

基于模糊隶属关系的 Mean Shift 图像跟踪改进算法

沈悦^① 吴小俊^②

^①(江苏科技大学电子信息学院 镇江 212003)

^②(江南大学信息工程学院 无锡 214122)

摘要: 鉴于Mean Shift算法中核心步骤计算所得的新目标模板中心位置几乎都非整数的问题, 该文提出了模糊隶属度的概念, 并在此基础上, 改进了Mean Shift算法在图像跟踪中的实现步骤。实验表明, 该算法计算量小, 且能很好的减少原算法中不必要的误差。

关键词: 图像跟踪; Mean Shift; 模糊隶属度

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2008)11-2688-05

Improvement Algorithm of Mean Shift Based on Fuzzy Subordination Relations

Shen Yue^① Wu Xiao-jun^②

^①(School of Electronics and Information, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

^②(School of Information Engineering, Southern Yangtze University, Wuxi 214122, China)

Abstract: In view of the fact that the center position of new target template obtained by the computation of the core step of Mean Shift algorithm is nearly all non-integer, in this paper, a concept of fuzzy membership is proposed, the steps in image tracking using Mean Shift algorithm are improved. The results of experiments show that computational load and error are less than that of the original algorithm.

Key words: Image tracking; Mean Shift; Fuzzy membership

1 引言

“Mean Shift”概念最早是由Fukunaga等人^[1]于1975年提出, 在1995年, Cheng^[2]将这个�方法引入计算机视觉领域。Chmaniciu和Meer^[3-5]对Mean Shift在图像滤波、分割和跟踪中的使用都做了较早的论述。Mean Shift算法的理论是建立在N维连续空间基础之上的, 而算法迭代过程中的核心步骤(下文式(1))计算所得潜在目标中心位置的行列表示难免为非整数(实际过程中几乎都为非整数), 这就与一般意义上灰度图像的像素点都是基于矩阵, 且行列表示都为整数的基本要求不相匹配。如果分别计算该结果临近像素点的Bhattacharyya距离, 虽然准确性较高, 但时间复杂度太大。鉴于此, 本文引入模糊隶属度的概念, 实验结果显示了本算法的优越性。

2 Mean Shift跟踪算法

Mean Shift是一种基于外部特征的跟踪算法, 能对非刚性目标进行实时跟踪^[6-8], 其涉及到的一些概念, 如像素的概率密度分布 $\hat{p}_u(y_0)$ 和 \hat{q}_u , Bhattacharyya距离等, 在文献[6]中已经做了较详细的介绍, 在此不再赘述, 其中算法实现

跟踪的方法步骤如下(下文中称该步骤为原算法):

(1)初始化当前图像潜在目标的中心位置为 y_0 , 分别计算 $\hat{p}_u(y_0)$ 以及 $\rho[\hat{p}_u(y_0), \hat{q}_u]$;

(2)更新权值 w_i ;

(3)确定潜在目标中心的新位置:

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

并根据结果计算 $\rho[\hat{p}_u(y_1), \hat{q}_u]$;

(4)当 $\rho[\hat{p}_u(y_1), \hat{q}_u] < \rho[\hat{p}_u(y_0), \hat{q}_u]$ 时, 则 $y_1 = 0.5(y_0 + y_1)$;

(5)若 $\|y_0 - y_1\| < \varepsilon$ 则迭代结束, y_1 即为潜在目标窗口中心, 否则令 $y_1 = y_0$, 返回到(1)。

3 非整数结果的分析与处理

从上文的实现步骤(3)中, 其结果 y_1 几乎都非整数, 其并不能直接代入其后步骤的计算当中。那么, 选择怎样的整数像素点作为其替代者, 不妨从以下几方面考虑。

3.1 候选点的选定

不难理解, 一般情况下(y_1 行列都非整数), 可以作为替代者的像素点(行列皆为整数), 必然是与 y_1 的Euclid距离最近的4个位置之一, 不妨称这样的4个点(位置)为候选点,

