

文章编号: 0253-2239(2010)11-3164-07

基于随机蕨的光电成像末端制导目标初始化

陈 冰 赵亦工 李 欣

(西安电子科技大学模式识别与智能控制研究所, 陕西 西安 710071)

摘要 为实现光电成像末端制导中的自适应目标初始化, 针对末端制导景象匹配中图像存在尺度、旋转、灰度和3D视角差异, 及传统方法运算量较大的问题, 基于随机蕨分类器构造了一种新的景象匹配算法。算法首先利用基准图像进行分类器训练, 然后基于该分类器对实时图像进行特征匹配。为剔除误匹配特征对, 对初始匹配特征对中的对应区域分别进行尺度不变特征变换(SIFT)特征描述, 基于马氏距离准则进行误匹配特征对剔除。根据顺序抽样一致性算法(PROSAC)对剩余的匹配特征对估计两图像的外极几何关系, 最终根据外极几何关系求得目标在实时图像中的位置和尺寸信息。仿真结果表明, 该算法能够在光电成像末端制导过程中实现稳定的目标初始化, 在极端条件下的稳定性优于原随机蕨分类器算法。

关键词 模式识别; 景象匹配; 随机蕨分类器; 尺度不变特征变换(SIFT)描述符; 误匹配特征对剔除; 外极几何关系中图分类号 TP391.41 文献标识码 A doi: 10.3788/AOS20103011.3164

Target Initialization Based on Random Ferns during Electro-Optical Imaging Terminal Guidance

Chen Bing Zhao Yigong Li Xin

(Research Institute of Pattern Recognition and Intelligent Control, Xidian University, Xi'an, Shaanxi 710071, China)

Abstract In order to solve the problems of image differences in scale, rotation, grayscale and 3D viewing angle, and achieve adaptive target initialization during electro-optical imaging terminal guidance, a new scene matching framework based on random Ferns classifier was constructed. The process of realizing the method includes classifier off-line training which yields fast run-time performance was performed. Candidate match regions between reference image and run-time image were found by the classifier. Scale invariant feature transform (SIFT) descriptors of corresponding regions in each candidate matches were computed and false matches feature pairs rejecting was performed based on Mahalanobis distance criterion. Epipolar geometry of the two images was estimated by applying PROSAC to the central locations of the corresponding regions in the final matches. Target location and size in run-time image were computed based on the epipolar geometry. Simulation results show that the proposed method provides robust target initialization during electro-optical imaging terminal guidance and is more stable than original Ferns methods under severe conditions.

Key words pattern recognition; scene matching; random Ferns classifier; scale invariant feature transform (SIFT) descriptor; false matches feature pairs rejecting; epipolar geometry

1 引 言

稳定准确的目标初始化是光电成像末端制导过程中对目标实现稳定跟踪的前提。导弹进入目标区域时, 常利用存储于导引头中的基准图像与拍摄到的实时图像进行景象匹配, 求得实时图像中目标的

位置和尺寸, 从而为末端制导目标跟踪提供准确的目标初始化信息。由于基准图像与实时图像在成像设备, 拍摄时间、拍摄角度、拍摄高度和方向等方面的不同, 两图像在尺度、旋转角、灰度及3D视角方面均可能有较大差异, 这种情况下的景象匹配属于

收稿日期: 2010-01-05; 收到修改稿日期: 2010-05-14

作者简介: 陈冰(1984—), 男, 博士研究生, 主要从事模式识别及图像处理等方面的研究。E-mail: ice32bit@yahoo.cn

导师简介: 赵亦工(1960—), 男, 博士, 教授, 主要从事模式识别及图像处理等方面的研究。

E-mail: ygzha@xidian.edu.cn

本文电子版彩色效果详见中国光学期刊网(<http://www.opticsjournal.net>)

宽基线图像匹配问题^[1]。

近年来,宽基线图像匹配算法不断涌现,大多沿袭了 Schmid 等^[2]构造的理论框架,通常分为四步:1)提取稳定的局部特征;2)对特征进行不变性描述;3)特征匹配;4)基于外极几何约束,求得两图像间的对应关系。较经典的有 Harris 仿射不变性特征(Harris-affine)^[3]、尺度不变性特征变换(SIFT)^[4~6]、快速稳健性特征(SURF)^[7]和最稳定极值区域(MSER)^[8]等。参考文献[3~7]中的算法为获取特征的各种不变性,运算量较大,难以满足实时性,MSER 虽然能够满足实时性,但对图像模糊和尺度差异的稳健性较差^[9]。在成像环境较复杂且对运算量要求严格的末端制导领域,上述宽基线图像匹配算法均不适用。为降低运算量, Lepetit 等^[10]把宽基线图像匹配视为分类问题,将运算量较大的部分转移到分类器训练过程中,构造了一种名为随机蕨(Random ferns)的非层次结构特征,基于朴素贝叶斯分类算法对局部图像进行快速识别^[11,12]。目前,该方法已被用于增强现实^[13]和同时定位与地图创建^[14]等方面。

