

基于 UKF 的窗口自适应 Mean-Shift 算法

杨 帆, 郑春红, 杨 刚

(西安电子科技大学工程学院, 西安 710071)

摘 要: 传统的 Mean-Shift 跟踪算法窗口固定, 不能对尺度任意变化的目标进行有效跟踪。为此, 提出一种多尺度理论与无味卡尔曼滤波器(UKF)相结合的视频跟踪改进算法。利用多尺度理论统计跟踪窗内的信息量, 使用 UKF 对得到的信息量进行预测, 通过修正后的信息量计算窗口变化比例系数, 对尺度任意变化的目标进行跟踪。实验结果证明, 该算法能对尺度任意变换的目标进行有效跟踪。

关键词: 目标跟踪; 无味卡尔曼滤波器; Mean-Shift 算法; 信息度量

Mean-Shift Algorithm with Adaptive Window Based on UKF

YANG Fan, ZHENG Chun-hong, YANG Gang

(School of Electronic Engineering, Xidian University, Xi'an 710071, China)

【Abstract】 The traditional fixed bandwidth Mean-Shift tracking algorithm can not have an effective tracking for any changes in targets. An novel method is proposed that is multi-scale space theory combined with Unscented Kalman Filter(UKF). UKF filter is introduced to predict the information in the tracking window which is calculated by the multi-scale space theory. The proportion of the target image area is got by the modified information. It is implemented by the combination of the Mean-Shift tracking algorithm and UKF to track targets. Experimental result confirms the effectiveness of the improved algorithm.

【Key words】 object tracking; Unscented Kalman Filter(UKF); Mean-Shift algorithm; information measure

DOI: 10.3969/j.issn.1000-3428.2011.14.052

1 概述

目标跟踪是计算机视觉研究领域的热点问题, 在视频监控、目标识别、人机界面等应用领域中, 往往需要在复杂的环境下对运动物体进行有效的跟踪。Mean-Shift 算法作为一种高效的模式匹配算法, 已经成功地应用在实时性要求较高的目标跟踪算法中。

文献[1]分别以核窗宽确定的原始跟踪窗口做正负10%增量修正后的跟踪窗口进行跟踪, 并选择较大的 Bhattacharyya 系数所对应的核窗宽作为最佳核窗宽, 在实际中, 该方法只能对尺寸逐渐缩小的目标有较好的跟踪效果。文献[2]利用视频前后2帧中目标的角点匹配估计刚性物体的仿射模型参数, 并由此参数来更新核窗宽, 由于依赖于角点匹配, 不适合非刚性物体的跟踪。文献[3]将信息度量方法推广到由视频图像构成的多尺度图像空间中。但是, 当跟踪窗内的背景不断变化时, 多尺度理论中的信息量计算方法容易将变化的背景也当做有用的目标信息进行统计, 使文献[3]的方法在使用中受到较大限制。

本文将多尺度图像信息量的度量方法和无味卡尔曼滤波器(Unscented Kalman Filter, UKF)相结合, 并引入到 Mean-Shift 跟踪的窗口更新中, 根据信息量的变化自动选取跟踪窗口的大小。在文献[3]的基础上, 计算图像的信息量, 介绍 UKF 算法, 提出一种基于 UKF 的窗口自适应 Mean-Shift 算法。

2 多尺度图像信息量度量

人在不同的距离观察同一图像获得的感受是不一样的, 如远距离观察时, 仅能看到图像的轮廓, 而近距离观察时, 能看到更多的细节, 这就是尺度效应。而视频跟踪中尺度不断变化的被跟踪目标也满足尺度效应, 因此, 如果能找到有效的信息量度量方法来区分目标的不同尺度, 即可根据信息

量的变化来自动选择跟踪窗口的尺度。

2.1 图像特征点

Marr 视觉理论^[4]认为, 视觉处理得到的初始简图中要素的数量可以作为一个十分好的信息度量标准。基于这一思想, 文献[5]定义了2类特征点, 并用图像中两类特征点的数量来度量尺度空间中图像的信息量。

