

基于光流场分割和 Canny 边缘提取 融合算法的运动目标检测

张泽旭¹, 李金宗¹, 李宁宁²

(1. 哈尔滨工业大学电子与通信工程系, 黑龙江哈尔滨 150001; 2. 北京空间飞行器总体设计部, 北京 100087)

摘要: 在摄像机运动的情况下, 提出了一种基于光流场分割和 Canny 边缘算子融合技术的运动目标检测方法. 这种方法可分为三步: 第一步利用运动的内极线约束和 C-均值聚类算法完成目标区域的分割, 并获得分割图; 第二步在分割图中利用 Canny 边缘算子获得细化的目标区域边缘图; 第三步根据光流场中的流速值完成分割图和边缘图的融合, 并检测出完整的运动目标. 实验表明, 这种方法可以有效地从复杂自然场景的图像序列中检测出完整的运动目标.

关键词: C-均值聚类; 边缘融合; 内极线约束; 延伸焦点; 光流

中图分类号: TN911.73 **文献标识码:** A **文章编号:** 0372-2112 (2003) 09-1299-04

Detection of Moving Object Using a Fusion Method Based on Segmentation of Optical Flow Field and Edge Extracted by Canny's Operator

ZHANG Ze-xu¹, LI Jin-zong¹, LI Ning-ning²

(1. Dept of Electronic & Communication Engineering, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China;

2. Beijing Spacecraft General Design Department, Beijing 100087, China)

Abstract: We proposed a method on detection of moving object using a fusion algorithm based on segmentation of optical flow field and edge extracted by Canny's operator in the image sequence acquired by a moving camera. This method can be performed with three steps. In the first step, the segmentation of object area was finished and segmentation map was obtained using motion epipolar constraint and C-Mean value cluster algorithm. In the second step, edge map of object area was established by Canny's edge operator. In the third step, a whole object was detected by a novel fusion method based on segmentation map and edge map. The experiments have shown encouraging results using two sequences of natural images with complex backgrounds.

Key words: C-Mean cluster; edge fusion; epipolar constraint; focus of expansion; optical flow

1 引言

运动目标检测一直是机器视觉、图像理解和精确制导领域非常重要的研究内容. 在摄像机运动的条件下, 尤其当场景非常复杂时, 仅仅依靠单一的检测算法很难检测出完整的运动目标. 对于多个运动目标的情况, 运动目标的检测变得更为复杂. Thompson 等人利用运动目标的光流方向与运动内极线约束确定的背景光流方向的差异来检测运动目标^[1], 但是在较为复杂的自然背景中, 仅利用内极线约束很难获得完整的运动目标. Sasa 等人提出了一种利用光流场运动信息完成目标分割的方法^[2], 但是仅应用到简单背景和静止摄像机的情况下. Smith 等人在 Asset-2 系统的实现中, 利用基于特征的光流场来实现目标的聚类^[3], 很明显, 稀疏的光流场尽管提高了运算速度, 但很难得到完整的运动目标. Adiv 通过利用仿射变换的六个参数完成对多个运动目标的光流场分割^[4], 但这种分割的计算花费是相当大的.

本文在摄像机运动的情况下, 提出了基于光流场分割和

Canny 边缘检测算子的融合算法来完成运动目标的检测. 这种方法包括光流场分割、Canny 边缘提取以及分割图和边缘图融合等三个步骤, 其原理框图如图 1 所示.

2 光流场的计算

我们利用 Nagel 提出的光流算法^[5,6]来计算流速(帧间位移矢量). 这种方法认为在垂直于灰度变化方向上的流速变化率应尽量的小, 即在光流基本方程的基础上施加“方向平滑约束”条件, 因而在一定程度上缓解光流计算中对图像边缘的过度平滑^[7]. 定义相邻两时刻图像帧为 $E(x, y, t_1)$ 和 $E(x, y, t_2)$, 下文缩写为 $E_1(x)$ 和 $E_2(x)$, 其中 $x = (x, y)$, 考虑光流基本方程

