

# Mean-Shift 跟踪算法中核函数窗宽的自动选取\*

彭宁嵩<sup>1,2+</sup>, 杨杰<sup>1</sup>, 刘志<sup>1</sup>, 张风超<sup>1</sup>

<sup>1</sup>(上海交通大学 图像处理与模式识别研究所,上海 200030)

<sup>2</sup>(河南科技大学 电子与信息学院,河南 洛阳 471039)

## Automatic Selection of Kernel-Bandwidth for Mean-Shift Object Tracking

PENG Ning-Song<sup>1,2+</sup>, YANG Jie<sup>1</sup>, LIU Zhi<sup>1</sup>, ZHANG Feng-Chao<sup>1</sup>

<sup>1</sup>(Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200030, China)

<sup>2</sup>(Institute of Electronics and Information, He'nan University of Science and Technology, Luoyang 471039, China)

+ Corresponding author: Phn: +86-21-62934243, Fax: +86-21-62932035, E-mail: pengningsong@sjtu.edu.cn

Received 2004-05-02; Accepted 2004-10-09

**Peng NS, Yang J, Liu Z, Zhang FC. Automatic selection of kernel-bandwidth for Mean-Shift object tracking. *Journal of Software*, 2005,16(9):1542–1550. DOI: 10.1360/jos161542**

**Abstract:** Classic Mean-Shift based tracking algorithm uses fixed kernel-bandwidth, which limits the performance when the object scale exceeds the size of the tracking window. Based on the analysis of similarity of object kernel-histogram in different scales, i.e. the Bhattacharyya coefficient, a theorem is found and proved i.e. the changes of object scale and position within the kernel will not impact localization accuracy of Mean-Shift based tracking algorithm. Using this theorem an automatic bandwidth selection method is proposed based on backward tracking and object centroid registration. The proposed method is applied to track vehicle changing in size with encouraging results.

**Key words:** Mean-Shift; object tracking; kernel-bandwidth selection; Bhattacharyya coefficient; affine model

**摘要:** 传统核窗宽固定的 Mean-Shift 跟踪算法不能很好地对逐渐增大尺寸的目标进行有效的跟踪.在分析同一目标在不同尺度下核直方图基于 Bhattacharyya 系数相似性的基础上,发现并证明了在核窗宽固定的条件下,目标在其窗宽范围内进行缩放、平移运动并不影响 Mean-Shift 跟踪算法空间定位的准确性.在此基础上,提出了一种基于后向跟踪、形心配准的核窗宽自动选取算法.对尺度渐大的车辆进行的跟踪实验验证了该算法的有效性.

**关键词:** Mean-Shift;目标跟踪;核窗宽选取;Bhattacharyya 系数;仿射模型

中图法分类号: TP391 文献标识码: A

目标跟踪广泛应用于监控、视频编码以及军工领域.如何有效地在视频序列中对感兴趣目标进行有效跟踪,是计算机视觉中一个极具挑战性的课题.Mean-Shift<sup>[1,2]</sup>作为一种高效的模式匹配算法,由于不需要进行穷尽

\* Supported by the National Natural Science Foundation of China under Grant No.30170274 (国家自然科学基金)

作者简介: 彭宁嵩(1973 - ),男,辽宁沈阳人,博士生,讲师,主要研究领域为计算机视觉,图像处理,模式识别;杨杰(1964 - ),男,教授,博士生导师,主要研究领域为数据挖掘,目标检测与识别;刘志(1978 - ),男,博士生,主要研究领域为计算机视觉,图像处理;张风超(1975 - ),男,博士生,主要研究领域为视频分析,模式识别.

搜索,已经成功地应用在对实时性要求较高的目标跟踪领域中<sup>[3-6]</sup>.在 Mean-Shift 跟踪算法中,核窗宽的大小起着非常重要的作用.因为它不但决定了参与 Mean-Shift 迭代的样本数量,而且也反映了跟踪窗口的大小.通常,核窗宽由初始跟踪窗口的尺寸决定,而且在整个跟踪过程中不再发生变化<sup>[3,4]</sup>.然而,当目标存在明显尺度变化的时候,尤其是当目标尺寸逐渐增大以至超出核窗宽范围的时候,固定不变的核窗宽常常会导致目标的丢失.目前,还没有很好的方法用来解决核窗宽的自动选取问题.文献[5]提出基于不变矩的方法,但是矩特征的计算严重影响了 Mean-Shift 跟踪算法的实时性.文献[6]采用正负 10% 的增量分别对核窗宽进行修正.该方法需要在当前帧中用 3 个不同大小的核进行 3 次独立的 Mean-Shift 跟踪计算,选择出较大的 Bhattacharyya 系数所对应的窗宽为最佳核窗宽.当目标逐渐缩小尺寸时,该方法可以得到较好的效果.但是,当目标逐渐增大尺寸时,核窗宽很难被扩大,反而经常越变越小.这是因为基于 Bhattacharyya 系数的相似性度量经常会在较小的跟踪窗口中达到局部最大<sup>[7]</sup>.Collins 提出通过增加一个额外的尺度核,在定义的离散尺度空间中进行 Mean-Shift 迭代,从而找到最佳的核窗宽<sup>[7]</sup>.然而,由于所采用的 Epanechnikov 核的导数为常数<sup>[1]</sup>,因此 Mean-Shift 迭代过程等效于在空间定位的基础上对尺度空间进行平均操作.所以,该方法在本质上和文献[6]中提出的方法一样,具有类似的缺陷.

