A real-coded evolutionary algorithm-based registration approach for forensic identification using the radiographic comparison of frontal sinuses

Óscar Gómez†‡, Pablo Mesejo†‡, Óscar Ibáñez†‡, Andrea Valsecchi†‡ and Óscar Cordón†‡
*Dept. of Computer Science and Artificial Intelligence, University of Granada, Spain.
†Andalusian Research Institute DaSCI, University of Granada, Spain.
‡Panacea Cooperative Research S. Coop., Ponferrada, Spain.
Email: ogomez@decsai.ugr.es, pmesejo@decsai.ugr.es, oscar.ibanez@panacea-coop.com, valsecchi.andrea@panacea-coop.com, ocordon@decsai.ugr.es

Abstract—Comparative radiography is the forensic anthropology technique in which ante-mortem (AM) and post-mortem (PM) radiographic materials (e.g., X-ray images or CTs) are compared in order to determine the identity of a deceased human being. One of the most commonly used anatomical structures in comparative radiography are the frontal sinuses. The frontal sinuses are osseous cavities located in the skull, which are used in forensic identification tasks due to their singularity and high identification power. In order to automate the comparison of frontal sinuses in AM and PM materials, it is necessary to perform the registration of these materials (i.e., it is necessary to carry out the alignment of these anatomical regions). However, the manual alignment of these structures is a time-consuming and subjective process. In order to tackle this problem, this paper presents an automatic frontal sinuses registration method in comparative radiography using real-coded evolutionary algorithms (RCEAs). The task is formulated as a 2D-3D image registration problem using a 9 Degrees of Freedom perspective transformation model; two RCEAs (DE and MVMO-SH) are compared in the minimization of the registration cost function, and the best of them (MVMO-SH) is applied to an identification scenario including 50 X-ray images and 50 CTs. The results obtained show that the proposed automatic registration system is able to filter more than 80% of the sample.

Index Terms—Biomedical Image Registration, Forensic Identification, Comparative Radiography, Frontal Sinuses, Real-Coded Evolutionary Algorithms, Mean Variance Mapping Optimization, Differential Evolution

I. INTRODUCTION

Comparative radiography (CR) [1] is a forensic identification technique based on the comparison of skeletal structures in ante-mortem (AM) and post-mortem (PM) radiographs. Since the discovery of X-rays by Roentgen in 1895 [2], forensic experts have made use of radiographic images as evidence in their endeavour (e.g. bullet analysis [3], age estimation [4], and forensic identification [5]). During the first decades of the twentieth century, the use of X-rays as a method of positive identification gradually consolidated in scientific literature. In fact, in 1949, CR techniques played a crucial role in the identification of people involved in the Noronic ship’s disaster, proving their importance for identification and being included in many mass disaster identification protocols [6]. Nowadays, CR is still employed in many forensic identification scenarios. For instance, the Michigan State University Forensic Anthropology Laboratory (MSUFAL) performed 193 identifications using this approach between 2002 and 2015 [7].

Several bones and cavities have been reported as useful for candidate short-listing or positive identification based on their individuality and uniqueness [8]. In particular, frontal sinuses (see Fig. 1) are widely recognized as a useful and reliable method of identification [9], fulfilling the Daubert criteria [10]1. Frontal sinuses are only absent in 4% of the population and are maintained unchanged during the rest of the life [12]. Although, rarely, some external factors such as traumatisms can change slightly their morphology, frontal sinuses are considered as a skeleton fingerprint. Their utilization for CR-based identification was first reported in 1926 by comparing their morphology in AM and PM radiographs [13]. Nowadays, CR identification based on frontal sinuses is widely accepted by the forensic community, and many works have reported their utility via image comparison to establish positive identification [14]–[16].

CR techniques have lower cost and time requirements in comparison to DNA analysis, which are crucial factors in mass disaster victim identification scenarios. However, the application of CR requires the superimposition of the AM and PM data for their visual comparison by producing PM radiographs simulating the AM ones in scope and projection. This is a time-consuming trial-and-error process, that relies completely on the skills and experience of the analyst. Furthermore, the utility of the method is reduced because of the errors related to analysts’ fatigue and subjectivity. There is thus a need to automate CR-based identification methods.

