

一种基于空间直方图的 Mean-Shift 跟踪算法*

胡彬, 赵欢, 郑敏

(湖南大学 计算机与通信学院, 长沙 410082)

摘要: 经典的 Mean-Shift 跟踪算法以颜色直方图为特征对目标进行跟踪。颜色直方图反映的是图像中颜色的组成情况, 即出现了哪些颜色以及各种颜色出现的次数。颜色直方图具有旋转不变性、缩放不变性等优点, 经常用于图像检索, 即通过比较颜色直方图的差异来衡量两幅图像在颜色全局分布上的差异。但是颜色直方图不能反映颜色的空间分布特征, 当跟踪目标与背景色颜色相近时可能造成错误跟踪, 导致跟踪失败。考虑目标颜色空间分布特征, 将空间分布信息融入颜色直方图中, 提出了基于空间颜色直方图的 Mean-Shift 跟踪方法, 全面地反映了颜色的整体分布信息和空间分布信息。在 VC 6.0 上利用新方法和经典 Mean-Shift 跟踪方法分别对发生遮挡的视频和快速运动的视频进行跟踪, 实验结果表明, 提出的新方法能够更好地对目标进行跟踪, 避免跟踪过程中的目标丢失。

关键词: 目标跟踪; Mean-Shift 跟踪; 颜色直方图; 空间颜色直方图; 相似性度量

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2010)06-2394-04

doi: 10.3969/j.issn.1001-3695.2010.06.114

Improved Mean-Shift tracking algorithm based on spatial histogram

HU Bin, ZHAO Huang, ZHENG Min

(School of Computer & Communication, Hunan University, Changsha 410082, China)

Abstract: The color-histogram indicates the color components of the target object-the sorts of color and the number of every kind of color. The global color histogram method has some excellencies-the rotation inflexibility and scale inflexibility and so on, so it is often used in the image retrieving-comparing two images through computing the difference of the color histogram of the two images. But the color histogram does not involve the spatial distributing information of the color. It may bring on failed tracking while the colors of target object is similar with the background colors. This paper proposed an improved tracking method which based on spatial color histogram, which involved both the colors number information and the spatial distribution information of the colors information. This paper applied new method and the classic Mean-Shift tracking method on the object occlusion of video and fast-motion video using VC 6.0. The experiment shows that the new method can tracking the target object better than the classic Mean-Shift tracking algorithm.

Key words: object tracking; Mean-Shift tracking; color histogram; spatial color histogram; similarity match

0 引言

目标跟踪一直是计算机视觉研究领域的热点问题, 其应用领域有视频监控、人机用户接口、虚拟现实等。如何对视频序列中的目标进行稳健、有效的跟踪是目标跟踪的研究重点。颜色信息由于对目标的平移、旋转和尺度变化具有不变性以及对象遮挡和姿态变化不敏感的特点而成为图像处理中的一种重要特征。因此, 基于颜色信息的目标跟踪方法引起了许多研究者的关注^[1-8], 在视频监控、智能人机交互、图像和视频编码以及智能机器人等许多研究领域得到了广泛的应用。

均值漂移 (Mean-Shift) 是 Fukunaga 等人^[1]提出的一种非参数概率密度梯度估计算法, 在统计相似性计算与连续优化方法之间建立了一座桥梁。Cheng^[2]发展了这一理论, 将其运用到模式分类与图像分割中, 此后, 均值漂移被广泛用于模式分类、图像分割以及目标跟踪等方面。Comaniciu 等人^[3,4]提出了一种加权直方图 (weighted histogram) 建模目标的颜色概率

