

# 利用 C\_SURF 配准的空基视频运动目标检测

刘 慧<sup>1,2</sup> 李清泉<sup>3</sup> 高春仙<sup>2</sup> 曾 喆<sup>4</sup>

1 武汉大学电子信息学院,湖北 武汉,430079

2 厦门大学水声通信与海洋信息技术教育部重点实验室,福建 厦门,361005

3 武汉大学测绘遥感信息工程国家重点实验室,湖北 武汉,430079

4 中国石油大学(华东)地球资源与信息学院,山东 青岛,266555

**摘 要:**针对传统的车辆检测算法的性能易受低空移动平台影响造成相机自运动以及外界的干扰等问题,提出了一种基于改进的 C\_SURF 彩色特征稳像和光流法向量相结合的方法来解决低空视频中的运动车辆检测问题。通过图像稳像消除了相机的自运动和外界干扰问题,提高了运动车辆的检测性能。实验结果显示,该方法不仅在检测车辆方面可以获得更好的检测性能,在复杂的背景环境下也能有效地检测运动车辆。

**关键词:**C\_SURF 算子;车辆检测;光流法向量;KLT 算子;视频监控

中图分类号:TN391

文献标志码:A

基于低空视频的监控系统的主要目的是收集某个区域的静止或动态图像,然后人工或机器分析得到感兴趣的信息。动态图像能够提供静态图像不能提供的信息,因此在低空检测中显得更为重要。运动目标检测是低空检测中的一个关键问题。同时,低空平台的移动性也给运动目标检测算法带来了极大挑战。

传统的车辆检测方法,如混合高斯模型法<sup>[1]</sup>、核密度估计法<sup>[2]</sup>、HMM 模型法<sup>[3]</sup>、特征空间分解法<sup>[4]</sup>等,其本质都是基于背景统计模型,以“相机和背景是静止的”为理论假设前提,所以这些方法不能解决背景也是动态变化的运动车辆检测问题。Barron<sup>[5]</sup>针对背景运动的情况提出了光流法,主要利用运动车辆矢量流特征和运动背景矢量流特征的差异来检测运动车辆,但由于算法的复杂度高,抗噪性能较差,影响了光流法在实际中的运用。SASTRY<sup>[6]</sup>等利用基于特征的图像稳像技术将动态背景转化为静态背景,弥补了相机的自运动。由于这些特征都不具备颜色、旋转、尺度不变性,导致图像配准的精度不高,且采用的图像配准算法多是将彩色图像转化为灰度图后再配准,色彩信息的丢失也会引起误配准,从而影响检测精度。尺度、旋转不变的 SIFT 算法<sup>[7]</sup>及其改进算法 CSIFT (colored scale invariant feature

transform)<sup>[8]</sup>和 SURF 算法<sup>[9]</sup>的提出,为上述算法的改进提供了理论基础。

基于此,在空基视频中采用基于改进的 C\_SURF (colored speeded-up robust features) 的彩色图像稳像和光流法向量相结合的技术来检测运动车辆。首先,求出彩色图像各个位置处的颜色不变量,再提取特征点及其描述算子,然后采用最小二乘法求出特征点匹配对<sup>[6]</sup>,对图像进行稳像,最后计算出参考帧和匹配帧的光流法向量<sup>[6]</sup>,并根据法向量的值检测运动车辆。实验结果表明,和其他方法相比,该方法在检测运动车辆方面可以获得更好的检测效率和检测精度;即使在复杂的背景环境下,也能有效地检测运动车辆。该方法对空基视频的运动目标检测有一定的理论和应用价值。

## 1 C\_SURF 算法原理

SURF<sup>[9]</sup>是 Bay 等人在 SIFT<sup>[7]</sup>算法的基础上提出的一种对图像旋转、平移、缩放和噪声具有良好鲁棒性的快速的特征提取算法。SURF 算法在特征点检测部分采用盒子滤波器来构建尺度图像,并利用积分图像简化 Hessian 特征点计算;在特征描述部分,采用一个 64 维向量作为特征量,

