j_0}^{k,i} w_{l,j_0}$. This weighting scheme is *adaptive*, since the blurring level it introduces around the source cloud decreases at the same time as new iterations refine the point's membership w.r.t. the median cloud. Let \mathcal{N}_p^* denote the subset of \mathcal{N}_p such that $w_{l,j_0} > 0$. Optionally, in order to take into account non-stationary Gaussian noise, we can apply a third corrective coefficient χ_{\min} / χ_l to w_{l,j_0} , where χ_{\min} is the smaller M-Estimate of the noise standard deviation of the clouds involved in \mathcal{N}_p^* .

Dealing with time-correlated contaminations (*e.g.*, pathologies or non-rigid parts) is one of the main objectives of our algorithm. These contaminations are much less likely to undergo random evolution throughout the sequence than segmentation-based contaminations. Some of these large contaminations may appear stable in a noticeable percentage of the instances and may then give rise to secondary attracting pools able to bias the global optimization. Therefore, in order to prevent the emergence of these pools, an additional *quorum rule* must be applied. This rule adaptively dilates the noise envelope so as to cover at least 50% of the nearest target points indexed in \mathcal{N}_p . Thus, whatever the source point $p_{k,i}$, among its $K-1$ target points q_{l,j_0} , at least $(K-1)/2$ must express a membership $w_{l,j_0}^{k,i} \leq 0.5$ versus $p_{k,i}$. This leads to a simple modification of the previous centroid relationship where λ' is replaced by the adaptive expression $\max(\lambda', (\sqrt{2} / (\sqrt{2} - 1))^{1/2} \chi_k^{-1} \text{med}_l \{|p_{k,i} - q_{l,j_0}|\})$. This quorum rule, in conformity with a democratic acceptance of the notion of consensus, ensures that the location of the virtual target point $r_{k,i}$ systematically takes into account the influence of at least 50% of the $K-1$ target points. Thus, the actual merging operator is said to perform *median consensus* filtering while computing the virtual target points.


```

1.  $\mathcal{X}_{\text{new } k} \leftarrow \mathcal{X}_{\text{old } k} \leftarrow \infty, \hat{T}_k \leftarrow I, k=1, \dots, K; \hat{w}_{k,i} \leftarrow 1, k=1, \dots, K, i=1, \dots, N_k; Itg \leftarrow 0;$ 
2. Repeat
3.    $\mathcal{X}_{\text{new } k} \leftrightarrow \mathcal{X}_{\text{old } k}, k=1, \dots, K; Itg \leftarrow Itg + 1;$ 
4.   For  $k=1, \dots, K$  do
5.      $\mathcal{X}'_{\text{new } k} \leftarrow \mathcal{X}'_{\text{old } k} \leftarrow \mathcal{X}_{\text{old } k}; \hat{W}_k^0 \leftarrow \hat{W}_k; P_k^0 \leftarrow P_k; It \leftarrow 0;$ 
6.     Repeat
7.        $Q \leftarrow \emptyset; \mathcal{X}'_{\text{new } k} \leftrightarrow \mathcal{X}'_{\text{old } k}; It \leftarrow It + 1;$ 
8.       For  $i=1, \dots, N_k$  do
9.          $\mathcal{N} \leftarrow \emptyset;$ 
10.        For  $l=1, \dots, K$  if  $l \neq k$  do  $j_0 = \arg \min_{j=1, \dots, N_l} |p_{k,i}^0 - p_{l,j}|, \mathcal{N} \leftarrow \mathcal{N} \cup \{(p_{l,j_0}, \hat{w}_{l,j_0}, \mathcal{X}_{\text{old } l})\};$ 
11.         $\lambda'' \leftarrow \max(\lambda', (\sqrt{2}/(\sqrt{2}-1))^{1/2} \mathcal{X}'_{\text{old } k} \text{med}_{m \in \mathcal{N}} \{|p_{k,i}^0 - p_m|\});$ 
12.        For  $m \in \mathcal{N}$  do  $w'_m \leftarrow (1 - (|p_{k,i}^0 - p_m| / \lambda'' \mathcal{X}'_{\text{old } k})^2) \text{if } |p_{k,i}^0 - q_m| \leq \lambda'' \mathcal{X}'_{\text{old } k}, w'_m \leftarrow 0$  else;
13.         $\mathcal{N}^* \leftarrow \{(p_m, w_m, \mathcal{X}_m) \in \mathcal{N} / w_m > 0\}; \mathcal{X}_{\min} \leftarrow \min_{m \in \mathcal{N}^*} \{\mathcal{X}_m\};$  for  $m \in \mathcal{N}$  do  $w''_m \leftarrow \mathcal{X}_{\min} / \mathcal{X}_m;$ 
14.         $\bar{w} \leftarrow \sum_{m \in \mathcal{N}^*} w_m w''_m; r_{k,i} \leftarrow \sum_{m \in \mathcal{N}^*} w_m w''_m p_m / \bar{w}$  if  $\bar{w} > 0, r_{k,i} \leftarrow \text{med}_{m \in \mathcal{N}^*} \{p_m\}$  else;
15.         $\tilde{q}_{k,i} \leftarrow K^{-1} p_{k,i}^0 + (K-1) K^{-1} r_{k,i}; Q \leftarrow Q \cup \{(p_{k,i}^0, \tilde{q}_{k,i})\};$ 
16.        End do
17.        For  $m \in Q$  do  $\hat{e}_m \leftarrow |q_m - p_m|; \hat{\sigma}_k \leftarrow 1.5 \text{med}_{m \in Q} \{\hat{e}_m\};$  For  $m \in Q$  do  $\hat{w}_{k,m}^0 \leftarrow (1 - (\hat{e}_m / \lambda \hat{\sigma}_k)^2) \text{if } \hat{e}_m \leq \lambda \hat{\sigma}_k, \hat{w}_{k,m}^0 \leftarrow 0$  else;
18.         $\mathcal{X}'_{\text{new } k} \leftarrow (\sum_{m \in Q} \hat{w}_{k,m}^0 \hat{e}_m^2 / \sum_{m \in Q} \hat{w}_{k,m}^0)^{1/2}; \bar{p} \leftarrow \sum_{m \in Q} \hat{w}_{k,m}^0 p_m / \sum_{m \in Q} \hat{w}_{k,m}^0; \bar{q} \leftarrow \sum_{m \in Q} \hat{w}_{k,m}^0 q_m / \sum_{m \in Q} \hat{w}_{k,m}^0;$ 
19.        For  $m \in Q$  do  $p_m^c \leftarrow p_m - \bar{p}, q_m^c \leftarrow q_m - \bar{q}; H \leftarrow \{h_{n,l} = \sum_{m \in Q} \hat{w}_{k,m}^0 p_m^c q_m^c, n=1, \dots, 3; l=1, \dots, 3\}; USV^T \leftarrow \text{SVD}(H);$ 
20.         $\hat{R} \leftarrow V U^T; \hat{i} \leftarrow \bar{q} - \hat{R} \bar{p}; \hat{T} \leftarrow \{\hat{R}, \hat{i}\}; P_k^0 \leftarrow \hat{T} P_k^0; \hat{T}_k \leftarrow \hat{T} \hat{T}_k;$ 
21.        While  $|\mathcal{X}'_{\text{new } k} - \mathcal{X}'_{\text{old } k}| \geq \mu \mathcal{X}'_{\text{old } k}$  and  $It < It_{\max}$ 
22.        End do
23.        For  $k=1, \dots, K$  do  $\hat{W}_k \leftarrow \hat{W}_k^0, P_k \leftarrow P_k^0, \mathcal{X}_{\text{new } k} \leftarrow \mathcal{X}'_{\text{new } k};$ 
24. While  $\exists k / |\mathcal{X}_{\text{new } k} - \mathcal{X}_{\text{old } k}| \geq \mu \mathcal{X}_{\text{old } k}$  and  $Itg < Itg_{\max}$ 