我们在分析同一目标在不同尺度下的核直方图基于 Bhattacharyya 系数相似性的基础上,发现并证明了如下定理:在固定核窗宽的条件下,目标在其窗宽范围内进行缩放、平移运动并不影响 Mean-Shift 跟踪算法空间定位的准确性.在此基础上,提出了一种核窗宽自动选取算法,以便对不断增大尺寸的刚性目标进行有效的跟踪.假设目标在连续帧中的运动满足仿射模型.如果我们可以准确地找出仿射模型中的伸缩幅值,那么核窗宽就可以根据此伸缩幅值进行更新,从而减小尺度定位的误差.然而,求解仿射模型参数往往需要进行特征点对的回归计算.误匹配的特征点对会严重影响回归的精度<sup>[8]</sup>.利用发现的定理,首先,我们采用后向跟踪的方法对连续两帧图像中的目标形心进行配准.然后,在配准的基础上抽取特征点对,并进行回归计算.这样可以有效地消除误匹配,保证了回归的精度.

本文第 1 节首先简要介绍 Mean-Shift 跟踪算法,在分析该算法缺陷的同时,给出我们所发现的定理.第 2 节详细介绍后向跟踪和目标形心配准的方法,以及如何有效地选择特征点对.实验结果在第 3 节给出.最后是结论和展望.

## 1 Mean-Shift 跟踪算法性能分析

### 1.1 Mean-Shift跟踪算法

设  $A$  是嵌入在  $n$  维欧式空间  $X$  中的有限集合.在  $x \in X$  处的 Mean-Shift 矢量定义为<sup>[2]</sup>

$$ms = \frac{\sum_a k(a-x)w(a)a}{\sum_a k(a-x)w(a)} - x, \quad a \in A \quad (1)$$

其中,  $k$  是核函数,  $w$  是权重.在  $x$  处计算出的 Mean-Shift 矢量  $ms$  反向指向卷积曲面

$$J(x) = \sum_a g(a-x)w(a) \quad (2)$$

的梯度方向<sup>[1,2]</sup>,其中  $g$  是  $k$  的影子核<sup>[2]</sup>.沿着  $ms$  方向不断移动核函数中心位置直至收敛就可以找到邻近的模式匹配位置<sup>[1,2]</sup>.

为了下面叙述方便,我们首先给出以下几个定义.

**定义 1.** 在一帧图像中,目标所在的图像区域称为目标区域,用  $F$  表示.  $F$  以外的图像区域称为背景区域,用  $B$  表示.包含  $F$ ,且面积最小的圆形区域的圆心称为目标的形心.称同时包含目标图像区域  $F$  和  $B$  的圆形区域  $T$  为跟踪窗口.假设  $F$ ,  $B$  各自对应的颜色直方图中的非零子项位置不重合,也即目标与背景有着明显的颜色差异.在车辆监控等很多实际场合,上述假设基本上都是成立的.

**定义 2.** 给定跟踪窗口  $T$ , 设  $\{X_i\}_{i=1,\dots,n}$  是以其圆心为原点的像素坐标,则  $T$  包含图像的核直方图  $P = \{p_\mu\}_{\mu=1,\dots,m}$  定义为  $p_\mu = C \sum_{i=1}^n k(\|X_i/r\|^2) \delta[q(X_i) - \mu]$  <sup>[3]</sup>,其中,  $\delta$  是 Kronecker delta 函数.映射  $q: R^2 \rightarrow \{1,\dots,m\}$  把相应位置处像素的颜色进行  $m$  级量化.  $C$  为归一化常数.通过约束条件  $\sum_{\mu=1}^m p_\mu = 1$  可得  $C =$

$1/\sum_{i=1}^n k(\|X_i/r\|^2)$ .  $r$  称为核函数  $k$  的窗宽,同时也是跟踪窗口  $T$  的半径.

定义 3. 两个具有  $m$  个分量的核直方图  $P_i$  和  $P_j$  的相似性用 Bhattacharyya 系数  $\rho = \sum_{l=1}^m \sqrt{p_l^i p_l^j}$ ,  $i \neq j$  表示. 其中,  $p_l^i$  和  $p_l^j$  分别是两个核直方图中对应分量的值.

当目标图像采用核直方图建模时,给定模型与候选图像,核直方图之间的相似性可以通过 Bhattacharyya 系数来度量. Comaniciu 等人基于上述思想设计出了相应的  $J(x)$ , 并证明了 Mean-Shift 迭代的收敛性<sup>[3]</sup>, 从而使跟踪问题转化为 Mean-Shift 模式匹配问题.

## 1.2 “固定核窗宽-可变目标尺度”条件下核直方图间 Bhattacharyya 系数分析

采用固定核窗宽的 Mean-Shift 跟踪算法<sup>[3,4]</sup>在跟踪存在尺度变化的目标时有以下特点:(1) 当目标不断增大尺寸并且大于核窗宽时,不仅会导致尺度定位偏差,同时也导致空间定位偏差<sup>[7]</sup>. 在这种情况下,如何兼顾尺度、空间定位的准确性,是本文要解决的主要问题,即跟踪窗口的尺寸在适应目标尺寸增大的同时,使其中心对准目标的形心.(2) 当目标不断缩小尺寸时,由于核窗宽固定,即跟踪窗口大小不变,最终必定会导致极大的尺度定位误差. 虽然跟踪窗口在包含目标区域的同时逐渐混入了很多的背景区域,但是,跟踪窗口的中心始终指示着目标的形心. 也就是说,空间定位误差很小.

