

基于 ORB 特征的快速目标检测算法

李小红 谢成明 贾易臻 张国富
(合肥工业大学计算机与信息学院 合肥 230009)

摘要: 在动态场景运动目标检测下提出了一种新颖的快速目标检测算法,针对 SURF 算法不能满足实时性的需要,提出基于 ORB(oriented FAST and rotated BRIEF) 特征的特征点匹配算法,接着采用八参数旋转模型,结合最小二乘法求解全局运动参数进行运动补偿,最后使用帧差法来获得运动目标。在此过程中采用 PROSAC(progressive sample consensus) 算法来去除外点。实验结果表明,该算法不仅保持了 SURF 本身的优越性,而且提高了检测速度,可以实时准确的检测出运动目标。

关键词: 目标检测; 运动补偿; ORB; 旋转模型; PROSAC

中图分类号: TP391 文献标识码: A 国家标准学科分类代码: 520.6040

Rapid moving object detection algorithm based on ORB features

Li Xiaohong Xie Chengming Jia Yizhen Zhang Guofu

(School of Computer Science & Information, Hefei University of Technology, Hefei 230009, China)

Abstract: A novel method for detecting object under dynamic scene is proposed. Aiming at SURF algorithm can not meet the needs of real-time application, a feature point matching algorithm based on the ORB (oriented FAST and rotated BRIEF) algorithm is presented. It uses the rotation model of eight parameters, combines with the least squares method to solve the global motion parameters for motion compensation, and the frame difference method is last used to get a moving target. In this process, PROSAC (progressive sample consensus) algorithm is used to eliminate the outliers generated in the process of feature point matching. The experiments demonstrate that, this method not only remains the advantages of SURF but also improves the detection rate, and can effectively detect moving object in real-time and accurately.

Keywords: moving object detection; motion compensation; ORB; rotational model; PROSAC

1 引言

运动目标检测技术一直是智能视频监控领域的重要研究方向之一,而视频监控中运动目标检测主要分为静态场景和动态场景,静态场景由于摄像机固定,背景相对固定,其检测相对要简单很多,目前已经有了很成熟的算法,这些算法中主要分为三大类: 背景差分法、帧间差分法及光流法。而动态场景由于摄像机的不断运动使得背景不停的在变化,而且运动目标也在运动,从而图像中就

存在这两种运动的叠加,所以在运动的背景中提取运动的目标一直是研究的难点,国内外学者在这方面也做了很多研究^[1-7]。其基本思想都是: 首先为摄像机的运动建立运动模型,通过对运动模型的求解来得到图像的全局运动参数,其次根据得到的参数对图像进行运动补偿,最后利用帧差法实现对运动目标的检测。

常用的运动估计的方法主要有以下几种: 基于图像块匹配的方法^[8-9]、基于空间特征的方法^[10]以及基于变换域的方法^[11]等几种。1) 基于

收稿日期: 2012-40 Received Date: 2012-40

* 基金项目: 国家自然科学基金(61174170, 61004103, 61100127)、教育部博士点基金(20100111110005)、中国博士后科学基金(2010M500111)、安徽省自然科学基金(090412058, 11040606Q44)资助项目

图像块匹配算法，其基本思想是对整幅图像平面进行划分，得到很多互不相交的小方块(blocks)并同时假设块内的像素点具有相同的运动矢量，用块的平移运动来表示相邻两帧间的所有仿射运动。2) 基于空间特征的算法，它是在参考帧和搜索帧中找到匹配的特征结构，对这两个特征结构运用某种特征相似性匹配法则进行匹配，最后根据匹配结果求出运动矢量。3) 基于变换域方法，它的主要思想是通过变换域的转化来实现运动参数的估计，通常是把时空域中的估计转化到频率域来计算。其中块匹配法和特征点匹配法是目前最为常用的两种算法。

