

文章编号:1007-2780(2014)06-0997-06

改进的基于码本模型目标检测算法

刘翔^{1*}, 杨鑫², 王蕾³

- (1. 武汉科技大学 工程训练中心, 湖北 武汉 430065;
2. 中国电力工程顾问集团 中南电力设计院, 湖北 武汉 430071;
3. 武汉科技大学 工程训练中心, 湖北 武汉 430065)

摘要:为克服基于码本模型的目标检测算法在光照变化条件下对检测结果产生的影响,提出了一种融合卡尔曼滤波思想的改进码本算法。该方法在计算图像像素亮度变化时选用了 YUV 颜色空间,使空间坐标轴与亮度变化方向一致,使亮度变化的计算量平均下降了 44.9%。同时在建立码本模型时为每个像素初始化一个卡尔曼滤波器,该滤波器利用前后两帧图像亮度信息预测与修正当前像素值,对光照变化有较好的适应性。仿真实验结果表明,该算法与 YUV 码本模型、RGB 码本模型以及 GMM 算法相比在亮度变化的条件下对噪声的抑制作用更强,体现出更好的自适应性。

关键词:目标检测;码本模型;卡尔曼滤波

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/YJYXS20142906.0997

Object detection algorithm based on improved codebook model

LIU Xiang^{1*}, YANG Xin², WANG Lei³

- (1. *Engineering Training Centre of Wuhan University of Science and Technology, Wuhan 430065, China;*
2. *Centre Southern China Electric Power Design Institute of China Power Engineering Consulting Group, Wuhan 430071, China;*
3. *Engineering Training Centre of Wuhan University of Science and Technology, Wuhan 430065, China)*

Abstract: In order to overcome the influence of illumination changes on detecting results by using codebook-based object detection algorithm, an improved codebook algorithm that fuses Kalman filtering is proposed in this paper. The algorithm calculates the pixel brightness change in YUV colour space, which corresponds brightness changes with spatial coordinates and makes the calculation of brightness process fallen by an average of 44.9%. Meanwhile, it initializes a Kalman filter for each pixel at the time when the codebook model is established and the filter uses two successive image's brightness information to predict and correct current pixel value. This method has good adaptability on brightness changes. The experimental results show that this algorithm reflects a better adaptability under the condition of brightness changes, compared with ordinary YVU codebook model, RGB codebook model and GMM method.

Key words: object detecting; codebook model; Kalman filtering

收稿日期:2013-12-18;修订日期:2014-03-27.

* 通信联系人, E-mail:liuxiang@wust.edu.cn

1 引言

计算机视觉是当今科学领域中一个极富有挑战性的重要研究领域。运动目标检测是计算机视觉研究中各种复杂后期处理(如:目标跟踪、目标分类识别、行为理解等)的基础,因此是该领域中一大研究热点。目标检测的任务是将视频图像分割成目标前景和静态背景两个部分,然后再将目标前景部分提取出来。尽管从上世纪 80 年代以来在计算机视觉基础研究领域取得了许多进展,但至今为止仍然没有一种对于各种不同场景普遍适用的检测方法。学者们在此项研究中提出的方法主要有:光流法^[1-2]、帧间差分法^[3-5]和背景减除法^[6-7]等。

从应用研究的情况看来,背景减除法是一种快速、简单和最为常用的方法。其难点在于建立一个具有较强自适应性的背景模型,以及随着时间推移对该模型进行更新维护。平均背景法^[8]以一定时间(帧数)内像素的平均值作为背景,其背景模型的质量与所选用的视频图像的帧数有较强的依赖关系。当帧数选得较大时可以获得质量较高的背景模型,但却消耗了过多的内存。除此以外,该方法的自适应性也较差。当环境发生变化后,检测的效果不甚理想。高斯背景模型法^[9]对每个像素进行高斯统计建模,然后再将新输入的像素点与模型中均值进行比较,根据它们之间的偏差程度来划分前景与背景。但当背景呈现出多模特征时检测效果较差。混合高斯背景模型法^[10]用多个高斯分布对每个像素进行建模,如果场景中的像素点满足任意一个高斯分布则为背景,否则即为前景,最后再用新像素点的信息更新原有模型中的参数。该方法虽克服了高斯背景模型法的缺点,但算法比较复杂计算量大。基于码本的背景模型法^[11]给场景中背景的获得带来了另外一种思路。该方法的基本思想是得到每个像素的时间序列模型,它能较好地处理时间的起伏,但对内存消耗较大。