图 1 给出了两段视频序列. 其中,跟踪窗口(白色圆圈)的大小始终是不变的,即核窗宽固定. 第 1 行序列是目标不断增大尺寸的跟踪结果,第 2 行是目标缩小尺寸的跟踪结果. 在每行的第 1 帧我们都用初始跟踪窗口锁定目标车辆,使其形心与窗口中心重合.

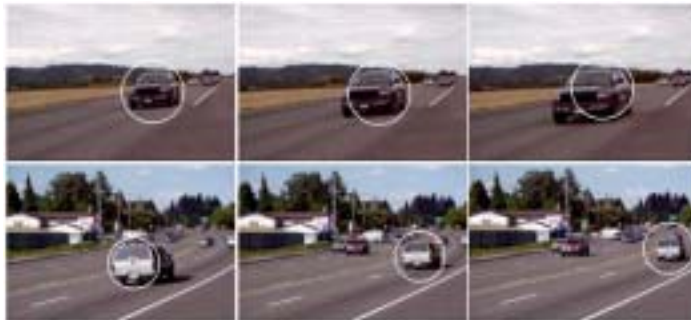


Fig.1 Mean-Shift tracking results with fixed kernel-bandwidth (from left to right)

图 1 固定核窗宽条件下 Mean-Shift 跟踪效果(从左到右)

图 2 表示的是以目标形心  $40 \times 40$  邻域范围内的每一个点为中心的跟踪窗口所对应的 Bhattacharyya 系数. 其中,图 2(a)和图 2(b)分别对应图 1 的第 1 行和第 2 行中间的那一帧图像. 图 2(a)中的曲面存在多处局部极值,这是因为跟踪窗口湮没在目标图像中,造成在多个位置处存在相似的 Bhattacharyya 系数. 以上原因也是导致 Mean-Shift 跟踪算法出现空间定位偏差的主要原因. 因此,从图 1 的第 1 行我们看到,跟踪窗口中心不能始终锁定目标的形心. 在图 1 第 2 行的序列中,被跟踪的车辆逐渐远离镜头,尺寸在逐渐变小. 可以清楚地看到,跟踪窗口的中心始终指向目标的形心. 从图 2(b)中我们发现,某位置处的 Bhattacharyya 系数是其与目标形心距离的单调递减函数,且在目标形心处达到最大值. 我们可以用下面的定理来描述这种现象.

定理. 假定两帧图像  $i$  和  $i+1$  中存在有不同尺度的同一个目标. 给定 3 个半径相同的跟踪窗口  $T_1, T_2$  和  $T_3$ , 它们的核直方图分别为  $P_1, P_2$  和  $P_3$ . 若满足:(1)  $T_1$  是图像  $i$  中的跟踪窗口,  $T_2$  和  $T_3$  是图像  $i+1$  中的跟踪窗口;(2)  $F_1$  形心与  $T_1$  圆心重合,  $F_2$  形心与  $T_2$  圆心的距离小于  $F_3$  形心与  $T_3$  圆心的距离,则  $P_1$  与  $P_2$  的 Bhattacharyya 系数大于与  $P_3$  的 Bhattacharyya 系数.

该定理的证明见附录. 该定理表明,在第  $i+1$  帧中,以目标形心为中心的跟踪窗口(设为  $T'_1$ )所包含的图像与  $T_1$  中的图像最相似. 由 Mean-Shift 理论<sup>[1,2]</sup>,只要在帧  $i+1$  中目标的伸缩、平移均处于在  $T_1$  的范围内(核窗宽范围内),则通过 Mean-Shift 迭代可以求得帧  $i+1$  中最佳的匹配窗口  $T'_1$ . 也就是说,跟踪窗口的中心将始终锁定目标

的形心.实验表明,在大多数情况下,只要目标的大部分区域能出现在  $T_i$  范围内,Mean-Shift 跟踪算法则均能锁定目标中心.

