Improving GMM registration with class encoding

Solmaz Panahi, Jeremy Chopin, Matej Ulicny & Rozenn Dahyot

Department of Computer Science, Maynooth University, Ireland

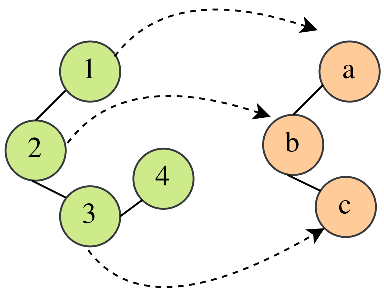
Abstract

Point set registration is critical in many applications such as computer vision, pattern recognition, or in fields like robotics and medical imaging. This paper focuses on reformulating point set registration using Gaussian Mixture Models while considering attributes associated with each point. Our approach introduces class score vectors as additional features to the spatial data information. By incorporating these attributes, we enhance the optimization process by penalizing incorrect matching terms. Experimental results show that our approach with class scores outperforms the original algorithm by [Jian and Vemuri, 2011] in both accuracy and speed.

Keywords: Point set registration, Graph matching, Gaussian mixture models (GMMs)

1 Introduction

Registration of point sets, that involves aligning multiple sets of points in a common coordinate system, is a fundamental task in computer vision and pattern recognition. Gaussian Mixture Models (GMMs) are a powerful representation for point distributions and the Euclidean distance between GMMs is a robust cost function to minimize for estimating the transformation between point sets to perform registration [Jian and Vemuri, 2011]. Point sets can be understood as graphs with points as nodes but without edges linking the nodes. Consequently graph matching algorithms have also been proposed to register graphs affected by spatial deformation [Zhou and De la Torre, 2013]. In this paper, we show experimentally that adding class scores as attributes benefit registration with GMMs [Jian and Vemuri, 2011], providing a more accurate registration while reducing computational time.



	a	b	c
1	1	0	0
2	0	1	0
3	0	0	1
4	0	0	0

Figure 1: Example between two graphs with nodes $\mathcal{V}_1 = \{1, 2, 3, 4\}$ and $\mathcal{V}_2 = \{a, b, c\}$ (left). Graphs Matching aims at estimating binary matrix X of correspondences (shown right) between $\mathcal{V}_1 = \{1, 2, 3, 4\}$ and $\mathcal{V}_2 = \{a, b, c\}$ given a global affinity matrix K.

2 State of the art

[Chopin et al., 2023] recently introduced a deep learning pipeline for image segmentation augmented with a graph matching technique for post-processing to correct errors of over-segmentation. Their proposed graph encoding uses class scores as features on the vertices (or nodes) of a graph that are associated with segmented regions. The graph matching is performed between a graph model of perfect segmentation and a new observed one assuming that both are spatially aligned, hence the approach is sensitive to rotation effects for instance. To address this problem, graph matching can be extended with registration methods such as ICP (Iterative Closest Point) for matching graphs subject to global rigid and non-rigid geometric constraints [Zhou and De la Torre, 2013]. As an alternative to ICP registration, GMM-reg have been proposed as a robust approach able to estimate spatial transformations even with outlier correspondences [Jian and Vemuri, 2011]. Here we investigate extending GMM registration using class scores as node attributes and relate it to graph matching.

Notations. We define two graphs noted $\mathcal{G}_1 = \{\mathcal{V}_1, \mathcal{E}_1\}$ and $\mathcal{G}_2 = \{\mathcal{V}_2, \mathcal{E}_2\}$ such that (ignoring subscripts) $\mathcal{V}$ is the set of vertices (a.k.a. nodes), and $\mathcal{E}$ the set of edges. Graph matching aims to estimate a binary matrix $\mathrm{X} = [X_{ij} \in \{0,1\}]$ of correspondences between nodes $\mathcal{V}_1$ and $\mathcal{V}_2$, given a global affinity matrix K encoding node and edge similarities between the two graphs (cf. Fig 1).

Graph Matching. For each pair of nodes $(i, j) \in \mathcal{V}_1$ and pair $(l, s) \in \mathcal{V}_2$ the given affinity $K_{is,jl}$, graph matching finds a mutual assignment between elements of the sets $\mathcal{V}_1$ and $\mathcal{V}_2$ to maximize the total score for all pairs of assignments [Zhou and De la Torre, 2013]:

$$\hat{\mathrm{X}} = \arg\max_{\mathrm{X}} \sum_{i,j \in \mathcal{G}_1} \sum_{s,l \in \mathcal{G}_2} K_{ij,sl}\ X_{is}\ X_{jl} \tag{1}$$

subject to some constraints [Huang et al., 2021] $\forall i \in \mathcal{G}_2 : \sum_{s \in \mathcal{G}_1} X_{is} \leq 1$ and $\forall s \in \mathcal{G}_1 : \sum_{i \in \mathcal{G}_2} X_{is} \leq 1$. When nodes capture spatial information, graph matching can be extended to take into account a deformation T allowing to register graph $\mathcal{G}_1$ on graph $\mathcal{G}_2$ [Zhou and De la Torre, 2013].

