

# GMSAC——一种鲁棒的基于高斯混合模型的基础矩阵估计算法

徐帆 方磊 田文 王宏远

(华中科技大学电子与信息工程系, 武汉 430074)

**摘要** 基础矩阵的鲁棒性估计是计算机视觉领域的一个基本问题。为了提高基础矩阵的估计精度,首先指出了现有的鲁棒性算法——RANSAC和MLESAC理论上的缺陷和实际应用中的问题;然后通过详细分析局外点复杂的成因,同时运用混合高斯分布代替均匀分布分别对不同成因的局外点进行了有针对性的建模,并提出了一种鲁棒性更强的算法——GMSAC。实验结果表明,相比于MLESAC算法,GMSAC算法提供了更高的模型似然度和计算精度。

**关键词** 基础矩阵 局外点 随机抽样一致性算法 随机抽样最大似然算法 混合高斯模型

中图分类号:TP301.6 文献标识码:A 文章编号:1006-8961(2008)09-1789-07

## GMSAC—A Gaussian Mixture Based Robust Estimator for Fundamental Matrix

XU Fan, FANG Lei, TIAN Wen, WANG Hong-yuan

(Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074)

**Abstract** A new method is presented for robustly estimating the fundamental matrix from image correspondence. Starting from RANSAC and MLESAC algorithm, we addressed some problems posed from both practical and theoretical standpoints and propose a new algorithm GMSAC. GMSAC adopts the same sampling strategy and maximization likelihood theory as the previous approaches. But instead of uniform distribution used by MLESAC, GMSAC choose Gaussian mixture to model the outliers. Due to the complex nature of outliers' occurrence, Gaussian mixture is more suitable for the distribution of outliers. We make a detailed analysis on the formation of outliers, and model different types of outliers respectively. Results are given for several image sequences, and it is demonstrated that this method gives results superior to the original MLESAC.

**Keywords** fundamental matrix, outlier, RANSAC, MLESAC, Gaussian mixture model(GMM)

### 1 引言

如何精确计算同一个场景的两幅2维图像之间的极线约束关系是计算机视觉领域的一个基本问题,因为在获得两图之间的约束关系之后,即可用来进行场景的3维重建<sup>[1]</sup>以及摄像机自标定<sup>[2]</sup>等多方面的应用。极线约束的数学表现形式是 $3 \times 3$ 的基础矩阵 $\mathbf{F}$ <sup>[3]</sup>,基础矩阵常用的求解方法有线性的7点算法<sup>[4]</sup>和正规化8点算法<sup>[5]</sup>,以及最小化重投

影误差<sup>[6]</sup>的非线性算法,但是由于误匹配等原因造成图像匹配点数据中存在局外点,使得计算结果往往偏差很大。

局外点是指一些游离于整体数据之外,而与参数模型严重不符的数据点,这些错误的观测数据是由于“仪器的故障”而引入,并遵从于一些潜在的参数模型的数据。计算机视觉应用中,局外点产生的原因十分复杂,两图之间重复的图像纹理、重复的场景几何结构、大幅度的变换和位移等,都会对图像间的特征点匹配造成致命的影响。因此,如何过滤掉

收稿日期:2006-11-20;改回日期:2007-04-09

第一作者简介:徐帆(1983~)男。现为华中科技大学电子与信息工程系博士研究生。主要研究方向为多视图的3维建模、计算机视觉中的鲁棒性算法。E-mail:drowtterr@gmail.com

通讯作者:王宏远,男。教授,博士生导师。研究方向为视频及信号处理、图像压缩。E-mail:wthywl@public.wh.hb.cn

局外点,而保留正确的局内点数据,是计算机视觉以及其他领域一直在研究的一个问题。

Fischler 和 Bolles 于 1981 年提出的随机抽样一致性算法 (random sample consensus, RANSAC)<sup>[7]</sup> 是计算机视觉领域应用最广泛的鲁棒性方法,它是采用随机抽样的方法对局外点进行甄别。Torr 等人提出的 MSAC (M-estimator sample and consensus) 算法<sup>[8]</sup> 和 MLESAC (maximum likelihood estimation by sample and consensus) 算法<sup>[9]</sup> 引入了针对局内点和局外点的混合概率分布模型,并在 RANSAC 随机抽样的架构下实现了模型参数的最大似然估计。

MLESAC 算法的一个主要缺点是由于对局外点成因不明而使用了均匀分布建模。Torr 等人随后提出的 MAPSAC (maximum a posterior by sample and consensus) 算法<sup>[10]</sup> 和 IMPsAC (importance sampling using RANSAC) 算法<sup>[11]</sup>, 虽都对基础矩阵的先验概率进行了研究,并实现了贝叶斯最大似然估计,但仍然使用均匀分布来对局外点进行建模。陈付幸和王润生以及 Brandt 指出,局外点应遵从混合高斯分布<sup>[12,13]</sup>,但在文献[12]、[13]中主要关注于模型参数的求解问题,而模型参数集则存在一定局限性。