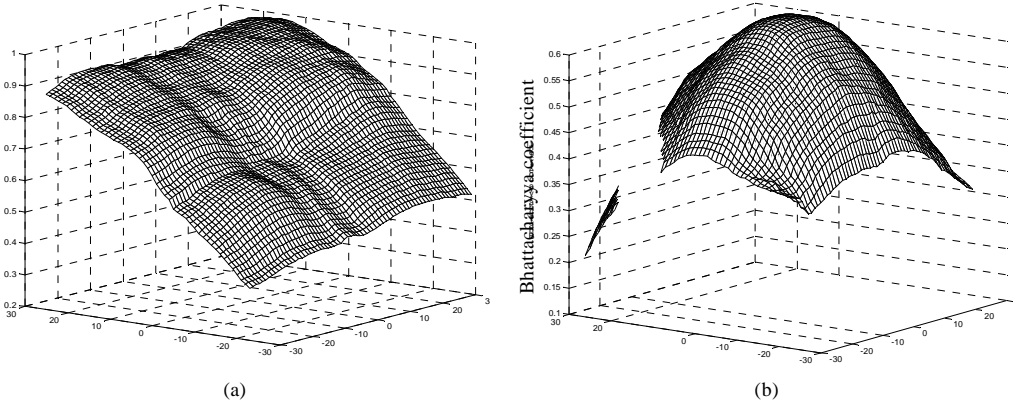


Fig.2 Bhattacharyya coefficient of the windows around the object centroid

图2 目标中心附近窗口的 Bhattacharyya 系数

当目标不断增大尺寸且大于核窗宽时,Mean-Shift 跟踪算法会产生空间定位偏差.在得到这个偏差位置以后,我们可以逆向进行处理,即后向跟踪.这时,视频逆序列就等效为目标在逐渐缩小尺寸.这样,Mean-Shift 跟踪算法可以准确锁定逆向帧中相应的偏差点,这使得我们补偿这个偏差成为可能.下一节我们将详细介绍如何进行目标形心配准补偿空间定位偏差,以及在此基础上如何选取核窗宽消除尺度定位偏差.

## 2 核窗宽的自动选取

基于连续帧中刚性运动物体满足仿射模型的假设,首先,利用后向跟踪的方法对目标形心进行配准补偿空间定位偏差.在配准的基础上,相邻帧中跟踪窗口里面的特征点坐标被归一化到以目标形心为原点的坐标系中.这样,与直接从两个未配准的跟踪窗口中匹配特征点相比,我们的方法可以有效地减少误匹配,从而为后面精确估算目标仿射模型中的伸缩幅值奠定了良好的基础.最后,核窗宽依据伸缩幅值进行更新.

### 2.1 仿射模型和特征点匹配

这里只考虑两种在实际中经常遇到的运动:平移和伸缩.因此,目标的仿射模型由以下形式给出:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_x \\ e_y \end{pmatrix} \quad (4)$$

其中,  $(x, y)$  和  $(x', y')$  分别是同一个目标特征点在帧  $i$  和  $i+1$  中的位置.  $e = \{e_x, e_y\}$  是平移参数,  $s = \{s_x, s_y\}$  是目标水平、垂直方向的伸缩幅值.利用  $s$ , 核窗宽可以按照如下方法进行更新:

$$r = r \cdot \max(s_x, s_y) \quad (5)$$

对于刚性物体,角点能够很好地刻画其空间结构,而且易于检测.因此,我们采用两帧中匹配的角点作为样本对仿射模型参数进行估算.

假定帧  $i$  中的跟踪窗口  $T_i$  里有  $N$  个角点,帧  $i+1$  中的跟踪窗口  $T_{i+1}$  里面有  $N'$  个角点.  $T_i$  和  $T_{i+1}$  半径相同,它们的中心均与对应目标的形心重合.也就是说,目标的形心在相邻两帧里已经过了配准.给定一个位于  $T_i$  中的角点  $P_c|T_i$ ,在  $T_{i+1}$  中,它的对应角点  $P_c|T_{i+1}$  应当满足:

$$I(P_c|T_{i+1}) = \min \{ |I(P_j|T_i) - I(P_j|T_{i+1})| \}_{j=1,2,\dots,n} \quad (6)$$

其中,  $I$  是像素亮度.  $n$  个候选点位于帧  $T_{i+1}$  中以  $P_c|T_i$  位置为中心的一个给定的小窗口  $G$  里面.由于经过了配准和归一化处理,  $G$  的尺寸可以取很小从而使得  $n < N'$ .对  $T_i$  中的每一个角点  $P_c|T_i$ ,我们用式(6)找到它在帧  $T_{i+1}$  中的匹配点.显然,与具有时间复杂度  $O(NN')$  的最大似然模板匹配方法<sup>[8,9]</sup>相比,我们这种局部匹配的方法,其时间

复杂度是  $O(Nn)$ , 实现起来也非常容易.

## 2.2 目标形心配准

在这一节里,我们将详细叙述如何对相邻两帧的目标形心进行配准.

假定在帧  $i$  中,以  $o_i$  为形心的目标被以  $c_i = o_i$  为中心的初始跟踪窗口  $T_i$  选定.当增大尺寸的目标出现在帧  $i+1$  时,针对当前的目标形心  $o_{i+1}$ ,应当存在偏移  $d = c_{i+1} - o_{i+1}$ ,其中  $c_{i+1}$  是帧  $i+1$  里面的跟踪窗口  $T_{i+1}$  的中心,这个偏差是由固定核窗宽的 Mean-Shift 跟踪算法造成的空间定位偏差.因为当前目标的尺寸大于  $T_{i+1}$  的尺寸,所以  $T_{i+1}$  内只包含了目标的某些部分.为了配准两帧中目标的形心,首先,我们生成一个新的核函数直方图用来表示被  $T_{i+1}$  包含的这部分图像,即部分当前目标.实际上,  $c_{i+1}$  指示的就是这个部分目标的形心.从帧  $i+1$  到  $i$ ,这块图像区域在缩小尺寸,我们就可以精确地在帧  $i$  中用 Mean-Shift 跟踪算法找到它的形心  $c'_{i+1}$ .这样,在  $c_i$  和  $c'_{i+1}$  之间,存在另外一个偏移  $d' = c_i - c'_{i+1}$ .假定相邻帧间的运动很小,我们就可以用  $d'$  近似来补偿  $d$ .最后,帧  $i+1$  中目标的形心位置的估算如下:

$$o_{i+1} \approx c_{i+1} - d' \quad (7)$$

因此,当给定一个目标图像以后,我们就可以利用式(7)在当前帧中对目标形心进行配准,然后再用式(6)匹配角点.这种处理方法可以有效地消除误匹配点.图3是通过后向跟踪对目标形心进行配准的示意图.

