

视觉基础矩阵估计方法的性能比较与分析

蔡涛¹ 段善旭¹ 李德华²

(华中科技大学电气与电子工程学院 武汉 430074)¹

(华中科技大学图像识别与人工智能研究所 图像信息处理与智能控制教育部重点实验室 武汉 430074)²

摘要 基础矩阵描述了单个场景的2幅视图之间的对应关系,在计算机视觉领域中扮演着极其重要的角色。首先讨论了三维视觉系统中的极线几何理论,随后论述了几类基础矩阵的估计方法,并对这些方法作了详细的比较和评价,最后实现了各种算法且使用仿真数据以及真实图像数据对各自的性能作了分析和总结。比较结果说明:1)如果图像特征点定位精确并且匹配无误,则线性方法的结果相当好;2)迭代算法可以解决定位的高斯噪声,但是当数据存在错误匹配时,性能很差;3)鲁棒算法能够同时解决定位误差和误匹配两类问题。此外,模拟实验和真实图像实验的结果表明,基于特征分析的正交回归最小二乘法的计算结果优于经典的线性最小二乘法。

关键词 计算机视觉,极线几何,基础矩阵,鲁棒估计,图像匹配

中图分类号 TP242.6+2 **文献标识码** A

Performance Comparison and Analysis of Fundamental Matrix Estimating Methods for Computer Vision Applications

CAI Tao¹ DUAN Shan-xu¹ LI De-hua²

(College of Electrical & Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China)¹

(State Education Commission Key Laboratory for Image Processing and Intelligent Control, Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan 430074, China)²

Abstract The fundamental matrix (F matrix) relates corresponding points across two different viewpoints and defines the basic relationship between any two images of the same scene. Therefore, the F matrix plays an important role in most computer vision applications. Some important computing methods for the F matrix were introduced and analyzed after describing the epipolar geometry in computer vision. At last, these methods were implemented and their performances were evaluated systematically based on simulated data and practical images. The test results proved that 1) the linear methods will work well on precisely located point-pairs without no mismatch; 2) the iterative nonlinear methods can conquer the Gaussian noise in positions of point-pairs, however, have poor performance for mismatched points; 3) the robust methods can resolve the problems brought by noise and mismatching. Furthermore, the results also showed that the eigen-analysis based orthogonal regression methods outperform the conventional least squares methods.

Keywords Computer vision, Epipolar geometry, Fundamental matrix, Robust estimation, Image matching

1 引言

从2D图像中获取景物的3D信息是计算机视觉研究的重要内容^[1]。一般而言,必须采用2台摄像机组成的立体视觉系统,或用单摄像机从不同角度获取两幅或多幅图像,才可能重建景物的3D结构。在没有任何有关成像系统的先验知识的前提下,某个景物的两幅图像通过所谓极线几何^[2]相互联系,它是立体视觉中最基本的几何描述;如果两幅图像的内参数(如焦距、主点坐标等)已知,且使用归一化图像坐标表示图像点位置,那么联系两幅图像的极线几何通过本质矩阵^[3](Essential Matrix, E矩阵)描述;否则,只能使用像素坐标表示点位置,此时联系两幅图像的就是基础矩阵(Fundamental

Matrix, F矩阵),其包含了用于建立两幅图像之间对应关系的所有必要几何信息。

对F矩阵的估计是3D重建^[3]、运动估计^[4]、摄像机定标^[5]、匹配和跟踪^[6]的基础。已有的研究成果表明^[7],可以通过对F矩阵进行分解建立起相对某个射影基底的双摄像机投影矩阵,并在射影几何意义下实现对匹配点的3D重建,该重建结果同真实的欧氏重建之间相差一个射影变换。虽然不能从射影重建中得到任何关于长度、角度等欧氏度量,但是射影重建中仍然存在着诸如共面性、共线性以及线段交比等丰富信息,这些信息对于机器人漫游和物体识别等应用来说已经足够。在许多实际应用中,比如从视频序列重建3D场景,由于镜头的焦距连续变化,通常意义下的摄像机定标已不可

到稿日期:2009-01-22 返修日期:2009-04-10 本文受国家自然科学基金(69775022)和863计划(863-306-ZT04-06-3)资助。

蔡涛(1974—),男,博士,讲师,主要研究方向为信息检测技术与计算机视觉等, E-mail: tsaitow_cn@sina.com; 段善旭(1970—),男,博士,教授,主要研究方向为智能信息处理与电机控制等; 李德华(1946—),男,学士,教授,主要研究方向为人工智能与计算机视觉等。

能,故不能提取任何欧氏几何信息。但是,如果摄像机满足针孔模型,我们仍有可能得到射影重建。更进一步,如果将场景中的某些先验知识引入射影重建,则可以得到有关场景更为具体的结构。例如,Hartley等^[7]在得到射影重建后使用8个参考点信息获得了欧氏几何结构以及摄像机参数。再如,Mohr等^[8]将诸如点位置、平行关系和垂直平面等约束直接嵌入到最优化过程中,从而实现欧氏重建。Zhang^[9]曾指出,即便图像已标定,但如果使用未标定图像的约束信息作为中间步骤,也可以获得更可靠的欧氏重建。Maybank^[10]和Faugeras^[11]证明,如果摄像机内参数不变,那么可以在不使用任何标定物条件下利用射影不变量进行通常意义的摄像机定标并实现欧氏重建,这就是通常所说的摄像机自定标技术。Lourakis和Deriche在文献^[12]和文献^[13]中详细讨论了内参数不变和可变两种情况下如何利用 F 矩阵的奇异值分解方法实现摄像机的自定标。