收稿日期:2013-05-12

项目来源:国家自然科学基金资助项目(41101355);中央高校基本科研业务费专项资金资助项目(12CX04002A)。

第一作者:刘慧,博士生,主要从事图像处理、机器视觉等研究。E-mail:liuhui\_edu@126.com

通讯作者:李清泉,博士,教授。E-mail:qqli@whu.edu.cn

在不影响检测精度的情况下大大降低了算法的复杂度,其特征检测速度要比 SIFT 算法快 3 倍左右,且对阈值不敏感。为了使 SURF 算法能更好地用于彩色图像处理,本文结合色彩特性和几何结构特征,基于 CSIFT 算法的思想,提出了改进的 C-SURF 算法。

### 1.1 颜色不变量

Kubelka-Munk<sup>[10]</sup>理论描述了物体的光谱辐射特性,其模型表达为:

$$E(\lambda, \mathbf{X}) = e(\lambda, \mathbf{X}) [1 - \rho_f(\mathbf{X})]^2 R_\infty(\lambda, \mathbf{X}) + e(\lambda, \mathbf{X}) \rho_f(\mathbf{X}) \quad (1)$$

式中,  $\lambda$  为波长;  $\mathbf{X}$  为二维矢量  $(x, y)$ , 表示观测位置;  $e(\lambda, \mathbf{X})$  为光谱强度;  $\rho_f(\mathbf{X})$  表示  $\mathbf{X}$  处 Fresnel 反射系数;  $R_\infty(\lambda, \mathbf{X})$  表示材料反射率;  $E(\lambda, \mathbf{X})$  为观测处的反射谱。多数情况下,  $e(\lambda, \mathbf{X})$  在各波长上保持不变, 与位置有关, 将  $e(\lambda, \mathbf{X})$  改写成  $i(\mathbf{X})$  形式, 则式(1)变为:

$$E(\lambda, \mathbf{X}) = i(\mathbf{X}) \{ [1 - \rho_f(\mathbf{X})]^2 R_\infty(\lambda, \mathbf{X}) + \rho_f(\mathbf{X}) \} \quad (2)$$

对  $\lambda$  分别求一阶导数和二阶导数, 然后相除得:

$$H = \frac{E_\lambda}{E_\lambda} = \frac{\partial E / \partial \lambda}{\partial^2 E / \partial \lambda^2} = \frac{\partial R_\infty(\lambda, \mathbf{X}) / \partial \lambda}{\partial^2 R_\infty(\lambda, \mathbf{X}) / \partial \lambda^2} = f[R_\infty(\lambda, \mathbf{X})] \quad (3)$$

式中,  $H$  为颜色不变量的一种表述, 同观察位置、表面朝向、光强大小、反射系数等均无关。在符合人眼视觉系统和 CIE-1964-XYZ 标准的条件下, 彩色图像的  $R, G, B$  分量和  $(E, E_\lambda, E_\lambda)$  的关系近似为:

$$\begin{bmatrix} E \\ E_\lambda \\ E_\lambda \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.6 & 0.17 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

通过式(3)和式(4), 可以将彩色图像矩阵转换为彩色图像的颜色不变量矩阵。

### 1.2 尺度不变性

对参考图像  $I_1(x, y)$  和待配准图像  $I_2(x, y)$ , 分别求出颜色不变量图像  $H_1(x, y)$  和  $H_2(x, y)$ 。以  $I_1, I_2$  作为 SURF 特征点的提取, 以  $H_1, H_2$  作为 SURF 特征点的描述, 并将提取的特征信息分别作为  $I_1, I_2$  的特征信息在彩色图像上进行标识。

SURF 算法在图像的尺度空间中使用 Hessian 矩阵进行极值检测, 并将该极值作为候选特征点。图像  $I$  的尺度空间定义为:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (5)$$