本文主要针对局外点的建模问题进行了研究和改进。本文首先指出局外点引入的主要途径是特征点的位置误差和误匹配;随后通过详细分析局外点的成因,说明由不同途径引入的局外点模型的期望存在着差异;最后针对不同途径引入的局外点,运用不同参数集的混合高斯模型(GMM)进行建模,以期获得更高的模型似然度。本文同时讨论了一些算法的实现问题,并给出了算法的详细步骤,最后提出了一种新的鲁棒性方法——GMSAC (Gaussian mixture based sample consensus)。

### 1.1 RANSAC

RANSAC 的算法思想是通过在所有数据中循环随机抽取最小数据集来求解模型参数,希望能够在抽样中规避局外点,以取得最优解,RANSAC 有如下等式:

$$l = \frac{\log(1 - P)}{\log(1 - (1 - \varepsilon)^M)} \quad (1)$$

该式表示数据中局外点的比例是  $\varepsilon$  的情况下,抽样  $l$  次可以保证在置信度  $P$  下,至少有一次抽样全是局内点,其中  $M$  是最小样本的体积。

应用到基础矩阵计算的 RANSAC,需要首先设定一个门限  $T$ ;然后从所有共  $n$  对匹配点中选取 8

对匹配点,应用 8 点线性算法来计算基础矩阵  $F$ ;接着代入  $F$  计算每一对匹配点的重投影误差  $e_i^2$ ,如果  $e_i^2 < T^2$ ,则判为局内点;如果  $e_i^2 > T^2$ ,则判为局外点;最后选取局内点个数最多的一次  $F$  作为输出。RANSAC 实际上就是最小化如下的代价函数:

$$C_{\text{RANSAC}} = \sum_{i=1}^n \rho(e_i^2), \rho(e_i^2) = \begin{cases} 0 & \text{如果 } e^2 < T^2 \\ 1 & \text{如果 } e^2 \geq T^2 \end{cases} \quad (2)$$

RANSAC 的一大缺陷是阈值  $T$  的选取,因为如果  $T$  选取得太高,则所有的匹配点都会被判为局内点,若  $T$  选取得太低,则所有匹配点都会被判为局外点。由于阈值  $T$  的设置没有理论依据,并且也没有明确的物理意义,从而导致在实际应用中针对不同的场景,需要设置不同的阈值,才能取得相对理想的效果。

### 1.2 MLESAC

MLESAC 是在 RANSAC 的基础上建立的显式概率模型,其误差分布遵从一个混合概率密度函数,即

$$p(e_i | \Theta, F) = \alpha \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right) \exp\left(-\frac{e_i^2}{2\sigma^2}\right) + (1 - \alpha) \frac{1}{W} \quad (3)$$

局内点的分布遵从高斯分布,而局外点的建模则采用均匀分布,其中  $W$  是匹配窗口的宽度。由于 MLESAC 仍然使用随机抽样的方法,因此其求解问题就转化为了一个最大似然估计问题,相当于最小化如下代价函数:

$$C = - \sum_{i=1}^n \log\left(\alpha \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right) \exp\left(-\frac{e_i^2}{2\sigma^2}\right) + (1 - \alpha) \frac{1}{W}\right) \quad (4)$$

MLESAC 在 RANSAC 随机抽样的基础上引入了统计学的思想,即通过似然度来对随机抽样产生的模型假说进行评估。MLESAC 在实际应用中最大的一个缺陷在于,对局外点建模的无力,而且使用均匀分布没有考虑当前应用场景的任何先验信息,这就直接导致了对局外点辨识度的下降,因此如何寻找一种更优化的更适合于基础矩阵计算的局外点概率模型是 MLESAC 当前亟待解决的问题。

## 2 局外点成因分析及建模

### 2.1 局外点成因

匹配点中的局外点是在特征点提取和匹配环节

引入的,其成因主要有以下两点:

### (1) 错误的特征点位置

由于图像的分辨率和特征点探测器的性能局限,图像上的像素坐标并不是特征点的真实位置,这类误差称为特征点的位置误差。尽管图像的分辨率限制是数字图像处理中不能规避的问题,但由于这类误差都在 1pixel 以内,可以认为是一个小方差的高斯噪声,其通过最小二乘法就可以很好地求解,因此对计算精度影响不大,可被认为是数据中的局内点;可是由于特征点探测器算法在图形理解力上的欠缺和不完备,因而造成的特征点位置误差可能会达到 3pixels 以上,这些大误差的匹配特征点对基础矩阵的计算精度影响较大,可见这类位置误差较大的特征点是局外点的主要来源。