```

Fig. 6. Pseudocode of the IMCP algorithm. Parameters: (i) outlier error scales λ, λ' (typ. $\lambda = \lambda' = 3$), (ii) minimal error gain μ of global or local iterations (typ. 0.1%), (iii) local and global iteration bounds It_{\max}, Itg_{\max} . Inputs: set of K clouds $\{P_k, k=1, \dots, K\}$, with points $P_k \equiv \{p_{k,i}, i=1, \dots, N_k\}$. Outputs: (i) membership maps $\{\hat{W}_k, k=1, \dots, K\}$ with $\hat{W}_k \equiv \{\hat{w}_{k,i}, i=1, \dots, N_k\}$ and (ii) rigid pose estimations $\{\hat{T}_k, k=1, \dots, K\}$ of the K initial cloud locations, both w.r.t. to the location of their common virtual MCS.

Fig. 6 shows the whole pseudocode of the IMCP algorithm operating in point mode. The specification of the full content of the core loop (including lines 5 to 21) is expected to be multithread-safe.

IV. VALIDATION OF THE IMCP ALGORITHM

A. Parameters overview

The IMCP algorithm only brings five parameters into play: the relative gain μ tuning the local and global convergence, (typically 0.1%), maximum bounds for the number of local iterations It_{\max} and global iterations Itg_{\max} (typ. ∞ , both), the scale factor λ associated with the robust error norm, and its duplicate λ' in the rule managing the emergence of the local consensus. These parameters can be classified in order of growing importance. The relative gain μ should be about 0.1%. The iteration bounds, especially useful in case of meaningless initial registrations, should be set to act as algorithmic fuses. However, it should be noted that setting It_{\max} to 1 leaves both performance and overall computational cost mostly unchanged. This enforces redundancy coalescence but, in turn, the asymptotic part in the convergence process is then solved less efficiently. The auxiliary scale factor λ' should take a value close to λ and, by default, we state $\lambda' = \lambda$. The only parameter choice that could prove critical is λ . But tests show that default values such as $\lambda = 3$ in point

mode, and $\lambda = 5$ in polyhedral mode, give satisfactory results. Finally, due to its adaptive features, the default parameters of the IMCP algorithm will seldom require any modifications.

B. Validation methodology

The IMCP algorithm is assumed to operate within an analysis framework that also enables a rough pre-alignment to be performed. As regards bone structures, this initial solution is usually provided through alignment of the principal axes of the point clouds. Therefore, in this context, the purpose of the IMCP is to correct the predictive error introduced by the initial registrations. For validation purposes, rather than applying the IMCP correction transforms to their respective instances, these are kept in their native location. Indeed, the inverse correction transforms are applied to their respective pre-alignment coordinates system – *i.e.*, their trihedron. In this way, if the accumulations of the correction transforms become optimal, the trihedrons become superimposed. Thus, rather than simulating artificial movements while testing various arbitrary pre-defined transforms, this validation approach makes the visual validation of the result easier, and allows us to focus on the most difficult initial configurations – *i.e.*, correcting the overall misalignment biases resulting from by the deficiencies embedded in the various instances. The tests discussed below refer to the sequence depicted in Fig. 7. This is the most difficult artificial one that was tested so far. The trihedrons,

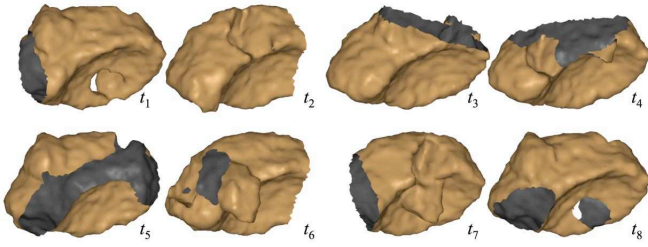


Fig. 7. This is a complex artificial sequence designed so as to embed the major types of difficulty we expect to experience in real sequences. This sequence involves eight instances presenting uncorrelated truncations, Gaussian noise, and clippings. The standard deviation of the noise is kept constant throughout the sequence. As these shapes are assumed to come from some 3D segmentation procedure, their high frequency components are smoothed and spread. These instances also involve both large and coherent contaminations, and some are intentionally correlated, that are expected to account for errors coming from some unreliable automatic segmentation procedures as well as some time evolving pathologies.