从微分算子的局部极值角度统一这两类特征点。假设在2维图像平面内, 点 P 为图像 $f(x, y)$ 内某像素点, $1 \sim 8$ 为点 P 的8邻接像素点 $N_8(P)$ 。现以 P 为极点建立极坐标系, 设极角为 l , $l \in L$, $L = \{k\pi/4, k = -4, -3, -2, -1, 0, 1, 2, 3, 4, k \in Z\}$, 则对 P 的8个邻接像素点进行如下描述:

$$f_l(x, y) = \begin{cases} f(x - \sin(l), y + \cos(l)) & l = \frac{m\pi}{2} \\ f(x - \sqrt{2}\sin(l), y + \sqrt{2}\cos(l)) & l = \frac{m\pi}{4}, l \neq \frac{m\pi}{2} \end{cases} \quad (1)$$

在极角 l 方向上, 取步长 $\Delta h = 1$, 定义 $f(x, y)$ 的方向微分算子:

$$\nabla_l f(x, y) = \frac{f_l(x, y) - f(x, y)}{\Delta h} = f_l(x, y) - f(x, y) \quad (2)$$

下面给出文献[5]中2类特征点在式(2)定义的方向微分算子下的定义:

定义 1 设2维离散平面上点 $P(x, y)$, $\forall l_1, l_2 \in L$, $L = \{k\pi/4, k = -4, -3, -2, -1, 0, 1, 2, 3, 4, k \in Z\}$, 当 $l_1 - l_2 = m\pi$, $m = \pm 1$ 时, $\nabla_{l_1} f(x, y) \cdot \nabla_{l_2} f(x, y) > 0$, 则点 P 称为第1类特征点。

定义 2 设2维离散平面上点 $P(x, y)$, $\forall l_1, l_2 \in L$, $L =$

作者简介: 杨 帆(1986—), 男, 硕士研究生, 主研方向: 图像处理, 目标跟踪, 目标识别; 郑春红, 副教授; 杨 刚, 教授

收稿日期: 2011-01-03 **E-mail:** yangfan860615@163.com

$\{k\pi/4, k = -4, -3, \dots, 0, 1, 2, 3, k \in \mathbb{Z}\}$, 当 $l_1 - l_2 = m\pi$, $m = \pm 1$ 时, $\nabla_{l_1}(f_{l_1}(x, y) - f_{l_2}(x, y)) \cdot \nabla_{l_2}(f_{l_1}(x, y) - f_{l_2}(x, y)) > 0$, 则点 P 称为第 2 类特征点。

2.2 图像信息量的度量方法

由于第 1 类和第 2 类特征点都是视觉上的细节重要特征点, 因此这 2 类特征点的集合从视觉上能反映图像的重要信息, 由这 2 类特征点可以定义图像 $f(x, y)$ 的信息量。

定义 3 设图像 $f(x, y)$ 的第 1 类特征点个数为 I_1 , 第 2 类特征点的个数为 I_2 , 则图像的信息量 I 定义为 I_1 和 I_2 之和。

文献[5]证明了定义 3 给出的信息度量方法在 Gauss 尺度空间中的有效性, 文献[3]将它推广到由视频图像构成的多尺度图像空间中。

2.3 窗口变化比例系数的计算

根据实际视频中目标的尺度变化速度, 每隔 N 帧统计 1 次被跟踪的目标信息量, 然后根据信息量的变化对跟踪窗进行更新。假设第 n 帧跟踪窗内的信息量为 I_1 , 对跟踪窗乘以 $1 \pm \alpha$ ($0 \leq \alpha \leq 0.5$) 后的区域的信息量为 I_2 、 I_3 (前者对应“+”, 后者对应“-”)。类似地, 对第 $n+N$ 帧统计得到的信息量分别为 I_4 、 I_5 、 I_6 (此时的跟踪窗与第 n 帧的跟踪窗尺寸相等), 并设尺度变化比例系数为 T 。

假如满足 $I_5 - I_2 \geq 0$, 则判断目标尺度可能增加, 否则, 判断目标尺度可能减小, 计算公式如下:

$$T = \begin{cases} \lg(\eta \frac{I_3 - (I_2 - I_1)}{I_1}) & \frac{I_5}{I_4} > \frac{I_2}{I_1} \\ \lg(\eta \frac{I_4}{I_1}) & 0.95 I_1 > I_4 \\ 0 & \text{else} \end{cases} \quad (3)$$

