LETTER

Multiple Gaussian Mixture Models for Image Registration

Peng YE[†a)], Fang LIU[†], *Nonmembers*, *and* Zhiyong ZHAO[†], *Member*

SUMMARY Gaussian mixture model (GMM) has recently been applied for image registration given its robustness and efficiency. However, in previous GMM methods, all the feature points are treated identically. By incorporating local class features, this letter proposes a multiple Gaussian mixture models (M-GMM) method for image registration. The proposed method can achieve higher accuracy results with less registration time. Experiments on real image pairs further proved the superiority of the proposed method.

key words: *multiple Gaussian mixture models, Gaussian mixture model, image registration*

1. Introduction

The accurate registration of images taken under different imaging conditions like multi-source, multi-sensor, or multi-modality is still an open case [1]. As a fundamental step, the result of image registration affects subsequent steps greatly for tasks like image navigation and image fusion. Registration methods based on Gaussian mixture model (GMM) have recently been proposed and used in medical and multi-sensor image registration for their robustness and efficiency [2]–[5]. However, in previous GMM methods, all the feature points are treated identically. This lack of discrimination between feature points not only hinders the registration accuracy but also unnecessarily raises the computation load through the use of a general parameter searching space. Through incorporating local image information, this letter proposes a multiple Gaussian mixture models (M-GMM) method for image registration. The proposed method achieves higher accuracy with less registration time, which is confirmed by experiments on real multi-source image pairs. The main contributions of this letter are twofold: 1. An image registration method based on multiple GMMs instead of one GMM is proposed. By using multiple GMMs to represent structural features in a scene, the proposed method can be more flexible and more accurate; 2. Local texture features around feature points, which are neglected in previous GMM methods, are incorporated with structural features in the proposed method to obtain higher accuracy results and less registration time.

Manuscript received December 27, 2013.
Manuscript revised March 12, 2014.
[†]The authors are with College of Electronic Science and Engineering, National University of Defense Technology, 410073 China.
a) E-mail: ye.peng.email@gmail.com
DOI: 10.1587/transinf.E97.D.1927

2. Brief Review of GMM Registration Method

GMM registration methods are various [2]–[5]. Without loss of generosity, coherent point drift (CPD) [2] is chosen as an example in this letter. Denoting $X = \{x_n;\ n = 1, \ldots, N\}$ the data points, $Y = \{y_m;\ m = 1, \ldots, M\}$ the reference points and GMM centroids, the GMM probability density function is defined as:

$$p(x_n) = \omega p(x_n \mid y_{M+1}) + (1-\omega)\sum_{m=1}^{M} P(y_m)p(x_n \mid y_m) \tag{1}$$

where $p(x_n \mid y_m) = \frac{1}{(2\pi\sigma^2)^{D/2}}\exp\left(-\frac{\|x_n - y_m\|^2}{2\sigma^2}\right)$. $p(x_n \mid y_{M+1})$ is an additional uniform distribution with weight ω added to the mixture models to account for noise and outliers. D is the dimension of the point sets. Equal isotropic covariance σ^2 and equal membership probabilities $p(y_m) = 1/M$ for all GMM components ($m = 1, \ldots, M$) are used. The two point sets are correctly aligned when the objective function $\sum_{n=1}^{N} p(x_n)$ is maximized.

3. Multiple GMM for Image Registration

As in Eq. (1), for each x_n, the posterior possibility $p(x_n \mid y_m)$ is computed for all y_m. In fact in lots of cases $p(x_n \mid y_m)$ should be zero. For example, in Fig. 1, point A and point B’ have very different texture background — plants and man-made building. They belong to different classes. Therefore there is no chance point B’ would be corresponding to point A (The image pair is partially shown for displaying convenience). The posterior possibility $p(x_{B'} \mid y_A)$ is

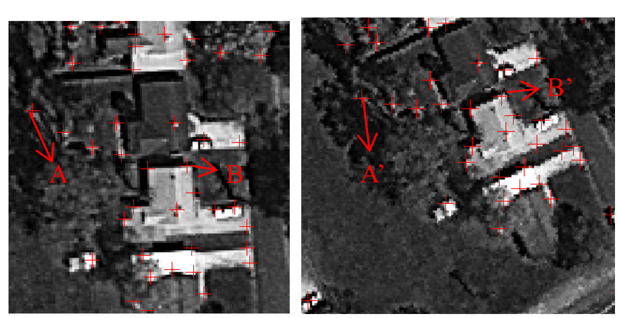


Fig. 1 Multi-source image pair (partial) to be registered.

