Journal of Digital Imaging

Automatic Correspondence on Medical Images: A Comparative Study of Four Methods for Allocating Corresponding Points

T. L. Economopoulos,[1] P. A. Asvestas,[2] and G. K. Matsopoulos[1]

The accurate estimation of point correspondences is often required in a wide variety of medical image-processing applications. Numerous point correspondence methods have been proposed in this field, each exhibiting its own characteristics, strengths, and weaknesses. This paper presents a comprehensive comparison of four automatic methods for allocating corresponding points, namely the template-matching technique, the iterative closest points approach, the correspondence by sensitivity to movement scheme, and the self-organizing maps algorithm. Initially, the four correspondence methods are described focusing on their distinct characteristics and their parameter selection for common comparisons. The performance of the four methods is then qualitatively and quantitatively compared over a total of 132 two-dimensional image pairs divided into eight sets. The sets comprise of pairs of images obtained using controlled geometry protocols (affine and sinusoidal transforms) and pairs of images subject to unknown transformations. The four methods are statistically evaluated pairwise on all image pairs and individually in terms of specific features of merit based on the correspondence accuracy as well as the registration accuracy. After assessing these evaluation criteria for each method, it was deduced that the self-organizing maps approach outperformed in most cases the other three methods in comparison.

KEY WORDS: Point extraction, automatic point correspondence, iterative closest points, template matching, correspondence by sensitivity to movement, self-organizing maps, features of merit, registration accuracy

INTRODUCTION

Point correspondence is defined as a procedure for allocating homologous points on two or more images. This is of particular importance in medical imaging, where point correspondence methods are used in a wide variety of applications, such as the geometric alignment of two or more medical data sets[1] (X-rays, computed tomography (CT), magnetic resonance, positron emission tomography (PET), dental, etc.) obtained from the same patient (intra-subject registration) or from different subjects (inter-subject registration). Other crucial applications include mammogram screening,[2] 3D reconstruction,[3] determination of camera location,[4] motion analysis,[5] structure from motion,[6] and the determination of a subject's point of gaze.[7]

Generally, point correspondence algorithms comprise of two steps; the detection or definition of control points and the estimation of the correspondences. The first step concerns either the manual definition of control points on one image[1] or their detection by an automatic algorithm.[8] Control points are usually points of geometrical interest such as L-shaped corners or T-shaped or Y-shaped junctions. Once the control points are defined, the estimation of the corresponding points is feasible. Generally, this estimation involves the

[1]*From the School of Electrical and Computer Engineering, National Technical University of Athens, 9, Iroon Polytechniou str, Zografos, 15780 Athens, Greece.*

[2]*From the Department of Medical Instruments Technology, Faculty of Technological Applications, Technological Educational Institute of Athens, Ag. Spyridonos Street, Egaleo, 122 10 Athens, Greece.*

Correspondence to: G. K. Matsopoulos, School of Electrical and Computer Engineering, National Technical University of Athens, 9, Iroon Polytechniou str, Zografos, 15780 Athens, Greece; tel: +30-210-7722288; fax: +30-210-7723557; e-mail: gmatso@esd.ece.ntua.gr

Online publication 3 March 2009

doi: 10.1007/s10278-009-9190-z

optimization of an objective function, starting from an initial guess and eventually achieving a solution through an iterative process. The objective function measures either the similarity between image patches[8–15] or the distance between sets of points.[16–19] The wide diversity in the implementation of point correspondence methods gave birth to numerous novel techniques. Some of these techniques are derivations from simple concepts (such as pattern matching),[8,11,20,21] while others are considerably more complex,[22–25] involving statistical analysis and neural networks.

Previous comparative studies consider point correspondence methods as part of image registration schemes and compare them directly with other conventional image registration methods.[26,27] Particularly, in,[26] a survey of various feature-based and area-based registration methods is presented for two-dimensional multimodal data. On the other hand, Matabosch et al.[27] are concerned with the comparison of algorithms for aligning surfaces on three-dimensional data sets.

In this paper, a comprehensive study is presented in order to evaluate and compare the accuracy of four point correspondence algorithms. Those include the template matching (TM),[10] the iterative closest points (ICP),[16] the correspondence by sensitivity to movement (CSM),[15] and the self-organizing maps (SOMs) algorithms.[25] Most of the methods included in this study are well-established schemes with several descendant methods and variations. For instance, the template-matching algorithm is cited well over 5,000 times (Google Scholar, 2008); the iterative closest point algorithm is cited 1,975 times (Scorpus, 2008), while sensitivity to movement and the self-organizing maps are cited in 24 and ten works, respectively (Scorpus, 2008). Each of the selected techniques presents a unique approach for allocating correspondences, featuring numerous distinctive aspects. The four algorithms are described in the following sections. All methods were extensively tested over a total of 132 two-dimensional (2D) pairs of retinal and dental images. The particular image pairs were acquired from fully controlled geometry protocols (known affine and sinusoidal transformations) as well as from semi-controlled geometry protocols (unknown transformations). Eight sets of pairs were formed. The accuracy of all four algorithms was evaluated qualitatively, in terms of visual assessment, as well as quantitatively in terms of several features of merit (FOM),[28] combined with statistical evaluation. In cases where the actual correspondences were not known due to the image acquisition process, the features of merit used for quantitative evaluation were based on image registration accuracy. On the other hand, when the actual correspondences were known in advance, the features of merit employed were based on point accuracy of correspondence.

The rest of the paper is organized as follows. The "Methods" section describes the protocols used to acquire the eight data sets featured in this study. It also examines each of the four tested point correspondence algorithms and the methodology used to compare them. The "Results" section presents the results and the techniques used to obtain them and finally the "Discussion" section evaluates some of the major factors affecting the assessed point correspondence methods.

METHODS

Data Acquisition

Eight data sets of medical images were used in this study. Four sets (sets I–IV) were created from images using fully controlled geometry procedures. Those four data sets consisted of ten reference images each. The corresponding images were produced by applying local (sinusoidal) or global (affine) transformation models on the reference images. The parameters utilized to produce those sets are described in the following sections. Hence, a total of 40 controlled geometry image pairs, divided into four sets, were contributed to this study. The first two sets (sets I and II) consist of retinal images while the rest of the produced pairs (sets III and IV) consist of dental medical data. An additional 20 retinal image pairs (set V) were considered. Those pairs were employed to evaluate the performance of the four compared methods on images subject to unknown transformations.

All retinal images were acquired using the IMAGEnet 1024 system, which is a fully functional digital imaging system for acquisition, analysis, storage, and retrieval of retinal images. Digital red-free images of size 512×512 pixels and pixel size of 10 μm were directly obtained using a charge-coupled camera that was mounted on the Topcon TRC-50IX, providing 50° angle of coverage, 39-mm working distance, and a green filter, which causes the retinal blood vessels to appear dark.

Furthermore, 72 dental image pairs, acquired under semi-controlled geometry, were included in the study through the following protocol: a dry mandible was mounted on a device which permitted the object and the film to be rotated vertically and horizontally relative to the central part of the X-ray beam. The reference radiograph was taken with the central ray of the X-ray beam perpendicular to the long axes of the teeth as judged subjectively and with no resulting overlaps of adjacent tooth surfaces. Corresponding images were then obtained by moving the object either vertically or horizontally relative to the X-ray beam at angles of 0°, 3°, and 6°. This corresponds to motion about the *x*- and *y*-axis in three-dimensional space. In two dimensions, those movements could be approximated by affine transformations with unknown parameters. Three dry mandibles were used in order to produce a total of 72 *in vitro* dental pairs grouped into three further sets of 24 image pairs each (sets VI–VIII). All dental radiographs were digitized with a flat scanner (Agfa Arcus II) producing 8-bit grayscale image files. The focus of the object and the object-to-film distance were kept constant at 40 cm and 0.5 cm, respectively. The size of the *in vitro* radiographs used in the study was 428×310 pixels. A summary of all eight data sets used in this study is shown in Table 1.

Point Correspondence

Point correspondence algorithms can be broadly divided in two categories: those that require the extraction of control points from both images (reference and corresponding) and those that require the extraction of control points from one image only (reference). The main representative of the first category is the ICP algorithm, where a least-squares approach is used iteratively in order to find the parameters that best describe the transformation between the two point sets. On the other hand, the algorithms of the second category use either block-matching or feature-matching approaches in order to find pairs of homologous points. This section provides a description of each method in comparison, focusing on its distinct characteristics.

Control Points Extraction

The first step in any point correspondence approach is to allocate and extract an initial set of control points from the reference image. There are numerous methods for extracting points of interest from medical images.[8,29–30] As the efficiency of those methods varies with the nature of the subject images, two techniques were employed throughout this study to obtain the initial control points:

(a) The method described in[29] which is optimized for detecting bifurcations in retinal images
(b) The general purpose method adopted by Likar and Pernus[8] for obtaining points around edges and ridges on dental images, which is suitable for working with a large number of points

In the first case, the number of extracted bifurcation points is determined automatically by the algorithm, while for the dental images, a fixed number of 200 control points were extracted from each reference image.

Table 1. Overview of the Data Sets Used in the Study

Set	Number of image pairs	Modality	Acquisition	Transformation
Set I	10	Retinal		Affine
Set II	10	Retinal		Sinusoidal
Set III	10	Dental		Affine
Set IV	10	Dental	Controlled geometry (reference transformed)	Sinusoidal
Set V	20	Retinal	Semi-controlled geometry	Unknown
Set VI	24			
Set VII	24			
Set VIII	24	Dental	Semi-controlled geometry	Unknown

Automatic Correspondence Methods in Comparison

As mentioned, this study examines four well-established point correspondence algorithms; namely the template matching, the iterative closest point, the correspondence by sensitivity to movement, and the self-organizing maps algorithms. In the following sections, an overview of the four algorithms is provided, emphasizing on their individual properties.

Template-Matching Method

TM is one of the most commonly used methods for allocating correspondences.[10] The reason for this is that the technique itself is rather flexible and quite simple to implement. Numerous variations of the method have been devised, some of which favor accuracy, while others target at improving execution time.[12–14] Whatever the case, the basic principle of all template-matching schemes is the same: a sub-image (template) from the reference image is compared with a pool of possibly transformed sub-images from the corresponding image. In particular, a sub-image of fixed size (template) is initially generated from a selected area of the reference image I_R. In our case, the initial control points are considered to be the central pixels of each sub-image. Then, for each control point, the template propagates along each pixel of the corresponding image I_C. The similarity between the template and the sub-image of the corresponding image underneath it is usually assessed through a measure of match. The pixel of the corresponding image, for which the measure of match obtains its optimum value, is considered to be the corresponding point of the control point defining the center of the current template. The generic template-matching scheme employed for the purposes of this study utilizes a correlation function as a measure of match. In addition, the template is compared with transformed sub-images of the corresponding image using the similarity transformation, which is modeled by four parameters for rotation (one parameter), scaling (one parameter), and translation along the *x*- and *y*-axis (two parameters, one for each axis). Finally, the size of the templates was fixed to 21×21 pixels.