### (2) 特征点误匹配

现今应用最成熟的匹配算法的原理是计算特征点邻域的灰度互相关函数,但这种基于图像灰度相似性的策略,在遇到场景中存在复杂的纹理和重复的几何结构时,误匹配就会不可避免地产生。虽然由误匹配引入的局外点数量较少,但误差极大,对基础矩阵的估计有着致命的影响。

由于引入的途径不同,两类局外点在分布特性上也有不同,其中第 1 类局外点的概率分布的期望值还是在特征点的理想坐标范围内,此期望与局内点分布相似,只是方差较大;而误匹配局外点则比较复杂,通过简单的分析即可初步确定它的分布特性,它的概率密度除了在理想坐标处有一波峰外,在那些容易造成误匹配的“误匹配疑似点”处也有凸起,这个特性会在后面的叙述中得到详细的分析。因此相对于统一建模,充分地使用这一先验信息,对两类局外点分别建模无疑是一种更好的选择。

## 2.2 联合图像中的匹配点

匹配点分布在 4 维联合图像空间(joint image)<sup>[14]</sup>中,相对于 MLESAC 算法对误差  $e$  的建模,选择在联合图像空间下,对匹配点进行建模显得更加直观和恰当。

假设共有  $n$  对匹配点,其坐标向量集分别为  $\{m_1, \dots, m_n\}$  和  $\{m'_1, \dots, m'_n\}$ , 有  $m_i = [u_i, v_i]^T$ ,  $m'_i = [u'_i, v'_i]^T$ , 组成 4 维联合图像中的观测数据  $X = \{x_1, \dots, x_n\}$ , 其中  $x_i = [u_i, v_i, u'_i, v'_i]^T$ 。

下面首先在联合图像空间中对局内点数据进行

建模,如 2.1 节所述,局内点应遵从一个小方差的高斯分布,即

$$p(x_i | \Sigma, \hat{x}_i) = \frac{1}{\sqrt{(2\pi)^4 |\Sigma|}} \times \exp\left(-\frac{(x_i - \hat{x}_i)^T \Sigma^{-1} (x_i - \hat{x}_i)}{2}\right) \quad (5)$$

其中,  $\Sigma$  为协方差矩阵,根据计算机视觉领域的经典假设,各维数据相互独立且等方差,有  $\Sigma = \sigma^2 I_4$ 。式中,  $\hat{x}_i = [\hat{u}_i, \hat{v}_i, \hat{u}'_i, \hat{v}'_i]^T$  为匹配点没被噪声影响的理想精确坐标,代入式(8)有

$$p(x_i | \theta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^4 \exp\left(-\frac{(x_i - \hat{x}_i)^T (x_i - \hat{x}_i)}{2\sigma^2}\right) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^4 \exp\left(-\frac{d^2(m_i, \hat{m}_i) + d^2(m'_i, \hat{m}'_i)}{2\sigma^2}\right) \quad (6)$$

其中,  $d^2(m_i, \hat{m}_i) + d^2(m'_i, \hat{m}'_i)$  就是重投影误差  $e_i^2$ 。若重投影误差的计算依赖于基础矩阵  $F$ , 则局内点的概率分布密度函数如下:

$$p_{in}(x_i | \Theta_{in}, F) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^4 \exp\left(-\frac{e_i^2}{2\sigma^2}\right) \quad (7)$$

局内点的参数集  $\Theta_{in}$  为高斯分布的标准差  $\sigma$ 。

## 2.3 位置误差建模

特征点探测器所造成的误差分布仍然是一个以理想坐标为中心分布,由于形成误差的内因十分复杂,不同的算法以及场景配置都会影响到特征点的探测精度,因此这里使用由  $t$  个参数互不依赖的单一高斯分布组成的混合高斯模型来建模:

$$p_1^{out}(x_i | \Theta_1^{out}, F) = \sum_{k=1}^t \lambda_k \left(\frac{1}{\sqrt{2\pi\sigma_k^2}}\right)^4 \exp\left(-\frac{e_i^2}{2\sigma_k^2}\right) \quad (8)$$

$$\sum_{k=1}^t \lambda_k = 1$$

式中,  $\Theta_1^{out} = \{\lambda_1, \dots, \lambda_t, \sigma_1, \dots, \sigma_t\}$  是位置误差局外点分布的参数集,  $e_i^2$  是重投影误差。

## 2.4 误匹配建模

如图 1 所示,场景中存在大量的重复结构和纹理,这里考察图 1(a) 的两个容易产生误匹配的特征点  $A_1$  和  $B_1$  (使用  $\bullet$  标识),它们在图 1(b) 中的对应正确匹配点位置是  $a_1$  和  $b_1$  (正确匹配点使用  $\blacktriangle$  标识)。由于图 1(b) 中特征点  $a_1, a_2, a_3$  的邻域灰度分布几乎一致,因此  $A_1$  很容易误匹配到  $a_2$  或  $a_3$  上(疑似误匹配点使用  $\blacktriangledown$  标识),而与其他点形成误匹配的概率几乎为 0。