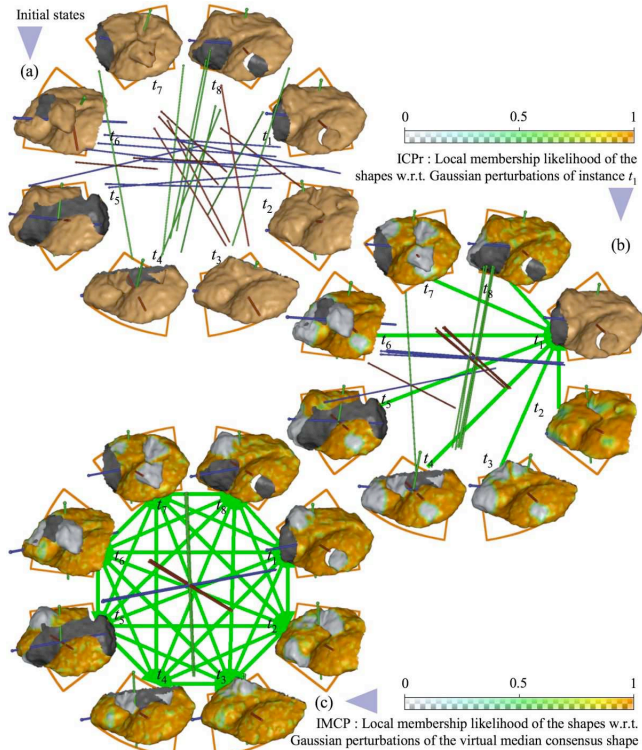


Fig. 8. Improvements obtained through median consensus registration. Ring (a) depicts the artificial sequence introduced in Fig. 7. Initial registrations are provided from the inertia trihedrons of the shells. In order to enable an accurate assessment of the registration results, all of the instances are built at the same location as their common native model. Therefore, as soon as these trihedrons become kinematically equivalent, they would appear exactly superposed. Hence, the transform estimates are not applied to shapes. On the contrary, their inverse is applied to their corresponding trihedron. Therefore, the mutual scattering of these trihedrons – as drawn in the middle of the rings – should account for the matching noise throughout the sequence. Ring (b) depicts the pair-wise ICPr alignments w.r.t. instance t_1 . Apart from one major error, all of the alignments remain approximate. In contrast with ICPr results, the median consensus through IMCP alignments (ring (c)) does not convey perceptible errors. See also Fig. 10 and 11 for a comparative analysis of the scattering level of these registrations.

superposed in their initial locations (Fig. 8.a), give an overview of the pre-alignment errors to correct.

So as not to unduly grant the IMCP with the advantages coming from the robust norm, section IV.D numerically compares its results with those of the robust ICPr algorithm

described in section II.B. The parameters (λ, μ) remain common to the two algorithms – as well as supporting the polyhedral mode. However, before making numerical comparisons, one can already perform some meaningful visual comparisons. Fig. 8.b shows the pairwise ICPr correction ($\lambda = 5$) versus one instance arbitrarily chosen. It is clear that a major error is still present, and that, in the other cases, the superposition of the trihedrons remains approximate. In this sequence, the instances present various levels of difficulty. Therefore, in practical circumstances, the relevance of the pairwise ICPr result will also depend on how judicious the reference choice was. This should be the one that seems the most representative (see also section V.D). Fig. 8.c clearly shows the improvements achieved by the IMCP algorithm ($\lambda = \lambda' = 5$). This result underlines the relevance of the concept of IMCP registration: while, at the same time, discarding the arbitrary part involved in classical pairwise approaches, we can obtain a nearly perfect result that confirms gains in both robustness and accuracy.

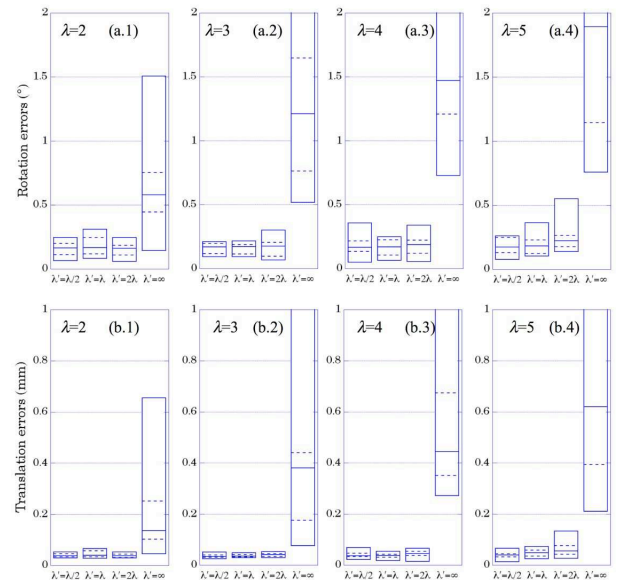


Fig. 9. IMCP estimation errors w.r.t. outlier scale l and consensus mode setting l' . An infinite value implies no consensus at all. The current working cases are similar to the one depicted in Fig. 8.c. Upon convergence, the relative location of an instance is expressed through the rigid transform aligning its trihedron on the virtual trihedron that accounts for the mean of the eight trihedrons. Two numbers then account for the error amplitude of an alignment transforms: the rotation angle around its quaternion vector (top row) and the modulus of its translation vector (bottom row). Each rectangular box plot encloses eight sorted values, including the minimum, median, and maximum values. In order to grasp the significance of the translation errors, we should keep in mind the major extent of the structure – close to 2 cm – as well as the mean edge length of the mesh – 1 mm.

