Approximate Least-Squares Attributed Graph Matching via Bayesian Inference

M.A. van Wyk[1], B.J. van Wyk[1,2] and T. Janse van Rensburg[1]

[1] Rand Afrikaans University, Johannesburg, South Africa

[2] Kentron, a Division of Denel, Centrurion, South Africa

mavw@ing.rau.ac.za, ben.van.wyk@kentron.co.za, tjvr@ing.rau.ac.za

Abstract

In this paper we present a novel derivation of the polynomial time approximate Least-Squares Graph Matching (LSGM) algorithm for solving the attributed graph matching (AGM) problem. Here the algorithm is shown to follow from a Bayesian inference framework. This algorithm is robust against random differences existing between the two graphs to be matched. However, mathematical analysis of the algorithm shows that it is only suitable for solving full-graph matching problems.

Key-Words: - Graph matching, Attributed relational graphs, Least-squares matching, Combinatorial optimisation, Bayesian inference.

1. Introduction

Attributed graphs, introduced by Tsai and Fu [1], have proven to be very useful for structural description of objects. In turn, structural descriptions of objects enable objects to be compared, classified and matched based on their structural descriptions. This approach has proven to be very useful in many applications of computer vision where, typically, it is required to match objects from different images usually acquired from different sensors located at different positions [2].

The problem of matching attributed graphs is an instance of the graph isomorphism problem which is combinatorial in nature with NP computational complexity (consult reference [3] for the relevant definitions). Several algorithms have been proposed to perform graph matching during recent years. Tsai and Fu [1] proposed a tree search technique but due to the combinatorial nature of the problem this approach is impractical for all but the smallest of graphs. Other proposed algorithms include relaxation algorithms [4, 5], a symmetric polynomial transform algorithm [6], a linear programming algorithm [7] and an eigen-decomposition (EGM) algorithm [8]. These are all polynomial time algorithms for matching weighted graphs and not suited for sub-graph matching.

Recently, Gold and Rangarajan [9] proposed the Graduated Assignment Graph Matching (GAGM) algorithm which combines graduated non-convexity, two-way assignment constraints and sparsity to achieve large improvements in accuracy and speed in comparison to other matching algorithms. The computational complexity of the GAGM algorithm is $O(mm')$ where m and m' are the number of edges in the two graphs. Van Wyk, Durrani and Van Wyk [10, 11] recently proposed a Reproducing kernel Hilbert space Interpolator-based Graph Matching (RIGM) algorithm for solving attributed full- and sub-graph matching problems. The complexity of this algorithm is $O(n^3)$ where n is the number of vertices of the largest of the two graphs. The particular attribute of the RIGM algorithm is that it memory requirement is particularly small. Several more new graph matching algorithms have recently been proposed [12].

The outline of the paper is as follows. Sections 2 and 3, respectively, review AGM preliminaries and Bayesian inference followed by the Bayesian-based derivation of the LSGM algorithm in section 4. Section 5 presents numerical results followed by the conclusion.

2. AGM Preliminaries

The focus of this paper is on matching of undirected and directed attributed graphs. An *attributed graph* having n vertices, r edge attributes per edge and s attributes per vertex will be represented by the symbol

$$G = (V, E, \{\mathbf{A}_i\}_{i=1}^{r}, \{\mathbf{B}_j\}_{j=1}^{s})$$

where V is the set of vertices of the graph G, r is its set of edges, $\mathbf{A}_i \in \mathbf{R}^{n\times n}$ is the edge attribute adjacency matrix associated with the i th edge and $\mathbf{B}_j \in \mathbf{R}^{n\times 1}$ is the vertex attribute vector associated with the j th vertex. A *weighted graph* is a special case of an attributed graph for which $r = 1$ and $s = 0$.

Consider a reference graph and duplicate graph represented by $G' = (V', E', \{\mathbf{A}'_i\}_{i=1}^{r}, \{\mathbf{B}'_j\}_{j=1}^{s})$ and $G = (V, E, \{\mathbf{A}_i\}_{i=1}^{r}, \{\mathbf{B}_j\}_{j=1}^{s})$, respectively. The number of vertices of G' is $n' := |V'|$ and that of G is $n := |V|$ where we shall assume without loss of generality that $n' \geq n$. *Attributed Graph Matching* (AGM) refers to the process of matching each vertex of G uniquely with a vertex of G' such that the correspondence between attribute values is as consistent as possible. The criterion defining this consistency is a cost function which we shall introduce shortly. *Full-graph matching* (FGM) refers to matching of two graphs having the same number of vertices (i.e. $n' = n$) whereas *sub-graph matching* (SGM) refers to matching G to some sub-graph of G' (i.e. $n' > n$). G is *isomorphic* to a sub-graph of G' if there exists an $n \times n'$ permutation sub-matrix $\mathbf{P}$ such that