本文基于随机蕨提出了一种新的光电成像末端制导目标初始化方法。通过对基准图像进行随机的仿射变换模拟目标在任意视角条件下的成像,以获得大量的训练样本,从中提取稳定的角点来构造随机蕨,通过训练得出基于随机蕨的朴素贝叶斯分类器(简称随机蕨分类器)以实现快速稳定的宽基线图像匹配,求得最终的目标初始信息。另外,针对随机蕨分类器可能产生大量的误匹配特征对,导致外极几何关系求取不稳定的问题。利用 SIFT 描述符对初始匹配的特征对进行描述,并根据马氏距离准则构造了一种误匹配特征对剔除策略,以提高正确匹配的特征对在初始匹配特征对中的比例,确保能够精确获取外极几何关系。

2 末端制导目标初始化算法原理框图

基于随机蕨分类器构造的光电成像末端制导目标初始化算法的算法框架如图 1 所示。

3 基于随机蕨的朴素贝叶斯分类

3.1 算法思想

设基准图像中的特征点数为 H ,将以特征点 k 为中心的局部区域 patch_k 所有可能的视图情况均归为一类,将各特征区域所属的类用 $c_i, i=1, \dots, H$

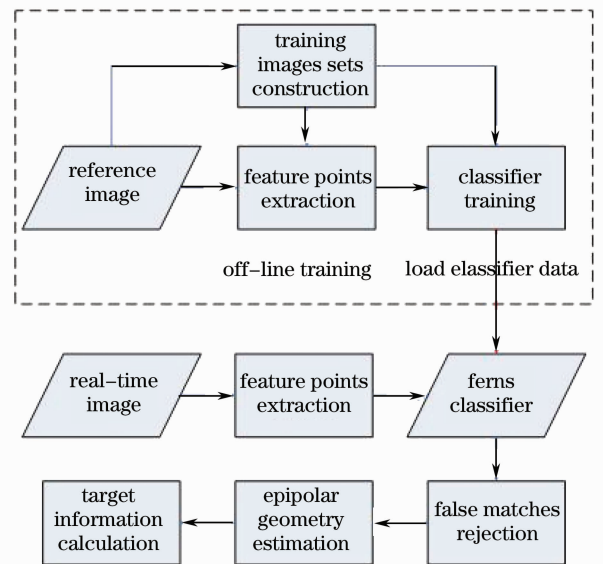


图 1 光电成像末端制导目标初始化算法原理框图
Fig.1 Block diagram of the terminal guidance target initializing algorithm

表示。则可将查找实时图像中特征点 k^{input} 在基准图像上的对应特征点视为对 $\text{patch}_k^{\text{input}}$ 的分类问题^[12]。

设 $f_j, j=1, \dots, N$ 为基于 $\text{patch}_k^{\text{input}}$ 的二元特征集,则 $\text{patch}_k^{\text{input}}$ 的所属类别为

$$\hat{c}_i = \arg \max_{c_i} P(C = c_i | f_1, f_2, \dots, f_N), \quad (1)$$

式中 C 为表示类的随机变量,由贝叶斯公式可得

$$P(C = c_i | f_1, f_2, \dots, f_N) = \frac{P(f_1, f_2, \dots, f_N | C = c_i)P(C = c_i)}{P(f_1, f_2, \dots, f_N)}, \quad (2)$$

设先验概率 $P(C)$ 为均匀分布,(2)式分母部分与类别无关,则(1)式可简化为

$$\hat{c}_i = \arg \max_{c_i} P(f_1, f_2, \dots, f_N | C = c_i), \quad (3)$$

设 $\text{patch}_k^{\text{input}}$ 的尺寸为 $L \times L$, (取 $L=32$), 则二元特征 f_j 的值取决于 $\text{patch}_k^{\text{input}}$ 中在分类器训练时随机生成的两像素位置 $d_{j,1}$ 和 $d_{j,2}$ 的灰度大小 $I_{d_{j,1}}$ 和 $I_{d_{j,2}}$ 比较的结果,即

$$f_j = \begin{cases} 1 & I_{d_{j,1}} < I_{d_{j,2}} \\ 0 & \text{otherwise} \end{cases}, \quad (4)$$

由于 f_j 的形式很简单,因此需要有足够多的特征以确保实现正确分类,取 $N \approx 300$,即随机选取 300 对像素位置进行灰度大小比较,则对(3)式中联合分布概率的完整表达形式的求解需要的存储量为 $H \times 2^N$ 字节。为降低存储量,并保证 f_j 之间具有足够的相关性,将 f_j 分为 M 组,每组包含 $S=N/M$ 个二元特征,并假设不同组的二元特征之间相互独立,组内二元特征之间具有相关性,将这些组定义为蕨

特征(Ferns),则(3)式中的条件概率可近似为

$$P(f_1, f_2, \dots, f_N | C = c_i) = \prod_{m=1}^M P(F_m | C = c_i), \quad (5)$$

