

基于 mean-shift 的快速跟踪算法设计与实现

孙 剑, 周凤岐, 周 军

(西北工业大学航天学院, 西安 710072)

摘要: 介绍了一种基于 mean-shift 跟踪算法, 分析了 mean-shift 中各参数对实时性能的影响。并利用 DM642EVM 作为视频信号处理器, 设计了一套运动目标全物理仿真实验系统。

关键词: 霍夫变换; 直线检测; 快速算法

Design and Realization of A Fast Tracing Algorithm Based on Mean-shift

SUN Jian, ZHOU Fengqi, ZHOU Jun

(College of Astronautics, Northwestern Polytechnic University, Xi'an 710072)

【Abstract】 This paper presents a moving object tracking algorithm. The tracing algorithm is based on the mean-shift tacking method which is based on the histograms of moving objects in consecutive image frames. And a physical simulation system for tracing moving objects is designed employing DM642EVM as video signal processor.

【Key words】 Hough transformation; Line detection; Fast algorithm

运动目标的检测与跟踪是应用视觉领域的一个重要的课题。目标跟踪涉及到图像处理、模式识别、人工智能以及自适应控制多领域问题的综合解决, 其中对于复杂背景(或环境)下的图像处理一直是视觉信息处理和理解的难点。在实际应用中, 复杂的现场环境以及跟踪应用的实时性要求, 对算法都提出了更高的要求。因此, 研究复杂背景下运动目标的检测和跟踪的课题有重要意义。本文的研究目的是建立一个视频图像监控系统, 根据目标的运动方向, 系统将自动控制摄像机转动, 将目标锁定在视野中。mean-shift技术的框架最早是由Fukunaga和Hostetler提出的。但直到1997年它的优良性能才被注意到^[1]。

1 基于 mean-shift 的跟踪算法

假设有 d 维空间 R^d 中的样本集合 $\{x_i\} i=1..n$, 在点 x 处, 使用核 $K(x)$, 窗半径 h 的多变量核密度估计由下式计算:

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (1)$$

一个核 K 的 profile 是 $[0, \infty) \rightarrow R$ 的函数 k , 它使得 $K(x) = k(\|x\|^2)$ 。

使用 profile 来表示多变量核密度估计, 即

$$\hat{f}_{h,K}(x) = \frac{1}{nd^d} \sum_{i=1}^n k\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \quad (2)$$

进而可以得到核密度梯度估计:

$$\hat{\nabla} f_{h,K}(x) \equiv \nabla \hat{f}_{h,K}(x) = \frac{2}{nh^{d+2}} \sum_{i=1}^n (x-x_i) k'\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \quad (3)$$

如果 $k(x)$ 在 $[0, \infty)$ 上除了有限个点外可导, 那么可以定义函数 $g(x) = -k'(x)$ 。由 $g(x)$ 可以导出新的核 $G(x) = Cg(\|x\|^2)$, C 是归一化因子。使用该核的多变量核密度估计为

$$\hat{f}_G(x) = \frac{C}{nd^d} \sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \quad (4)$$

可以推导出(推导过程请参考文献[1]):

$$\hat{\nabla} f_{h,K}(x) = \hat{f}_{h,G}(x) \frac{2/C}{h^2} M_{h,G}(x) \quad (5)$$

其中

$$M_{h,G}(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \quad (6)$$

被称为 mean-shift 向量, 即使用核 G 作为权值的加权平均, 与 x (核的中心) 的差。 $k(x)$ 称为 $g(x)$ 的影子 (shadow), 式(5)稍加变形, 就有

$$m_{h,G}(x) = \frac{1}{2} h^2 C \frac{\hat{\nabla} f_{h,K}(x)}{\hat{f}_{h,G}(x)} \quad (7)$$

上式表明在 x 处, 使用核 G 计算的 mean-shift 向量正比于使用核 K 得到的归一化的密度梯度估计。归一化要依靠使用核 G 计算的 x 处的密度估计。因此 mean-shift 向量总是指向密度增大的最大方向。