鉴于极线几何的重要性,对 F 矩阵的估计业已成为计算机视觉领域的一个研究热点。但是,仅根据测量数据(图像的像素坐标)来求解 F 矩阵,同计算机视觉中其它问题一样,其对数据噪声极其敏感,因而寻找视觉 F 矩阵的鲁棒性算法具有十分重要的理论意义和应用价值。本文首先简单阐述了极线几何的基本原理;随后对现有文献中主要的 F 矩阵计算方法进行了较全面的介绍和综合分析;最后利用仿真数据和真实图像对文中介绍的方法在不同条件下进行了实验并给出了实验分析结果。

2 基础矩阵计算问题描述

对于从线性摄像机(针孔模型)获得同一景物的两幅未校正图像而言,它们之间的基本关系就是极线几何约束。如图1所示,若两幅图像 I 和 I' 中的点对 $\{(m_i, m'_i)\}_{i=1}^n$ 是3D场景中 M_i 的投影点,则必然满足极线约束方程:

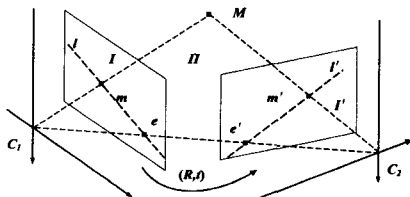


图1 极线几何约束示意图

$$\bar{m}_i^T F \bar{m}_i = 0, \text{ 且 } \text{rank}(F)=2 \text{ 和 } \det(F)=0 \quad (1)$$

其中, $\bar{m}_i = (u_i, v_i, 1)^T$ 和 $\bar{m}'_i = (u'_i, v'_i, 1)^T$ 分别是 m_i 和 m'_i 的齐次坐标表示。实际上,式(1)是关于 F 的9个未知参数的

线性齐次方程,若设 $F = \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix}$, 且 $f = [f_1, f_2, \dots, f_9]^T$, 则式(1)可以改写为 F 各个元素的线性方程 $u_i^T f = 0$, 其中 $u_i = (u_i, v_i, 1)^T$ 是9维向量。如果有 n 对匹配点,则可将 n 个线性方程堆叠起来,最后得到超定方程组:

$$U_n f = 0_{n \times 1} \quad (n \geq 9) \quad (2)$$

其中, U_n 是一个 $n \times 9$ 矩阵且 $U_n = [u_1, u_2, \dots, u_n]^T$ 。所有 F 矩阵的计算方法都是建立在式(2)的基础上,但需要注意 F 在相差一个常数因子意义下唯一,同时 $\text{rank}(F)=2$, 故虽然 F 的未知参数有9个,但其自由参数实际为7个。

3 基础矩阵估计方法分类

3.1 线性优化方法

通过求解线性方程计算 F 矩阵具有复杂度低、计算速度快的优点,虽然其受匹配点定位噪声及错误匹配点的影响很大,但仍是其它高性能算法的必要基础。

3.1.1 7点算法^[14]

如前所述, F 的自由参数为7个,故只需利用7对匹配点的图像数据就可以确定 F ,此时式(2)中 $n=7$ 且 $\text{rank}(U_7)=7$ 。通过对 U_7 做奇异值分解(SVD),可以得到其对应零奇异值的两个向量 f_1 和 f_2 (分别对应 F 的两个解 F_1, F_2),二者张成 U_7 的零空间(null space)。因式(2)是齐次方程,故 F 的解集应为 $F = \alpha F_1 + (1-\alpha) F_2$,其中 α 为标量因子。又由于 $\det(F)=0$,则可得到关于 α 的3次多项式方程 $f(\alpha) = \det(\alpha F_1 + (1-\alpha) F_2) = 0$,它最多有3个实解。如果某个解的多项式绝对值 $|f(\alpha)|$ 比其它解小得多,就可认为是两幅图像间的运动;否则,3个解都是可能的,必须全部考虑。

3.1.2 8点算法^[3,15]

实际应用中,通常选取的匹配点数会多于7个。如果忽略 F 的秩2约束,则 F 的求解等效于最优化问题 $\min_F \sum_i (\bar{m}_i^T F \bar{m}_i)^2$, 可等价于 $\min_F \|U_n f\|^2$ 。同样, f 在相差一个常数因子的条件下是唯一的。为避免该优化问题出现平凡解 $f=0$,需对 f 的元素施加某些约束,从而可使用最小二乘法求解。不失一般性,假设 f 的最后一个元素(即 $f_9 = F_{33}$)不为零,可令 $f_9 = -1$,且设 c_9 是 U_n 最后一列向量,于是有:

$$\|U_n f\|^2 = \|U_n' f' - c_9\|^2 = f'^T U_n'^T U_n' f' - 2c_9^T U_n' f' + c_9^T c_9 \quad (3)$$

