

# Smooth Point-set Registration using Neighboring Constraints

Gerard Sanromà<sup>a,1,\*</sup>, René Alquézar<sup>b</sup>, Francesc Serratosa<sup>a</sup>, Blas Herrera<sup>a</sup>

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

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

---

## Abstract

We present an approach for Maximum Likelihood estimation of correspondence and alignment parameters that benefits from the representational skills of graphs. We pose the problem as one of mixture modelling within the framework of the Expectation-Maximization algorithm. Our mixture model encompasses a Gaussian density to model the point-position errors and a Bernoulli density to model the structural errors. The Gaussian density components are parameterized by the alignment parameters which constrain their means to move according to a similarity transformation model. The Bernoulli density components are parameterized by the continuous correspondence indicators which are updated within an annealing procedure using Softassign. Outlier rejection is modeled as a gradual assignment to the null node. We highlight the analogies of our method to some existing methods.

We investigate the benefits of using structural and geometrical information by presenting results of the full rigid version of our method together with its pure geometrical and its pure structural versions. We compare our method to other point-set registration methods as well as to other graph matching methods which incorporate geometric information. We also present a non-rigid version of our method and compare to state-of-the art non-rigid registration methods.

---

\*Corresponding author

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

<sup>1</sup>Fax: (+34) 977 55 97 10

Results show that our method gets either the best performance or similar performance than state-of-the art methods.

*Keywords:* graph matching, point-set registration, correspondence problem, expectation-maximization, softassign

---

## 1. Introduction

Alignment of point-sets is frequently used in *pattern recognition* when objects are represented by sets of coordinate points. The idea behind is to be able to compare two objects regardless the effects of a given transformation model on their coordinate data. This is at the core of many object recognition applications where the objects are defined by coordinate data (e.g., medical image analysis, shape retrieval, ...), learning shape models (Dryden and Mardia, 1998; Cootes et al., 1995) or reconstructing a scene from various views (Hartley and Zisserman, 2000).

Given that the correspondences are known, there is an extensive work done towards the goal of finding the alignment parameters that minimize some error measure. To cite a few, Dryden and Mardia (1998); Kendall (1984) deal with isometries and similarity transformations; Berge (2006); Umeyama (1991) deal with Euclidean transformations (i.e. excluding reflections from isometries); Haralick et al. (1989) deals with similarity and projective transformations; and Hartley and Zisserman (2000) deals exclusively with projective transformations.

However, the point-set alignment problem is often found in the more realistic setting of unknown point-to-point correspondences. This problem becomes then a *registration problem*, this is, one of jointly estimating the alignment and correspondence parameters. Although non-iterative algorithms exist for specific types of transformation models (Ho and Yang, 2011), this problem is usually solved by means of non-linear iterative methods that, at each iteration, estimate correspondence and alignment parameters. Despite being more computationally demanding, iterative methods are more appealing to us than the direct ones due to its superior tolerance to noise and outliers.

We distinguish between two families of approaches at solving this problem. Ones are based on the *Expectation-Maximization* (EM) algorithm (Dempster et al., 1977), and the others use *Softassign* (Gold and Rangarajan, 1996; Gold et al., 1998; Rangarajan et al., 1997). The former ones have the ad-

vantage of offering statistical insights of such decoupled estimation processes while the latter ones benefit from the well-known robustness and convergence properties of the Softassign embedded within deterministic annealing procedures.

Myronenko and Song (2010) proposed *Coherent Point Drift* (CPD), a point-set registration method using the EM algorithm that is defined for rigid, affine and non-rigid transformations. Gold et al. (1998); Rangarajan et al. (1997) proposed *Robust Point Matching* (RPM), a method using Softassign that is defined for affine and rigid transformations. Later, Chui and Rangarajan (2003, 2000) presented *TPS-RPM*, its extension to non-rigid transformations.

*Graph matching* approaches allow for neighboring relations between points into the point-set registration problem. *Graduated Assignment* by Gold and Rangarajan (1996) is a remarkable graph matching method using Softassign. Cross and Hancock (1998) presented an approach for graph matching and point-set alignment using the EM algorithm that was defined for affinities and projectivities. One limitation of this approach is the high computational demand of the dictionary-based structural model. Luo and Hancock (2003) proposed an EM-like approach for graph matching and point-set alignment based on a cross-entropy measure. They proposed a model of structural errors based on a Bernoulli distribution. This model was defined for rigid-body transformations.

We propose a joint structural graph matching and point-set registration method whose main contributions are the following:

- We try to bridge the gap between the EM-based and the Softassign-based approaches by formulating the graph matching problem within a principled statistical framework, while benefiting from the desirable properties of the Softassign and deterministic annealing ensemble.
- Correspondence problem is approximated as a succession of linear assignment problems which are solved using Softassign. This way, we are able to use continuous correspondence variables as opposed to other approaches that use discrete ones (Cross and Hancock, 1998; Luo and Hancock, 2003).
- Outlier rejection is modelled as a smooth assignment to the null node within the annealing procedure.

- The proposed model can be easily adapted to allow only for geometric or structural information. We show how it can be seen as a more general framework with clear connections to other well-known methods.
- Although the proposed method deals with rigid transformations we show how it can be embedded into a non-rigid deformation procedure thus obtaining similar or better performances than state-of-the-art non-rigid registration methods.

The outline of this paper is the following. In section 2 we formulate the matching problem as one of mixture modelling with missing data and propose our mixture model. In section 3 we derive the EM algorithm for our model. Section 4 presents the methodology used to reject outliers. In section 5 we highlight parallelisms of the proposed method with some other existing methods. In section 6 we present some experiments and results, and finally in section 7 conclusions are given.

## 2. A Mixture Model

Consider two graph representations  $\mathcal{G} = (\mathcal{U}, D, \mathcal{X})$  and  $\mathcal{H} = (\mathcal{V}, M, \mathcal{Y})$  extracted from two images.

The node-sets  $\mathcal{U} = \{u_a, a \in \mathcal{I}\}$  and  $\mathcal{V} = \{v_\alpha, \alpha \in \mathcal{J}\}$  contain the symbolic representations of the nodes, where  $\mathcal{I} = 1, \dots, |\mathcal{U}|$  and  $\mathcal{J} = 1, \dots, |\mathcal{V}|$  are their index-sets.

The vector-sets  $\mathcal{X} = \{\mathbf{x}_a, a \in \mathcal{I}\}$  and  $\mathcal{Y} = \{\mathbf{y}_\alpha, \alpha \in \mathcal{J}\}$ , contain the column vectors  $\mathbf{x}_a = (x_a^V, x_a^H)^\top$  and  $\mathbf{y}_\alpha = (y_\alpha^V, y_\alpha^H)^\top$  of the two-dimensional coordinates (vertical and horizontal) of each node, where  $\top$  denotes the transpose operator.

The adjacency matrices  $D$  and  $M$  contain the edge-sets, encoding some kind of relation between pairs of nodes (e.g., connectivity or spatial proximity). Hence,  $D_{ab} = \begin{cases} 1 & \text{if } u_a \text{ and } u_b \text{ are linked by an edge} \\ 0 & \text{otherwise} \end{cases}$  (the same applies for  $M_{\alpha\beta}$ ).

We use continuous correspondence indicators  $S$  so, we denote as  $s_{a\alpha} \in S$ , the probability of node  $u_a \in \mathcal{U}$  being in correspondence with node  $v_\alpha \in \mathcal{V}$ .