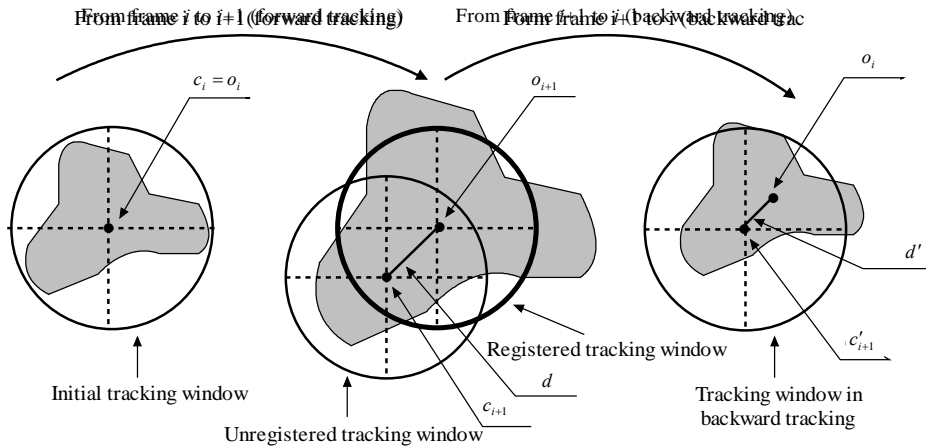


Fig.3 Illustrations of backward tracking and object centroid registration

图3 后向跟踪和目标形心配准示意图

## 2.3 算法描述

在通过回归计算得到目标仿射模型以后,就可以利用公式(5)对当前核窗宽进行更新.更新后的核窗宽一方面用于修正当前跟踪窗口的尺寸从而减小尺度定位偏差,另一方面,在下一帧跟踪中用来决定 Mean-Shift 迭代过程中样本的数量.这样,系统就能很好地适应目标尺度的变化,克服了固定核窗宽的限制性.具体算法如下:

算法. 核窗宽自动选取.

Begin

1. 在帧  $i$  选定目标得到初始跟踪窗口  $T_i$  并在帧  $i+1$  中进行 Mean-Shift 跟踪得到  $T_{i+1}$ ;
2. 以  $T_{i+1}$  为初始跟踪窗口在帧  $i$  中进行 Mean-Shift 跟踪得到  $T'_i$ ;
3. 根据  $T_i$  与  $T'_i$  中心点的位置差移动  $T_{i+1}$ , 并扩大其尺寸:  $r = r \cdot \varepsilon, \varepsilon > 1$ ;
4. 从  $T_i, T_{i+1}$  中提取角点, 进行匹配;
5. 分别对匹配点的横、纵坐标进行回归得到  $s_x$  和  $s_y$ ;
6. 用  $r = r \cdot \max(s_x, s_y)$  更新  $T_{i+1}$  大小.

End

算法第 3 步中的  $\varepsilon$  是为了扩大配准后的跟踪窗口尺寸,以便尽可能多地得到匹配点对.鉴于实际应用中多使用矩形区域作为显示窗口.此时核函数的作用域(跟踪窗口范围)就是该矩形的外接圆区域.这时,算法第 6 步应更改为  $r = r\sqrt{(s_x)^2 + (s_y)^2}$ .

### 3 实验结果

我们对不断增大尺寸车辆的视频序列进行了跟踪测试.视频序列中每一帧均为位图格式的彩色图像.我们对 RGB 分量分别进行 32 级量化处理得到的核直方图大小为  $32 \times 32 \times 32$ .图 4 是基于 Mean-Shift 理论的跟踪算法的跟踪效果比较.第 1 行序列采用的是固定窗宽.随着目标尺寸的增加,跟踪窗口湮没于其中.不但造成了尺度上的偏差,同时也没有对目标形心进行准确定位;第 2 行是采用正负 10% 的增量估算核窗宽的方法.不但空间定位效果不佳,而且跟踪窗口尺寸甚至还在变小;最后一行是我们提出的后向跟踪、形心配准的方法.从中可以明显看出,我们的方法同时具备了尺度、空间上的定位准确性.算法的具体参数设置见表 1.有关角点提取的算子和参数可参见文献[9].

