Hybrid Genetic Algorithm and Procrustes Analysis for Enhancing the Matching of Graphs Generated from Shapes

Gerard Sanromà[1,⋆], Francesc Serratosa[1], and René Alquézar[2]

[1] Departament d'Enginyeria Informàtica i Matemàtiques,
Universitat Rovira i Virgili, Spain
{gerard.sanroma,francesc.serratosa}@urv.cat
[2] Departament de Llenguatges i Sistemes Informàtics,
Universitat Politècnica de Catalunya, Spain
alquezar@lsi.upc.es

Abstract. Typically, graphs generated via skeletonization of shape images are small and present low structural constraints. This fact constitutes a source of ambiguities for structural matching methods. Hybrid Genetic Algorithms have been effectively used for graph matching. This paper presents a new method which combines Hybrid Genetic Search with an enhanced model for graph matching. This enhanced model is based on the cliques model by Wilson and Hancock but introduces Procrustes Analysis over positional information in order to eliminate ambiguities. Comparative results are presented of the performance of the Hybrid Genetic algorithm with both the original cliques model and the enhanced model.

Keywords: graph matching, hybrid genetic algorithm, procrustes analysis, shape analysis, cliques model.

1 Introduction

The aims of graph matching is to associate nodes in a data graph with nodes in a model graph [1][2][3]. Nodes are associated so that a compatibility measure is maximized, usually computed using only the structural constraints (edges) provided by the graphs under match. While some authors exploit the relations of edge existence/absence between pairs of nodes [1][3], Wilson and Hancock use the overall adjacency relations of each node [2].

There is a variety of ways to represent an object by the means of a graph. For instance, given an image of a shape, we can extract the edges through a skeletonization, and place the nodes at end points, junction points and high curvature points. Graphs generated from shapes (e.g. handwritten letters) through skeletonization are not necessarily large and they also present low structural constraints (i.e. low connectivity between nodes). Note that the best connected node,

⋆ Corresponding author.

N. da Vitora Lobo et al. (Eds.): SSPR&SPR 2008, LNCS 5342, pp. 298–307, 2008.

has at most three edges, due to the nature of the skeletonization process. Structural matching methods often fail to match graphs under these circumstances. Purely structural criterions often generate ambiguous compatibility measures. Ambiguity refers to the fact of assigning the same amount of probability to a number of matching alternatives, most of them intuitively wrong.

The model by Wilson and Hancock (Section 3) provides good means for matching graphs with high or even moderate structural constraints, since it exploits the overall adjacency relations of each node [2]. It has demonstrated to be powerful when applied to large graphs of road maps and also Delaunay graphs. Delaunay triangulations, even on reduced point-sets, provide structural information enough in order to eliminate ambiguities. In the case of large graphs, local ambiguities are eliminated when evaluating the context.

Graphs generated from shapes are typically small and present low structural constraints. Delaunay triangulation is not a feasible choice for augmenting the structural constrains because the information about the shape contour provided by the edges from the skeletonization would be lost. There is a need of an extra source of information.

Landmark positions are traditionally used in Statistical Shape Analysis [4]. The main inconvenience of using positional information is that it depends on the position, orientation and scale of the point-sets. Procrustes analysis (Section 2), provide a framework to compare such point-sets in a way invariant to Euclidean similarity transformations. Our aim is to use Procrustes methods to augment the structural criterion proposed by Wilson and Hancock [2] with a measure of positional error between two point-sets and thereby, eliminate ambiguities with a more fine-grained compatibility measure.

Wilson and Hancock [2] propose an optimization procedure based on probabilistic relaxation. It consists on a gradient ascent process that updates the matching configuration through iterations. Graph-matching applications are characterized for presenting highly irregular search spaces. This fact makes gradient-ascent optimization processes, potentially able of getting trapped in local optima. Genetic Algorithms (GA) are optimization processes effectively capable of locating global optimal solutions [5]. A population of individuals is evolved in GA by the means of mutation and crossover operations in a manner similar than the natural evolution does. Convergence to the global optimum is guaranteed by selection of the most fit individuals to pass to the next generation. The fitness function, that measures the adaptation of each individual, is a key point on the genetic search.