$$(\nabla E)^T \cdot u + E_t = 0 \quad (1)$$

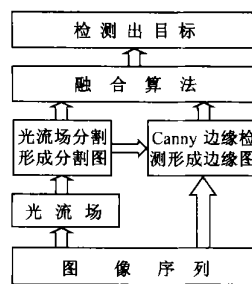


图1 运动目标融合检测
算法原理框图

$\nabla E = (E_x, E_y)^T$, E_x, E_y, E_t 为在点 (x, y) 处沿 x, y, t 方向的偏导数, $u = (u, v)^T$, u, v 分别是在 $t = t_2 - t_1$ 时间间隔内沿 x, y 方向的位移. 根据“方向平滑约束”, Nagel 提出如下的对 u, v 的附加约束条件 $\text{trace}((\nabla u)^T W (\nabla u)) = \min$ (2) 其中 $W = F / \text{trace}(F)$ 是加权矩阵, 它根据视频所含灰度值的空间变化情况, 来补偿运动场中的变化, 其中 F 为

$$F = \begin{bmatrix} E1_y & -E1_x \\ -E1_x & E1_y \end{bmatrix} \begin{bmatrix} E1_y & -E1_x \\ -E1_x & E1_y \end{bmatrix}^T + b^2 \begin{bmatrix} E1_{yy} & -E1_{xy} \\ -E1_{xy} & E1_{xx} \end{bmatrix} \begin{bmatrix} E1_{yy} & -E1_{xy} \\ -E1_{xy} & E1_{xx} \end{bmatrix}^T \quad (3)$$

综合式(1)、(3), 光流场的计算归结为求如下的变分问题:

$$dxdy((E2(x) - E1(x - u))^2 + \text{trace}((\nabla u)^T W (\nabla u))) = \min \quad (4)$$

其中 α 为控制方向平滑程度的权值. 通过求解欧拉方程, 可以得到图像上每一点处的松弛迭代形式的解.

3 光流场的分割

3.1 运动内极线约束

假设一个摄像机相对一个固定场景运动, 并通过透视投影将景物成像到图像平面上. 如果坐标系固定到摄像机上, 那么可以认为场景是相对于摄像机运动的, 景物的运动可以用图像平面的流速来描述, 此时速度是物体表面投影到图像平面的像素坐标、摄像机相对于物体表面的运动和摄像机与物体表面距离的函数^[8]

$$u(x) = u_t(x) + u_r(x) \quad (5)$$

$$u_t(x) = [(-U + xW) / r(x), (-V + yW) / r(x)]^T \quad (6)$$

$$u_r(x) = [Axy - B(x^2 + 1) + Cy, A(y^2 + 1) - Bxy - Cx]^T \quad (7)$$

式中, $u(x)$ 是在图像像素坐标 $x = (x, y)$ 处的流速, x, y 已被焦距归一化, $u_t(x)$ 是平移分量, 其中 $r(x)$ 是 x 所对应的场景中的点的第三维坐标, 而 $u_r(x)$ 为旋转分量; $T = (U, V, W)^T$ 和 $C = (A, B, C)^T$ 分别为摄影机的空间三维平移速度和旋转角速度.

如果摄像机相对于场景只有平移运动而没有旋转运动, 那么场景将产生一种独特的光流场形态, 即场景投影到象平面上像素点的运动好像是从图像平面的一个固定点沿延伸直线产生出来, 这一点为延伸焦点 FOE(Focus Of Expansion). 我们将这种由 FOE 所确定的光流场的运动形态称为运动的内极线约束. 从方程(5)可以得到 FOE 的位置

$$(x, y)_{FOE} = (U/W, V/W) \quad (8)$$

可见延伸焦点的位置取决于平移的方向, 而不是速度大小. 因此, 在场景中相对静止的图像像素 (x, y) 处的光流方向可以根据摄像机运动的方向由运动内极线约束确定, 即

$$FOE = \tan^{-1}[(V - W_y) / (U - W_x)] \quad (9)$$

当 $W = 0$ 时, 延伸焦点在图像坐标的远处, 此时对应的光流场是平行的. 而图像中光流与 FOE 有很大差异的像素处对应着运动的目标区域, 由此, 可以检测出运动目标.

当摄像机有旋转运动时, 情况变得更为复杂. 因为 $u_r(x)$ 仅同旋转参数有关, 而与图像的形状属性无关, 所以可以通过

对 $u_r(x)$ 的估计来预测 $u_t(x)$. 在图像平面上的每一像素处, 可以由观察到流速中减去 $u_r(x)$, 从而得到流速的平移成分 $u_t(x)$. 由 $u_t(x)$ 所建立的流场满足运动的内极线约束条件, 因而可以由该约束确定出运动的目标区域. 但是由于光流的计算误差和对摄像机旋转参数的估计误差, 仅利用内极线约束很难完整地确定出运动目标区域, 还需要对利用内极线约束获得的光流场进行动态聚类分割.