2007-05-08收到, 2007-12-17改回

国家自然科学基金(60572034), 教育部新世纪优秀人才计划项目(2006)和江苏省自然科学基金(BK2006081)资助课题

如图 1 所示(圆代表一般像素点,五角星代表 y_1 , 四角星代表候选点)。

以上情况几乎占到实验结果的 99%以上,但是特殊情况依然值得考虑:

(1)当 y_1 的行或列之一为整数,且 y_1 漂移方向非水平和垂直时,则将 y_1 沿漂移方向延伸一段非常小的距离 p ,再根据 y_1 新的位置,按照上文一般情况的描述选择候选点,如图 2 所示为该情况之一(箭头表示漂移方向)。

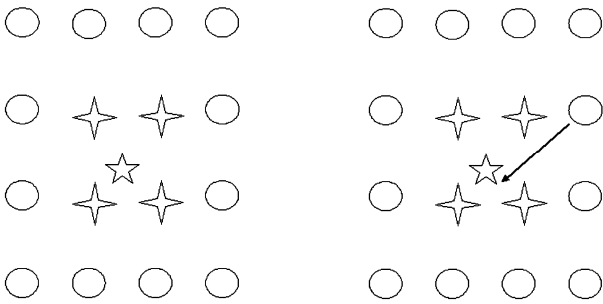


图 1 一般情况

图 2 特殊情况 1

(2)当 y_1 的行或列之一为整数,且 y_1 漂移方向为水平或垂直时,则候选点退化为两个,分别为水平或垂直方向上与 y_1 的 Euclid 距离最近的两个点,如图 3 所示为该情况之一(箭头表示漂移方向)。

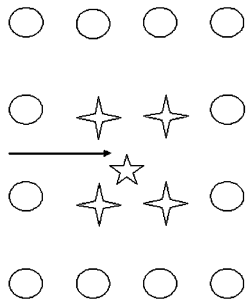


图 3 特殊情况 2

(3)当 y_1 的行与列都为整数时(这种概率微乎其微),则替代点就是 y_1 本身。

3.2 中心漂移方向与鼓励漂移

Cheng^[2]的文章中已经证明,每次迭代中,中心的漂移总是指向概率密度增大最大的方向,即梯度方向;而 Chmaniciu 和 Meer^[6]则用 Bhattacharyya 距离来判断漂移方向的合理性,并通过实验证明,每一次迭代中超过 99.9%的情况下,原算法第 4 步是没有必要的。根据以上论断不难看出,中心漂移的方向对于 Mean Shift 算法至关重要,在候选点中,与 y_1 有最近似漂移方向的候选点理应有最大的可能成为最终的替代点。

但凡收敛过程总有一个收敛幅度,由于图像噪声的存

在,不能保证每次都有大幅度的漂移是很正常的。例如原中心 y_0 的位置(行列表示)为(10,10),而非整数中心位置 y_1 的位置为(10.3,10.4),如果仅以简单的四舍五入方法来处理 y_1 ,则 y_1 与 y_0 相同,根据上文 Mean Shift 的算法步骤应认为迭代结束,这是很不合理的。为了避免由于对型如(0.3, 0.4)的“小漂移”的忽视而造成整个系统提前收敛到一个非理想值,重视甚至放大“小漂移”的作用,即视其为有效漂移是值得考虑的,我们不妨称这样的策略为鼓励漂移。

4 基于模糊隶属关系的改进算法

根据以上情况,本文引入基于模糊隶属关系的改进算法。本算法根据前文提到的漂移方向和“小漂移”作用为基本因素,为候选点分配属于替代点的模糊隶属度,并在此基础上改进了 Mean Shift 算法在图像跟踪中的实现步骤。