式中,  $\sigma$  为尺度因子;  $G(x, y, \sigma)$  是二维高斯函数;  $*$  代表二维卷积;  $L(x, y, \sigma)$  表示不同尺度空间下

的图像。

图像  $I$  的点在尺度  $\sigma$  处的 Hessian 矩阵定义为:

$$\mathbf{H}(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (6)$$

式中,  $L_{xx}, L_{xy}, L_{yy}$  代表拉普拉斯高斯算子。SURF 算法采用加了权重的盒子滤波近似代替  $L_{xx}, L_{xy}, L_{yy}$ 。最后得到  $\mathbf{H}$  矩阵的判别式为:

$$\text{Det}(\mathbf{H}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (7)$$

求取 Hessian 矩阵行列式的相应图像后, 在相应图像上采用 3D 非最大抑制算法, 可以求取不同尺度的特征点。

#### 1.2.1 尺度空间的建立

尺度空间的建立与 SIFT 算法类似, 也是由若干个 octave 组成, 每个 octave 由若干个 intervals 组成。和 SIFT 算法不同的是, SURF 算法是原始图像保持不变, 只改变盒子滤波器的大小。因此这种算法允许不同尺度空间的图像同时处理, 不需对图像进行二次抽样。

#### 1.2.2 $D_{xx}, D_{xy}, D_{yy}$ 的计算

采用积分图像计算下式:

$$\begin{cases} D_{xx} = (r - 2 \times b, c - l - b, r + 2 \times b, c + l + b) - 3 \times (r - 2 \times b, c - b, r + 2 \times b, c + b) \\ D_{xy} = (r - l - b, c - 2 \times b, r + l + b, c + 2 \times b) - 3 \times (r - b, c - 2 \times b, r + b, c + 2 \times b) \\ D_{yy} = (r - l, c - l, r - 1, c - 1) + (r + 1, c + 1, r + l, c + l) - (r - l, c + 1, r - 1, c + l) - (r + 1, c - l, r + l, c - 1) \end{cases} \quad (8)$$

式中,  $(r, c)$  为像素点的行和列;  $l$  为叶子长度即滤波器大小的  $1/3$ ;  $b = \lceil l/2 \rceil$  ( $\lceil \cdot \rceil$  为取整操作)。

### 1.3 旋转不变性

1) 主方向分配。在检测完特征点后, 首先将特征点领域由  $20 \times 20$  分成  $4 \times 4$  小块, 每个小块为  $5 \times 5$  大小, 同时采用积分图像与 Harr 小波进行卷积得到该小块内像素点的  $dx$  和  $dy$ , 然后获得每一个像素点的角度  $\theta = \arctan(dy/dx)$ , 对这些角度以  $\pi/3$  为窗口进行统计, 统计出来最多的角度值即为该特征点的主方向。

2) 特征描述算子形成。区间划分同主方向分配, 计算 25 个像素点在 Harr 小波于  $x$  和  $y$  上的响应值  $dx$  和  $dy$ , 形成 4 维向量  $\sum dx, \sum |dx|, \sum dy$  和  $\sum |dy|$ , 从而形成  $4 \times 4 \times 4 =$

64 的描述算子。

2 实验结果及分析

2.1 实验数据和仿真平台

实验以飞艇为数据采集平台,飞艇上安装了一个 720 像素×576 像素分辨率的摄像机,在距离地面约 120 m 的空中进行拍摄,飞艇速度约为 30 km/h,摄像机拍摄速度为 25 帧/s。图 1(d)、1(e)显示了两段视频的序列影像,分别为 2. avi 和 gs. avi。图 1(f)是 Anand 文中采用的的视频 TucsonBlvd\_original. avi。从车辆的数目和背景复杂度来看,视频 2 中含有的车辆最少,且背景简单,不含有建筑物的信息,因此相对而言,视觉误差很小<sup>[11]</sup>;视频 gs 中含有的车辆增多,且背景复杂,存在建筑物信息带来的视觉误差;视频 TucsonBlvd\_original 不仅车辆多,背景复杂,相对前两个视频而言,分辨率较低。仿真的硬件平台®<sup>TM</sup>为:CPU 2.1 GHz, RAM 2 GB,仿真软件为 VC++ 6.0。