设 $\{x_i^*\}_{i=1..n}$ 目标所在区域的归一化的点集, 区域以 0 为中心。一个各向同性, 具有单调递减凸 profile 函数的核, 为各点加权, 点离中心越远, 权值越小。由于靠外的像素值最不可靠, 容易被遮掩或背景影响, 这样的加权增强了密度估计的健壮性。定义函数 $b: R^2 \rightarrow \{1..m\}$, 对于每个像素 x_i^* , $b(x_i^*)$ 是该像素特征在量化的特征空间中的量化序号。如果点的坐标已按照 (h_x, h_y) 归一化, 则特征 $u = 1..m$ 在目标模式中出现的概率 $\hat{q}_u = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u]$, δ 是 Kronecker 函数, 常数 C 是归一化函数, 使得 $\sum_{u=1}^m \hat{q}_u = 1$, 则

作者简介: 孙 剑(1976—), 男, 博士生, 主研方向: 数字伺服, 图像处理; 周凤岐、周 军, 教授、博导

收稿日期: 2006-01-01 **E-mail:** sjxa@163.com

$$C = \frac{1}{\sum_{i=1}^n k(\|x_i\|^2)}$$

设候选目标模式为 $\{x_i\}_{i=1, \dots, n_h}$ ，以 y 为中心。使用相同的核 profile，半径 h ，特征 $u = 1 \dots m$ 在候选目标当中出现的概率为

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_h} k(\|\frac{y-x_i}{h}\|^2) \delta[b(x_i) - u] \quad (8)$$

C_h 是归一化因子，使得 $\sum_{u=1}^m \hat{p}_u = 1$ ，

$$C_h = \frac{1}{\sum_{i=1}^{n_h} k(\|\frac{y-x_i}{h}\|^2)}$$

h 决定了候选目标的尺度。 \hat{q}_u 和 $\hat{p}_u(y)$ 就是描述目标和候选目标的特征。

使用 Bhattacharyya 系数：

$$\rho(y) = \rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (9)$$

来描述目标和候选目标的相似程度，该系数越大，目标和候选目标也越相似。定义它们之间的距离为

$$d(y) = \sqrt{1 - \rho[\hat{p}(y), \hat{q}]} \quad (10)$$

当有了目标及候选目标的描述，并且有了衡量它们相似程度（距离）的准则后，目标跟踪就变成了在当前帧中搜索对应于目标的新位置，该位置使得以 y 为自变量的距离函数取得最小值。搜索从前一帧中目标的位置开始，在其周围邻域内寻找。

最小化式(10)，等价于最大化式(9)，假设前一帧中目标的位置为 \hat{y}_0 ，首先应计算出当前帧中 \hat{y}_0 位置处候选目标的 $\{\hat{p}_u(\hat{y}_0)\}_{u=1, \dots, m}$ ，如果将 $\rho[\hat{p}(y), \hat{q}]$ 在目标预测位置 \hat{y}_0 处展开，可以获得 $\rho[\hat{p}(y), \hat{q}]$ 的线性近似：

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_0) \hat{q}_u} + \frac{1}{2} \sum_{u=1}^m \hat{p}_u(y) \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}} \quad (11)$$

该近似在候选目标 $\{\hat{p}_u(y)\}_{u=1, \dots, m}$ 与初始的 $\{\hat{p}_u(\hat{y}_0)\}_{u=1, \dots, m}$ 相差不大时，是令人满意的，一般说来相邻帧间这样的假设是合理的。应该强制使条件 $\hat{p}_u(\hat{y}_0) > 0$ 成立，如果某个 u 对应的 $\hat{p}_u(\hat{y}_0)$ 不满足，就不使用它的 $\hat{p}_u(\hat{y}_0)$ 。代式(8)入式(11)可以得到：

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_0) \hat{q}_u} + \frac{C_h}{2} \sum_{i=1}^{n_h} w_i k(\|\frac{y-x_i}{h}\|^2) \quad (12)$$

其中

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}} \quad (13)$$

这样，最小化 $d(y)$ ，就变成最大化式(12)中的第 2 项，而该项代表着当前帧中 y 位置处的使用 $k(x)$ 计算的密度估计，只不过数据多了加权值 w_i 。这样就可以使用 mean-shift 过程寻找邻域内该密度估计的极大值(模式)。在这个过程中，核从当前位置 \hat{y}_0 移向新位置 \hat{y}_1 ：

$$\hat{y}_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g(\|\frac{\hat{y}_1 - x_i}{h}\|^2)}{\sum_{i=1}^{n_h} w_i g(\|\frac{\hat{y}_1 - x_i}{h}\|^2)} \quad (14)$$