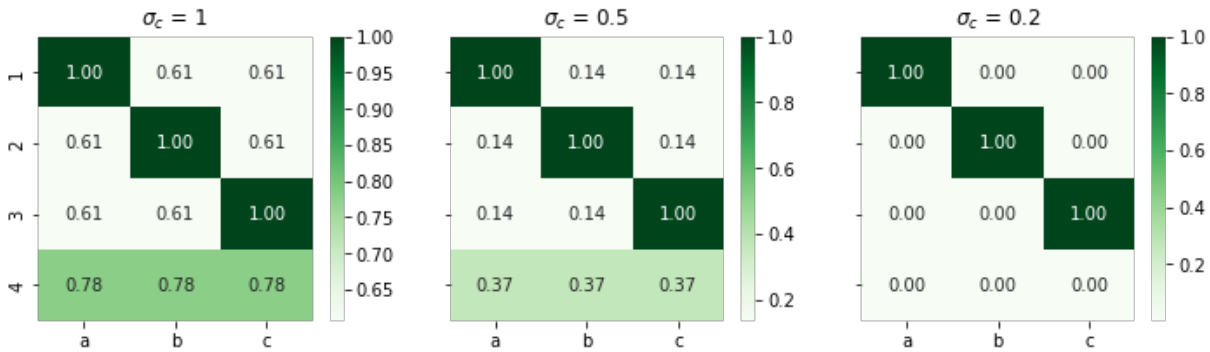


Figure 2: σ_c effect when computing weights $w_{i,s}$ between nodes of graphs shown Fig. 1. Selecting smaller values for σ_c imposes a strict constraint on the optimization process by penalizing the matching of nodes that belong to different classes. The matrix of weights converges towards the binary matrix X of correspondences shown in Fig. 1.

Registration with GMMs. [Jian and Vemuri, 2011] have proposed to estimate deformation T between point sets $\mathcal{V}_1$ and $\mathcal{V}_2$ by minimising the Euclidian distance between two Gaussian Mixtures Models (GMMs) fitted on each point set. Here rigid transformation (rotation, translation) is considered in $\mathbb{R}^2$, in which case the estimation of T is performed as:

$$\hat{T} = \arg\max_T \sum_{i=1}^{|\mathcal{V}_1|} \sum_{s=1}^{|\mathcal{V}_2|} \mathcal{N}(0; T(v_1^{(i)}) - v_2^{(s)}, \Sigma) \tag{2}$$

where T is a transform function of parameter $\theta = [t_1, t_2, \phi]$ representing the translation and rotation, $\mathcal{N}(x; \mu, \Sigma)$ indicates the normal distribution for random vector x with mean μ and covariance Σ. For simplicity, we have chosen isotropic covariance $\Sigma = \sigma^2 \mathrm{I}_2$ in this work (I_2 identity matrix in $\mathbb{R}^2$). $v_1^{(i)}$ is the spatial coordinate in $\mathbb{R}^2$ used as attribute for the node i in $\mathcal{V}_1$ (resp. $v_2^{(s)}$ is the spatial coordinate in $\mathbb{R}^2$ used as attribute for the node s in $\mathcal{V}_2$).

3 Registration with class attributes

Following [Chopin et al., 2023], we propose to extend Equation (2) by concatenating a class vector (noted c) to the spatial coordinate (v) as part of the attribute describing the nodes such that the estimation becomes:

$$\hat{T} = \arg\max_T \sum_{i=1}^{|\mathcal{V}_1|} \sum_{s=1}^{|\mathcal{V}_2|} \exp\left(\frac{-\left\| T\left(v_1^{(i)}\right) - v_2^{(s)} \right\|^2}{4\sigma^2} \right) \times \exp\left(\frac{-\| c_1^{(i)} - c_2^{(s)} \|^2}{4\sigma_c^2} \right) \tag{3}$$

In formula (3), the term with class scores can be interpreted as weights w_{is} that does not depend on the transformation T to be estimated:

$$w_{is} = \exp\left(\frac{-\left\| c_1^{(i)} - c_2^{(s)} \right\|^2}{4\sigma_c^2} \right) \tag{4}$$

We note that if $c_1^{(i)} = c_2^{(s)}$ the weight is equal to one $w_{is} = 1$. When $c_1^{(i)} \neq c_2^{(s)}$ (nodes with different class scores), then the weight is less than one $w_{is} < 1$. This implies that incorrect matches are penalized. We show experimentally (Section 4) that this leads to better estimates $\hat{T}$ with faster convergence of the algorithm.

Effect of σ_c. When we choose σ_c small enough, the weights w_{is} for $i \neq s$ become close to zero. Figure 2 shows the effect of different values of σ_c.