基于图像块匹配的方法由于其复杂度小、计算简单而得到很大的应用，但它不适合旋转、缩放或者仿射变换等图像的非平移运动，而基于空间特征的方法中特征点之间的匹配不需假设块内像素运动一致的条件，摆脱了块匹配方法的局限性。因此，便采用基于空间特征的算法。

SURF 算子被公认是既保持了 SIFT 算子的良好性能，又较大的改进了实时性，但还是不能满足实时跟踪的要求。因此本文针对 SURF 算法在速度上的不足，提出了基于 ORB 特征算子的特征点匹配算法，在保持了 SURF 良好性能的前提下，提高匹配速度，确保算法的实时性。

2 ORB 特征算子

近年来局部不变特征已经成为国内外的研究热点^[12]，其中 SIFT 和 SURF 是最著名的两种局部不变特征检测算子。针对 SIFT 和 SURF 在实时性上的缺陷，E Rublee 等人在 2011 年的 ICCV(IEEE International Conference on Computer Vision) 上提出了 ORB^[13] 特征检测算子，在速度上有个极大的提升。ORB(oriented FAST and rotated BRIEF) 是建立在著名的 FAST^[14-15] 特征检测和 2010 年提出的 BRIEF^[16] 特征描述子基础上。

2.1 特征点检测

ORB 采用的是 FAST 算子检测特征点，然后再给检测到的特征点加个方向信息，构成 oFAST，解决了 FAST 算子不带有方向性的严重缺陷。

FAST 检测的角点定义为：在像素点的周围邻域内有足够的像素点与该点处于不同的区域。

考虑图像中任意一个像素点和以它为中心的一个区域，通常情况下，该区域选择圆形区域。如图 1 所示，要判断 P 点是否为角点：将 P 点的灰度值和它邻域内 16 个像素点的灰度值进行比较，若圆圈上存在 n 个连续的像素点的灰度值大于 P 点的灰度值加 t 或者小于 P 点的灰度值减 t(t 是阈值)，则 P 点为角点。

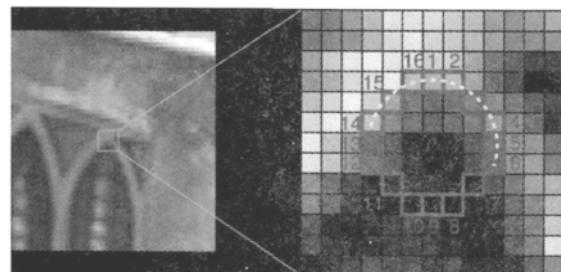


图 1 FAST 角点检测

Fig. 1 FAST corner detection

这里 n 一般选 12、9 等值，实验证明，n 取 9 效果最好，ORB 中采用的是 FAST-9。FAST 本身的缺点是：1) FAST 没有角点响应函数，并且具有很大的边缘响应；2) FAST 不产生多尺度特征。

针对这两个缺陷，ORB 的解决方法是：1) 若想得到 N 个关键点，先降低阈值得到多于 N 个点，再用 Harris 角点检测对其进行排序，取前 N 个点。2) 使用图像中的尺度金字塔，在金字塔每一层得到 FAST 特征。

为了给检测到的角点加个方向信息，ORB 使用了一个简单有效的角点定向方法——灰度质心法，灰度质心法假定角点的灰度和质心之间存在一个偏移，这个向量可以用于表示一个方向。

Rosin 定义邻域矩为：

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y)$$

质心为：

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right)$$

特征点与质心的夹角定义为 FAST 特征点的方向：

$$\theta = \text{atan2}(m_{01}, m_{10})$$

为了提高方法的旋转不变性，需要确保 x 和 y 在半径为 r 的圆形区域内，即 $x, y \in [-r, r]$ ，r 等于邻域半径。

2.2 特征点描述

ORB 中使用 BRIEF 描述子对检测到的特征点进行描述，并解决了 BRIEF 本身不具有旋转不变性的首要缺陷。

BRIEF 描述子计算简单、快速，其基于这样的思想：图像邻域能够用相对少量的强度对比来表达。

定义 $S \times S$ 大小的图像邻域 P 的准则 τ 为：

$$\tau(p; x, y) := \begin{cases} 1, & p(x) < p(y) \\ 0, & \text{其他} \end{cases}$$