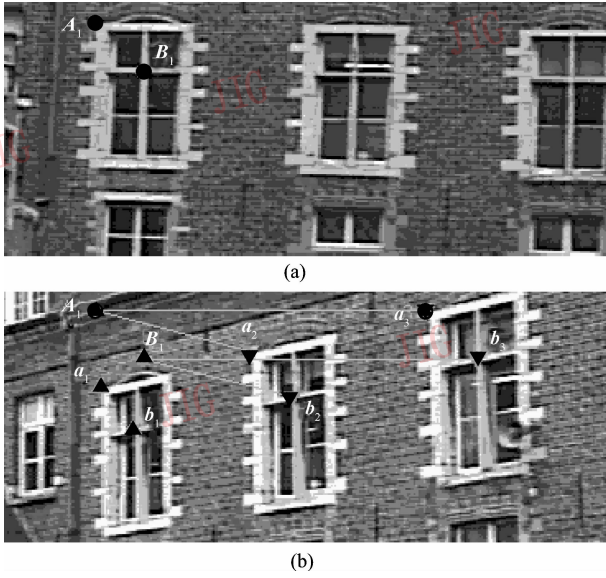


图 1 匹配场景局部图

Fig. 1 A segment of the image match

根据如上所述,可以使用两个分别以  $a_2$  与  $a_3$  为中心的高斯分布来对  $A_1$  的误匹配建模。假设  $A_1$  通过匹配算法后在图 1 下图得到的匹配点是  $A_2$ , 则可以预见  $A_2$  有很大的几率位于  $a_1, a_2$  或  $a_3$  的一个小邻域内。联合图像空间下的点  $x_1 = [A_1^T, A_2^T]^T$  的误匹配概率密度函数如下式所示:

$$p_2^{\text{out}}(x_1 | \Theta_2^{\text{out}}, F) = \sum_{k=1}^2 \lambda_k \left( \frac{1}{\sqrt{2\pi\sigma_k^2}} \right)^4 \times \exp\left( -\frac{d^2(x_1, \mu^k)}{2\sigma_k^2} \right), \sum_{k=1}^2 \lambda_k = 1 \quad (9)$$

式(9)为两个高斯分布组成的混合概率密度函数,这里把这两个分布称为误匹配疑似分布。两个误匹配疑似分布的期望  $\mu_1^1 = [A_1^T, a_2^T]^T$ ,  $\mu_1^2 = [A_1^T, a_3^T]^T$  是联合图像空间中的误匹配疑似点,它们的几何表现形式为图 1(b)中的线段  $A_1 a_2$  和  $A_1 a_3$ 。  $k = 1, 2$  表示  $x_1$  有两个误匹配疑似分布。但这里存在以下几个问题:

(1) 实际应用中如何得到误匹配疑似点的坐标,只有得到了这些点的坐标,才能有针对性地建模;

(2) 不同的特征点,其误匹配疑似点也不同,是否能寻找一种适用于所有特征点的建模方式,否则模型会由于参数过多、过于复杂而不能求解。

下面逐一分析上述问题。

在不进行深层次的图像分析的情况下,是不可

能对图像中的重复结构有所了解的,而且误匹配疑似点的个数和坐标也无从获知。解决这个问题的方法是在目前的最大似然框架下,通过引入误匹配疑似点坐标向量  $\mu$  作为模型参数来进行估计。

下面讨论  $\mu$  的参数化问题。2.1 节的分析指出误匹配的形成在很大程度上是由于图像中重复的结构所致,这种重复的结构不仅带来了误匹配,也带来了“重复”的误匹配。在两图变化不是非常大的情况下,误匹配疑似点往往能保持稳定的相对位置关系。如图 1(b)所示,与点  $A_1$  和  $B_1$  的误匹配疑似点对应的线段两两几乎平行,且长度几乎相等。利用不同特征点的误匹配疑似点之间客观存在的相关性,同时使用特征点坐标来参数化  $\mu$ ,就可使第 2 个问题得到很好的解决。

现在回到整体框架,对于一对匹配点的坐标向量  $m_i$  和  $m'_i$ ,有  $\mu^k = (m_i, m_i + \gamma_k)$ ,其中  $\gamma_k$  为误匹配疑似点相对  $m_i$  的位移量,这里称为误匹配向量,有:

$$d^2(x_i, \mu^k) = d^2(m_i, m_i) + d^2(m'_i, m_i + \gamma_k) = d^2(m'_i, m_i + \gamma_k) = d^2(m'_i - m_i, \gamma_k) \quad (10)$$