The typical template-matching approach for allocating point correspondences is a rather simple technique to implement. Moreover, the algorithm incorporates exhaustive search tactics, which render it quite accurate. On the other hand, exhaustive searches are somewhat inefficient when it comes to time-critical applications. As a result, all template-matching methods usually require a considerable amount of time to converge.

Iterative Closest Point Algorithm

The ICP algorithm[16] is broadly used for allocating corresponding points due to its robustness, simplicity, and fast execution time. The algorithm itself is a flexible framework which can be customized according to the application needs. As a result, numerous variants of the ICP have been devised over the years.[19] The key property of ICP is that it requires the extraction of control points from both images. Then, an iterative procedure is applied: for each control point of the reference image, the closest control point from the corresponding image is obtained. Then, using a standard least-squares approach, the parameters of the desired transformation are calculated and the control points from the corresponding image are transformed using the specific parameters. The iterations are terminated when a suitable criterion has been satisfied (e.g., the number of iterations has reached a maximum value, the mean square error of the least-squares approach is below a threshold, etc.). In our case, an elastic transformation based on the thin-plate spline (TPS) deformation was adopted to transform the candidate corresponding points.[31] The TPS transform comprises of a global affine component and a local elastic component, thus providing ICP with additional flexibility to cope with elastic deformations on the subject images. Moreover, the average square distance between the point sets was considered as a measure of match for the ICP implementation used throughout this study. The ICP algorithm eventually converges to an optimal set of corresponding points, provided that proper points of interest were selected on both the reference and the corresponding image.

The basic idea of ICP is to calculate correspondences between two or more point clouds. In our case, the first point cloud is considered to be the set of initial points on the reference image, while the second one is a set of candidate points on the corresponding image. All ICP variants adopt a trial-and-error approach to allocate correspondences. As a result, the accuracy of the estimated corresponding points should improve after each iteration (trial). In general, the more iterations are completed, the more accurate are the results obtained. In the present study, the maximum number of iterations was set to 500 in order to maximize the efficiency of the algorithm.

Correspondence by Sensitivity to Movement Algorithm

The CSM algorithm is an advanced variant of the template-matching[10] scheme for allocating point correspondences. Template-matching-based techniques estimate corresponding points by matching a region from the reference image to a region from the corresponding image (one-to-one match) and hence a unique correspondence is calculated per control point. On the other hand, CSM considers several candidate corresponding points for each control point, using a weighting scheme.[15]

The algorithm comprises of three distinct stages. Firstly, a number of candidate corresponding points are allocated for each control point from the reference image. These points are bound to a specific region on the corresponding image. A similarity measure is then calculated for each candidate point. In our case, a simple correlation scheme was employed. The candidate points together with their similarity measures form the match map of the current control point. From the match map of the control point, a "tentative" point is calculated as follows:

$$\tilde{q}_i = \frac{\sum_{i=1}^{M} K_i \overrightarrow{v_i}}{\sum_{i=1}^{M} K_i} \tag{1}$$

where $\overrightarrow{v_i}$ is the vector from the control point to point i in the match map. K_i is a constant defined by $K_i = \frac{MoM_i}{1+\left|\overrightarrow{v_i}\right|^2}$, where MoM_i is the similarity value (measure of match) for point i. The sums involved in Eq. 1 are over the entire match map of M elements

During the second stage of the algorithm, the control points on the reference image are slightly displaced towards all eight possible neighboring positions and a new match map is estimated in each case. Each new tentative point is calculated as before, while taking distances from the control point to each point in the match map into account. Consequently, a cloud of tentative corresponding points is estimated for each control point. The final stage of CSM involves analyzing the distribution of those tentative points for each control point. If the tentative points are scattered along a line, the point closest to the line is considered as the corresponding point; otherwise, the centroid of the scatter is selected.

Self-Organizing Maps Algorithm

The SOMs is a neural network algorithm, which is able to train itself in an unsupervised manner, through an iterative process.[17] The SOMs model was introduced by Kohonen[25] and comprises of a layer of m neurons arranged in a one-dimensional or two-dimensional grid. In our case, each control point P_j from the reference image is considered as a neuron with weight vector w_j, which holds the parameters of a local transformation. The local transformation preferred in this study is a similarity transformation. The network is firstly initialized by setting the weight vector w_j of every neuron j to the parameters of the identity similarity transformation. After the network is properly initialized, the training iterations begin. At each iteration n, a random signal $s(n)$ is generated and presented to the network. The random signal generator is similar to the one used in simulated annealing.[17] The components of the random signal are random numbers within a predefined range and correspond to the parameters of a local similarity transformation. The generated random signal is then applied to all network neurons. The neuron j with weight vector w_j is declared as the winning neuron, according to the rule below:

$$j = \arg\max_i \left\{ MoM\left(\mu_{Ai}(I_R), \mu_{T_{s(n)}(A_i)}(I_C) \right) \right\} \tag{2}$$

where A_i is a square region centered around P_i; $T_{s(n)}(A_i)$ is a similarity transformation of the region A_i with parameter vector $s(n)$; $\mu_{Ai}(I)$ denotes the restriction of an image I to the region A_i and MoM is the preferred measure of match, which in our case is the correlation coefficient.

This practically means that the wining neuron is the one that achieves the maximum measure of match subsequent to the application of signal $s(n)$.

After the winning neuron j is found, the neuron itself as well as its neighboring neurons i have their weight vectors modified according to the following equation:

$$w_i(n+1) = w_i(n) + h_{ij}(n)[s(n) - w_i(n)] \quad (3)$$

where h_{ij} is a Gaussian-type function which depends on the distance between the winning neuron and its neighboring neurons and n. At this point, a single iteration is complete; hence, a new random signal is presented to the network and the process described above is repeated. The training of the network terminates after a fixed number of iterations is reached (for example, n= 5,000). After the training is completed, the weight vectors of the neurons contain the optimal local transformation parameters. The control points obtained from the reference image are then transformed according to those parameters in order to produce their estimated corresponding points.

Comparison Details

The performance of each featured point correspondence algorithm depends on a number of parameters. The optimal parameters were selected after a series of trials on several image pairs from all available sets, in order to enhance the validity of the study. Those parameters are summarized in Table 2. As can be seen in Table 2, identical parameters were used for common properties of the four methods in order to conduct a meaningful and fair comparison. The major parameters affecting each scheme are further examined in the "Discussion" section.

Evaluation Methods

The four point correspondence methods were compared pairwise using several FOMs and by conducting statistical analysis.[28, 32] As shown in Table 3, three quantitative features of merit were used throughout this study: one when the actual corresponding points were known (FOM^1) and two when no prior knowledge of actual correspondences was available (FOM^2, FOM^3). The preferred FOMs were based on the root mean squared distance (RMSD), the mean edge distance, and the inverse mutual information metrics (FOM^1, FOM^2, and FOM^3, respectively). All FOMs are described in some further detail later. The first step in the evaluation process was to calculate all features of merit for all image pairs. Then, for each set, FOM and pair of methods, a statistical comparison was performed by means of Student's paired t test. The null hypothesis was that the two compared methods did not differ significantly in terms of the particular feature of merit used. A 95% confidence level was considered to assess the null hypothesis in our case. In order to examine the degree of variation between two methods, the statistical relevance metric[28] was employed:

$$r^i_{m1,m2}(S) = \left[1 - \frac{FOM^i_{m2}(S)}{FOM^i_{m1}(S)}\right] \times 100 \quad (4)$$

where $r^i_{m1,m2}(S)$ defines the statistical relevance of the improvement of method $m2$ over method $m1$ for an image pair S, using the feature of merit $i(\mathrm{FOM}^i(S))$. This scheme assumes that method $m2$ is strictly better than method $m1$; in other words, $m2$ exhibits a lower feature of merit than $m1$ for the particular image pair or simply $\mathrm{FOM}^i_{m2}(S) < \mathrm{FOM}^i_{m1}(S)$.

After all individual statistical relevance metrics have been calculated, an average statistical rele-

Table 2. Parameters Used for the Qualitative and Quantitative Evaluation of the TM, ICP, CSM, and SOMs Algorithms

Parameter	TM	ICP	CSM	SOMs
Measure of match	Correlation coefficient	Mean square distance	Correlation coefficient similarity	
Transformation	Similarity	Thin-plate splines		
Search optimization	Simulated annealing	–	Simulated annealing	–
Maximum displacement (pixels)	80	–	80	
Maximum rotation (°)	25	-	25	
Maximum scaling	10%	–	10%	
Stopping number of iterations	–	500	–	5,000
Template size (pixels)	21 × 21	–	–	–
Match-map size (pixels)	–	–	15 × 15	–

Table 3. The Features of Merit (FOMs) Used to Evaluate TM, ICP, CSM, and SOMs for Images Subject to Both Known and Unknown Transformations

FOMi	Preferred Metric
Root mean squared distance ($i = 1$)	$\mathrm{FOM}^1 = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|Q_i - \tilde{Q}_i\right\|^2}$
Mean edge distance ($i = 2$)	$\mathrm{FOM}^2 = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|E_i - \tilde{E}_i\right\|}$
Inverse mutual information ($i = 3$)	$\mathrm{FOM}^3 = \frac{1}{H(I_R)+H(I_{GTR})-H(I_R,I_{GTR})}$

vance was estimated for all images in a set between two methods. The average statistical relevance was recorded only for the cases where the null hypothesis was rejected, namely the two methods differed significantly as per the specific FOM. Using this scheme, all four examined point correspondence methods were compared pairwise.

RESULTS

As mentioned, eight data sets were used to qualitatively and quantitatively evaluate the four point correspondence methods in comparison. The first four sets comprise of images acquired through controlled geometry procedures while the remaining four sets were obtained using semi-controlled geometry protocols. The performance of the four point correspondence methods is evaluated separately for the two aforementioned cases.

Controlled Geometry Data

Four data sets were acquired by transforming the reference images using known transformations. Each reference image was transformed using either a known global transformation (*affine*) or a known local transformation (*sinusoidal*). Thus, each point (x, y) on the reference images was transformed to its new coordinates (x', y') in order to form a total of 40 corresponding images. The global transformation, which was employed for sets I and III, was an affine one:

$$\begin{aligned} x' &= s\,\cos\theta(x - x_c) - s\,\sin\theta(y - y_c) + t_x + x_c \\ y' &= s\,\sin\theta(x - x_c) + s\,\cos\theta(y - y_c) + t_Y + y_c \end{aligned} \tag{5}$$

where s is the scaling factor; θ is the rotation angle and t_x and t_y are the displacements along the horizontal and the vertical axis, respectively. The coordinates (x_c, y_c) define the central point of the reference image. In our case: s=1.02, θ=10°, t_x=5 and t_y=−5.