式中： $p(x)$ 为平滑后的图像邻域 P 在 $x = (u, v)^T$ 处的像素灰度值。

选择 n_d 个 (x, y) 位置对，唯一的定义了二进制准则，BRIEF 描述子就为 n_d 维的二进制比特串：

$$f_{nd}(p) := \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i, y_i).$$

n_d 可以为 128、256、512 等。选不同值需要在速度、存储效率和识别率之间权衡。

BRIEF 中图像邻域的准则仅考虑单个像素，所以对噪声敏感。为了解决这个缺陷，ORB 中每个测试点采用的是 31×31 像素邻域中的 5×5 子窗口，其中子窗口的选择服从高斯分布，再采用积分图像加速计算。

BRIEF 本身是无向的，不具有旋转不变性，ORB 的解决方法是尝试给 BRIEF 添加一个方向。在位置 (x_i, y_i) 处，对于任意 n 个二进制准则特征集，定义一个 $2n$ 矩阵： $S = \begin{pmatrix} x_1, x_2, \dots, x_n \\ y_1, y_2, \dots, y_n \end{pmatrix}$

使用邻域方向 θ 和对应的旋转矩阵 R_θ ，构建 S 的一个校正版本 S_θ ： $S_\theta = R_\theta S$

于是 Steered BRIEF 描述子就为： $g_n(p, \theta) := f_{nd}(p) | (x_i, y_i) \in S_\theta$ 。

得到 Steered BRIEF 后，接着执行一个贪婪搜索，从所有可能的像素块对中找到 256 个相关性最低的像素块对，即得到最终的 rBRIEF。

3 快速目标检测算法

3.1 求解全局运动参数

多年来人们对于如何选取合适的模型来描述摄像机的运动进行了大量的研究，针对应用场合的不同以及摄像机本身的特点提出了很多的全局

运动模型，按照描述的参数多少可以划分为：六参数仿射模型、八参数旋转模型、四参数模型、三参数模型和二参数平移模型。

由于摄像机有平移、旋转、缩放和倾斜等可能存在的运动，在摄像机旋转很大的情况下，线性的仿射参数模型不能很好的描述摄像机的运动，而且需要对视频序列实时的检测出运动目标，针对这两点我们选择通用性强的八参数旋转模型来描述图像的像素点运动。

设 $t-1$ 帧和 t 帧图像中点 p_{t-1} 和 p_t 对应的坐标为 $X_{t-1} = (x_{t-1}, y_{t-1})$ 和 $X_t = (x_t, y_t)$ 。 $U = (u, v)^T$ 为 p_{t-1} 的运动矢量，则 t 帧 X_t 可由式 $X_t = X_{t-1} + U$ 计算得到。则 U 满足下式：

$$\begin{aligned} U &= C(X_{t-1}) A \\ A &= (a, b, c, d, e, f, g, h)^T \\ C(X_{t-1}) &= \begin{bmatrix} x & y & 1 & 0 & 0 & 0 & x^2 & xy \\ 0 & 0 & 0 & x & y & 1 & xy & y^2 \end{bmatrix} \end{aligned}$$

摄像机仿射参数矩阵 A 主要是由摄像机的旋转和平移参数决定，在实际中摄像机的旋转角度和平移速度都是未知的，因此在求解参数矩阵的时候就需要寻找其他的途径。

在该模型下定义参数矩阵矢量 $A = (a, b, c, d, e, f, g, h)^T$ ，即为所求的八参数旋转参数矩阵。使用特征点对估计摄像机运动参数的思想就是：在相邻两帧中分别搜索特征点，再对特征点对匹配，得到 $F = \{f_1, f_2, \dots, f_n, \dots, f_N\}$ ，表示匹配点对的集合，其中 $f_n = (X_{t-1,n}, X_{t,n})$ 为第 n 对匹配点，并用最小二乘法求最优解。