式中 $F_m = [f_{\sigma(m,1)}, f_{\sigma(m,2)}, \dots, f_{\sigma(m,S)}]$, $m = 1, \dots, M$, 表示第 m 个蕨, $\sigma(m, j)$ 表示范围为 $1, \dots, N$ 的随机函数。则 $\text{patch}_{k^{\text{input}}}$ 的所属类别为

$$\hat{c}_i = \arg \max_{c_i} \prod_{m=1}^M P(F_m | C = c_i). \quad (6)$$

要对(6)式进行求解,只需估计每个蕨 F_m 和类 c_i 的类条件概率 $P(F_m | C = c_i)$ 。与(3)式相比,(6)式的存储量为 $M \times 2^S$ 字节,取 $M = 40, S = 11$,在保证算法具有较高性能的同时,每一类只需 8×10^4 字节的存储量,远小于 $N = 440$ 时 2^N 字节的运算量,有利于工程实现。

3.2 分类器训练

3.2.1 训练集建立

为使分类器对图像投影变形、光照变化、图像模糊以及噪声具有较强的稳健性,需构造能反映上述变形的训练集对分类器进行训练。

当摄像头距离目标较远时,目标上的局部区域可被视为平面,则通过对当前视角下的目标图像进行仿射变换能够模拟目标在其他视角条件下的情况,设仿射变换矩阵为 \mathbf{A} ,可被分解为如下形式:

$$\mathbf{A} = \mathbf{R}_\theta \mathbf{R}_\phi^{-1} \mathbf{S} \mathbf{R}_\phi, \quad (7)$$

式中 \mathbf{R}_θ 和 \mathbf{R}_ϕ 是两个角度参数分别为 θ 和 ϕ 的旋转矩阵, $\mathbf{S} = \text{diag}[\lambda_1, \lambda_2]$ 为尺度矩阵。分别在 $[-\pi, +\pi]$ 范围内随机选取 θ 和 ϕ , 在 $[0.6, 1.5]$ 范围内随机选取 λ_1 和 λ_2 , 即可构造出随机的仿射变换矩阵 \mathbf{A}_{rand} , 设基准图像为 \mathbf{I}_0 , 则与 \mathbf{I}_0 具有随机的图像尺度、旋转角和 3D 视角差异的图像 \mathbf{I}_{rand} 为

$$\mathbf{I}_{\text{rand}} = \mathbf{A}_{\text{rand}} \mathbf{I}_0. \quad (8)$$

为提高算法对光照变化的稳健性,随机改变训练图像 \mathbf{I}_{rand} 的平均灰度,以模拟不同光照条件下的图像,将改变平均灰度后的图像记作 $\mathbf{I}_{\text{scale}}$ 。为提高算法对噪声和图像模糊的稳健性,对 $\mathbf{I}_{\text{scale}}$ 添加均值为 0, 方差为 25 的高斯白噪声,得到加噪图像 $\mathbf{I}_{\text{noise}}$, 然后对 $\mathbf{I}_{\text{noise}}$ 进行随机的高斯平滑,最终求得训练图像 $\mathbf{I}_{\text{train}}$, 并将由上述不同随机参数生成的训练图像所构成的集合记作 B_{train} , 文中集合 B_{train} 内的训练样本数为 10000, 即随机生成的训练图像数为 10000。

3.2.2 特征点提取

要使分类器足够稳定,必须提取在较大程度投影变形、光照变化、图像模糊以及噪声存在情况下仍

然稳定存在的特征点。利用加速分割测试特征(FAST)方法^[15]对基准图像 \mathbf{I}_0 进行特征点提取,设提取到的某特征点为 \mathbf{k} , 则 \mathbf{k} 应在上述干扰存在的情况下仍然具有较高的被检测概率 $P(\mathbf{k})$ 。

设对 \mathbf{I}_0 采用第 3.2.1 节中的方法得到的新视角模拟图像为 \mathbf{I}_i , 由(8)式可得

$$\mathbf{I}_i = \mathbf{A}_{\text{rand}} \mathbf{I}_0. \quad (9)$$

利用 FAST 方法对图像 \mathbf{I}_i 进行特征点提取,设 $\bar{\mathbf{k}}$ 为在 \mathbf{I}_i 中求得的某特征点,根据 \mathbf{A}_{rand} 对其做逆仿射变换,可求得该特征点在基准图像 \mathbf{I}_0 中对应的特征点 \mathbf{k}' 的位置:

$$\mathbf{k}' = \mathbf{A}_{\text{rand}}^{-1} \bar{\mathbf{k}}. \quad (10)$$

因此,随机产生 N_{total} 幅模拟图像,通过对是否在当前模拟图像 \mathbf{I}_i 中提取到了某特征点 \mathbf{k} 所对应的特征点 $\bar{\mathbf{k}}$ 进行计数,即可求得特征点 \mathbf{k} 的被检测概率为

$$P(\mathbf{k}) = \frac{N_{\text{detected}}}{N_{\text{total}}}, \quad (11)$$

式中 N_{detected} 为在 N_{total} 幅模拟图像中成功提取到与特征点 \mathbf{k} 向对应的特征点 $\bar{\mathbf{k}}$ 的总图像数。令 $N_{\text{total}} = 5 \times 10^3$, 取被检测概率 $P(\mathbf{k})$ 最高的 400 个特征点作为分类器的初始类进行训练,则在第 3.1 节中,各特征区域所属的类 $c_i, i = 1, \dots, H$ 即对应于分类器的初始类, $H = 400$ 。