C. Optimal settings of the consensus emergence

Considering the processing of the sequence shown in Fig. 8, Fig. 9 summarizes the statistical analysis of IMCP alignment errors as a function of λ and λ' . Here, each of the error boxes (depicting the extrema, the quartiles, and the median values) accounts for the dispersion of the 8 error estimations. To measure the mutual dispersion of these 8 instances, a trihedron is built that represents the mean location of the 8 alignment trihedrons. The misalignment deviation for an instance is then quantified by its translation modulus and angle

of rotation versus the mean trihedron location.

Despite the fact that the precision can only decrease as λ increases, we can note that the precision of the simultaneous matching remains nearly constant in the range $\lambda \in [2, 5]$. Similarly, the consensus parameter setting does not affect the results as long as λ' remains close to λ . In other respects, decreasing λ' by values lower than λ rapidly becomes ineffective since the quorum rule induces an adaptive increase of λ' each time this is required. On the other hand, increasing λ' by values higher than λ progressively inhibits any expression of the quorum rule. Fig. 9 shows that its inhibition ($\lambda' \rightarrow \infty$) significantly lessens performance. Indeed, in the range $\lambda \in [2, 5]$, we can consider that the λ' setting becomes less significant and we set $\lambda' = \lambda$. However, choosing $\lambda' = \lambda/2$ may become more suitable whenever λ must be set to values higher than 5. For low scale factor values (*e.g.*, $\lambda \leq 2$), it would still be possible to obtain a slight precision gain due to actual rejections of object parts having a fuzzy rejection status. But this gain would be obtained at the expense of long convergence time and would also give false rejections of some useful information. To sum up, in the framework of this study, an optimal choice for λ would range from 4 to 5, which comes down choosing orders of magnitude comparable to those usually prescribed for the use of the Biweight M-estimator (*e.g.*, see [29]).

D. Testing IMCP versus ICPr

Fig. 10 summarizes the statistical analysis of errors linked with both IMCP and pairwise ICPr alignments. The measurement protocol is similar to that on which Fig. 9 is based. We can note that the ICPr results are heavily dependent on the choice of instance taken as the reference. Only two instances lead to acceptable results throughout the whole sequence. Moreover, even if the IMCP algorithm is made to compete with the best ICPr working case (*i.e.*, reference t_8), the precision provided by the IMCP approach is higher by one order of magnitude than that obtained by ICPr. A more meaningful cluster-based comparison is depicted in Fig. 11. It no longer make reference to an average location. Henceforth, it maps the dispersion of the cluster of residual misalignments observed along the $K(K-1)/2$ non-oriented edges in the network embedding the K instances. While addressing the scoring of the ICPr approach, the best choice for its reference instance cannot be known beforehand. Thus, the corresponding cluster (Fig. 11.b) has to expand to $K^2(K-1)/2$ edges so as to fully account for the poor reliability of this standard approach.

The test case depicted in Fig. 8 corresponds to a difficult configuration for which the convergence rate is rather slow. Indeed, a Core2Duo-2.33GHz processor operating through a singlethread implementation needs one minute. However, as the algorithm is intrinsically parallel, a multithread implementation could easily divide this time by the number of cores. Moreover, IMCP-based result improvements clearly counterbalance the processing cost. A simple protocol improvement would involve reducing the initial difficulty level by applying a pairwise pre-alignment using the ICPr

algorithm, since this latter converges in less than one second. However, other experiment contexts, not discussed here, show that this strategy may sometimes worsen the difficulty level encountered by the IMCP algorithm; comparing the ICPr cluster (Fig. 11.b) to the initial cluster (Fig. 11.a) makes this foreseeable. A more promising strategy would be to process through a hierarchical IMCP whenever multi-resolution meshes can be made available.

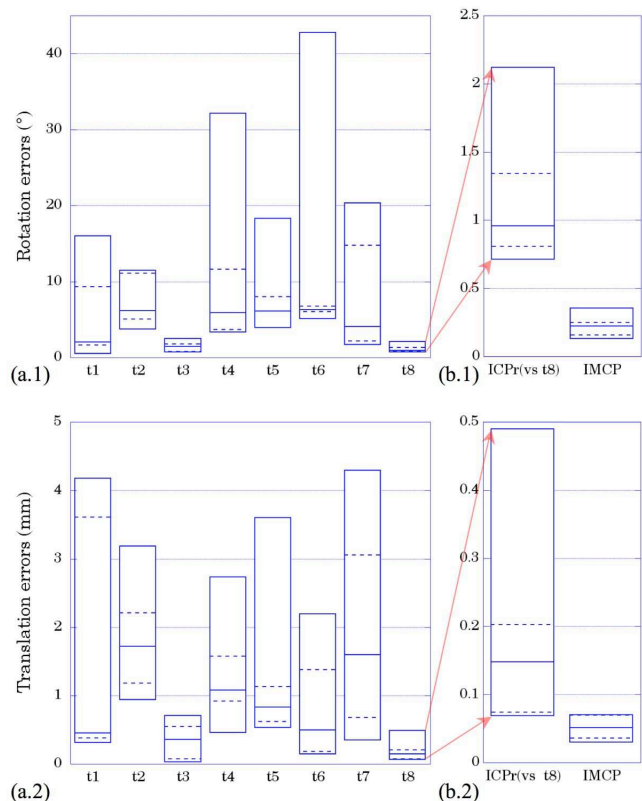


Fig. 10. Comparison of ICPr and IMCP estimations w.r.t. the artificial sequence using the same analyzing technique as Fig. 9. Boxes (a.1) t_1 and (a.2) t_1 both account for the ICPr test case depicted in Fig. 8.b where instance t_1 is arbitrarily chosen as the common reference instance. In order to provide an objective comparison with IMCP results, as the optimal reference index is an unknown parameter, we must take into account all of the possible pairwise ICPr registration configurations w.r.t. the retained reference index. These are drawn by boxes t_1 to t_8 . Even if a user were lucky enough by choosing t_8 as the reference index, Fig. 10.b.1 and 10.b.2 show that the IMCP errors – here those related to Fig. 8.c – remain much more homogeneously distributed and outperform the mean errors of the best ICPr working case with one order of magnitude.