其中, f' 是 f 前8个元素构成的向量, U_n' 是 U_n 前8列构成的 $n \times 8$ 矩阵。为求式(3)的最小值,令 $\frac{\partial \|U_n f\|^2}{\partial f} = 0$, 得 $f' = (U_n'^T U_n')^{-1} U_n'^T c_9$ 。由于事先不知道 F 的哪个元素不为零,因此有必要对 F 矩阵的每一个元素做上述计算,最后在9种可能情况下,使得 $\|U_n f\|^2$ 最小的 f 就是所需的结果。

此外,可以施加 $\|f\|=1$ 约束,此时 f 的每个元素具有同等地位。于是,上述最优化问题成为标准的约束优化问题:

$$\min_f \|U_n f\|^2, \text{ 且满足条件 } \|f\|=1 \quad (4)$$

引入Lagrange乘子 λ ,可将其转化为对指标函数 $\Phi(f) = \|U_n f\|^2 + \lambda(1 - \|f\|^2)$ 的无约束优化问题。同样,令 $\frac{\partial \Phi(f)}{\partial f} = 0$, 则 $U_n^T U_n f = \lambda f$, 故 λ 是对称矩阵 $M \equiv U_n^T U_n$ 的特征值,而 f 是其对应的特征向量。

上述2种方法在统计学中分别称为线性回归估计和正交回归估计,在某些文献中也分别称为线性最小二乘(Linear Least Square, LinLS)和线性特征分析(Linear Eigen Analysis, LinEig)。

3.1.3 秩2约束方法^[16]

即便采用了大量匹配点数据,线性算法仍对噪声极为敏感。其中的重要原因是忽略了 F 的秩2约束。最常用的解决方法是:求解在某种度量意义下距离 F 最为接近的秩2矩阵作为最终估计结果。若 F 的奇异值分解为 $F = USV^T$, 其中 $S = \text{diag}([\sigma_1, \sigma_2, \sigma_3])$ 且 $\sigma_1 > \sigma_2 > \sigma_3$, 而 U 和 V 为正交规范矩阵,分析表明^[17]在Frobenius范数意义下,距离 F 最为接近的

秩 2 矩阵为 $\hat{F} = U \hat{S} V^T$, 其中 $\hat{S} = \text{diag}([\sigma_1, \sigma_2, 0])$ 。通常将此方法作为 8 点方法的后处理过程, 使得最终的 F 满足秩 2 约束。

实际上, 可以在对式(4)的优化过程中施加秩 2 约束, 并且仍然得到解析解。不失一般性, 设 $f = [g^T, f_8, f_9]^T$, 其中 g 是 f 前 7 个元素组成的向量, 另设 c_8, c_9 是 U_n 最后两列向量, 而 B 是 U_n 前 7 列向量构成的 $n \times 7$ 矩阵。从 $U_n f = 0$, 有 $Bg = -f_8 c_8 - f_9 c_9$ 。假设 B 的秩为 7, 那么可以使用最小二乘法得到:

$$g = -f_8 (B^T B)^{-1} B^T c_8 - f_9 (B^T B)^{-1} B^T c_9 \quad (5)$$

上述求解依赖于 2 个自由参数 f_8 和 f_9 。使用约束 $\det(F) = 0$, 得到关于 f_8 和 f_9 的 3 自由度齐次方程, 以保证至少得到 f 的一个解。如果有 3 个实数解, 则选择在约束 $\|f\| = 1$ 下能够最小化 $\|U_n f\|$ 的解。具体实现时, 可以对 f 的 36 种配对选择作同样的计算, 并从 108 个可能的解中选择能够最小化 $\|U_n f^2\|$ 的一个。

3.2 非线性迭代方法

线性方法能降低计算时间, 但是当存在误匹配点或者有较大误差时, 求解精度相当差。线性方法主要是极小化 $\sum_i (\bar{m}_i^T F \bar{m}_i)^2$, 直观上该目标函数是合理的, 但没有代表多少物理意义。为了获得更好结果, 须考虑使用非线性迭代算法。 F 矩阵的迭代计算方法主要有 3 种: 最小化点到极线距离、基于梯度的方法和参数空间优化方法。

3.2.1 最小化点到极线距离^[17]

线性最小二乘法使用了优化准则 $\min_F \sum_i (\bar{m}_i^T F \bar{m}_i)^2$ 。众所周知, 该准则为代数距离, 不具有任何物理意义。一个很自然的想法是最小化点到其对应极线的距离, 即:

$$\min_F \sum_i (d^2(\bar{m}_i, F^T \bar{m}_i') + d^2(\bar{m}_i', F \bar{m}_i)) \quad (6)$$

其中, $d(\cdot, \cdot)$ 是点到直线的距离计算函数。令 $l_i = F^T \bar{m}_i' = [l_1, l_2, l_3]^T$ 和 $l_i' = F \bar{m}_i = [l_1', l_2', l_3']^T$, 则式(6)等价于:

$$\min_F \sum_i w_i^2 (m_i^T F m_i)^2 \quad (7)$$

其中,

$$w_i = [(l_1^2 + l_2^2)^{-1} + (l_1'^2 + l_2'^2)^{-1}]^{1/2} \quad (8)$$

求解式(6)通常采用加权线性最小二乘法(WLS), 但需事先确定权重 w_i 。为此, 可采用线性迭代算法, 即初始时令 $w_i = 1$, 根据式(7)求解 F 最小二乘解, 然后根据式(8)确定 w_i , 再根据式(7)求解 F 的加权最小二乘解。如此迭代, 直至收敛。