3.2.3 训练

在求得训练集合 B_{train} 和分类器初始类 c_i 之后,在 $L \times L$ 范围内(取 $L = 32$),按照均匀分布随机选取 $M \times S$ ($M = 40, S = 11$) 对像素 $d_{j,1}$ 和 $d_{j,2}$ 的位置。对每个类 c_i , 将这 $M \times S$ 对像素位置上的灰度值分别代入(4)式,计算类 c_i 的 M 个蕨中的 S 个二元特征 f_j 的值,并据此对(6)式中的每个蕨 F_m 和类 c_i 的类条件概率 $P(F_m | C = c_i)$ 进行估计。

设 F_m 中由各二元特征序列所组成的以 2 为底的指数的值为 x , 则蕨 F_m 能取到的最大值为 $x_{\text{max}} = 2^S$, 将 $P(F_m | C = c_i)$ 写作

$$p_{x,c_i} = P(F_m = x | C = c_i), \quad (12)$$

则可以采用(14)式为约束条件对每个 $p_{x,c_i}, x = 1, 2, \dots, x_{\text{max}}$ 进行估计,估计公式为

$$p_{x,c_i} = \frac{N_{x,c_i} + N_r}{N_{c_i} + x_{\text{max}} N_r}, \quad (13)$$

$$\sum_{x=1}^{x_{\text{max}}} p_{x,c_i} = 1, \quad (14)$$

式中 N_{x,c_i} 为类 c_i 的训练样本中 F_m 取值为 x 的样本数, N_{c_i} 为类 c_i 的总样本数, N_r 为常数项,取 $N_r = 1$ 。

4 误匹配特征对的剔除

4.1 模型拟合

随机蕨分类器能够快速求得两图像间的初始匹配特征对,该匹配特征对将被用来做模型拟合,求取基准图像和实时图像间的外极几何关系。外极几何关系可用基础矩阵 \mathbf{F} 表示。设 $\mathbf{x}_i = [x, y, 1]^T$ 和 $\mathbf{x}'_i = [x', y', 1]^T$ 分别为某对匹配特征对中对应区域的中心位置分别在基准图像和实时图像中的齐次坐标,则匹配特征对的对应关系式为

$$\mathbf{x}\mathbf{F}\mathbf{x}' = 0, \quad (15)$$

式中 \mathbf{F} 为 3 阶秩为 2 的基础矩阵,通常采用随机抽样一致性算法(RANSAC)^[16,17] 剔除误匹配特征对,估计基础矩阵 \mathbf{F} 的最优解。然而,当误匹配特征对的数目在总的匹配特征对中所占的比例过大时,RANSAC 算法的性能将急剧下降。由于随机蕨分类器算法中缺少对误匹配特征对的剔除策略,将产

生大量的误匹配特征对。

为降低误匹配特征对较多给后期的模型拟合带来的影响,Ozuysal 等基于顺序抽样一致性算法(PROSAC)^[18] 来进行模型拟合,剔除误匹配特征对,该方法优先选取匹配置信度较高的匹配特征对进行模型拟合,有利于拟合过程的快速收敛,提高了在误匹配特征对所占比率较高时的算法性能。

然而,当基准图像与实时图像之间的差异过大时,会导致误匹配特征对在初始匹配特征对中所占比率过高,此时 PROSAC 算法也无法对基础矩阵 \mathbf{F} 进行正确求解。如图 2(a)所示的两幅室外场景的城市图像,由于两图像间的尺度差异过大,产生的误匹配特征对过多[如图 2(b)所示],导致模型拟合时求得的最终匹配特征对的数目为 0,PROSAC 算法失效。

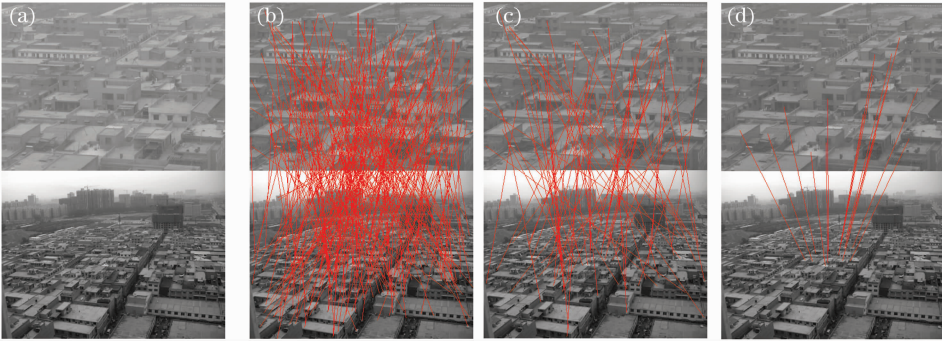


图 2 添加误匹配特征对剔除策略后的宽基线图像匹配结果示例。(a)原图像;(b)初始匹配特征对;(c)误匹配特征对剔除后的匹配特征对;(d)最终匹配结果

Fig. 2 Examples of wide baseline image matching after adding the false matches feature pair rejecting strategy.