V. DISCUSSION

To counteract the shortcomings of the nearest neighbor operator, a widespread strategy leads to providing the points with attributes that are invariant with respect to the geometric transformation, for example, in the case of rigid transforms, by introducing intrinsic attributes depending on the differential geometry. This allows a strategy to be devised that could lessen, *a priori*, the rate of irrelevant point correspondences. However, this common approach would only transfer our requirement for robustness to the technique used to extract these invariants and, furthermore, would lead us to a less generic algorithmic strategy. More importantly, as our working hypothesis considers shape instances corrupted by

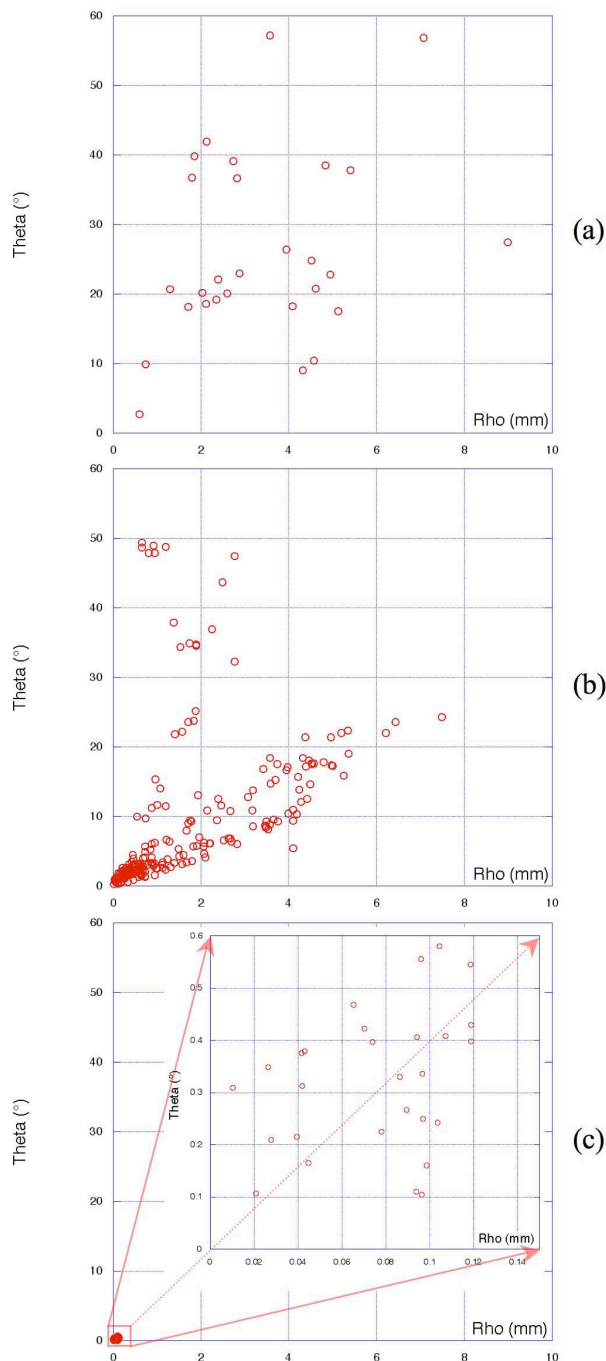


Fig. 11. Bivariate comparison of ICP and IMCP estimations from the artificial sequence. Unlike Fig. 10, where statistical comparisons were made on a univariate basis, we can account for the relative dispersions of the registered instances through bivariate scatter-plots, the translation modulus being termed “Rho” and the rotation angle around quaternion vector being termed “Theta”. The mutual dispersion of the cloud of eight instances makes reference to the full $8(8-1)/2=28$ non-oriented edges of the network linking the eight nodes of the graph, each edge being valued by a pair $\{\text{Rho}, \text{Theta}\}$. Fig. 11.a (resp. 11.c) depicts the scatter-plot of the 28-edge network associated with the initial (resp. IMCP registered) state as shown by Fig. 8.a (resp. 8.c). As a way to provide a statistical counterpart to the reference-index-choice uncertainty, intrinsic to any pair-wise approach, we have to collect observations from the eight possible working configurations. Thus, the ICP results give rise to the cloud of 8×28 -points in Fig. 11.b. Unlike ICP, the IMCP results seem perfect as long as Fig. 11.c is seen through the same drawing scale. Thus, its cluster has to be magnified.

considerable contaminations and noise, it becomes unrealistic to count on reliable estimations of normals and, thus, second

order attributes. Therefore, as this rules out the use of the point-to-plane metric, the point-to-point metric linked to the simplest nearest neighbor operator remains the most suitable.

At first sight, insofar as reliable estimations of normals could have been made available, the alternative global registration approach described in [41] could seem to meet our processing requirements. However, in our processing context, even with the availability of reliable normals, using [41] would still not remain sound. Indeed, since this previous work manages distance field on 3D grids through its *first-order* approximation, it requires bounding the computation of the grid-nodes to the neighborhood of the shape boundaries. Thus, valid matching vectors whose length exceeds a few grid steps become *de facto* labeled as outlier matches. Therefore, performing an accurate and explicit reconstruction of the root shape would require starting from a good pre-alignment. This is to be compared with our poor pre-alignments (see the error magnitudes in Fig. 11.a) making this assumption unrealistic. Thus, as part of a point matching process, the root shape has to remain implicit. Its explicit reconstruction is deferred to an optional and final post-processing step (not discussed here), where robust estimations of the normals then become achievable.