$$\mathbf{A}_i = \mathbf{P}\mathbf{A}'_i\mathbf{P}^T, \qquad i = 1, \ldots, r$$

and

$$\mathbf{B}_j = \mathbf{P}\mathbf{B}'_j, \qquad j = 1, \ldots, s$$

where $\mathbf{A}_i \in \mathbf{R}^{n\times n}$, $\mathbf{B}_j \in \mathbf{R}^{n\times 1}$, $\mathbf{A}'_i \in \mathbf{R}^{n'\times n'}$ and $\mathbf{B}'_j \in \mathbf{R}^{n'\times 1}$. By modelling non-idealities in the construction process of the duplicate graph by additive noise the above relationships may be expressed, more realistically, as

$$\mathbf{A}_i = \mathbf{P}\mathbf{A}'_i\mathbf{P}^T + \varepsilon\mathbf{N}_i, \qquad i = 1, \ldots, r$$

and

$$\mathbf{B}_j = \mathbf{P}\mathbf{B}'_j + \varepsilon\mathbf{M}_j, \qquad j = 1, \ldots, s$$

where $\mathbf{N}_i$ is an $n \times n$ noise matrix, $\mathbf{M}_j$ is an $n \times 1$ noise matrix and ε is related to the noise power and is assumed to be independent of the indices i and j. Now the general AGM problem can be expressed as the combinatorial optimisation problem,

$$\min_{\mathbf{P}} \left(\sum_{i=1}^{r} W_i \left\| \mathbf{A}_i - \mathbf{P}\mathbf{A}'_i\mathbf{P}^T \right\|^q + \sum_{j=1}^{s} W_{j+r} \left\| \mathbf{B}_j - \mathbf{P}\mathbf{B}_j \right\|^q \right)$$

where $\|\cdot\|$ represents some norm, $\mathbf{P} \in \mathrm{Per}(n, n')$ where $\mathrm{Per}(n, n')$ denotes the set of all $n \times n'$ permutation sub-matrices and $\{W_i\}_{i=1}^{r+s}$ is a set of non-negative weights satisfying

$$\sum_{i=1}^{r+s} W_i = 1.$$

Typical values for q are 1 and 2. The above minimisation problem is combinatorial in nature due to the constraint placed on the argument $\mathbf{P}$, namely that it must be a permutation sub-matrix. In summary, the *inexact* attributed graph matching problem (attributed sub-graph matching problem, respectively) is about finding the permutation matrix (permutation sub-matrix, respectively) that minimises the above cost function.

3. Bayesian Inference

In the context of inference or hypothesis testing Bayes' theorem is usually considered in the following form

$$\mathrm{prob}(\text{hypothesis} \mid \text{data}, I) \propto \mathrm{prob}(\text{data} \mid \text{hypothesis}, I) \times \mathrm{prob}(\text{hypothesis} \mid I)$$

where $\mathrm{prob}(A|B)$ is the probability that event A occurred given that event B has occurred. Here I represents all implicit and explicit relevant *background information* at hand. The quantity $\mathrm{prob}(\text{hypothesis} \mid I)$ is usually called the *prior probability*; it represents our state of knowledge (or ignorance) about the truth of the hypothesis. In the context of the above form of Bayes' theorem this is then modified by the *data* (or the experimental evidence) through the *likelihood function*, $\mathrm{prob}(\text{data} \mid \text{hypothesis}, I)$, to obtain the *posterior probability*, $\mathrm{prob}(\text{hypothesis} \mid \text{data}, I)$, which represents our state of knowledge of the truth of the hypothesis in the light of the available data.

For those instances where the hypothesis is parameterised by a set of parameters, say $\mathbf{H} = \{H_i\}$ we may express the posterior probability as $\text{prob}(\mathbf{H} \mid \text{data}, I)$. For these cases the hypothesis maximising the posterior probability (density) is obtained by solving the equation

$$\nabla_{\mathbf{H}} \text{ prob}(\mathbf{H} \mid data, I) = \mathbf{0}$$

for the parameter values $\{H_i\}$.