(a) original image; (b) initial matches feature pair; (c) after false matches feature pair rejection; (d) final matches

要从根本上解决该问题,需要对基于随机蕨分类器的宽基线图像匹配算法的算法框架进行改进,在进入模型拟合步骤之前添加误匹配特征对的剔除策略,降低误匹配特征对在初始匹配特征对中所占的比例。

4.2 误匹配特征对剔除策略

通过对传统的宽基线图像匹配算法的大量仿真和分析,发现在 SIFT, SURF 等传统方法求得的初始匹配特征对中,误匹配特征对较少,原因在于传统方法大多采用类似于欧氏距离比值准则^[19] 的机制进行特征匹配,只保留匹配置信度较高的特征对。而基于随机蕨分类器采用的方法类似于最小欧氏距离准则,保留了所有待匹配特征的最近邻作为与其匹配的特征,该准则缺少对匹配特征对的选取策略,势必会引入大量的误匹配特征。为解决此问题,将

传统方法与随机蕨分类器相结合,提出了一种误匹配特征对剔除策略。

设基准图像 \mathbf{I} 和实时图像 \mathbf{I}' 之间的初始匹配特征对构成的集合为

$$D = [(P_i, Q_i) \mid P_i \leftrightarrow Q_i, P_i \subset \mathbf{I}, Q_i \subset \mathbf{I}', i = 1, 2, \dots, n].$$

式中 P_i 和 Q_i 分别表示两图像的初始匹配特征对中对应的特征区域, n 表示初始匹配特征对数。算法步骤如下:

1) 分别构造特征区域 P_i 和 Q_i 的特征描述符 \mathbf{V}_P 和 \mathbf{V}_Q , 特征描述符的构造可选取目前比较成熟的算法,如 SIFT 方法,也可以根据算法对各种图像形变不变性的不同需求自己构造^[16]。文中采用维数为 128 的 SIFT 描述符对特征区域进行描述。

2) 对集合 D 中对应匹配的特征区域 P_i 和 Q_i ,

$i = 1, 2, \dots, n$, 计算其对应特征描述符 \mathbf{V}_{P_i} 和 \mathbf{V}_{Q_i} 之间的马氏距离 (Mahalanobis distance) $d_i(\mathbf{V}_P, \mathbf{V}_Q)$, 马氏距离公式为

$$d_i(\mathbf{V}_P, \mathbf{V}_Q) = \sqrt{(\mathbf{V}_{P_i} - \mathbf{V}_{Q_i})^T \boldsymbol{\Sigma}^{-1} (\mathbf{V}_{P_i} - \mathbf{V}_{Q_i})}, \quad (16)$$

式中 $\boldsymbol{\Sigma}$ 表示 \mathbf{V}_{P_i} 和 \mathbf{V}_{Q_i} 的协方差矩阵。 $d_i(\mathbf{V}_P, \mathbf{V}_Q)$ 能够度量特征描述符 \mathbf{V}_{P_i} 和 \mathbf{V}_{Q_i} 的差异程度。

3) 设集合 D 中所有对应特征区域之间的马氏距离构成的集合为

$E = [d_i(\mathbf{V}_P, \mathbf{V}_Q) \mid (P_i, Q_i) \in D, i = 1, 2, \dots, n]$, 若集合 D 中的某特征区域对 (P_i, Q_i) 被判定为误匹配特征对, 记作 $P_i \not\leftrightarrow Q_i$, 若被判定为匹配特征对, 记作 $P_i \leftrightarrow Q_i$, 则基于马氏距离的误匹配特征对剔除公式为

$$\begin{cases} P_i \leftrightarrow Q_i & d_i(\mathbf{V}_P, \mathbf{V}_Q) < T_d \\ P_i \not\leftrightarrow Q_i & d_i(\mathbf{V}_P, \mathbf{V}_Q) \geq T_d \end{cases}, \quad (17)$$

式中 T_d 为误匹配特征对剔除阈值, 将其定为

$$T_d = \mu + k\sigma, \quad (18)$$

式中 μ 为集合 E 中马氏距离的均值, σ 为其标准差, k 为权值因子, 取 $k = 3$ 。

由于误匹配特征对剔除策略针对的对象是由随机森林分类器初始匹配成功的特征区域对, 只对初始匹配的区域对进行特征描述和马氏距离度量, 不涉及在高维特征空间中进行特征向量匹配搜索的问题, 因此, 上述算法步骤消耗的运算量很小。在进行误匹配特征对剔除之后, 基于 PROSAC 方法对剩下的匹配特征对进行模型拟合, 有利于拟合过程的快速收敛和模型拟合精度的提高。