actually 0. Generally speaking, if x_n and y_m belong to different classes, the posterior possibility between them should be 0. However previous methods still compute these matching possibilities. The computation of these possibilities not only waste computation energy but also produces an inaccurate parameter searching space, decreasing the possibility of achieving global optimal parameters. In this letter M-GMM registration method is proposed. For input feature sets, each feature point is firstly classified by its local image information. Texture feature around the feature points is used in this letter to determine the feature point's class. For feature points from different classes, the matching possibilities between them are 0. Therefore the GMM point registrations only proceed between data points and reference points of the same class. Each class would generate a GMM registration. For C classes, we get C GMMs. For data points and reference points X, Y, by classification of C classes, we obtain $X = \{X^c;\ c = 1, \dots, C\}$, $Y = \{Y^c;\ c = 1, \dots, C\}$. For the c-th class, data point set X^c has K points and reference point set Y^c has L points, denoting as $X^c = \{x_k^c;\ k = 1, \dots, K\}$, $Y^c = \{y_l^c;\ l = 1, \dots, L\}$. Notice that each class would contain different amount of feature points. Feature points of each class would generate a separate GMM point registration and could proceed in parallel. The c-th class feature sets would generate a GMM registration as Eq. (1), which is shown in Eq. (2):

$$p(x_k^c) = \omega p(x_k^c \mid y_{L+1}^c) + (1-\omega) \sum_{l=1}^{L} P(y_l^c) p(x_k^c \mid y_l^c) \quad (2)$$

where $p(x_k^c \mid y_l^c) = \frac{1}{(2\pi\sigma^2)^{D/2}} \exp\left(-\frac{\|x_k^c - y_l^c\|^2}{2\sigma^2}\right)$. Through these C GMMs, the parameter searching space is more accurate, leading to smaller computation load. Since the results of these C classes' registrations probably have small difference, a final adaptation registration should be done with all feature points. The new objective function is defined as Eq. (3):

$$\sum_{n=1}^{N} p(x_n) = \sum_{c=1}^{C} p(X^c) = \sum_{c=1}^{C} \sum_{k=1}^{K} p(x_k^c) \quad (3)$$

When the feature points cannot be classified into different classes, meaning there is only one class for all points, M-GMM degenerates into GMM. In this case, C =1, Eq. (3) also degenerates into the case of GMM. GMM is just a special case of M-GMM, where local features fail to provide valuable information. Since it is unlikely that local features cannot provide any valuable information in most cases, it is therefore rather necessary than complementary to incorporate local information to discriminate between feature points. However, these local information are mostly neglected in previous GMM methods.

4. Experimental Results

17 pairs of multi-source images are used to test the proposed method. The test image set consists of 10 pairs of multi-source images downloaded from different satellite map websites (map.google.com, map.baidu.com, map.sogou.com, etc.), 7 pairs of multi-sensor images of the same area (map.google.com and SPOT5). In our experiments, Harris corner is chosen as the feature detector and a 20×20 local area around each corner is used to extract texture feature. Gray-Level Co-Occurrence Matrix (GLCM) is used as the texture descriptor. The corners are classified by k-nearest neighbors (KNN) method into two classes: man-made building and others. The geometric transformation is set as affine transformation. Figure 2 shows the results of M-GMM method compared with those of GMM method. Figure 2 (a) shows the accuracy of M-GMM and GMM by root mean square error (RMSE) in pixels. Figure 2 (b) shows the registration time cost of M-GMM and GMM. It could be seen that the proposed method achieved higher accuracy with less registration time. Still some GMM results also achieved high accuracy and time cost similar to M-GMM. This is because in these cases, features of certain class are dominant while features of other classes are nominal. Therefore GMM method produces results similar to that of M-GMM. Figure 3 (7th in Fig. 2) shows one example pair of this case. It could be seen that man-made building features are dominate in both images. The number of the feature points of the man-made building class is close to the number of the overall feature points. Therefore M-GMM produces results similar to GMM's on this image pair. Other representative results are presented in Figs. 4 and 5. Figure 4 (1st in Fig. 2) shows the result of the whole image of Fig. 1. As shown in the red boxes, due to larger differences, the result of GMM is more blurred when the reference image and the transformed image are overlapped. Figure 5 (11th in Fig. 2) shows the results on multi-sensor image pair. The multi-sensor image registration results are more blurred than mono-sensor multi-source

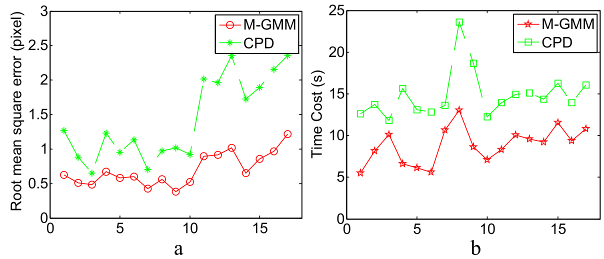


Fig. 2 Results of M-GMM and GMM.