It is satisfied that

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

where,  $1 - \sum_{\alpha} s_{a\alpha}$  is the probability of node  $u_a$  being an outlier.

Our aim is to recover the correspondence indicators  $S$  and the alignment parameters  $\Phi$  that maximize the *observed-data* likelihood of the data-graph  $P(\mathcal{G}|S, \Phi)$ . Within this setting, constraints on the data-graph  $\mathcal{G}$  are evaluated on the model-graph  $\mathcal{H}$ . To make this problem tractable, we introduce the *hidden variables*, namely, the corresponding model graph nodes  $v_{\alpha} \in \mathcal{V}$ .

By assuming that the observations are independent and identically distributed, the observed-data likelihood writes

$$P(\mathcal{G}|S, \Phi) = \prod_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} P(u_a, v_{\alpha}|S, \Phi) \quad (2)$$

Following a similar development than Luo and Hancock (2001) we factorize, using the Bayes rules, the *complete-data* likelihood in the right hand side of equation (2) into terms depending on individual correspondence indicators, in the following way.

$$P(u_a, v_{\alpha}|S, \Phi) = K_{a\alpha} \prod_{b \in \mathcal{I}} \prod_{\beta \in \mathcal{J}} P(u_a, v_{\alpha}|s_{b\beta}, \Phi) \quad (3)$$

where  $K_{a\alpha} = [1/P(u_a|v_{\alpha}, \Phi)]^{|\mathcal{I}| \times |\mathcal{J}| - 1}$ . If we assume that conditional dependence of data-graph node  $u_a$  can only be taken into account in the presence of the correspondence matches  $S$ , then  $P(u_a|v_{\alpha}, \Phi) = P(u_a)$ . Further assuming equiprobable priors  $P(u_a)$ , we can safely discard these quantities in the maximization of equation (2), since they do not depend either on  $S$  nor  $\Phi$ .

We propose a measure for the complete-data likelihood of equation (3) that combines a model of structural errors based on a Bernoulli distribution augmented with a model of geometric errors based on a Gaussian distribution.

With regards to the structural relations, Luo and Hancock (2001) proposed to model the likelihood of an observed relation given the hypothesis on the correspondences using a Bernoulli distribution with parameters  $S$ . This is, given two corresponding pairs of nodes  $u_a, u_b \in \mathcal{U}$  and  $v_{\alpha}, v_{\beta} \in \mathcal{V}$ , they assumed that there will be edge-discordance (i.e.,  $D_{ab} = 0 \vee M_{\alpha\beta} = 0$ ) with a fixed (low) probability of error  $P_e$ . Otherwise, there will be edge-concordance with probability  $1 - P_e$ . This is,

$$P(u_a, v_{\alpha}|s_{b\beta}) = \begin{cases} (1-P_e) & \text{if } D_{ab}=1 \wedge M_{\alpha\beta}=1 \wedge s_{b\beta}=1 \\ P_e & \text{otherwise} \end{cases} \quad (4)$$

With regards to the geometrical measurements, it is reasonable to consider that point-position errors between corresponding points follow a Gaussian density. In the case of no correspondence, we use a fixed probability  $\rho$  that will model the outlier process. This is,

$$P(u_b | s_{b\beta}, \Phi) = \begin{cases} P_{b\beta}^{(\Phi)} & \text{if } s_{b\beta}=1 \\ \rho & \text{otherwise} \end{cases} \quad (5)$$

where  $P_{b\beta}^{(\Phi)}$  is a Gaussian measurement on the point-position errors with parameters  $\Phi$ . This is,

$$P_{b\beta}^{(\Phi)} = \frac{1}{2\pi|\Sigma|^{1/2}} \exp \left[ -\frac{1}{2} \|\mathbf{x}_b - \mathcal{T}(\mathbf{y}_\beta; \Phi)\|_\Sigma^2 \right] \quad (6)$$

where  $\mathcal{T}(\mathbf{y}_\beta; \Phi)$  represents the geometric transformation of model point  $\mathbf{y}_\beta$  according to alignment parameters  $\Phi$ , and  $\|\mathbf{d}\|_\Sigma^2 = \mathbf{d}^\top \Sigma^{-1} \mathbf{d}$  is the squared Mahalanobis distance to covariance matrix  $\Sigma$ , with  $\mathbf{d}$  a column vector. As opposed to the standard Gaussian modeling approach, here the means are parameterized by the alignment parameters which enforce prior knowledge about the transformation that exists between the two sets of points.

We propose a more fine-grained likelihood measure than that of equation (4) by considering that it is appropriate to weight the likelihood of an observed relation with the geometric likelihood term defined in equation (5).

In the case of discrete correspondence indicators (i.e.,  $s_{b\beta} = \{0, 1\}$ ), the proposed density writes

$$P(u_a, v_\alpha | s_{b\beta}, \Phi) = \begin{cases} (1-P_e)P_{b\beta}^{(\Phi)} & \text{if } D_{ab}=1 \wedge M_{\alpha\beta}=1 \wedge s_{b\beta}=1 \\ P_e P_{b\beta}^{(\Phi)} & \text{if } (D_{ab}=0 \vee M_{\alpha\beta}=0) \wedge s_{b\beta}=1 \\ P_e \rho & \text{if } s_{b\beta}=0 \end{cases} \quad (7)$$

We extrapolate to continuous correspondence indicators by exploiting each case of equation (7) as exponential indicators. This is,

$$P(u_a, v_\alpha | s_{b\beta}, \Phi) = \left[ (1 - P_e) P_{b\beta}^{(\Phi)} \right]^{D_{ab} M_{\alpha\beta} s_{b\beta}} \left[ P_e P_{b\beta}^{(\Phi)} \right]^{(1 - D_{ab} M_{\alpha\beta}) s_{b\beta}} \left[ P_e \rho \right]^{(1 - s_{b\beta})} \quad (8)$$

By using the density measurement of equation (8), the final expression for the complete-data likelihood of equation (3), expressed in the exponential

form, is

$$P(u_a, v_\alpha | S, \Phi) = \exp \left[ \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta} D_{ab} M_{\alpha\beta} \ln \left( \frac{1 - P_e}{P_e} \right) + s_{b\beta} \ln \left( \frac{P_{b\beta}^{(\Phi)}}{\rho} \right) + \ln(P_e \rho) \right] \quad (9)$$

### 3. Expectation Maximization

The EM algorithm has been previously used by other authors to solve the graph matching problem (Cross and Hancock, 1998; Luo and Hancock, 2001). We seek the optimal alignment parameters  $\Phi^*$  and the correspondence indicators  $S^*$  that maximize our observed-data log-likelihood, i.e.,  $\ln P(\mathcal{G} | S, \Phi)$ . This is,

$$\{\Phi^*, S^*\} = \arg \max_{\Phi, S} \sum_{a \in \mathcal{I}} \ln \left( \sum_{\alpha \in \mathcal{J}} P(u_a, v_\alpha | S, \Phi) \right) \quad (10)$$

Dempster et al. (1977) proposed to replace equation (10) by its *conditional expectation conditioned by the observed data*. It has been proven that maximizing the conditional expectation is equivalent at maximizing the observed-data log-likelihood. Accordingly, we seek the parameters  $S^{(n+1)}, \Phi^{(n+1)}$  that maximize the following objective function

$$\begin{aligned} \{\Phi^{(n+1)}, S^{(n+1)}\} &= \arg \max_{\Phi, S} E_{\mathcal{V}} [\ln P(\mathcal{G} | S, \Phi) | \mathcal{G}] \\ &= \arg \max_{\Phi, S} \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} P(v_\alpha | u_a, S^{(n)}, \Phi^{(n)}) \ln P(u_a, v_\alpha | S, \Phi) \end{aligned} \quad (11)$$

where  $P(v_\alpha | u_a, S^{(n)}, \Phi^{(n)})$  are the posterior probabilities of the missing data given the most recent available parameters  $S^{(n)}, \Phi^{(n)}$ .