Several works have been presented that use Hybrid GAs for optimization in graph-matching processes. Hybrid GAs arise from the combination of GAs with further optimization procedures (e.g. gradient ascent). The two ones most related with the work presented here [6][7], consist on the implementation of a hybrid GA to match graphs according with the model by Wilson and Hancock [2] (Section 3). We demonstrate that better graph-matching results are achieved applying hybrid genetic search over the augmented model presented in this paper (Section 4), than over the original one.

In Section 2 we describe Procrustes analysis, in Section 3, the cliques model by Wilson and Hancock [2] is described, in Section 4 we define our contribution to the cliques model in order to improve the matching of graphs generated from shapes, in Section 5 details on the genetic search procedure used are given, in Section 6 results are presented and finally, in Section 7 some conclusions are given.

2 Procrustes Analysis

Consider two correspondent point-sets X_1 and X_2 arranged in $k \times m$ matrices of cartesian coordinates of k points in m dimensions.

Procrustes Analysis involves the least squares superimposition of two point-sets using the similarity transformations. Estimation of the similarity parameters γ, Γ and β is carried out by minimizing the squared Euclidean distance

$$D^2_{PA}(X_1, X_2) = \inf_{\Gamma,\beta,\gamma} \parallel X_2 - \beta X_1 \Gamma - 1_k \gamma^T \parallel^2 , \tag{1}$$

where $\parallel X \parallel = \text{trace}\left(X^T X\right)^{\frac{1}{2}}$ is the Euclidean norm, $\beta \in \Re^+$ is a positive scaling factor, $\Gamma \in SO(m)$ is a $m \times m$ rotation matrix and $\gamma \in \Re^m$ is a translation m-vector.

Following Dryden and Mardia [4], to obtain the minimum of Equation (1), first we substitute X_1, X_2 by their centered versions normalized to unit size $Z_1 = \frac{CX_1}{\|CX_1\|}$ and $Z_2 = \frac{CX_2}{\|CX_2\|}$, where $C = I_k - \frac{1}{k} 1_k 1_k^T$ is the centering matrix which applies a translation that brings the centroids of X_1 and X_2 to the origin.

Afterwards, the minimizing parameters are found at

$$\hat{\gamma} = 0 , \quad \hat{\beta} = \frac{\text{trace}\left(Z_2^T Z_1\right)}{\text{trace}\left(Z_1^T Z_1\right)} , \quad \hat{\Gamma} = UV^T , \tag{2}$$

where $V \Lambda U^T = \text{svd}\left(Z_2^T Z_1\right)$ is the *singular value decomposition* of $Z_2^T Z_1$.

Alternatively, we may use the centered non-normalized versions $X_1^C = CX_1$ and $X_2^C = CX_2$ with $\hat{\beta} = 1$ if we want to preserve the original size of the point-sets when calculating the distance measure.

3 A Gradient Ascent Approach to Graph Matching

In this Section, the graph matching technique used in the hill-climbing step of the hybrid genetic algorithm is described. The aim of graph matching is to associate nodes in a data-graph $G_1 = (V_1, E_1, A_1)$, where V_1 is a set of nodes, E_1 is a set of arcs and $A_1 = \left\{ x_u^{(1)}, \forall u \in V_1 \right\}$ is a set of symbols associated with the nodes, against those in a model-graph $G_2 = (V_2, E_2, A_2)$, where $A_2 = \left\{ x_v^{(2)}, \forall v \in V_2 \right\}$. The matching is represented by a function $f : V_1 \rightarrow V_2 \cup \phi$ for the nodes in the data graph G_1 to those in the model graph G_2 augmented with a null-label.