The automation of the CR’s superimposition process is complex and computationally expensive (see Section II for further details). This is due to several reasons, such as the unknown set-up of the AM radiograph, the fact that image

1The Daubert criteria [11] determine whether evidence is admissible in a court of law. An identification method fulfills the Daubert criteria when: (1) it is testable and peer reviewed; (2) it possesses known potential error rates; and (3) it is accepted by the forensic community.
intensities are not reliable or even not captured, and the high multimodality of the search space configured by the registration framework, among others. These reasons make classic 3D-2D image registration (IR) techniques [18] not suitable for CR, and more sophisticated techniques should be considered in order to solve it, as real-coded evolutionary algorithms (RCEAs). In particular, the current state-of-the-art approach for CR follows an evolutionary 3D-2D IR methodology based on Differential Evolution (DE) [19]. In [19], several numerical optimization methods based on both linear search (Nelder-Mead, BFGS, LBFGS) and trust-region (Levenberg-Marquardt, BOBYQA) were tested to solve our IR problem, but their performance was insufficient. DE is taken as comparative reference in the experimentation of this paper, but we also introduce a competitor approach termed Mean-variance mapping optimization (MVMO) [20]. The reason to include this RCEA can be found on its excellent performance and robustness when tackling computationally expensive real-coded optimization problems as 3D-2D IR for CR an instance of these kinds of problems. MVMO is in top positions in expensive optimization competitions, such as IEEE CEC 2013 [21], 2014 [22], 2015 [23], 2016 [24], and 2018 [25].

This paper presents an automatic frontal sinuses registration method in CR using RCEAs. The contribution of this research work is twofold. First, this is the first time that the problem is posed as a 3D-2D IR problem using a 9 Degrees of Freedom (9-DoF) perspective transformation model (a more complete model than the one previously employed [19]). Second, we employ a novel RCEA (MVMO-SH) that allows us to achieve state-of-the-art results in the IR-based CR problem and, therefore, up to our knowledge, it represents the most accurate automatic CR identification system available nowadays.

The rest of the paper is structured as follows. Section II briefly reviews the current state of the art in IR for CR. Section III describes our proposal to tackle the 3D bone scan-2D radiograph superimposition problem (9-DoF registration model and MVMO-SH). Section IV presents experiments and results obtained. The main conclusions are in Section V.

II. BACKGROUND AND RELATED WORKS

IR [18] is the process of aligning two or more images of the same or different dimensionalities into one coordinate system. 3D-2D IR approaches (e.g., Computed Tomographies (CTs)-radiograph or 3D surface mode-radiograph in CR) are classified into intensity-based and feature-based. Intensity-based methods compare intensities [18] of a 2D projection of the volumetric image with a fixed 2D image. However, one crucial consideration in any automatic method for CR is that the intensity level could have changed between the AM and the PM images as the bone density changes within the individual through time (due to factors as aging, osteoporosis, and the PM interval). Additionally, the intensity information varies depending on the X-ray acquisition protocols, devices employed, and cannot be acquired by 3D surface scans, which are being increasingly used by forensic labs [26] to scan PM "clean" bones due to their great availability and low cost, while just a few of them can afford a CT scan.

The former limitations lead to feature-based approaches. Feature-based IR methods minimize the distance between geometrical features to be segmented in both images. Forensic anthropologists consider bone morphology (silhouette) a reliable marker for performing the CR-based identification to compare radiographs of frontal sinuses [27], clavicles [28], and patellae [29]. However, these works cannot be considered as automatic IR methods since they only compare silhouettes using elliptical Fourier analysis [30]. The former one is based on the comparison of AM radiograph with a PM radiograph. The latter two, in contrast, are based on the comparison of the AM radiograph with a set of a predefined 2D projected images obtained through the rotation of 3D surface models acquired with a 3D laser range scanner. However, none of these completely automate the search for the best possible 2D projection of the PM 3D surface model of the bone.