The basic idea is to alternate between Expectation and Maximization steps until convergence is reached. The expectation step involves computing the 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 observed data.

#### 3.1. Expectation

In the expectation step, the posterior probabilities of the missing data (i.e., the corresponding model-graph  $v_\alpha$  estimates) are computed using the current parameter estimates  $S^{(n)}, \Phi^{(n)}$ .

The posterior probabilities are computed, according to the Bayes rule, using the following expression

$$\begin{aligned}
P(v_\alpha | u_a, S^{(n)}, \Phi^{(n)}) &= \frac{P(u_a, v_\alpha | S^{(n)}, \Phi^{(n)})}{\sum_{\alpha'} P(u_a, v_{\alpha'} | S^{(n)}, \Phi^{(n)})} \\
&= \frac{\exp \left[ \sum_{b,\beta} s_{b\beta} D_{ab} M_{\alpha\beta} \ln \left( \frac{1-P_e}{P_e} \right) \right]}{\sum_{\alpha'} \exp \left[ \sum_{b,\beta} s_{b\beta} D_{ab} M_{\alpha'\beta} \ln \left( \frac{1-P_e}{P_e} \right) \right]} \stackrel{def}{=} \omega_{a\alpha}^{(n)} \quad (12)
\end{aligned}$$

Note that when substituting the complete-data likelihoods by their expressions of equation (9), the last two terms in the summations are cancelled out by the quotient since they do not depend either on nodes  $u_a$  or  $v_\alpha$ . As we have stated in equation (5), the hidden corresponding nodes  $v_\alpha$  do not affect the point-position errors. As consequence, the missing data posteriors are revealed as strictly structural measurements. Point-position errors, which are conditionally dependant on the correspondence indicators  $s_{b\beta}$ , will affect the ML estimate of the correspondence parameters, as we will see later.

### 3.2. Maximization

It is a well-established strategy to implement the maximization step into a series of *conditional maximization* steps (Horn et al., 2011). Then, it turns into an instance of the *expectation conditional maximization* (ECM) (Meng and Rubin, 1993) algorithm which still shares the desirable convergence properties of EM. According to ECM, maximization of equation (11) can be decomposed into three steps. First, maximize over the alignment parameters, next compute empirical covariances using the newly estimated alignment parameters  $\Phi^{(n+1)}$ , and finally maximize over the correspondence indicators while using the newly estimated empirical covariances  $\Sigma^{(n+1)}$  and alignment parameters  $\Phi^{(n+1)}$ .

#### 3.2.1. Maximum Likelihood Similarity Alignment Parameters

We seek the alignment parameters  $\Phi^{(n+1)}$  that maximize equation (11). We use the expressions in equations (12) and (9) for the posterior probability and conditional likelihood terms, respectively. Discarding the terms constant with respect to the alignment parameters we obtain the following expression

$$\Phi^{(n+1)} = \arg \max_{\Phi} \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} \omega_{a\alpha}^{(n)} \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} \ln \left( \frac{P_{b\beta}^{(\Phi)}}{\rho} \right) \quad (13)$$

Rearranging and further removing other terms constant with respect to the alignment parameters, we get

$$\begin{aligned}\Phi^{(n+1)} &= \arg \max_{\Phi} \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} \ln \left( \frac{P_{b\beta}^{(\Phi)}}{\rho} \right) \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} \omega_{a\alpha}^{(n)} \\ &= \arg \max_{\Phi} \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} \ln P_{b\beta}^{(\Phi)}\end{aligned}\quad (14)$$

$$= \arg \min_{\Phi} \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} \|\mathbf{x}_b - \mathcal{T}(\mathbf{y}_\beta; \Phi)\|_{\Sigma^{(n)}}^2 \quad (15)$$

In going from equation (14) to (15) we turned the maximization into a minimization by substituting the geometrical probability term  $P_{b\beta}^{(\Phi)}$  by its expression in (6) while discarding the constant terms.

Note that the alignment parameters do not depend on the posterior probability terms  $\omega_{a\alpha}^{(n)}$  but on the correspondence variables  $s_{b\beta}^{(n)}$ . This is because, as stated in equation (5), point position errors are evaluated on the basis of the correspondence variables instead of the missing-data posteriors.

From equation (15), we seek the optimal rotation matrix  $R^*$ , scaling parameter  $\eta^*$  and translation 2-vector  $\mathbf{t}^*$  that minimize the following quantity

$$\min_{R, \eta, \mathbf{t}} \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} \|\mathbf{x}_b - (\eta R \mathbf{y}_\beta + \mathbf{t})\|_{\Sigma^{(n)}}^2 \quad (16)$$

subject to determinant of  $R$  equals 1.

As noticed by Horaud et al. (2011), values  $s_{b\beta}^{(n)}$  define a spatial mapping of points in  $\mathcal{X}$ . This way, equation (16) can be simplified by introducing the virtual observation  $\mathbf{w}_\beta \in \mathcal{W}$  and its weight  $\varphi_\beta$  that are assigned to a model point  $\mathbf{y}_\beta$

$$\mathbf{w}_\beta = \frac{1}{\varphi_\beta} \sum_{b \in \mathcal{I}} s_{b\beta}^{(n)} \mathbf{x}_b \quad (17)$$

$$\varphi_\beta = \sum_{b \in \mathcal{I}} s_{b\beta}^{(n)} \quad (18)$$

In order to simplify the problem we will assume that the covariances are isotropic, namely,  $\Sigma^{(n)} = \sigma^2 I_2$  where  $I_2$  is the  $2 \times 2$  identity matrix. This way, the Mahalanobis distance reduces to the Euclidean distance.