3.2 基于 C-均值聚类算法的矢量场分割

C-均值聚类算法是基于误差平方和准则的动态聚类算法^[9]. 误差平方和聚类准则函数, 定义为

$$J_c = \sum_{j=1}^c \sum_{k=1}^{n_j} \|x_k - m_j\|^2 \quad (10)$$

$$m_j = \frac{1}{n_j} \sum_{k=1}^{n_j} x_k, j = 1, 2, \dots, c \quad (11)$$

其中 x_k 为混合样本集 $X = \{x_1, x_2, \dots, x_n\}$ 中的样本, 该样本集被聚合成 c 个分离的子集, 分别包含 n_1, \dots, n_c 个样本. m_j 是第 j 个子集中样本的均值.

实验中, 我们定义的样本为光流场像素的坐标, 误差平方和准则为欧氏距离准则, C-均值聚类算法在初始划分的基础上, 使用迭代算法, 逐步优化聚类结果, 使准则函数 J_c 达到极小值, 获得 c 个类型. 然后, 比较各个类型里的样本数目, 若样本数目过少, 则认为是虚警而被消除掉; 若还存在多个类型, 认为存在多个运动目标, 实验中我们完成了对单个和多个运动的目标区域的分割, 获得很好的分割图.

4 Canny 边缘检测

Canny 边缘检测算子充分反映了最优边缘检测器的数学特性, 对于不同类型的边缘, 均具有良好的信噪比, 优异的定位性能、对单一边缘产生多个响应的低概率性和对虚假边缘响应的最大抑制能力^[10]. 这些性能对具有复杂背景的自然图像的边缘提取有很好的效果.

对于二维图像, Canny 认为阶跃型边缘的最优边缘检测器的形状与高斯函数的一阶导数类似, 利用二维高斯函数的圆对称性和可分解性, 可以计算高斯函数在任意方向上的方向导数与图像的卷积. 设二维高斯函数

$$G(x, y) = \frac{1}{2\pi} \exp\left[-\frac{x^2 + y^2}{2}\right] \quad (12)$$

在某一方向上 $G(x, y)$ 的一阶导数为

$$G_n = \frac{\partial G}{\partial n} = n \cdot \nabla G \quad (13)$$

式中, $n = [\cos \theta, \sin \theta]^T$ 是单位方向向量; $\nabla G = [G_x, G_y]^T$ 是梯度矢量. 我们将图像 $f(x, y)$ 与 G_n 作卷积, 同时改变 n 的方向, 当 $\partial(G_n * f(x, y)) / \partial n = 0$ 时, 可求解出当 $G_n * f(x, y)$ 取得最大值时的 n

$$n = \frac{\nabla G * f(x, y)}{\nabla G * f(x, y)} \quad (14)$$

很显然 n 正交于检测边缘的方向, 即在该方向上, $G_n * f(x, y)$ 有最大输出响应

$$|G_n * f(x, y)| = |\nabla G * f(x, y)| \quad (15)$$

在实际应用中, 式(12)中的 G 影响原始模板截断成有限的尺

寸 $N = b\sqrt{2} + 1$ 的大小,其中 b 为权值。

利用分割算法获得光流分割场后,在这些分割区域里包含所有我们所关心的运动目标。因此,将在这些分割区域里利用 Canny 算子提取边缘,一方面可以大大限制背景干扰,另一方面可以有效地提高运行的速度。这样在分割图的基础上,我们得到可靠运动目标区域的边缘图,假定数目为 n ,边缘像素集用 C 表示,即 $C = (c_1, c_2, \dots, c_n)$,样本为像素坐标。

5 融合算法

通过光流场的分割,可得运动区域的分割图,但由于光流的计算误差以及利用运动内极线约束分割时的估计误差,得到的分割图很难对运动目标进行精确的定位。利用 Canny 边缘检测算子可以得到较精确的目标边缘图,但是结果与门限的选择有关,若门限较低,则边缘图可能包含较多的复杂背景干扰,为了抑制背景干扰,应该选择高门限,而门限较高又会损失目标弱的边缘。因此,我们在光流场分割的结果中进一步提取目标的边缘,并且与 Canny 边缘图融合,以便得到较完整的目标区域。这个边缘融合算法是建立在像素级上的。

