

A New Graph Matching Method for Point-Set Correspondence using the EM Algorithm and Softassign

Gerard Sanromà^{a,1,*}, René Alquézar^b, Francesc Serratosa^a

^a*Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Av. Països Catalans, 26 Campus Sescelades, 43007 Tarragona, Spain.*

^b*Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Parc Tecnològic de Barcelona. C/ Llorens i Artigas 4-6. 08028 Barcelona, Spain.*

Abstract

Finding correspondences between two point-sets is a common step in many vision applications (e.g., image matching or shape retrieval). We present a graph matching method to solve the point-set correspondence problem, which is posed as one of mixture modelling. Our mixture model encompasses a model of structural coherence and a model of affine-invariant geometrical errors. Instead of absolute positions, the geometrical positions are represented as relative positions of the points with respect to each other. We derive the Expectation-Maximization algorithm for our mixture model. In this way, the graph matching problem is approximated, in a principled way, as a succession of assignment problems which are solved using Softassign. Unlike other approaches, we use a true continuous underlying correspondence variable. We develop effective mechanisms to detect outliers. This is a useful technique for improving results in the presence of clutter. We evaluate the ability of our method to locate proper matches as well as to recognize object categories in a series of registration and recognition experiments. Our method compares favourably to other graph matching methods as well as to point-set registration methods and outlier rejectors.

Keywords: correspondence problem, graph matching, affine registration, outlier detection, expectation maximization, softassign

*Corresponding author

Email addresses: gsanroma@gmail.com (Gerard Sanromà), ralquezar@iri.upc.edu (René Alquézar), francesc.serratosa@urv.cat (Francesc Serratosa)

¹Fax: (+34) 977 55 97 10

1. Introduction

The correspondence problem in computer vision tries to determine which parts of one image correspond to which parts of another image. This problem often arises at the early stages of many computer vision applications such as 3D scene reconstruction, object recognition, pose recovery and image retrieval, among others. So, it is of basic importance to develop effective methods that are both robust -in the sense of being able to deal with noisy measurements- and general -in the sense of having a wide field of application-.

The typical steps involved in the solution of the correspondence problem are the following. First, a set of tentative feature matches is computed. These tentative matches can be further refined by a process of outlier rejection that eliminates the spurious correspondences or alternatively, they can be used as starting point of some optimization scheme to find a different, more consistent set.

Tentative correspondences may be computed either on the basis of correlation measures or feature-descriptor distances.

Correlation-based strategies compute the matches by means of the similarity between the image patches around some interest points. Interest points (that play the role of the images' parts to be matched) are image locations that can be robustly detected among different instances of the same scene with varying imaging conditions. Interest points can be corners (intersection of two edges) [1] [2] [3], maximum curvature points [4] [5] [6] or isolated points of maximum or minimum local intensity [7].

On the other hand, approaches based on feature-descriptors use the information local at the interest points to compute descriptor-vectors. Those descriptor-vectors are meant to be invariant to geometric and photometric transformations. So that, corresponding areas in different images present low distances between their feature-descriptors. A recent paper by Mikolajczyk and Schmid [8] evaluate some of the most competent approaches.

Despite the invariance introduced during the detection/description and the matching phases, the use of local image contents may not suffice to get a reliable result under certain circumstances (e.g., regular textures, multiple instances of a given feature across the images or, large rigid/non-rigid deformations). Figure 1 shows an example of a matching by correlation of a scene under rotation and zoom.

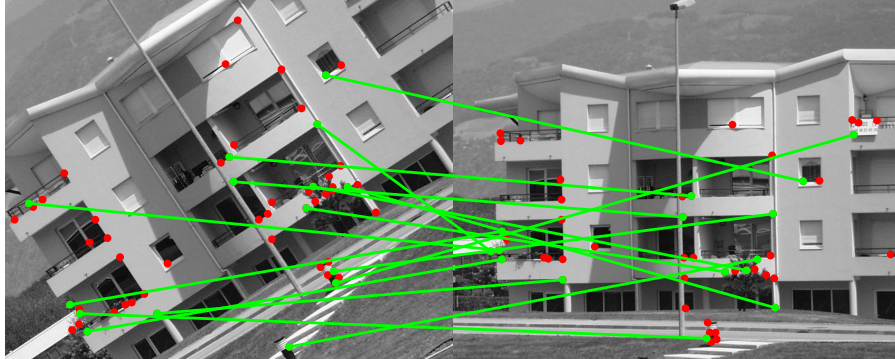


Figure 1: Two sample images belonging to the class *Resid* from ref. [9] with superposed Harris corners [1]. The green lines represent the tentative correspondences computed by matching by correlation. The red dots are unmatched points. There are several misplaced correspondences.

It is a standard procedure to exploit the underlying geometry of the problem to enforce the global consistency of the correspondence-set. This is the case of the model fitting paradigm RANSAC [10] which is extensively used in computer vision to reject outliers. It selects random samples of correspondences from a tentative set and use them to fit a geometric model to the data. The largest consensus obtained after a number of trials is selected as the inlier class. Another effective outlier rejector is based on a Graph Transformation [11]. This is an iterative process that discards one outlying correspondence at a time, according to a graph-similarity measure. After each iteration, the graphs are reconfigured in order to reflect the new state of the remaining correspondences. The process ends up with two isomorphic graphs and the surviving correspondences constitute the inlier class.

The main drawback of these methods is that their ability to obtain a dense correspondence-set strongly depends on the reliability of the tentative correspondences. Since they are unable either to generate new correspondences or to modify the existing ones, an initial correspondence-set with few successes may result in a sparse estimate. This is illustrated in figure 2.

Other approaches such as Iterative Closest Point (ICP) [12] that fall into the optimization field, attempt to simultaneously solve the correspondence and the alignment problem. Despite they are able to modify the correspondences at each iteration, simple nearest neighbour association is prone to local minima, specially under bad initial alignment estimates.

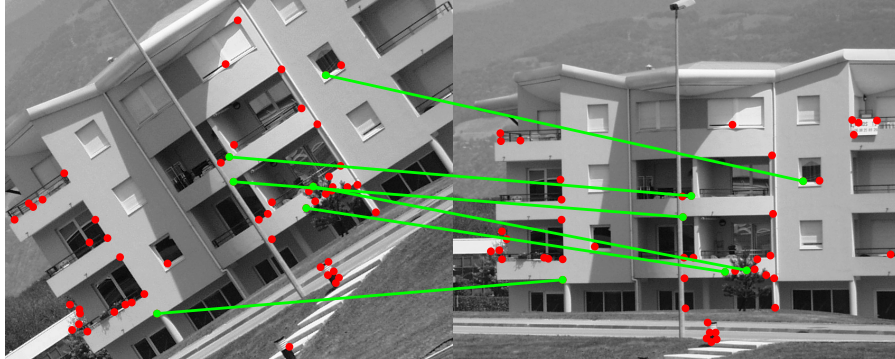


Figure 2: The green lines represent the resulting RANSAC inliers from the initial correspondence-set from figure 1. Only a few inliers are found by RANSAC. This may not be suitable in the cases when a more dense correspondence-set is needed.

Attributed Relational Graphs (more generally, graphs) are representational entities allowing for attributes in the nodes and relations among them in the edges. Attributed Graph Matching methods are optimization techniques that contemplate these two types of information to compute the matches and therefore, do not rely on simple nearest neighbour association. In the following section, we review the process of solving the correspondence problem in computer vision using graph techniques.

1.1. The Correspondence Problem in Computer Vision using Graphs

The first step at solving the correspondences between two images is to extract their graph representations.

In the case of general images, a commonly adopted representation is to associate feature points to nodes and generate the edge relations following either a Delaunay triangulation [13] or a k -nearest-neighbor strategy [11].

In the case of binary shape images, it is common to extract the graphs using the shapes' medial axis or *skeleton* [14] [15]. Some approaches to chinese character recognition represent the strokes in the nodes [16] [17]. However, it is more usual to represent the skeletal end-and-intersection-points in the nodes, and their links in the edges. Some approaches use *Shock-graphs* [18] [19] [20] or *Attributed Skeletal Graphs* [21]. These are types of graphs which are closely related to the skeletal representations and therefore, cannot be applied to more general computer vision problems.

Labeling objects of a scene using their relational constraints is at the core of all general-purpose graph-matching algorithms. An early attempt to discrete labeling was by Waltz [22]. Rosenfeld *et al.* [23] developed a model to relax the Waltz’s discrete labels by means of probabilistic assignments. They introduced the notion of compatibility coefficients and laid the bases of *probabilistic relaxation* in graph matching. Hummel and Zucker [24] firmly positioned the probabilistic relaxation into the continuous optimization domain by demonstrating that finding consistent labellings was equivalent at maximizing a local average consistency functional. Thus, the problem could be solved with standard continuous optimization techniques such as gradient ascent. Gold and Rangarajan [25] developed an optimization technique, *Graduated Assignment*, specifically designed to the type of objective functions used in graph matching. They used a Taylor series expansion to approximate the solution of a quadratic assignment problem by a succession of easier linear assignment problems. They used Softassign [26][27][28] to solve the linear assignment problems in the continuous domain. The key ingredients of their approach were two-way constraints satisfaction and a continuation method to avoid poor local minima.

Another family of approaches, also in the continuous optimization domain, uses statistical estimation to solve the problem. Christmas *et al.* [29] derived the complete relaxation algorithm, including the calculation of the compatibility coefficients, following the Maximum A Posteriori (MAP) rule. Hancock *et al.* [30][31] used *cliques*, a kind of graph sub-entities, for graph matching. Furthermore, they proposed a new principled way of detecting outliers that consists in measuring the net effects of a node deletion in the retriangulated graph. Accordingly, an outlier is a node that leads to an improvement in the consistency of the affected cliques after its removal. Nodes are regularly tested for deletion or reinsertion following this criterion.

Hancock *et al.* [32] [31] formulated the problem of graph matching as one of probability mixture modeling. This can be thought of as a *missing data* problem where the correspondence indicators are the parameters of the distribution and the corresponding nodes in the model-graph are the hidden variables. They used the Expectation-Maximization (EM) Algorithm [33] to find the Maximum Likelihood (ML) estimate of the correspondence indicators. Hancock *et al.* [34] [31] presented approaches to jointly solve the correspondence and alignment problems. They did so by exploiting both the geometrical arrangement of the points and their structural relations.

The advantages of posing graph matching as a joint correspondence and

alignment problem, are twofold. On one hand, structural information may contribute to disambiguate the recovery of the alignment (unlike purely geometric approaches). On the other hand, geometrical information may aid to clarify the recovery of the correspondences in the case of structural corruption (unlike structural graph matching approaches).