4.1 漂移相似度

定义 1 4 候选点中,与原中心 y_0 位置的 Euclid 距离最近者称为起始点,记为 D_1 ;与 D_1 同行不同列者称为横侧点,记为 D_2 ;与 D_1 同列不同行者称为纵侧点,记为 D_3 ;与 D_1 对角者称为漂移点,记为 D_4 。两候选点中,与原中心 y_0 位置的 Euclid 距离最近者称为起始点,记为 D_1 ;另一点称为漂移点,记为 D_4 。

定义 2 当 $\|y_1 - D_i\| < r (i=1,2,3,4)$ 时,有 $y_1 = D_i$,其中 r 是一个很小的值,作用相当于上文实现步骤中第(5)步的 ϵ ,称为隶属半径。

定义 3 对于图像矩阵,以行水平增大方向为 x 轴正方向,列垂直增大方向为 y 轴负方向定义二维坐标。以原中心 y_0 为起点,指向目标位置(Target)的矢量称为漂移方向。漂移方向的量化值称为斜率,其表示如下:

(1)当 y_0 与 Target 坐标不同行且不同列时,斜率表示为两者坐标 y 轴方向差值与 x 轴方向差值的商。

(2)当 y_0 与 Target 坐标同行或同列时,斜率表示为一个范围,如图 4 所示,图中阴影由圆弧和两条切线包围,表示斜率范围,该范围中的任意值都可表示为两点之间的斜率。

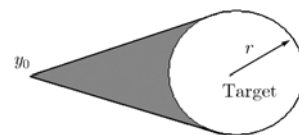


图 4 斜率表示范围

(3)当 y_0 与 Target 坐标重合时,斜率为 0。

所谓漂移相似度,就是指候选点 D_i 相对于 y_0 的漂移方向与 y_1 相对于 y_0 的漂移方向的相似程度,以两者斜率的比值来实现,当候选点为 4 个时,漂移相似度表示为

$$S_{D_i} = \text{gama}(l_{D_i} / l_m) \tag{2}$$

其中 l_{D_i} 表示候选点 D_i 相对于 y_0 漂移方向的斜率, l_m 表示

y_1 相对于 y_0 漂移方向的斜率, gama 函数将漂移相似度的值域定义在 $[0,1]$ 上, 表示为

$$\text{gama} = \begin{cases} x, & 0 \leq x \leq 1 \\ 1/x, & x > 1 \\ 0, & \text{其他} \end{cases} \quad (3)$$

当 l_{D_i} 非定值时(如定义 3 中第(2)种情况), 则在其范围内选择一个斜率使得 S_{D_i} 最大。

而当候选点仅有两个时, 即 l_m 为 0 或 ∞ 时, 令 $S_{D_1} = S_{D_4} = 1$; $S_{D_2} = S_{D_3} = 0$ 。

4.2 “小漂移”隶属度

定义 4 令 y_1 行坐标与 D_1 行坐标之差的绝对值为行绝对漂移, 记为 x' ; 令 y_1 列坐标与 D_1 列坐标之差的绝对值为列绝对漂移, 记为 y' ; x' 和 y' 合称绝对漂移。

对于游离于各候选点之间的绝对漂移, 其漂移方向、具体的位置以及其在 y_0 与 y_1 之间整体漂移中所占的比重, 都直接影响了其对各候选点作用的大小。综合以上各点因素, 不妨将绝对漂移相对各候选点的隶属关系表示如下:

$$\left. \begin{aligned} H_{D_1} &= (1 - W(d)K(x'))(1 - W(d)K(y')) \\ H_{D_2} &= (1 - W(d)K(x'))W(d)K(y') \\ H_{D_3} &= W(d)K(x')(1 - W(d)K(y')) \\ H_{D_4} &= W^2(d)K(x')K(y') \end{aligned} \right\} \quad (4)$$