4. LSGM Algorithm via Bayesian Inference

4.1 Bayesian Framework

For the purpose of matching two attributed graphs, say $G' = (V', E', \{\mathbf{A}'_i\}_{i=1}^r, \{\mathbf{B}'_j\}_{j=1}^s)$ (the reference graph) and $G = (V, E, \{\mathbf{A}_i\}_{i=1}^r, \{\mathbf{B}_j\}_{j=1}^s)$ (the duplicate graph) by means of Bayesian inference, we write Bayes's theorem for the graph matching problem, namely

$$\text{prob}(\mathbf{P} \mid G, G') \propto \text{prob}(G, G' \mid \mathbf{P}) \times \text{prob}(\mathbf{P})$$

subject to $\mathbf{P} \in \text{Per}(n, n')$.

We have, until now, not made any assumption about the distribution of the errors in the expressions in Section 2 relating the two graphs. Assuming all these errors to be independent and identically distributed normal random variables and adopting a policy of *naïveté* expressed by the uniform prior probability density, yields the following expression for the posterior probability density function

$$p(\mathbf{P} \mid G, G') \propto \mathrm{e}^{-J(\mathbf{P})}$$

where

$$J(\mathbf{P}) = \sum_{i=1}^{r} \left\| \mathbf{A}_i - \mathbf{P}\mathbf{A}'_i \mathbf{P}^T \right\|_F^2 + \sum_{j=1}^{s} \left\| \mathbf{B}_j - \mathbf{P}\mathbf{B}'_j \right\|_2^2 \qquad (1)$$

with $\|\cdot\|_F$ the Frobenius matrix norm and $\|\cdot\|_2$ the Euclidean vector norm. Assuming ε to be known [13, pp. 28, 54, 69], the posterior pdf maximisation problem to be solved here can be expressed as the following constrained minimisation problem,

$$\min_{\mathbf{P} \in \text{Per}(n, n')} J(\mathbf{P}).$$

This is precisely the minimisation problem that needs to be solved for obtaining the approximate least-squares graph matching algorithm.

4.2 The LSGM Algorithm [14, 15]

Inspection reveals that the argument of the Frobenius norm in the expression for $J(\mathbf{P})$ is not linear in the matrix $\mathbf{P}$ thereby complicating the analysis. However, for the case $n' = n$ this difficulty can be circumvented by transforming (1) using the following identity

$$\left\| \mathbf{A} - \mathbf{P}\mathbf{A}'\mathbf{P}^T \right\|_F = \left\| \mathbf{A}\mathbf{P} - \mathbf{P}\mathbf{A}' \right\|_F \qquad (2)$$

which holds for any orthogonal matrix $\mathbf{P}$ and hence also for any $\mathbf{P} \in \text{Per}(n, n')$. This now gives the equivalent cost function

$$J_*(\mathbf{P}) = \sum_{i=1}^{r} \left\| \mathbf{A}_i \mathbf{P} - \mathbf{P}\mathbf{A}'_i \right\|_F^2 + \sum_{j=1}^{s} \left\| \mathbf{B}_j - \mathbf{P}\mathbf{B}'_j \right\|_2^2. \qquad (3)$$

For the case $n' > n$ the identity (2) becomes the following inequality

$$\begin{aligned} \left\| \mathbf{A}\mathbf{P} - \mathbf{P}\mathbf{A}' \right\|_F^2 &= \left\| \mathbf{A} - \mathbf{P}\mathbf{A}'\mathbf{P}^T \right\|_F^2 + \sum_{j \notin R(\sigma)} \left\| \mathbf{a}'_j \right\|_2^2 \\ &\geq \left\| \mathbf{A} - \mathbf{P}\mathbf{A}'\mathbf{P}^T \right\|_F^2 \end{aligned} \qquad (4)$$

for $\mathbf{P}$ any element of $\text{Per}(n, n')$. Here $\sigma: \{1, \ldots, n'\} \to \{1, \ldots, n'\}$ is the sub-permutation induced by $\mathbf{P}$, $R(\sigma)$ is its range and $\mathbf{a}'_j$ is the j th column vector of $\mathbf{A}'$. It is evident for the case $n' > n$ that, if $\mathbf{P}$ minimises $J_*(\cdot)$ and $J_*(\mathbf{P}) = 0$, then $\mathbf{P}$ also minimises $J(\cdot)$. However, if $\mathbf{P}$ minimises $J_*(\cdot)$ but $J_*(\mathbf{P}) \neq 0$ then $\mathbf{P}$ does not necessarily minimise $J(\cdot)$. Conversely, even if $\mathbf{P}$ minimises $J(\cdot)$ and $J(\mathbf{P}) = 0$, there is no guarantee that $\mathbf{P}$ minimises $J_*(\cdot)$. Despite this problem that occurs for the case $n' > n$ we shall proceed with our investigation of the minimisation problem based on $J_*(\cdot)$.