2.2 SURF 与 CSURF 算法性能分析

表 1 是 C\_SURF 和 CSIFT、SURF 的算法检测性能比较。从表 1 中可以看出,与 SURF 算法相比,由于 C\_SURF 算法考虑了颜色信息,因此检测出来的特征点数量较少,其匹配的特征点数量也较少,而基于 C\_SURF、SURF、CSIFT 计算的两帧图像之间的仿射变换参数没有太大的差别,但是 C\_SURF 算法的计算效率是最优的,其速度比 SURF 算法快了近 1 s,比 CSIFT 算法快了近 3 s。因此本文提出的 C\_SURF 算法在不改变匹配精度的情况下,效率是最优的。

表 1 C\_SURF、CSIFT 和 SURF 算法的性能对比

Tab. 1 Performance Comparison of C\_SURF, CSIFT and SURF

| 匹配特征点数量 |     | 仿射变换参数  |  |  | 检测时间/s |
|---------|-----|---|--|--|--------|
| C_SURF  | 235 | $H = \begin{bmatrix} 1.000\ 6 & -0.000\ 6 & 2.151\ 2 \\ -0.000\ 8 & 0.999\ 8 & -12.052\ 5 \\ 0 & 0 & 1 \end{bmatrix}$ |  |  | 3.012  |
|         |     |   |  |  |        |
| CSIFT   | 295 | $H = \begin{bmatrix} 0.996\ 5 & -0.001\ 9 & 2.939\ 9 \\ -0.003\ 2 & 0.995\ 8 & -11.197\ 6 \\ 0 & 0 & 1 \end{bmatrix}$ |  |  | 6.202  |
|         |     |   |  |  |        |
| SURF    | 363 | $H = \begin{bmatrix} 1.001\ 0 & -0.000\ 9 & 2.134\ 8 \\ -0.000\ 3 & 0.998\ 1 & -11.828\ 6 \\ 0 & 0 & 1 \end{bmatrix}$ |  |  | 4.02   |
|         |     |   |  |  |        |

2.3 实验结果分析

图 2 是本文算法和 Anand C 提出的基于改进的 KLT 算法检测结果比较。其中图 2(a)中红

色的矩形框是运动车辆的检测结果,而图 2(b)中红色椭圆框中圈出的彩色如黄色、绿色等框代表运动车辆检测结果。从图 2 中可以看出,即使在图像分辨率低、背景复杂性高且运行车辆较多的情况下,相比较 Anand C 提出的算法,本文方法也能较好地检测出运动车辆。