图 2 为添加了误匹配特征对剔除策略后的宽基线图像匹配示例结果。图 2(a) 为两幅基于室外场景的可见光城市图像, 由于对摄像头设置了不同的视场范围, 两图之间存在很大的尺度差异; 图 2(b) 是基于随机森林分类器得到的初始匹配特征结果, 共得到 241 对匹配特征对, 对应的匹配特征对用红线表示, 从图中可以看出, 误匹配特征对所占的比例非常高, 直接利用此结果进行模型拟合, 共求得 0 对匹配特征对, 模型拟合失败, PROSAC 算法失效。图 2(c) 为基于剔除的方法对初始匹配特征对进行误匹配特征对剔除之后的结果, 共剩余 73 对匹配特征对。剔除误匹配特征对之后, 剩余的匹配特征对中正确匹配的特征对所占的比例明显提高。图 2(d) 为基于 PROSAC 对剩余匹配特征对进行模型拟合时, 最终得到的 15 对匹配特征对, 所有匹配特征对均为正确匹配, 可以基

于上述正确匹配特征对求得两图像间的外极几何关系。由此可得, 添加了误匹配特征对剔除策略之后, 算法性能得到了明显的提高。

5 实验结果与分析

5.1 室外真实场景下的末端制导目标初始化

为对算法的实用性进行验证, 基于实地录取的可见光城市图像进行末端制导目标初始化仿真实验。图 3(a) 为两幅对同一场景在不同时间拍摄的可见光城市图像, 上图拍摄于 2008 年 6 月 24 日, 下图拍摄于 6 月 25 日, 图像尺寸为 $400 \text{ pixel} \times 300 \text{ pixel}$ 。由于天气条件不同, 两图像在灰度上具有较大差异, 又因为拍摄角度的不同, 图中建筑物具有一定的 3D 视角差异。

以图 3(a) 的上图为基准图像进行分类器训练, 以下图为实时图像进行图像匹配。如图 3(b) 所示, 图中的绿色十字表示分别从基准图像和实时图像中提取到的各特征区域的中心位置。在基准图像中提取到的特征区域数为 400 个, 由于实时图像的拍摄时间为傍晚, 光线条件较差, 因此图像偏暗, 得到的特征区域不如基准图像中的特征区域稳定, 共提取到 332 个特征区域。图 3(c) 为基于随机森林分类器所得到的初始特征匹配结果, 共得到 197 对匹配的特征区域对, 用红色连线表示, 从图中可以看出, 利用原始的随机森林分类器得到的初始匹配特征对中存在大量的误匹配特征。图 3(d) 为基于本文方法进行误匹配特征对剔除后的匹配特征对, 共剩余 107 对匹配特征对, 很明显剔除掉的特征绝大多数均为误匹配特征对, 剩下的匹配特征对中绝大多数均为正确匹配的特征对, 图 3(e) 为基于 PROSAC 方法得到剔除剩余的误匹配特征对后得到的匹配结果, 共求得 93 对匹配的匹配特征对, 且全部匹配正确, 能够据此精确地估计两图像间的外极几何关系。图 3(f) 为基于最终匹配的特征区域对估计求得的两图像的外极几何关系, 红线为图像的外极线。图 3(g) 为末端制导目标初始化结果, 图中的红色框体为对基准图像中坐标为 $(268, 62)$ 、尺寸为 $60 \text{ pixel} \times 43 \text{ pixel}$ 的初始目标, 绿色框体为根据外极几何关系在实时图像中求得的坐标为 $(201, 84)$ 、尺寸为 $59 \text{ pixel} \times 42 \text{ pixel}$ 的对应目标。可以看出, 由于外极几何关系估计准确, 能够得到基准图像中目标建筑物在实时图像中的对应位置和尺寸, 为室外真实场景下的末端制导跟踪提供稳定的初始信息。