3.2 PROSAC 算法

在采用特征点对进行运动补偿的方法时，得到的匹配特征点对有可能会有误匹配的情况发生又或者其他各种原因使特征点的数值偏离真实值，称这些点为外点（outliers）。这些点虽然相对于得到的整个特征点来说很少，但是如果在进行全局运动估计的时候引入了这些外点，则会给全局运动估计带来很大的错误。因此必须采用方法去除外点，将混在一起的外点和内点（inliers）从数据集中区分开来，选择正确的内点进行参数的估计。

有效数据占大多数，无效数据只是很少量时，可以通过最小二乘法或类似的方法来确定模型的参数和误差，但是当无效数据很大时最小二乘法

就失效了，因此在这里采用 PROSAC(progressive sample consensus 改进的样本一致性) [17] 算法来去除误匹配点，该方法通过重复迭代然后去除外点的匹配特征点对，采用最小二乘估计求得的参数矩阵就比较接近真实值。

PROSAC 算法是 RANSAC 算法的改进，而且鲁棒性和计算效率比 RANSAC 更高。RANSAC 算法是随机抽取样本，没有考虑到样本之间的差异性。但在实际情况下样本之间存在好坏差异，即有的样本内点概率高，有的样本内点概率低。与 RANSAC 算法不同，PROSAC 算法不是从所有数据中抽取样本，而是事先通过一个线性排序过程，将样本按质量好坏进行一个降序排列，再从具有较高质量的数据子集中抽取。经过若干次的假设——验证后，得到最优解。

在本文中，将 ORB 特征点匹配对的最近邻和次近邻的比值作为匹配质量的定量表示。比值越小，则匹配的质量越好，对匹配质量进行从好到差的排序。采用的 PROSAC 算法的流程图如图 2 所示。

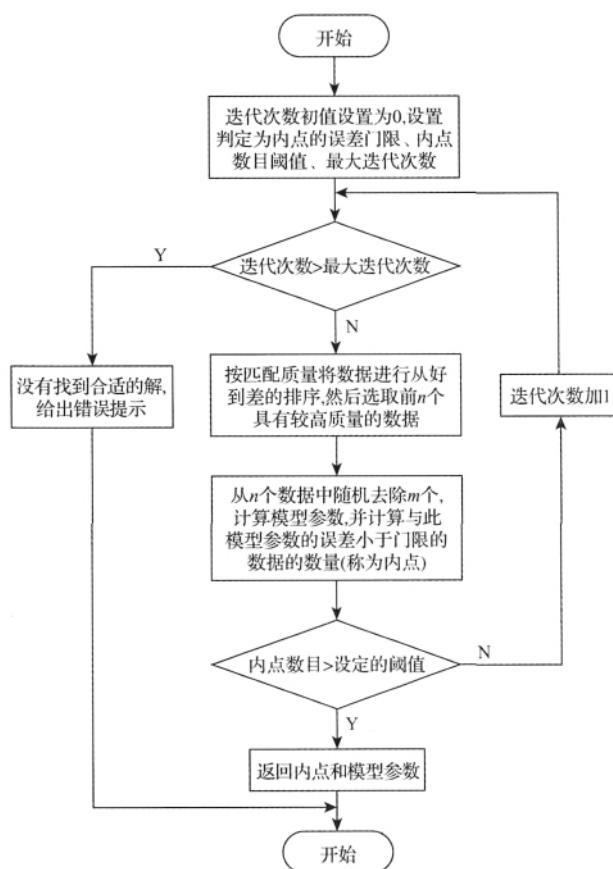


图 2 PROSAC 算法
Fig. 2 PROSAC algorithm

4 实验结果和分析

本节将本算法和传统算法进行了比较，包括：块匹配算法和 SURF 算法，并对结果进行比较和分析。块匹配算法和 SURF 算法是解决动态场景下运动目标检测问题的两个具有代表性的方法。实验平台在 Inter Core2 Duo CPU 内存 2.4 G 的 PC 上使用 VC6.0 调试。在实验中我们采用隔帧处理的方法来提高目标的检测效率。