The local deformation, applied in sets II and IV, was a sinusoidal one:

$$\begin{aligned} x' &= x + a\,\sin\left(\tfrac{y}{T}\right) \\ y' &= y + a\,\cos\left(\tfrac{x}{T}\right) \end{aligned} \tag{6}$$

where a=8 and T=128.0. The two transformations defined in Eqs. 5 and 6 are demonstrated on a retinal image (pair 6) from set I in Figure 1. Figure 1a presents the reference image itself, while Figure 1b, c shows the two corresponding images as these were produced after applying the *affine* and the *sinusoidal* transforms, respectively. As can be seen in Figure 1, the local and global transformations used in this study have been selected to represent realistic medical data. Hence, only relatively small rotations, translations, and deformations were considered.

As mentioned, the adopted comparison procedure assumes a predefined set of control points on the reference image in order to estimate their correspondences. In case of sets I–IV, those points were allocated using two automatic point extraction algorithms: one for extracting 200 points of interest from dental images[8] (sets III and IV) and a separate one for allocating bifurcation points on the retinal data[29] (sets I and II). In the latter case, the number of extracted points depended on the morphology of the retinal images (number of bifurcations). Once the control points have been defined for each reference image, their actual corresponding points are calculated using Eq. 5 for global and Eq. 6 for local transformations, respectively.

Qualitative evaluation was performed by means of visual assessment. In this case, the actual and the estimated points are superimposed on the corresponding images. It follows that the better the estimated points match with their actual counterparts, the higher is the accuracy of the method. Typical examples are presented in Figure 2 up to Figure 4, which demonstrate retinal (set I) and dental (set III) image pairs, respectively. In all cases, actual corresponding points are presented as solid circles, while estimated points are defined by transparent circles. In particular, as can be seen in

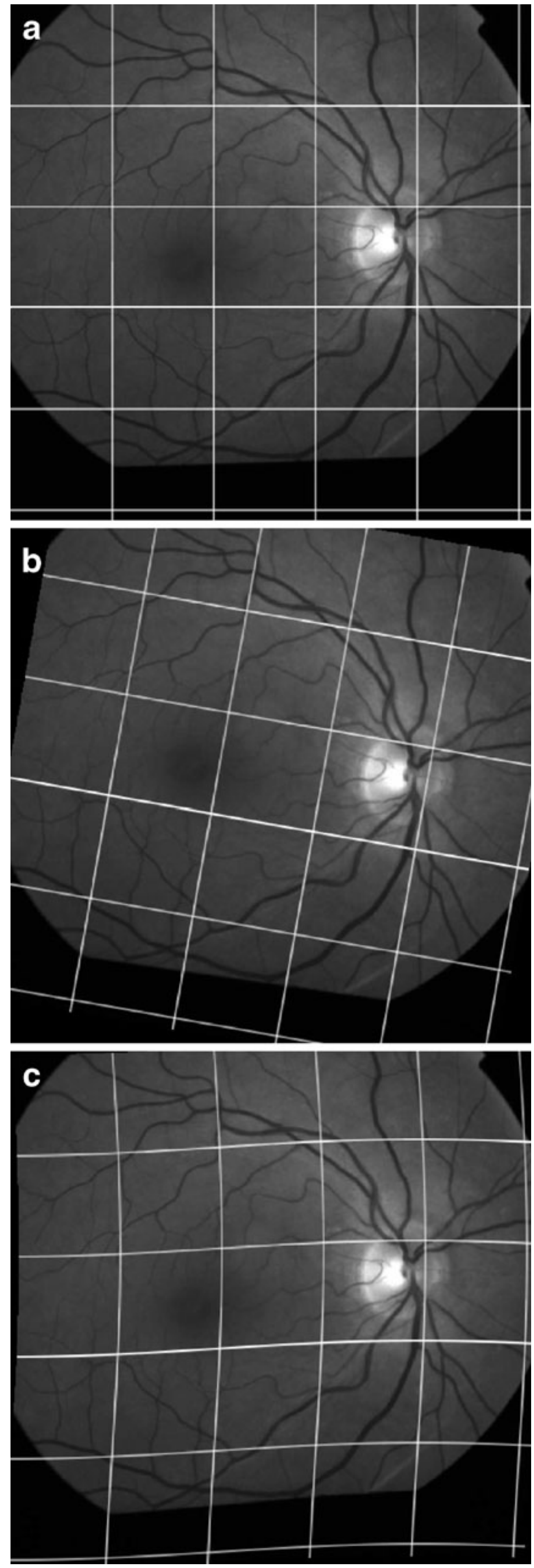


Fig 1. **Generation of test images using a controlled geometry protocol. a Reference retinal image from set I (pair 6); b *Affine* transformation. c *Sinusoidal* deformation. A *grid* is superimposed on all images to show the effect of the transformation.**

Figure 2, the SOMs method outperforms on average all other methods by achieving superior match between the estimated and actual corresponding points. This is more clearly visible when examining a zoomed section of Figure 2, shown in Figure 3. From the example illustrated on those two figures, it may be concluded that SOMs (Figs. 2e and 3d) clearly outperform both ICP (Figs. 2c and 3b) and CSM (Figs. 2d and 3c). On the other hand, TM (Figs. 2b and 3a) follows SOMs closely for the particular example, but a more detailed examination in Figure 3 reveals that it is marginally less accurate than SOMs. The SOMs algorithm also achieves superior point correspondence accuracy compared to the other three methods in comparison for the dental image set. A typical example is depicted in Figure 4, using 200 control points. As can be seen there, SOMs (Fig. 4e) achieve a better match compared to TM (Fig. 4b), ICP (Fig. 4c), and CSM (Fig. 4d).

In this case, where images subject to known transformations are considered, the actual correspondences can be trivially calculated. Therefore, in order to quantitatively assess the four methods, the corresponding points obtained using each technique were compared against the actual correspondences using the RMSD between the estimated and the actual corresponding points.[33] The RMSD is defined as follows:

$$FOM^1 = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|Q_i - \tilde{Q}_i\right\|^2} \qquad (7)$$

where Q_i and $\tilde{Q}_i(i = 1, 2, \ldots, N)$ are the actual and the estimated corresponding points, respectively. Obviously, the lower is the FOM^1 measurement, the better is the correspondence obtained.

Using the feature of merit described in Eq. 7, the four methods were compared pairwise for all images in sets I–IV. The results concerning all sets subject to known transformations are shown in Table 4 in terms of the average statistical relevance metric, using FOM^1. In Table 4, a "+" sign next to a measurement hints that the first method is systematically better than the second one, while a "–" sign suggests the opposite. A "*" symbol denotes that the two compared methods do not exhibit significant statistical difference. It follows that the larger are the values shown in Table 4, the greater is the difference between the two compared methods. As can be seen in Table 4, SOMs outperform in most cases TM, ICP, and CSM for the particular feature of merit. The SOMs approach is especially advantageous over the dental images (sets II and IV), where it systematically outperforms all other three methods. On the other hand, CSM outperforms both TM and ICP in the cases where retinal images are involved (sets I–II). Finally, TM and ICP perform equally on average, leading to inconclusive statistical analysis in most cases, which is denoted by a "*" in Table 4.

The four methods were also quantitatively assessed independently through their average FOM^1 values. The results are shown in Table 5, where all image pairs from sets I–IV are examined. As can be seen in Table 5, SOMs achieve lower average FOM^1 values than all other three methods in comparison. Moreover, SOMs exhibit a very low standard deviation of the FOM^1 metric in most cases, which indicates that the particular method achieves uniformly distributed performance.

Semi-controlled Geometry Data

The remaining four data sets considered in this paper were subjected to unknown transformations. As mentioned previously, 20 retinal image pairs were acquired through a semi-controlled geometry protocol (set V). In addition, three more sets (sets VI–VIII) of dental images, consisting of 24 pairs each, were obtained from a separate *in vitro* study. As a result, a total of 92 image pairs were available in order to test the four methods in comparison on images with unknown transformations.

For set V of retinal images, the ideal corresponding points were known beforehand (bifurcation points). In this case, points were extracted in both images using the methodology described in.[29] Therefore, the methods illustrated before can be used to assess the four point correspondence methods both qualitatively, through visual assessment, and quantitatively, using the feature of merit quoted in Eq. 7. In Table 6, all four methods are compared pairwise using the statistical methodology described previously. As can be seen in Table 6, SOMs clearly outperform both TM and CSM on average, followed by ICP. TM is the worst-performing method for the particular set of retinal images. Moreover, all 20 image pairs of set V are examined individually, in terms of their recorded FOM^1 metric, in Table 7. The conclusions drawn previously are confirmed there, as SOMs achieve a lower average FOM^1 measurement, followed by ICP, CSM, and finally TM.

For the remaining three sets of images obtained through a semi-controlled geometry protocol (sets VI–VIII), the actual corresponding points are unknown. Therefore, in this case, the four methods were evaluated through an image-registration-oriented approach for both qualitative and quantitative assessment. The initial points of interest were first extracted using the algorithm proposed in.[8] The four methods (TM, ICP, CSM, and SOMs) were then applied to each image pair from sets VI–VIII. The estimated corresponding points produced by each method were subsequently used to calculate an affine transformation to register the corresponding images, through the least-squares method in conjunction with singular value decomposition.[34] The preferred affine transformation is shown below.

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & dx \\ a_3 & a_4 & dy \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \qquad (8)$$

where a_1, a_2, a_3, a_4, dx, and dy define the parameters of the affine transformation. The registered images were re-sampled using the bilinear interpolation method.[35] Finally, the performance of the four methods was evaluated by means of the registration quality of the produced registered corresponding images.

The registration quality of the four methods in comparison was visually assessed by superimposing the edges of the reference image on the aligned corresponding image. The edges were detected by applying the Canny edge detector.[36] An example is shown in Figure 5, featuring an image from set VII. As can be seen in Figure 5, the registered image produced using the corresponding points estimated through the SOMs method (Fig. 5h) is more accurate than the other three methods in comparison (Fig. 5b, d, f). The outline of the edges in Figure 5h fits the aligned corresponding image more accurately than the images obtained using TM (Fig. 5b), ICP (Fig. 5d), and CSM (Fig. 5f). This may also be verified by calculating the absolute difference of the reference image with respect to its registered counterpart. This case is also illustrated in Figure 5.

The images shown in Figure 5c, e, g, i were obtained by calculating the absolute difference between the intensity of the reference image and the intensity of the transformed corresponding image, on a pixel-by-pixel basis. In this case, the darker the resulting difference is, the better is the registration quality. Yet again, SOMs (Fig. 5i) perform better than all the other three methods in comparison, in terms of registration quality. However, in Figure 5, it can be seen that although TM (Fig. 5b, c) and CSM (Fig. 5f, g) perform generally worse than SOMs, they clearly outperform ICP (Fig. 5d, e) for the particular example.