分布 (又称为基于核的概率密度表达), 该方法统计落入每一个直方图区间像素的加权个数, 其中权重是落入该区间的像素二维空间坐标和目标中心的距离函数, 距离越小权重越大。这种方法用 Bhattacharyya 距离来度量颜色模型之间的相似度, 通过一阶梯度下降算法将该距离测度极小化, 从而在搜索范围内得到目标的位置。该方法被称之为 Mean-Shift (MS) 算法, 能够通过较少的几次迭代收敛到目标的真实位置, 由于 MS 算法速度快, 适用于实时视频跟踪中对目标进行跟踪而受到了研究者的重视。Collins 借鉴 Lindeberg 的尺度空间理论, 将其融入 Mean-Shift 算法中用来解决其不能很好地跟踪尺度变化的目标的缺点^[5]。实验结果表明该算法实时性很强, 可以实现对运动人脸的快速跟踪, 同时对目标遮挡也有很好的鲁棒性。

颜色直方图只包含了该图像中某一颜色值出现的频数, 仅仅反映了物体的全局色彩信息, 不能反映颜色的空间分布特征。任一幅图像都能惟一地给出一幅与它对应的直方图, 但不同的图像可能有相同的颜色分布, 从而就具有相同的直方图, 因此直方图与图像是一对多的关系, 仅仅以目标的颜色直方图

收稿日期: 2009-10-23; 修回日期: 2009-11-28 基金项目: 高等学校科技创新工程重大项目培育资金资助项目 (070866)

作者简介: 胡彬 (1985-), 男, 湖南邵阳人, 硕士研究生, 主要研究方向为计算视频处理等 (hhubin_112@163.com); 赵欢 (1967-), 湖南长沙人, 教授, 主要研究方向为计算机视频、音频处理; 郑敏 (1984-), 女, 湖南长沙人, 硕士研究生, 主要研究方向为计算机视频处理。

为模型进行 Mean-Shift 迭代跟踪不能全面地反映目标颜色的空间分布信息,在跟踪过程中有可能发生跟踪失败。

1 经典 Mean-Shift 跟踪方法

1.1 目标模型描述

视频目标跟踪是 Mean-Shift 算法的一个重要应用方面,大多数文献都沿用 Comanicu 的方法^[3,4],采用 Bhattacharyya 系数作为相似性度量标准,然后利用 Mean-Shift 算法来确定最佳目标位置。Mean-Shift 跟踪算法以目标的颜色直方图作为目标特征的描述模型。如果目标特征和候选目标特征分别表示为

$$q = \{q_u\}_{u=1,\dots,m}, p(y) = \{p_u(y)\}_{u=1,\dots,m}$$

其中: y 表示的是候选目标区域的中心位置,满足 $\sum_{u=1}^m q_u = 1, \sum_{u=1}^m p_u = 1; m$ 为特征值 bin 的个数。采用 Bhattacharyya 系数来衡量目标模型与候选目标的相似性,即 $\rho(y)u \equiv \rho[p(y), q]$ 。目标模型的特征分布表示为

$$q_u = C \sum_{i=1}^n k(\|x_i^* - u\|^2) \delta[b(x_i^*) - u]$$

其中: $\delta(\cdot)$ 是 Kronecker delta 函数; $\{x_i^*\}_{i=1,\dots,n}$ 表示的是目标区域里面的像素位置,归一化后目标模型中心像素的位置为 $0; b: R^2 \rightarrow \{1, \dots, m\}$ 是像素点到像素特征 bin 的映射函数, C 为归一化常数。

相应的候选目标模型的特征分布表示为

$$p_u = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \delta[b(x_i^*) - u]$$

其中: C_h 是归一化常数; n_h 为候选目标区域中的像素总数; $k(\cdot)$ 为核函数; h 为带宽。

1.2 基于相似度量函数的目标跟踪

相似函数定义了两个模型之间的相似程度,在 Mean-Shift 跟踪算法中,为了得到目标与候选目标相似度,定义目标模型与候选模型的计算公式为

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

其中: $\rho[p(y), q] = \sum_{u=1}^m \sqrt{p_u(y)q_u}$ 为 q 和 $p(y)$ 之间的 Bhattacharyya 系数。

要使得式(1)取得最小值就必须保证 Bhattacharyya 系数 $\rho[p(y), q]$ 取最大值。在 $p_u(y_0)$ 处进行一阶 Taylor 式展开,忽略一些极小项,可以得到 $\rho[p(y), q]$ 的线性近似式如下:

$$\rho_{\text{approx}}[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0)q_u} + \frac{1}{2} \sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}}$$