其中, 参数 η 的作用是消除背景对于信息量的影响, 其值接近于 1, 背景越复杂与 1 的差别越大。计算得到当前跟踪窗变化比例系数 T 后, 就可以得到更新后的跟踪窗口的尺寸。

3 UKF 算法

在实际计算中, 信息量的变化往往是非线性的, 而 UKF 就是一种以 UT 变换为基础, 采用 Kalman 线性滤波框架、递归式贝叶斯估计确定性采样的滤波方法。UKF 不是对非线性模型做近似, 而是对状态的概率密度函数做近似, 因此, UKF 算法^[6-8]要比 EKF、KF 算法的估计精度高。

假设一般的非线性系统为:

$$\begin{cases} x_k = f(x_{k-1}, u_k) + w_k \\ y_k = h(x_k) + v_k \end{cases} \quad (4)$$

其中, x_k 为 k 时刻修正的信息量; u_k 是 k 时刻已知的输入量, 即为 k 时刻信息量的测量值; y_k 是 k 时刻的观测量; w_k 是过程噪声; v_k 是观测噪声, 本文假设两者都是零均值的高斯噪声; f 是状态方程; h 是测量方程。

UKF 算法具体如下:

(1) 初始化

$$\begin{aligned} \bar{x}_0 &= E[x_0] \\ P &= E_0[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] \end{aligned}$$

(2) Sigma 点的计算

$$\begin{aligned} \chi_{k-1} &= [\bar{x}_{k-1}, \bar{x}_{k-1} \pm \sqrt{(n+\lambda)P_{k-1}}]; \\ \hat{\chi}_{k|k-1} &= f(\chi_{k-1}) \\ \bar{x}_{k|k-1} &= \sum_{i=0}^{2n} W_i^m \hat{\chi}_{i,k|k-1} \end{aligned}$$

$$\begin{aligned} P_{k|k-1} &= \sum_{i=0}^{2n} W_i^c (\hat{\chi}_{i,k|k-1} - \bar{x}_{k|k-1})(\hat{\chi}_{i,k|k-1} - \bar{x}_{k|k-1})^T \\ \chi_{k|k-1} &= [\bar{x}_{k|k-1}, \bar{x}_{k|k-1} \pm \sqrt{(n+\lambda)P_{k|k-1}}] \\ \gamma_{k|k-1} &= h(\chi_{k|k-1}) \\ \bar{y}_{k|k-1} &= \sum_{i=0}^{2n} W_i^m \gamma_{i,k|k-1} \end{aligned}$$

在 Sigma 点的计算过程中, 有: $W_0^m = \frac{\lambda}{n+\lambda}$; $W_0^c = \frac{\lambda}{n+\lambda} + (n - \alpha^2 + \beta)$; $W_i^m = W_i^c = \frac{1}{2(n+\lambda)}$, $\lambda = n(\alpha^2 - 1)$ 。

(3) UKF 更新

$$\begin{aligned} P_{y_k y_k} &= \sum_{i=0}^{2n} W_i^c (\gamma_{i,k|k-1} - \bar{y}_{k|k-1})(\gamma_{i,k|k-1} - \bar{y}_{k|k-1})^T \\ P_{x_k y_k} &= \sum_{i=0}^{2n} W_i^c (\chi_{i,k|k-1} - \bar{x}_{k|k-1})(\gamma_{i,k|k-1} - \bar{y}_{k|k-1})^T \\ K_k &= P_{x_k y_k} P_{y_k y_k}^{-1} \\ P_k &= P_{k|k-1} - K_k P_{y_k y_k}^{-1} K_k^T \\ \bar{x}_k &= \bar{x}_{k|k-1} + K_k (y_k - \bar{y}_{k|k-1}) \end{aligned}$$

其中, $\sqrt{(n+\lambda)P_{k|k-1}}$ 是 $(n+\lambda)P_{k|k-1}$ 平方根的第 i 行或第 i 列向量, 它可以通过 Cholesky 分解得到; α 为尺度参数, 决定 Sigma 点的分布范围; β 表示 x 先验分布的因子 (对于高斯分布来说, 通常设为 2); W_i 是与第 i 个点相对应的权重, 上标 m 与均值相对应; c 与方差相对应。