As stated in the introduction, the current consensual technique focuses on tracking the inertia trihedron linked to the binary mask of each bone. It has been extensively probed and compared with marker-based measurements [16]. The relative locations of these trihedrons provide the initial guess required by our algorithm. So, the purpose of our validation step is to show that our algorithm improves both accuracy and robustness w.r.t. inertia-trihedron-based tracking. As such an assessment does not require the availability of a real sequence, we build an artificial sequence involving more difficult problems than those that can be expected from a real sequence. Indeed, since the IMCP accuracy rises to ± 0.1 mm and $\pm 0.4^\circ$, assessing accuracy w.r.t. real sequences would have required the availability of a quasi-perfect ground truth. Conversely, robustness assessments can still operate through an approximate ground truth.

VI. CONCLUSION AND PERSPECTIVES

The notion of median consensus provides noticeable improvements in terms of accuracy and robustness for any problems addressing 3D shape tracking along time and, in particular, problems relating to the analysis of articular movements. Due to the IMCP robustness, pre-processing tasks – like shape segmentation – can involve less reliable techniques such as automated ones. Moreover, the algorithm proves simple to implement and combines well-known techniques. Furthermore, it involves a small number of parameters for which a default setup is proposed. Because widely dissimilar contexts can be tackled without any readjustment, this setup proves to be flexible. The algorithm assessment was made on synthetic sequences built so that they could account for much more difficult cases than those we were expecting to cope within a real sequence. Fig. 12.b, where some instances exhibit considerable truncations, shows

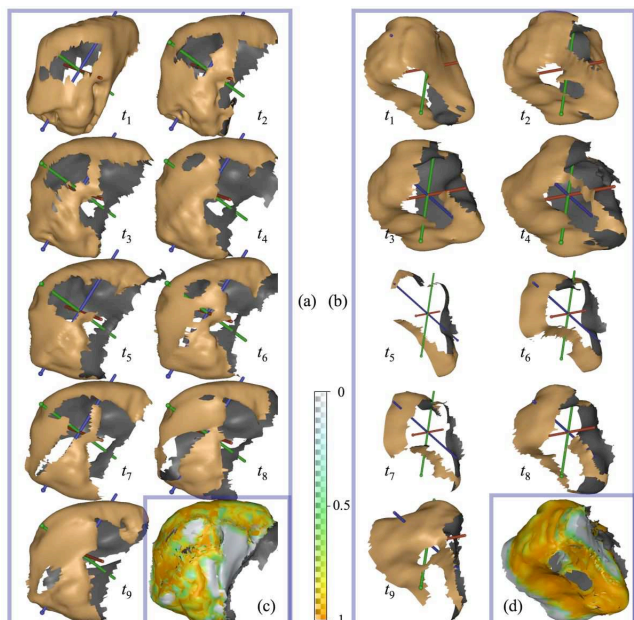


Fig. 12. IMCP tracking of inhomogeneous real sequences. Fig. 3 already depicted the carpus movements under consideration. Box (a) refers to the trapezoid sequence, whereas box (b) refers to the proximal M3 sequence. Upon simultaneous registrations, Fig. 12.c (resp. 12.d) represents a superposition – *not* a fusion – with texture colors accounting for the local membership likelihood, ranging from 0 to 1, to the virtual MCS of the trapezoid (resp. metacarpal M3). In spite of poor initial guess locations, accounted for by the trihedrons, and very sparse – e.g. t_5 t_6 t_7 – residual shape informations, the IMCP algorithm was still able to recover the relevant colocations for each M3 instance.

that the IMCP algorithm successfully deals with each instance of an actual difficult sequence. However, as it may become too hard to achieve whenever the number of instances become large, the convergence criterion still needs future work to make it versatile enough. Possible applications go way beyond just image sequence analysis. Since the IMCP algorithm *simultaneously* performs an implicit synthesis of the actual rigid evolving shape, one useful application could be to link the algorithm to a post-processing step aiming at *explicit* reconstruction of the MCS in super-resolution. Applying accurate morphological analysis techniques [2, 45] would then

ACKNOWLEDGMENT

The authors thank Janet Ormrod – Senior Lecturer with GET-ENST Bretagne, Department LCI – for carefully proof-reading the English writing of the paper.

REFERENCES

- [1] J.K. Udupa, B.E. Hirsch, H.J. Hillstrom, G.R. Bauer, J.B. Kneeland, "Analysis of in vivo 3-D internal kinematics of the joints of the foot", *IEEE Trans. on Biomedical Engineering*, Vol. 45, No. 11, pp.1387–1396, Nov. 1998.
- [2] J.J. Jacq, C. Roux, "Geodesic morphometry with applications to 3D morpho-functional anatomy", *Proceedings of the IEEE*, Vol. 91, No. 10, pp. 1680–1698, Oct. 2003.
- [3] J.J. Jacq, C. Roux, "Automatic detection of articular surfaces in 3-D image through minimal subset random sampling", In J. Troccaz, E. Grimson, and R. Mösges editors, *Computer Vision, Virtual Reality and Robotics in Medicine. LNCS 1205*, Springer, pp. 73–82, Mar. 1997.
- [4] J.J. Jacq, C. Roux, E. Stindel, C. Lefevre, "Analytical surface recognition in three-dimensional (3D) medical images using genetic matching. Application to the extraction of spheroidal articular surfaces