In order to quantitatively evaluate the examined point correspondence techniques on sets VI–VIII, alternative features of merit had to be defined, as the actual correspondences are unknown for the particular data sets. As mentioned, the FOMs in this case were based on the registration quality of the registered images which were produced using the estimated corresponding points from each method. As a result, the preferred primary FOM for evaluating dental images subject to unknown transformations is the mean edge distance between the reference and the transformed corresponding image:

$$FOM^2 = \frac{1}{N}\sum_{i=1}^{N}\left\|E_i - \tilde{E}_i\right\| \quad (9)$$

where $E_i(i=1,2,\ldots,N)$ are all points along the edges of the reference image I_R and $\tilde{E}_i$ are their closest respective points from the edges of the transformed image I_{GTR}. The edges of both images were extracted through a Canny edge detector.[36] Generally, a small value of the edge distance metric corresponds to a close match between the reference and the transformed corresponding image. This, in turn, suggests that the smaller is the value of the edge distance obtained, the higher is the quality of the intensity-based image registration and thus the more accurate are the correspondences obtained. In addition to Eq. 9, a secondary FOM was used in this case, defined by the inverse mutual information between the reference and the registered image. The particular metric is shown below:

$$FOM^3 = \frac{1}{H(I_R) + H(I_{GTR}) - H(I_R, I_{GTR})} \quad (10)$$

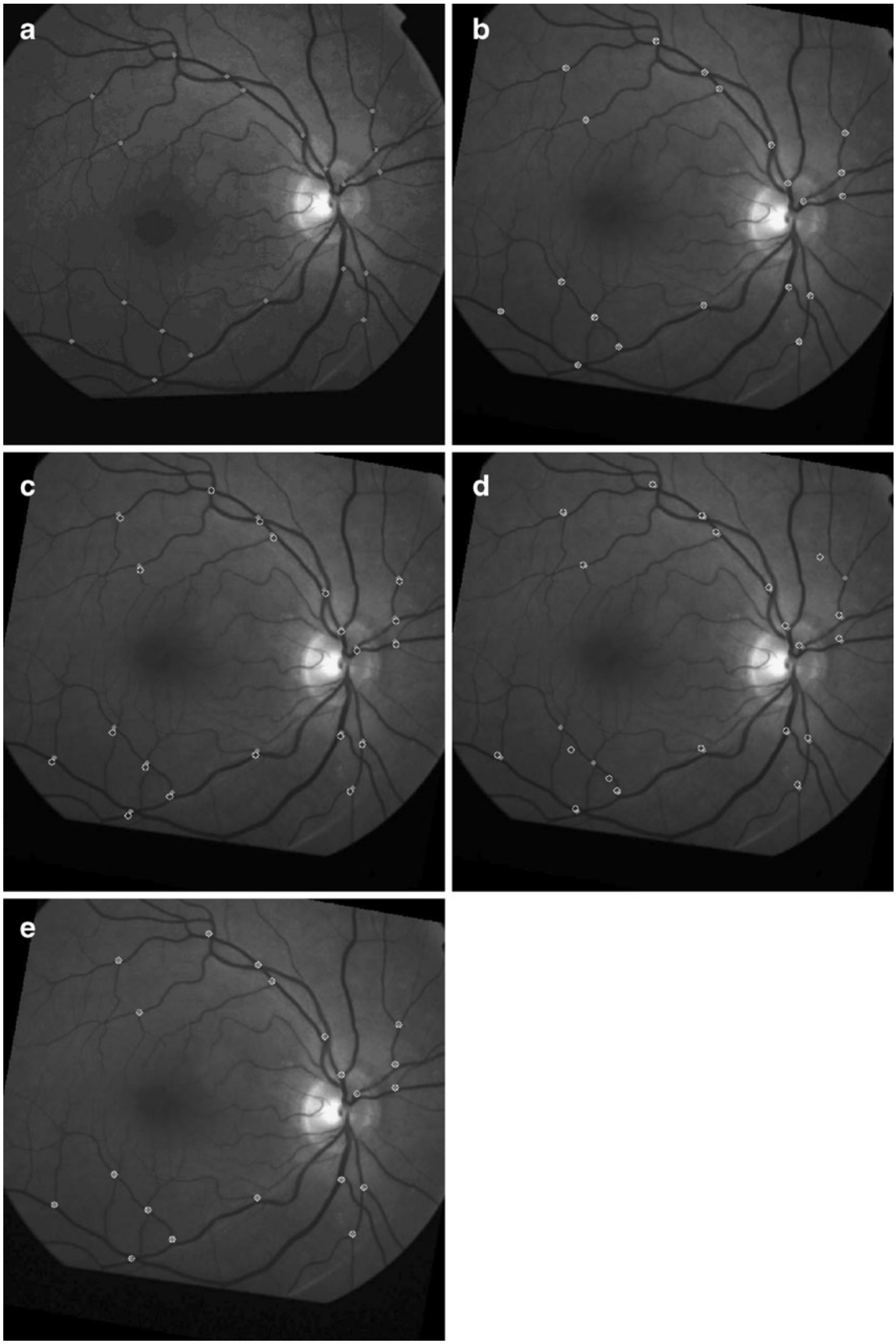
where $H(I_{\mathrm{R}})$ and $H(I_{\mathrm{GTR}})$ are the marginal entropies of the reference and the transformed image, respectively, while $H(I_{\mathrm{R}},I_{\mathrm{GTR}})$ represents their joint

Fig 2. Visual assessment of the four point correspondence methods, in terms of point accuracy on a retinal image pair (pair 6) from set I. a Reference image with the initial points of interest superimposed. Corresponding image subject to affine transformation with both the actual (*solid circles*) and the estimated (*transparent circles*) corresponding points superimposed for b TM, c ICP, d CSM, and e SOMs.

entropy. In this case, a high mutual information measurement (and hence a low value in Eq. 10) represents a successful match between the reference and the registered corresponding image. The two particular FOMs were preferred as they are both able to cope with any contrast differences between the reference and the corresponding images caused by the acquisition process of sets VI–VIII.

Having suitably defined the features of merit to be used with the semi-controlled geometry data sets, the four point correspondence techniques were compared pairwise according to the statistical evaluation scheme described previously. Thus, in spite the fact that different FOMs were used for images subject to known and unknown transformations, all comparisons were performed on a common platform, namely the statistical relevance metric.[28] The results obtained for sets VI–VIII, in terms of the average statistical relevance, are shown in Table 8 for both FOMs. Yet again, a "+" sign next to a measurement hints that the first method is systematically better than the second one; a "–" sign suggests the opposite while the "*" symbol denotes that the two compared methods do not exhibit significant statistical difference. By examining this table, it may be concluded that the SOMs approach is the preferred method for allocating correspondences in the particular dental images. The reason for this is that SOMs present significant performance difference compared to the TM, ICP, and CSM for the vast majority of cases (especially for sets VII and VIII). This is reflected by the positive values shown in Table 8 concerning SOMs, when compared to any of the three other methods. On the other hand, the worst-performing technique for the dental images obtained through a semi-controlled geometry protocol seems to be CSM, as both TM and ICP outperform the particular approach.

The conclusions drawn previously may be also confirmed by examining the average mutual information and its standard deviation, for all dental sets subject to unknown transformations. Those measurements are illustrated


a
b
c
d
e

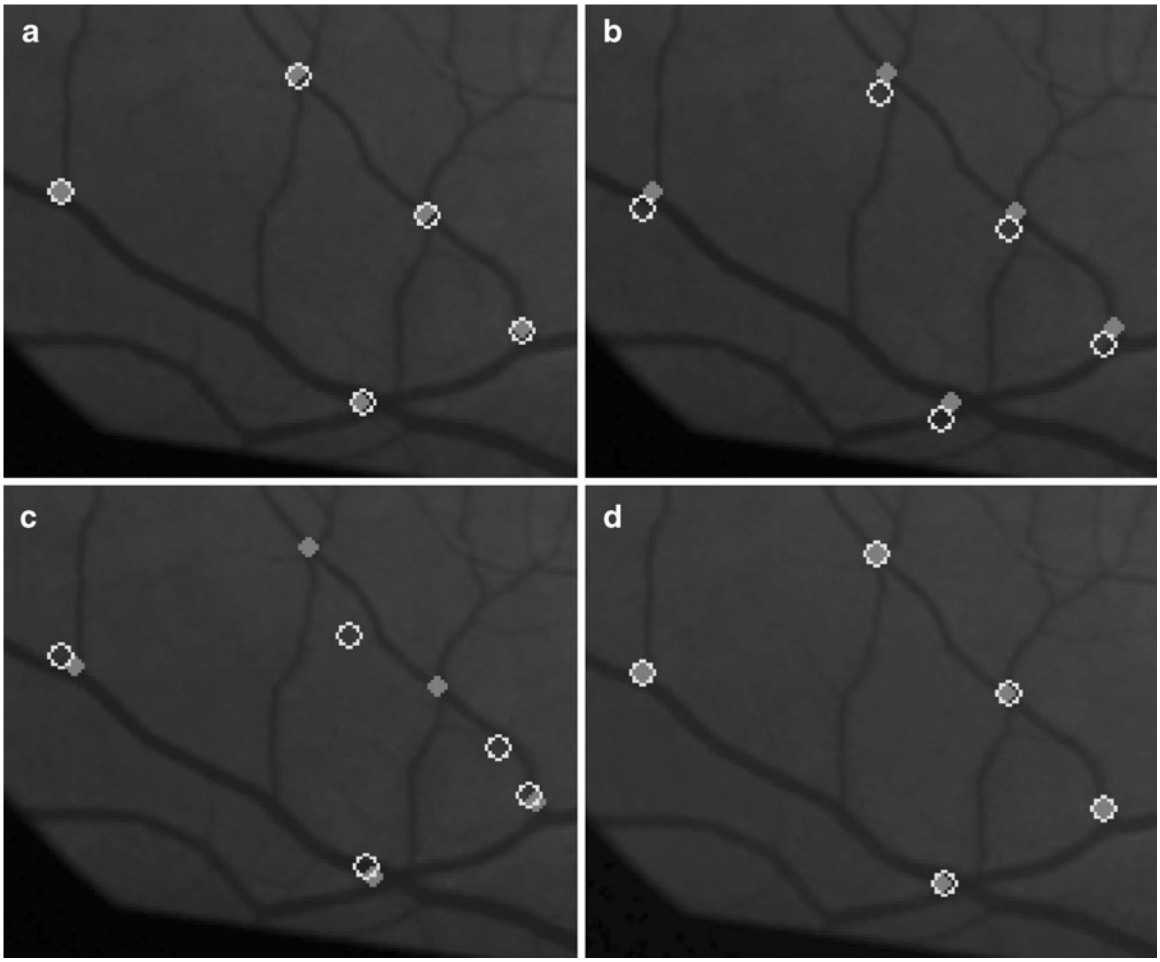


Fig 3. Zoomed sections of the images shown in Fig. 2. Proportional zoomed sections of the corresponding image with both the actual (*solid circles*) and the estimated (*transparent circles*) corresponding points superimposed for a TM, b ICP, c CSM, and d SOMs.

in Table 9. There, SOMs exhibits a higher mean mutual information measurement over TM, ICP, and CSM for sets VII–VIII. In addition, the much lower standard deviation measurements obtained for SOMs in most cases suggest that the technique is quite robust as it exhibits consistent performance.