In [2] the graphs are matched using *cliques*. For a given node indexed $\jmath$ from G_1, the clique $C_\jmath = \left\{u_1, \ldots, u_{|C_\jmath|}\right\}$ is defined as $C_\jmath = \jmath \cup \{\imath | \, (\imath, \jmath) \in E_1\}$. The matched realization of the clique $C_\jmath \subseteq V_1$ is denoted as

$$\Gamma_\jmath = \left\{f(u_1), \ldots, f(u_{|C_\jmath|})\right\} \ . \tag{3}$$

For each clique on the model graph $S_\imath$, a set of *structure-preserving mappings* (SPMs) are generated in order to obtain a dictionary $\Theta_\jmath = \left\{S_\imath^k\right\}$ onto which we evaluate each data graph clique $C_\jmath$. The generation of the SPMs is a key point consisting in two steps: (1) the model graph clique $S_\imath$ is padded with dummy nodes in order to accommodate size differences with $C_\jmath$ due to node loss and merging and, (2) the nodes of the padded model graph relations are then permuted so as to preserve the cyclicity of the non-dummy nodes. This process effectively preserves the adjacency structure of the model graph nodes while leaving dictionary invariant to potential scene translations, scalings or rotations [2].

The final gradient-ascent rule to update the matching configuration according to Ref. [2] is

$$f(u) = \underset{v^{(2)} \in V_2 \cup \phi}{\arg\max} P\left(f(u^{(1)}) = v^{(2)} | x_u, x_v\right) P\left(f\right) \ , \tag{4}$$

where $P\left(f(u^{(1)}) = v^{(2)} | x_u, x_v\right)$ is the probability of match between nodes $u^{(1)} \in V_2$ and $v^{(2)} \in V_2$ given measurements only relative to the nodes under match, and $P\left(f\right)$ is the joint *prior* which gauges the overall consistency of the matching configuration. The most critical of these is a means of modeling the joint prior $P\left(f\right)$. In the following we describe the developement of this structural consistency measure.

This process commences by modelling the consistency of match of an individual clique $C_\jmath$ onto G_2, i.e. $\Gamma_\jmath$. The dictionary of SPMs $\Theta_\jmath$ represents the consistent labellings available for gauging the quality of match. As demanded by the Bayes rule:

$$P\left(\Gamma_\jmath\right) = \sum_{S_\imath^k \in \Theta_\jmath} P\left(\Gamma_\jmath | S_\imath^k\right) P\left(S_\imath^k\right) \ . \tag{5}$$

Assuming independence in the matching errors between adjacent nodes of the same clique, the conditional probabilities become

$$P\left(\Gamma_\jmath | S_\imath^k\right) = \prod_{r=1}^{|S_\imath^k|} P\left(f(u_r) | v_r\right) \ . \tag{6}$$

Sources of errors are assumed to occur with uniform probability P_e. Under these assumptions, the distribution rule under the product of (6) is

$$P\left(f(u_r) | v_r\right) = \left(1 - P_e\right)^{\delta_{u_r v_r}} P_e^{1 - \delta_{u_r v_r}} \ , \tag{7}$$

where δ_{ij} is the Kronecker delta which is a function of two variables, usually integers, which is 1 if they are equal, and 0 otherwise.

Combining this distribution with Equation (6) we obtain the following expression for the conditional probabilities:

$$P\left(\Gamma_J|S_i^k\right) = P_e^{H\left(\Gamma_J,S_i^k\right)}\left(1-P_e\right)^{E\left(\Gamma_J,S_i^k\right)} \ , \tag{8}$$

where $H\left(\Gamma_J, S_i^k\right)$ is the number of conflicts between Γ_J and S_i^k (i.e. the Hamming distance), and $E\left(\Gamma_J, S_i^k\right)$ is the number of coincidences between Γ_J and S_i^k.