其中 K 称为漂移函数, 是定义域和值域都在 $[0,1]$ 上的, 单调增的凸函数。由多次的实验得出以下结论: 当 D_1 相对于 y_0 有较大漂移时, 绝对漂移的作用与之相比相对较小; 反之则较大。因此引入了权值函数 W , 它是定义在 $[0, \infty)$ 上, 值域在 $(0,1]$ 上的单调减函数。而作用其上的参数 $d = \|D_1 - y_0\|$, W 的性质体现了绝对漂移相对整体漂移所占的比重。

4.3 改进算法

本文候选点的隶属度是基于漂移相似度和“小漂移”隶属度的, 且同时拥有较大的两者的值是候选点较高隶属度的保证, 但是前者是从整体上对候选点的评估, 而后者仅局限于候选点内部的绝对漂移作用, 因此在最终的隶属关系上应该以前者为主, 后者为辅。鉴于此, 本文以两者的和为隶属度的量度, 又引入漂移系数 α , 该系数取值范围为 $[0,1]$, 体现了两者对隶属度的主辅影响。

$$G_{D_i} = S_{D_i} + \alpha H_{D_i} \quad (5)$$

α 的取值是改进算法的重点, 本文将在下一节中通过实验结果对此做进一步的分析。

以式(5)为基础的多次实验证明: 超过 90%以上的情况下, 拥有最大隶属度的候选点同时也拥有最大的 Bhattacharyya 距离, 同时考虑到 3.2 节中提及的文献[6]中的论断, 对图像跟踪步骤进行改进, 具体步骤为

- (1)初始化当前图像潜在目标的中心位置为 y_0 , 并计算 $\hat{p}_u(y_0)$;
- (2)更新权值;

(3)根据式(1)更新潜在目标中心的新位置 y_1 ;

(4)若 $\|y_1 - \text{round}(D)\| < r$, 则 $y_1 = D$, 则迭代结束; 否则进入(5);

(5)计算候选点, 若本次候选点与上次相同(不考虑先后), 则进入(7), 否则进入(6);

(6)根据式(5)分别计算各候选点的隶属度, 选其最大者作为 y_1 , 若 $y_0 = y_1$ 则迭代结束; 否则令 $y_0 = y_1$, 然后转到(1);

(7)分别计算各候选点的 Bhattacharyya 距离, 选其最大者作为 y_1 , 迭代结束。

5 实验结果与分析

为了验证本算法的优越性, 以4幅标准人脸图像的人眼跟踪测试及3组视频序列的跟踪为例, 在P4 2.6GHz的CPU, 256M内存的PC机上, 用MATLAB编程实现并分别进行比较。

图5为人眼跟踪测试实验结果, 3幅人脸图像的大小均为 $100 \cdot 100$, 跟踪测试中取 (h_x, h_y) 为(27,19), 眼睛模板的直方图量化层数取255。图5(a)为原算法的跟踪结果, 其他3组均为本文算法, 其中图5(b)中 α 取0.2, 图5(c)中 α 取0.5, 图5(d)中 α 取0.8。

图6为行人视频序列跟踪实验结果, 3组视频序列中前两组为笔者自行拍摄, 最后一组乃网上下载, 前两组每帧大小均为 $240 \cdot 320$, 最后一组为 $384 \cdot 512$, 模板直方图量化层数均取255。其中前两幅图像取自视频序列1, 中间两幅取自视频序列2, 最后两幅取自视频序列3。