The quantitative measurements obtained for sets VI, VII, and VIII are not directly related to point correspondence accuracy. Instead, the estimated correspondences were utilized in order to produce suitable registered images, through the methodology described earlier. Consequently, it was the registration quality of those images assessed rather than the quality of the corresponding points themselves. However, the idea is that the quality of the estimated correspondences will be reflected on the produced registered corresponding images. The reason for this is that if a poor set of corresponding points is obtained, inconsistent transformation parameters will be calculated, which in turn will produce a poorly registered corresponding image. Likewise, accurate corresponding points will result in the computation of accurate transformation parameters, which will ultimately generate a perfectly registered corresponding image, exhibiting a low FOM measurement.

Fig 4. Visual assessment of the four point correspondence methods, in terms of point accuracy for a dental image pair (pair 8) from set III. a Reference image with the initial points of interest superimposed. Corresponding image with both the actual (*solid dark circles*) and the estimated (*transparent bright circles*) corresponding points superimposed for b TM, c ICP, d CSM, and e SOMs.

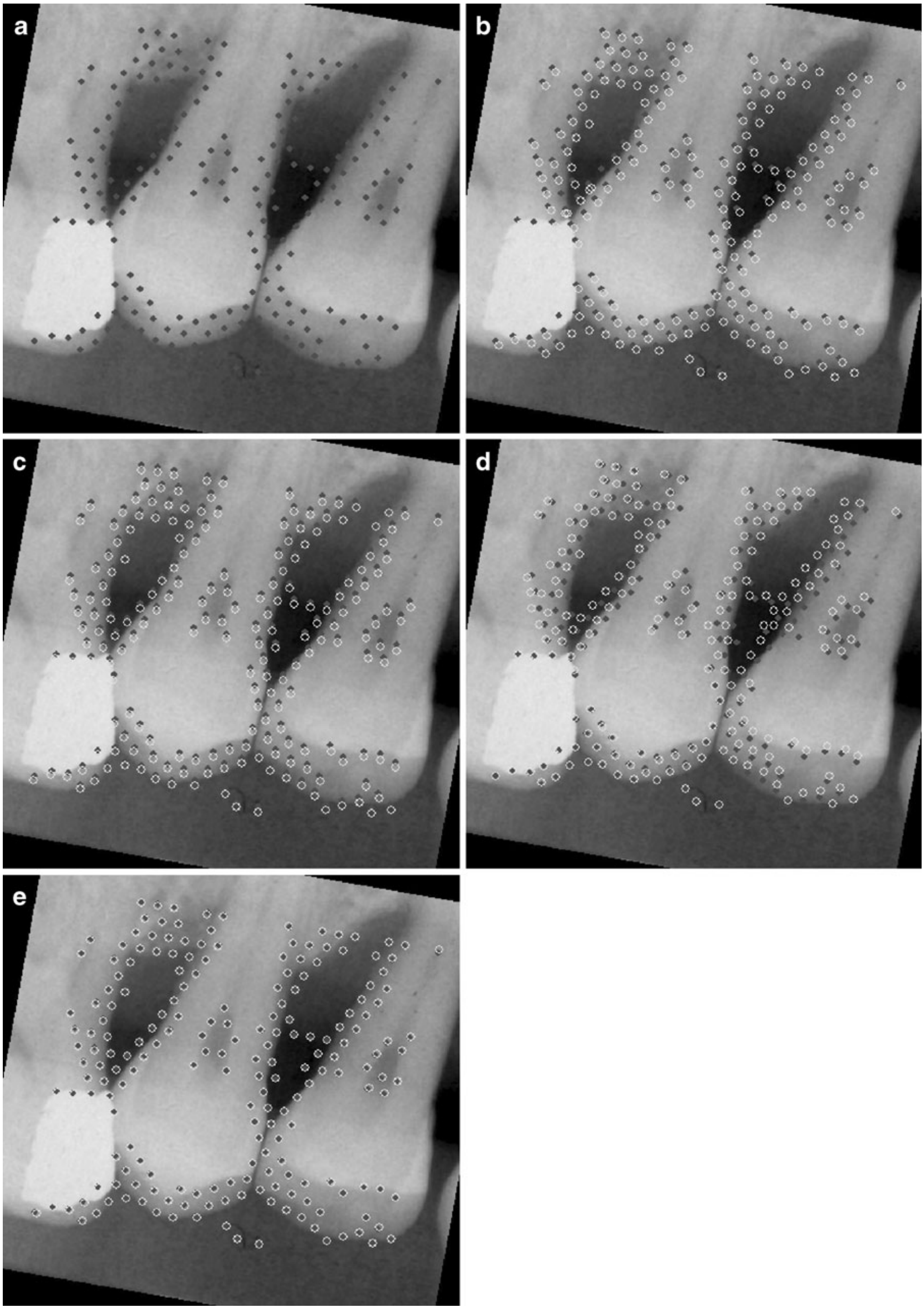

a
b
c
d
e

Table 4. Average Statistical Relevance of the Difference in Performance of the Four Automatic Point Correspondence Methods (Compared Pairwise) for All Retinal and Dental Data Subject to Known Global and Local Transformations (Sets I–IV), Using FOM[1]

Method pair	Set I (retinal–affine)	Set II (retinal–sinusoid)	Set III (dental–affine)	Set IV (dental–sinusoid)
SOMs/TM	37.183+	30.140+	54.710+	56.840+
SOMs/ICP	58.889+	51.075+	60.214+	55.894+
SOMs/CSM	*	*	57.900+	60.789+
TM/ICP	*	*	*	*
TM/CSM	23.372−	26.843−	*	*
ICP/CSM	46.358−	44.588−	*	10.283+

Noisy Data

The four point correspondence methods studied in this paper were also evaluated in the presence of noise. A retinal image pair from set I (pair 6) was used for the purposes of this test. One of the two images of the pair (the corresponding image) was corrupted as follows: firstly, Gaussian noise has been added with standard deviations of 0.001, 0.002, 0.005, and 0.01, creating five noisy corresponding images. Secondly, Gaussian blurring has been added to the same image with radii equal to 2, 4, 6, and 8 pixels, creating four more distorted images. The particular types of noise were employed to simulate two common occasions, occurring during the radiographic capture process. Gaussian noise is used in order to replicate electronic interference, while Gaussian blurring imitates focal abnormalities. The level of noise added to distort the corresponding image may be quantified through the peak signal to noise ratio (PSNR).[37] The PSNR of all noisy images was calculated using the following equations:

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_1^2}{MSE}\right) \tag{11}$$

Table 5. Performance of the Four Automatic Correspondence Methods in Terms of Point Accuracy Using FOM[1] for Images Subject to Known Transformations (Sets I–IV)

		Affine				Sinusoidal			
Pairs		TM	ICP	CSM	SOMs	TM	ICP	CSM	SOMs
	1	0.953	6.483	1.903	1.114	0.842	4.165	2.144	3.404
	2	5.399	6.551	3.891	2.032	0.971	6.291	2.944	2.552
	3	7.648	5.093	1.402	2.002	8.498	13.566	1.913	1.802
	4	6.887	4.816	1.697	1.559	7.501	4.791	2.112	2.248
	5	5.441	5.417	1.901	1.716	6.910	5.029	2.126	2.992
	6	7.774	4.896	1.495	2.062	9.408	4.992	1.786	2.594
	7	7.156	5.947	4.313	1.245	10.490	5.355	4.387	2.149
	8	11.473	5.544	7.397	5.205	13.798	6.121	7.725	2.718
	9	1.014	3.835	2.399	2.340	0.994	4.793	2.936	1.992
	10	1.030	4.931	2.916	2.051	6.491	4.603	2.562	3.185
	Mean	5.478	5.351	2.931	2.133	6.590	5.971	3.064	2.564
Retinal images	STD	3.511	0.827	1.859	1.147	4.417	2.748	1.806	0.525
	1	3.912	4.100	4.057	1.243	4.504	4.464	5.693	2.393
	2	2.941	4.535	7.276	1.628	3.112	4.471	4.992	1.550
	3	5.591	4.905	5.193	1.207	5.881	4.415	5.845	1.989
	4	5.489	4.686	4.783	1.476	6.248	4.791	5.731	2.407
	5	5.147	4.746	3.002	2.625	6.802	4.995	5.159	1.863
	6	4.362	4.458	4.757	1.026	4.709	4.985	4.851	2.104
	7	4.714	4.659	4.496	1.750	5.909	4.812	5.550	2.167
	8	2.101	4.404	3.901	2.406	2.224	4.472	4.479	1.421
	9	4.813	4.200	4.098	3.442	6.298	4.779	4.902	2.267
	10	4.891	4.548	5.175	1.049	5.076	4.505	5.208	2.422
	Mean	4.396	4.524	4.674	1.785	5.077	4.669	5.241	2.058
Dental images	STD	1.122	0.246	1.128	0.796	1.477	0.227	0.450	0.354

Table 6. Average Statistical Relevance of the Difference in Performance of the Automatic Point Correspondence Methods for All Pairs of Methods (Compared Pairwise) for the Retinal Data Subject to Unknown Transformations (Set V), Using FOM[1]

Method pair	Set V (retinal–unknown transformation)
SOMs/TM	33.105 +
SOMs/ICP	*
SOMs/CSM	20.022 +
TM/ICP	25.754−
TM/CSM	17.843−
ICP/CSM	*

$$MSE = \frac{1}{N}\sum_{i=1}^{N} \|I_R(i) - I_{CN}(i)\|^2 \qquad (12)$$

where MAX_I is the maximum pixel value of the image, that is, 255 for grayscale images represented by 8 bits/pixel. MSE is the mean square error between the reference image I_R and the noisy corresponding image I_{CN} and N is the number of pixels. The PSNR, which is measured in decibels, is inversely proportional to the level of noise present in the distorted image. This practically implies that the greater is the value of the PSNR, the weaker is the noise signal added. For the purposes of this study, the images distorted with Gaussian noise exhibited PSNR values spanning from 30.4 to 20.5 dB, while for the blurred images they ranged from 37.7 to 30.2 dB.

In order to evaluate TM, ICP, CSM, and SOMs under the presence of noise, the FOM[1] feature of merit presented in Eq. 7 was employed. As a result, the estimated corresponding points were compared against the actual correspondences in each case. As can be seen in Table 10, all methods in comparison are generally affected by noise. In particular, the FOM[1] measurements increase with the level of the noise added to the corresponding image. This holds for both Gaussian noise and blurring. Blurring seems to have a milder effect on the point correspondence efficiency, but the level of noise exhibited in this case is generally lower than the one observed in Gaussian noise. This is reflected by the higher PSNR values calculated for the blurred images, compared to the measurements obtained for images distorted with Gaussian noise. By examining Table 10, it may be concluded that all methods are equally affected by any type of noise. However, SOMs seem to have an edge over TM, ICP, and CSM, as in all cases it achieves much lower FOM[1] measurements for its estimated corresponding points. For large levels of noise, nevertheless, all methods are rendered impractical. This is shown in Table 10, where the thresholds for the minimum practical PSNR measurements, as these are evaluated through visual inspection, are 27.4 and 33.1 for Gaussian noise and blurring, respectively. Table 10 illustrates that SOMs is more resistant to noise, but adding more noise of any type will eventually hinder the efficiency of the algorithm.