Assuming that each of the SPMs is equi-probable, i.e. $P\left(S_i^k\right) = \frac{1}{|\Theta_J|}$ and using the natural exponential form for expressing Equation (8), the final model for the clique matching probabilities is

$$P\left(\Gamma_J\right) = \frac{1}{|\Theta_J|}\sum_{S_i^k\in\Theta_J}\exp\left[-k_{er}H\left(\Gamma_J,S_i^k\right) - k_{eq}E\left(\Gamma_J,S_i^k\right)\right] \ , \tag{9}$$

where $k_{er} = \log\left[\frac{1}{P_e}\right]$ and $k_{eq} = \log\left[\frac{1}{(1-P_e)}\right]$.

Finally, the joint prior is computed by averaging the clique matching probabilities over the nodes of the data graph, i.e.

$$P\left(f\right) = \frac{1}{|V_1|}\sum_{J\in V_1} P\left(\Gamma_J\right) \ . \tag{10}$$

The matching configuration is updated using the MAP decision rule given in (4) and according to a Discrete Relaxation scheme [2]. The strategy is to set P_e to an initial high value to reflect a poor labelling, and reduce it through iterations.

In the next Section we describe our contributions to the model above, aimed to improve the matching of graphs generated from shapes.

4 Augmenting the Matching Criterion with Procrustes Analysis

Graphs generated from shapes are typically small and they present low structural constraints. These two facts together are a potential source of ambiguity for purely structural matching methods. Our aim is to augment the measure of structural consistency reported in [2] (Equation 10) with a measure of similarity based on the positions of the nodes. By doing so, we assume that corresponding nodes in different graphs keep similar positional relations with their neighboring nodes. Note that, although articulated shapes would break this assumption in some extent, the study of such effects is over the scope of this paper.

To start developing our model, we augment the set of unary measurements associated with the nodes with the coordinates of the feature points. Hence, $A = \{(x_i, \overrightarrow{\alpha_i}), \forall i \in V\}$, where $\overrightarrow{\alpha_i} = (a_1, \ldots, a_m)$ is a row m-vector of cartesian coordinates in m dimensions, associated with each node.

The underpinning idea is to weight the contributions of the consistently mapped nodes (i.e, those such that $f(u) = v$) with a gaussian error derived

from the Euclidean distance and the expected variance-covariance, after the Procrustes superimposition. By introducing such errors, we augment the measure of structural consistence drawn up in Section 3 with a measure of shape similarity between cliques. Thus, different cliques with the same cardinality on a given graph are no longer susceptible to contribute the same amount to the energy functional.

Our expression for the marginal probabilities of (7) is

$$P\left(f(u_r), \alpha_{u_r} | v_r, \alpha_{v_r}\right) = P_{match}\left(\overrightarrow{\alpha'_{u_r}}, \overrightarrow{\alpha_{v_r}}\right)^{\delta_{u_r v_r}} \rho^{1-\delta_{u_r v_r}} , \tag{11}$$

where δ_{ij} is the Kronecker delta, ρ is a constant controlling the inconsistent matching process in a similar manner than P_e did in Section 3 and, P_{match} means the probability of matching nodes $u_r^{(1)}$ with $v_r^{(2)}$ according with the similarity between the shapes of the cliques they belong to. We must take care to set ρ to be always lower than the probability of matching two corresponding nodes $u_r^{(1)}$ and $v_r^{(2)}$ according to P_{match}. Notice that, while weighting the contributions of the consistently mapped nodes, we aim to preserve the idea by Wilson and Hancock [2] of penalizing structural inconsistencies with low contributions to the overall measure. Otherwise, the gradient-ascent scheme would lead to a structurally inconsistent state.