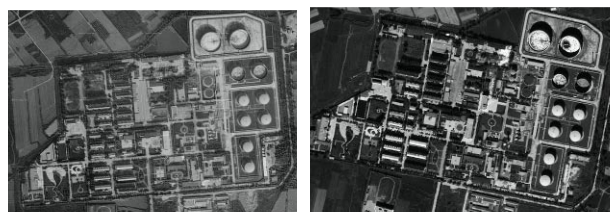
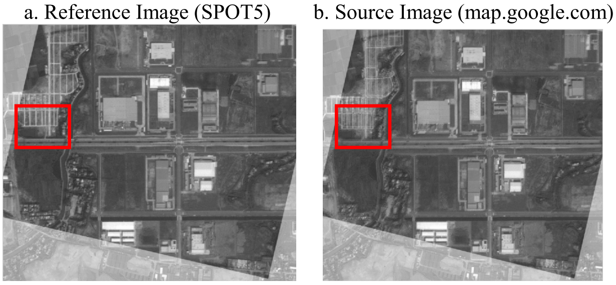


Fig. 3 Image pair with similar results from M-GMM and GMM.

a. Result of M-GMM b. Result of GMM

Fig. 4 Results of M-GMM and GMM on Fig. 1.

a. Reference Image (SPOT5) b. Source Image (map.google.com)



c. Result of M-GMM d. Result of GMM

Fig. 5 Multi-sensor results.

image results because of more outliers brought by different sensors. The quantitative evaluation of registration error by RMSE is shown as in Fig. 2 (a). For multi-sensor images, features from the man-made building class are more robust than features of other classes. Therefore they produce good initial transformation parameters for the overall registration. While in the case of GMM, the outliers brought by the different sensors greatly hinder the registration accuracy.

5. Discussion and Conclusions

This letter proposed an M-GMM method for image registration. By discrimination between feature points, the proposed method achieved higher accuracy with less registration time. In this letter only texture features are used. Further works could be done on finer classification of feature points. Also different classes could play different weights in the new objective function. Though the formulation of M-GMM in this letter is based on CPD, it could be seen that M-GMM is not deeply coupled with CPD. The idea of using multiple GMMs instead of one GMM could be easily extended to other mixture model based registration methods [3]–[5].

It should be noted that the classification time was not addressed in this letter. This is because, many applications like sequential image registration, can perform texture feature extraction and classification offline. In our observation, even combined with the time cost of the classification step, the time cost of M-GMM is comparable with that of GMM. The time cost of M-GMM fluctuates around that of GMM because of difference in feature point number. Nevertheless the time cost of the classification step should be carefully addressed in subsequent research.

References

[1] J. Le Moigne, N.S. Netanyahu, and R.D. Eastman, eds., Image Registration for Remote Sensing, Cambridge University Press, Cambridge & New York, 2011.

[2] A. Myronenko and X. Song, "Point set registration: Coherent point drift," IEEE Trans. Pattern Anal. Mach. Intell., vol.32, no.12, pp.2262–2275, Dec. 2010.

[3] P. Wang, Z.G. Qu, P. Wang, Y.H. Gao, and Z.K. Shen, "A coarse-to-fine matching algorithm for FLIR and optical satellite image registration," IEEE Geos. Remote Sens. Lett., vol.9, no.4, pp.599–603, July 2012.

[4] B. Jian and B.C. Vemuri, "Robust point set registration using Gaussian mixture models," IEEE Trans. Pattern Anal. Mach. Intell., vol.33, no.8, pp.1633–1645, Aug. 2011.

[5] R. Horaud, F. Forbes, M. Yguel, G. Dewaele, and J. Zhang, "Rigid and articulated point registration with expectation conditional maximization," IEEE Trans. Pattern Anal. Mach. Intell., vol.33, no.3, pp.587–602, March 2011.