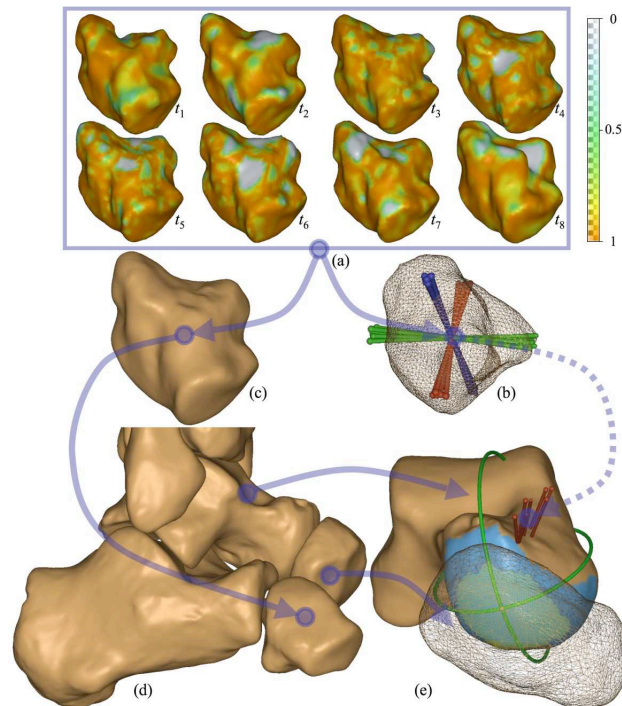


Fig. 13. The IMCP algorithm simultaneously provides two types of improvements: (i) accurate kinematics and (ii) accurate shape description. This picture illustrates MRI-based tarsus kinematics. Once they are registered, each cuboid instance receives its MCS membership map (Fig. 13.a). Fig. 13.c represents an explicit reconstruction of the MCS from the cuboid sequence. Fig. 13.b shows the inertia trihedron errors cancelled through simultaneous registration. After applying the same process to the other bones, we obtain an IMCP-based reconstruction covering the full tarsus (Fig. 13.d). The available shape accuracy then enables us to compute the global symmetries of joint surfaces [45]. Fig.13.e depicts the talo-navicular joint (blue) and its global symmetries (green osculatory circles). On the other hand, availability of accurate and independent kinematics measurements enables us to materialize the helical axes sequence accounting for each step of the movement of the navicular w.r.t. the talus (red axes, Fig. 13.e). Thus, applying this full framework would enable us to materialize inter-relations between global symmetries of the joint surface and the set of finite helical axes linking consecutive positions of the congruent bones.

enable the shape characteristics of articular surfaces to be placed in relation to the observed movements. The preliminary results shown in Fig. 13 constitute a first step in this direction.

- in 3D computed tomography data sets", *International Journal of Imaging System and Technology*, Vol. 11, No. 1, pp. 30–43, May 2000.
- [5] H.K. Hahn, H.O. Peitgen, "IWT – Interactive Watershed Transform: A hierarchical method for efficient interactive and automated segmentation of multidimensional grayscale images", in *Proc. of Medical Imaging, SPIE 5032*, San Diego, Feb. 2003.
- [6] J. Crisco and R. McGovern, "Efficient calculation of mass moments of inertia", *J. of Biomechanics*, Vol. 31, pp. 97–101, 1998.
- [7] C. Neu, R. McGovern and J. Crisco, "Kinematic accuracy of three surface registration methods in a three-dimensional wrist bone study", *J. of Biomechanical Engineering*, Vol. 122, pp. 528–533, 2000.
- [8] G. Marai, D. Laidlaw, C. Demiralp, S. Andrews, C. Grimm and J. Crisco, "Estimating joint contact areas and ligament lengths from bone kinematics and surfaces", *IEEE Transactions on Biomedical Engineering*, Vol. 51, No. 5, 2004.
- [9] C. Neu, J. Crisco and S. Wolfe, "In vivo kinematic behavior of the radio-capitate joint during wrist flexion-extension and radio-ulnar deviation", *J. of Biomechanics*, Vol. 34, pp. 1429–1438, 2001.
- [10] D. Moore, K. Hogan, J. Crisco, E. Akelman, M. Dasilva and A.-P. Weiss, "Three dimensional in vivo kinematics of the distal radioulnar joint in malunited distal radius fractures", *J. Hand Surg.*, Vol. 27, pp. 233–242, 2002.