The next step is to solve this new minimisation problem. Firstly, observe the following equivalence,

$$\mathbf{E} = \mathbf{AP} - \mathbf{PA}' \quad \Leftrightarrow \quad \mathrm{vec}(\mathbf{E}) = \mathbf{M}_{\mathbf{AA}'}\mathrm{vec}(\mathbf{P})$$

where $\mathbf{M}_{\mathbf{AA}'} \in \mathbf{R}^{nn' \times nn'}$ has the form

$$\mathbf{M}_{\mathbf{AA}'} = \mathbf{M}_{\mathbf{A}} + \mathbf{M}_{\mathbf{A}'}$$

with

$$[\mathbf{M}_{\mathbf{A}}]_{i,j-(i-1)\bmod n+i-1} := a_{(i-1)\bmod n+1,j}$$

for $i = 1,\ldots,nn'$ and $j = 1,\ldots,n$, $\mathbf{A} := (a_{ij})$ and

$$[\mathbf{M}_{\mathbf{A}'}]_{i,(i-1)\bmod n+n(j-1)+1} := -a'_{j,[(i-1)-(i-1)\bmod n]/n+1}$$

for $i = 1,\ldots,nn'$ and $j = 1,\ldots,n'$. All remaining elements of these two matrices are zero. Using this equivalence and defining $\mathbf{p} := \mathrm{vec}(\mathbf{P})$ we obtain

$$\begin{aligned} J_*(\mathbf{p}) &= \sum_{i=1}^{r} \left\| \mathbf{M}_{\mathbf{A}_i\mathbf{A}'_i}\mathbf{p} \right\|_2^2 + \sum_{j=1}^{s} \left\| \mathbf{B}_j - \mathbf{N}_{\mathbf{A}_j}\mathbf{p} \right\|_2^2 \\ &= \mathbf{p}^T\mathbf{X}\mathbf{p} - 2\mathbf{y}^T\mathbf{p} + z \qquad (5) \end{aligned}$$

where

$$\mathbf{X} := \sum_{i=1}^{r} \mathbf{M}^T_{\mathbf{A}_i\mathbf{A}'_i}\mathbf{M}_{\mathbf{A}_i\mathbf{A}'_i} + \sum_{j=1}^{s} \mathbf{N}^T_{\mathbf{B}'_j}\mathbf{N}_{\mathbf{B}'_j},$$

$$\mathbf{y} := \sum_{j=1}^{s} \mathbf{N}^T_{\mathbf{B}'_j}\mathbf{B}_j, \qquad z := \sum_{j=1}^{s} \mathbf{B}_j^T\mathbf{B}_j .$$

The matrix $\mathbf{N}_{\mathbf{B}'_k} \in \mathbf{R}^{n \times nn'}$ is defined by

$$[\mathbf{N}_{\mathbf{B}'}]_{i,i+n(j-1)} := b'_j, \quad \mathbf{B} := (b_j)$$

for $i = 1,\ldots,n$, and $j = 1,\ldots,n'$ and all remaining elements zero. Now, using the standard least-squares procedure we obtain the solution to the *unconstrained* minimisation problem as

$$\hat{\mathbf{p}} = \mathbf{X}^{-1}\mathbf{y} \qquad (6)$$

assuming $\mathbf{X}$ to be invertible and consequently

$$J_*(\hat{\mathbf{p}}) = z - \mathbf{y}^T\hat{\mathbf{p}} .$$

By expressing an arbitrary $\mathbf{p}$ in terms of the (unconstrained) optimal solution as $\mathbf{p} = \hat{\mathbf{p}} + \delta\mathbf{p}$ we now obtain