We present a new graph matching approach aimed at finding the correspondences between two sets of coordinate points. The main novelties of our approach are:

- Instead of individual measurements, our approach uses relational information of two types: structural and geometrical. This contrasts with other approaches that use absolute geometrical positions [31][34].
- It maintains a true continuous underlying correspondence variable throughout all the process. Although there are approaches that relax the discrete assignment constraints through the use of statistical measurements, their underlying assignment variable remains discrete [30][31][32][34].
- We face the graph matching problem as one of mixture modelling. To that end, we derive the EM algorithm for our model and approximate the solution as a succession of assignment problems which are solved using Softassign.
- We develop effective mechanisms to detect and remove outliers. This is a useful technique in order to improve the matching results.

Figure 3 shows the results of applying our method to the previous matching example.

Although they are more effective, Graph Matching algorithms are also more computationally demanding than other approaches such as the robust estimator RANSAC. Suboptimal Graph Matching algorithms, such as the ones treated in this paper, often present an $\mathcal{O}(N^4)$ complexity. However, Graph Matching algorithms can be very useful at specific moments during a real-time operation, e.g. when the tentative correspondence-sets are insufficient for further refinement or when drastic discontinuities appear in the video flow that cause the tracking algorithms to fail. When these circumstances are met, it may be advisable to take a couple of seconds in order to conveniently redirect the process. We present computational time results that demonstrate that our algorithm can match a considerable amount of points in an admissible time using a C implementation.

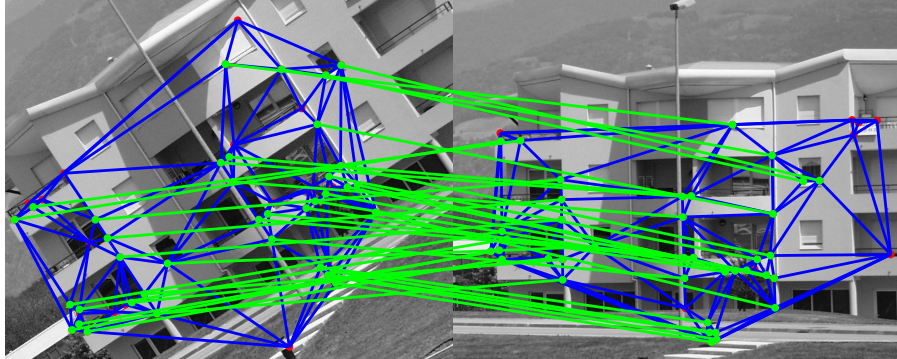


Figure 3: Superposed on the images there are the extracted graphs. Blue lines within each image represent the edges, generated by means of a Delaunay triangulation on the nodes. The nodes correspond to the Harris corners. The green lines represent the resulting correspondences of applying our method, using as starting point the correspondence-set of figure 1. Our approach arrives at a correct dense correspondence-state, while still leaving a few unmatched outliers in both images.

The outline of this paper is as follows. In section 2, we formalize some concepts such as graphs representations and correspondence indicators. The mixture model is presented in section 3. In section 4, we give the details on the optimization procedure using the EM algorithm. The mechanisms for outlier detection are presented in section 5. Before the experimental validation is provided in section 7, we briefly overview some methods related to ours in section 6. Last, discussion about the results and concluding remarks are given in sections 8 and 9.

2. Graphs and Correspondences

Consider two graph representations $G = (\mathbf{v}, D, \mathbf{p})$ and $H = (\mathbf{w}, E, \mathbf{q})$, extracted from two images (e.g., figure 3).

The node-sets $\mathbf{v} = \{v_a, \forall a \in \mathcal{I}\}$ and $\mathbf{w} = \{w_\alpha, \forall \alpha \in \mathcal{J}\}$ contain the symbolic representations of the nodes, where $\mathcal{I} = 1 \dots |\mathbf{v}|$ and $\mathcal{J} = 1 \dots |\mathbf{w}|$ are their index-sets.

The vector-sets $\mathbf{p} = \{\vec{p}_a = (p_a^H, p_a^V), \forall a \in \mathcal{I}\}$ and $\mathbf{q} = \{\vec{q}_\alpha = (q_\alpha^H, q_\alpha^V), \forall \alpha \in \mathcal{J}\}$, contain the column vectors of the two-dimensional coordinates (horizontal and vertical) of each node.

The adjacency matrices D and E contain the edge-sets, representing some kind of structural relation between pairs of nodes (e.g., connectivity or spatial

proximity).

Hence, $D_{ab} = \begin{cases} 1 & \text{if } v_a \text{ and } v_b \text{ are linked by an edge} \\ 0 & \text{otherwise} \end{cases}$ (the same applies for $E_{\alpha\beta}$).

We deal with *undirected unweighted* graphs. This means that the adjacency matrices are symmetric ($D_{ab} = D_{ba}, \forall a, b \in \mathcal{I}$) and its elements can only take the $\{0, 1\}$ values. However, our model is also applicable to the *directed weighted* case.

The variable S represents the state of the correspondences between the node-sets \mathbf{v} and \mathbf{w} . Therefore, we denote the probability that a node $v_a \in \mathbf{v}$ corresponds to a node $w_\alpha \in \mathbf{w}$ as $s_{a\alpha} \in S$.

It is satisfied that

$$\sum_{\alpha \in \mathcal{J}} s_{a\alpha} \leq 1, \forall a \in \mathcal{I} \quad (1)$$

where the probability of node v_a being an outlier equals to

$$1 - \sum_{\alpha \in \mathcal{J}} s_{a\alpha} \quad (2)$$

2.1. Geometrical Relations

Similarly as it is done with the structural relations, instead of its individual measurements, our aim is to consider the geometrical relations between pairs of nodes. To that end, we define the new coordinate vectors $\vec{x}_{ab} = (\vec{p}_b - \vec{p}_a), \forall a, b \in \mathcal{I}$ and $\vec{y}_{\alpha\beta} = (\vec{q}_\beta - \vec{q}_\alpha), \forall \alpha, \beta \in \mathcal{J}$, that represent the coordinates of the points \vec{p}_b and \vec{q}_β relative to \vec{p}_a and \vec{q}_α , respectively. Accordingly, we define a new descriptor \mathbf{x}_a for node v_a , as the translated positions of the remaining points so that their new origin is at point \vec{p}_a , i.e., $\mathbf{x}_a = \{\vec{x}_{ai}, i \in \mathcal{I}\}$. Similarly for graph H , $\mathbf{y}_\alpha = \{\vec{y}_{\alpha j}, j \in \mathcal{J}\}$. This is illustrated in figure 4.

Affine invariance is introduced at the level of node descriptors so, we consider different affine registration parameters $\Phi_{a\alpha}$ for each possible correspondence $v_a \rightarrow w_\alpha$. Since geometrical information is used in a relational way, affine registration does not depend on any translation parameter. Affine registration parameters $\Phi_{a\alpha}$ are then defined by the 2×2 matrix

$$\Phi_{a\alpha} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}.$$

We denote the whole set of affine registration parameters as $\Phi = \{\Phi_{a\alpha}, \forall a, \alpha\}$.

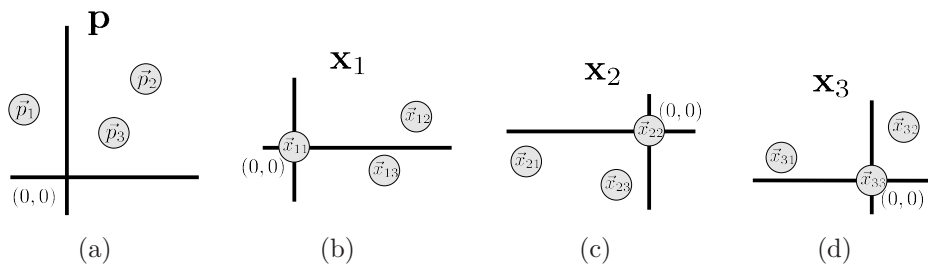


Figure 4: The entire point-set \mathbf{p} (a) and, the descriptors \mathbf{x}_1 (b), \mathbf{x}_2 (c) and \mathbf{x}_3 (d), that represent the spatial distribution of the point-sets around their new origins \vec{p}_1 , \vec{p}_2 and \vec{p}_3 , respectively.

3. A Mixture Model

Our aim is to recover the set of correspondence indicators S that maximize the incomplete likelihood of the relations in the observed graph G . Since the geometrical relations are compared in an affine invariant way, we contemplate the affine registration parameters Φ . Ideally, we seek the correspondence indicators that satisfy

$$S = \arg \max_{\hat{S}} \left\{ \max_{\hat{\Phi}} P(G|\hat{S}, \hat{\Phi}) \right\} \quad (3)$$

The mixture model reflects the possibility that any single node can be in correspondence with any of the reference nodes. The standard procedure to build likelihood functions for mixture distributions consists in factorizing over the observed data (i.e., observed graph nodes) and summing over the hidden variables (i.e., their corresponding reference nodes). We write,

$$P(G|S, \Phi) = \prod_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} P(v_a, w_\alpha | S, \Phi_{a\alpha}) \quad (4)$$

where $P(v_a, w_\alpha | S, \Phi_{a\alpha})$ represents the probability that node v_a corresponds to node w_α given the correspondence indicators S and the registration parameters $\Phi_{a\alpha}$. We are assuming conditional independence between the observed nodes.

Following a similar development than Luo and Hancock [32] we factorize, using the Bayes rules, the conditional likelihood in the right hand side of equation (4) into terms of individual correspondence indicators, in the

following way.

$$P(v_a, w_\alpha | S, \Phi_{a\alpha}) = K_{a\alpha} \prod_{b \in \mathcal{I}} \prod_{\beta \in \mathcal{J}} P(v_a, w_\alpha | s_{b\beta}, \Phi_{a\alpha}) \quad (5)$$

where

$$K_{a\alpha} = \left[\frac{1}{P(v_a | w_\alpha, \Phi_{a\alpha})} \right]^{|\mathcal{I}| \times |\mathcal{J}| - 1} \quad (6)$$

If we assume that the observed node v_a is conditionally dependant on the reference node w_α and the registration parameters $\Phi_{a\alpha}$ only in the presence of the correspondence matches S , then $P(v_a | w_\alpha, \Phi_{a\alpha}) = P(v_a)$.

If we assume equiprobable priors $P(v_a)$, then we can safely discard these quantities in the maximization of equation (3), since they do not depend neither on S or Φ .

The main aim of equation (5) is to measure the likelihood of the correspondence between nodes $v_a \in \mathbf{v}$ and $w_\alpha \in \mathbf{w}$, by evaluating the compatibility of the pairwise relations emanating from them, by means of the correspondence indicators $s_{b\beta}$.