假设通过光流场的分割所得到的运动区域分割图中共有 m 个像素,所有像素点坐标的集合为 $X_s = (x_1, x_2, \dots, x_m)$ 。分割区域里的各点均有流速,则有 m 个流速矢量 $u_i = (u_i, v_i)^T$, $i = 1, 2, \dots, m$ 。令

$$f_i = \sqrt{(u_i^2 + v_i^2)}, i = 1, 2, \dots, m \quad (16)$$

可以构成流速模值的混合样本集 $F = \{f_1, f_2, \dots, f_m\}$ 。

在光流分割图内,利用 Canny 边缘检测算子可以得到边缘像素集 $C = \{c_1, c_2, \dots, c_n\}$,由于选择较高的门限,对于干扰有很大抑制,同时也损失了部分边缘像素。在边缘像素集 $C = (c_1, c_2, \dots, c_n)$ 的基础上,执行 C-均值聚类算法,可以将其分成 1 和 2 两类: 1 为目标边缘类,样本集为 $X_t = (x_1^{(t)}, x_2^{(t)}, \dots, x_{k_1}^{(t)})$,样本数为 k_1 ; 2 为背景边缘类,样本集为 $X_b = (x_1^{(b)}, x_2^{(b)}, \dots, x_{k_2}^{(b)})$,样本数为 k_2 , $k_1 + k_2 = n$ 。显见, $n < m$, 并且 $X_t \subset X_s$, $X_b \subset X_s$ 且 $X_t \cap X_b = \emptyset$ (17)

由于样本集 X_t 包含的边缘像素不够完整,希望从样本集 X_s 中提取部分边缘像素进行补充。为此,在混合样本集 F 中,可以确定与 X_t 相对应的流速模值的边缘子集 $F_t = (f_1^{(t)}, f_2^{(t)}, \dots, f_{k_1}^{(t)})$,该集合反映了目标边缘像素处流速的大小,因此,可以从中获得边缘流速的门限 thresh ,一种可以选择的门限值为

$$\text{thresh} = (f_1^{(t)} + f_2^{(t)} + \dots + f_{k_1}^{(t)}) / k_1 + \quad (18)$$

另一种选择的方法是将 F_t 中样本的最小值作为门限,即

$$\text{thresh} = \min(f_1^{(t)}, f_2^{(t)}, \dots, f_{k_1}^{(t)}) / \quad (19)$$

为微调参数。实验表明,选用式(18)比较合适。这样,对样本集 F 中的样本可以根据 thresh 分成两类,即

$$\begin{cases} f_i & F_e, & f_i > \text{thresh} \\ f_i & \text{otherwise}, & f_i < \text{thresh} \end{cases}, 1 \leq i \leq m \quad (20)$$

显然,样本子集 $F_e = (f_1^{(e)}, f_2^{(e)}, \dots, f_{m_1}^{(e)})$ 的每个样本都是边缘像素的流速模值,样本数为 m_1 ,另一类为其他背景像素的流速模值,样本数为 m_2 , $m_1 + m_2 = m$ 。根据 F_e ,不难得到相对应的像素坐标集 $X_e = (X_1^{(e)}, X_2^{(e)}, \dots, X_{m_1}^{(e)})$ 。

理论与实验分析表明, X_e 包含的目标边缘同光流场的精度有很大的关系,对由本文的光流算法计算的光流场, X_e 即包含大部分强的目标边缘,也很好恢复了弱的目标边缘。而 X_t 仅包含比较强的目标边缘。这样,根据 X_e 和 X_t ,使有关的像素融合,可以得到完整的目标区域的边缘集 X_o ,融合算法主要操作为

$$X_o = X_t \cup X_e \quad (21)$$

6 仿真实验

6.1 单个运动目标检测

图 2 是图像中只包含单个运动目标的实验结果,其中 (a) 为一幅从自然图像序列中抽取的 256×256 的灰度图像(第 15 帧),图像序列中背景是斑驳的操场,一辆微型车从图像的左上角向右下方移动,为追踪目标,摄像机也缓慢地移动; (b) 是在原图像上给出的基于光流的分割区域图,白色曲线形成的闭合区域为分割区域,这个区域是根据第 15 帧和 16 帧图像计算的光流场(图 2(e))及其分割算法得到的光流分割场(图 2(f))确定的; (c) 是在光流分割区域里由 Canny 算子计算出的边缘图; (d) 是执行融合算法后最后结果,目标被完整地检测出来。在分割图的计算中, C-均值聚类门限为 3.0; 利用 Canny 算子计算边缘时, $\sigma = 1.25$, 门限值为 0.105; 融合过程中目标的流速模值门限为 1.25 像素/帧。