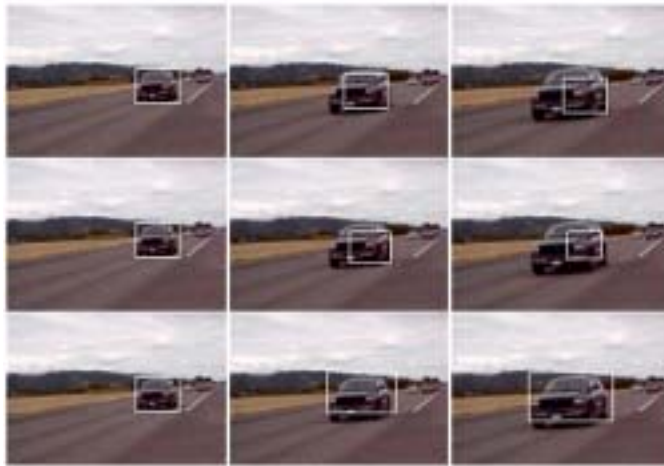


Fig.4 Comparison of Mean-Shift tracking results (left to right)

图 4 Mean-Shift 跟踪效果比较(从左到右)

Table 1 Parameters of our method

表 1 本文提出算法的参数

$\sigma$ value in corner operator	Threshold in corner operator	Size of the $G$	Max iterates in Mean-Shift
1	2 000	0.1	20

图 5(a)和图 5(b)分别对应图 4 最后一行的前两帧.图 5(b)中的虚线框是用固定核窗宽进行跟踪的结果,配准后的跟踪窗口见细线框所示.图 5(c)~图 5(e)分别显示了从初始跟踪窗口、未配准的跟踪窗口和配准后的跟踪窗口中提取的角点.在图 6 中,所有的角点坐标被归一化放置在以跟踪窗口中心为原点的坐标系里面,其中每条线段两端的点表示匹配点对.图 6(a)显示的是从初始跟踪窗口中提取的角点(+号)和从未配准跟踪窗口中提取的角点(点号)的匹配关系.图 6(b)显示的是从初始跟踪窗口中提取的角点(+号)和从配准的跟踪窗口中提取的角点(点号)的匹配关系.与图 5(a)相比,很明显,图 5(b)中的线段几乎都有着整齐的形态,即几乎所有的线段都是向外伸展的(从+号点指向点号点).这表明,一方面,误匹配点对通过形心配准的方法得到了显著消除;另一方面,明显地表示出目标的尺度膨胀趋势.利用普通的最小二乘法对图 6(b)中的匹配点进行回归可以较精确地求得缩放幅值.图 5(b)中的粗线框就是经过更新后的核窗宽对应的新的跟踪窗口.

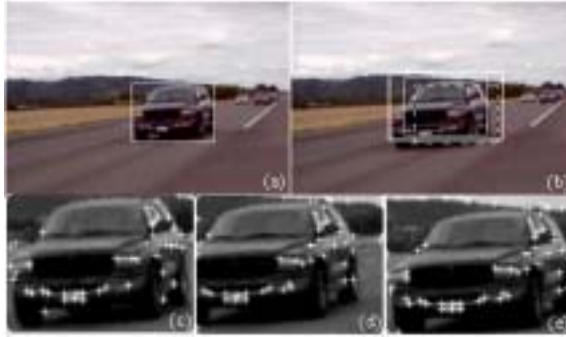


Fig.5 Tracking windows and the corners extracted from them

图 5 跟踪窗口和从中提取的角点

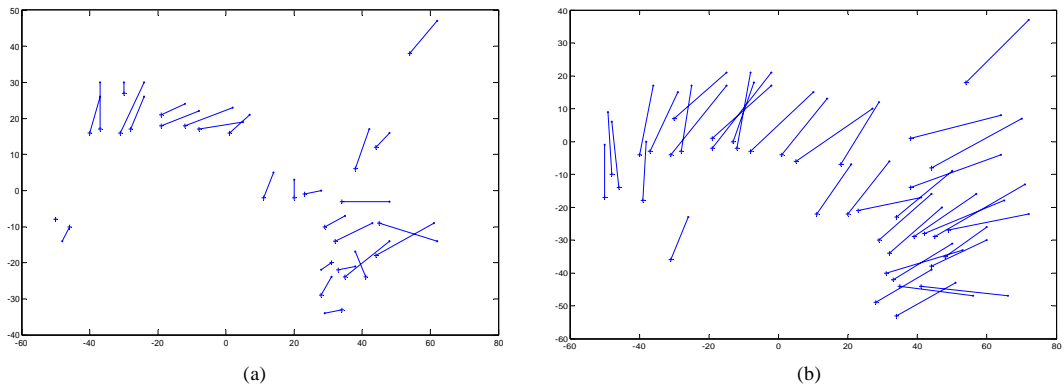


Fig.6 Comparison of corner correspondences

图 6 匹配角点比较

#### 4 结论和展望

Mean-Shift 跟踪算法以其出色的搜索效率、无须参数的初始化过程以及对边缘遮挡鲁棒的性能较好地兼顾了目标跟踪算法对实时性和稳健性的要求。然而,在目标不断增大尺寸的情况下,固定的核窗宽严重影响了跟踪效果。我们在“固定核窗宽-可变目标尺度”的情况下对跟踪窗口在不同位置处的核函数直方图的相似性,即 Bhattacharyya 系数进行了研究,发现并证明了在核窗宽固定的条件下,目标在其窗宽范围内进行缩放、平移运动并不影响 Mean-Shift 目标跟踪算法的定位准确性。利用以上发现,提出了后向跟踪、形心配准的思想。克服了使用固定核窗宽的 Mean-Shift 跟踪算法的一些缺陷。