从图 5 和图 6 的实验结果不难看出, 原算法虽然实现了

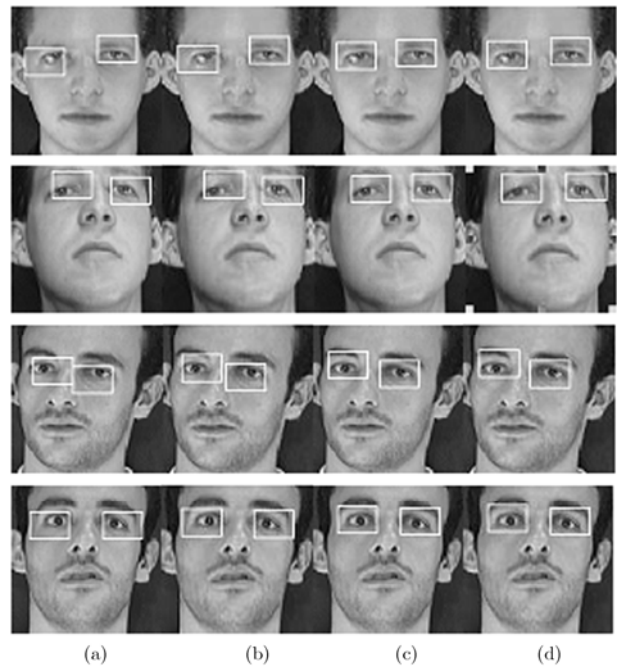


图5 人眼跟踪实验结果对比

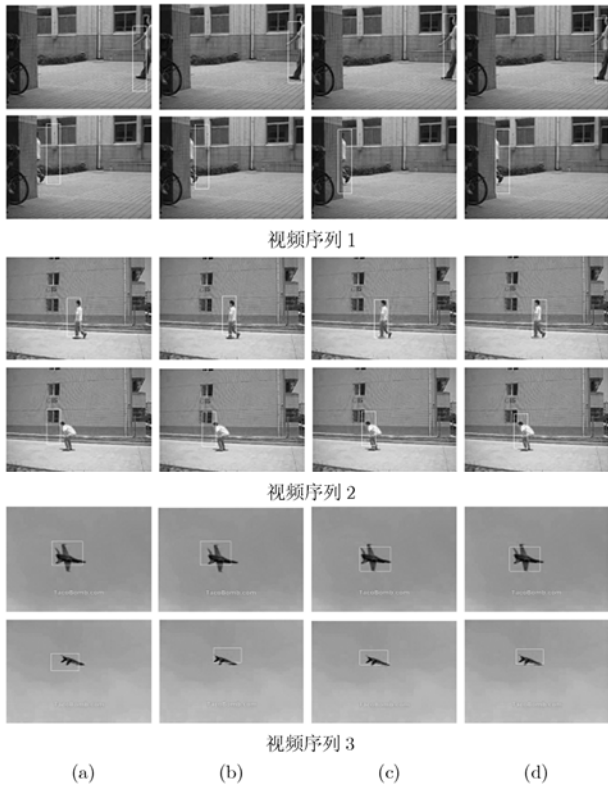


图 6 视频序列跟踪结果对比

对目标大体的跟踪,但是显得不够精确,观察窗口往往与目标之间有一定的偏移;相比较而言,本文算法中的(c)和(d)两组,在完成跟踪的基础上,尽可能地实现了使目标处于观察窗口的中心位置,一定程度上提高了跟踪的精确性。在图 6 的视频序列中,目标分别发生了遮挡,缩放和旋转,故而两种算法在跟踪精度上都受到了一定影响,然而与原算法相比,本文算法仍要稍好。从两组实验结果整体而言, α 取 0.5 与 0.8 时的跟踪效果相差无几(其系统耗时并不相同),然而都比 α 取 0.2 时更加理想,因此 α 的选值对于跟踪效果很重要,值得讨论。

图 7 通过大量实验数据绘制而得。当 α 取值不大于 0.6 时,系统开销相对平均值较低(b),但是在小于约 0.3 处,累计有效的漂移量却明显低于平均值,这是由于迭代次数过少所致(a),容易造成系统提前收敛于一个非理想值;而当 α 取值大于约 0.3 时,系统的累计有效漂移量与平均值相当(a),然大于约 0.7 处,系统开销明显增大(b),这是由于过大的“小漂移”隶属度易造成目标在像素点之间来回振荡,从而导致时间上的负担加剧。综上分析, α 的取值在(0.3, 0.6)之间为宜,过大则常有耗时的迭代(d),而过小则累计有效漂移量不足,导致跟踪结果不够精确(c),图 6 的实验结果也表明了这点。