$$\begin{aligned} J_*(\mathbf{p}) &= J_*(\hat{\mathbf{p}}) + \delta\mathbf{p}^T\mathbf{X}\delta\mathbf{p} \\ &= J_*(\hat{\mathbf{p}}) + \delta\mathbf{p}^T\left(\sum_{i=1}^{nn'} \lambda_i\mathbf{v}_i\mathbf{v}_i^T\right)\delta\mathbf{p} \\ &= J_*(\hat{\mathbf{p}}) + \left(\sum_{i=1}^{nn'} \lambda_i\cos^2\theta_i(\delta\mathbf{p})\right)\left\|\delta\mathbf{p}\right\|_2^2 \qquad (7) \end{aligned}$$

where we have used the Spectral Decomposition Theorem for symmetric matrices, $\mathbf{v}_i$ is the normalised eigenvector corresponding to the eigenvalue λ_i of $\mathbf{X}$ and

$$\cos\theta_i(\delta\mathbf{p}) := \frac{\delta\mathbf{p}^T\mathbf{v}_i}{\left\|\delta\mathbf{p}\right\|_2}.$$

The above implies that we need to study the eigenvalues of $\mathbf{X}$. Since any matrix of the form $\mathbf{M}^T\mathbf{M}$ is symmetric with non-negative eigvenvalues we conclude that the eigenvalues of $\mathbf{X}$ are non-negative. We shall assume that the eigenvalues of $\mathbf{X}$ are strictly positive.

Now, consider the constrained minimisation problem

$$\min_{\substack{\mathbf{p} \\ \mathrm{devec}(\mathbf{p}) \in Per(n,n')}} J_*(\mathbf{p}) = J_*(\hat{\mathbf{p}}) + \min_{\delta\mathbf{p}} \delta\mathbf{p}^T\mathbf{X}\delta\mathbf{p} \qquad (8)$$

where $\mathrm{devec}(\cdot)$ is the inverse operation of $\mathrm{vec}(\cdot)$. The assumption about the positivity of the eigenvalues of $\mathbf{X}$ implies that the quadratic form in (8) is positive definite. Also, note that

$$\begin{aligned} \left\|\delta\mathbf{p}\right\|_2^2 &= \left\|\mathbf{p}\right\|_2^2 + \left\|\hat{\mathbf{p}}\right\|_2^2 - 2\mathbf{p}^T\hat{\mathbf{p}} \\ &= n + \left\|\hat{\mathbf{p}}\right\|_2^2 - 2\mathbf{p}^T\hat{\mathbf{p}}. \qquad (9) \end{aligned}$$

Now, for small values of the noise magnitude parameter ε, the unconstrained solution $\hat{\mathbf{p}}$ will be close to the $\mathrm{vec}(\mathbf{P}_o)$, with $\mathbf{P}_o$ the original permutation sub-matrix. This implies that the distance (as defined by the norm) between $\hat{\mathbf{p}}$ and any other vectorised permutation sub-matrix will be large, e.g. for $\varepsilon = 0$ either $\left\|\delta\mathbf{p}\right\|_2^2 = 0$ or $\left\|\delta\mathbf{p}\right\|_2^2 \geq 2$. Therefore, we expect for small ε that $\left\|\delta\mathbf{p}\right\|_2^2$ in (7) to be much more sensitive to perturbations in $\mathbf{p}$ [for $\mathrm{devec}(\mathbf{p}) \in \mathrm{Per}(n,n')$] than the $\{\cos\theta_i\}_{i=1}^{nn'}$ are. Thus, if for small ε, we consider these cosines each to be constant, then the cost $J_*(\mathbf{p})$ in (7) is minimised by minimising $\|\delta\mathbf{p}\|_2^2$. From (9) we see that

$$\min_{\substack{\mathbf{p} \\ \mathrm{devec}(\mathbf{p})\in \mathrm{Per}(n,n')}} \|\delta\mathbf{p}\|_2^2 = \left(n + \|\hat{\mathbf{p}}\|_2^2\right) - 2\max_{\substack{\mathbf{p} \\ \mathrm{devec}(\mathbf{p})\in \mathrm{Per}(n,n')}} \mathbf{p}^T\hat{\mathbf{p}}.$$