最终误匹配局外点的概率密度函数如下:

$$p_2^{\text{out}}(x_i | \Theta_2^{\text{out}}, F) = \sum_{k=1}^q \lambda_k \left( \frac{1}{\sqrt{2\pi\sigma_k^2}} \right)^4 \times \exp\left( -\frac{d^2(m'_i - m_i, \gamma_k)}{2\sigma_k^2} \right), \sum_{k=1}^q \lambda_k = 1 \quad (11)$$

其中,  $q$  为误匹配疑似分布的个数,  $\Theta_2^{\text{out}} = \{\lambda_1, \dots, \lambda_q, \sigma_1, \dots, \sigma_q, \gamma_1, \dots, \gamma_q\}$  为误匹配局外点分布参数集。

## 2.5 总体概率模型

最终匹配点的概率分布是由式(7)的局内点分布和式(8)、(11)的局外点分布加权组成的混合高斯分布,即

$$p(x_i | \Theta, F) = \lambda_{\text{in}} p_{\text{in}}(x_i | \Theta_{\text{in}}, F) + \lambda_1^{\text{out}} p_1^{\text{out}}(x_i | \Theta_1^{\text{out}}, F) + \lambda_2^{\text{out}} p_2^{\text{out}}(x_i | \Theta_2^{\text{out}}, F) \quad (12)$$

对总共  $m + t + 1$  个高斯分布进行统一编号,化简权值参数  $\lambda$ , 记为  $\alpha$ , 有

$$p(x_i | \Theta, F) = \sum_{k=0}^t \alpha_k \left( \frac{1}{\sqrt{2\pi\sigma_k^2}} \right)^4 \exp\left( -\frac{e_i^2}{2\sigma_k^2} \right) + \sum_{k=t+1}^{t+q} \alpha_k \left( \frac{1}{\sqrt{2\pi\sigma_k^2}} \right)^4 \exp\left( -\frac{d^2(m'_i - m_i, \gamma_k)}{2\sigma_k^2} \right) \quad (13)$$

$$\Theta = \{\alpha_0, \dots, \alpha_{t+1}, \sigma_0, \dots, \sigma_{t+q}, \gamma_{t+1}, \dots, \gamma_{t+q}\},$$

$\sum_{k=0}^{t+q} \alpha_k = 1$ , 其中编号0为局内点概率分布的参数。

### 3 GMSAC 算法

#### 3.1 算法步骤

基于式(13)的概率模型,本文提出了一种新的鲁棒性方法——GMSAC,其算法步骤如下:

- (1) 首先正规化所有的匹配点数据;
- (2) 先在共  $n$  组匹配点中选取 8 组匹配点,再使用 8 点算法计算基础矩阵  $\mathbf{F}$ ;
- (3) 使用  $\mathbf{F}$  计算梯度误差  $\mathbf{d} = \{e_1^2, \dots, e_n^2\}$ ;
- (4) 使用 EM 算法对式(13)进行最大似然估计来求解模型参数集  $\Theta$ ;
- (5) 计算似然度  $L$ , 取代价函数  $C_{\text{GMSAC}} = -L$ 。计算作为局内点的或是局外点每一对匹配点的概率密度  $p_{\text{in}}$  和  $p_{\text{out}}$ , 如果  $p_{\text{in}} > p_{\text{out}}$ , 则该匹配点记为局内点,反之,则记为局外点;

(6) 计算局外点所占比率  $\varepsilon = \sum_{i=1}^n p_{\text{out}}(x_i)$ , 并根据式(1)计算更新所需的循环数,跳至步骤(1)循环运行,最小化代价函数  $C_{\text{GMSAC}}$  并取对应的基础矩阵  $\mathbf{F}$  和局内点输出。

(7) 以输出的  $\mathbf{F}$  作为初始值,对所有的输出局内点做非线性优化,即最小化重投影误差,并进一步对  $\mathbf{F}$  进行最大似然估计。

#### 3.2 算法实现与执行效率

对混合概率密度函数的参数集  $\Theta$  进行最大似然估计最常用的是 EM(expectation maximization) 算法<sup>[15]</sup>。EM 算法首先需要预估一个初始参数集  $\Theta_0$ , 随后引入未知的随机变量  $\mathbf{y} = \{y_1, \dots, y_n\}$  表示第  $i$  个数据  $\mathbf{x}_i$  属于混合概率分布模型中的第  $y_i$  个分布。EM 算法分为两步,循环直至参数收敛:

##### (1) E-步

利用初始参数集  $\Theta_0$  计算概率  $p(y_i | \mathbf{x}_i, \Theta_0, \mathbf{F})$ ,  $i = 1, \dots, n; y_i = 0, \dots, m + q + 1$ 。