假设 $p_u(y_0) > 0$, 对于所有的 $u = 1, \dots, m$ 成立,则有

$$\rho_{\text{approx}}[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0)q_u} + \frac{C_h}{2} \sum_{i=1}^{n_h} w_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \tag{2}$$

其中: $w_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(y_0)}} \delta[b(x_i) - u]$ 。式中第一部分为常量,要使得 $\rho[p(y), q]$ 最大,式(2)的第二部分就得取最大值。结合均值漂移算法计算步骤,可得 Mean-Shift 算法的迭代公式如下:

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)} \tag{3}$$

其中: $g(x) = -k'(x)$ 。如果使用 Epanechnikov 核,那么式(3)可进一步化简为

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \tag{4}$$

首先初始化 y_0 , 计算 w_i , 并重复式(4)的计算,直到 $\|y_{i+1} - y_i\| < \varepsilon$ 或大于预定的最大迭代次数(一般设置为 5 次),迭代结束, ε 为预先设定的大于 0 的正数(一般取值为 1 个像素),得到的最终位置 y_{i+1} 即为目标中心所在的位置。

2 改进的 Mean-Shift 算法

2.1 空间颜色直方图

2.1.1 考虑图像的边缘特征

图像的边缘信息是图像一个重要特征,是人眼识别图像的一个重要因素,但是颜色直方图并没有考虑图像的边缘信息。本文将属于边缘点的颜色和其他不属于边缘点的同类颜色区别对待,对颜色直方图进行扩展。

设 M 为定义的颜色数目,加入边缘点特征后将颜色特征点扩展到 $2M$ 个,本文定义从图像像素这 $2M$ 个特征空间的函数如下:

设 x 为图像中的一个像素,分别定义其 RGB 颜色空间值为 R_x, G_x, B_x , 则映射函数为

$$K = Hx \times M + Rx \times L_1 + Gx \times L_2 + Bx \tag{5}$$

其中: Hx 取值为 0 或者 1 (当该像素点为边缘像素点时 Hx 取值为 1, 否则 Hx 取值为 0); L_1, L_2 为量化值。通过上述公式,将图像颜色特征空间从 M 扩展到 $2M$ 空间。

2.1.2 考虑像素坐标位置

颜色直方图是图像的一种重要特征,然而其完全丢弃了图像颜色的空间分布信息。为了有效利用图像颜色的空间分布信息, Rao 等人^[7]提出采用环形颜色直方图来描述颜色的空间分布特征, Clique 等人^[8]提出采用 SCH 来描述具有相同颜色的像素在图像空间的分布特征。借鉴 Rao 和 Clique 等人提出的像素在图像空间的分布特征,针对 Mean-Shift 跟踪算法固有的缺点,将像素分布的坐标信息考虑到直方图中。

首先本文给出同一种颜色像素质心的定义:

定义 1 设 I 表示任意一幅图像, $I(x, y)$ 为像素 (x, y) 处的值:

$$A_i = \{(x, y) | (x, y) \in I, I(x, y) = i, 1 \leq i \leq m\}$$

为图像中颜色为 i 的所有像素的集合,其中 m 表示图像颜色总的数目, $|A_i|$ 表示集合 A_i 中像素的数目,为图像中颜色为 i 的所有像素的质心。 x_i 和 y_i 定义为

$$x_i = \frac{1}{|A_i|} \sum_{(x,y) \in A_i} x, y_i = \frac{1}{|A_i|} \sum_{(x,y) \in A_i} y \tag{6}$$

定义 2 设 p 为 (x, y) 处的颜色为 i 的一个像素点,该点到颜色 i 的质心距离以及到质心的均值分别定义为

$$d_n(p, C_i) = \sqrt{(x-x_i)^2 + (y-y_i)^2}, \mu_i = \frac{\sum_{n=1}^{|A_i|} d_n(p, C_i)}{|A_i|} \tag{7}$$