图 3 是对 MPEG-4 标准测试序列 foreman 的实验结果，图像分辨率为 352×288 。图(a) 为原序列的第 132 帧；图(b) 为块匹配算法得到的检测结果，可以看到背景的干扰没有被很好的去除；图(c) 和图(d) 分别是 SURF 算法和 ORB 算法得到的检测结果，可以看到二者都很好的去除了背景的干扰，而且 ORB 算法比 SURF 算法得到的轮廓更加清晰。此组实验说明：在目标与摄像头存在高速旋转的情况下，本算法可以得到更佳的结果。

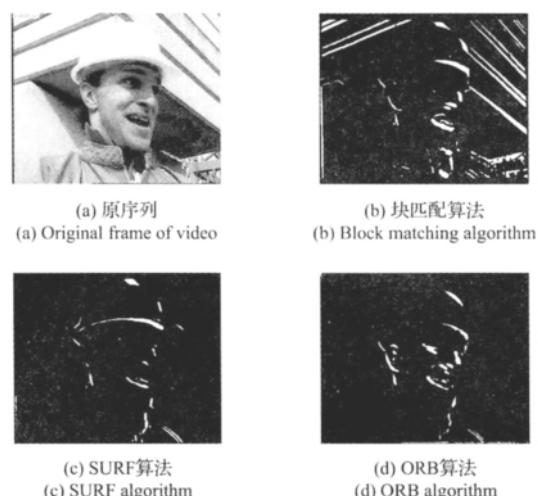


图 3 序列 1 的实验结果

Fig. 3 Experiment result of video sequence 1

图 4 是对 MPEG-4 标准测试序列 coastguard 的实验结果，图像分辨率为 352×288 。图(a) 为原序列的第 28 帧；图(b) 为块匹配算法得到的检测结果，由于该标准序列的摄像机的运动主要是水平方向的平移，因此块匹配能很好的检测出运动；图(c) 为 SURF 算法得到的检测结果；图(d) 为本文 ORB 算法得到的检测结果，比块匹配算法和 SURF 算法得到的结果更加清晰。此组实验说明：除了摄像机旋转速度很大的情况下，在平移运动方面，本

文算法同样适用。

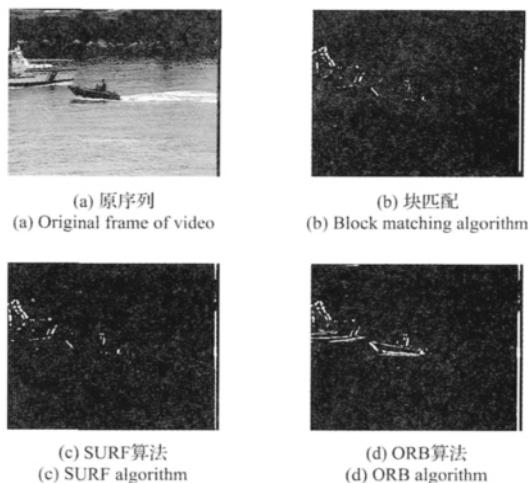


图 4 序列 2 的实验结果

Fig. 4 Eperiment result of video sequence 2

图 5 是对 MPEG-4 标准测试序列 stefan 的实验结果, 图像分辨率为 352×288 。图(a)为原序列的第 24 帧; 图(b)为块匹配算法得到的检测结果, 由于块匹配只能对平移运动的摄像机有很好的效果, 在这组实验结果可以看到还存在背景观众的残影, 而且检测到的目标也不够清晰, 在摄像机其他的运动方式下得到的效果比较差; 图(c)和图(d)分别是 SURF 和本文 ORB 算法得到的检测结果, 二者依然很好的检测出前景目标。此组实验说明: 对于复杂的摄像机运动情况, 块匹配已经失效, 本算法依然有效, 可以很好的检测出目标。