4 Mean-Shift 跟踪窗口自适应算法

传统 Mean-Shift 算法的窗口是固定不变的, 在解决尺度变化目标的跟踪问题时效果不佳。当目标在不断运动时, 其背景也在不断的变化, 利用多尺度理论中的信息量计算方法时, 容易将不断变化的背景也当做目标信息进行统计, 使文献[3]的方法在使用中受到很大的限制。

本文提出的窗口自适应算法的主要思想是将多尺度理论统计出的信息量作为 UKF 滤波器的输入值, 对得到的信息量进行预测, 再使用修正后的信息量计算窗口变化比例系数 T 。

窗口更新系数 T 的具体计算步骤如下:

(1) 初始化帧数 $n=0$, 尺度变化系数 $T=0$, 取跟踪窗口尺度更新间隔帧数 N ;

(2) 读取一帧图像, $n=n+1$, 如果 $n=1$, 手动选定跟踪窗口, 按 2.3 节中的算法得到信息量 $I_1 \sim I_3$, 否则进入步骤(2);

(3) 如果帧数 n 是 N 的整数倍, 进入步骤(4), 否则 $T=0$, 转入步骤(2);

(4) 根据第 n 帧的跟踪窗口计算第 $n+N$ 帧窗口内的信息量, 得到 $I_4' \sim I_6'$ 。

利用 UKF 滤波器, 结合 $I_1 \sim I_3$, 将 $I_4' \sim I_6'$ 分别作为输入的状态量进行预测, 得到修正后的 $I_4 \sim I_6$, 再利用 2.3 节中的算法得到 T_{n+N} 更新数据: $I_1 = I_4$, $I_2 = I_5$, $I_3 = I_6$, 转到步骤(2)。

5 实验结果及分析

为验证本文提出的改进算法的有效性, 分别选取 55 帧连续变小的车辆图像和 95 帧连续变大的行人图像进行跟踪实验, 并每隔 3 帧计算一次窗口的变化比例系数。

图 1 给出了对逐渐变小车辆的跟踪中计算得到的跟踪窗内信息量的修正结果, 其中, “*” 表示采用文献[5]方法得出的跟踪窗内信息量值; 曲线表示利用 UKF 算法计算得出的修正值。

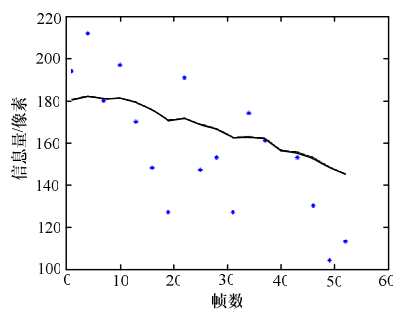


图1 连续变小车辆跟踪窗内信息量修正结果

图2给出了采用本文方法和文献[3]方法对图1结果中跟踪窗宽的比较,其中,“+”表示文献[3]方法得到的窗宽结果;曲线表示本文方法得到的窗宽结果。

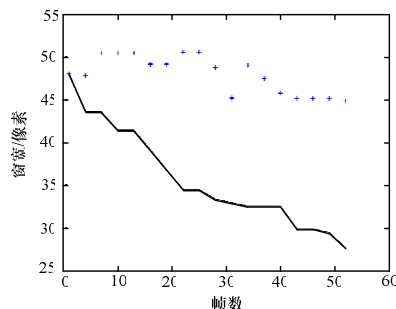


图2 连续变小车辆跟踪窗宽度比较

图3给出了对逐渐变大行人的跟踪中计算得到的跟踪窗内信息量的修正结果,其中,“*”表示采用文献[5]方法得出的跟踪窗内信息量值;曲线表示利用UKF算法计算得出的修正值。