定义 3 定义颜色 i 的像素到质心的距离方差为 σ_i^2 则

$$\sigma_i^2 = \frac{\sum_{n=1}^{|A_i|} (d_n(p, C_i))^2}{|A_i|} \tag{8}$$

本文定义新的直方图模型为 $\langle b_i, u_i, \sigma_i^2 \rangle$, 计算公式为

$$q(y) = b_i(y) G(u_i, \sigma_i^2)$$

融入颜色的空间分布信息到新的直方图模型中。

如下两幅图像,图 2(a) 为图 1(a) 交换两个像素块的位置

而得到的。两幅图像的像素颜色的类型与相同特征颜色的个数完全一样,只是在空间分布上发生了变化。首先计算原图像的颜色直方图和改进过后的空间颜色直方图(图 1)。

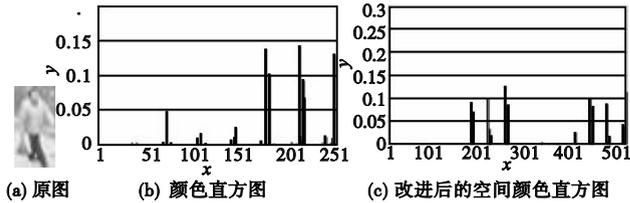


图1 原图及改进前后的直方图

下面计算交换像素块后的图像颜色直方图和改进过后的空间颜色直方图(图 2)。比较图 1 和 2 可以发现,图 1(a)和图 2(a)的颜色直方图是一样的,但是图 1(a)和图 2(a)的空间直方图却有着明显的区别。由此可见,空间颜色直方图能够体现出像素的空间分布信息,能够更全面地表征图像。

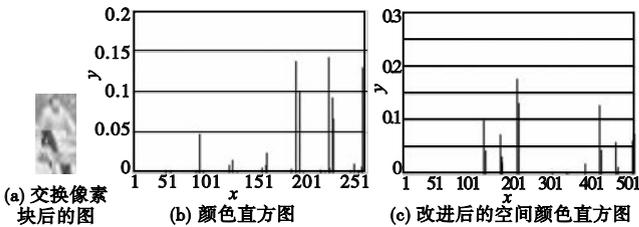


图2 交换像素块图及其改进前后直方图

2.2 色模型及相似性度量

基于上述空间—颜色直方图模型,给定由 N 个像素组成的目标参考图像,其中 x_i 是第 i 个像素的空间坐标,这个像素在图像中的位置可以用这个像素到其颜色质心的距离即 $d_u(x_i, C_u)$ 来体现,更新过后的 p_u 按如下方式定义

$$p_u(t(x); b_u, \mu_u, \sigma_u^2) = b_u G(\mu_u, \sigma_u^2) \tag{9}$$

其中: $G(\mu_u, \sigma_u^2)$ 是均值向量为 μ_u , 方差矩阵为 σ_u^2 的高斯分布, μ_u, σ_u^2 的定义见定义 1 和 2; $b_u(y)$ 为色彩特征分布计算表达式。

考虑一个由 N' 个像素组成的中心坐标为 y 的候选图像区域的颜色分布:

$$p(y) = b_{u'}(y) G(\mu_{u'}, \sigma_{u'}^2) \tag{10}$$

其中: $\mu_{u'}, \sigma_{u'}^2$ 与式(9)中的 μ_u, σ_u^2 具有相似的形式,而 $b_{u'}(y)$ 为经典 Mean-Shift 算法的目标色彩直方图。

$q_u(t(x); b_u, \mu_u, \sigma_u^2)$ 与 $p_u(y)$ 之间的相似性根据 Bhattacharyya 距离来度量 $\rho[q_u, p_u(y)] = \sum_{u=1}^m \sqrt{q_u, p_u(y)}$ 。通过积分得到

$$\rho[q_u, p_u(y)] = C_u (b_u b_{u'})^{\frac{1}{2}} \tag{11}$$

其中:

$$C_u = \left(\frac{2 |R_u \parallel R_{u'} \parallel}{|R_u + R_{u'}|} \right)^{\frac{1}{2}} \exp\left(-\frac{1}{4} (\mu_u - \mu_{u'})^2 (R_u + R_{u'})^{-1}\right)$$