- [11] S. Wolfe, C. Neu, J. Crisco and N. Haven, "In vivo scaphoid, lunate, and capitate kinematics in flexion and in extension", *J. Hand. Surg.*, Vol. 25, pp. 860–869, 2000.
- [12] J. Crisco, R. McGovern and S. Wolfe, "Non-invasive approach to measuring in vivo three-dimensional carpal kinematics", *J. of Orthopaedic Research*, Vol. 17, pp. 96–100, 1999.
- [13] B. Hirsch, J. Udupa and D. Roberts, "Joint kinematics can be determined by computerized three dimensional imaging of MRI data", *Anat Rec.*, Vol. 45, p. 226, 1990.
- [14] J.K. Udupa, B.E. Hirsch, S. Samarasekera and R. Goncalves, "Joint kinematics via three-dimensional MR imaging", *In Visual Biomed Comp., SPIE*, Vol. 1808, pp. 664–669, 1992.
- [15] B. Hirsch, J. Udupa and S. Samarasekera, "New method of studying joint kinematics from three-dimensional reconstructions of MRI data", *J. Am. Podiatr. Med. Assoc.*, Vol. 86, No. 1, pp. 4–15, 1996.
- [16] S. Siegler, J.K. Udupa, S.I. Ringleb, C.W. Imhauser, B.E. Hirsch, D. Odhner, P.K. Saha, E. Okereke and N. Roach, "Mechanics of the ankle and subtalar joints revealed through a 3D quasi-static stress MRI technique", *J. of Biomechanics*, Vol. 38, No. 3, pp. 567–578, 2005.
- [17] P.J. Besl, N.D. McKay. "A method for registration of 3-D shapes". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, pp. 239–256, 1992.
- [18] P.K. Saha, J.K. Udupa, A.X. Falcão, B.E. Hirsch, and S. Siegler, "Iso-shaping rigid bodies for estimating their motion from image sequences", *IEEE Trans. on Medical Imaging*, Vol. 23, No. 1, pp. 63–72, Jan. 2004.
- [19] G.E. Marai, D.H. Laidlaw, J.J. Crisco, "Super-resolution registration using tissue-classified distance fields", *IEEE Trans. on Medical Imaging*, Vol. 25, No. 2, Feb. 2006, pp. 177–187.
- [20] T. Cresson, J.J. Jacq, V. Burdin, C. Roux, "Performing accurate rigid kinematics measurements from 3D in vivo image sequences through median consensus simultaneous registration", in *Proc. of the IEEE Engineering in Medicine and Biology 27th Annual Conference*, Shanghai, Sept. 2005.
- [21] X. Pennec, "Multiple registration and mean rigid shapes: application to the 3-D case", in *Proc. 16th Leeds Annual Statistical Research Workshop*, pp. 178–185, Leeds, UK, Jul. 3-5, 1996.
- [22] M.A. Audette, F.P. Ferrie, and T.M. Peters. "An algorithmic overview of surface registration techniques for medical imaging". *Medical Image Analysis*, 4(3):201–217, Sept. 2000.
- [23] A. Lorusso, D.W. Eggert, R.B. Fisher. "A comparison of four algorithms for estimating 3D rigid transformation". Internal Report, Dpt of Artificial Intelligence, Univ. of Edinburgh, 28p, 1996.
- [24] D.W. Eggert, A. Lorusso, R.B. Fisher. "Estimating 3-D rigid body transformation: a comparison of four major algorithms". *Machine Vision and Applications*, No. 9, pp. 272–290, 1997.
- [25] K.S. Arun, T.S. Huang, S.D. Blostein. "Least-squares fitting of two 3-D point sets". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 5, pp. 698–700, 1987.
- [26] P.J. Huber, "Robust estimation of a location parameter", *Ann. Math. Stat.*, 35, pp 73-101, 1964.
- [27] P.J. Huber, "Robust statistics", John Wiley & Sons, 1981.
- [28] M.J. Beaton, J.W. Tukey, "The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data", *Technometrics*, 16, pp. 147–185, 1974.
- [29] R.M. Haralick, H. Joo, C.N. Lee, X. Zhuang, V.G. Vaidya, M.B. Kim, "Pose estimation from corresponding point data". *IEEE Trans. on System, Man and Cybernetics*, Vol. 19, No. 6, pp. 1426–1446, Nov. 1989.
- [30] Y. Chen, G. Medioni, "Object modelling by registration of multiple range images", *Image and Vision Computing*, 10(3), pp. 145–155, Apr. 1992.
- [31] J.M. Rouet, J.J. Jacq, and C. Roux, "Genetic algorithms for robust 3-D MR-CT registration", *IEEE Trans. on Information Technology in Biomedicine*, Vol. 4, No. 2, pp. 126–136, Jun. 2000.
- [32] G. Godin, P. Boulanger, "Range image registration through invariant computation of curvature", in *Proc. ISPRS Workshop From Pixels to Sequences*, Zurich, pp. 170–175, 1995.
- [33] Z. Zhang. "Iterative point matching for registration of free form curves and surfaces". *Int. J. of Computer Vision*, Vol. 12, No. 2, pp. 119–152, 1994.
- [34] J. Guehring, "Reliable 3D surface acquisition, registration and validation using statistical error models", in *Proc. of the 3rd IEEE International Conference on 3-D Digital Imaging and Modeling*, 2001, pp. 224–231.
- [35] K. Nishino, K. Ikeuchi, "Robust simultaneous registration of multiple range images", in *Proc. of 5th Asian Conf. on Computer Vision (ACCV'02)*, pp. 454–461, Melbourne, Australia, Jan. 2002.
- [36] P.J. Neugebauer, "Geometrical cloning of 3D Objects via simultaneous registration of multiple range images", in *Proc. of Int. IEEE Conf. on Shape Modeling and Application*, Mar 1997, pp. 130–139.
- [37] R. Benjema and F. Schmitt, "A solution for the registration of multiple 3D point sets using unit quaternions", in *Proc. ECCV '98*, Freiburg, Germany, Jun. 1998, pp. 34–50.
- [38] J. Williams and M. Bennamoun, "Simultaneous registration of multiple corresponding point sets", *Computer Vision and Image Understanding*, 81, pp. 117–142, 2001.
- [39] A. J. Stoddart and A. Hilton, "Registration of multiple point sets", in *Proc. IEEE Int. Conf. on Pattern Recognition (ICPR'96)*, Vienna, Aug. 1996, Vol. A, pp. 40–44.
- [40] D. W. Eggert, A. W. Fitzgibbon, and R. B. Fisher, "Simultaneous registration of multiple range views for use in reverse engineering of CAD models", *Comput. Vision Image Understanding*, 69, Mar. 1998, pp. 253–272.
- [41] T. Masuda, "Registration and integration of multiple range images by matching signed distance fields for object shape modeling", *Computer Vision and Image Understanding*, 87, pp. 51–65, 2002.
- [42] S.J. Cunningham, A.J. Stoddart, "N-View point set registration: a comparison", *British Machine Vision Conference*, Nottingham, UK 1999.
- [43] E. Maschino, Y. Maurin, P. Andrey, "Joint registration and averaging of multiple anatomical surface models", *Computer Vision and Image Understanding*, Vol. 101, pp. 16–30, 2006.
- [44] K. Kafadar, "John Tukey and robustness", *Statistical Science*, Vol. 18, No. 3, pp. 319–331, 2003.
- [45] S. Allaire, J.J. Jacq, V. Burdin, C. Couture, C. Roux, "Type-constrained robust fitting of quadrics with application to the 3D morphological characterization of saddle-shaped articular surfaces", in *Proc. of the 11th IEEE International Conference on Computer Vision (ICCV'07), workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, 14–20 Oct., 2007, Rio de Janeiro.