本文主要针对静态背景环境下使用码本模型进行目标检测的问题进行了探讨。文章的主要结构如下:首先简要介绍了码本算法的基本原理和其优势,然后指出了码本算法中存在的问题(亮度计算与亮度随时间或受噪声影响变化的问题),重

点介绍了针对这些问题对原算法进行的改进(颜色空间转换与使用 Kalman 滤波修正像素值),最后在仿真结果的基础上分析了改进算法的优点和不足之处。

2 基于码本模型的目标检测方法

基于码本模型的目标检测方法最早是由 Kim 等人提出的,它可以分为背景建模和目标检测两个阶段。建模阶段先对场景中的每个像素点构建一个初始码本模型,然后通过一段时间内输入视频对应位置上像素值的亮度变化和色度失真进行聚类,最终得到场景中背景像素的码本模型;检测阶段再对新输入的像素值与模型中的码元进行聚类,从而将背景与前景区分开。

相对于其他的目标检测算法,码本模型算法的优势^[12]在于:

- (1)能够很好地实现在树枝摇摆、水面波纹、旗帜飘扬等多模动态背景条件下的检测;
- (2)算法减少了计算量,降低了存储空间,对一段 30 帧/s 的 5 min 视频,平均每个像素的码本仅需要 6.5 个码字;
- (3)算法将亮度和色度分开,解决了光线变化对背景分割的影响;
- (4)算法具有很好的自适应性,在背景分割的同时,引入缓冲码本进行背景的自适应更新。

3 码本模型的改进算法

Kim 等人提出的码本模型是在 RGB 空间上对像素颜色失真进行聚类并以此作为产生码字的依据。RGB 空间模型如图 1 所示, X 、 Y 、 Z 三个方向分别代表红、绿、蓝。亮度(灰度)轴沿该立方体经过原点 O 的对角线方向,如图 1 中 OA 所示。要对颜色失真进行聚类必须计算该模型下像素值的颜色亮度,由图 1 可知,在此空间亮度轴上任意一点 X 的亮度值为:

$$OX = \sqrt{X'C^2 + X'D^2 + X'X^2}, \quad (1)$$

其中: $X'C$ 、 $X'D$ 和 $X'X$ 分别表示 X 在 X 、 Z 、 Y 三个方向上的投影。

由于 RGB 颜色空间与人眼感知的差异较大,并且计算颜色失真需要进行开方与平方运算,增加了算法的时间复杂性。一种改进的方法是将像

素的颜色空间从 RGB 映射到 YUV 空间,将亮度信息与色度信息分离保证亮度变化与空间坐标轴方向一致,在计算亮度失真时只需使用加法和减法,大大地降低了计算量。基于此考虑,本文选取在 YUV 空间构造码本模型的方法进行前景检测。

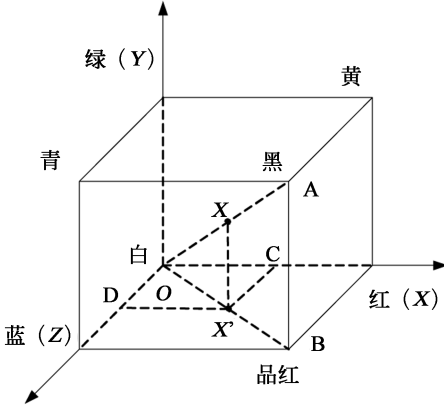


图1 RGB颜色模型空间

Fig.1 RGB color space model

尽管基于码本模型的目标检测算法在分割前景与背景时对光线变化影响分割效果上有一定的抑制作用,但在光线长时间较大范围变化的情况下,这种检测算法可能会失效。因为此时的像素值可能已经不是目标的特征值,用这个值进行聚类计算会发生错误的划分。为了获得更加接近目标特征值,本文在码本模型中建立一个与码元更新同步的 Kalman 滤波器,利用 Kalman 滤波器的收敛性,使滤波器的输出值始终接近于目标的特征值,减少因亮度变化引起的错分概率,提高检测目标的准确度与完整性。

3.1 Kalman 滤波器

Kalman 滤波^[13-15]技术在雷达、GPS 导航、计算机视觉等领域有着广泛的应用。Kalman 滤波器是一种线性滤波器,它假设系统的状态方程和观测方程是线性的,系统噪声和观测噪声是独立的高斯过程。系统状态的均值和协方差矩阵用一组观测方程和更新方程来迭代计算。