所以定义目标模型和候选模型的相似度为

$$\rho[q, p(y)] = \sum_{u=1}^d \rho(q_u, p_u(y)) \tag{12}$$

直方图只包含了该图像中某一颜色值出现的频数,而丢失了某像素所在的空间位置信息。本文提出的新的颜色直方图模型考虑了颜色的空间分布信息,能够更好地反映图像的特征。

2.3 改进后的算法

设在前一帧中目标的中心位于 y_0 , 在当前帧中, Mean-Shift 算法基于梯度下降方法寻找式(12)的极值。类似于经典

Mean-Shift 算法将式(12) 在 $p(y_0)$ 进行一阶泰勒展开,相对于 y 求偏导数并令其为零,得到如下迭代等式:

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\frac{\|y_0 - x_i\|}{h}\right)}{\sum_{i=1}^{n_h} w_i g\left(\frac{\|y_0 - x_i\|}{h}\right)} \tag{13}$$

$$w_i = \sum_{u=1}^m C_u \sqrt{\frac{b_u}{b_u(y_0)}} \delta[b(x_i) - u]$$

其中: C_u 与上面定义相同。如果使用 Epanechniko 核,那么式(13)可进一步化简为同式(4)的形式。整个跟踪流程如下:

a) 逐帧采集视频,若为第一帧则转到步骤 b), 否则转到步骤 c)。

b) 选定跟踪目标,通过 2.2 节描述的模型计算其空间—颜色直方图 M , 转到步骤 c)。

c) 取初始目标中心为上一帧确定的目标中心, 设为 y_0 。转到步骤 d)。

d) 通过 2.2 节描述的模型,计算中心为 y_0 的候选目标的空间—颜色归一化直方图 $p(y_0)$ 。计算 M 与 $p(y_0)$ 的匹配系数。转到步骤 e)。

e) 利用 Mean-Shift 迭代公式,计算得到候选目标的中心位置 y_1 。通过 2.2 节描述的模型,计算中心为 y_1 的候选目标的空间—颜色归一化直方图 $p(y_1)$ 。计算 M 与 $p(y_1)$ 的匹配系数。转到步骤 f)。

f) 判断 $\|y_1 - y_0\|$ 是否小于阈值,是的话停止迭代, y_1 即为要找的最优目标中心点, 转到步骤 a); 否则, 比较 $\rho[p(y_0), M]$ 与 $\rho[p(y_1), M]$ 。若 $\rho[p(y_0), M] \leq \rho[p(y_1), M]$, 则设 $y_0 = 0.5(y_1 + y_0)$, 否则 $y_0 = y_1$ 。转到步骤 d)。

经典的 Mean-Shift 算法的主要时间消耗在计算选定区域和候选区域的颜色直方图计算式(4), 时间复杂度为 $O(MN)$, 其中 M, N 为选定区域的长度和宽度。算法改进后, 式(4)并没有变复杂, 主要是在计算直方图的时候需要额外计算像素点的颜色均值和方差, 均值需要 MN 次计算, 方差需要 MN 次计算, 总共增加 $2MN$ 次计算时间, 总的时间复杂度可以用 $O(MN) + 2MN$ 表示, 其中 $O(MN) + 2MN$ 的时间复杂度即为 $O(MN)$ 。所以虽然稍微增加了计算量, 但是并没有增加算法的时间复杂度, 满足实时性要求。

3 实验

图像的采样速率为 30 fps, RGB 颜色空间为 512。机器配置为 Pentium® 4 3.0 GHz, 512 MB 内存, 在 Windows XP 系统下 Visual C++ 编程实现。

实验 1 相似背景下发生遮挡的视频目标跟踪

该视频为一组人行行走视频序列, 视频中目标与背景颜色十分相似。在第 128 帧时目标被对面而来的行人遮挡, 跟踪的难度较大。实验表明, 经典 Mean-Shift 算法在发生遮挡后(第 140 帧)导致目标跟踪失败, 而改进后的算法却能很好地抵抗遮挡导致的跟踪失败, 具有良好的跟踪效果(图 3)。