其中 $g(x) = -k'(x)$ ， $k(x)$ 的导数在 $x \in [0, \infty)$ 上除了有限个点外都存在。

2 对跟踪算法的分析

影响 mean-shift 收敛速度主要有以下几个方面：(1) 特征 \hat{q}_u 的选取即对灰度级的划分，灰度级划分地越细，计算速度越慢，相对跟踪效果越好，抗干扰能力越强。反之则收敛速度变快抗干扰能力减弱。(2) 核函数 $g(x)$ 选取的不同计算耗时也不同。(3) 窗口 h 的大小， h 的大小会影响跟踪器的带宽，在不考虑计算耗时的影响下， h 越大则跟踪器的带宽

越大，相反则会变小。但是， h 越大则计算速度越慢。(4) 在每一次迭代过程中 w_i 之间的差对均值平移的速度影响很大，差值越大则平移的速度越快。随着迭代次数的增加 w_i 将趋近于 1。这时平移步长变得很小趋近于 0。

通过对上面的分析，发现对于不同的跟踪对象需要在跟踪精度，抗干扰能力和实时性方面做必要的折中以找到合适的参数。其中 \hat{q}_u ， $g(x)$ ， h 的选取都是在实时性和跟踪精度，抗干扰能力 3 方面折中的结果。

3 模版匹配算法与 mean-shift 算法的比较

模版匹配算法又称相关匹配跟踪算法，其最一般的数学表达式如下：

$$D(i, j) = \frac{1}{M^2} \sum_{l=1}^M \sum_{m=1}^M |S(i+l, j+m) - W(l, m)| \quad (15)$$

算法的实质是在图像找到一块区域，这块区域与模版之间的对应像素灰度差的绝对值的和为最小。mean-shift 算法是利用目标图像像素灰度距离中心点的距离和作为特征。将候选目标的特征与模版特征进行比较，再通过 mean-shift 过程使得跟踪算法收敛于平衡点。与模版匹配法相比，mean-shift 跟踪算法有以下几个优点。

(1) 模版匹配算法是在整幅图像上搜索出与原模版最匹配的区域，耗时很大。

而 mean-shift 算法是假设目标在下一帧的位置距离当前位置有一个很小的距离，利用迭代的算法收敛到平衡点，所以相对耗时较少。

(2) 由于 mean-shift 算法对各像素点灰度按到中心点的距离加权，距离中心点越远，权值越小，通常目标的边缘（距离中心点较远）最容易受到干扰，所以，mean-shift 算法的抗干扰能力要比模版匹配要强。

(3) mean-shift 算法中对跟踪框进行了归一化处理，在目标的形状发生缩放，旋转的情况下，仍然得到很好的跟踪效果。

(4) 模版匹配相关算法对于刚体目标具有很好的跟踪性能，但是对于非刚体目标，鲁棒性则很差。而 mean-shift 算法则对刚体和非刚体都有很好的跟踪效果。

4 运动目标跟踪全物理仿真实验系统

硬件结构：运动目标跟踪全物理仿真系统结构框如图 1 所示。图像传感器采用 MINTRON 1881EX 型摄像头。图像信号处理采用闻亭公司的 DM642EVM 视频图像处理平台，二轴转台采用四通步进电机做执行机构，电机控制芯片选用 TI 公司生产的 TMS320LF2407A。该芯片可以同时控制两路步进电机工作。

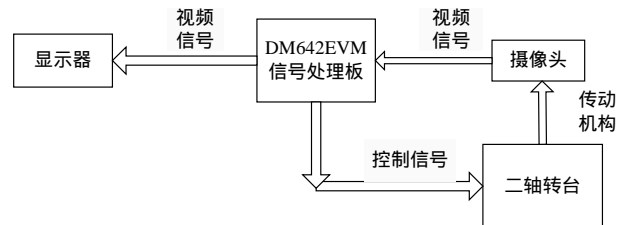


图 1 运动目标跟踪全物理仿真实验系统硬件结构

软件结构：系统软件系统结构如图 2。为了提高系统的实时性，对于核心算法部分均采用汇编语言编写。图 3 是实时跟踪系统所采集的视频图像中的一帧图像。其中后面的框示意 mean-shift 初始迭代的位置，前面的框示意收敛的平衡位置。

(下转第 77 页)