Another important consideration about 3D-2D IR methods for medical domains [31] is that most of them are designed for a controllable set-up. Therefore, they can assume a calibrated case where the parameters related to the perspective distortions...
are known, and only considering the parameters related to
translations and rotations (6 DoF), and with a initialization
pose close to the ground truth (GT) pose (i.e. a maximum tar-
get registration error [] of 16 mm in [31], etc). However, these
assumptions are not suitable for CR since the AM radiograph
was taken in an unknown pose with an unknown radiograph
device. Therefore, the search for the optimal solution in the CR
scenario is more complex. There are a few exceptions such as
Feldman et al. [32], that proposed a 3D-2D IR method based
on the silhouette that does not rely on assumptions about the
initial pose by using free-form curves and surfaces, but it is
only applicable in the calibrated case (6-DoF).

IR methods based on RCEAs, a.k.a. evolutionary IR meth-
ods, have demonstrated to overcome some of these drawbacks
in others IR problems [33]. In particular, Gómez et al. [19]
proposed an evolutionary 3D-2D IR approach for CR based on
the bone or cavity silhouette. It automatizes the search of the
best possible 2D projection of the PM 3D surface model of the
bone (either obtained using a 3D scanner or segmented
from a PM CT), and it does not consider any assumption on
the initialization or the main parameter related to the
perspective distortions in radiographs (i.e. the source to image
distance, a.k.a. SID). This proposal is based on the use of DE
[34], a modification of the DICE metric [35] that considers
occlusion regions (which are regions hard to segment either
because of the fuzzy borders of the bone or occlusions caused
by other overlapped structures), and a simple perspective
transformation (with 7 parameters: 3 translations; 3 rotations;
and the SID). However, [19] does not model all sources of
perspective distortion in a radiograph. In particular, the
perspective transformation employed on that work does not
model angled radiographs. In angled radiographs (e.g., Waters’
view radiographs), the ray that joins the center of the image
receptor and the X-ray generator is not perpendicular, resulting
in a displacement of the principal point. This displacement
does perspective distortions. The method presented in [19]
also showed the following drawbacks: (1) the robustness of
the DE algorithm, especially with clavicles and patellae, that
in some runs led to bad superimpositions due to the stochastic
nature of DE and the highly multimodal search space tackled
(see [19] for a study of the multimodal landscape of the
problem); and (2) the large amount of time required to obtain
a superimposition with DE (on average, 1800 seconds). This
long time is motivated by the high computational time required
by each evaluation (on average, it takes 0.250 seconds for
a projection of 1290 × 1050 pixels in a standard computer),
uncovering the computationally expensive optimization nature
of the CR problem, and the high number of evaluations
required by the optimizer to converge to a solution.

To overcome the aforementioned limitations, in this paper,
we first extend the registration model (from a 7-DoF to a
9-DoF model) by including two more parameters to correct
for perspective distortions (corresponding to the X and Y
axis). This will allow us to model posterior-anterior and
Waters’ views for frontal sinuses. Second, we substitute DE
by a RCEA specialized on expensive optimization real-coded
problems (MVMO-SH).

III. METHODOLOGY

The evolutionary IR method requires the five following
components (these are further detailed in [19] and depicted
in Fig. 2): (1) the model (PM 3D surface model of the sinus
cavity) and the scene image (AM radiograph, where the silhou-
ette of the bone/cavity is segmented as well as the occlusion
region, a.k.a. the region where the segmentation expert cannot
distinguish if there is bone/cavity or not); (2) the perspective
transformation responsible of generating a 2D image from a
3D object; (3) the expert knowledge of the problem that
delimits the target transformation (i.e. radiographs acquisition
protocols [36]); (4) a similarity metric which measures the
resemblance of a 2D projection with the original 2D image
(overlapping); and (5) a RCEA, which looks for the best
parameters for the transformation to minimize the error of the
similarity metric.