3.1. A Probability Density Function

In the following, we propose a density function for measuring the conditional likelihood of the individual relations in the right hand side of equation (5).

For the sake of clarity, we will define our density function in different stages. First, we will propose separate structural and geometrical models in the case of binary correspondence indicators, i.e., $s_{b\beta} = \{0, 1\}$, $\forall b \in \mathcal{I}, \forall \beta \in \mathcal{J}$. Next, we will fuse these separate relational models into a combined one and, last we will extrapolate to the case of continuous correspondence indicators.

Regarding the structural relations, we draw on the model in [32] [34]. It considers that structural errors occur with a constant probability P_e . This is, given two corresponding pairs of nodes $v_a \rightarrow w_\alpha, v_b \rightarrow w_\beta$, we assume that there will be lack of edge-support (i.e., $D_{ab} = 0 \vee E_{\alpha\beta} = 0$) with a constant probability P_e . Accordingly, we define the following likelihood function

$$P(D_{ab}, E_{\alpha\beta} | s_{b\beta}) = \begin{cases} (1 - P_e) & \text{if } s_{b\beta} = 1 \wedge D_{ab} = 1 \wedge E_{\alpha\beta} = 1 \\ P_e & \text{otherwise} \end{cases} \quad (7)$$

With regards to the geometrical relations we consider that, in the case of correspondence between nodes v_b and w_β , an affine-invariant measurement

of the relative point errors $P(\vec{x}_{ab}, \vec{y}_{\alpha\beta} | \Phi_{a\alpha})$ (for brevity $P_{a\alpha b\beta}$) is appropriate in gauging the likelihood of the relation \vec{x}_{ab} . We use a multivariate Gaussian distribution to model this process. We write

$$P_{a\alpha b\beta} = \frac{1}{2\pi|\Sigma|^{1/2}} \exp \left[-\frac{1}{2} (\vec{x}_{ab} - \Phi_{a\alpha} \vec{y}_{\alpha\beta})^T \Sigma^{-1} (\vec{x}_{ab} - \Phi_{a\alpha} \vec{y}_{\alpha\beta}) \right] \quad (8)$$

where Σ is a diagonal variance matrix and, $\Phi_{a\alpha} \vec{y}_{\alpha\beta}$ are the transformed coordinates $\vec{y}_{\alpha\beta}$ according to the affine registration parameters $\Phi_{a\alpha}$, a 2×2 matrix of affine scale and rotation parameters. Note that \vec{x}_{ab} and $\vec{y}_{\alpha\beta}$ are already invariant to translation (figure 4).

In the case of no correspondence between nodes v_b and w_β , we assign a constant probability ρ that controls the outlier process (see section 5). Therefore, the conditional likelihood becomes

$$P(\vec{x}_{ab}, \vec{y}_{\alpha\beta} | s_{b\beta}, \Phi_{a\alpha}) = \begin{cases} P_{a\alpha b\beta} & \text{if } s_{b\beta} = 1 \\ \rho & \text{if } s_{b\beta} = 0 \end{cases} \quad (9)$$

Now it is turn to define a combined measurement for the structural and geometrical likelihoods. To this end, we fuse the densities of equations (7) and (9) into the following expression

$$P(v_a, w_\alpha | s_{b\beta}, \Phi_{a\alpha}) = \begin{cases} (1 - P_e) P_{a\alpha b\beta} & \text{if } s_{b\beta} = 1 \wedge (D_{ab} = 1 \wedge E_{\alpha\beta} = 1) \\ P_e P_{a\alpha b\beta} & \text{if } s_{b\beta} = 1 \wedge (D_{ab} = 0 \vee E_{\alpha\beta} = 0) \\ P_e \rho & \text{if } s_{b\beta} = 0 \end{cases} \quad (10)$$

The above density function is defined only in the case of binary correspondence indicators $s_{b\beta}$. We extrapolate it to the continuous case by exploiting, as exponential indicators, the conditional expressions of equation (10) in the following way,

$$P(v_a, w_\alpha | s_{b\beta}, \Phi_{a\alpha}) = \left[(1 - P_e) P_{a\alpha b\beta} \right]^{D_{ab} E_{\alpha\beta} s_{b\beta}} \left[P_e P_{a\alpha b\beta} \right]^{(1 - D_{ab} E_{\alpha\beta}) s_{b\beta}} \left[P_e \rho \right]^{(1 - s_{b\beta})} \quad (11)$$

Figure 5 shows an illustrative plot of the density function of equation (11).

Substituting equation (11) into (5) (and discarding the observed node priors $P(v_a)$), the final expression for the likelihood of the correspondence

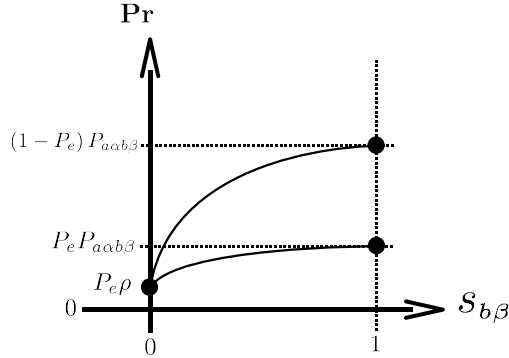


Figure 5: Density function of equation (11), an extension of the function of equation (10) to continuous correspondence indicators. Each solid curve represent either the case of edge-support (i.e., $D_{ab} = 1 \wedge E_{\alpha\beta} = 1$) or lack of it (i.e., $D_{ab} = 0 \vee E_{\alpha\beta} = 0$). At the extrema of each curve (i.e., $s_{b\beta} = \{0, 1\}$), represented with black dots (\bullet), we find the three cases of equation (10).

between nodes v_a and w_α , expressed in the exponential form, is

$$P(v_a, w_\alpha | S, \Phi_{a\alpha}) = \exp \left\{ \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta} \left[D_{ab} E_{\alpha\beta} \ln \left(\frac{1 - P_e}{P_e} \right) + \ln \left(\frac{P_{a\alpha b \beta}}{\rho} \right) \right] + \ln(P_e \rho) \right\} \quad (12)$$

This is, the exponential of a weighted sum of structural and geometrical compatibilities between the pairwise relations emanating from nodes $v_a \in \mathbf{v}$ and $w_\alpha \in \mathbf{w}$. The weights $s_{b\beta}$ play the role of selecting the proper reference relation (w_α, w_β) that it is appropriate in gauging the likelihood of each observed relation (v_a, v_b).

These structural and geometrical coefficients (i.e., $D_{ab} E_{\alpha\beta} \ln \left(\frac{1 - P_e}{P_e} \right)$ and $\ln \left(\frac{P_{a\alpha b \beta}}{\rho} \right)$) are equivalent to the compatibility coefficients of the probabilistic relaxation approaches [23][24][29]. In this way, the structural and geometrical compatibilities are posed in a principled, balanced footing.

4. Expectation Maximization

The EM algorithm has been previously used by other authors to solve the Graph Matching problem [31] [32]. It is useful to find the parameters that maximize the expected log-likelihood for a mixture distribution. In our case,

we use it to find the correspondence indicators that maximize the expected log-likelihood of the observed relations, given the optimal alignments. From equations (3) and (4), we write,

$$S = \arg \max_{\hat{S}} \left\{ \max_{\hat{\Phi}_{a\alpha}} \left\{ \sum_{a \in \mathcal{I}} \ln \left[\sum_{\alpha \in \mathcal{J}} P \left(v_a, w_\alpha | \hat{S}, \hat{\Phi}_{a\alpha} \right) \right] \right\} \right\} \quad (13)$$

Dempster *et al.* [33] showed that maximizing the log-likelihood for a mixture distribution is equivalent at maximizing a weighted sum of log-likelihoods, where the weights are the missing data estimates. This is posed as an iterative estimation problem where the new parameters $S^{(n+1)}$ are updated so as to maximize an objective function depending on the previous parameters $S^{(n)}$. Then, the most recent available parameters $S^{(n)}$ are used to update the missing data estimates that, in turn, weigh the contributions of the log-likelihood functions. Accordingly, this utility measure is denoted

$$\Lambda (S^{(n+1)} | S^{(n)}) = \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} P (w_\alpha | v_a, S^{(n)}, \Phi_{a\alpha}) \ln P (v_a, w_\alpha | S^{(n+1)}, \Phi_{a\alpha}) \quad (14)$$

where the posterior probabilities of the missing data given the most recent available parameters $P (w_\alpha | v_a, S^{(n)}, \Phi_{a\alpha})$ weigh the contributions of the conditional log-likelihood terms.

The basic idea is to alternate between Expectation and Maximization steps until convergence is reached. The expectation step involves computing the a posterior probabilities of the missing data using the most recent available parameters. In the maximization phase, the parameters are updated in order to maximize the expected log-likelihood of the incomplete data.

4.1. Expectation

In the expectation step, the posterior probabilities of the missing data (i.e., the reference graph measurements w_α) are computed using the current parameter estimates $S^{(n)}$.

The posterior probabilities can be expressed in terms of conditional likelihoods, using the Bayes rule, in the following way

$$P (w_\alpha | v_a, S^{(n)}, \Phi_{a\alpha}) = \frac{P (v_a, w_\alpha | S^{(n)}, \Phi_{a\alpha})}{\sum_{\alpha'} P (v_a, w_{\alpha'} | S^{(n)}, \Phi_{a\alpha'})} \equiv R_{a\alpha}^{(n)} \quad (15)$$

Substituting our expression of the conditional likelihood of equation (12) into equation (15), the final expression for the posterior probabilities becomes,

$$R_{a\alpha}^{(n)} = \frac{\exp \left\{ \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} D_{ab} E_{\alpha\beta} \ln \left(\frac{1-P_e}{P_e} \right) + s_{b\beta}^{(n)} \ln \left(\frac{P_{a\alpha b\beta}}{\rho} \right) + \ln (P_e \rho) \right\}}{\sum_{\alpha' \in \mathcal{J}} \exp \left\{ \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} D_{ab} E_{\alpha'\beta} \ln \left(\frac{1-P_e}{P_e} \right) + s_{b\beta}^{(n)} \ln \left(\frac{P_{a\alpha' b\beta}}{\rho} \right) + \ln (P_e \rho) \right\}} \quad (16)$$

4.2. Maximization

Maximization is done in two steps. First, optimal registration parameters $\Phi_{a\alpha}$ are computed for each $P(v_a, w_\alpha | S, \Phi_{a\alpha})$. Last, global correspondence indicators are updated using the optimal $\Phi_{a\alpha}$'s.