Kalman 滤波器的递归过程可由公式(2)~(7)表示。设 $\mathbf{x}_{k-1}$ 是 $k-1$ 时刻的 n 维状态向量, $\mathbf{P}_{k-1}$ 是 $k-1$ 时刻的误差协方差矩阵。

k 时刻的状态预测值为:

$$\mathbf{x}'_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{w}_k. \quad (2)$$

k 时刻的测量值为:

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}'_k + \mathbf{v}_k. \quad (3)$$

k 时刻的先验估计为:

$$\mathbf{P}'_k = \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q}_{k-1}. \quad (4)$$

Kalman 增益为:

$$\mathbf{K}_{gk} = \mathbf{P}'_k\mathbf{H}_k^T (\mathbf{H}_k\mathbf{P}'_k\mathbf{H}_k^T + \mathbf{R}_k)^{-1}. \quad (5)$$

修正后的目标状态:

$$\mathbf{x}_k = \mathbf{x}'_k + \mathbf{K}_{gk}(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}'_k). \quad (6)$$

k 时刻的后验估计为:

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_{gk}\mathbf{H}_k)\mathbf{P}'_k. \quad (7)$$

将式(6)结果作为下一时刻的状态向量代入(2)式,将(7)式结果作为下一时刻的先验估计代入(4)式不断重复(2)~(7)式的计算就是 Kalman 滤波器的工作过程。上述各式中 $\mathbf{A}$ 为系统传递矩阵, $\mathbf{w}_k$ 称为过程噪声, $\mathbf{H}$ 是测量矩阵, $\mathbf{v}_k$ 是测量噪声, $\mathbf{Q}_{k-1}$ 是过程噪声 $\mathbf{w}_k$ 的协方差矩阵, $\mathbf{R}_k$ 是测量噪声 $\mathbf{v}_k$ 的协方差矩阵, $\mathbf{I}$ 是单位矩阵。

3.2 码本模型的建立

设 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 是用于建立背景模型的 n 帧视频序列中同一位置上像素值的集合,其中 $\mathbf{x}_i$ ($i=1, 2, \dots, n$) 是一个包含 YUV 三个颜色分量的向量。接下来为每个像素建立一个与之对应的码本模型: $\mathbf{CB}_j = \{c_j, k_j, L_j, t_j, p_j\}$, ($j=1, 2, \dots, m$) 其中 c_j ($j=1, 2, \dots, m$) 是该码本中的一个码字,它是由 6 个基本参数组成的一个六元组: $c_j = \langle I^{\text{high}}, I^{\text{low}}, I^{\text{max}}, I^{\text{min}}, t_{\text{update}}, t_{\text{stale}} \rangle$ 这 6 个参数的含义依次为:此码元各通道的阈值学习上限、阈值学习下限、像素中各通道的最大值、最小值,此码元最后一次更新的时间以及此码元最长不更新时间。 k_j ($j=1, 2, \dots, m$) 是码书对应的 Kalman 滤波器结构。该结构中主要参数组成一个九元组: $k_j = \langle \mathbf{x}^{\text{pre}}, \mathbf{x}^{\text{cor}}, \mathbf{T}, \mathbf{M}, \mathbf{Q}, \mathbf{R}, \mathbf{p}^{\text{pre}}, \mathbf{p}^{\text{post}}, \mathbf{K}_g \rangle$ 。 $\mathbf{x}^{\text{pre}}$ 为预测状态向量, $\mathbf{x}^{\text{cor}}$ 为修正后的状态向量, $\mathbf{T}$ 为传输矩阵, $\mathbf{M}$ 为量测矩阵, $\mathbf{Q}$ 是过程噪声矩阵, $\mathbf{R}$ 是测量噪声矩阵, $\mathbf{p}^{\text{pre}}$ 是先验误差估计矩阵, $\mathbf{p}^{\text{post}}$ 是后验误差估计矩阵, $\mathbf{K}_g$ 是 Kalman 增益矩阵。 L_j 是该码本中码元的个数, t_j 是该码本现在的时间,可理解为当前帧序号, p_j 是该码书对应像素点的像素值。

3.3 算法流程

背景码本学习建立阶段的主要步骤:

①输入第一帧图像,为图像中每个像素点新建一个码本并将码本置空。初始化 Kalman 滤波

器结构中的 $\mathbf{x}^{\text{pre}}, \mathbf{T}, \mathbf{M}, \mathbf{p}^{\text{pre}}, \mathbf{Q}, \mathbf{R}$ 等参数。用此时的像素值初始化码字 c 中的 $I^{\text{high}}, I^{\text{low}}, I_{\text{max}}$ 和 I_{min} 。