By introducing the virtual observation of equations (17) and (18) and the isotropic covariances, minimization of equation (16) can be expressed in the simpler form

$$\min_{R, \eta, \mathbf{t}} \sum_{\beta \in \mathcal{J}} \varphi_{\beta}^{(n)} \sigma^{-2} \|\mathbf{w}_{\beta} - (\eta R \mathbf{y}_{\beta} + \mathbf{t})\|^2 \quad (19)$$

We cannot directly estimate the similarity parameters from (19) as done by Umeyama (1991) since the means and variances of  $\mathcal{W}$  and  $\mathcal{Y}$  are biased by the weights  $\varphi_{\beta}^{(n)}$ . We follow the Ansatz by Rangarajan et al. (1997) and compute the weighted means and variances of the point-sets. This is,

$$\bar{\mathbf{w}} = \frac{\sum_{\beta} \varphi_{\beta}^{(n)} \mathbf{w}_{\beta}}{\sum_{\beta} \varphi_{\beta}^{(n)}} \quad (20)$$

$$\bar{\mathbf{y}} = \frac{\sum_{\beta} \varphi_{\beta}^{(n)} \mathbf{y}_{\beta}}{\sum_{\beta} \varphi_{\beta}^{(n)}} \quad (21)$$

$$\sigma_w^2 = \frac{\sum_{\beta} \varphi_{\beta}^{(n)} \|\mathbf{w}_{\beta} - \bar{\mathbf{w}}\|^2}{\sum_{\beta} \varphi_{\beta}^{(n)}} \quad (22)$$

$$\sigma_y^2 = \frac{\sum_{\beta} \varphi_{\beta}^{(n)} \|\mathbf{y}_{\beta} - \bar{\mathbf{y}}\|^2}{\sum_{\beta} \varphi_{\beta}^{(n)}} \quad (23)$$

$$\Sigma_{wy} = \frac{\sum_{\beta} \varphi_{\beta}^{(n)} (\mathbf{w}_{\beta} - \bar{\mathbf{w}}) (\mathbf{y}_{\beta} - \bar{\mathbf{y}})^{\top}}{\sum_{\beta} \varphi_{\beta}^{(n)}} \quad (24)$$

(note that the variances  $\sigma^{-2}$  appearing in equation (19) cancel out).

Optimal parameters  $R^*$ ,  $\eta^*$  and  $\mathbf{t}^*$  are found as follows (Umeyama, 1991).

$$R^* = UV^{\top} \quad (25)$$

$$\eta^* = \frac{1}{\sigma_y^2} \text{tr}(\Lambda) \quad (26)$$

$$\mathbf{t}^* = \bar{\mathbf{x}} - \eta^* R^* \bar{\mathbf{y}} \quad (27)$$

where  $U, \Lambda, V$  come from the *singular value decomposition*  $\Sigma_{wy} = U\Lambda V^{\top}$ .

### 3.2.2. Empirical Covariances

We compute the variances using the newly estimated registration parameters  $\Phi^{(n+1)}$  according to the following expression.

$$\sigma^2 = \frac{\sum_{b, \beta} s_{b\beta}^{(n)} (\mathbf{x}_b - \mathcal{T}(\mathbf{y}_{\beta}; \Phi^{(n+1)}))^{\top} (\mathbf{x}_b - \mathcal{T}(\mathbf{y}_{\beta}; \Phi^{(n+1)}))}{\sum_{b, \beta} s_{b\beta}^{(n)}} \quad (28)$$

and set isotropic covariance matrix as  $\Sigma^{(n+1)} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$ .

### 3.2.3. Maximum Likelihood Correspondence Indicators

One of the key points in our work is to approximate the solution of the graph matching problem by a succession of easier assignment problems. As it is done in Robust Point Matching (Gold et al., 1998; Rangarajan et al., 1997; Chui and Rangarajan, 2003, 2000) and Graduated Assignment (Gold and Rangarajan, 1996), we use *Softassign* to solve the assignment problems in a continuous way.

According to the EM development, we compute the correspondence indicators  $S^{(n+1)}$  that maximize equation (11). Substituting equations (12) and (9) into (11) and discarding the constant term ( $\ln(P_e\rho)$ ), we obtain

$$S^{(n+1)} = \arg \max_S \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} \omega_{a\alpha}^{(n)} \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta} \left[ D_{ab} M_{\alpha\beta} \ln \left( \frac{1-P_e}{P_e} \right) + \ln \left( \frac{P_{b\beta}^{(n+1)}}{\rho} \right) \right] \quad (29)$$

where  $P_{b\beta}^{(n+1)}$  is the Gaussian of the point errors of equation (6) using the recently estimated alignment parameters  $\Phi^{(n+1)}$  and covariance matrix  $\Sigma^{(n+1)}$ .

Rearranging terms we obtain the following assignment problem (Gold and Rangarajan, 1996)

$$S^{(n+1)} = \arg \max_S \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta} B_{b\beta} \quad (30)$$

where

$$\begin{aligned} B_{b\beta} &= \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} \omega_{a\alpha}^{(n)} \left[ D_{ab} M_{\alpha\beta} \ln \left( \frac{1-P_e}{P_e} \right) + \ln \left( \frac{P_{b\beta}^{(n+1)}}{\rho} \right) \right] \\ &= \ln \left( \frac{P_{b\beta}^{(n+1)}}{\rho} \right) \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} \omega_{a\alpha}^{(n)} + \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} \omega_{a\alpha}^{(n)} D_{ab} M_{\alpha\beta} \ln \left( \frac{1-P_e}{P_e} \right) \end{aligned} \quad (31)$$

$$\simeq \ln \left( \frac{P_{b\beta}^{(n+1)}}{\rho} \right) + \sum_{a \in \mathcal{I}} \sum_{\alpha \in \mathcal{J}} \omega_{a\alpha}^{(n)} D_{ab} M_{\alpha\beta} \ln \left( \frac{1-P_e}{P_e} \right) \quad (32)$$

is the benefit value for the assignment  $u_b \rightarrow v_\beta$ .

We have observed a better stability of the algorithm when removing the summation  $\sum_{a,\alpha} \omega_{a\alpha}^{(n)}$  in going from equation (31) to (32) which acts as a constant amplification term, especially when dealing with large graphs.

Notice how the two motivations underpinning our work, namely, the pure geometric and the pure structural, are clearly identified in the benefit measure of equation (32).

Computation of the correspondence indicators with Softassign consists in two steps:

1. Correspondence indicators are updated with the exponential of the benefit coefficients. This is,

$$s_{b\beta} = \exp [\mu B_{b\beta}] \quad (33)$$

where  $\mu$  is a control parameter.

2. Two-way constraints are imposed by alternatively normalizing across rows and columns the matrix of exponentiated benefits. This is known as the *Sinkhorn normalization* (Sinkhorn, 1964) and, it is applied either until convergence or a predefined number of times. We have observed an improvement in the performance of the algorithm when applying Sinkhorn normalization to the missing data posteriors of equation (12) as well.

Softassign is run within an annealing procedure that increases the value of  $\mu$  at each maximization step. Starting from low values of  $\mu$ , the correspondence indicators  $s_{b\beta}$  are gradually pushed from continuous to discrete values as  $\mu$  increases.

#### 4. Outlier Rejection

It is important to develop techniques aimed at detecting and rejecting outliers in order to minimize their influence.

We consider that a node  $u_b \in \mathcal{U}$  (or  $v_\beta \in \mathcal{V}$ ) is an outlier if there is not any node  $v_\beta, \forall \beta \in \mathcal{J}$  (or  $u_b, \forall b \in \mathcal{I}$ ) with a matching benefit  $B_{b\beta}$  higher than a predefined threshold.

Outlier detection is handled as an assignment to (or from) the *null* node. Considering that the null node has no edges, and that all the probabilities  $P_{b\beta}^{(\Phi)}$  involving the null node are equal to  $\rho$ , then the benefit values of equation (32) corresponding to the null assignments are equal to zero. We create an augmented benefit matrix  $\tilde{B}$  by adding to  $B$  an extra row and column of zeros representing the benefits of the null assignments (i.e.,  $B_{b\emptyset}, \forall b \in \mathcal{I}$  and  $B_{\emptyset\beta}, \forall \beta \in \mathcal{J}$ ).