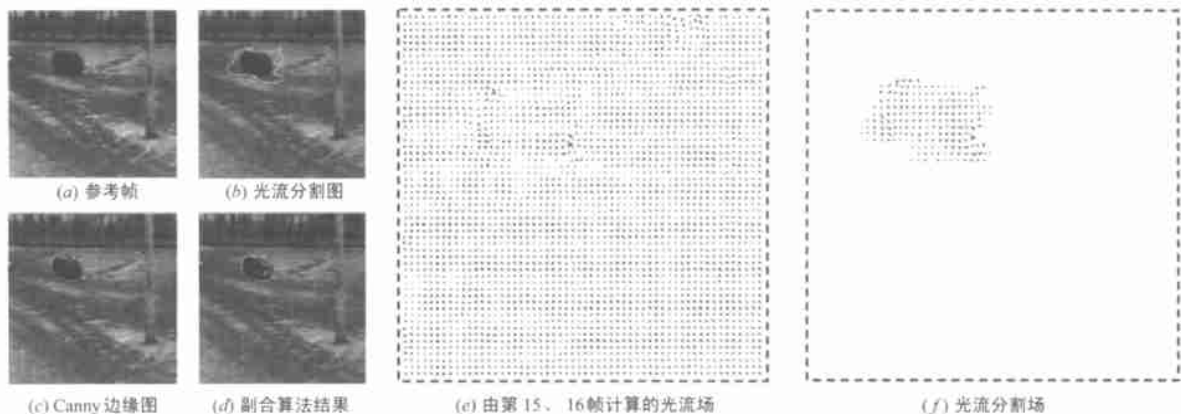


图 2 图像中只包含单个运动目标的实验结果

6.2 多运动目标检测

图 3 是图像中包含多个运动目标的实验结果,其中 (a) 为熟悉的 taxi 图像序列的第 15 帧,有 3 个有运动的目标,已在图中标出,其中目标 3 部分被树遮掩;(b) 是在原图像上给出的基于光流的分割区域图,白色曲线形成的闭合区域为分割区域,这个区域是根据第 15 帧和 17 帧图像计算的光流场(图

3(e)) 及其分割算法得到的光流分割场(图 3(f)) 确定的;(c) 是在分割区域里由 Canny 算子计算出的边缘图;(d) 是执行融合算法的最后结果,检测出了 3 个完整的运动目标.在分割图的计算中,C-均值聚类门限为 2.5;利用 Canny 算子计算边缘时, $\sigma = 1.0$,门限值为 0.045;融合过程中对 3 个目标的流速模值门限分别为 1.35 像素/帧、1.76 像素/帧和 1.52 像素/帧.