By Schwarz's inequality $\mathbf{p}^T\hat{\mathbf{p}}$ is maximised when $\mathbf{p}$ and $\hat{\mathbf{p}}$ are collinear. Since $\mathbf{p}$ may only contain 0 and 1 as components, clearly $\mathbf{p}^T\hat{\mathbf{p}}$ is maximised if the 1-components of $\mathbf{p}$ coincide with the largest (positive) components of $\hat{\mathbf{p}}$ subject to $\mathrm{devec}(\mathbf{p})$ being an $n \times n'$ permutation sub-matrix. Thus, we have to maximise the sum of n of the components of $\hat{\mathbf{p}}$ subject to the condition that the positions of these components coincide with the 1-components of a vectorised $n \times n'$ permutation sub-matrix. This constrained maximisation problem, namely

$$\tilde{\mathbf{p}} := \arg\max_{\substack{\mathbf{p} \\ \mathrm{devec}(\mathbf{p})\in \mathrm{Per}(n,n')}} \mathbf{p}^T\hat{\mathbf{p}} \qquad (10)$$

can be solved by means of the (maximising) *bipartite matching algorithm* [11] applied to the matrix $\hat{\mathbf{P}} := \mathrm{devec}(\hat{\mathbf{p}})$. The $n \times n'$ permutation sub-matrix $\tilde{\mathbf{P}} := \mathrm{devec}(\tilde{\mathbf{p}})$ produced by the bipartite matching algorithm satisfies

$$\|\tilde{\mathbf{p}} - \hat{\mathbf{p}}\|_2 \equiv \|\tilde{\mathbf{P}} - \hat{\mathbf{P}}\|_F \le \|\mathbf{P} - \hat{\mathbf{P}}\|_F$$

for all $\mathbf{P} \in \mathrm{Per}(n,n')$ and hence it is the closest permutation sub-matrix to $\hat{\mathbf{P}}$ in the Euclidean sense. This completes the derivation of the LSGM algorithm.

4.3 Computational Complexity

Because matrix inversion of an $m \times m$ real matrix has a computational complexity of $O(m^3)$ and because the inversion of an $nn' \times nn'$ matrix is required in (6) the complexity of this LSGM algorithm is $O(n^3 n'^3)$ compared to $O(n^2 n'^2)$ for the GAGM for matching of complete graphs.

4.4 Sub-Graph Matching

To see that the LSGM algorithm is deficient as a sub-graph matching algorithm consider the equality in (4) but expressed in terms of column vectors of the matrices $\mathbf{A}$ and $\mathbf{A}'$, namely

$$\|\mathbf{AP} - \mathbf{PA}'\|_F^2 = \sum_{i=1}^{n} \|\mathbf{a}_i - \mathbf{a}'_{\sigma(i)}\|_2^2 + \sum_{j \notin R(\sigma)} \|\mathbf{0} - \mathbf{a}'_j\|_2^2$$

The right most term can be interpreted as a constraint placed on the original norm (the first term). The cost function $J_*(\cdot)$ in (3) contains the sum of expression of this kind and when minimised these constraints promote certain column vectors to be matched to $\mathbf{0}$.

5. Numerical Results

In order to evaluate the performance of the LSGM algorithm the following procedure was used. Firstly, the parameters n', n, r, s and ε were fixed. For every trial a new reference graph was generated randomly with all attributes distributed uniformly between 0 and 1. An $n \times n'$ permutation sub-matrix $\mathbf{P}$ was generated randomly and then used to permute the rows and columns of the attribute adjacency matrices of G'. This operation also discards certain rows and columns when $n' > n$. Next, an independently generated noise matrix (vector, respectively), weighted by ε, was added to each edge attribute adjacency matrix (vertex attribute vector, respectively) to obtain the duplicate graph G. The elements of each noise matrix/vector were uniformly distributed in the interval $[-1/2, 1/2]$. The reason for using uniformly distributed noise instead of normally distributed noise as assumed in the derivation was because conventionally graph matching algorithms have always been evaluated using uniformly distributed noise. The different graph matching algorithms were then used to estimate an approximation to the original permutation sub-matrix, $\mathbf{P}$.

Figure 1 shows the performance of the LSGM algorithm compared to the performance of the EGM, GAGM and RIGM algorithms for $n' = 10$, $n = 10$, $r = 3$ and $s = 3$. The probability of a correct vertex-vertex assignment was calculated for each value of ε based on 500 trials. The superior performance of the LSGM algorithm is evident. Comparison of the cost associated with the solution produced by the LSGM algorithm to the cost associated with the actual permutation sub-matrix used to permute and prune G' showed that the LSGM's solution starts to deviate from the actual permutation sub-matrix for values of ε greater than 0.4. This is the point where the hypotheses upon which the derivation of the algorithm is based start to break down.

Similar result were obtained for other values of n', n, r and s.