对于存在有剧烈全局运动、低帧率、目标剧烈运动的视频序列。当选定了初始跟踪窗口以后,当前帧在相同位置处的窗口很可能不包含所要跟踪的目标。这必定会导致 Mean-Shift 跟踪算法丢失目标<sup>[4]</sup>。我们可以先对当前帧做降分辨率处理,使图像缩小尺寸<sup>[10]</sup>。这样,发生剧烈运动的目标可能被包含在未改变尺寸的跟踪窗口当中。由于图像尺寸缩小,目标尺寸也同比例缩小。利用本文中的定理就可以在降分辨率的帧图像中准确定位目标,然后再把锁定的位置恢复到原分辨率坐标下完成跟踪。在下一步的工作中,我们将努力解决这个问题。

#### References:

- [1] Fukunaga K, Hostetler LD. The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Trans. on Information Theory, 1975,21(1):32-40.
- [2] Cheng Y. Mean shift, mode seeking and clustering. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1995,17(8): 790-799.

- [3] Comaniciu D, Ramesh V, Meer P. Real-Time tracking of non-rigid objects using mean shift. In: Werner B, ed. IEEE Int'l Proc. of the Computer Vision and Pattern Recognition, Vol 2. Stoughton: Printing House, 2000. 142–149.
- [4] Yilmaz A, Shafique K, Shah M. Target tracking in airborne forward looking infrared imagery. Int'l Journal of Image and Vision Computing, 2003,21(7):623–635.
- [5] Bradski GR. Computer vision face tracking for use in a perceptual user interface. In: Regina Spencer Sipple, ed. IEEE Workshop on Applications of Computer Vision. Stoughton: Printing House, 1998. 214–219.
- [6] Comaniciu D, Ramesh V, Meer P. Kernel-Based object tracking. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003, 25(5):564–575.
- [7] Collins RT. Mean-Shift blob tracking through scale space. In: Danielle M, ed. IEEE Int'l Conf. on Computer Vision and Pattern Recognition, Vol 2. Baltimore: Victor Graphics, 2003. 234–240.
- [8] Olson CF. Maximum-Likelihood image matching. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002,24(6):853–857.
- [9] Hu W, Wang S, Lin RS, Levinson S. Tracking of object with SVM regression. In: Jacobs A, Baldwin T, eds. IEEE Int'l Conf. on Computer Vision and Pattern Recognition, Vol 2. Baltimore: Victor Graphics, 2001. 240–245.
- [10] Mohammad GA. A fast globally optimal algorithm for template matching using low-resolution pruning. IEEE Trans. on Image Processing, 2001,10(4):626–533.

附录:定理证明.

不失一般性,假设:(1)  $F_2$  和  $F_3$  是  $F_1$  均匀缩小尺寸后的目标区域;(2)  $F_i, i=1,2,3$  均由  $u$  个不同颜色量化值的子区域组成,即  $F_i = \{f_j\}_{j=1,\dots,u}$ ; (3)  $B_i, i=1,2,3$  分别由  $v_i$  个不同颜色量化值的子区域组成,即  $B_i = \{b_j\}_{j=1,\dots,v_i}$ ; (4) 在以  $m$  级量化的核直方图  $P_i, i=1,2,3$  中,和子区域  $f_j$  对应的子项表示为  $\{fp_j^i\}_{j=1,\dots,u}$ , 对应  $b_j$  的子项表示为  $\{bp_j^i\}_{j=1,\dots,v_i}$ . 其中  $u+v \leq m, v = \max(v_1, v_2, v_3)$ .

由定义 2,核直方图用连续形式表示:

$$\begin{cases} fp_j^i = C_i \iint_{\sigma=f_j} k(\|X/r\|^2) d\sigma, i=1,2,3; j=1,\dots,u \\ bp_j^i = C_i \iint_{\sigma=b_j} k(\|X/r\|^2) d\sigma, i=1,2,3; j=1,\dots,v_i \\ C_i = 1 / \left( \sum_{j=1}^u \iint_{\sigma=f_j} k(\|X/r\|^2) d\sigma + \sum_{j=1}^{v_i} \iint_{\sigma=b_j} k(\|X/r\|^2) d\sigma \right) \end{cases} \quad (A1)$$

由积分中值定理得

$$\begin{cases} fp_j^2 = C_2 \iint_{\sigma=f_j} k(\|X/r\|^2) d\sigma = C_2 \cdot S_{f_j}^2 \cdot k\left(\frac{\xi_{f_j}^2}{r}\right)^2 \\ fp_j^3 = C_3 \iint_{\sigma=f_j} k(\|X/r\|^2) d\sigma = C_3 \cdot S_{f_j}^3 \cdot k\left(\frac{\xi_{f_j}^3}{r}\right)^2 \end{cases} \quad (A2)$$