4.2.1. Maximum Likelihood Affine Registration Parameters

We are interested in the registration parameters that lead to the maximum likelihood, given the current state of the correspondences $S^{(n)}$. In other words, the node descriptors \mathbf{x}_a and \mathbf{y}_α must be optimally registered before we can estimate the next correspondence indicators $S^{(n+1)}$. It is important that the registration do not modify the origins of the node descriptors, since these are the locations of the evaluated nodes v_a and w_α . As consequence, the registration parameters $\Phi_{a\alpha}$ are a 2×2 matrix of affine rotation and scaling parameters (without translation).

Therefore, we recover the Maximum Likelihood (ML) registration parameters $\Phi_{a\alpha}$, directly from equation (12). This is,

$$\begin{aligned} \Phi_{a\alpha} &= \arg \max_{\hat{\Phi}_{a\alpha}} \left\{ \ln P \left(v_a, w_\alpha | S^{(n)}, \hat{\Phi}_{a\alpha} \right) \right\} = \\ & \arg \max_{\hat{\Phi}_{a\alpha}} \left\{ \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} \ln \left(\frac{\hat{P}_{a\alpha b\beta}}{\rho} \right) + D_{ab} E_{\alpha\beta} s_{b\beta}^{(n)} \ln \left(\frac{1-P_e}{P_e} \right) + \ln (P_e \rho) \right\} \end{aligned} \quad (17)$$

We discard all the terms constant w.r.t the registration parameters and obtain the following equation

$$\Phi_{a\alpha} = \arg \max_{\hat{\Phi}_{a\alpha}} \left\{ \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} \ln \left(\hat{P}_{a\alpha b\beta} \right) \right\} \quad (18)$$

Now, we substitute the geometrical likelihood term by its expression of equation (8). We discard the constant terms of the multivariate Gaussian function and cancel the exponential and the logarithm functions, thus turning the maximization problem into a minimization one, by removing the minus sign of the exponential. We get the following expression

$$\Phi_{a\alpha} = \arg \min_{\hat{\Phi}_{a\alpha}} \left\{ \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} \left(\vec{x}_{ab} - \hat{\Phi}_{a\alpha} \vec{y}_{\alpha\beta} \right)^T \Sigma^{-1} \left(\vec{x}_{ab} - \hat{\Phi}_{a\alpha} \vec{y}_{\alpha\beta} \right) \right\} \quad (19)$$

We seek the matrix of affine parameters $\Phi_{a\alpha} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$ that minimize the weighted sum of squared Mahalanobis distances between the relative points \vec{x}_{ab} and the transformed relative points $\hat{\Phi}_{a\alpha} \vec{y}_{\alpha\beta}$. The coefficients $s_{b\beta}$ weigh the contribution of each pairwise distance in a way that the resulting registration will tend to minimize the distances between the relative positions of those v_b and w_β with the larger correspondence indicators.

Further developing equation (19), we obtain the following objective function

$$\mathcal{E} = \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta} \left[\left(x_{ab}^H - \phi_{11} y_{\alpha\beta}^H - \phi_{12} y_{\alpha\beta}^V \right)^2 / \sigma_H^2 + \left(x_{ab}^V - \phi_{21} y_{\alpha\beta}^H - \phi_{22} y_{\alpha\beta}^V \right)^2 / \sigma_V^2 \right] \quad (20)$$

where $\vec{x}_{ab} = (x_{ab}^H, x_{ab}^V)$ and $\vec{y}_{\alpha\beta} = (y_{\alpha\beta}^H, y_{\alpha\beta}^V)$ contain the horizontal and vertical coordinates of v_b and w_β relative to v_a and w_α , respectively, and $\Sigma = \text{diag}(\sigma_H^2, \sigma_V^2)$ is a diagonal matrix of variances.

Minimization of equation (20) w.r.t. the affine parameters ϕ_{ij} is done by solving the set of linear equations $\delta\mathcal{E}/\delta\phi_{ij} = 0$.

4.2.2. Maximum Likelihood Correspondence Indicators

One of the key points in our work is to approximate the solution of the graph matching problem by means of a succession of easier assignment problems. Following the dynamics of the EM algorithm, each one of these problems is posed using the most recent parameter estimates. As it is done in Graduated Assignment [25], we use the *Softassign* [26][27][28] to solve the assignment problems in a continuous way. The two main features of the *Softassign* are that, it allows to adjust the level of discretization of the solution by means of a control parameter and, it enforces two-way constraints

by incorporating a method discovered by Sinkhorn [26]. The two-way constraints guarantee that one node of the observed graph can only be assigned to one node of the reference graph, and vice versa. In the case of continuous assignments, this is accomplished by applying alternative row and column normalizations (considering the correspondence variable S as a matrix). Moreover, Softassign allows us to smoothly detect outliers in both sides of the assignment (see section 5).

According to the EM development, we compute the correspondence indicators $S^{(n+1)}$ that maximize the utility measure of equation (14). In our case, this equals to

$$S^{(n+1)} = \arg \max_{\hat{S}} \left\{ \Lambda \left(\hat{S} | S^{(n)} \right) \right\} = \arg \max_{\hat{S}} \left\{ \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} R_{a\alpha}^{(n)} \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} \hat{s}_{b\beta} \left[D_{ab} E_{\alpha\beta} \ln \left(\frac{1-P_e}{P_e} \right) + \ln \left(\frac{P_{a\alpha b\beta}}{\rho} \right) \right] + \ln (P_e \rho) \right\} \quad (21)$$

where $R_{a\alpha}^{(n)}$ are the missing data estimates.

Rearranging, and dropping the terms constant w.r.t the correspondence indicators, we obtain

$$S^{(n+1)} = \arg \max_{\hat{S}} \left\{ \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} \hat{s}_{b\beta} \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} R_{a\alpha}^{(n)} \left[D_{ab} E_{\alpha\beta} \ln \left(\frac{1-P_e}{P_e} \right) + \ln \left(\frac{P_{a\alpha b\beta}}{\rho} \right) \right] \right\} \quad (22)$$

which, as it can be seen in the following expression, presents the same form as an assignment problem [25]

$$S^{(n+1)} = \arg \max_{\hat{S}} \left\{ \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} \hat{s}_{b\beta} Q_{b\beta}^{(n)} \right\} \quad (23)$$

where the $Q_{b\beta}^{(n)}$ are the benefit coefficients for each assignment.

Softassign computes the correspondence indicators in two steps. First, the correspondence indicators are updated with the exponentials of the benefit coefficients

$$s_{b\beta} = \exp (\mu Q_{b\beta}) \quad (24)$$

where μ is a control parameter. Second, two-way constraints are imposed by alternatively normalizing across rows and columns the matrix of exponentiated benefits. This is known as the *Sinkhorn normalization* and it is applied either until convergence of the normalized matrix or a predefined number of times.

Note that, the correspondence indicators $s_{b\beta}$ will tend to discrete values ($s_{b\beta} = \{0, 1\}$) as the control parameter μ of equation (24) approaches to ∞ .

We also apply the Sinkhorn normalization to the posterior probabilities of the missing data so that they are more correlated with the correspondence indicators.

Since the matrices may not be square (i.e., different number of nodes in the observed and reference graphs), in order to fulfill the law of total probability, we complete the Sinkhorn normalization process with a normalization by rows.

Figure 6 shows the pseudo-code implementation of our method.

```

while  $\mu \leq \mu_f$  do
  while ( $S^{(n)}$  does not converge)  $\wedge$  (iterations  $\leq$  max) do
    # ML Affine Parameters
     $\Phi_{a\alpha} \leftarrow \arg \max_{\hat{\Phi}_{a\alpha}} \left\{ \ln P \left( v_a, w_\alpha | S^{(n)}, \hat{\Phi}_{a\alpha} \right) \right\}, \forall a \in \mathcal{I} \alpha \in \mathcal{J}$ 

    # Expectation
     $R_{a\alpha}^{(n)} \leftarrow \frac{P(v_a, w_\alpha | S^{(n)}, \Phi_{a\alpha})}{\sum_{\alpha'} P(v_a, w_{\alpha'} | S^{(n)}, \Phi_{a\alpha'})}, \forall a \in \mathcal{I} \alpha \in \mathcal{J}$ 

     $R^{(n)} \leftarrow \text{Sinkhorn} (R^{(n)})$ 

    # Maximization
     $Q_{b\beta}^{(n)} \leftarrow \sum_{a,\alpha} R_{a\alpha}^{(n)} \left[ D_{ab} E_{\alpha\beta} \ln \left( \frac{1-P_e}{P_e} \right) + \ln \left( \frac{P_{a\alpha b\beta}}{\rho} \right) \right], \forall b \in \mathcal{I} \beta \in \mathcal{J}$ 

     $s_{b\beta}^{(n+1)} \leftarrow \exp \left( \mu Q_{b\beta}^{(n)} \right), \forall b \in \mathcal{I} \beta \in \mathcal{J}$ 
     $S^{(n+1)} \leftarrow \text{Sinkhorn} (S^{(n+1)})$ 

     $S^{(n)} \leftarrow S^{(n+1)}$ 
  end
   $\mu \leftarrow \mu \times (1 + \epsilon)$ 
end

```

Figure 6: The outer loop gradually increase the Softassign parameter μ , thereby pushing from continuous to discrete solutions. This reduces the chances of getting trapped in local minima [25]. The body contains the pseudo-code of the E and M steps. Each iteration of the inner loop performs one step of the EM algorithm.

5. Outlier Detection

A node in one graph is considered to be an outlier if it has no correspondent node in the other graph.

Consider, for example, the case of figure 3. The rightmost nodes in the right image are outliers originated from the detection of features in the non-overlapping parts of the images. On the other hand, the unmatched nodes in the overlapping parts are outliers originated by differences in the feature detection patterns.

Outliers can dramatically affect the performance of a matching and therefore, it is important to develop techniques aimed at minimizing their influence [35].

According to our purposes, a node $v_b \in \mathbf{v}$ (or $w_\beta \in \mathbf{w}$) will be considered an outlier to the extent that there is no node $w_\beta, \forall \beta \in \mathcal{J}$ (or $v_b, \forall b \in \mathcal{I}$) which presents a matching benefit $Q_{b\beta}^{(n)}$ above a given threshold.

From equations (22) and (23), the benefit values have the following expression

$$Q_{b\beta}^{(n)} = \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} R_{a\alpha}^{(n)} \left[D_{ab} E_{\alpha\beta} \ln \left(\frac{1 - P_e}{P_e} \right) + \ln \left(\frac{P_{a\alpha b\beta}}{\rho} \right) \right] \quad (25)$$

Note that, the value of ρ controls whether the geometrical compatibility term contributes either positively (i.e., $\rho < P_{a\alpha b\beta}$) or negatively (i.e., $\rho > P_{a\alpha b\beta}$) to the benefit measure.