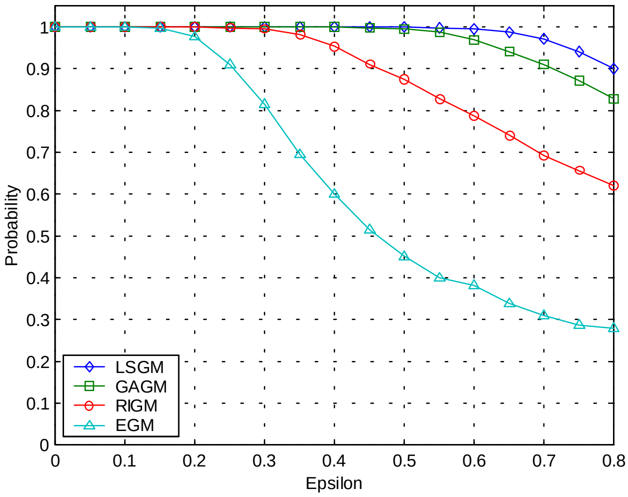

Figure 1. Probability of correct vertex-vertex matching vs. ε for (10,3,3) attributed graphs.

6. Conclusion

A novel derivation of the approximate least-squares algorithm for the matching of attributed graphs was presented. The approach taken was that of Bayesian inference. The computational complexity of this polynomial time algorithm is $O(n^3 n'^3)$. The numerical analysis of the LSGM algorithm showed that it generates the correct solution up to large deformations (corresponding to an ε-value of 0.4) of the reference graph. Even though this algorithm is deficient as a sub-graph matching algorithm it was found to perform sub-graph matching with at least some degree of success.

7. References

[1] W.H. Tsai and K.S. Fu, "Error-Correcting Isomorphism of Attributed Relational Graphs for Pattern Analysis," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-9, 1979, pp. 757-768.

[2] J.P. Maré and M.A. van Wyk, "Graph Matching in an Automatic Target Recognition System," *Pattern Recognition and Analysis Society of South Africa's PRASA'98 Workshop*, 1998, pp. 101-106.

[3] Wikipedia encyclopedia, http://wikipedia.com.

[4] L. Kitchen, "Relaxation applied to Matching Quantitative Relational Structures," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-10, 1980, pp. 96-110.

[5] W.J. Christmas, J. Kittler and M. Petrou, "Structural Matching in Computer Vision using Probabilistic Relaxation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 8, 1993, pp. 749-764.

[6] H.A.L. Almohamad, "Polynomial Transform for Matching Pairs of Weighted Graphs," *Applications of Mathematical Modelling*, Vol. 5, No. 4, 1991, pp. 216-222.

[7] H.A.L. Almohamad and S.O. Duffuaa, "A Linear Programming Approach for the Weighted Graph Matching Problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 5, 1993, pp. 522-525.

[8] S. Umeyama, "An Eigendecomposition Approach to Weighted Graph Matching Problems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 5, 1988, pp. 695-703.

[9] S. Gold and A. Rangarajan, "A Graduate Assignment Algorithm for Graph Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 5, 1993, pp. 377-388.

[10] M.A. Van Wyk, T.S. Durrani and B.J. van Wyk, "A RKHS Interpolator-Based Graph Matching Algorithm," To appear in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, July 2002.

[11] M.A. Van Wyk and T.S. Durrani, "A Framework for Multiscale and Hybrid RKHS-Based Approximators," *IEEE Transactions on Signal Processing*, Vol. 48, No. 12, 2000, pp. 3559-3568.

[12] B.J. van Wyk, *Kronecker Product , Successive Projection and Related Graph Matching Algorithms*, Ph.D. Dissertation, University of the Witwatersrand, Johannesburg, South Africa, 2002.

[13] D.S. Sivia, *Data Analysis: A Bayesian Tutorial*, Oxford University Press, 1996.

[14] M.A. van Wyk and J. Clark, "An Algorithm for Approximate Least-Squares Attributed Graph Matching," in *Problems in Applied Mathematics and Computational Intelligence*, N.E. Mastorakis (ed.), WSES Press, 2001, pp. 67-72.

[15] M.A. van Wyk, "Graph Matching: Analytical and Preliminary Simulation Results," Kentron Doc. No. 10428-05000-705-003, 1998.

[16] J. Clark and D.A. Holton, *A First Look at Graph Theory*, World Scientific, 1996.