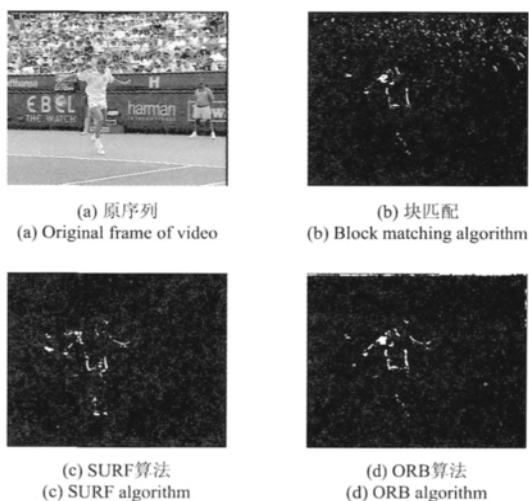


图 5 序列 3 的实验结果

Fig. 5 Experiment result of video sequence 3

表 1 是 3 种算法的执行时间比较。从表中可以看出, 在实时性方面 ORB 算法较块匹配算法和 SURF 有很大的提高。大约是块匹配算法的 3 倍, 是 SURF 算法的 6 倍。而且本算法在摄像机的各种运动形式下均能得到较高的检测准确率, 得到的最终目标也更清晰。

表 1 3 种算法执行时间比较表
Table 1 Comparison of execution time of three algorithms

算法	算法运算时间(ms/帧)		
	序列 1	序列 2	序列 3
块匹配	84.64	88.99	84.76
SURF 算法	178.95	173.1	180.2
ORB 算法	31.6	29.3	30.2

5 结 论

该文提出了一种基于 ORB 特征算子的特征点匹配算法以解决运动摄像机下的目标检测问题。首先为图像的全局运动建立旋转参数模型, 其次通过特征点匹配算法在相邻帧之间建立特征点对, 通过 PROSAC 算法去除外点, 并通过最小二乘法求解旋转参数, 最后用帧差法得出目标。实验结果证明该算法可以实时、准确的检测出复杂场景中的运动目标。

参考文献

- [1] MAHADEVAN V , VASCONCELOS N. Background subtraction in highly dynamic scenes [C]. IEEE Conference on Computer Vision and Pattern Recognition , 2008: 1-6.
- [2] VISENTINI I. On stable dynamic background generation technique using gaussian mixture models for robust object detection. advanced video and signal based surveillance [C]. IEEE International Conference on AVSS08 , 2008: 41-48.
- [3] LIU L , WANG Y J. Robotic dynamic target recognition and tracking based on the monocular vision [C]. Control Conference , 2008: 193-197.
- [4] 逯新华, 史忠科. 基于动态模板匹配的空中运动目标检测与跟踪控制 [J]. 电子测量与仪器学报 , 2010 , 24(10): 935-941.
LU X H , SHI ZH K. Detection and tracking control for air moving target based on dynamic template matching [J]. Journal of Electronic Measurement and Instru-