3.2.2 基于梯度的方法^[18]

设 $f_i = \bar{m}_i^T F \bar{m}_i$, 由于每个 f_i 的方差并不一样, 因此最小化 $\sum_i f_i^2$ 并不能得到 F 的最优估计。只有当每项具有相同方差时, 最小二乘估计才能得到最优结果。因此, 我们可以最小化加权平方和 $\min_F \sum_i f_i^2 / \sigma_i^2$, 其中 σ_i^2 是 f_i 的方差。在 1 阶近似条件下, 可认为 f_i / σ_i 为标准高斯分布, 其方差相同且为 1。由于特征点都是使用相同方法提取, 因此有理由认为图像点的定位误差为独立同分布高斯噪声, 即协方差矩阵为 $\Delta_{m_i} = \Delta_{m_i'} = \text{diag}(\sigma^2, \sigma^2)$, 于是 f_i 的方差由下式给出:

$$\begin{aligned} \sigma_{f_i}^2 &= (\partial f_i / \partial m_i)^T \Delta_{m_i} (\partial f_i / \partial m_i) + (\partial f_i / \partial m_i')^T \Delta_{m_i'} (\partial f_i / \partial m_i') \\ &= \sigma^2 [l_1^2 + l_2^2 + l_1'^2 + l_2'^2] \end{aligned}$$

最终, 优化问题成为 $\min_F \sum_i (\bar{m}_i^T F \bar{m}_i)^2 / g_i^2$, 其中 $g_i = (l_1^2 + l_2^2 + l_1'^2 + l_2'^2)^{1/2}$ 。后续的计算过程与点到极线距离最小方法类似。

3.2.3 基于参数空间优化的方法^[2]

将 F 的列记为 c_1, c_2, c_3 , 则有如下分析结果:

$$\text{rank}(F) = 2 \Leftrightarrow \begin{cases} (\exists j_0, j_1, j_2 \in [1, 3]) (\exists \lambda_1, \lambda_2 \in \mathbb{R}) \\ c_{j_0} + \lambda_1 c_{j_1} + \lambda_2 c_{j_2} = 0 \\ (\text{not } \exists \lambda \in \mathbb{R}) c_{j_1} + \lambda c_{j_2} = 0 \end{cases}$$

也就是说, 矩阵 F 的某一列是其它两列的线性组合, 同时该两列不是线性相关的。因此, 可以将 F 参数化为 $(c_{j_1}, c_{j_2}, \lambda_1, \lambda_2)$ 。若构造一个 3 维向量, 使其第 j_1 个和第 j_2 个元素分别是 λ_1 和 λ_2 , 而第 j_0 个元素为 1, 那么显然该向量是 F 的特征向量, 对于 F 矩阵而言它就是极点的齐次坐标。为了使该问题为对称形式, 亦即对于另一幅图像中的极点也应该有这样的计算过程, 需要在行方向上做类似列方向的分解, 这就将参数化的集合划分为 9 种映射, 分别对应着选择某一列和某一行作为其它两列和两行的线性组合。设左右两图像中的极点(有限距离)的齐次坐标分别为 $\bar{e}_1 = (x_1, y_1, 1)^T$ 和 $\bar{e}_2 = (x_2, y_2, 1)^T$, 则当 $i_0 = 3$ 和 $j_0 = 3$ 时, F 的参数化表示为:

$$F = \begin{bmatrix} a & b & -ax_1 - by_1 \\ c & d & -cx_1 - dy_1 \\ -ax_2 - cy_2 & -bx_2 - dy_2 & (ax_1 + by_1)x_2 + (cx_1 + dy_1)y_2 \end{bmatrix}$$

上述参数化满足约束 $F \bar{e}_1 = 0$ 和 $F^T \bar{e}_2 = 0$ 以及 $\det(F) = 0$ 和 $\text{rank}(F) = 2$ 。注意到 F 在相差一个比例因子下相同, 可以将所有参数 $(a, b, c, d, x_1, y_1, x_2, y_2)$ 中的某一个固定(本文令 $d = 1$)。至此, F 的计算过程为: 使用 8 点算法计算 F 的初始估计; 对 F 进行参数化, 得到参数的初始值; 使用数值方法(如 Levenberg-Marquardt 算法)优化式(6)。整个处理过程中, F 的参数有可能发生变化, 即 F 的参数初始选择并不一定适合最后的基础矩阵。

3.2.4 Gold Standard 算法

此算法由 Hartley^[14] 提出, 用它求解基础矩阵的最大似然估计值 $\hat{F}$ 。假设匹配点坐标只存在高斯噪声, 则通过最小化 $\sum_i (d(m_i, \hat{m}_i)^2 + d(m_i', \hat{m}_i')^2)$ 得到 $\hat{F}$, 其中 $\{m_i \leftrightarrow m_i'\}$ 是观测的匹配点对, 而 $\{\hat{m}_i \leftrightarrow \hat{m}_i'\}$ 是“真实”匹配点位置的估计坐标且满足 $\hat{m}_i^T \hat{F} \hat{m}_i' = 0$ 。Hartley 利用线性算法计算初始秩 2 矩阵 $\hat{F}$, 由此构造出 2 台虚拟射影摄像机, 利用三角测量方法确定其 3D 点坐标并求得其 2 个投影匹配的坐标, 最后利用 Levenberg-Marquardt 算法求解上述优化问题, 具体实现不再赘述。