DISCUSSION

In this study, four algorithms for estimating automatic point correspondences were considered: the TM, the ICP, the CSM, and the SOMs algorithms. The four algorithms were evaluated using a large variety of two-dimensional data, consisting of retinal and dental radiographs obtained by both controlled and semi-controlled geometry protocols. After quite an extensive qualitative and quantitative analysis, it was concluded that the SOMs approach outperformed in

Table 7. Performance of the Four Automatic Correspondence Methods in Terms of Point Accuracy Using FOM[1] for Retinal Images Subject to Unknown Transformations (Set V)

Pairs	TM	ICP	CSM	SOMs
1	1.265	2.460	1.181	2.598
2	6.875	5.187	4.744	3.857
3	8.959	2.764	7.303	2.049
4	5.275	5.093	4.161	3.665
5	8.112	9.902	8.819	7.587
6	2.290	2.306	2.349	6.416
7	9.079	7.265	3.221	3.751
8	7.846	3.465	3.692	2.000
9	9.840	7.542	8.519	6.005
10	5.600	1.568	4.725	4.494
11	6.133	2.628	4.728	2.931
12	1.586	1.804	1.460	1.536
13	9.347	2.397	3.590	2.090
14	7.764	5.582	6.522	2.458
15	2.361	3.002	4.549	1.204
16	8.647	5.860	7.331	6.464
17	7.054	1.992	5.488	6.171
18	6.728	2.969	6.450	8.068
19	6.157	4.907	5.052	1.128
20	5.069	4.670	3.861	1.285
Mean	6.299	4.168	4.887	3.788
STD	2.649	2.244	2.122	2.248

most cases TM, ICP as well as CSM in terms of the statistical relevance scheme and the features of merit utilized in this study.

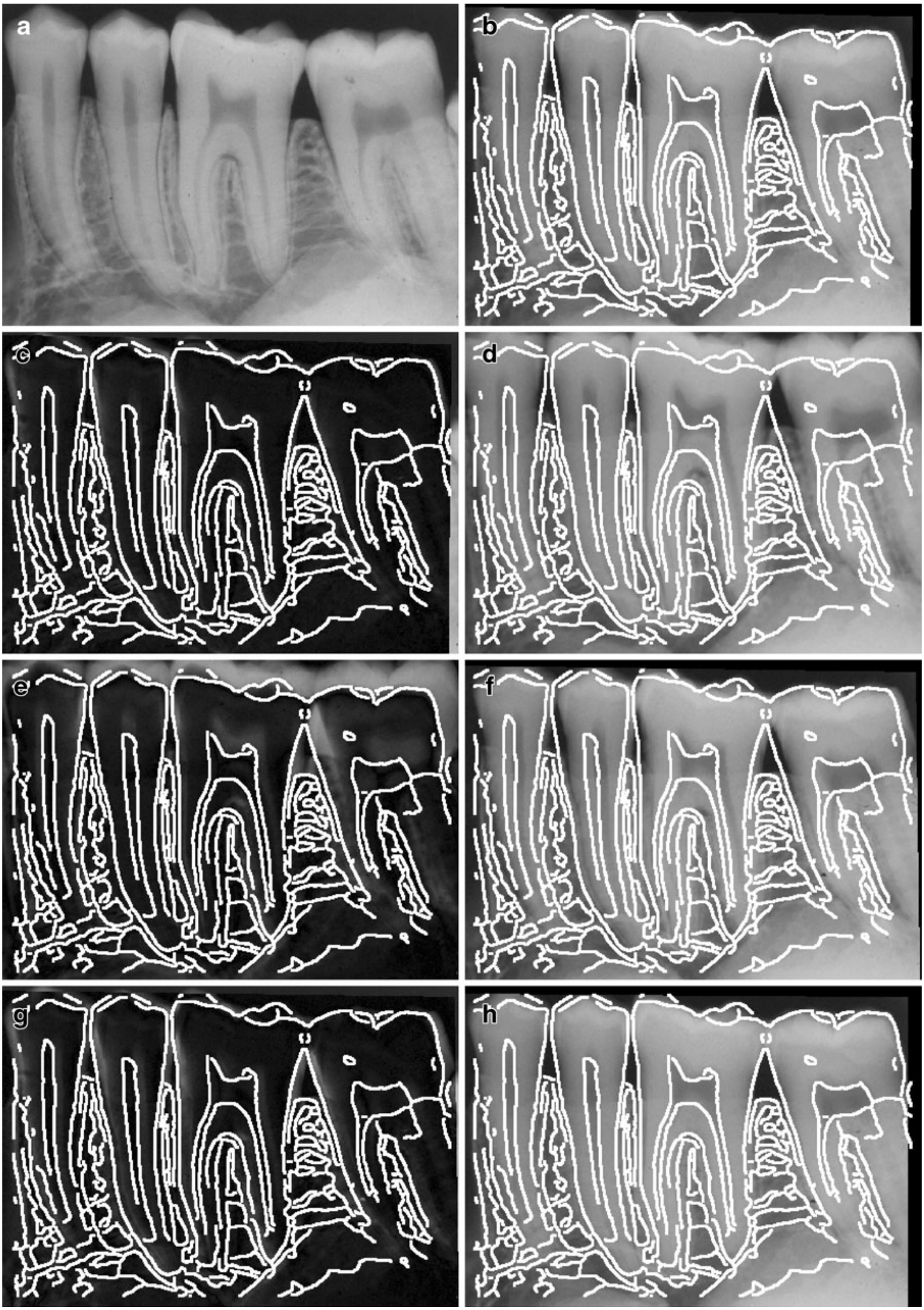
All four methods in comparison depend on several parameters which control key aspects of their algorithms. The preferred values for the parameters used to evaluate all methods are quoted in Table 2. When possible, the same values were employed for common properties of the four methods, throughout the study. Nevertheless, the values shown in Table 2 were obtained after several trials using images from all available sets in order to ensure the best possible performance out of each method. Furthermore, as can be seen in Table 2, there are numerous parameters unique to each method. The effect of varying those values has to be examined comprehensively to assess the results obtained by the four methods in comparison. For this purpose, the same image pair (pair 6 from set I) was used for all methods in order to obtain comparable results. In addition, each method exhibits its own strengths and weaknesses, which are reviewed in some further detail throughout this section.

The most important characteristic of the template-matching scheme is the size of the template itself. As mentioned previously, TM searches for correspondences by comparing selected areas from the corresponding image to a fixed pattern (the template) from the reference image. The size of the pattern is crucial for the performance of the particular method. In general, using large templates allows for a wider search space for corresponding points. Templates of increased size are especially advantageous over images exhibiting large displacements. On the other hand, large templates have a negative impact in execution time and hence are unsuitable for time-critical applications. Moreover, employing large templates on small images renders TM highly inefficient as execution time increases without any noticeable gain over point correspondence accuracy. In order to demonstrate this, a retinal image from set I (pair 6) was evaluated using the TM approach with varying template sizes. In the particular example, the method was tested using template sizes ranging from 3×3 to 33×33 pixels. The results are shown in Figure 6, in terms of FOM^1, which is presented in Eq. 7. As can be seen in Figure 6a, small template sizes hinder the correspondence accuracy of the algorithm for the particular image. On the

Fig 5. Visual assessment of the four point correspondence methods, in terms of registration quality. a Reference dental image drawn from set VII. Transformed corresponding image with the edges of the reference image superimposed for b TM, d ICP, f CSM, and h SOMs. Absolute difference between the reference image and the transformed corresponding image with the edges of the reference image superimposed for c TM, e ICP, g CSM, and i SOMs.

other hand, large templates cause the execution time to increase dramatically (Fig. 6b), without having a proportional positive effect on the scheme's accuracy (Fig. 6a). For the particular example, TM performs as expected with window sizes equal or greater than 9×9 pixels. However, as can be seen in Fig. 6a, the method exhibits its optimal performance, in terms of FOM^1, when 21×21 pixels templates are used. This also holds for the majority of the test images considered. Hence, 21×21 window sizes were preferred throughout this study in order to balance between point correspondence accuracy and efficiency, in terms of execution time.

The implementation of the iterative closest points algorithm used throughout this study is a classic implementation of ICP,[16] featuring an improved elastic transformation function. The basic idea of ICP is to calculate correspondences between two or more point clouds.[16] In our case, the first point cloud is considered to be the set of initial points on the reference image, while the second one is a set of candidate points on the corresponding image. Therefore, the particular approach requires control points to be extracted from both the reference and the corresponding image. ICP then tries to match those two sets of points in an iterative manner as described earlier. All ICP variants adopt a trial-and-error approach to allocate correspondences. This implies that the scheme is repeated for a finite number of iterations. However, there is a certain number of iterations beyond which minimal accuracy improvements are recorded. Since an increased number of trials negatively impacts on the efficiency of ICP, the preferred number of trials is usually selected such that accuracy and execution time are balanced. An example is shown in Figure 7, where an image pair from set I (pair 6) is considered. ICP was applied to the particular image pair for a varying number of iterations, ranging from 50 to 1,000. As can be seen in Figure 7a, no significant performance gain is observed after 500 iterations, in terms of FOM^1.


a
b
c
d
e
f
g
h

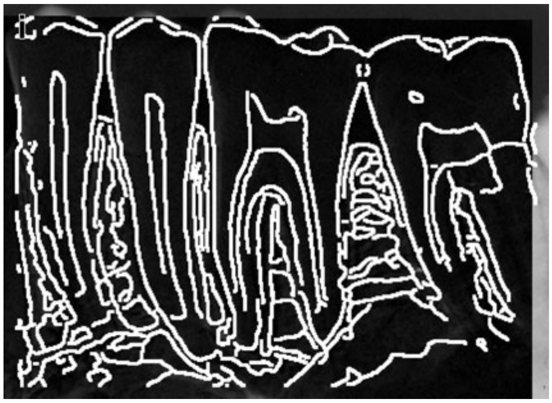


Fig 5. (continued)

Since convergence time increases steadily with the number of iterations (Fig. 7b), the particular value was considered to be sufficient and therefore was selected for evaluating ICP throughout the study.

Although the ICP method can be highly efficient and fast, it is heavily dependent upon the performance of the preferred point extraction method. The reason for this is that the algorithm itself is not able to calculate the point correspondences but it merely links points from one set to another. Therefore, ICP variants are generally not suitable for allocating point correspondences unless there is a predefined set of candidate corresponding points on the corresponding image. Those points may be either defined automatically[8,29–30] or simply manually.