视频序列中, 第 128 帧发生遮挡, 到 146 帧结束。发生遮挡的时候 BH 系数会降低, 在遮挡效果达到最大化的时候 BH 最小。根据目标模板与跟踪模板之间的距离定义: $d(y) = \sqrt{1 - \rho[p(y), q]}$, 在 BH 系数达到最小的时候, 模板之间的距离最大。当遮挡结束的时候, 随着 BH 系数的增大, 模板之间

距离逐渐变小,回到遮挡之前的水平。模板距离如图 4 所示。

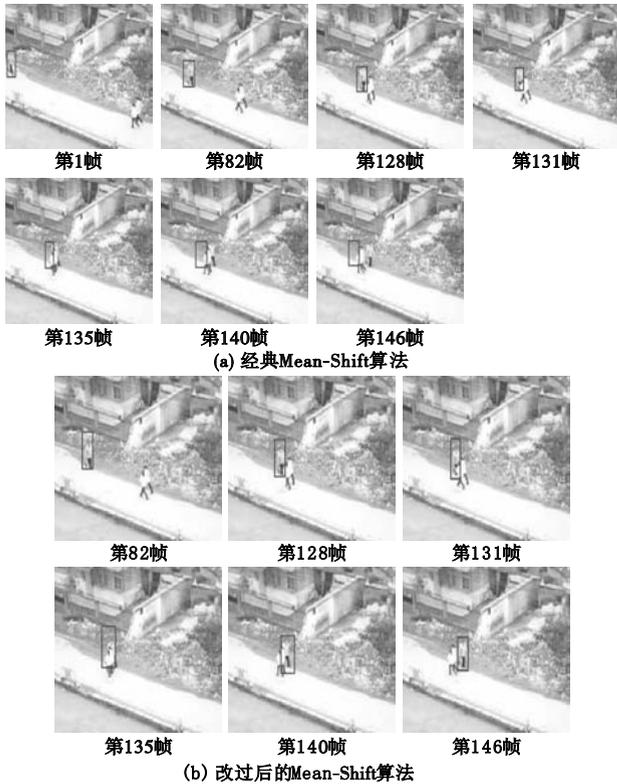


图3 相似背景下发生遮挡的视频目标跟踪:经典的Mean-Shift算法与改进后的Mean-Shift算法跟踪效果比较

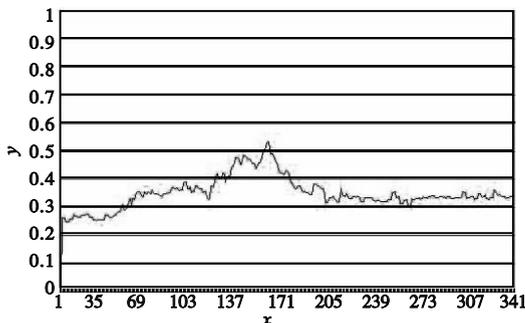


图4 改进算法跟踪过程中模板距离

实验 2 相似背景下发生快速运动的目标跟踪

该视频为狼追逐羚羊的一个视频,视频中目标与背景颜色十分相似,羚羊的运动速度较快,在第 10~101 帧的时候,经典 Mean-Shift 算法还能够较好地跟踪羚羊,但是到了第 104 帧,就逐渐出现跟踪定位不准,最终导致跟踪丢失,到第 106 帧的时候已经完全丢失了。改进后的基于空间直方图的跟踪算法在第 101 帧以后也能较好地为目标进行跟踪,如图 5 所示。

该视频目标跟踪中的目标与模板的距离参数在 0.5 上下波动,如图 6 所示。

4 结束语

本文提出的基于改进后的颜色直方图的 Mean-Shift 跟踪方法,考虑了目标边缘特征点的信息以及像素的空间位置分布信息,相比原来的经典 Mean-Shift 跟踪算法,能够在比较苛刻的条件下对目标进行更加稳定的跟踪,能够满足实时性的要求。下一步研究的重点是如何对目标颜色进行聚类分类以忽略小比例颜色对目标直方图的影响以及如何对多目标进行有效跟踪。