3.3 鲁棒估计方法

由于错误数据的存在将严重影响 F 的计算精度, 致使计算结果是无用的, 因此要求算法具有很好的稳定性。最小二乘法假设数据噪声为零均值分布, 且假定误差分布满足一个给定的单参数模型。实际研究表明, 后者假设是不合理的。在有些情况下, 即使采样数据中仅存在一个错误匹配, 最小二乘法得到的估计结果也会受到影响。因此, 需采用鲁棒估计方法, 其中常见的有 M-Estimators, 最小平方中值(Least Median Squared, LMedS)和 RANSAC(RANdom SAMple Consensus)。M-Estimators 对定位精度不高的匹配点的稳定性较好, 在匹配点中所含错误数据比例较小(20%左右)时也能

获得较好的估计结果。LMedS 则对这两类匹配点都有较好的稳定性,甚至当错误数据的比例较大(接近 50%)时也能获得稳定的估计结果。而 RANSAC 在理论上可以在错误率大于 50%时仍获得稳定的估计结果。

3.3.1 M-Estimators

在 M-Estimators 方法中,定义残差 r_i 为第 i 对匹配点相对各自极线的距离偏差,即 $r_i^2 = d^2(\bar{m}_i, F^T \bar{m}_i') + d^2(\bar{m}_i', F \bar{m}_i)$ 。标准最小二乘法设法最小化 $\sum_i r_i^2$,但是当有误差匹配(outliers)时,极少的错误值就可能使 $\sum_i r_i^2$ 很大,故最小二乘法不稳定。相比之下,M-Estimators 将残差的平方 r_i^2 替换成另一个关于残差的函数 $\rho(r_i)$,该函数是对称正定函数,其最小值在零处。M-Estimators 可以用加权最小二乘法来实施,即 $\rho(r_i) = w_i r_i^2$ 。比较常用的加权函数有 Huber 和 Tukey 等。实验表明,M-Estimators 采用多次迭代并逐步消除误差较大的数据点的方法,在定位误差较大的情况下有较好的稳定性。但如果样本点中存在错误匹配点,则由 M-Estimators 得到的结果的稳定性并不令人满意,因此必须寻求更为稳定的估计算法。

3.3.2 LMedS

LMedS 方法通过寻求非线性优化问题 $\min_i \{\text{med}(r_i^2)\}$ 来估计参数,其中 r_i 仍然是指第 i 对匹配点相对各自极线的距离偏差。也就是说,对全部数据的所有可能的残差平方的中值求解时,这个方法给出的中值最小。研究表明,该方法对数据的非高斯噪声和错误数据都非常鲁棒。由于仅考虑残差的中值,理论上即使有近 50%的错误数据,LMedS 也能给出正确的结果。然而,与 M-Estimators 不同,LMedS 不能转化成加权最小二乘法,它必须通过在可能的参数空间里搜索来解决,参数空间可由数据组合产生。而这个空间太大,只能利用蒙特卡罗(Monte Carlo)方法分析随机抽取的子空间^[17]。

3.3.3 RANSAC

RANSAC 最初由 Fischler 和 Bolles^[20] 提出,用于解决参数鲁棒估计问题。Torr 率先将其用于 F 矩阵的估计^[21]。在很大程度上,该方法同 LMedS 类似。不同的是,它需要事先指定某个门限值,以判断某对匹配点是否与 F 矩阵相容,而最终估计的结果就是获得相容匹配个数最多的 F 矩阵。研究表明,RANSAC 能够在错误匹配超过 50%时获得较好的结果。但是,在某些情况下,RANSAC 容易陷入局部极小。

4 图像数据的预处理

4.1 数据归一化处理

研究表明,对获得的图像点坐标做归一化处理,可以使 F 矩阵的计算更稳定。为此,一些学者做了相关研究。数据归一化过程中,分别使用两个矩阵 T 和 T' ,将 m_i 和 m_i' 变换到 $\hat{m}_i = T m_i$ 和 $\hat{m}_i' = T' m_i'$,然后使用归一化数据计算 $\hat{F}$,最终使 $F = T^T \hat{F} T'$ 。Zhang^[18] 将所有数据归一化到 $[-1, 1]$ 之间,并且点集的重心位于坐标原点。另外,Hartley^[15] 提出将点集进行缩放和平移,最终使每一个点到原点的距离均值为 $\sqrt{2}$ 。换句话说,处理后所有点的“平均”值齐次坐标为 $[1, 1, 1]^T$ 。实验表明,Hartley 的方法可以得到更为精确的结果。

4.2 随机样本抽取

LMedS 和 RANSAC 实际上是随机优化方法,其计算效

率可由类似蒙特卡罗法来达到。然而,这样产生的某个样本的 7 个点相互之间可能很接近。这种情况应该避免,因为这些点求出的 F 矩阵很不稳定,其结果无多大用处。为了获得更高的稳定性和效率,Zhang 等^[18] 根据分块技术(bucketing technique)开发了规则的随机抽取法(random selection method)。其过程简述如下:

首先计算图像点坐标的最小和最大值(min, max)。然后,将图像区域均匀分成 $b \times b$ 块(图 2(a)中取 $b=8$)。每块内的点构成点匹配的子集,而不含任何点的块被舍去。为了产生 7 个点的样本,先随机抽取互不相同的 7 块,然后在每块中随机抽取一个匹配。假设错误匹配在空间上均匀分布,为保证所有的匹配被抽取的概率相同,采用了蒙特卡罗采样方法。如果共有 l 块,将区域 $[0, 1]$ 分成 l 个间隔,使得第 i 个间隔的宽度等于 $\frac{n_i}{\sum_i n_i}$,其中 n_i 为第 i 块的点匹配数目。在块的随机抽取中,如果产生的一个 $[0, 1]$ 均匀分布随机数落在第 i 个间隔上,那么第 i 块就被选中(如图 2(b)所示)。

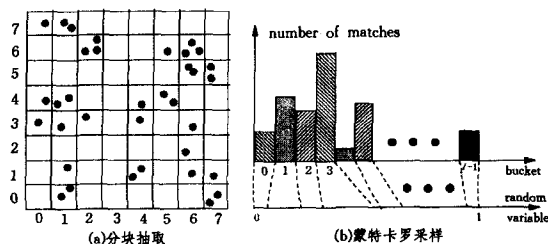


图2 随机抽取方法

5 计算实验与结论

5.1 实现的方法及评价指标

为了比较、评价以上各种算法,本文使用模拟数据和真实图像数据进行了相关实验。其中实现的方法包括前述 3 类的共 11 种:

- (1)线性方法
8 点线性算法(1-LinEig, 2-LinLS);
- (2)非线性方法
极线距离加权迭代法(3-IterEig, 4-IterLS);
梯度加权迭代法(5-GradEig, 6-GradLS);
非线性参数空间优化法(7-NonParam);
- (3)鲁棒方法
M 估计器(8-M-Est);
RANSAC(9);
最小平方中值(10-LMedS);
LMedS + M-Est(11)。

在第 1 至第 6 种方法的缩写中,后缀“Eig”表示采用线性特征分析技术,而“LS”表示采用线性二乘技术。另外,LMedS + M-Est 表示先使用 LMedS 方法剔除错误数据并估计误差协方差,然后使用 M-Estimator 方法求精。为了保证数值求解的稳定性,本文在上述算法实现中对输入数据都做了归一化处理。对于各种算法的性能比较,本文采用所有匹配点到对应极线距离的均方根(Root of Mean Squared-RMS)作为评价指标函数,即:

$$rms = \left(\frac{1}{n} \sum_i (d^2(\bar{m}_i, F^T \bar{m}_i') + d^2(\bar{m}_i', F \bar{m}_i)) \right)^{1/2}$$

5.2 数值仿真实验

(1)实验方法

模拟透视成像摄像机,其内参数为 $K = \begin{bmatrix} 600 & 0 & 256 \\ 0 & 600 & 256 \\ 0 & 0 & 1 \end{bmatrix}$ 。被观测目标是 200 个 3D 点,随机分布在边长为 360 单位的立方体中,立方体中心距离摄像机光心 600 个单位。两台摄像机相对运动参数为:平移向量 $t=[40, 0, 0]^T$ 、旋转轴方向为 $[0, 0, 1]^T$ 、旋转角度为 10° 。根据摄像机成像模型,将 3D 点投影到两个成像平面上,然后叠加零均值高斯噪声,其方差 $\sigma(\text{noise sigma})$ 从 0 到 1 变化,步长 0.1,同时错误数据所占百分比 $r(\text{outlier ratio})$ 从 0% 到 50% 变化,步长 10%。