##### (2) M-步

按下式估算参数集  $\Theta$ , 再取  $\Theta_0 = \Theta$  返回 E-步循环计算。

$$\alpha_k = \frac{1}{n} \sum_{i=1}^n p(y_i = k | \mathbf{x}_i, \Theta_0, \mathbf{F})$$

$$\sigma_k^2 = \frac{1}{4} \frac{\sum_{i=1}^n e_i^2 p(y_i = k | \mathbf{x}_i, \Theta_0, \mathbf{F})}{\sum_{i=1}^n p(y_i = k | \mathbf{x}_i, \Theta_0, \mathbf{F})}$$

$$\gamma_k = \frac{\sum_{i=1}^n (\hat{m}_i - m_i) p(y_i = k | \mathbf{x}_i, \Theta_0, \mathbf{F})}{\sum_{i=1}^n p(y_i = k | \mathbf{x}_i, \Theta_0, \mathbf{F})} \quad (14)$$

但 EM 算法在实际应用中有如下两个缺点:

- (1) 区域极值问题:EM 算法输出结果的好坏很大程度上依赖于参数初值的选取;
- (2) 模型协调性问题:式(14)中,如果方差  $\sigma_k^2$  的估计值为 0(超过了计算软件的最小精度值),则会使似然度趋于无穷大<sup>[16]</sup>。

同时 GMSAC 相对于 MLESAC,由于模型复杂度有所上升,因此算法效率也是不容忽视的问题。

区域极值问题是统计学一直在研究的一个问题,本文无意在此展开讨论,此方面的深入研究可参见文献[17]。本文在实验中采用了标准的 EM 算法,大量的实验结果表明,在同等 EM 算法的支持下,GMSAC 算法能够取得大大优于 MLESAC 算法的似然度,并可取得良好的计算精度。

方差估计为 0 的问题其出现的原因是由于 E-步计算得出的  $p(y_i = k | \mathbf{x}_i, \Theta_0, \mathbf{F}) = 0$  的缘故。针对这个问题,GMSAC 采取了一种抛弃型策略,即一旦出现  $\sigma_k^2 = 0$  的情况,则抛弃第  $k$  个模型,继续 EM 算法计算剩余参数输出。这个策略在一定程度上降低了算法复杂度。

另外需要指出的是,由于 GMSAC 算法和 MLESAC 算法以及传统 RANSAC 算法的效率瓶颈都在于重投影误差的计算,因此 GMSAC 算法相对于 MLESAC 算法的计算时间并没有很大的上升,计算中通常使用重投影误差的一阶近似<sup>[18]</sup> 梯度误差(又称 Sampson 误差)来提高算法的运行效率,另外,借助 RANSAC 算法预检验<sup>[19, 20]</sup> 也能大大加速算法的执行。

#### 3.3 建模参数选取策略

若建模使用的高斯分布数量越多,则最终输出的模型似然度越大,但也会带来更大的计算复杂度。这是一个普遍存在的权衡问题,具体到 GMSAC 算法,可得到以下认识:

(1) 位置误差局外点建模的高斯分布数  $t$ :通过大量实验可见,  $t$  取 3 ~ 5 是一个比较均衡的取值,其能够取得良好的算法效率和模型似然度输出。本文实验取  $t = 4$ 。

(2) 误匹配局外点建模的高斯分布数  $q$ : GMSAC 算法包含数据正规化步骤,若选取  $q = 4$  并

设置误匹配疑似向量  $\gamma$  的初值为正规化数据的密集区  $[\pm 1, \pm 1]^T$ , 则可以取得理想的算法输出。本文实验采用上述设置。

## 4 实验结果

### 4.1 实验配置

实验数据采用窄基线的标准图像序列 Castle 和作者自行拍摄的 Desktop 宽基线图像序列。Castle 序列是采用子像素精度的 Harris 角点探测器<sup>[21]</sup>来提取特征点, 而 Desktop 序列是采用 SIFT 特征点探测器<sup>[22]</sup>来提取特征点, 为了不失一般性, 特征点匹配采用全图范围内匹配, 最后将匹配上的特征点作为实验输入。除 GMSAC 算法外, 还对下列两种算法做了对比试验:

#### (1) MLESAC 算法

按照文献[9]所描述的算法, 使用如式(3)的概率模型, 且只使用 EM 算法估计  $\alpha$ 。

#### (2) MLESAC-Mod 算法

根据本文第 3 节和文献[10], 对 MLESAC 算法进行了一些改良, 即使用联合图像空间中的概率分布来建模, 即

$$p(\mathbf{x}_i | \Theta, \mathbf{F}) = \sum_{i=1}^n \alpha \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^4 \times \exp\left(-\frac{e_i^2}{2\sigma^2}\right) + (1-\alpha) \frac{1}{V} \quad (15)$$

局外点建模仍然使用均匀分布, 这里的  $V$  等于联合图像的体积, 在实际应用中即是两图的面积之积。计算中使用 EM 算法同时对参数  $\alpha$  和  $\sigma$  进行收敛。