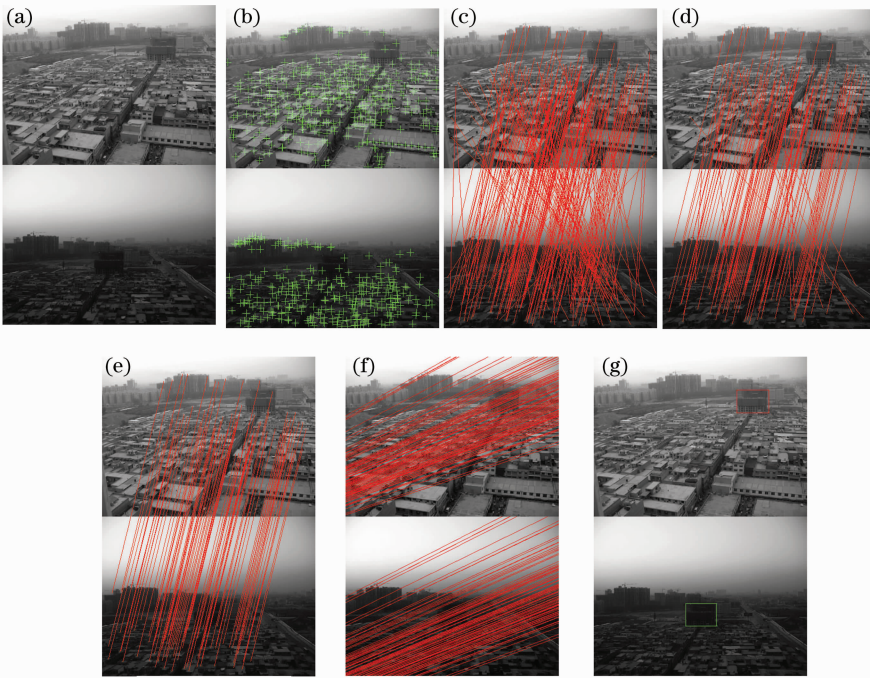


图 3 室外场景下的末端制导目标初始化示例。(a) 原图像；(b) 特征提取；(c) 特征匹配；

(d) 误匹配特征对剔除；(e) 最终匹配结果；(f) 两图像的外极几何关系；(g) 求得的目标位置及尺寸

Fig. 3 Examples of terminal guidance target initialization in outdoor environments. (a) original image; (b) feature extraction; (c) feature matches; (d) after false matches feature pair rejection; (e) final matches; (f) epipolar geometry of two images; (g) target location and size

5.2 在极端条件下的稳定性对比实验

为体现本算法在极端条件下与改进前的基于随机蕨分类器的宽基线图像匹配算法相比的优越性，

利用标准测试图像进行对比仿真实验。如图 4 所示,图 4(a)中上下两幅图像之间存在较大的尺度和旋转角差异,基于随机蕨分类器共求得 273 对初

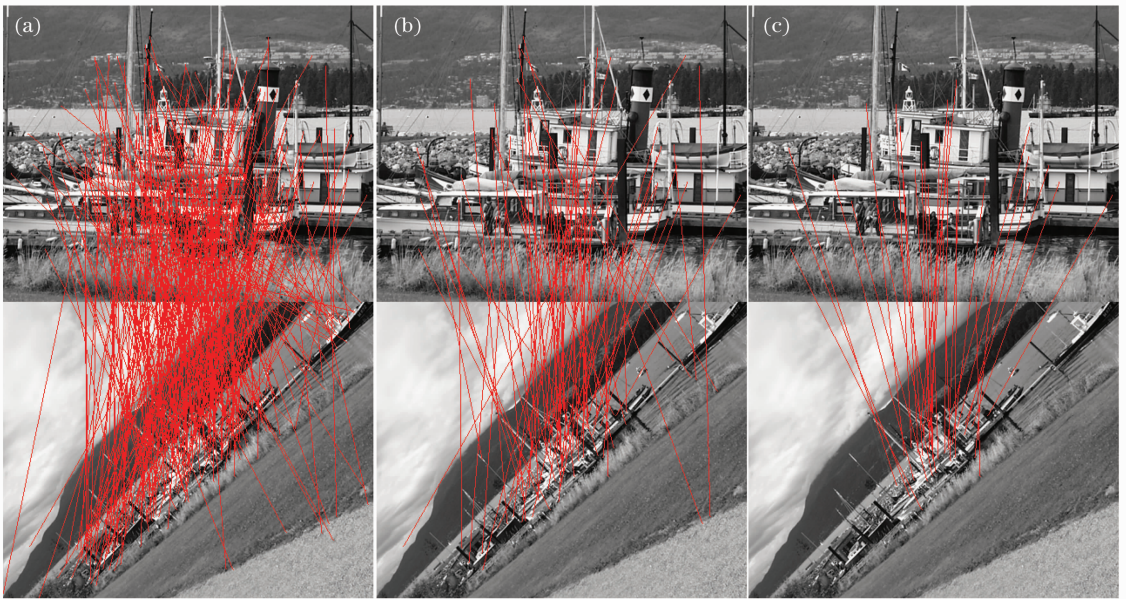


图 4 本文算法在极端条件下的稳定性实验结果。(a) 基于随机蕨求得的初始匹配特征对；(b) 误匹配特征对剔除后剩余的匹配特征对；(c) 最终匹配结果

Fig. 4 Results of stability experiment under severe condition using our method. (a) initial matching feature pair based on random ferns; (b) matches feature pair after false matches feature pair rejecting; (c) final matches

始匹配特征对,其中误匹配特征对所占的比例明显较高,给模型拟合造成了较大困难,基于 PROSAC 算法直接利用初始匹配特征进行模型拟合时,最终得到的匹配特征对数为 0,无法求得两图像间正确的外极几何关系,算法失效;图 4(b)为对初始匹配特征对进行误匹配特征对剔除后的结果,共剩余 75 对匹配特征对,与图 4(a)相比,剩余的匹配特征对中误匹配特征对所占的比例明显降低,因此,误匹配特征对剔除策略将大多数的误匹配特征对成功剔除,有利于模型拟合中对外极几何关系的正确估计。图 4(c)为基于 PROSAC 算法利用剩余的匹配特征对进行模型拟合,剔除掉剩余的误匹配特征对后的结果,最终得到 32 对匹配特征对,所有的匹配特征对均匹配正确,能够据此对外极几何关系进行精确估计。由此可见,本算法在极端条件下的稳定性要强于原始的基于随机森林分类器的宽基线图像匹配算法。