We model the outlier detection process as an assignment to (or from) the *null* node. We consider that the null node has no edges at all and, all the geometrical terms $P_{a\alpha b\beta}$ involving it are equal to ρ . Under these considerations, the benefit values of equation (25) corresponding to the null assignments are equal to zero. We therefore create an augmented benefit matrix $\tilde{Q}^{(n)}$ by adding to $Q^{(n)}$ an extra row and column of zeros. This extra row and column represent the benefits of the *null* assignments (i.e., $Q_{b\emptyset}, \forall b \in \mathcal{I}$ and $Q_{\emptyset\beta}, \forall \beta \in \mathcal{J}$).

We apply the Softassign (exponentiation and Sinkhorn normalization) to the augmented benefit matrix $\tilde{Q}^{(n)}$. When performing Sinkhorn normalization we keep in mind that the *null* assignments are special cases that only satisfy one-way constraints. This is, there may be multiple assignments to *null* in both graphs. Finally, the extra row and column are removed leading to the resulting matrix of correspondence parameters $S^{(n+1)}$. This process is illustrated in figure 7.

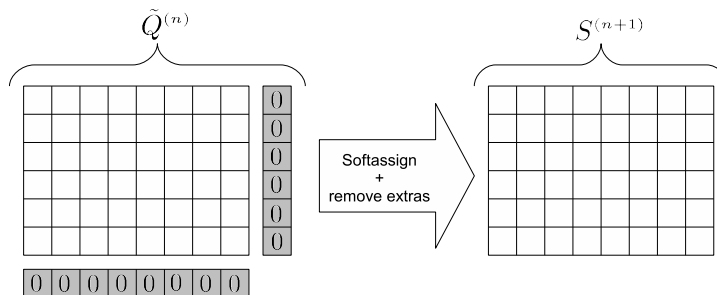


Figure 7: The Softassign and outlier detection process.

As the control parameter μ of the Softassing increases, the rows and columns of $S^{(n+1)}$ associated to the outlier nodes, tend to zero. This fact reduces the influence of these nodes in the maximization phases of the next iteration that, in turn, lead to even lower benefits, and so on.

It is now turn to define the value of the constant ρ . Since ρ is to be compared with $P_{\alpha\alpha\beta}$, it is convenient to define it in terms of a multivariate Gaussian measurement of a distance threshold. This is,

$$\rho = \frac{1}{2\pi|\Sigma|^{1/2}} \exp \left[-\frac{1}{2} \vec{d}^T \Sigma^{-1} \vec{d} \right] \quad (26)$$

where $\Sigma = \text{diag}(\sigma_H^2, \sigma_V^2)$ is the same diagonal variance matrix as we use in equation (8) and $\vec{d} = (d^H, d^V)$ is a column vector with the horizontal and vertical thresholding distances.

Cancelling the constant terms in the numerator and denominator of the geometrical compatibility term and expressing the thresholding distances as a quantity proportional to the standard deviations of the data, (i.e., $\vec{d} = (N\sigma_H, N\sigma_V)$), the expression of ρ to be compared with $P_{\alpha\alpha\beta}$ is

$$\rho = \exp \left\{ -\frac{1}{2} \left[\left(\frac{N\sigma_H}{\sigma_H} \right)^2 + \left(\frac{N\sigma_V}{\sigma_V} \right)^2 \right] \right\} = \exp(-N^2) \quad (27)$$

So, we define ρ as a function of the number N of standard deviations permitted in the registration errors, in order to consider a plausible correspondence.

6. Related Methods

In this section we briefly overview to existing methods related to ours that are used in the experiments.

Cross and Hancock perform graph matching with a *Dual-Step* EM algorithm [31]. They update alignment parameters according to the following rule.

$$\Phi^{(n+1)} = \arg \max_{\hat{\Phi}} \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} P(w_\alpha | v_a, \Phi^{(n)}) \zeta_{a\alpha}(f^{(n)}) \ln P(v_a, w_\alpha | \hat{\Phi}) \quad (28)$$

where $\zeta_{a\alpha}(f^{(n)})$ is the expected value for the assignment probabilities regarding the structural consistence of the mappings defined by the correspondences $f^{(n)}$, $P(v_a, w_\alpha | \hat{\Phi})$ is a multivariate Gaussian measurement of the point position errors given the alignment parameters and, $P(w_\alpha | v_a, \Phi^{(n)})$ is the posterior probability term computed in the E-step.

Correspondence variables are estimated in a discrete way with the following Maximum a Posteriori rule.

$$f^{(n+1)}(a) = \arg \max_{\alpha} P(w_\alpha | v_a, \Phi^{(n)}) \zeta_{a\alpha}(f^{(n)}) \quad (29)$$

Luo and Hancock's present a *Unified* framework for alignment and correspondence [34] that maximizes a cross-entropy measure in an EM-like fashion. They estimate Maximum Likelihood alignment and correspondence parameters according to the following rules.

$$\Phi^{(n+1)} = \arg \max_{\hat{\Phi}} \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} Q_{a\alpha}^{(n)} \ln P(v_a, w_\alpha | \hat{\Phi}) \quad (30)$$

and

$$S^{(n+1)} = \arg \max_{\hat{S}} \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} P_{a\alpha}^{(n)} \ln Q(v_a, w_\alpha | \hat{S}) \quad (31)$$

where $Q_{a\alpha}^{(n)}$ and $P_{a\alpha}^{(n)}$ are the posterior probability terms according to structural and positional observations and $P(\bullet)$ and $Q(\bullet)$ are the corresponding probability density functions. Correspondence variables are updated in a discrete way.

7. Experiments and Results

We assess the performance of our method in terms of registration accuracy and recognition ability.

We have roughly set the value of $\rho = \exp(-1.6^2)$ for our method in all the experiments. We have not found this parameter to be specially application dependant since this same value has offered a fair performance in all the variety of experiments presented.

The parameters for the rest of the methods have been likewise set so as to show a good performance. We have tried to be as efficient as possible in the implementations of all the methods. Unless otherwise noted, all the computational time results refer to Matlab® run-times. All the experiments have been conducted on an Intel® Xeon® CPU E5310 at 1.60GHz.

7.1. Registration Experiments

These experiments are aimed at testing both the ability of our model to discriminate correct matching hypothesis as well as the accuracy of the optimization method to locate them.

Performance is assessed by either the correct correspondence rate or the *mean projection error* depending on whether the graphs are synthetically generated (with known ground truth correspondences) or extracted from real images (with known ground truth homography). We compare to other graph matching methods as well as to known point-set registration methods and outlier rejectors.

The parameter P_e is more application dependant than ρ since it establishes the scale of the structural contribution of our model which is to be added to the geometric contribution in order to set up the consistency measure of equation (25). Specifically, the value of the structural contribution depends on this scale parameter as well as the mean node degree (i.e., the mean number of incident edges upon each node).

All the graphs used in this section have been generated by means of Delaunay triangulations over point-sets, where each point has been assigned to a node. We have conducted matching experiments on randomly generated graphs and have experimentally found that a value of $P_e = 0.03$ performs well for this type of graphs. Therefore, we have used this value for our method in all the experiments in this section.

This section is divided as follows. In sections 7.1.1 and 7.1.2 we use synthetic graphs to evaluate specific aspects of our model. In section 7.1.3

we use real images.

7.1.1. Synthetic Non-Rigid Deformations

In the next set of experiments we evaluate the matching ability in the presence of non-rigid deformations. We have matched randomly generated patterns of 15 points with deformed versions of themselves. Deformations have been introduced by applying random Gaussian noise to the coordinate positions of the points.

In the synthetic experiments we assess the performance of each method through the correct correspondence rate. To see how this performance measure is related to our model we measure the ratio $\mathcal{L}_{EM-Soft}/\mathcal{L}_{gtr}$, between the value of the log-likelihood function at the solution found by the EM algorithm and that at the ground truth matching.

From equation (4) the expression of log-likelihood function according to our model is the following.

$$\mathcal{L} = \sum_{a \in \mathcal{I}} \ln \left[\sum_{\alpha \in \mathcal{J}} P \left(v_a, w_\alpha | \hat{S}, \hat{\Phi}_{a\alpha} \right) \right] \quad (32)$$

Figure 8 shows that even though there is an increasing trend in the disagreement between the model hypothesis and the established ground truth as the deformation increases, such a disagreement remains close to the optimum value of 1 for deformations up to 20%.

We have compared the correct correspondence rates of our method (*EM-Soft*) to that of the graph matching + point-set alignment methods *Dual-Step* and *Unified*. The *Dual-Step* has been implemented with an affine geometrical model as well as the capability of detecting outliers. Such an outliers-detection capability increases considerably its required computational time but, evaluating the performance of this feature is an important aspect in our experiments. All the approaches have been initialized with the resulting correspondences of a simple nearest neighbour association. Figure 9 shows the correct correspondence rates with respect to the amount of noise. The mean computational times are: 14.6 sec. (*EM-Soft*), 124.5 sec. (*Dual-Step*) and 0.91 sec. (*Unified*).

The computational time obtained with a C implementation of our method is 0.24 sec.

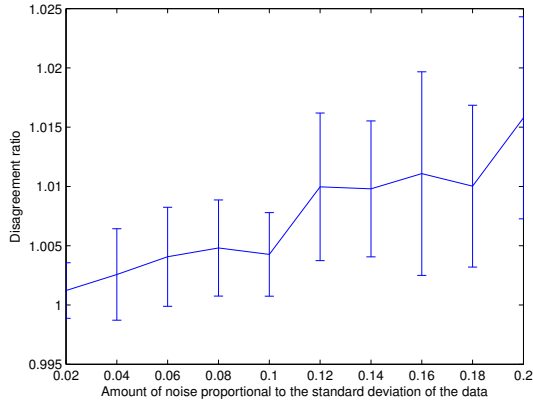


Figure 8: Ratio of the log-likelihood of the suboptimal solution found by our method to that of the ground truth solution, according to our model. As the deformation increases the likelihood of the ground truth solution falls below other (partially incorrect) solutions. Each location on the plots is the mean of 25 experiments (5 random patterns of points by 5 random deformations).

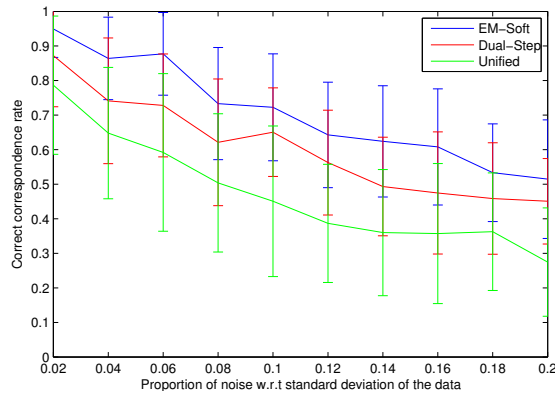


Figure 9: Correct correspondence rate with respect to the amount of noise in the point positions (expressed proportionally to the variance of the data). Each location on the plots is the mean of 25 experiments (5 random patterns of points by 5 random deformations).