- ment , 2010 , 24(10) : 935-941.
- [5] 王进花 , 曹洁. 一种基于特征融合的点特征目标跟踪算法 [J]. 电子测量与仪器学报 , 2010 , 24(6) : 537-541.
WANG J H , CAO J. Object tracking algorithm based on the feature fusion points [J]. Journal of Electronic Measurement and Instrument , 2010 , 24(6) : 537-541.
- [6] 李劲菊 , 朱青 , 王耀南. 一种复杂背景下运动目标检测与跟踪算法 [J]. 仪器仪表学报 , 2010 , 31(10) : 2242-2247.
LI J J , ZHU Q , WANG Y N. Moving target detection and tracking algorithm based on complex background [J]. Chinese Journal of Scientific Instrument , 2010 , 31(10) : 2242-2247.
- [7] 周许超 , 屠大维 , 陈勇 , 等. 基于相位相关和差分相乘的动态背景下运动目标检测 [J]. 仪器仪表学报 , 2010 , 31(5) : 980-983.
ZHOU X CH , TU D W , CHEN Y , et al. Moving object detection under dynamic background based on phase-correlation and differential multiplication [J]. Chinese Journal of Scientific Instrument , 2010 , 31 (5) : 980-983.
- [8] HONG H , CHOU Y. Motion estimation for video standards [J]. Journal of VLSI signal processing , 1997(17) : 113-136.
- [9] TAO T F , HAN CH ZH , WU Y Q. Motion estimation based on an improved block matching technique [J]. Chinese Optics Letter , 2006 , 14(4) : 208-210.
- [10] MEULEMANS P , WILSON R. Correcting velocity measurements by tracking of linear features [J]. IEEE Colloquium on Motion Analysis and Tracking , 1999 (10) : 1-6.
- [11] ERDEM C , KARABULUT Z , YANMAZ E. Motion estimation in the frequency domain using fuzzy c-planes clustering [J]. IEEE Transactions on Image Processing , 2001 , 10(12) : 1873-1879.
- [12] 孙浩 , 王程 , 王润生. 局部不变特征综述 [J]. 中国图象图形学报 , 2011(2) : 141-151.
SUN H , WANG CH , WANG R SH. Summarize of local invariant features [J]. Chinese Journal of Scientific Instrument , 2011(2) : 141-151.
- [13] RUBLEE E , RABAUD V , KONOLIGE K , et al. ORB: an efficient alternative to SIFT or SURF [C]. IEEE International Conference on Computer Vision , 2011: 2564-2571.
- [14] ROSTEN E , DRUMMOND T. Fusing points and lines for high performance tracking [C]. IEEE International Conference on Computer Vision , 2005: 1508-1515.
- [15] ROSTEN E , DRUMMOND T. Machine learning for high-speed corner detection [C]. In European Conference on Computer Vision , 2006: 430-443.
- [16] CALONDER M , LEPETIT V , FU P. Brief: binary robust independent elementary features [C]. European Conference on Computer Vision , 2010: 778-792.
- [17] CHUM O , MATAS J. Matching with PROSAC-progressive sample consensus [C]. IEEE Conference on Computer Vision and Pattern Recognition(CVPR) , 2005.

作者简介

李小红 , 1970 年出生 , 现为合肥工业大学计算与信息学院副教授 , 主要研究方向为图像处理与分析、视频图像编码和计算机网络通信等。

E-mail: xh_litong@sina.com

Li Xiaohong was born in 1970 , associate professor in School of Computer and Information , Hefei University of Technology. Her main research interests include image processing and analysis , video image encoding and computer network communication etc.

谢成明 , 1988 年出生 , 2010 年于合肥工业大学获得学士学位 , 现为合肥工业大学在读硕士研究生 , 主要研究方向为目标检测与跟踪。

E-mail: xiechengming1988@163.com

Xie Chengming was born in 1988 , received B. Sc. from Hefei University of Technology in 2010. He is currently a M. Sc. candidate of Hefei University of Technology. His main research direction is target detection and tracking.

贾易臻 , 1989 年出生 , 2011 年于合肥工业大学获得学士学位 , 现为合肥工业大学在读硕士研究生 , 主要研究方法为数字图像分析与处理。

E-mail: 563031490@qq.com

Jia Yizhen was born in 1989 , received B. Sc. from Hefei University of Technology in 2011. He is currently a M. Sc. candidate of Hefei University of Technology. His main research direction is digital image analysis and processing.

张国富 , 1979 年出生 , 现为合肥工业大学计算与信息学院副教授 , 主要研究领域为复杂系统、联盟博弈、群智能等。

E-mail: zgf@hfut.edu.cn

Zhang Guofu was born in 1979 , associate professor in School of Computer and Information , Hefei University of technology. His main research interests include complex systems , coalitional games and swarm intelligence etc.