6 结 论

基于随机森林分类器构造了一种新的光电成像末端制导目标初始化方法。基于随机森林分类器进行景象匹配,将宽基线图像匹配视为分类问题,把运算量较大的特征匹配部分转移到分类器训练过程中,解决了传统宽基线图像匹配算法运算量较大的问题。针对随机森林分类器求得的误匹配特征对较多的缺陷,利用 SIFT 方法对初始匹配特征对进行特征描述,基于马氏距离准则对上述匹配特征对进行二次匹配,剔除掉部分误匹配特征对,提高了算法在极端环境条件下的稳健性。利用 PROSAC 方法剔除剩余的误匹配特征对、估计两图像的外极几何关系;最终根据外极几何关系求得目标在实时图像中的位置及尺寸信息。实验结果表明,该方法能够在光电成像末端制导阶段实现稳定的目标初始化,为后期的末端制导跟踪打下了坚实基础。

参 考 文 献

- 1 T. Tuytelaars, L. V. Gool. Matching widely separated views based on affine invariant regions [J]. *International Journal of Computer Vision*, 2004, **59**(1): 61~85
- 2 C. Schmid, R. Mohr. Local grayvalue invariants for image retrieval [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, **19**(5): 530~535
- 3 K. Mikolajczyk, C. Schmid. Scale & affine invariant interest point detectors [J]. *International Journal of Computer Vision*, 2004, **60**(1): 63~86
- 4 D. G. Lowe. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, **60**(2): 91~110
- 5 Zhang Ruijuan, Zhang Jianqi, Yang Cui *et al.*. Study on color image registration technique based on CSIFT [J]. *Acta Optica Sinica*, 2008, **28**(11): 2097~2103
张锐娟, 张建奇, 杨 翠 等. 基于 CSIFT 的彩色图像配准技术研究 [J]. *光学学报*, 2008, **28**(11): 2097~2103
- 6 Tian Ying, Yuan Weiqi. Ear recognition based on fusion of scale invariant feature transform and geometric feature [J]. *Acta Optica Sinica*, 2008, **28**(8): 1485~1491
田 莹, 苑玮琦. 尺度不变特征与几何特征融合的人耳识别方法 [J]. *光学学报*, 2008, **28**(8): 1485~1491
- 7 H. Bay, T. Tuytelaars, L. V. Gool SURF: Speeded up robust features [C]. *European Conference on Computer Vision*, 2006, **1**: 404~417
- 8 J. Matas, O. Chum, M. Urban *et al.*. Robust wide baseline stereo from maximally stable extremal regions [C]. In *Proceedings of the British Machine Vision Conference*, 2002, 384~393
- 9 P. E. Forssen, D. G. Lowe. Shape descriptors for maximally stable extremal regions [C]. *International Conference on Computer Vision (ICCV)*, Brazil, 2007, 59~73
- 10 V. Lepetit, P. Fua. Keypoint recognition using randomized trees [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(9): 1465~1479
- 11 M. Ozuysal, P. Fua, V. Lepetit. Fast keypoint recognition in ten lines of code [C]. *Conference on Computer Vision and Pattern Recognition*, 2007, 1~8
- 12 M. Özuysal, M. Calonder, V. Lepetit *et al.*. Fast keypoint recognition using random ferns [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(3): 448~461
- 13 C. Scherrer, J. Pilet, P. Fua *et al.*. The haunted book [C]. *Proceedings of International Symposium on Mixed and Augmented Reality*, 2008, 163~164
- 14 S. Hinterstoisser, O. Kutter, N. Navab *et al.*. Real-time learning of accurate patch rectification [C]. *Conference on Computer Vision and Pattern Recognition*, 2009, 2945~2952
- 15 E. Rosten, T. Drummond. Machine learning for high-speed corner detection [C]. *European Conference on Computer Vision*, 2006, 430~443
- 16 R. Harley, A. Zisserman. *Multiple View Geometry in Computer Vision* [M]. New York: Cambridge University Press, 2003. 290~293
- 17 Liu Guixi, Liu Dongmei, Liu Fengpeng *et al.*. A robust registration algorithm based on feature points matching [J]. *Acta Optica Sinica*, 2008, **28**(3): 454~461
刘贵喜, 刘冬梅, 刘凤鹏 等. 一种稳健的特征点配准算法 [J]. *光学学报*, 2008, **28**(3): 454~461
- 18 O. Chum, J. Matas. Matching with PROSAC-progressive sample consensus [C]. *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, 1:220~226
- 19 Chen Bing, Zhao Yigong, Li Xin. A new approach to scene matching during electro-optical imaging terminal guidance [J]. *Acta Optica Sinica*, 2010, **30**(1): 163~168
陈 冰, 赵亦工, 李 欣. 一种新的光电成像末端制导景象匹配方法 [J]. *光学学报*, 2010, **30**(1): 163~168