②对新输入图像中的像素 x_i 将其带入已初始化的 Kalman 滤波器中计算出 $\mathbf{x}^{\text{cor}}, \mathbf{p}^{\text{post}}, \mathbf{K}_g$ 等参数。其中, $\mathbf{x}^{\text{cor}}$ 将作为下帧输入图像中该位置像素的 $\mathbf{x}^{\text{pre}}$ 。

③将 $\mathbf{x}^{\text{cor}}$ 各通道分量值与码字 c 中 I_{max} 和 I_{min} 进行比较, 如果 $\mathbf{x}^{\text{cor}}$ 各通道分量均在 I^{high} 与 I^{low} 之间则认为 $\mathbf{x}^{\text{cor}}$ 与码字匹配。最后再根据 $\mathbf{x}^{\text{cor}}$ 的值更新 I_{max} 和 I_{min} 。

$$I_{\text{min}j} = \begin{cases} x_j^{\text{cor}} & x_j^{\text{cor}} < I_{\text{min}j} \\ I_{\text{min}j} & x_j^{\text{cor}} \geq I_{\text{min}j} \end{cases} \quad (8)$$

$$I_{\text{max}j} = \begin{cases} x_j^{\text{cor}} & x_j^{\text{cor}} > I_{\text{max}j} \\ I_{\text{max}j} & x_j^{\text{cor}} \leq I_{\text{max}j} \end{cases} \quad (9)$$

下标 j 表示这些值在各通道上的分量, 本文中指 Y, U, V 三个分量。

④更新阈值学习参数各通道的上下限

$$I_{\text{low}j} = \begin{cases} I_{\text{low}j} - \epsilon & x_j^{\text{cor}} + cb\text{Bounds} < I_{\text{low}j} \\ I_{\text{low}j} & \text{otherwise} \end{cases} \quad (10)$$

$$I_{\text{high}j} = \begin{cases} I_{\text{high}j} + \epsilon & x_j^{\text{cor}} - cb\text{Bounds} > I_{\text{high}j} \\ I_{\text{high}j} & \text{otherwise} \end{cases} \quad (11)$$

⑤如果 $\mathbf{x}^{\text{cor}}$ 与当前码本中任一码字都不匹配则利用 $\mathbf{x}^{\text{cor}}$ 新建一个码字, 同时使码字长度增加一个单位。

⑥删除码本中一段时间内没有更新的码字以精简码字的长度。

目标检测阶段的主要步骤:

①根据背景建立阶段学习得到的上、下边界确定检测阶段各通道学习的上、下限范围。

②如果新输入像素各通道分量均在此上、下限范围之内, 则此像素与背景匹配; 反之, 则将该像素与前景匹配。

③根据匹配结果将前景从图像中分割出来。

4 实验结果与分析

为了验证本文算法的有效性, 实验中选取了 3 段测试视频(分辨率 320×240)分别使用 YUV 空间的 codebook 算法、RGB 空间的 codebook 算法、GMM 算法以及本文的算法进行比较。本文

的实验硬件环境为: 2 GB 内存, Intel(R) Core™ i3-2310M CPU @ 2.10 GHz; 软件环境为: Microsoft Visual Studio 6.0; 编程语言为: C++。3 次实验的结果分别为图 2、图 3 和图 4 所示。