The CSM technique is a method based on the principles of template matching. As mentioned, CSM firstly performs minor movements to the control points on the reference image, thus producing a tentative corresponding point from a match map for each such movement. The size of the match map refers to the area within which acceptable tentative corresponding points are considered. The size of the match map, usually expressed by a square area, is of particular importance for CSM, as only candidate points within the match map are estimated. In order to assess the optimal size of the match map, several trials were conducted featuring varying sizes ranging from 3×3 to 33×33 pixels. The performance of CSM in each case was evaluated in terms of the measured FOM^1 metric between the estimated and actual correspondences, using a retinal image from set I (pair 6). The results shown in Figure 8a indicate that the optimal match-map size is 15×15 pixels, which was adopted as the preferred value throughout the study. As can be seen in Figure 8a, either excessively small or large match maps hinder the efficiency of the CSM algorithm. The reason for this is that small match maps simply do not provide sufficient space for candidate correspondences, while large ones may well confuse the algorithm by producing spurious candidates, thus scattering the tentative corresponding points over a large area. On the other hand, the size of the match map does not have an immense effect in the execution time of the algorithm. This is illustrated in Figure 8b. Although the execution time for CSM does rise as the match-map size increases, the effect is minimal compared to TM, which also features a similar template scheme.

By examining the distribution of candidate corresponding points over minor displacements

Table 8. Average Statistical Relevance of the Difference in Performance of the Automatic Point Correspondence Methods (Compared Pairwise) for All Pairs of Methods For the Dental Image Pairs Subject to Unknown Transformations (Sets VI–VIII)

	Method pair	Set VI	Set VII	Set VIII
	SOMs/TM	12.633+	15.600+	8.793+
	SOMs/ICP	*	14.345+	9.396+
	SOMs/CSM	25.272+	24.822+	20.085+
	TM/ICP	17.115−	*	*
	TM/CSM	12.912+	10.588+	12.153+
FOM^2 (mean edge distance)	ICP/CSM	27.007+	*	10.939+
	SOMs/TM	9.340+	14.269+	16.776+
	SOMs/ICP	*	26.067+	14.189+
	SOMs/CSM	24.307+	28.387+	41.744+
	TM/ICP	*	14.257+	*
	TM/CSM	15.842+	16.511+	30.125+
FOM^3 (inverse mutual information)	ICP/CSM	22.423+	*	28.098+

Table 9. Performance of the Four Automatic Correspondence Methods in Terms of Registration Accuracy Using the Mutual Information (1/FOM^3) for Dental Images Subject to Unknown Transformations (Average Values over Sets VI–VIII)

Data sets	TM	ICP	CSM	SOMs
Semi-controlled geometry set VI (24 pairs)	1.257 ± 0.180	1.396 ± 0.270	1.046 ± 0.184	1.382 ± 0.053
Semi-controlled geometry set VII (24 pairs)	1.266 ± 0.264	1.112 ± 0.494	1.053 ± 0.232	1.467 ± 0.143
Semi-controlled geometry set VIII (24 pairs)	1.171 ± 0.238	1.217 ± 0.408	0.824 ± 0.238	1.397 ± 0.140

of the control points, CSM ensures that the estimated corresponding points are as reliable as possible. In effect, the particular method initially performs a template-matching scheme for defining the match maps and then it examines the reliability of the estimated correspondences by assessing their sensitivity to the movement of their respective control point. In that way, only reliable corresponding points may be extracted. Therefore, CSM favors robustness over execution time, as it is generally slower than TM and ICP. Moreover, although CSM is an improvement over the original template-matching scheme, it does not always guarantee superior performance over its ancestor method. In fact, there are cases where the efficiency of CSM degrades due to the characteristics of the subject image pair. Poor-quality blurred images are especially unfavorable for CSM.

As with any other iterative method, the self-organizing map algorithm is greatly affected by the number of iterations executed. Again, the more iterations are performed, the better is the accuracy of the corresponding points, up to a certain number where no further noticeable improvement can be accounted. For the particular image pairs included in this study, it was observed that SOMs did not improve point accuracy noticeably for more than 5,000 iterations. In some instances (especially in smaller images), performance was saturated for even less iterations, but the value of 5,000 was universally adopted to suffice all cases. As can be seen in Figure 9, the SOMs algorithm was tested on a retinal image pair from set I (pair 6) using a variable number of iterations (from 500 to 15,000). There, it is quite clear that employing more than 5,000 iterations does not improve the accuracy noticeably (Fig. 9a). Moreover, the fact that the convergence time of the particular approach raises dramatically as the number of trials increases (Fig. 9b), employing a larger number of iterations would hinder the efficiency of the method, as the execution time of the method increases without any performance gain.

After evaluating all methods in comparison, SOMs were considered to be the most effective method on average. This holds for sets of images subject to both known and unknown transformations. Throughout the study, it was illustrated that when SOMs were compared pairwise to all other three methods in comparison, the statistical relevance of its performance gain was greater on average in most cases than any other method.

Table 10. Performance of the Four Automatic Correspondence Methods in Terms of Registration Accuracy Using FOM^1 in the Presence of Noise (on Retinal Pair 6 from Set I)

	Noise parameter	PSNR (dB)	TM	ICP	CSM	SOMs
	0	–	7.774	4.896	1.495	2.062
	0.001	30.4	8.690	7.829	19.224	4.990
	0.002	27.4	13.863	11.351	21.988	9.064
	0.005	23.4	33.247	25.726	34.017	19.121
Gaussian noise (σ)	0.01	20.5	38.301	31.113	39.228	23.673
	0	–	7.774	4.896	1.495	2.062
	2	37.7	8.585	6.799	7.126	4.089
	4	33.1	12.285	10.548	19.541	12.339
	6	31.3	24.336	20.017	23.295	19.885
Gaussian blurring (radius)	8	30.2	33.379	24.265	26.406	22.390

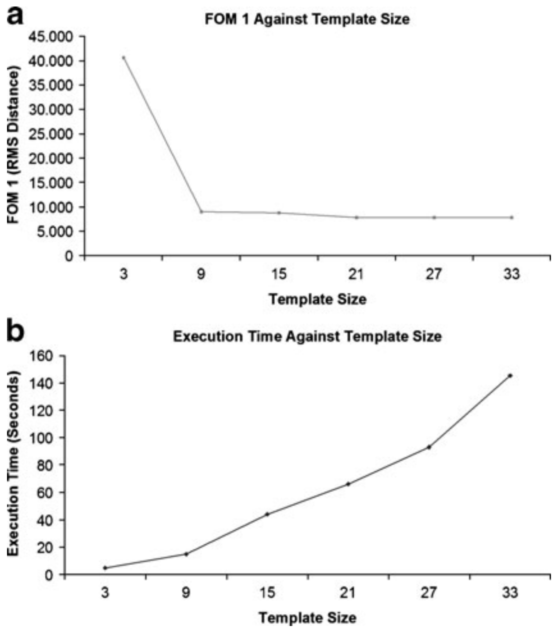


Fig 6. Performance of the template-matching technique for varying the size of the template. The template was varied from 3 to 33 pixels wide and the tests were performed on a retinal image from set I (pair 6). The performance was assessed in terms of a the FOM[1] and b the execution time.

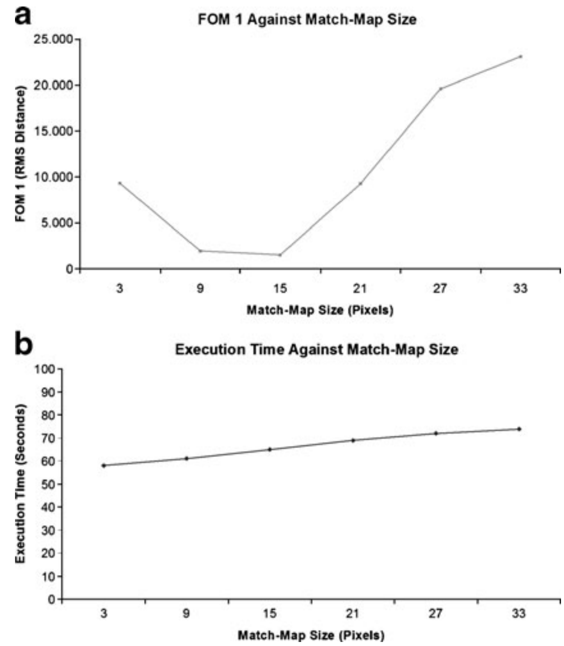


Fig 8. Performance of the correspondence by sensitivity to movement algorithm for varying the size of the match map. The match map was varied from 3 to 33 pixels wide and the tests were performed on a retinal image from set I (pair 6). The performance was assessed in terms of (a) the RMS distance and (b) the execution time.

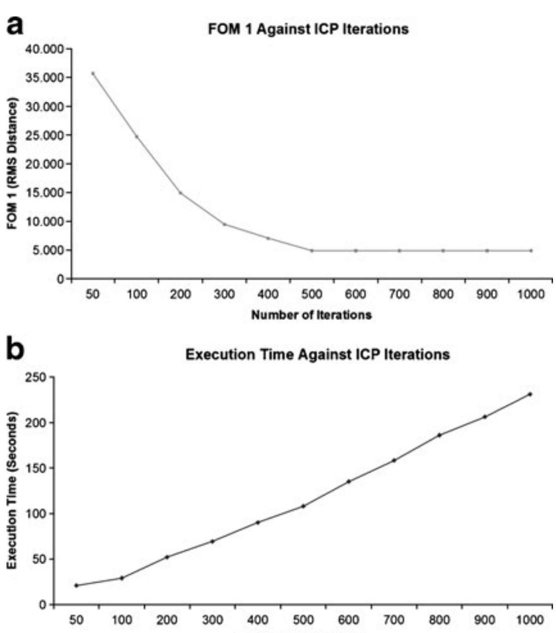


Fig 7. Performance of the ICP method for varying the number of iterations. A retinal image from set I was used (pair 6) and the method was applied with the number of iterations ranging from 50 to 1,000. The effects on performance were recorded in terms of a the RMS distance and b the execution time.

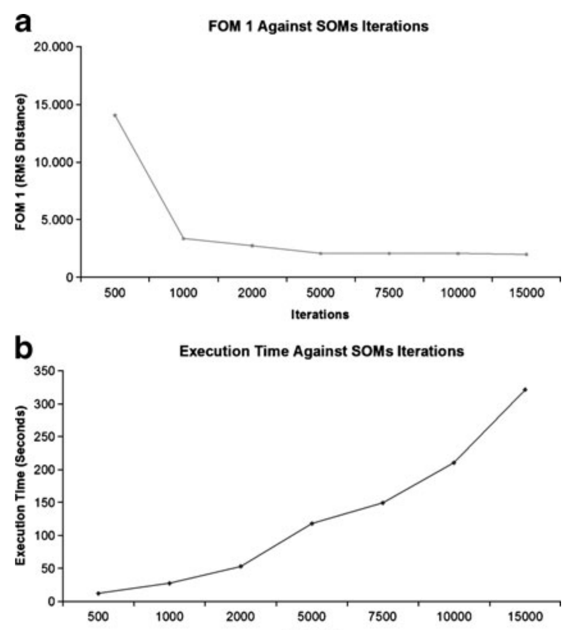


Fig 9. Performance of the self-organizing maps method for a varying number of iterations. A retinal image from set I was used (pair 6) and the method was applied with a varying number of stopping iterations ranging from 500 to 15,000. The effects on the performance were recorded in terms of (a) the RMS distance and (b) the execution time.