Note that  $\rho$  establishes the threshold at which the terms  $\ln(P_{b\beta}^{(\Phi)}/\rho)$  contribute positively (i.e.,  $\rho < P_{b\beta}^{(\Phi)}$ ) or negatively (i.e.,  $\rho > P_{b\beta}^{(\Phi)}$ ) to the benefit measure.

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

As the control parameter  $\mu$  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.

We still have to define the value of the outlying threshold  $\rho$ . From the first term of equation (32), we see that  $\rho$  is to be compared with  $P_{b\beta}^{(n+1)}$ . We consider therefore convenient to define it in terms of a multivariate Gaussian of a distance threshold. This is,

$$\rho = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}\|\mathbf{d}\|_{\Sigma}^2\right] \quad (34)$$

where,  $\Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$ , is an isotropic covariance variance matrix and,  $\mathbf{d} = (d^V, d^H)^{\top}$  is a column vector with the vertical and horizontal thresholding distances.

If we cancel out the Gaussian constant terms in the numerator and denominator of the quotient  $P_{b\beta}^{(\Phi)}/\rho$ , and we express the thresholding distance proportionally to the standard deviations of the data (i.e.,  $\mathbf{d} = (N\sigma, N\sigma)$ ), the expression of  $\rho$  to be compared with  $P_{b\beta}^{(n+1)}$  becomes

$$\rho = \exp\left\{-\frac{1}{2}\left[\left(\frac{N\sigma}{\sigma}\right)^2 + \left(\frac{N\sigma}{\sigma}\right)^2\right]\right\} = \exp(-N^2) \quad (35)$$

So,  $\rho$  is defined as a function of the number  $N$  of standard deviations permitted in the alignment errors, in order to consider a plausible correspondence.

## 5. Related Methods

In the following we highlight the relationships between the proposed method and some existing methods.

Gold and Rangarajan (1996) presented Graduated Assignment, a structural graph matching method that updates the correspondence variables  $s_{a\alpha} \in S$  following an annealing scheme in the following way

$$s_{a\alpha}^{(n+1)} = \exp \left[ \mu \sum_{b \in \mathcal{I}} \sum_{\beta \in \mathcal{J}} s_{b\beta}^{(n)} C_{a\alpha b\beta} \right] \quad (36)$$

where  $\mu$  is an annealing parameter that is gradually increased and  $C_{a\alpha b\beta}$  is the edge-compatibility coefficient for the assignment  $(a, b) \rightarrow (\alpha, \beta)$ . We have used the commonly adopted value  $C_{a\alpha b\beta} = c D_{ab} M_{\alpha\beta}$  that assigns a positive scalar  $c$  in the case of edge-concordance, and 0 otherwise. The updating equation (36) is followed by a Sinkhorn normalization on the matrix of correspondence variables  $S$ .

There is a noticeable parallelism between equation (36) from Graduated Assignment and equations (32) and (33) from our method. If we disregard any geometric measurement in our method by setting  $P_{b\beta}^{(n+1)} = \rho$  for all  $b, \beta$  we obtain a pure structural version of our method which is equivalent to the aforementioned implementation of Graduated Assignment given the identifications  $c = \ln [(1 - P_e) / P_e]$  and  $s_{b\beta}^{(n)} = \omega_{b\beta}^{(n)}$ .

A particular case of our method with a pure geometric motivation consists on using an ambiguous structural model, this is, set the value  $P_e = 0.5$ . This particular case reduces to iteratively computing the correspondence and alignment parameters according to the following steps: (1) from equations (32), (33) and (35), update  $S$  with the following expression

$$\begin{aligned} s_{b\beta} &= \exp \left[ \mu \left( - \|x_b - \mathcal{T}(y_\beta; \Phi)\|_\Sigma^2 - \ln \rho \right) \right] \\ &= \exp \left[ \mu \left( - \|x_b - \mathcal{T}(y_\beta; \Phi)\|_\Sigma^2 + N^2 \right) \right] \end{aligned} \quad (37)$$

(2) normalize  $S$  across rows and columns having into account the extra row and column of the null assignments, (3) compute alignment parameters according to equation (16), and (4) increase  $\mu$  and repeat steps (1-3) until  $\mu$  reaches a predefined threshold.

It is worth pausing at this point to consider the analogies of this particular case of our method to a well-known method by Rangarajan et al. (1997),

namely, the *Softassign Procrustes*. The essential difference with the Softassign Procrustes algorithm is that they use the squared Euclidean distance instead of the squared Mahalanobis distance. This way, their "robustness parameter", which is the analog of our  $N^2$  term, is to be compared to the Euclidean distance. Unfortunately, they do not address the estimation of this parameter in their paper. On the contrary, we pose it in terms of the standard deviation which is a well-defined parameter in our method.

## 6. Experiments and Results

We have performed matching experiments with synthetic and real data. Experiments with synthetic data consists on matching point-sets extracted from a fish and a Chinese character templates under nonrigid deformations, noise and outliers. Experiments with real data consists on matching point-sets extracted from images from various scenes across different zooms and rotations.

### 6.1. Synthetic Data

We have performed matching experiments on the dataset by Chui and Rangarajan (2000). This dataset contains perturbed instances of a fish and a Chinese character templates, consisting of 98 and 105 points, respectively. Perturbation levels range from mild to severe, with 100 different instances for each level. The types of perturbations are, (1) non-rigid deformations based on Gaussian radial basis functions (RBF) (Yuille and Grzywacz, 1989), (2) independent random noise and, (3) and a certain percentage of outliers ranging from 0% to 300%. A certain amount of ground-level non-rigid deformation is maintained in the random noise and outliers. See figure 1 for an example.

We use both rigid and non-rigid registration methods in the comparisons.

The non-rigid methods used in the comparisons are CPD (Myronenko and Song, 2010) and TPS-RPM (Chui and Rangarajan, 2003, 2000).

In order to compare to non-rigid methods we propose a non-rigid extension of our method consisting on an iterative process of correspondences estimation followed by Thin-Plate Splines (TPS) warping of the point-set using the computed correspondences. Correspondences are computed using the rigid version of our method (explained in this paper). The regularization parameter penalizing the deformations in TPS is set as in Chui and Rangarajan (2003, 2000). This is, it consists on an annealing parameter allowing

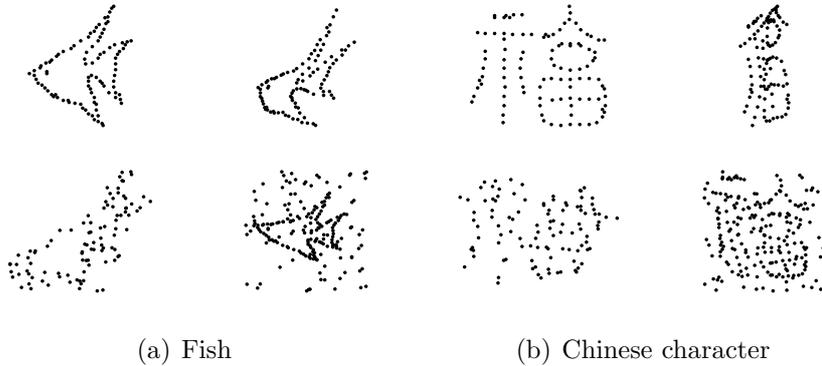


Figure 1: From top to down and left to right, the model templates and moderately perturbed instances due to non-rigid deformations, noise and outliers for (a) the fish template and, (b) the Chinese character template.

more intense deformations as the algorithm proceeds. In this way, at the beginning of the annealing procedure the rigid transformation model provides rough approximates of the registration that gradually become more accurate as we increase the intensity of the deformations and the model point-set approaches the form of the deformed point-set.