图 1 包含了采用本文提出的算法检测 3 段不同复杂度,不同分辨率的视频。从图 1 中可以

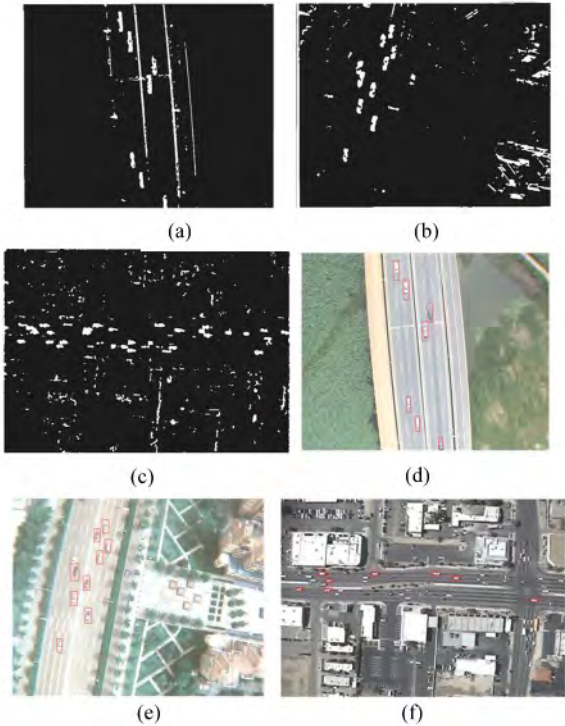


图 1 本文方法检测结果

Fig. 1 Vehicle Detect Results of the Proposed Method



图 2 本文方法和 Anand C 提出的算法比较

Fig. 2 Vehicle Detect Results of the Methods Proposed by Anand C and the Paper

看出,本文方法对不同图像分辨率、不同背景复杂度的运动车辆检测都是有效的。对于图像分辨率高、背景复杂度低、运动车辆较少的 2. avi 视频(图 1(a)、1(d))能够检测出全部的运动车辆;对于图像分辨率较低、背景复杂度较高、运动车辆较多的 gs. avi 视频(图 1(b)、1(e))基本能够检测出全部的运动车辆;对于图像分辨率低、背景复杂度较高、运动车辆多的 TucsonBlvd\_original. avi 视频(图 1(c)、1(f)),能够检测大部分的运动车辆。

表 2 是各种方法的检测性能对比,实验  $DR = TP/(TP+FN)^{[12]}$  (TP 表示正确检测出来的车

辆;FN 表示漏检测车辆)代表检测率; $FPR = FP/(TP+FP)$  (FP 表示错误检测出来的车辆)代表错误检测率。实验均对 25 帧检测数据进行了数学统计。从表 2 中可以看出,随着图像分辨率的降低,背景运动复杂度的增加以及运动车辆的增多,3 种算法的正确检测率在下降,错误检测率在上升。而相比较 GMM、LK 光流法,本文方法检测率较高而错误检测率较低,虽然由于算法的复杂性,导致了计算速度的降低,但是仍然能够快速地进行运动车辆检测。

表 2 检测性能对比

Tab. 2 Performance Comparison

| 检测方法 | DR/(%) |         |                          | FPR/(%) |         |                          | 速度/(s/帧) |
|------|--------|---------|--------------------------|---------|---------|--------------------------|----------|
|      | 2. avi | gs. avi | tucsonBlvd_original. avi | 2. avi  | gs. avi | tucsonBlvd_original. avi |          |
| 本文方法 | 94     | 58      | 24                       | 6       | 18      | 33                       | 1.135    |
| GMM  | 29     | 26      | 1                        | 45      | 75      | 79                       | 0.8      |
| LK 法 | 20     | 9       | 1                        | 91      | 80      | 93                       | 0.3      |

### 3 结 论

本文提出了一种基于 C\_SURF 提取特征的彩色图像配准算法和光流法向量相结合的算法,用于检测低空视频的运动车辆。基于彩色图像配准方法,充分利用图像的色彩信息,弥补了传统的基于灰度处理算法的不足,并将动态背景下运动车辆检测转换到静态背景下,同时结合高效的光流法向量算法检测运动车辆。实验验证了该方法的可行性和有效性,为空基视频的运动车辆检测研究提供一种新的解决途径,也为该领域的研究提供了一种可借鉴的方法。如何进一步提高该方法的效率是将来研究需要考虑的问题,可在特征选取和特征配准上做进一步探究。