7.1.2. Synthetic Addition of Random Points

The next set of experiments evaluates the matching ability in the presence of outliers. We have randomly added outlying points (with no correspondence in the other side) to *both* synthetic patterns of 15 points. We have preserved a proportion of ground-level non-rigid noise of 0.02 between the inliers of both patterns. In order to contribute positively to the correct correspondence rate,

outliers must not be matched to any point while, inliers must be assigned to its corresponding counterpart. The approaches compared in this experiment are those with explicit outlier detection mechanisms. These are *RANSAC* (affine) [10], Graph Transformation Matching (*GTM*) [11] and *Dual-Step* [31].

GTM is a powerful outlier rejector based on a graph transformation that holds a very intuitive idea. We use the same strategy as in [11] consisting in using k -NN graphs with $k = 5$ instead of Delaunay triangulations in order to present the results for the *GTM* method. However, similar results are obtained using Delaunay triangulations.

All the methods have been initialized with the resulting correspondences of a simple nearest neighbour association. Figure 10 shows the correct correspondence rate with respect to the number of outliers and figure 11 shows the computational times.

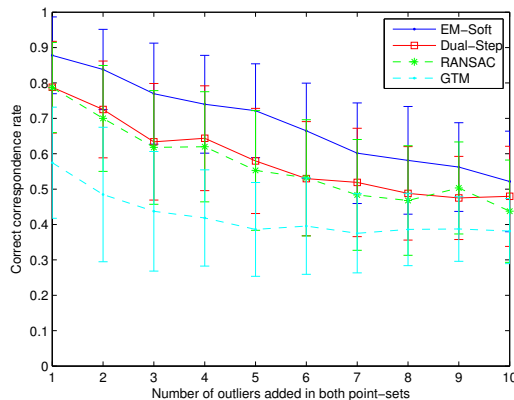


Figure 10: Correct correspondence rate with respect to the number of outliers in each side. Each location is the mean of 125 experiments (5 random inlier patterns by 5 random outlier patterns by 5 random ground-level non-rigid noise).

7.1.3. Real Images

We have performed registration experiments on real images from the database in ref. [9]. Point-sets have been extracted with the Harris operator [1]. Each pair of images shows two scenes related by either a zoom or a zoom + rotation. They belong to the classes *Resid*, *Boat*, *New York* and *East Park*. All the approaches use the same parameters as in the previous section. Figures 3 and 12 show the resulting correspondences found by our method as well as the tentative correspondences used as starting point.

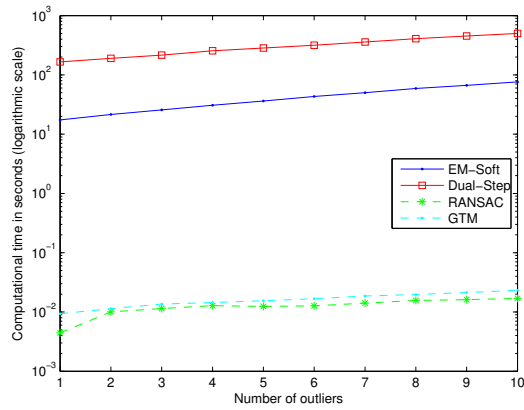
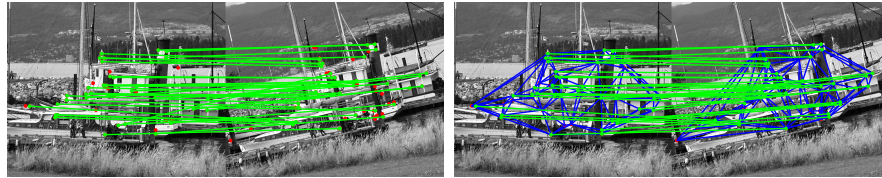
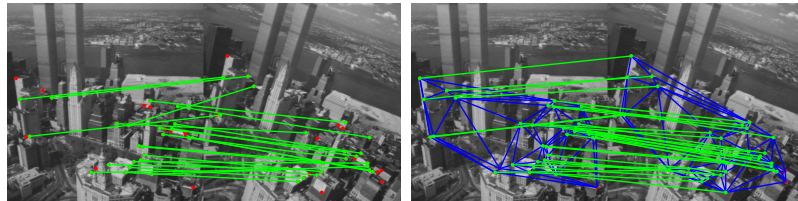


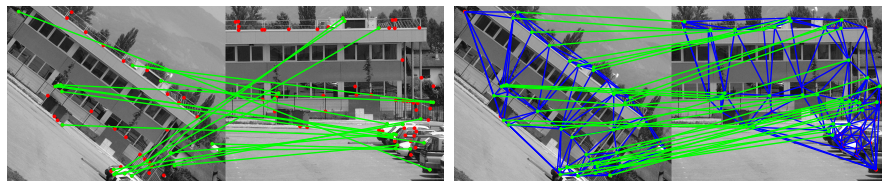
Figure 11: Plots of the computational times with respect to the number of outliers. The time (vertical) axis is in logarithmic scale.



(a) Boat



(b) New York



(c) East Park

Figure 12: Right column shows the results of our method using the matching by correlation results (left column) as starting point.

We have compared all the methods with outlier detection capabilities of the previous section. The graph-based approaches have been initialized with the matching by correlation results (*Corr*). We have used the matching by correlation implementation found in ref. [36]. In order to avoid the sparsity problem mentioned in figure 2, we have applied *ICP* [12] to the correlation results, as a previous step to the outlier rejectors *RANSAC* and *GTM*.

From the resulting correspondences, we have estimated the corresponding homographies with the DLT algorithm [37]. Since the ground truth homography between each pair of images is available, we have measured the mean projection error (MPE) of the feature-points in the origin images. Table 1 shows the mean projection errors as well as the proportion of matched points in the origin images. Table 2 shows the computational times in seconds. Table 4 shows the computational times of the Matlab and C implementations of our method.

In order to show how methods benefit from outlier rejection in a real world application, we have repeated the above experiments using the *Unified* method and the pure structural method Graduated Assignment (*GradAssign*) [25], both without explicit outlier rejection capabilities. We have also added modified versions of *EM-Soft* and *Dual-Step* so that outlier rejection is disabled (marked with an asterisk). Table 3 shows the results.

7.2. Recognition Experiments

In this section we assess the recognition ability of the underlying model used in our graph matching method in a series of shape retrieval experiments on the GREC database [38] and a database of 25 shapes. In these experiments, the structure of the graphs has been given rise by the morphology of the objects.

Due to numerical reasons, lower values of P_e are needed in the case of this morphologically-induced graphs than in the case of Delaunay triangulations. This is because in this case, resulting graphs are sparser and therefore structural contributions under equation (25) need to be amplified so as to play a role comparable to the geometric contributions. We have used the first 5 graphs from each class of the GREC database in order to tune the parameters of all the methods. Due to the similar nature of the graphs in both databases and to the lack of examples in the 25-shapes database in order to perform training, we have used the same parameters in both databases. We have used a value of $P_e = 3 \cdot 10^{-4}$ for our method.

Method	Resid		Boat		New York		East Park	
	MPE	%	MPE	%	MPE	%	MPE	%
<i>Corr</i>	835	27	24.5	76	31	67	463	43
<i>ICP</i>	40.3	87	21	100	19.4	100	88	100
<i>EM-Soft</i>	1.5	69	0.68	72	0.69	91	1.13	75
<i>Dual-Step</i>	1.3	72	1.7	62	0.7	91	153	25
<i>ICP+RANSAC</i>	12.3	54	10.7	64	10.9	76	98	41
<i>ICP+GTM</i>	32.5	61	10.5	70	2.9	70	327	45

Table 1: Mean Projection Error (MPE) and percentage of matched points in the origin images (%).

Method	Resid	Boat	New York	East Park
<i>Corr</i>	0.54	0.55	0.26	0.55
<i>ICP</i>	0.22	0.18	0.13	0.2
<i>EM-Soft</i>	378	449	73	438
<i>Dual-Step</i>	3616	3794	1429	3027
<i>ICP+RANSAC</i>	0.29	0.24	0.18	0.3
<i>ICP+GTM</i>	0.23	0.25	0.17	0.27

Table 2: Computational times (in seconds). The number of points of the origin (N_1) and destination (N_2) images in each case are: *Resid* $N_1 = 55, N_2 = 48$; *Boat* $N_1 = 50, N_2 = 61$; *New York* $N_1 = 34, N_2 = 34$; *East Park* $N_1 = 44, N_2 = 67$.

Method	Resid		Boat		New York		East Park	
	MPE	%	MPE	%	MPE	%	MPE	%
<i>EM-Soft*</i>	2619	87	25.3	100	23.9	100	56.8	95
<i>Dual-Step*</i>	26.2	83	19.5	100	1.15	91	332	100
<i>Unified</i>	39.8	69	12.3	86	3.04	88	1104	75
<i>GradAssig</i>	174	85	60.8	100	14.8	94	1716	100

Table 3: Mean Projection Error (MPE) and percentage of matched points (%) obtained without outlier rejection mechanisms.

Method	Resid	Boat	New York	East Park
<i>EM-Soft</i> (Matlab)	378	449	73	438
<i>EM-Soft</i> (C)	15.5	19.5	2.1	19.1

Table 4: Computational times of the Matlab and C implementations of our method.

Given a query graph G , we compute its similarity to a database graph H using the following measure

$$\mathcal{F}_{GH} = \frac{\max_{\hat{S}} \mathcal{F}(G, H; \hat{S})}{\max(\mathcal{F}_{GG}, \mathcal{F}_{HH})} \quad (33)$$

where, $\mathcal{F}(G, H; \hat{S}) = \ln P(G|\hat{S}, \Phi)$ is the incomplete log-likelihood of the observed graph G , assuming H as the missing data graph; $\mathcal{F}_{GG} = \mathcal{F}(G, G; I_G)$ and $\mathcal{F}_{HH} = \mathcal{F}(H, H; I_H)$, being I_G and I_H the identity correspondence indicators defining self-matchings. This results in a normalized measure $\mathcal{F}_{GH} \in [0, 1]$ that equals to one in the case of a self-matching between two identical graphs and, moves towards zero as they become different.

Note that the maximization in the numerator of equation (33) has the same form as the log-likelihood maximization of equation (13) performed by our EM algorithm (section 4).