其中,  $S_{f_j}^2, S_{f_j}^3$  分别是  $F_2$  和  $F_3$  的子区域  $f_j$  的面积,  $S_{f_j}^2 = S_{f_j}^3 \cdot \xi_{f_j}^2$  和  $\xi_{f_j}^3$  分别是  $F_2$  和  $F_3$  在子区域  $f_j$  上的某一点. 因为核函数  $k(\|X/r\|^2)$  是关于原点对称的单调递减偶函数<sup>[2]</sup>, 所以对于任意的  $\|X_1/r\| < \|X_2/r\|$ ,  $k(\|X_1/r\|^2) > k(\|X_2/r\|^2)$  成立. 根据已知条件, 区域  $F_3$  比区域  $F_2$  远离对应核函数  $k(\|X/r\|^2)$  的中心点. 显然有  $\left\| \frac{\xi_{f_j}^2}{r} \right\|^2 < \left\| \frac{\xi_{f_j}^3}{r} \right\|^2$ , 从而  $k\left(\frac{\xi_{f_j}^2}{r}\right)^2 > k\left(\frac{\xi_{f_j}^3}{r}\right)^2$ . 对于  $F_2$  和  $F_3$  存在相交区域的情况, 上述结论依然成立(证明略). 另外, 因为跟踪窗口半径相同, 即核窗宽相同, 所以有

$$C_1 = 1 / \iint_{\sigma=I_1} k(\|X/r\|^2) d\sigma = C_2 = C_3 \quad (A3)$$



综上所述,由(A2)可得  $fp_j^2 > fp_j^3$ . 进一步地,由核直方图的约束条件:

$$\sum_{j=1}^u fp_j^i + \sum_{j=1}^{v_i} bp_j^i = 1, i = 1, 2, 3 \tag{A4}$$

得  $\sum_{j=1}^{v_2} bp_j^2 < \sum_{j=1}^{v_3} bp_j^3$ . 因为  $F_2$  比  $F_1$  尺寸小,所以  $B_2$  的面积大于  $B_1$ . 因此,  $\sum_{j=1}^{v_1} bp_j^1 < \sum_{j=1}^{v_2} bp_j^2$ . 从而  $\sum_{j=1}^u fp_j^2 < \sum_{j=1}^u fp_j^1$ . 最终我们得出  $P_1, P_2$  和  $P_3$  中分量的大小关系如下:

$$\begin{cases} 1 > \sum_{j=1}^u fp_j^1 > \sum_{j=1}^u fp_j^2 > \sum_{j=1}^u fp_j^3 > 0 \\ 0 < \sum_{j=1}^{v_1} bp_j^1 < \sum_{j=1}^{v_2} bp_j^2 < \sum_{j=1}^{v_3} bp_j^3 < 1 \end{cases} \tag{A5}$$

由 Bhattacharyya 系数的定义可知,其大小实际上就是两组  $m$  维向量  $(\sqrt{p_1^i}, \dots, \sqrt{p_m^i})$  和  $(\sqrt{p_1^j}, \dots, \sqrt{p_m^j})$  夹角的余弦值. 夹角越大表明两个核直方图越不相似. 分别对背景区域的各分量以及目标区域的各分量开根号并进行矢量求和,得到 3 个新矢量  $V_i = \left( \sqrt{\sum_{j=1}^{v_i} bp_j^i}, \sqrt{\sum_{j=1}^u fp_j^i} \right)_{i=1,2,3}$ .  $\{V_i\}_{i=1,2,3}$  各矢量间的夹角同样刻画了  $\{P_i\}_{i=1,2,3}$  各核直方图之

间 Bhattacharyya 系数的大小. 结合如图 7 所示的几何关系以及式(A4)、式(A5)的约束,可以推出  $(V_3, V_1) > (V_2, V_1)$ , 如图所示,  $\beta > \alpha$ . 从而,  $P_1$  与  $P_2$  的 Bhattacharyya 系数大于  $P_1$  与  $P_3$  的 Bhattacharyya 系数.

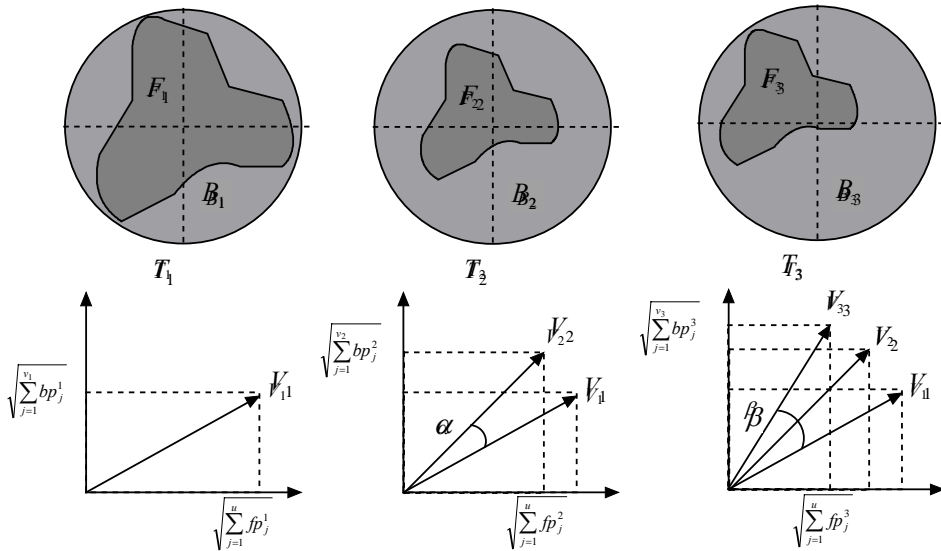


Fig.7 Tracking windows and similarity of their kernel-histogram

图 7 跟踪窗口以及所对应核直方图的相似性