As said, P_{match} is a gaussian error function between the two points, and it is defined as

$$P_{match}\left(\overrightarrow{\alpha'_{u_r}}, \overrightarrow{\alpha_{v_r}}\right) = \exp\left[-\frac{1}{2}\left(\overrightarrow{\alpha_{v_r}} - \overrightarrow{\alpha'_{u_r}}\right) \Sigma^{-1} \left(\overrightarrow{\alpha_{v_r}} - \overrightarrow{\alpha'_{u_r}}\right)^T\right] . \tag{12}$$

In the computation of $P_{match}\left(\overrightarrow{\alpha'_{u_r}}, \overrightarrow{\alpha_{v_r}}\right)$, the transformed coordinates $\overrightarrow{\alpha'_{u_r}}$ resulting from the Procrustes Analysis are used. Therefore, a point-set X_1 is built from the cartesian coordinates of the data graph nodes that have been correctly assigned in the current mapping S_i of the matching conditional $P\left(\Gamma_j | S_i\right)$ (i.e. the data graph nodes u_r under the circumstance $f(u_r) = v_r$). Thus, $X_1 = \{\overrightarrow{\alpha_{u_r}} | f(u_r) = v_r, r = 1, \ldots, |S_i|, f(u_r) \in \Gamma_j\}$ is a $k \times m$ matrix of cartesian coordinates of k points in m dimensions. Equivalently, a point set X_2 is built from the cartesian coordinates of the model graph nodes on which the data graph nodes have been correctly mapped to. Thus, $X_2 = \{\overrightarrow{\alpha_{v_r}} | f(u_r) = v_r, r = 1, \ldots, |S_i|, v_r \in S_i\}$ is a $k \times m$ matrix as well.

Finally, the new coordinates of the data graph $\overrightarrow{\alpha'_{u_r}}$ are taken from the transformed matrix $X_1^P = \beta X_1 \Gamma - 1_k \gamma^T$, after being aligned with X_2 using Procrustes Analysis (Section 2).

Collecting terms and using the natural exponential form, our model for the clique matching probabilities becomes

$$P\left(\Gamma_J\right) = \frac{1}{|\Theta_J|} \sum_{S_i \in \Theta_J} \exp\left[-k_{er} H - \frac{1}{2} \text{trace}\left(\left(X_2 - X_1^P\right) \Sigma^{-1} \left(X_2 - X_1^P\right)^T\right)\right] , \tag{13}$$

where $k_{er} = \log\left[\frac{1}{\rho}\right]$. Note that the second term in Equation (8), is replaced in our model by a product of terms as the one Equation (12). This product of exponential terms can be turned into a sum of the exponents when the clique matching probabilities are expressed in the exponential form. Finally, this sum can be expressed as a squared Euclidean distance similar than the one in Equation (1) but normalized by a variance-covariance matrix (i.e. a squared Mahalanobis distance).

The joint prior is computed as in Equation (10) by averaging the clique matching probabilities over the nodes of the data graph.

Details on the implementation of the hybrid genetic search are given in the following Section.

5 Hybrid Genetic Search

Genetic search has effectively proved to be capable of finding global optimal solutions in a wide variety of applications. A population of feasible solutions is evolved by the means of mutation and crossover operations.

A match $f : V_1 \rightarrow V_2$ from a data-graph G_1 to a model graph G_2 is represented as a list of length $|V_1|$ of labels drawn from V_2.

Mutation consists on randomly reassigning the matches on individual sites with uniform probability. These reassignments are done with random labels selected from V_2.

While the works in Refs. [6] and [7] use the measure of structural consistence reported by Wilson and Hancock (Equation (10)) as fitness measure, the present work use the augmented measure of matching consistence described in Section 4.

The approaches in Refs. [6] and [7], present a genetic search augmented with the gradient ascent approach reported by Wilson and Hancock [2] (Section 3) based on discrete relaxation. Mutation and crossover operators can leave the matches into inconsistent states (e.g. two different data-graph nodes assigned to the same model-graph node). The main aims of the gradient ascent step applied after the mutation and crossover operators are, firstly to return the matches to a consistent state, and secondly to push them towards local optima via hill-climbing.

In the present work, a gradient ascent step is conducted as stated by Equation (4) using the expression developed in Section 4 as the joint *prior*.