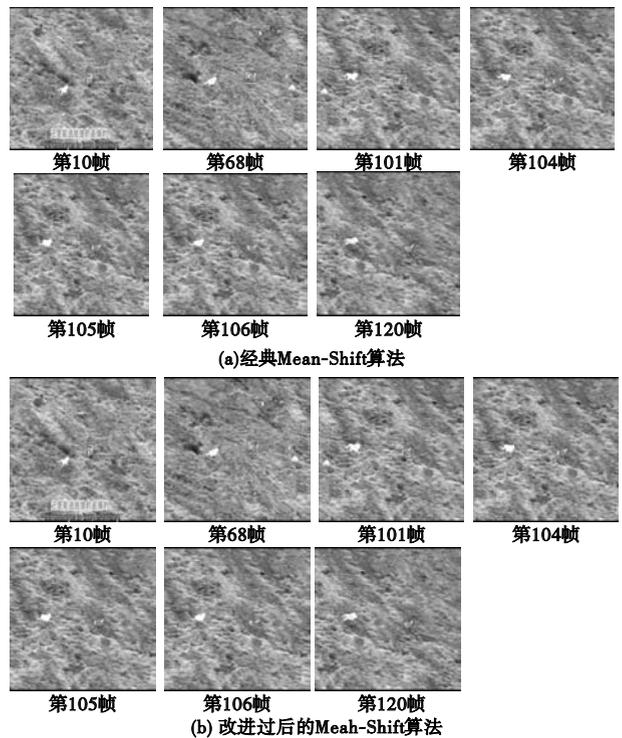


图5 相似背景下快速运动物体目标跟踪:经典的Mean-Shift算法与改进后的Mean-Shift算法跟踪效果比较

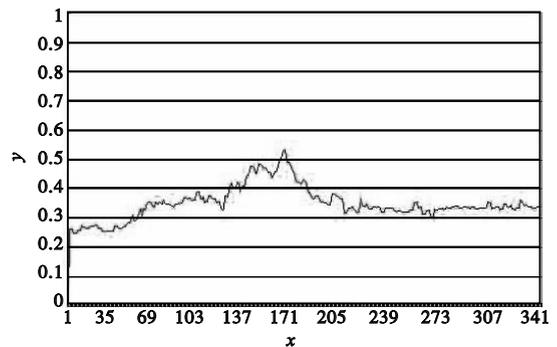


图6 改进算法跟踪过程中模板距离

参考文献:

- [1] FUKANAGA K, HOSTETLER L D. The estimation of the gradient of a density function, with applications in pattern recognition[J]. IEEE Trans on Information Theory, 1975, 21(1): 32-40.
- [2] CHENG Y. Mean-Shift, mode seeking, and clustering[J]. 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[C]//Proc of IEEE International Conference on Computer Vision and Pattern Recognition. Stoughton; Printing House, 2000: 142-149.
- [4] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.
- [5] 夏长俊,李翔,顾文锦,等.基于Canny边缘检测的Mean-Shift跟踪核窗宽确定方法[J]. 海军航空工程学院学报, 2008, 23(3): 308-314.
- [6] 文志强,蔡自兴.目标跟踪中巴氏系数误差的分析及其消除方法[J]. 计算机学报, 2008, 31(7): 1165-1174.
- [7] RAO A, SRIHARI R, ZHANG Z. Spatial color histogram for content based retrieval [C]//Proc of the 11th IEEE International Conference on Tools with AI. 1999: 183-186.
- [8] CINQUE L, LEVIALDI S, OLSEN K, et al. Color-based image retrieval using spatial-chromatic histogram[C]//Proc of IEEE International Conference on Multi-media Computing and System. 1999: 969-973.
- [9] 李培华.一种改进的Mean-Shift跟踪算法[J]. 自动化学报, 2007, 33(4): 316-348.