The projective transformation [37] behind an X-ray image is
a perspective transformation with 9 parameters (6 extrinsic pa-
rameters: 3 translations; 3 rotations; and 3 intrinsic parameters:
1 SID and 2 movements of the principal point). Notice that in
a radiograph the perspective distortion is mainly related to the
source to SID [38] (see Figure 2) instead of the focal distance.
However, angled radiographs, which are radiographs acquired
with procedures where the ray that joins the X-ray generator
and the center of the image receptor is not perpendicular,
also have perspective distortions due to the movement of the
principal point (see Fig. 2 for a graphical example).

Most works consider a calibrated scenario (only 6 pa-
rameters) and the SID is assumed as known which is not
the case for the CR problem [31]. Although the perspective
distortion can be small in many radiographs because of the
large distance between the X-ray generator and receptor (as in
chest radiographs), its consideration has shown to be crucial in
the IR endeavour. This has been shown in [19], where better
results were obtained using the perspective transformation than
the orthographic transformation, despite the more challenging
optimization problem involved.

The goodness of a projection of the PM 3D model with
respect to fixed 2D segmentation of the radiograph is measured
by a similarity metric, used as fitness function. This metric
measures the resemblance of a 2D projection with the original
2D image (overlapping). The most utilized metric to measure
the overlap of silhouettes is the DICE metric [35]. However,
this metric is not robust against occlusion and does not allow
partial matching. These drawbacks are overcome for the CR
problem using the Masked DICE metric [19] (see eq. 1),
designed ad-hoc for the CR problem. This metric incorporates
the information of an occlusion region segmented by the
expert into the DICE metric [35] increasing the robustness
of occlusions and partial matching.

\[
\text{Masked DICE} = \frac{2 \cdot |(I_A \setminus M) \cap (I_B \setminus M)|}{|I_A \setminus M| + |I_B \setminus M|} \tag{1}
\]
Fig. 2. Scheme of the proposal of 3D-2D IR for CR. Three main interconnected blocks are represented: (Right) the perspective transformation to obtain a projection of the 3D model with 9 parameters: translation \((t_x, t_y, \text{and} t_z)\), rotation \((r_x, r_y, \text{and} r_z)\), and perspective distortions \((\text{SID}, \beta_x, \text{and} \beta_y)\). In this figure, for the sake of clarity, we display distortions related to angled radiographs using the principal point. However, within the optimization process, instead of the location of this principal point we employ the angles of incidence of the principal ray on the center of the image receptor \((\beta_x \text{and} \beta_y)\). These angles can be used to calculate the principal point displacement; (Top left) The similarity metrics that compares the PM projection (colored in blue) and the AM segmentation (colored in red) considering an occlusion region (colored in gray); (Bottom left) the optimization process to estimate the 9 parameters of the registration transformation that are only weakly limited by the context and expert knowledge from the X-ray acquisition protocol [36].

where \(I_A\) is the set of pixels of object A (segmented bone) silhouette, \(I_B\) is the set of pixels of object B (PM project bone) silhouette, and M is the occlusion region. Lastly, Gomez et. al [19] studied the complexity of the search space of CR showing the multi-modality of the CR problem even in its simplest version (synthetic cases without occlusions).

A. MVMO-SH

The search space of the perspective transformation is complex and highly multi-modal [19]. Therefore, classic numerical optimization methods are not sufficient, and more sophisticated techniques should be considered in order to solve it satisfactorily as RCEAs [33], [39].

The best real-coded evolutionary algorithm (RCEA) for solving computationally expensive real-coded optimization problems according to the IEEE CEC competitions is the mean-variance mapping optimization (MVMO) optimizer [20]. MVMO has ranked in top positions in expensive optimization competitions, such as IEEE CEC 2013 [21], 2014 [22], 2015 [23], 2016 [24], and 2018 [25], showing an excellent performance and robustness. MVMO is a novel single-individual RCEA that considers a best solution archive, but its novelty lies within a new mapping function employed for mutating the offspring. This mapping function is based on the mean and variance of the best solution archive. MVMO has been numerically compared to other enhanced RCEAs showing a better performance in many problems, especially in terms of convergence speed. For instance, a powerful memetic variant called MVMO-SH (the “S” refers to the offspring approach based on single parent and multi-parent crossover, and the “H” for the hybridization of MVMO with the use of local search (LS) [40]) improves the global search performance of the classical MVMO. MVMO-SH considers a set of solutions (i.e. particles of a swarm) instead of just one, each having its own best solution archive and mapping function, and allows the exchange of information and dynamic reduction of the swarm size.