With regards to rigid registration we present the results of three graph matching methods and one point-set registration method. All the methods using geometric measurements implement a similarity transformation model which provide a rough approximation to the true transformations undergone by the pattern templates.

The graph matching methods are: the one presented in this paper, the Unified approach by Luo and Hancock (2003) and the Graduated Assignment by Gold and Rangarajan (1996), which is the equivalent of a pure structural version of our method as explained in section 5.

The point-set registration method is the approach presented in this paper using an ambiguous structural model (i.e.,  $P_e = 0.5$ ) which is closely related to the Softassign Procrustes by Rangarajan et al. (1997) as explained in section 5.

Ground truth matches are available between model templates and all the deformed templates. We assess the error differently in the case of rigid and non-rigid methods. In the non-rigid case, we compute the average Euclidean distance between corresponding points after the model template has been non-rigidly deformed to match the perturbed template. In the rigid case,

we take the average of the correspondence error where, for a given point in the model template, the correspondence error is the distance between the corresponding ground truth point in the perturbed template and the point which it is actually assigned to. Since this measure tends to penalize methods with higher number of correspondences we match all the points from the model template in all the methods. To do so, we apply the Hungarian method to the resulting matrix of matching coefficients of each method.

We have experimentally set the parameters  $P_e = 0.1$  for the proposed method and  $P_e = 0.03$  for the Unified approach by Luo and Hancock (2003) and the non-rigid extension of our method. With regards to the outlying threshold  $\rho$ , we have used the value  $N = 1$  from equation (35) in the proposed method and  $N = 3$  for the pure geometric version of our method. In the case of Graduated Assignment, we have experimentally set the compatibility coefficient in case of edge-concordance to  $c = 3.47$  (i.e.,  $P_e = 0.03$ ).

We have generated the graphs for the deformation and noise experiments following a mutual  $K$ -nearest-neighbour approach, with  $K = 5$ . This is, two points are joined by an edge if both points are among the 5 nearest neighbors of the other point.

In the outlier experiments we place edges between the pairs of points presenting the  $M$  lower pair-wise distances. With such a strategy we aim to concentrate most of the edges among the inliers which is useful in the present dataset, since outliers are comparably more scattered. Suppose that in average we want to place  $K'$  edges for each point. Therefore, we set  $M = (K' \cdot n) / 2$ , where  $n$  is the number of points. We have experimentally set  $K' = 4$  and  $K' = 2$  in the rigid and non-rigid methods using graphs, respectively.

Figures 2 and 3 show the error means and standard deviations for each method in the fish and Chinese character templates.

The proposed method outperforms the rest of the methods in the rigid-body alignment experiments. The Unified method by Luo and Hancock (2003), which also implements a combination of geometry and structure, shows a slight decrease in performance with respect to our method. This method updates correspondence and alignment parameters in an EM-like procedure like ours. However, Unified updates correspondences in a discrete way at each iteration. In the outlier experiments the Unified method experiences a decrease in performance with respect to our method, even though we have used the same approach to generate the structure as in ours.

Graduated Assignment and Softassign Procrustes (i.e., the pure struc-

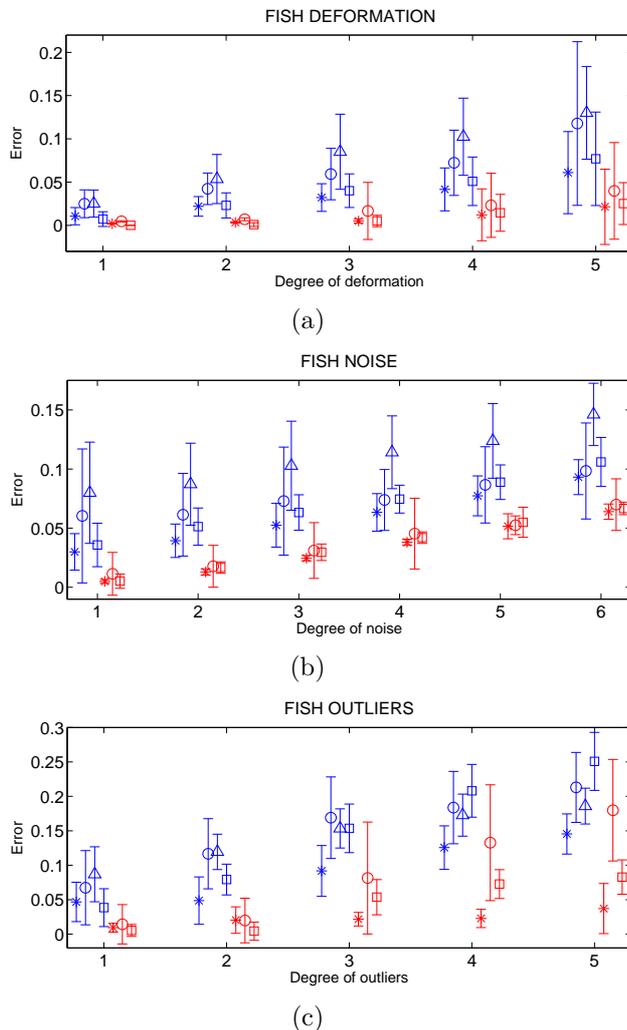


Figure 2: Results for the fish (a) deformation, (b) noise, and (c) outlier experiments. In blue the rigid methods: (\*) our method, (O) our method with  $P_e = 0.5$  (i.e., Softassign Procrustes), ( $\Delta$ ) Graduated Assignment and ( $\square$ ) Unified approach by Luo and Hancock (2003). In red the non-rigid methods: (\*) our method, (O) TPS-RPM and ( $\square$ ) CPD.

tural and geometric versions of our method) get considerably worse performance than the proposed method. This suggests that the performance of our method is due to the combination of both parts rather than being merit of one of its components.