Cumulative results are shown in Table 11, where the number of cases SOMs which performed better, equally, or worse than any of the other three methods is recorded. The measurements shown in Table 11 are drawn from Tables 4, 6, and 8, in terms of the average statistical relevance metric obtained in each case. As can be seen in Table 11, SOMs outperformed all other methods in comparison in 12 out of 15 cases for the root mean square distance feature of merit (FOM^1). Moreover, it did better than TM, ICP, and CSM altogether in eight out of nine cases using either the mean edge distance (FOM^2) or the inverse mutual information features of merit (FOM^3). Therefore, SOMs outperformed all other three methods in comparison in totally 28 out of 33 cases over all available FOMs, as can be seen in Table 11.

The performance of the TM algorithm may be improved by reducing the size of the sub-images, but this negatively affects point correspondence accuracy, especially if large translations are involved in the corresponding image. On the other hand, the performance of ICP depends on three key factors. Firstly, the particular method requires two point clouds to be extracted (from the reference and subsequent images), which implies that the accuracy of the algorithm is directly proportional to the accuracy of the preferred point extraction method. Secondly, it is possible that the two extracted point clouds do not match perfectly. For example, some correspondences may not be established, as suitable corresponding points simply do not exist in the candidate corresponding points cloud. In such a case, ICP will try the next best possible match, which will most likely produce a false correspondence. Finally, as mentioned, the accuracy of ICP also depends on the number of iterations performed. As far as CSM is concerned, the particular approach calculates several candidate corresponding points and then selects the optimal point by running a simple reliability test. By moving each control point over a specified area of the reference image, several scattered candidate points are estimated. In general, if those points are scattered over a relatively small area on the corresponding image, the candidate points are considered to be reliable. On the contrary, wildly scattered candidate points indicate an erroneous estimation process or an unsuitable control point.

The superiority of SOMs is mainly contributed to the ability of the particular method to calculate accurate point correspondences by limiting the amount of spurious corresponding points that mostly affect the other three methods in comparison, especially TM and CSM. This is achieved due to the characteristics of the algorithm itself. Points in SOMs are arranged into a neural network and, in order to calculate each corresponding point, the method takes into account the entire set of estimated points, thus minimizing the risk of considering false correspondences.

Nevertheless, the superiority of SOMs comes at a cost to convergence time, as it was by far the slowest method of all four methods examined. In general, the convergence time depends directly on the number of control points used. For example, for the dental images of the study, where 200 points were extracted, it took about 15 min to calculate point correspondences with SOMs. For the same images, TM estimated the correspondences in approximately 9 min, ICP in about 3 min, and CSM in 10 min. The particular timings are significantly reduced for the retinal images, where 13 to 30 bifurcation points were detected in any case. All tests featured in this study were performed on a typical desktop workstation (AMD Opteron 165, 1,800-MHz processor with 2 GB of random access memory).

This comparative study was conducted specifically using two-dimensional retinal and dental medical data. The aim of future research is twofold: firstly, to extend the current study to two-dimensional multimodal data sets (CT, magnetic resonance imaging, PET scans) and secondly to modify the four automatic point correspondence algorithms for operating on three-dimensional data sets.

Table 11. Number of Times that the SOMs Algorithm Performed Better than, Same as, and Worse than TM, ICP, and CSM Algorithms, When Compared Pairwise Using All Possible Combinations, for All Sets, and All Three FOMs

FOM^i	Better	Neutral	Worse
$i = 1$ (RMS distance)	12	3	0
$i = 2$ (mean edge distance)	8	1	0
$i = 3$ (inverse mutual information)	8	1	0
Total	28	5	0

CONCLUSIONS

This paper presented a comprehensive evaluative study comparing four commonly used algorithms for obtaining automatic point correspondences in two dimensions. The TM, the ICP, the CSM, and the SOMs algorithms were compared pairwise as well as independently. The methods were assessed both visually and quantitatively. The four methods were applied individually on 40 medical image pairs featuring known transformations and 92 more pairs subject to unknown transformations. In general, the SOMs approach outperformed in most cases the other three methods in comparison. SOMs achieved systematically superior point correspondence accuracy in most cases when compared pairwise to TM, ICP, and CSM, in terms of the statistical relevance measurement. The particular approach also outperformed the other three methods in comparison under the presence of noise in the corresponding images. On the other hand, TM and ICP performed evenly especially for images subject to known transformations, while CSM consistently underperformed when applied to dental images obtained through a semi-controlled geometry protocol.

REFERENCES

1. Maintz JBA, Viergever MA: A survey of medical image registration. Med Im Anal 2(1):1–36, 1998
2. Vujovic N, Brzakovic D: Establishing the correspondence between control points in pairs of mammographic images. IEEE Trans Image Proc 6(10):1388–1399, 1997
3. Heyden A, Berthilsson R, Sparr G: An iterative factorization method for projective structure and motion from image sequences. Image Vis Comput 17(13):981–991, 1999
4. Liu Y, Huang TS, Faugeras OD: Determination of camera location from 2-D to 3-D line and point correspondences. IEEE Trans PAMI 12(1):28–37, 1990
5. Zhang X, Liu Y, Huang TS: Motion analysis of articulated objects from monocular images. IEEE Trans PAMI 28(4):625–636, 2006
6. Micusık B, Pajdla T: Structure from motion with wide circular field of view cameras. IEEE Trans PAMI 28(7):1135–1149, 2006
7. Yu LH, Eizenman M: A new methodology for determining point-of-gaze in head-mounted eye tracking systems. IEEE Trans Biomed Eng 51(10):1765–1773, 2004
8. Likar B, Pernus F: Automatic extraction of corresponding points for the registration of medical images. Med Phys 26:1678–1686, 1999
9. Bedekar AS, Haralick RM: A Bayesian method for triangulation and its application to finding corresponding points. Int Conf Im Proc 2:2362, 1995
10. Pratt WK: Digital image processing. New York: Wiley, 1978
11. Di Stefano L, Mattocia S, Tombari F: An algorithm for efficient and exhaustive template matching. Im Anal Recognition 3211:408–415, 2004
12. Chen YS, Hung YP, Fuh CS: Fast block matching algorithm based on the winner update strategy. IEEE Trans Image Proc 10(8):1212–1222, 2001
13. Ghanbari M: The cross-search algorithm for motion estimation. IEEE Trans Commun 38(7):950–953, 1990
14. Chen JH, Chen CS, Chen YS: Fast algorithm for robust template matching with m-estimators. IEEE Trans Signal Process 51(1):230–243, 2003
15. Guest E, Berry E, Baldock RA, Fidrich M, Smith MA: Robust point correspondence applied to two- and three-dimensional image registration. IEEE Trans Patt Anal Mach Intell 23(2):165–179, 2001
16. Besl PJ, McKay ND: A method for registration of 3-D shapes. IEEE Trans Patt Anal Mach Intell 14(2):239–256, 1992
17. Matsopoulos GK, Asvestas PA, Mouravliansky NA, Delibasis KK: Multimodal registration of retinal images using self organizing maps. IEEE Trans Med Imag 23:1557–1563, 2004
18. Cao Z, Pan S, Li R, Balachandran R, Fitzpatrick MJ, Chapman WC, Dawant BM: Registration of medical images using an interpolated closest point transform: method and validation. Med Image Anal 8(4):421–427, 2004
19. Chetverikov D, Stepanov D, Krsek P: Robust Euclidean alignment of 3D points sets the: trimmed iterative closest point algorithm. Image Vis Comput 23:299–309, 2005
20. Georgescu B, Meer P: Point matching under large image deformations and illumination changes. IEEE Trans Pattern Anal Mach Intell 26(6):674–688, 2004
21. Economopoulos T, Matsopoulos GK, Asvestas PA, Gröndahl K, Gröndahl HG: Automatic correspondence using the enhanced hexagonal centre-based inner search algorithm for point-based dental image registration. Dentomaxillofac Radiol 37(4):185–204, 2008
22. Marti R, Zwiggelaar R, Rubin CME: Automatic point correspondence and registration based on linear structures. Intern J Pattern Recognit Artif Intell 16(3):331–340, 2002
23. Wang F, Vemuri BC, Rangarajan A, Eisenschenk SJ: Simultaneous nonrigid registration of multiple point sets and atlas construction. IEEE Trans Pattern Anal Mach Intell 30 (11):2011–2022, 2008
24. Delibasis KK, Asvestas PA, Mouravliansky NA, Economopoulos TL, Matsopoulos GK: Artificial immune network for automatic point correspondence in medical images. Conf Proc IEEE Eng Med Biol Soc. 840–843, 2007
25. Kohonen T: The self-organizing map. Proc IEEE 78 (9):1464–1480, 1990
26. Zitova B, Flusser J: Image registration methods: a survey. Image Vis Comput 21(11):977–1000, 2003
27. Matabosch C, Salvi J, Pinsach X, Pag J: A comparative survey on free-form surface registration. Image and Vision Computing 308–312, 2004

28. Grevera GJ, Udupa JK: An objective comparison of 3-D image interpolation methods. IEEE Trans Med Im 17(4):642–652, 1998

29. Laliberte F, Gagnon L, Sheng Y: Registration and fusion of retinal images an: evaluation study. IEEE Trans Med Imag 22(5):404–418, 2003

30. Rohr K: Extraction of 3D anatomical point landmarks based on invariance principles. Pattern Recognit 32:3–19, 1999

31. Bookstein FL: Principal warps: thin plate splines and the decomposition of deformations. IEEE Trans Pattern Anal Mach Intell 11:567–585, 1989

32. Devore JL: Probability and statistics for engineering and the sciences, Pacific Grove: Brooks/Cole, 1991

33. Zacharaki EI, Matsopoulos GK, Asvestas PA, Gröndahl K, Gröndahl HG: A digital subtraction radiography scheme based on automatic multiresolution registration. Dentomaxillofac Rad 33:1–14, 2004

34. Press WH, Teukolsky SA, Vetterling WT, Flannery BP: Numerical recipes in C: the art of scientific computing, Cambridge: Cambridge University Press, 1992

35. Gonzalez R, Woods R: Digital image processing. Addison Wesley: Reading, 1992, pp 414–428

36. Canny J: A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 8:679–714, 1986

37. Netravali AN, Haskell BG: Digital pictures: representation, compression, and standards, 2nd edition. New York: Plenum, 1995