4 Experimental results

We use the fish data from [Jian and Vemuri, 2011] in our experiments[1]. Fish data has 98 points for which 2D coordinates are provided and these are augmented with a unique class score vector $c = [0, 0, ..., 1, ..., 0]^T \in \mathbb{R}^{98}$ and $\sum_{i=1}^{98} c_i = 1$. Class scores used for $\mathcal{G}_1$ and $\mathcal{G}_2$ are the same while 2D coordinates are rotated from $\pm 24°$ to $\pm 96°$ with $\pm 24°$ intervals in $\mathcal{G}_2$. The experiments are done with $\sigma = \{2; 1; 0.5\}$ and $\sigma_c = 0.2$. The error is computed as the Euclidian distance $\|\hat{T} - T_{GT}\|$ between the estimated transformation $\hat{T}$ and ground truth T_{GT}. Figure 3 shows the experiments categorized by different values of σ. The results yields notable improvements in terms of accuracy and computational efficiency. Estimations on each experiments has been improved as well as decreasing the convergence time across different experiments.

5 Conclusion

Class scores have proven useful for graph matching for improving image segmentation results from deep learning [Chopin et al., 2023] and this work shows that class information also can provide better performance of registration with GMMs [Jian and Vemuri, 2011]. Future work will investigate if edge information can also be used efficiently for registration.

[1] The code will be available at `https://github.com/solmak97/GMMReg_Extension`

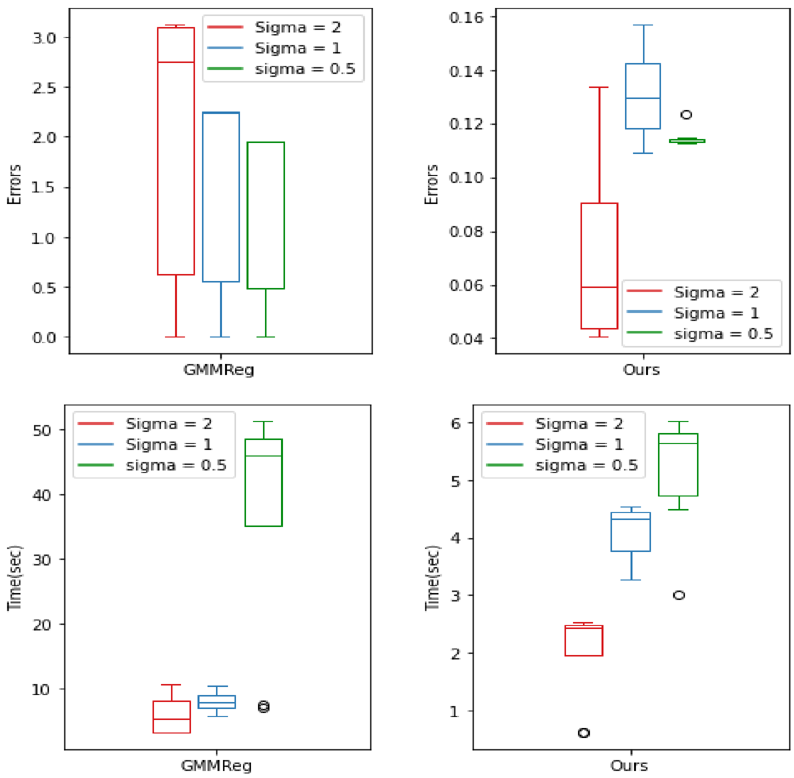


Figure 3: Experiment results on fish data with $\sigma_c = 0.2$. Distribution of errors $\|\hat{T} - T_{GT}\|$ are shown as box plot for $\sigma = \{2, 1, 0.5\}$ for both [Jian and Vemuri, 2011] algorithm (top left) and our approach (top right): note the difference of the order of amplitude in the error report on y-axis. Time for computations are likewise reported (bottom plots) showing our algorithm efficiency.

Acknowledgments

This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant number 18/CRT/6049/, the SFI Research Centres ADAPT (13/RC/2106 P2) and IForm (16/RC/3872), and is co-funded by the European Regional Development Fund. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

[Chopin et al., 2023] Chopin, J., Fasquel, J.-B., Mouchère, H., Dahyot, R., and Bloch, I. (2023). Model-based inexact graph matching on top of cnns for semantic scene understanding. *Computer Vision and Image Understanding (CVIU) journal*.

[Huang et al., 2021] Huang, X., Mei, G., Zhang, J., and Abbas, R. (2021). A comprehensive survey on point cloud registration. *CoRR*, abs/2103.02690.

[Jian and Vemuri, 2011] Jian, B. and Vemuri, B. C. (2011). Robust point set registration using Gaussian mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(8):1633–1645.

[Zhou and De la Torre, 2013] Zhou, F. and De la Torre, F. (2013). Deformable graph matching. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2922–2929.