MVMO-SH [21] begins with a initialization stage where the \(p\) particles (candidate solutions) of the swarm are randomly generated. The particles are normalized to the range [0, 1], which is a necessary condition to the latter mutation via mapping function (a key element in MVMO) and are only de-normalized for their fitness evaluation. Afterward, the following steps are performed for each generation (these are detailed in depth in [21]): (1) LS optimization of the particles with a probability \(p_{LS}\); (2) If a particle founds a better solution in terms of fitness than those in its solution archive, the new solution is added to the particle’s solution archive (notice that if the archive has reached its maximum size \(A_s\) the solution archive’s worst solution is removed); (3) Particles are sorted and divided into two groups according to their fitness value, the \(GP\) best ones are classified as “good particles” and the rest as “bad particles” \(GP\) is adapted along the process taking values between the 20% and 70% of \(p\). The good particles are modified via a custom single parent crossover
operation based on local best [22] and bad particles via a custom multi-parent crossover operation based on a subset of good particles [22]; (4) the particles are mutated using a mapping function. This mapping function is based on the mean and variance of each particle’s solution archive and a scaling factor \( f_s \) that modulates the function’s shape. The scaling factor usually begins with a small value \( f_{\text{start}} \) and progressively increases until reaching its maximum value \( f_{\text{end}} \) to progressively increase the algorithm’s accuracy.

To sum up, the most relevant parameters are: number of particles \( p \) (the recommended value is \( 15 \times \text{number variables} \)). If the number of particles chosen is equal to 1, MVMO-SH will perform as the standard MVMO, LS probability \( p_{\text{LS}} \), archive size \( A_s \), scaling factor start \( (f_{\text{start}}) \) and end values \( (f_{\text{end}}) \), initial value of the shape of all the variables at the beginning of the optimization \( d_r \) (values around 1-5 are suitable to guarantee good initial performance. In practice, it is usually set to 1), and parent selection method (random, neighbor group selection in single step or block steps, or sequential selection of the first variable and the rest randomly).

IV. EXPERIMENTS

The experimental study is divided into two parts. The first experiment is devoted to the comparison of DE and MVMO-SH with simulated CR problems of frontal sinuses. Meanwhile, the second experiment is devoted to study the identification capability of the proposed IR framework using the best resulting RCEA in real images of frontal sinuses.

The same stopping criteria is established for the two RCEAs to allow a fair comparison in terms of computational resources. The optimization process ends when at least one of the following three conditions hold: (1) the maximum number of evaluations is reached. This value is set to 50,000 evaluations (it includes the evaluations performed by the LS methods); (2) the optimization process has got stuck. It is considered that the optimization process has stagnated when it has performed 10,000 evaluations without improving the fitness of the best solution; and (3) the optimization process has achieved a good solution/superimposition. A solution is considered of good quality when it shows an error lower than 0.001 in terms of fitness (i.e. the 99.9% of the pixels are correctly overlapped).

All experiments have been performed on the high-performance computing server Alhambra (Univ. of Granada), composed of 1808 cores Fujitsu PRIMERGY CX250/RX350/RX500 nodes, although on average only 50 cores were available for this experimentation.

A. Performance metrics

Two ground truth (GT) metrics are employed to objectively measure the quality of the superimpositions achieved by the RCEAs: GT DICE [35] and the mean reprojection distance error (mRPD) [41]. The GT DICE metric measures the overlap between the GT projection’s silhouette (equal to the simulated AM projection but without any occlusion) and the 2D projection’s silhouette achieved by the RCEA. However, the GT DICE metric and the fitness function (i.e. Masked DICE) are highly correlated (e.g. they are equal in cases without occlusions) and thus, to avoid any possible bias, the mRPD metric is also employed. mRPD is a standard metric for the evaluation of 3D-2D IR methods by computing the reprojection error between the transformation obtained by the RCEA and the GT transformation (see [19] for further details of the utilization of mRPD in the CR problem). Notice that these metrics can be employed only in simulated CR problems since in real CR problems the GT projection and the GT transformation are unknown.