(2)实验结果

将几种典型方法计算结果的误差分布曲面显示在图 3 中。

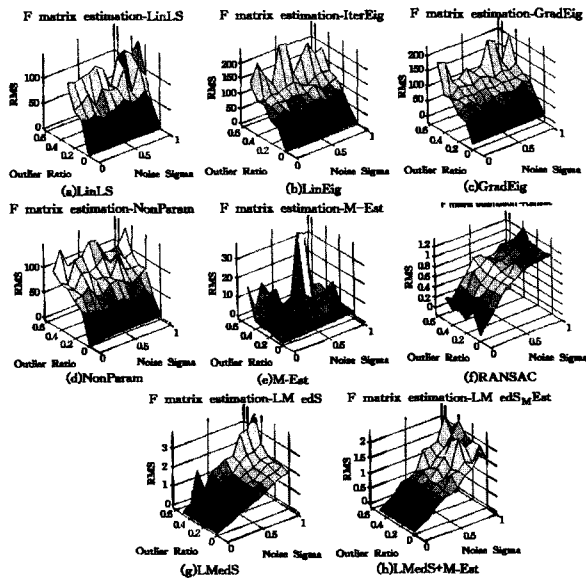


图 3 基础矩阵计算的仿真实验

5.3 真实图像实验

(1)实验方法

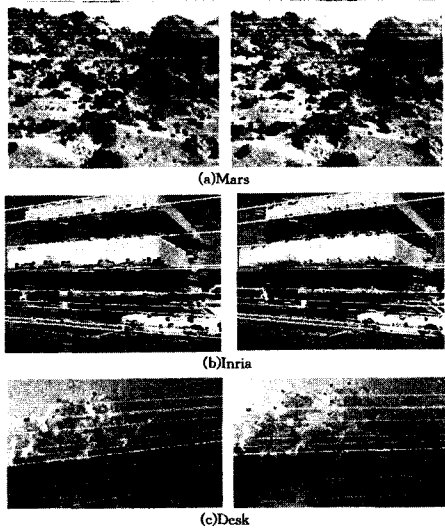


图 4 真实图像对和初始特征点匹配

首先,采用改进的 Harris 角点检测算法分别在左右两幅图像中检测特征点,然后,使用零均值相关系数(ZNCC)匹配方法获得初始匹配;最后,将匹配结果作为输入数据计算两幅图像的基础矩阵。

(2)实验结果

这里给出 3 组图像的实验结果。其中,Mars 为室外自然景物(512×512 ,图 4(a)),Inria 为室外人造建筑(512×256 ,图 4(b)),Desk 为室内人造物(800×600 ,图 4(c))。图 4 中显示了初始匹配特征点及其视差向量。图中还显示了通过 LMedS+M_Est 方法得到的部分极线。为便于比较,将数值仿真实验结果同真实图像的实验结果一并在表 1 中列出。

表 1 基础矩阵计算的实验结果

1-LinEig, 2-LinLS, 3-IterEig, 4-IterLS, 5-GradEig, 6-GradLS, 7-NonParam, 8-M-Est, 9-RANSAC, 10-LMedS, 11-M-Est+LMedS											
	线性方法				非线性方法				鲁棒方法		
Methods	1	2	3	4	5	6	7	8	9	10	11
$\sigma=0.0, r=0.0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\sigma=0.0, r=0.1$	68.798	61.568	57.874	63.449	62.217	63.449	57.440	0.270	0.391	0.000	0.000
$\sigma=0.1, r=0.0$	0.210	0.209	0.209	0.209	0.209	0.209	0.209	0.169	0.210	0.180	0.161
$\sigma=0.5, r=0.0$	0.993	1.016	0.989	1.008	0.989	1.008	0.988	0.803	0.820	0.868	0.785
$\sigma=0.5, r=0.1$	57.453	57.396	54.088	69.793	62.823	69.793	50.206	2.796	0.802	0.900	0.745
$\sigma=1.0, r=0.0$	1.885	2.026	1.902	2.002	1.901	2.002	1.870	1.709	1.028	1.758	1.493
$\sigma=1.0, r=0.1$	36.554	35.502	32.072	32.081	44.455	32.081	31.367	2.142	0.990	1.886	1.839
Mars	13.113	13.205	12.686	12.887	12.691	12.887	12.900	0.854	0.759	0.872	0.692
Inria	5.979	5.968	8.931	7.425	8.771	7.425	5.960	1.667	0.818	1.358	0.976
Desk	31.568	29.882	33.038	30.417	31.277	39.416	29.671	11.950	1.178	3.666	1.363

6 分析与结论

以仿真实验以及真实图像实验结果为依据,本文对不同的 F 矩阵计算方法有如下分析和结论。

(1)线性方法:如果特征点定位较精确并且没有错误匹配,那么可以得到很好的结果。但是,这类方法对错误匹配的鲁棒性非常差。

(2)非线性迭代方法:在一定程度上能够处理定位噪声的影响,但是实验结果表明改善效果不明显,同时,当存在错误匹配时,效率很低。

(3)参数空间优化方法:比线性方法和迭代方法好,但是同样不能处理错误匹配问题。

(4)鲁棒方法:能够同时处理数据噪声和错误匹配。当存在错误匹配时,M-Estimator 方法的性能有所降低,特别是对于 Desk 图像,由于初始匹配错误率较大导致 rms 较大。这一现象证实了 M-Estimator 对定位噪声是鲁棒的,但是对错误数据的鲁棒性较差。相比之下,RANSAC 和 LMedS 体现出较好的品质。

(5)如果使用 LMedS 方法预先剔除错误匹配,然后使用 M-Estimator,可以得到最好的结果。

(6)同线性最小二乘优化方法比较,基于特征分析技术的优化方法效果更好。

- [5] Prusinkiewicz P, Lindenmeyer A, et al. The Algorithmic Beauty of Plants [M]. Springer-Verlag, 1990
- [6] 王辉. 基于 L 系统的虚拟园林观赏树木生长建模研究[D]. 长沙: 中南林业科技大学, 2006: 7-9
- [7] Hearn D, Baker M P. 计算机图形学[M]. 蔡士杰, 吴春熔, 孙正兴, 译. 电子工业出版社, 2002
- [8] Farm G. Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide, 4th Edition [M]. Boston: Academic Press, 1996
- [9] 李钢, 刘华明. 基于 NURBS 的扫描曲面造型方法的研究[J]. 机械研究与应用, 2000, 13(3): 5-6
- [10] Cook R L. Shade trees[J]. Computer Graphics, 1984, 18(3):