所有实验均计算循环 100 次的实验结果平均值, 实验主要考察下列指标:

- (1) 算法输出的最大化模型似然度  $L$ ;
- (2) 实际局外点过滤效果;
- (3) 重投影误差单纯考察重投影误差均值  $E$  存在一定的偏向性, 而鲁棒性的原则则是在过滤局外点的同时, 尽可能地保留局内点, 本文采用的测试指标为  $Z = E/n_{in}$ , 其中  $n_{in}$  是输出模型局内点的数量。

### 4.2 实验 1 Castle 图像序列

Castle 图像序列共有 27 幅图像, 图像中包含有大量复杂的纹理和几何结构。本文选择基线适中的 Castle011 和 Castle014 进行实验, 两图匹配后共有 477 对匹配点。实验结果如表 1 所示, 由表 1 可见, GMSAC 算法在各方面均取得了优于 MLESAC 算法

的测试指标。局外点的过滤效果如图 2 所示, GMSAC 算法在极易出现误匹配的图像右下角取得了预期的良好的过滤效果, 可见误匹配疑似分布确实发挥了应有的作用。本实验中, 由于误匹配疑似分布的计算提高了计算消耗, 致使 GMSAC 算法的计算时间比 MLESAC 算法略高, 但作者认为在非实时应用的场景中, 通过适当牺牲执行效率来换取更精确的计算结果是值得的。

表 1 Castle 序列实验结果

Tab. 1 Test results for Castle sequence

算法	$L$	$E$	$n_{in}$	$Z$
MLESAC	4 926	0.694	333	0.002 5
MLESAC-Mod	5 395	0.226	315	0.001 5
GMSAC	6 831	0.109	301	0.001 1

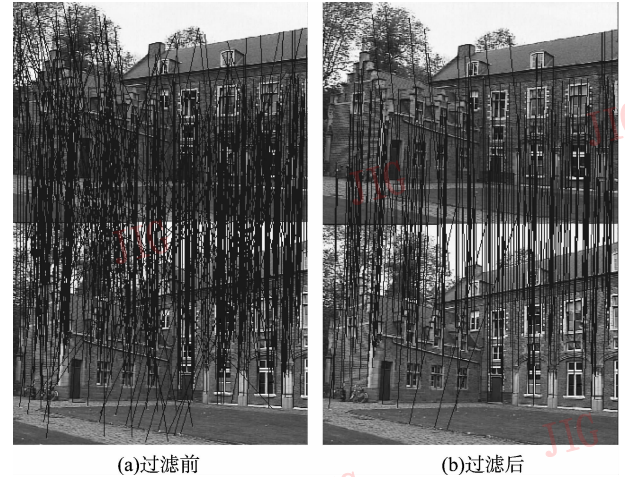


图 2 Castle 序列局外点过滤效果图

Fig. 2 Outlier removal for Castle sequence

### 4.3 实验 2 Desktop 图像序列

两图匹配后共有 152 对匹配点。实验结果如表 2 和图 3 所示, 由表 2、图 3 可见, GMSAC 配合基于相似变换不变量的 SIFT 特征点探测器在宽基线的环境下仍然能取得理想的效果, 并且相对于 MLESAC, 在宽基线的情况下, GMSAC 的优势体现得更加明显。

表 2 Desktop 序列实验结果

Tab. 2 Test results for Desktop sequence

算法	$L$	$E$	$n_{in}$	$Z$
MLESAC	992	1.718	112	0.011 7
MLESAC-Mod	1 575	0.566	90	0.008 3
GMSAC	2 113	0.197	78	0.005 6

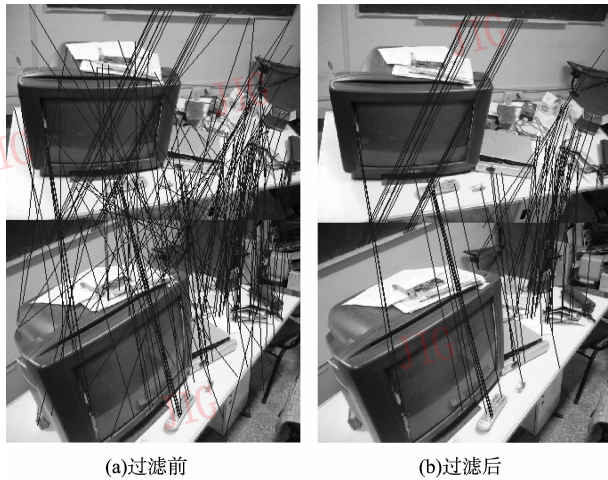


图3 Desktop序列局外点过滤效果图  
Fig. 3 Outlier removal for Desktop sequence

## 5 结论

本文通过详细分析局外点形成的原因,并针对不同原因使用不同参数集的混合高斯模型建模,提出了一套完整的鲁棒性算法——GMSAC。GMSAC的思想是通过更加直观准确的针对局外点的建模来获得更高的模型似然度,以提高基础矩阵的计算精度。实验证明,相对于MLESAC,GMSAC在算法消耗增加不大的前提下,能提供大大优于MLESAC的计算精度,且GMSAC特别适用于存在大量重复结构的场景计算。