### 参 考 文 献

- [1] Grimson W E L, Stauffer C, Romano R, et al. Using Adaptive Tracking to Classify and Monitor Activities in a Site[C]. The IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CA, USA, 1998
- [2] Jin Ting, Zhou Fugen, Bai Xiangzhi. Moving Object Detection in Airborne Video Using Kernel Density Estimation[J]. *Infrared and Laser Engineering*, 2011, 40(1): 153-158 (金挺, 周付根, 白相志. 利用核密度估计的空基视频运动目标检测[J]. 红外激光工程, 2011, 40(1): 153-158)
- [3] Rittscher J, Kato J, Joga S, et al. Probabilistic

- Background Model for Tracking[C]. Conference on Computer Vision(ECCV), Dublin, Ireland, 2000
- [4] Oliver N, Rosario B, Pentland A. A Bayesian Computer Vision System for Modeling Human Interactions[J]. *IEEE Trans. Patt. Analy. Mach. Intell.*, 2000, 22(8): 831-843
- [5] Barron J, Fleet D, Beauchemin S. Performance of Optical Flow Techniques[J]. *International Journal of Computer Vision*, 1994, 12(1): 43-77
- [6] Liu Hui, Li Qingquan, Zeng Zhe, et al. Vehicle Detection in Low Altitude Aircraft Video[J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(3): 316-320 (刘慧, 李清泉, 曾喆, 等. 利用低空视频检测道路车辆[J]. 武汉大学学报·信息科学版, 2011, 36(3): 316-320)
- [7] Lowe D G. Distinctive Image Features from Scale-invariant Keypoint [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110
- [8] Alaa E, Abdel Hakim, Aly A F. CSIFT: A SIFT Descriptor with Color Invariant Characteristics[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, USA, 2006
- [9] Bay H, Ess A, Tuytelaars, T, et al. Speeded-up Robust Features (SURF)[J]. *Computer Vision and Image Understanding*, 2008, 110(3): 346-359
- [10] Zhang Ruijuan, Zhang Jianqi, Yang Cui, et al. Study on Color Image Registration Technique Based on CSIFT[J]. *Acta Optica Sinica*, 2008, 28(11): 2 097-2 103 (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):2 097-2 103) 2009
- [11] Yu Q, Medioni G. Motion Pattern Interpretation and Detection for Tracking Moving Vehicles in Airborne Video[C]. IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009
- [12] Cao Xianbin, Lan Jinhe, Yan Pingkun, et al. Vehicle Detection and Tracking in Airborne Videos by Multi-motion Layer Analysis[J]. *Machine Vision and Applications*, 2011, 3: 921-935

## Moving Target Detection Using C\_SURF Registration

LIU Hui<sup>1,2</sup> LI Qingquan<sup>3</sup> GAO Chunxian<sup>2</sup> ZENG Zhe<sup>4</sup>

1 School of Electronic Information, Wuhan University, Wuhan 430079, China

2 Key Laboratory of Underwater Acoustic Communication and Marine Information Technology, Ministry of Education, Xiamen University, Xiamen 361005, China

3 State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

4 College of Geo-resources and Information, China University of Petroleum, Qingdao 266555, China

**Abstract:** Due to the high mobility, rapid deployment and a wide range of monitoring, vehicle detection and tracking system based on low-level mobile platform attract more and more attention. Cameral self-motion, outside interference and other reasons caused by low altitude mobile platforms impact the performance of traditional vehicle detection algorithms. To resolve the above problems, a new method on improved SURF color image stabilization is presented in this paper. From the experimental results, we can see, firstly, compared to other methods, the method proposed by the paper can achieve vehicle detection performance; secondly even in a complex background environment, the method in this paper can effectively detect moving vehicles.

**Key words:** C\_SURF; vehicle detection; normal vector of optical flow; KLT operator; video surveillance

**First author:** LIU hui, PhD candidate, specializes in the image processing and machine vision. E-mail: liuhui\_edu@126.com

**Corresponding author:** LI Qingquan, professor. E-mail: qqli@whu.edu.cn

**Foundation support:** The National Natural Science Foundation of China, No. 41101355; the Fundamental Research Funds for the Central Universities, No. 12CX04002A.