- [11] Hanrahan P, Lawson J. A language for shading and lighting calculations[J]. Computer Graphics, 1990, 24(4): 289-298
- [12] Pharr M, Hanrahan P. Geometry caching for ray-tracing displacement maps[C]//Proc. of the 7th Eurographics Workshop on Rendering. Porto, Portugal, 1996: 31-40
- [13] Schaufler G, Priglinger M. Efficient displacement mapping by image warping[C]//Proc of the 10th Eurographics Workshop on Rendering. Granada, Spain, Eurographics, 1999: 175-186
- [14] 付恺, 李春霞, 杨克俭, 等. Bump Mapping 原理及在 OpenGL 下的实现[J]. 交通与计算机, 2004, 22(2): 54-57
- [15] 杨刚. 全景图拼接算法的设计与实现[J]. 重庆工学院学报: 自然科学版, 2007, 21(19): 107-110

(上接第 247 页)

结束语 F 矩阵是许多计算机视觉应用中的重要参数, 其计算的准确性决定了后续的处理步骤能否成功。本文将常见的 F 矩阵估计方法划分为线性法、非线性法和鲁棒法 3 大类, 共计 11 种, 通过仿真数据和真实图像实验对各自的性能进行了评估。实际应用中, 由于视觉系统所采用的图像特征定位及匹配方法各有差异, 可根据具体情况合理选择本文介绍的 F 矩阵计算方法, 以实现高精度的 3D 信息获取。

参考文献

- [1] 马颂德, 张正友. 计算机视觉-计算理论与算法基础[M]. 北京: 科学出版社, 1998
- [2] Luong Q-T, Faugeras O D. The fundamental matrix: Theory, algorithms and stability analysis[J]. International Journal of Computer Vision, 1996, 1(17): 43-76
- [3] Longuet-Higgins H C. A computer algorithm for reconstructing a scene from two projections[J]. Nature, 1981, 293: 133-135
- [4] Mester R. On the mathematical structure of direction and motion estimation[C]//3rd International Symposium on Physics in Signal and Image Processing (PSIP). Grenoble, France, January 2003
- [5] Pollefeys M, et al. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters[J]. International Journal of Computer Vision, 1999, 32(1): 7-20
- [6] Zhang Z, Deriche R, Faugeras O, et al. A robust technique for matching two uncalibrated images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 82(2): 1129-1139
- [7] Hartley R, Gupta R, Chang T. Stereo from uncalibrated cameras [C]//Proceedings of the International Conference on Computer Vision and Pattern Recognition. IEEE, Urbana Champaign, IL
- [8] Mohr R, Boufama B, Brand P. Accurate projective reconstruction [C]//Mundy J, Zisserman A, eds. Applications of Invariance in Computer Vision. Vol. 825 of Lecture Notes in Computer Science, Berlin: Springer-Verlag, 1993: 257-276
- [9] Zhang Z. A new multistage approach to motion and structure estimation: From essential parameters to Euclidean motion via fundamental matrix[R]. 2910. INRIA Sophia-Antipolis, France
- [10] Maybank S J, Faugeras O. A theory of self-calibration of a moving camera[J]. International Journal of Computer Vision, 1992, 8(2): 123-151

- [11] Faugeras D, Luong Q T, Maybank S J. Camera self-calibration: theory and experiment[C]//Proceedings of the European Conference on Computer Vision. Santa Margherita L., 1992: 321-334
- [12] Lourakis M I A, Deriche R. Camera self-calibration using the singular value decomposition of the fundamental matrix: From point correspondences to 3D measurements[R]. 3748. INRIA Sophia-Antipolis, August 1999
- [13] Lourakis M I A, Deriche R. Camera self-calibration using the Kruppa equations and the SVD of the fundamental matrix: The case of varying intrinsic parameters [R]. 3911. INRIA Sophia-Antipolis, Mars, 2000
- [14] Hartley R I, Zisserman A. Multiple view geometry in computer vision[M]. Cambridge University Press, 2000
- [15] Hartley R I. In defense of the 8-point algorithm[C]//Proceedings of the 5th Proc. International Conference on Computer Vision. Boston, MA: IEEE Computer Society Press, 1995: 1064-1070
- [16] Faugeras Q, Mourrain B. On the geometry and algebra of the point and line correspondences between n images[C]//Proceedings of the 5th Proc. International Conference on Computer Vision. Boston, MA: IEEE Computer Society Press, 1995: 951-956
- [17] Zhang Z, Deriche R, Faugeras O, et al. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry[J]. Artificial Intelligence Journal, 1995, 78(10): 87-119
- [18] Zhang Z. Determining the epipolar geometry and its uncertainty: A review[J]. International Journal of Computer Vision, 1998, 27(2): 161-195
- [19] Zhang Z. Parameter Estimation Techniques: A tutorial with application to conic fitting[J]. Image and Vision Computing Journal, 1997, 15(1): 59-76
- [20] Fischler M A, Bolles R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography[J]. Comm. Assoc. Comp. Mach., 1981, 24(6): 381-395
- [21] Torr P H S. Motion segmentation and outlier detection [D]. Dept. of Engineering Science, University of Oxford, 1995
- [22] Zuliani M, Kenney C S, Manjunath B S. The MultiRANSAC algorithm and its application to detect planar homographies[C]//IEEE International Conference on Image Processing. September 2005