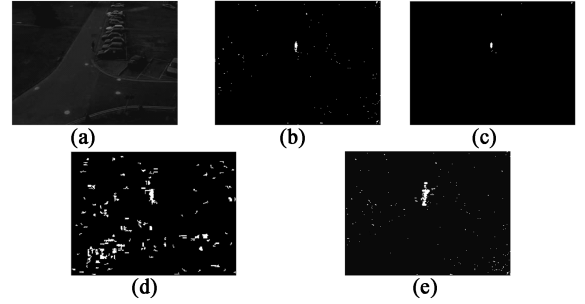


图 2 测试视频一的实验结果

Fig. 2 Experimental results of test video I

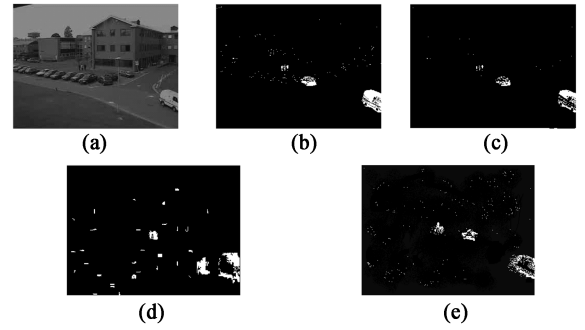


图 3 测试视频二的实验结果

Fig. 3 Experimental results of test video II

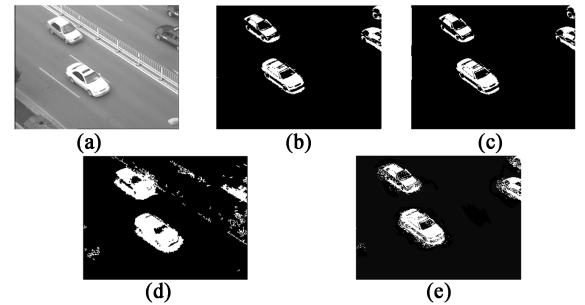


图 4 测试视频三的实验结果

Fig. 4 Experimental results of test video III

上面三幅图中(a)为测试视频原始帧序列, (b)为使用 YUV 模型 codebook 算法检测的结果, (c)为使用本文算法检测的结果, (d)为使用 GMM 算法检测的结果(高斯分布的个数取 3), (e)为使用 RGB 模型 codebook 算法的检测结果。图 2 为一对对比度较低的室外视频, 使用 YUV 模

型和 RGB 模型的 codebook 算法可以获得比较理想的前景目标,但由于室外光照变化的影响背景中混入了较多的噪声。使用本文提出的算法在 Kalman 滤波器的修正作用下,噪声得到了较好地抑制。由于视频本身对比度较低,这种修正作用也可能使部分前景误判为背景,但总体上仍能检测出目标。而使用 GMM 算法的检测效果较差,目标几乎淹没在噪声中。图 3 为一对比度较高的室外视频,比较其(b)、(c)、(d)、(e)各图可以发现本文的算法对噪声抑制较好,并且在这样的环境中各算法检出的目标前景信息损失情况明显优于图 1 所示的情况。图 4 为一背景较简单且对比度较高的室外视频,本文的算法与两种不同颜色模型下的 codebook 算法检测的效果近似相同,都对屏幕左下方草坪和树枝晃动引起的干扰有较好地抑制。使用 GMM 算法时,这样的干扰以及中央护栏上光线的漫反射干扰抑制效果不佳。表 1 中列出了几种算法的运行时间比较。

表 1 几种算法的运行时间比较 (单位:s)
Tab.1 Comparison of the running time of algorithms

	Codebook RGB 空间	Codebook YUV 空间	Codebook YUV& Kalman	GMM
测试视频一	0.26	0.16	2.57	0.29
测试视频二	0.33	0.27	3.63	0.42
测试视频三	0.12	0.08	2.18	0.17

参 考 文 献:

[1] Barron J, Fleet D. Performance of optical flow techniques [J]. *International Journal of Computer Vision*, 1994, 12 (1):42-77.

[2] 张水发,丁欢,张文生. 双模型背景建模与目标检测研究[J]. 计算机研究与发展, 2011, 48(11):1983-1990.
Zhang S F, Ding H, Zhang W S. Background modeling and object detection based on two-model [J]. *Journal of Computer Research and Development*, 2011, 48(11):1983-1990. (in Chinese)

[3] Lipton A, Fujiyoshi H, Ptail R S. Moving target classification and tracking from real-time video [C]. *Proc. IEEE Workshop on Application of Computer Vision*, 1998, 17(9):8-14.

[4] 吴君钦,刘昊,罗勇. 静态背景下的运动目标检测算法[J]. 液晶与显示, 2012, 27(5):682-686.
Wu J Q, Liu H, Luo Y. Algorithm of moving object detection in static background [J]. *Chinese Journal of Liquid Crystals and Displays*, 2012, 27(5):682-686. (in Chinese)

[5] 丁雪梅,王维雅,黄向东. 基于差分 and 特征不变量的运动目标检测与跟踪[J]. 光学精密工程, 2007, 15(4):570-576.
Ding X M, Wang W Y, Huang X D. New method for detecting and tracking of moving target based on difference and invariant [J]. *Optics and Precision Engineering*, 2007, 15(4):570-576. (in Chinese)

综上所述,本文所提出的改进码本算法与 RGB 码本算法、YUV 码本算法、GMM 算法相比在确保检测出足够完整的前景目标情况下对光照变化和环境扰动干扰有更好的抑制作用。但由于构造码本时需为像素建立 Kalman 结构,背景建立阶段要进行 Kalman 迭代运算因此使程序的实时性降低。