Gradient ascent methods are capable of correcting to around 70% of the initialization errors in the type of graphs used here. The choice of population size, as stated by Myers and Hancock [7], is made so that at least a given fraction (e.g. 30%) of the nodes are correctly assigned in a random initialization. Denoted as P_s, the probability that at least some fraction s of the nodes will have at least one correct assignment can be expressed as $P_s = \sum_{s|V_1| \leq k \leq |V_2|} \binom{|V_1|}{k} P_c^k \left(1 - P_c\right)^{|V_1|-k}$, where $P_c = 1 - \left(1 - \frac{1+r}{|V_2|+1}\right)^n$ is the probability of at least one correct assignment appearing in the initial population of size n with a fraction r of nodes corrupt.

6 Experiments

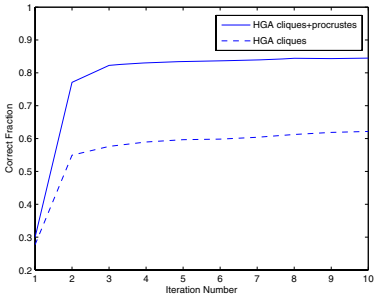
We present two types of experiments, evaluating both the matching efficiency and the convergence rate. We have used 84 graphs extracted from handwritten capital letters. Samples present moderate levels of structural corruption consisting on few added extra nodes and extra arcs (notice that we are not evaluating our model under severe structural corruption). Data graphs of each class are matched against a prototypical graph of that class. Comparative results are presented between the hybrid approach with the augmented cliques model reported in this paper (HGA cliques+procrustes) and the hybrid approach with the original cliques model (HGA cliques). Population is randomly initialized in all the experiments. A known ground truth consisting on the intuitively correct assignments is available for all the matchings.

For the matching efficiency experiments we have set the mutation and crossover probabilities to 0.4 and 0.5, respectively, and the population size such that a fraction of 0.3 assignments are correct in the initial population. Correct assignments fraction is obtained by comparison with the ground truth.

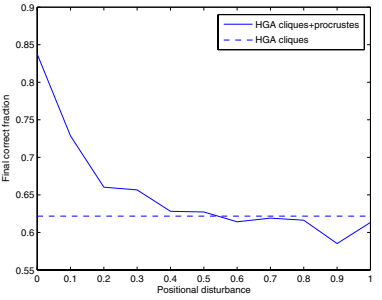
Figure 1.(a) shows the correct assignments fraction at each iteration. Fraction of correct assignments is computed at each iteration by averaging this quantity among the fittest individuals. Results at the first iteration correspond to initial population with neither gradient ascent nor genetic operator steps.

The next experiment evaluates the tolerance of our method to severe noise in the coordinates of the nodes. We have applied gaussian white noise to the (x, y) coordinates of the point positions associated with the nodes. The variance of such noise ranges from zero to the total variance of the data. Therefore, in the extreme case the variance due to noise is the same quantity than due to data. Figure 1.(b) shows correct assignments fraction obtained at the end of the algorithm for a given value of noise in the positional information. The algorithm finalizes when either the correct match is found or 20 iterations are reached. Since our method is the only one sensitive to this kind of noise, we have plotted the final correct fraction of the *HGA cliques* method as a base-line.

Results show that our method evolves significantly better through iterations. A correct fraction of 85% is reached at iteration 10 of our method while, *HGA cliques* is slightly over 60%. Because of ambiguities in the purely structural criterion, a number of global optima arise when using the *HGA cliques* method. Therefore, the average correct fraction among the fittest individuals degrades the results. Although a more fine-grained measure is used in the *HGA cliques+procrustes*, it does not always find the matching corresponding to the ground-truth. Is for this reason that averaged results are below 1. It is interesting to note that the correct fraction in the randomly initialized population (around 30%) is consistent with the criterion used in the choice of the population size. Results also show that our method outperforms *HGA cliques* while positional noise is under 75% approximately. It is interesting to note that positional noise degrades the effectiveness of our method just until a given threshold (approximately 20%). After that threshold, it stabilizes.