B. Experiment I: DE vs MVMO-SH on a simulated dataset

The dataset employed in Experiment I is formed by 150 simulated CR problems of frontal sinuses, each of them composed of a 3D surface model and a simulated radiograph of 480 × 600 pixels. All 3D surface models were manually segmented from CTs (see [19] for further details about the segmentation protocol). The simulated radiographs are randomly generated using a 9-DoFs perspective transformation within the ranges showed in Table I (these ranges have been set based on international acquisition protocols [36]). Each simulated radiograph has a different degree of occlusion in the frontal sinus silhouette: 50 simulated radiographs have no occlusion (0%); 50 simulated radiographs have an occlusion of around 20% of pixels; and the last 50 simulated radiographs have an occlusion of around 40% of pixels.

1) Experimental set-up: This experiment involves the application of the two optimizers (DE and MVMO-SH) using the 9-DoFs perspective distortions model. The parameters employed by both RCEAs are the following:

- DE: \( p = 100, F = 0.5, \) and \( P_r = 0.5 \) (tuned in [19]).
- MVMO-SH: \( p = d, A_s = 4, \) and \( F_{\text{end}} = 2.5 \) \([39]\).

10 independent runs are performed to avoid any possible bias caused by the stochastic component of the optimizer, resulting in 3,000 runs/superimpositions and around 200 computation hours using 50 cores of the said infrastructure.

2) Results: As can be seen in Table II, the best results are clearly obtained by MVMO-SH. In all metrics, this optimizer yields better (or equally good) results than DE. Importantly, statistical tests to compare the performance of DE vs MVMO-SH were performed. The Wilcoxon rank sum test provided a p-value of 2.546e-13 (GT DICE) and 2.411e-15 (mRPD), showing the existence of statistically significant differences between the performance of both optimizers, in favor of MVMO-SH. Furthermore, MVMO-SH is more robust to occlusions in the silhouette of the skeletal structure than DE, showing a superb performance even with occlusions up to the 40% of silhouette (the maximum degree of occlusion evaluated). Regarding execution time to reach one of the three convergence criteria, DE presented an average, median and standard deviation execution time of 1665.0, 1440.2 and 913.1843 seconds, respectively. In turn, MVMO-SH took 1056.3, 1111.5 and 396.3469 seconds, respectively. Therefore, MVMO-SH is a better and faster optimizer for this particular problem.

DE converged 11.1% of executions, finished because it had obtained a good superimposition 17.1% of times, and
TABLE I
PARAMETER RANGE OF EACH SKELETAL STRUCTURE ACCORDING TO INTERNATIONAL ACQUISITION PROTOCOLS [36] AND EXPERT KNOWLEDGE. EACH MILLIMETER CORRESPONDS TO 2 PIXELS IN THE DIGITAL IMAGE DOMAIN.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Frontal Sinuses</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_x$ (mm)</td>
<td>[-125, 125]</td>
</tr>
<tr>
<td>$t_y$ (mm)</td>
<td>[-150, 150]</td>
</tr>
<tr>
<td>$t_z$ (mm)</td>
<td>[900 - 200, 900 + 200]</td>
</tr>
<tr>
<td>$r_x$, $r_y$, and $r_z$ (degrees)</td>
<td>[-40°, 40°]</td>
</tr>
<tr>
<td>SID (mm)</td>
<td>[1000 - 100, 1000 + 100]</td>
</tr>
<tr>
<td>$\beta_x$ (degrees)</td>
<td>[-10°, 10°]</td>
</tr>
<tr>
<td>$\beta_y$ (degrees)</td>
<td>[-50°, 10°]</td>
</tr>
</tbody>
</table>

TABLE II
COMPARISON PERFORMANCE ACCORDING TO OPTIMIZER AND DEGREE OF OCCLUSION. THE LOWER THE VALUES THE BETTER.