5 结 论

提出了一种改进的码本模型目标检测算法。该方法主要有以下两个特点:(1)在 YUV 空间里建立码本模型,减少了计算亮度失真的计算量;(2)为视频序列中同一位置上的像素点序列构造 Kalman 滤波器模型,并用 Kalman 滤波器修正后的结果作为 codebook 背景学习的聚类中心,抑制了光照变化的影响。

但是本文算法还存在一些不足之处,(1)Kalman 滤波器是线性滤波器,在现实中对于复杂的背景观测方程往往是非线性的,采用线性方程的 Kalman 滤波器迭代结果可能会发散;(2)codebook 在进行背景学习前需要从 Kalman 滤波器的输出中获得数据,使得背景建立的时间增加,导致程序运行速度降低使目标检测的时间增加。这些问题将是后续工作中继续研究的对象。

- [6] 代科学,李国辉,涂丹,等. 监控视频运动目标检测减背景技术的研究现状和展望[J]. 中国图象图形学报,2006,11(7):919-927.
- Dai K X, Li G H, Tu D, *et al.* Prospects and current studies on background subtraction techniques for moving objects detection from surveillance video [J]. *Journal of Image and Graphics*, 2006,11(7):919-927. (in Chinese)
- [7] 张红颖,胡正. CenSurE 特征和时空信息相结合的运动目标检测[J]. 光学精密工程,2013,21(9):2452-2463.
- Zhang H Y, Hu Z. Moving object detection in combination of CenSurE and spatial-temporal information [J]. *Optics and Precision Engineering*, 2013,21(9):2452-2463. (in Chinese)
- [8] 袁宝红,张德祥,张玲君. 基于 OpenCV 的视频运动目标检测与跟踪[J]. 计算机系统应用,2013,22(5):90-93.
- Yuan B H, Zhang D X, Zhang L J. Detecting and tracking of moving target in video sequence based on OpenCV [J]. *Computer Systems and Applications*, 2013,22(5):90-93. (in Chinese)
- [9] 康晓晶,吴谨. 基于高斯背景建模的目标检测技术[J]. 液晶与显示,2010,25(3):454-459.
- Kang X J, Wu J. Object detecting technology based on Gauss background modeling [J]. *Chinese Journal of Liquid Crystals and Displays*, 2010,25(3):454-459. (in Chinese)
- [10] Stauffer C, Gimson W. Adaptive background mixture models for real-time tracking [C] // *Int. Conf. Computer Vision and Pattern Recognition*, 1999, 2:246-252.
- [11] Kim K, Chalidabhongse T H, Hardwood D, *et al.* Background modeling and subtraction by codebook construction [C] // *Proc. IEEE Int. Conf. on Image Processing*, Singapore, 2004:3061-3064.
- [12] 黄进,金炜东,秦娜. 基于三维高斯混合码本模型的运动目标检测算法[J]. 西南交通大学学报,2012,47(4):662-668.
- Huang J, Jin W D, Qin N. Moving objects detection algorithm based on three-dimensional gaussian mixture codebook model [J]. *Journal of Southwest Jiaotong University*, 2012,47(4):662-668. (in Chinese)
- [13] 杨冰,张为,王猛. 基于 Kalman 滤波器运动目标跟踪的火灾监测方法[J]. 信息技术,2013 (7):101-105.
- Yang B, Zhang W, Wang M. Fire detection method based on moving target tracking with Kalman filter [J]. *Information technology*, 2013 (7):101-105. (in Chinese)
- [14] 赵其杰,屠大维,高健,等. 基于 Kalman 滤波的视觉预测目标跟踪及其应用[J]. 光学精密工程,2008,16(5):937-942.
- Zhao Q J, Tu D W, Gao J *et al.* Kalman filter based vision predicting and object tracking method and its application [J]. *Optics and Precision Engineering*, 2008,16(5):937-942. (in Chinese)
- [15] 杨金显,袁赣南. 基于 MIMU/GPS 的组合导航设计及实验[J]. 光学精密工程,2008,16(2):285-294.
- Yang J X, Yuan G N. Design and experiment for INS based on MIMU/GPS [J]. *Optics and Precision Engineering*, 2008,16(2):285-294. (in Chinese)

作者简介:刘翔(1983—),男,湖北武汉人,硕士,主要研究方向为控制理论与控制工程和图像处理。E-mail:liuxiang@wust.edu.cn