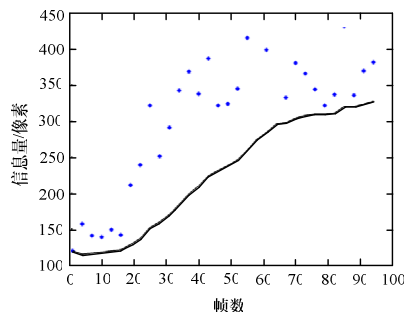


图3 连续变大行人跟踪窗内信息量修正结果

图4给出了采用本文方法和文献[3]方法对图3结果中跟踪窗宽的比较,其中,曲线表示本文方法得到的窗宽结果;“+”表示文献[3]方法得到的窗宽结果。

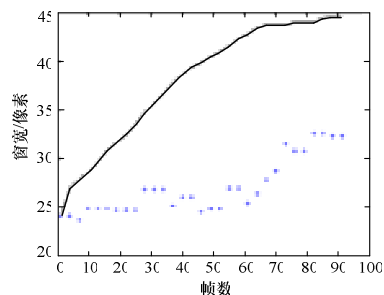


图4 连续变大行人跟踪窗宽比较

改进的 Mean-Shift 方法利用 UKF 算法对计算得到的信息量进行预测,利用修正后的信息量计算窗口更新系数。实验结果表明,这种方法在复杂背景条件下能够准确地计算出窗口更新系数,从而对尺度任意变化的目标进行有效的跟踪。本文算法对于光照变化是鲁棒的,因为跟踪窗口尺度的选取依赖于信息量的变化,而信息量仅与 2.1 节中的 2 类特征点相关,特征点都是式(2)的定义的方向微分算子的局部极值,因此光照的变化并不影响信息量的变化趋势。对于背景的变化,本文通过 UKF 算法来修正利用文献[3]方法计算得到的信息量,它是一种非线性滤波估计算法,可有效减少背景变化对于信息量计算的影响。

6 结束语

本文利用 UKF 滤波器解决多尺度空间的信息量度量问题,从而在一定程度上消除背景变化对计算窗口更新系数的影响。因为使用了 UKF 非线性滤波器,所以本文算法能够适应信息量的非线性变化,在相同条件下,能比线性滤波器获得更好的结果。但在实验过程中发现仍有许多方面需要进一步的研究和更加深入的探索,在进行窗口自适应选择时,本文应用多尺度信息量来计算窗口变化比例系数。在实际应用中,有很多因素都可以作为计算窗口变化比例系数的依据,例如像素值,互信息量等。如果能够找到一种计算简单,性能优良的因素作为窗口变化比例系数的计算依据,那么会进一步提高目标跟踪的精度并扩大它的应用范围。

参考文献

- [1] Comaniciu D, Ramesh V, Meer P. Kernel-based Object Tracking[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-575.
- [2] Zhu Shengli, Zhu Shan. Algorithm of Target Tracking Based on Mean Shift with Adaptive Bandwidth of Kernel Function[J]. Opto-Electronic Engineering, 2006, 33(8): 12-16.
- [3] 钱惠敏, 茅耀斌, 王执铨. 自动选择跟踪窗尺度的 Mean-Shift 算法[J]. 中国图象图形学报, 2007, 12(2): 246-249.
- [4] Sonka M, Hlavac V, Boyle R. Image Processing, Analysis, and Machine Vision[M]. 2nd ed. London, UK: Brooks/Cole Publishing, 1999.
- [5] 王郑耀, 程正兴, 汤少杰. 基于视觉特征的尺度空间信息度量[J]. 中国图象图形学报, 2005, 10(7): 922-928.
- [6] Julier S, Uhlmann J, Durrant-Whyte H F. A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators[J]. IEEE Trans. on Automatic Control, 2000, 45(3): 477-482.
- [7] Sunderhauf N, Lange S, Protzel P. Using the Unscented Kalman Filter in Mono-SLAM with Inverse Depth Parameterization for Autonomous Airship Control[C]//Proc. of IEEE International Workshop on Safety, Security and Rescue Robotics. Piscataway, New Jersey, USA: IEEE Press, 2007: 1-6.
- [8] 陈益, 赵高鹏, 刘娣. 简化 UKF 算法在摄像机标定中的应用[J]. 计算机工程, 2009, 35(19): 274-276.

编辑 金胡考