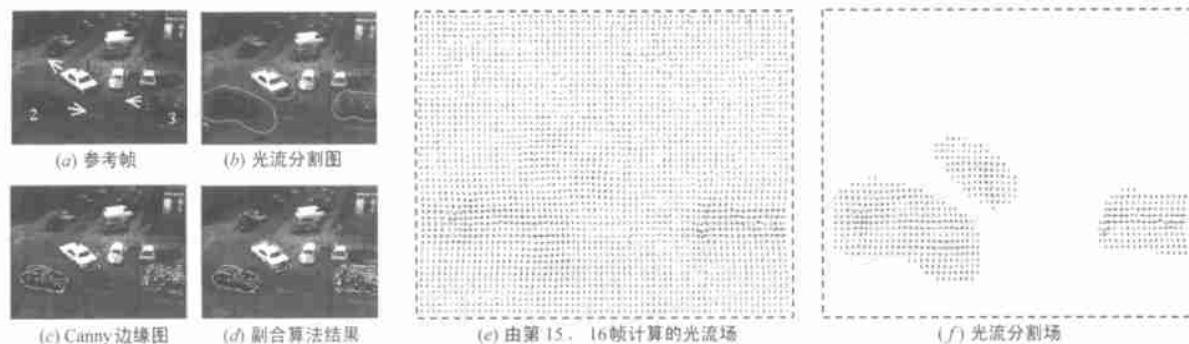


图 3 图像中包含多个运动目标的实验结果

6.3 计算时间分析

我们的仿真实验是在内存为 128 兆,主频 1.7G 的奔腾-IV 处理器的 PC 计算机上进行的,对单个运动目标的小汽车图像,检测出完整的目标需要 3.56 秒,其中光流场的计算为 2.82 秒.对含多目标的 taxi 图像,检测出 3 个完整的目标需要 3.74 秒,其中光流场的计算为 2.76 秒.可见,时间主要花费在光流场的计算上.但是,光流场的计算过程满足并行算法的要求,采取并行算法后,时间消耗会大幅度的减少,因此为实现运动目标的实时检测提供了很好的技术方向.

7 结论

在摄像机运动时,基于单一方法的运动目标的检测变得更为困难.针对这种情况,我们提出了基于光流场分割和 Canny 边缘提取的融合算法.分别对具有单个运动目标和多运动目标的两组自然图像序列进行了实验,均取得了满意的效果.同时,我们的融合算法允许边缘图的过检测,即降低背景干扰的同时,目标弱边缘的损失是允许的.通过对计算时间的分析,可以很好地看出我们提出的融合算法可以实时应用于单个和多个运动目标的检测,这对实际应用具有重大意义.

参考文献:

- [1] Thompson W B, Pong T C. Detecting moving object [J]. Int J Comp Vision, 1990, 4: 39 - 57.
- [2] Sasa G, Loncaric S. Spatio-temporal image segmentation using optical flow and clustering algorithm [A]. First Int'l Workshop on Image and Signal Processing and Analysis [C]. Pula, Croatia, 2000. 63 - 68.
- [3] Smith S M, Brady J M. ASSET-2: Real-Time Motion Segmentation and shape Tracking [J]. IEEE Trans, 1995, PAMI-17(17): 814 - 820.
- [4] Adiv G. Determining three-dimensional motion and structure from optical flow generated by several moving objects [J]. IEEE Trans, 1985, PAMI-7(4): 384 - 401.
- [5] Nagel H H. Displacement vectors derived from second-order intensity variations in image sequences [J]. Comp Vision Graph Image Process, 1983, 25: 85 - 117.

- [6] Nagel H H, Enkermann W. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences [J]. IEEE Trans, 1986, PAMI-5: 565 - 593.
- [7] Horn B K P, Schunck B G. Determining optical flow [J]. Artificial Intelligence, 1981, 17: 185 - 203.
- [8] Bruss A R, Horn B K P. Passive navigation [J]. Computer Vision, Graphics, and Image Processing, 1983, 21: 3 - 20.
- [9] 李金宗. 模式识别导论 [M]. 北京: 高等教育出版社, 1994. 316 - 321.
- [10] 郑南宁. 计算机视觉与模式识别 [M]. 国防工业出版社, 1998. 66 - 69.

作者简介:



张泽旭 男, 1971 年 3 月出生于吉林省吉林市, 现为空军第一飞行学院讲师, 2000 年毕业于哈尔滨工业大学电子与通信工程系, 获工学硕士学位. 现为该校博士研究生, 主要从事模式识别、图像处理、计算机视觉和 D3S 技术等方面的研究.



李金宗 男, 1940 年 3 月生于安徽省砀山, 教授, 博士生导师, 1964 年毕业于哈尔滨工业大学无线电工程系, 1985.3 - 87.3 在原西德研究图像序列分析技术, 回国后研究 D3S 技术的理论和应用, 其中涉及信号检测理论、图像处理和系列图像分析、模式识别、精密制导和寻的跟踪、三维视觉和空间交会对接等领域, 已完成六项部级专家会议鉴定, 出版专著“模式识别导论”、“离散正交变换导论”与“空间交会对接技术”等三部, 在国内外刊物和学术会议上发表论文 60 余篇.

李宁宁 女, 1970 年出生于黑龙江省哈尔滨, 1995 年毕业于哈尔滨工业大学研究生院, 获信号与信息处理学科硕士学位, 目前在中国空间飞行器总体设计部工作, 主要从事弱信号检测和图像处理、数据管理和通讯等方面的研究, 对亚像元检测技术有较深入的研究.