## 参考文献 (References)

- 1 Pollefeys M, van Gool L, Vergauwen M, *et al.* Visual modeling with a hand-held camera [J]. *International Journal of Computer Vision*, 2004, **59**(3): 207 ~ 232.
- 2 Meng Xiao-qiao, Hu Zhan-yi. Recent progress in camera self-calibration [J]. *Acta Automatica Sinica*, 2003, **29**(1): 110 ~ 124. [孟晓桥, 胡占义. 摄像机自标定方法的研究与进展 [J]. *自动化学报*, 2003, **29**(1): 110 ~ 124.]
- 3 Faugeras O, Luong Q T, Papadopoulos T. *The Geometry of Multiple Images* [M]. Cambridge, UK: The MIT Press, 2001.
- 4 Zhang Z Y. Determining the epipolar geometry and its uncertainty: A review [J]. *International Journal of Computer Vision*, 1998, **27**(2): 161 ~ 195.
- 5 Hartley R I. In defense of the eight-point algorithm [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, **19**(6): 580 ~ 593.
- 6 Hartley R I, Sturm P F. Triangulation [J]. *Computer Vision and Image Understanding*, 1997, **68**(2): 146 ~ 157.
- 7 Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. *Communications ACM*, 1981, **24**(6): 381 ~ 395.
- 8 Torr P H S, Murray D W. The development and comparison of robust methods for estimating the fundamental matrix [J]. *International Journal of Computer Vision*, 1997, **24**(3): 271 ~ 300.
- 9 Torr P H S, Zisserman A. MLESAC: A new robust estimator with application to estimating image geometry [J]. *Computer Vision and Image Understanding*, 2000, **78**(1): 138 ~ 156.
- 10 Torr P H S. Bayesian Model Estimation and Selection for Epipolar Geometry and Generic Manifold Fitting [J]. *International Journal of Computer Vision*, 2002, **50**(1): 35 ~ 61.
- 11 Torr P H S, Davidson C. IMPSAC: Synthesis of importance sampling and random sample consensus [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(3): 354 ~ 364.
- 12 Chen Fu-xing, Wang Run-sheng. Clustering algorithm for the fundamental matrix estimation [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2005, **17**(10): 2251 ~ 2256. [陈付幸, 王润生. 基础矩阵估计的聚类分析算法 [J]. *计算机辅助设计与图形学学报*, 2005, **17**(10): 2251 ~ 2256.]
- 13 Brandt S S. Maximum likelihood robust regression by mixture models [J]. *Journal of Mathematical Imaging and Vision*, 2006, **25**(1): 25 ~ 48.
- 14 Triggs B. Matching constraints and the joint image [A]. In: Grimson E ed: *Proceedings of the Fifth International Conference on Computer Vision* [C], Washington, DC, USA: IEEE Computer Society, 1995: 338 ~ 343.
- 15 Dempster A P, Laird N M, Rubin D B. Maximum likelihood from incomplete data via the EM algorithm [J]. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1977, **39**(1): 1 ~ 38.
- 16 Vapnik V N. *The Nature of Statistical Learning Theory* [M]. New York, NY, USA: Springer-Verlag, 1995.
- 17 Biernacki C, Celeux G, Govaert G. Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models [J]. *Computational Statistics & Data Analysis*, 2003, **41**(3 ~ 4): 561 ~ 575.
- 18 Zhang Z Y. On the optimization criteria used in two-view motion analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, **20**(7): 717 ~ 729.
- 19 Chen Fu-xing, Wang Run-sheng. Fast RANSAC with preview model parameters evaluation [J]. *Journal of Software*, 2005, **16**(08): 1431 ~ 1438. [陈付幸, 王润生. 基于预检验的快速随机抽样一致性算法 [J]. *软件学报*, 2005, **16**(08): 1431 ~ 1438.]
- 20 Fang L, Wang H Y, Tian W, *et al.* Accelerated maximum likelihood estimation sample consensus method [J]. *Advances in Systems Science and Applications*, 2007, **6**(4): 568 ~ 575.
- 21 Harris C, Stephens M. A combined corner and edge detector [A]. In: *Proceedings of 4th ALVEY Vision Conference* [C], Manchester, UK: Blackwell Publishing, 1988: 147 ~ 151.
- 22 Lowe D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, **60**(2): 91 ~ 110.