With regards to the non-rigid methods, our method shows slightly better

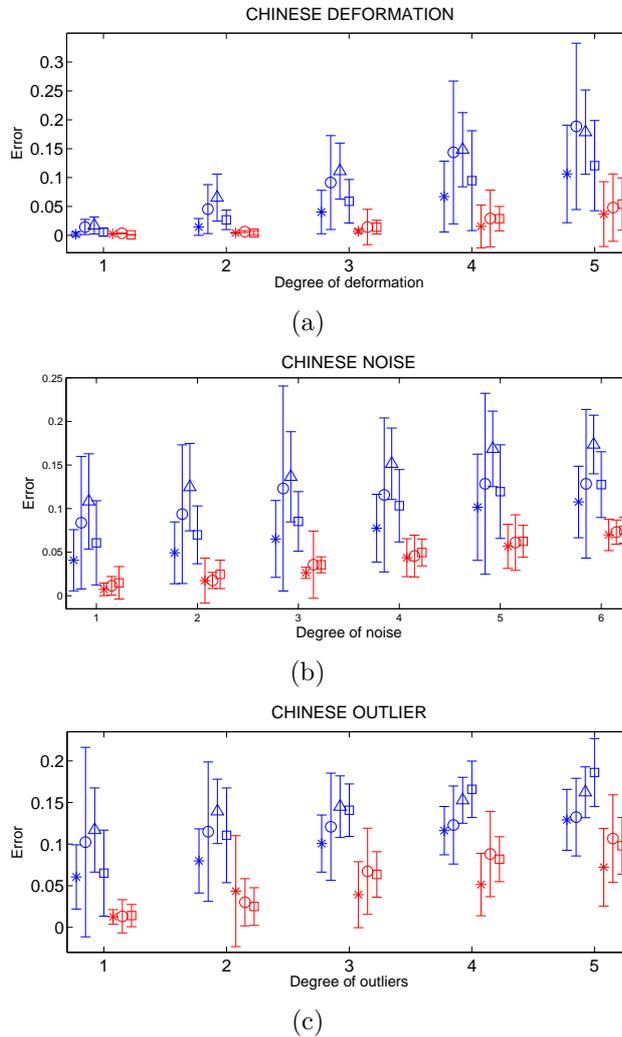


Figure 3: Results for the fish (a) deformation, (b) noise, and (c) outlier experiments. In blue the rigid methods: (\*) our method, (O) our method with  $P_e = 0.5$  (i.e., Softassign Procrustes), ( $\Delta$ ) Graduated Assignment and ( $\square$ ) Unified approach by Luo and Hancock (2003). In red the non-rigid methods: (\*) our method, (O) TPS-RPM and ( $\square$ ) CPD.

performance than Coherent Point Drift but with somewhat higher standard deviations in the deformation and noise experiments. In the outlier experiments our method outperforms the rest of the methods. TPS-RPM obtains comparable performance but with higher standard deviations, most of the times.

## 6.2. Real Data

We have performed image matching experiments with some databases from <http://www.featurespace.org> that hosts image databases which are commonly used for performance evaluation of local image detectors and descriptors. We have used the datasets *BOAT*, *EASTPARK*, *EASTSOUTH* and *RESID* from INRIA (France), each one containing a sequence of images showing a scene across different zooms and rotations.

Each sequence, containing between 10 and 11 images, can be ordered according to the variation in zoom. We perform a sort of narrow-baseline matching by using only adjacent image pairs from the ordered sequence. So, results for each dataset are the mean of 18 or 20 experiments. It is very difficult to handle the high amounts of clutter usually present in wide-baseline matching using non-discriminant features such as the arrangement of two sparse sets of points. Such a problem is more accurately driven with the use of discriminant features such as local image feature vectors.

Points are extracted with the *scale-invariant feature detector* by Lowe (2004) that locates points at the scale-space extrema of a *Difference-of-Gaussians* function. For each image, we keep the 50 points with the highest scales. We have chosen this point-set size since it represents the limit for our implementation of the Dual-Step method to execute in reasonable time. We have placed the edges between points by using a *Delaunay triangulation*.

We have compared the proposed method, the pure geometric version of our method, the Unified approach by Luo and Hancock (2003), the Dual-Step method by Cross and Hancock (1998) and the Graduated Assignment by Gold and Rangarajan (1996).

As in the synthetic experiments, we evaluate the results through the mean correspondence error. Since it is available the ground truth homography between each pair of images we can compute the ground truth projection onto the second image from a point in the first image. The mean correspondence error is then computed as the Euclidean distance between the ground truth projection of a point and the point where it has been actually assigned to.

All the methods have been initialized with a matching by correlation with a fixed window size. We have used the orientation of each point provided by the detector in order to achieve a certain invariance to rotations in the initialization. We have included the results of the matching by correlation in the comparisons.

As said, all the images are related by a similarity transformation. It is expected a moderate amount of structural corruption due to clutter. These

two facts lead us to reduce the specific weight of the structural measurements by increasing its uncertainty. Therefore, we have experimentally set  $P_e = 0.3$  for the proposed method and the Unified method. We have experimentally set the outlying parameter  $N = 0.5$  for the full version of our method and  $N = 1.25$  for the pure geometric one. The value of  $N = 1.25$  has been set so that it returns a similar number of correspondences than the full version. We have experimentally set the values  $P_e = 0.1$  and  $\rho = 0.0001$  for the Dual-Step method. We have experimentally set the compatibility coefficient of Graduated Assignment to  $c = 3.47$ .

Figure 4 shows the results obtained by each method in each database and figure 5 shows some matching examples.

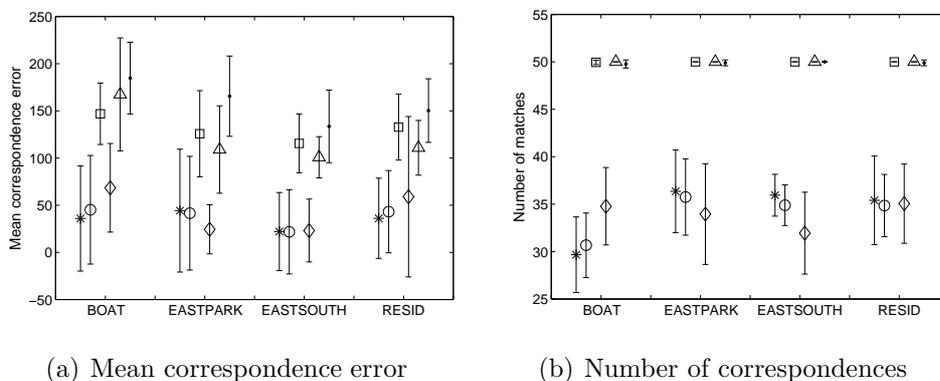


Figure 4: (a) Mean correspondence errors (in pixels) and (b) number of matches returned by (\*) our method, (O) our method with  $P_e = 0.5$  (i.e., Softassign Procrustes), ( $\square$ ) Unified approach by Luo and Hancock (2003), ( $\diamond$ ) Dual-Step by Cross and Hancock (1998), ( $\triangle$ ) Graduated Assignment and ( $\cdot$ ) matching by correlation (initialization).

Notice the superiority in terms of mean correspondence errors of the methods incorporating outlier rejection mechanisms. As you can see in figure 4(b), all the methods that do not reject outliers tend to match all available points. This penalizes in terms of mean correspondence errors due to the usual presence of clutter in the image matching experiments.

With regards to the methods incorporating outlier rejection, all of them present comparable accuracies with a similar number of matched points. Since all the images used in the experiments are related by a similarity transformation, the transformation-invariant geometric model used by our method has succeeded in accommodating the underlying transformations. Therefore,