Performance is assessed through precision-recall plots. We compute the pairwise similarities between all the graphs in the database thus obtaining, for each query graph, a list of retrievals ordered by similarity. Suppose that our database contains C classes with N graphs each. We can define a retrieval of depth r as the first r graphs from each ordered list. Note that the number of elements retrieved by a such an operation is rCN .

Precision is then defined as the fraction of retrieved graphs that are relevant in a retrieval of depth r . This is,

$$\text{precision} = \frac{\#\text{relevant}(r)}{rCN} \quad (34)$$

where $\#\text{relevant}(r)$ is the number of retrieved graphs that agree with the class of their respective queries, in a retrieval of depth r .

Recall is defined as the fraction of the relevant graphs that are successfully retrieved by retrieval of depth r . This is,

$$\text{recall} = \frac{\#\text{relevant}(r)}{CN^2} \quad (35)$$

where CN^2 is the maximum number of relevant graphs possible.

Precision-recall plots are generated by varying r in the range $[1 \dots CN]$.

7.2.1. GREC Graphs

We have performed retrieval experiments on the GREC subset of the IAM Graph Database Repository [38]. This subset is composed by 22 classes of 25 graphs each. Figure 13 shows an example graph of each class. Some

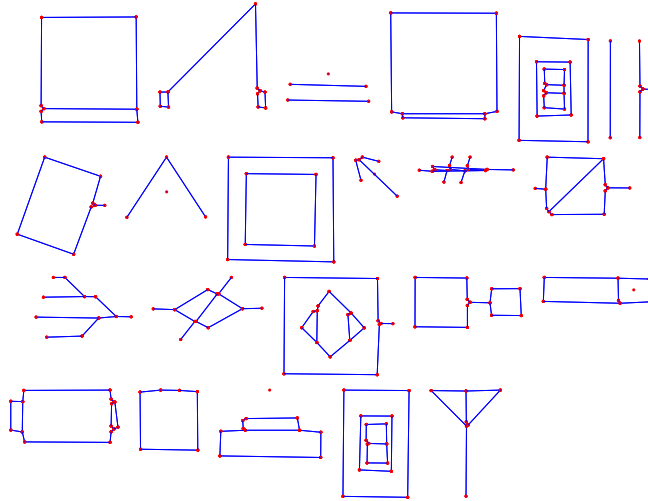


Figure 13: An example graph of each class of the GREC database [38]. Nodes are represented as red dots, while edges as blue lines.

classes show considerable inter-class similarities as well as significant intra-class variations such as missing or extra nodes, non-rigid deformations, scale differences and structural disruptions. See for example, the graphs in figure 14.

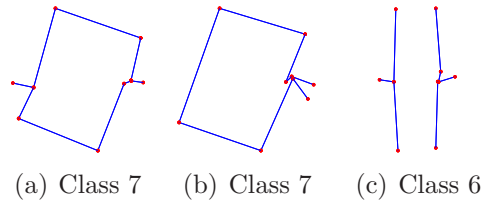


Figure 14: Compared in an affine-invariant way, graphs 14(a) and 14(c) show a similar node-set arrangement, although they are from different classes. They present however, slight differences between their structure. On the other hand, although graphs 14(a) and 14(b) are from the same class, we can see missing and extra nodes with respect to each other, while still having some differences between their structure. With these considerations, classification is not straightforward.

We have compared our method (*EM-Soft*) to the purely structural method *GradAssig* and the geometric + structural methods *Dual-Step* [31] and *Unified* [34]. We have included two additional pure geometric methods in order to provide evidence of the benefits of the combined methods. On one hand, we have used our method with an ambiguous structural model (i.e., $P_e = 0.5$). On the other hand, we have implemented a point-set registration algorithm (*EM-reg*) using the following EM update rule.

$$\Phi^{(n+1)} = \arg \max_{\Phi} \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} P(w_\alpha | v_a, \Phi^{(n)}) \ln P(v_a, w_\alpha | \hat{\Phi}) \quad (36)$$

where $P(v_a, w_\alpha | \hat{\Phi})$ is a multivariate Gaussian function of the point-position errors given the alignment parameters.

For each method, we have used the equivalent analog of the normalized similarity measure of equation (33) according to their models. All the approaches have been initialized by the tentative correspondences found as explained in appendix A. The methods that do not use correspondence parameters have been initialized by the alignment parameters that minimize the pairwise point-position errors according to the aforementioned correspondences.

Figure 15 shows the precision-recall plots obtained by varying the depth of the retrieval from 1 to 550 (the total number of graphs in the database).

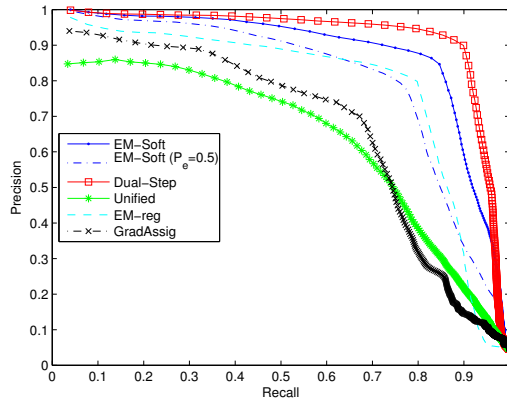


Figure 15: Precision-recall plots in the GREC database.

7.2.2. 25 Shapes Database

We have performed retrieval experiments on the database of 25 binary shape images of figure 16. Our aim here is to evaluate the recognition abil-



Figure 16: This database is divided into 6 classes: *Shark* (5 instances), *Plane* (4 instances), *Gen* (3 instances), *Hand* (5 instances), *Rabbit* (4 instances) and *Tool* (4 instances).

ities of several general-purpose graph matching approaches. Therefore, we have not used databases containing more specific types of deformations such as articulations because of the limitations imposed by the affine model assumptions.

We have used the skeleton pruning approach by Bai and Latecki [39] in order to obtain the skeletal representations. Graphs have been generated by placing the nodes at the end and intersection skeletal points and, the edges so as to fit to the rest of body of the skeleton. Figure 17 shows the graphs extracted from the above database.

All the approaches have been initialized with the tentative correspondences found as explained in appendix A.

We have implemented an affine-invariant template matching method in order to evaluate the benefits of using the structural abstractions instead of using directly the binary images. We evaluate the similarity between two registered binary images on the basis of their overlapping shape areas. Affine registration of the binary images is performed according to the tentative correspondences found by the method in appendix A.

Figure 18 shows the precision-recall plots of the graph matching approaches *EM-Soft*, *Dual-Step* and *Unified* and, the affine-invariant template matching (*TM*).

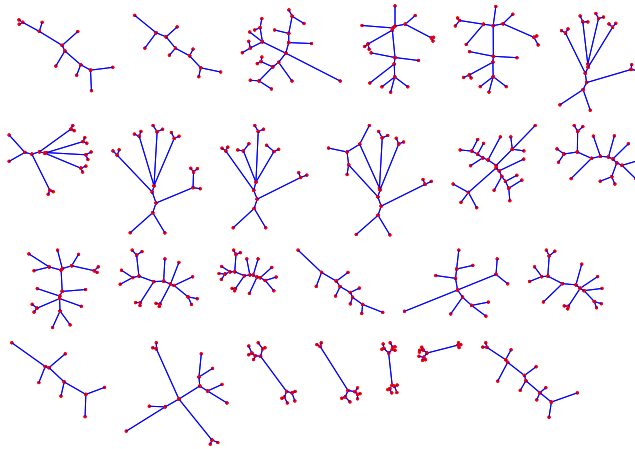


Figure 17: Graphs generated from the skeletons of the 25 shapes of figure 16.

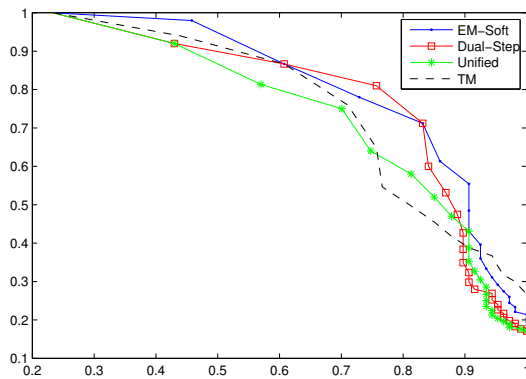


Figure 18: Precision-recall plots in the 25-shapes database.

8. Discussion

In the matching experiments under non-rigid deformations our method has shown to be the most effective among the compared graph-matching methods. Moreover, it shows a computational time in the typical range of the graph-matching methods. *Dual-Step* obtains a higher correct correspondence rate than *Unified*. However, its computational time is higher as well.

The matching experiments in the presence of outliers show that our method outperforms the compared ones. *Dual-Step* performs as effectively as *RANSAC*. Moreover, while outlier rejectors are specifically designed for these type of experiments, *Dual-Step* has a wider applicability.

The matching experiments on real images show that our method performs generally better than the others. *Dual-Step* performs, in most cases, similarly as ours but with higher computational times. It is worth mentioning that the considerable computational times required by *Dual-Step* are mainly due to the bottleneck that represents its outlier detection scheme. The graph-matching methods generally find more dense correspondence-sets. The ensembles *ICP+RANSAC* and *ICP+GTM* do not perform as effectively as the graph-matching methods but they do it faster.

Furthermore, we show how different methods benefit from outlier rejection in a real world application.

The efficiency shown by the C implementation of our method suggests that, while the outlier rejectors are appropriate for a regular real-time operation, it is feasible to use our method in specific moments when more effectiveness is required.

The dictionary-based structural model of the *Dual-Step* [31] has demonstrated to be the most effective in the retrieval experiments on the GREC database. Our method shows a performance decrease with respect to *Dual-Step*. *Unified* is unable to deal with the type of graphs used in this experiment.

Neither the pure geometric methods nor the pure structural (*GradAsig*) are as accurate as *Dual-Step* and *EM-Soft* in the precision-recall scores. This demonstrates the benefits of combining both sources of information as opposed to using them separately. Particularly revealing of this fact is the comparison between the two versions of our method.

In the 25-shapes database the proposed method and the *Dual-Step* obtain similar scores. The affine-invariant template matching method only retrieves correctly the most similar instances of each class. As we increment the depth of the retrieval and more significant intra-class variations appear, the direct comparison of templates experiments a decrease in performance with respect to our structural approach. This shows the benefits of using structural representations as opposed to template-based strategies in the present application. The limitations of the affine model assumptions prevents us from using shape databases presenting further deformations such as larger articulations.