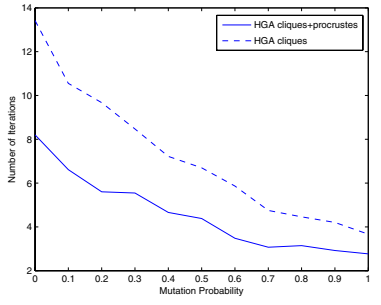


(a) Correct assignments fraction at each iteration of the algorithm. Results at the first iteration correspond to initial populaiton with no mutation, crossover or gradient ascent.

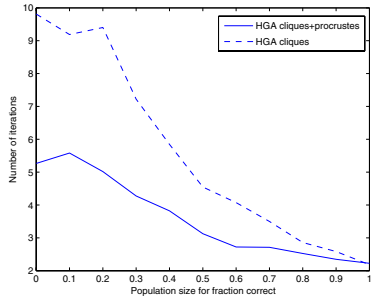


(b) Correct assignments fraction versus positional noise fraction. Since *HGA cliques* method is not sensible to this kind of noise the final correct fraction is plotted as a base-line.

Fig. 1. Matching efficiency results



(a) Iterations until convergence versus mutation probability



(b) Iterations until convergence versus correct fraction required in the choice of population size

Fig. 2. Convergence rate results

Convergence rate experiments evaluate the number of iterations needed to find the match corresponding to the ground-truth, under mutation probability and population size variations. Otherwise, the algorithm is stopped at iteration 20. Figures 2.(a) and 2.(b) show the number of iterations needed to converge versus mutation probability and population size, respectively. The mutation probability and the correct fraction required in the choice of the population size in the cases that are fixed, are set to 0.4 and 0.3, respectively.

There are a number of matching processes that do not find the correct match, as seen in Figure 1.(a). These processes contribute to the convergence rate with a number of iterations equal to the maximum permitted (20). Despite this fact, the number of iterations needed to find the correct match tend to decrease as the

mutation probability and population size increase. Anyway, our method always finds the correct match earlier.

7 Conclusions

We have presented a method for improving the matching of graphs that are small and present low structural constraints, such as those obtained from shapes (e.g. letters). We have used hybrid genetic search with an improved gradient ascent step based Wilson and Hancock's method [2] enhanced by the use of Procrustes analysis over positional information. Although we have evaluated our method with shapes it is applicable to any kind of object on which meaningful positional information can be extracted. The model presented integrates perfectly into the cliques framework. Results show a significant improvement in either matching ability and convergence rate of our method. Matching results of our method under severe positional disturbance are comparable to the ones obtained using the original hybrid approach without positional information.

Acknowledgements. This research was partially supported by Consolider Ingenio 2010, project CSD2007-00018, by the CICYT project DPI 2007-61452 and by the Universitat Rovira i Virgili (URV) through a predoctoral research grant.

References

1. Luo, B., Hancock, E.R.: Structural graph matching using the em algorithm and singular value decomposition. Pattern Analysis and Machine Intelligence 23(10) (October 2001)
2. Wilson, R.C., Hancock, E.R.: Structural matching by discrete relaxation. Pattern Analysis and Machine Intelligence 19(6) (June 1997)
3. Gold, S., Rangarajan, A.: A graduated assignment algorithm for graph matching. Pattern Analysis and Machine Intelligence 18(4) (April 1996)
4. Dryden, I.L., Mardia, K.V.: Statistical Shape Analysis. John Wiley and Sons, Chichester (1998)
5. Whitley, D., Beveridge, R., Graves, C., Mathias, K.: Test driving three 1995 genetic algorithms: New test functions and geometric matching. J. Heurist 1 (June 1995)
6. Cross, A.D.J., Wilson, R.C., Hancock, E.R.: Inexact graph matching using genetic search. Pattern Recognition 30(6) (1997)
7. Myers, R., Hancock, E.R.: Least-commitment graph matching with genetic algorithms. Pattern recognition 34 (October 1999)