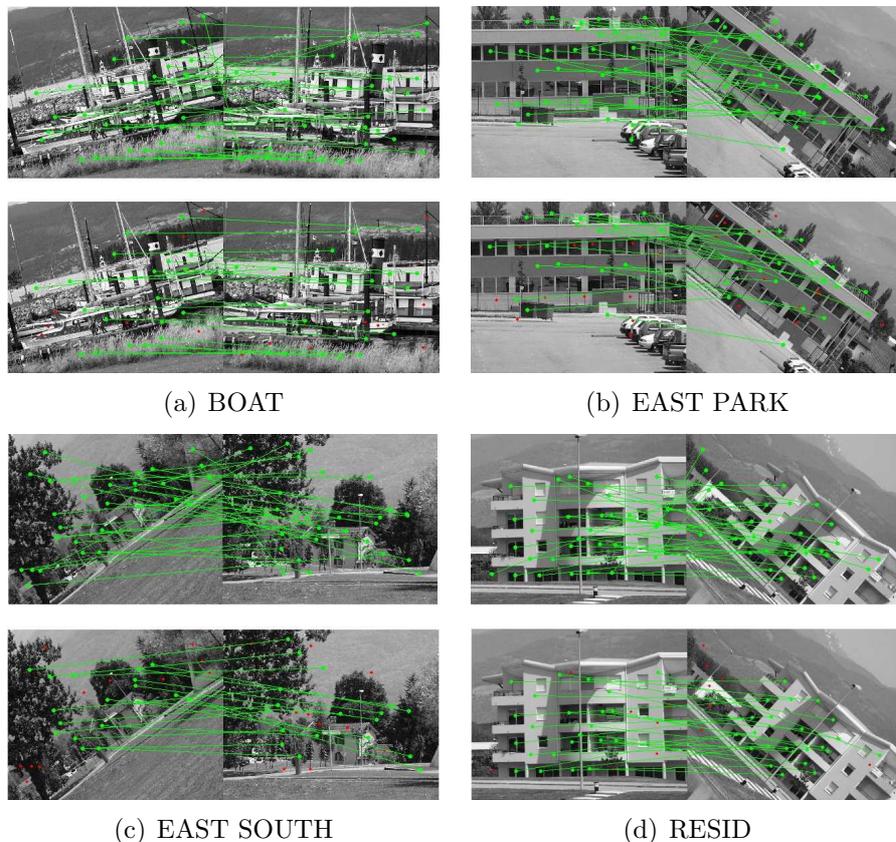


Figure 5: Some example matches from the BOAT, EAST PARK, EAST SOUTH and RESID databases. In each subfigure, top: match by correlation results (used as initialization), bottom: results after applying our method.

we have found no significant advantages from incorporating structural constraints. However, the full version of our method presents slightly better accuracies in the *BOAT* and *RESID* databases.

The Dual-Step method presents a similar accuracy than ours but requires considerably higher time to execute. The mean computational times and standard deviations obtained in all the databases by the Dual-Step method are  $799.31 \pm 78.99$  sec.; and those obtained by our method are  $1.68 \pm 0.2$  sec. All the methods have been implemented in Matlab and executed on an Intel Xeon CPU E5310 at 1.60GHz.

## 7. Conclusions

We have presented an ensemble approach for structural graph matching and point-set alignment that benefits from the additional representational facilities of graphs with respect to point-sets. We pose the problem as one of maximum likelihood estimation of correspondence and alignment parameters from a mixture distribution. Our mixture model assumes that point position errors and structural errors follow Gaussian and Bernoulli distributions, respectively. We derive the EM algorithm according to the proposed mixture model where alignment and correspondence parameters are estimated in conditional maximization steps. As opposed to other methods, our method uses a continuous correspondence variable. We use Softassign in order to compute the correspondence indicators at each iteration. An annealing procedure is implemented by updating a control parameter within Softassign at each maximization step. Outliers are gradually rejected on the basis of the number of standard deviations allowed in the alignment errors. We have performed matching experiments on synthetic and real data.

With regards to the synthetic experiments, the combination of both geometric and structural constraints proposed by our method has resulted in a superior performance than any of its parts separately. The proposed model of structure and geometry as well as the recovery of continuous correspondences through Softassign has resulted in superior performance than the Unified graph matching method by Luo and Hancock (2003) which also combines geometry and structure and updates correspondences in a discrete fashion.

The non-rigid extension proposed has shown comparable performance to state-of-the-art methods in the deformation and noise experiments. Our method outperforms the rest in the outlier experiments.

In the image matching experiments the methods with outlier rejection capabilities have performed the best, due to the usual presence of clutter in these types of experiments. There are no significant differences between the performance of the full and the pure geometric versions of our method in these experiments. This is because the similarity transformation model adjusts fairly well to the underlying geometry of the problem. The Dual-Step method by Cross and Hancock (1998) present a roughly similar performance than ours but takes a considerably higher time.

## Acknowledgements

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

## References

- Berge, J., March 2006. The rigid orthogonal procrustes rotation problem. *Psychometrika* 71 (1), 201–205.
- Chui, H., Rangarajan, A., 2000. A new algorithm for non-rigid point matching. In: *CVPR*. pp. 2044–2051.
- Chui, H., Rangarajan, A., February 2003. A new point matching algorithm for non-rigid registration. *Comput. Vis. Image Underst.* 89 (2-3), 114–141.
- Cootes, T. F., Taylor, C. J., Cooper, D. H., Graham, J., January 1995. Active shape models: their training and application. *Comput. Vis. Image Underst.* 61, 38–59.
- Cross, A., Hancock, E., NOV 1998. Graph matching with a dual-step em algorithm. *IEEE Trans. Pattern Analysis and Machine Intelligence* 20 (11), 1236–1253.
- Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B* 39 (1), 1–38.
- Dryden, I. L., Mardia, K., 1998. *Statistical shape analysis*. Wiley series in probability and statistics: Probability and statistics. J. Wiley.
- Gold, S., Rangarajan, A., April 1996. A graduated assignment algorithm for graph matching. *IEEE Trans. Pattern Analysis and Machine Intelligence* 18 (4).
- Gold, S., Rangarajan, A., Lu, C.-P., Pappu, S., Mjolsness, E., 1998. New algorithms for 2d and 3d point matching: pose estimation and correspondence. *Pattern Recognition* 31 (8), 1019–1031.
- Haralick, R. M., Joo, H., Lee, C., Zhuang, X., Vaidya, V. G., Kim, M. B., 1989. Pose estimation from corresponding point data. *Systems, Man and Cybernetics, IEEE Transactions on* 19 (6), 1426–1446.

- Hartley, R. I., Zisserman, A., 2000. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049.
- Ho, J., Yang, M.-H., January 2011. On affine registration of planar point sets using complex numbers. *Comput. Vis. Image Underst.* 115, 50–58.
- Horaud, R., Forbes, F., Yguel, M., Dewaele, G., Zhang, J., March 2011. Rigid and articulated point registration with expectation conditional maximization. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 587–602.
- Kendall, D. G., 1984. Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society* 16 (MAR), 81–121, pT: J.
- Lowe, D. G., January 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60 (2).
- Luo, B., Hancock, E., October 2001. Structural graph matching using the em algorithm and singular value decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23 (10).
- Luo, B., Hancock, E., OCT 2003. A unified framework for alignment and correspondence. *Computer Vision and Image Understanding* 92 (1), 26–55.
- Meng, X.-L., Rubin, D. B., 1993. Maximum likelihood estimation via the em algorithm: A general framework. *Biometrika* 80, 267–278.
- Myronenko, A., Song, X. B., 2010. Point set registration: Coherent point drift. *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (12), 2262–2275.
- Rangarajan, A., Chui, H., Bookstein, F. L., 1997. The softassign procrustes matching algorithm. In: *Proceedings of the 15th International Conference on Information Processing in Medical Imaging*. Springer-Verlag, pp. 29–42.
- Sinkhorn, R., 1964. Relationship between arbitrary positive matrices + doubly stochastic matrices. *Annals of Mathematical Statistics* 35 (2), 876–&.
- Umeyama, S., April 1991. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13, 376–380.

Yuille, A. L., Grzywacz, N. M., 1989. A mathematical analysis of the motion coherence theory. *International Journal of Computer Vision* 3 (2), 155–175.