9. Conclusions

We have presented a graph matching method aimed at solving the point-set correspondence problem. Our model accounts for relative structural and

geometrical measurements which keep parallelism with the compatibility coefficients of the Probabilistic Relaxation approaches. This contrasts with other approaches that use absolute coordinate positions [31][34]. Unlike other approaches [31][34][32][30], our underlying correspondence variable is continuous. To that end, we use Softassign to solve the individual assignment problems thus, enforcing two-way constraints as well as being able to control the level of discretization of the solution. Moreover, gradually pushing from continuous to discrete states reduces the chances of getting trapped in local minima [25]. We develop mechanisms to smoothly detect and remove outliers.

In contrast to other approaches such as *Unified* and *Dual-Step*, the proposed approach has the distinguished properties that it uses Softassign to estimate the continuous correspondence indicators as well as it is based on a model of relational geometrical measurements. Such properties demonstrate to confer the proposed approach a better performance in most of the experiments presented.

Our method is controlled by two parameters, namely, an outlying threshold probability ρ and a probability of structural error P_e . We have not found any particular dependence of the parameter ρ to a specific application and hence, we have used the same value in all the experiments. On the contrary, the parameter P_e scales the contribution of the structural component of our model which is to be compared to the geometric part. Since the value of the structural contribution depends on both this scaling factor and the mean degree of the nodes in the graphs, we have found a dependence of this parameter to the type of graphs used. Therefore, we have used two different values for graphs generated from Delaunay triangulations and morphologically-induced graphs. As far as we know, this dependence does not go further than to the types of graph-representations used.

Appendix A

The following is an adaptation of the Belongie *et al.*'s shape contexts [40] in order to match point-sets regardless of their orientations. Given two point-sets (in our case, the positions of the nodes), $\mathbf{p} = \{\vec{p}_a\}, \forall a \in \mathcal{I}$ and $\mathbf{q} = \{\vec{q}_\alpha\}, \forall \alpha \in \mathcal{J}$, we arbitrarily choose one of them (e.g., \mathbf{q}) and create M subsets at different orientations by applying M rigid-body rotations uniformly

distributed along the range $[0 \dots 360]$ degrees². We therefore obtain the subsets $\mathbf{q}^m, m \in [1 \dots M]$, corresponding to M different orientations of \mathbf{q} . Next, we compute the shape contexts, a kind of descriptor (log-polar histograms) that, for each point, encodes how the rest of the points are distributed around it. We do it for the $M + 1$ point-sets \mathbf{p} and $\mathbf{q}^m, m \in [1 \dots M]$. Using the χ^2 test statistic between the shape contexts [40], we compute the M matrices of matching costs C^m . Thus, $C_{a\alpha}^m$ indicates the cost of matching the point $\vec{p}_a \in \mathbf{p}$ to the point $\vec{q}_\alpha^m \in \mathbf{q}^m$. By applying the Hungarian algorithm [41] to each C^m , we compute the optimal assignments $f^m : \mathcal{I} \rightarrow \mathcal{J}$ from the points in \mathbf{p} to those in each one of the \mathbf{q}^m 's. We choose as the prevailing orientation m , the one with the minimum matching cost, i.e., $m = \arg \min_{\hat{m}} \left\{ \sum C_{a, f^{\hat{m}}(a)}^{\hat{m}} \right\}$ and, the resulting correspondences are those defined by f^m .

Acknowledgements

This research is supported by Consolider Ingenio 2010 (CSD2007-00018), by the CICYT (DPI 2010-17112) and by project DPI-2010-18449.

References

- [1] C. Harris, M. Stephens, A combined corner and edge detection, in: Proceedings of The Fourth Alvey Vision Conference, 1988, pp. 147–151.
- [2] J. Shi, C. Tomasi, Good features to track, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94), Seattle, 1994.
- [3] C. Tomasi, T. Kanade, Detection and tracking of point features, Tech. Rep. CMU-CS-91-132, Carnegie Mellon University (April 1991).
- [4] L. Kitchen, A. Rosenfeld, Gray-level corner detection, Pattern Recognition Letters 1 (2) (1982) 95–102.
- [5] L. Koenderink, W. Richards, Two-dimensional curvature operators, Journal of the Optical Society of America: Series A 5 (7) (1988) 1136–1141.

²We use $M = 12$

- [6] W. Han, M. Brady, Real-time corner detection algorithm for motion estimation, *Image and Vision Computing* (1995) 695–703.
- [7] E. Rosten, T. Drummond, Machine learning for high-speed corner detection, in: *In European Conference on Computer Vision*, Vol. 1, 2006, pp. 430–443.
- [8] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (10) (2005) 1615–1630.
- [9] <http://www.featurespace.org/>.
- [10] M. A. Fischler, R. C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM* 24 (6) (1981) 381–395.
- [11] W. Aguilar, Y. Frauel, F. Escolano, M. E. Martinez-Perez, A robust graph transformation matching for non-rigid registration, *Image and Vision Computing* 27 (2009) 897–910.
- [12] Z. Zhang, Iterative point matching for registration of free-form curves and surfaces, *International Journal of Computer Vision* 13 (2) (1994) 119–152.
- [13] R. Wilson, A. Cross, E. Hancock, Structural matching with active triangulations, *Computer Vision and Image Understanding* 72 (1) (1998) 21–38.
- [14] H. Blum, A transformation for extracting new descriptors of shape, in: *Models for the Perception of Speech and Visual Form*, MIT Press, 1967, pp. 363–380.
- [15] W.-B. Goh, Strategies for shape matching using skeletons, *Computer Vision and Image Understanding* 110 (3) (2008) 326–345.
- [16] K. Chan, Y. Cheung, Fuzzy-attribute graph with application to chinese character-recognition, *IEEE Transactions on Systems Man and Cybernetics* 22 (2) (1992) 402–410.

- [17] P. Suganthan, H. Yan, Recognition of handprinted chinese characters by constrained graph matching, *Image and Vision Computing* 16 (3) (1998) 191–201.
- [18] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, S. W. Zucker, Shock graphs and shape matching, *International Journal of Computer Vision*.
- [19] T. Sebastian, P. Klein, B. Kimia, Recognition of shapes by editing their shock graphs, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26 (5) (2004) 550–571.
- [20] A. Torsello, E. Hancock, A skeletal measure of 2d shape similarity, *Computer Vision and Image Understanding* 95 (1) (2004) 1–29.
- [21] C. Di Ruberto, Recognition of shapes by attributed skeletal graphs, *Pattern Recognition* 37 (1) (2004) 21–31.
- [22] D. Waltz, Understanding line drawings of scenes with shadows, in: *The Psychology of Computer Vision*, McGraw-Hill, 1975, p. pages.
- [23] A. Rosenfeld, R. A. Hummel, S. W. Zucker, Scene labelling by relaxation operations, *IEEE Transactions on Systems, Man and Cybernetics* (6) (1976) 420–433.
- [24] R. A. Hummel, S. W. Zucker, On the foundations of relaxation labling processes, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5 (3) (1983) 267–287.
- [25] S. Gold, A. Rangarajan, A graduated assignment algorithm for graph matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (4).
- [26] R. Sinkhorn, Relationship between arbitrary positive matrices + doubly stochastic matrices, *Annals of Mathematical Statistics* 35 (2) (1964) 876–&.
- [27] H. Chui, A. Rangarajan, A new point matching algorithm for non-rigid registration, *Computer Vision and Image Understanding* 89 (2-3) (2003) 114–141.
- [28] A. Rangarajan, S. Gold, E. Mjolsness, A novel optimizing network architecture with applications, *Neural Computation* 8 (5) (1996) 1041–1060.

- [29] W. Christmas, J. Kittler, M. Petrou, Structural matching in computer vision using probabilistic relaxation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (8) (1995) 749–764.
- [30] R. Wilson, E. Hancock, Structural matching by discrete relaxation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (6) (1997) 634–648.
- [31] A. Cross, E. Hancock, Graph matching with a dual-step em algorithm, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (11) (1998) 1236–1253.
- [32] B. Luo, E. R. Hancock, Structural graph matching using the em algorithm and singular value decomposition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23 (10).
- [33] A. P. Dempster, N. M. Laird, D. B. Rubin, Maximum likelihood from incomplete data via the em algorithm, *Journal of the Royal Statistical Society, Series B* 39 (1) (1977) 1–38.
- [34] B. Luo, E. Hancock, A unified framework for alignment and correspondence, *Computer Vision and Image Understanding* 92 (1) (2003) 26–55.
- [35] M. Black, A. Rangarajan, On the unification of line processes, outlier rejection, and robust statistics with applications in early vision, *International Journal of Computer Vision* 19 (1) (1996) 57–91.
- [36] <http://www.csse.uwa.edu.au/~pk/Research/MatlabFns/>.
- [37] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2003.
- [38] K. Riesen, H. Bunke, Iam graph database repository for graph based pattern recognition and machine learning, in: *Structural, Syntactic, and Statistical Pattern Recognition*, Vol. 5342 of *Lecture Notes in Computer Science*, 2008, pp. 287–297.
- [39] X. Bai, L. J. Latecki, W.-Y. Liu, Skeleton pruning by contour partitioning with discrete curve evolution, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (3) (2007) 449–462.

- [40] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, *IEEE Trans. Pattern Analysis and Machine Intelligence*. 24 (4) (2002) 509–522.
- [41] J. Munkres, Algorithms for the assignment and transportation problems, *Journal of the Society for Industrial and Applied Mathematics* 5 (1) (1957) 32–38.



Gerard Sanromà obtained his computer science engineering degree at Universitat Rovira i Virgili (Tarragona Spain) in 2004. He is now a Ph.D. student. His main areas of interest are computer vision, structural pattern recognition, graph matching and shape analysis.



René Alquézar received the licentiate and Ph.D. degrees in Computer Science from the Universitat Politècnica de Catalunya (UPC), Barcelona, in 1986 and 1997, respectively. From 1987 to 1991 he was with the Spanish company NTE as technical manager of R&D projects on image processing applications for the European Space Agency (ESA). From 1991 to 1994 he was with the Institut de Cibernètica, Barcelona, holding a pre-doctoral research grant from the Government of Catalonia. He joined the Department of Llenguatges i Sistemes Informàtics, UPC, in 1994 as an assistant professor, and since March 2001 he has been an associate professor in the same department. His current research interests include Neural Networks, Structural and Syntactic Pattern Recognition and Computer vision.



Francesc Serratosa received his computer science engineering degree from the Universitat Politècnica de Catalunya (Barcelona, Catalonia, Spain) in 1993. Since then, he has been active in research in the areas of computer vision, robotics and structural pattern recognition. He received his Ph.D. from the same university in 2000. He is currently associate professor of computer science at the Universitat Rovira i Virgili (Tarragona, Catalonia, Spain). He has published more than 40 papers and he is an active reviewer in some congresses and journals. He is the inventor of three patents on the computer vision and robotics fields.