<table>
<thead>
<tr>
<th>Optimizer</th>
<th>Occlusion</th>
<th>Masked DICE</th>
<th>GT DICE</th>
<th>mRPD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>median</td>
<td>std</td>
<td>mean</td>
</tr>
<tr>
<td>DE</td>
<td>0</td>
<td>0.004</td>
<td>0.001</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.008</td>
<td>0.001</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.012</td>
<td>0.001</td>
<td>0.038</td>
</tr>
<tr>
<td>MVMO-SH</td>
<td>0</td>
<td>0.002</td>
<td>0.001</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

achieved the maximum number of evaluations 71.8% of times. MVMO-SH converged 21.4% of executions, obtained a good superimposition 78.6% of times, and it never needed to achieve the maximum number of evaluations.

C. Experiment II: Identification capability of our 3D-2D IR-based CR framework with frontal sinuses on real data

The dataset employed was provided by the Hospital de Castilla-La Mancha (Spain) and is composed of 50 CTs and 50 radiographs where the frontal sinuses are visible. All CTs and radiographs were manually segmented by two MSc students (José Manuel Pérez Jiménez and Andrea Cerezo Vallecillo), both from the Physical Anthropology lab of Univ. of Granada.

1) Experimental set-up: This experiment studies the identification capability of the proposed 3D-2D IR-based CR framework using frontal sinuses and the best RCEA (MVMO-SH). We have performed a radiograph inter-expert study, where we have compared the two segmentations performed by the two aforementioned forensic experts on 50 radiographs against 50 CTs. In this experiment, we have only used the segmentation of the CTs performed by one of the MSc students, since the inter-expert segmentation error in CTs was practically non-existent. A total of 2,500 cross comparisons per segmentation.

Since previous experiments have already shown the robustness of MVMO-SH, and due to the large computational cost of employing again 10 repetitions, only 2 independent runs are performed. Each of the 5,000 runs takes on average 1,000 seconds, resulting in 2,777 hours of computation (or 166 computation days) that, performed on the 50 available cores of computing server Alhambra, required “only” around 56 computation days).

2) Results: Positive and negative cases have shown important differences in terms of fitness according to the Masked DICE Metric (see Fig. 3). However, this metric alone is not sufficient to precisely distinguish between positive and negative cases. Therefore, the results are reported using Cumulative Match Characteristic (CMC) curves (see Fig. 4) to study the identification capabilities of the proposal as done in [42] and in [19]. To focus on the identification reliability of the method only the best run (out of two) of each experiment is considered.

Fig. 3. Boxplots of the minimum error of positive and negative cases according to the Masked DICE metric.

Regardless of the X-ray image segmentation employed (see Fig. 4), the positive case ranks in the first position at least
identification decision support system that integrates different criteria, superimpositions, and skeletal structures.

ACKNOWLEDGMENT

This work was supported by the Spanish Ministry of Science, Innovation and Universities, and European Regional Development Funds (ERDF) under grant EXASOCO (PGC2018-101216-B-I00), and by the Regional Government of Andalusia under grant EXAISFI (P18-FR-4262). Mr. Gómez’s work was supported by Spanish MECD FPU grant [grant number FPU14/02380]. Pablo Mesejo is funded by the European Commission H2020-MSCA-IF-2016 through the Skeleton-ID Marie Curie Individual Fellowship [reference 746592]. Drs. Ibáñez and Valsecchi’s work is funded by Spanish Ministry of Science, Innovation and Universities-CDTI, Neotec program 2019 [reference EXP-0012269/SENE-20191236].

REFERENCES

Fig. 5. (Left) An example of a positive case, radiograph A compared against CT A; (Right) Example of negative cases, radiograph A compared against CTs A, B and C. Red lines display the segmented X-ray AM image, blue lines represent the projected 3D PM image (i.e., the registration result, or silhouette that we try to align to the red one), green lines represent perfect overlap between red and blue lines.