对两种算法计算量的量化比较见表 1。在相同漂移距离的前提下,本算法的平均迭代次数仅是原算法的一半,这是由于每次迭代中对非整数结果较正确的引导所致;同时因为

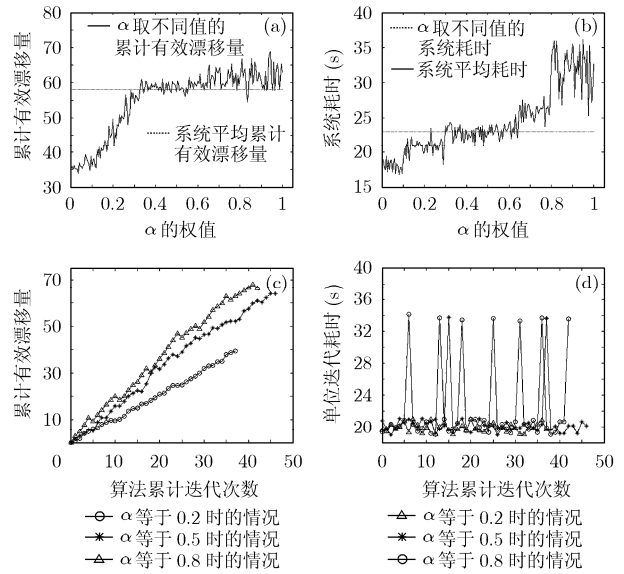


图 7 α 取不同值的实验结果

对不必要的 Bhattacharyya 计算的放弃,使得本算法平均迭代耗时也比原算法有较明显的降低;综合以上两者,本算法在相同漂移距离,相同迭代次数的前提下,耗时仅为原算法的三分之一,较好的减少了计算量。

表 1 两种算法的计算量比较

	原算法	本算法
平均单位漂移量 ¹⁾ 的迭代次数(n/m)	1.696	0.883
平均单位迭代的耗时(s/n)	29.697	20.760
平均单位迭代、单位漂移的耗时(s/n·m)	12.247	3.445

6 结束语

本文针对 Mean Shift 算法迭代过程中非整数中间结果的处理方法,引入了基于模糊隶属关系的改进算法。本算法以漂移相似度为主要参考依据,保证了最终结果的最大可靠性;以“鼓励漂移”策略为辅,较好的提高了跟踪的精确性;同时摒弃了原算法中不必要的 Bhattacharyya 距离计算,使得计算量也大为降低。实验结果表明,与原有算法相比,本文的算法有一定的优越性。

参考文献

- [1] Fukunage K and Hostetler D. The estimation of the gradient of a density function with application 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.

¹⁾两点之间的 Euclid 距离为 1 时称为单位漂移量,单位记为 m。

- [3] Comanicin D and Meer P. Mean shift analysis and application[J]. Proceedings of the Seventh IEEE International Conference Computer Vision[C], Kerkyra, Greece, 1999, 2: 1197-1203.
- [4] Comanicin D and Meer P. Mean shift: A robust application toward feature space analysis[J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 603-619.
- [5] Comanicin D and Meer P. Robust analysis of feature spaces: color image segmentation. Proc. 1997 IEEE conf. Computer Vision and Pattern Recognition[C]. San Juan. Puerto Rico: IEEE, 1997: 750-755.
- [6] Comanicin D, Ramesh V, and Meer P. Real-timetracking of non-rigid objects using mean shift. Proc of the IEEE Conference on Computer Vision and Pattern Recognition[C]. Hilton Head Island, USA, 2000, 2: 142-149.
- [7] Li Xiang-ru, Wu Fuc-hao, and Hu Zhan-yi. Convergence of a Mean Shift Algorithm[J]. *Journal of Software*, 2005, 16(3): 365-374.
- [8] Jia Jing-ping and Zhao Rong-chun. Tracking of objects in image sequences using bandwidth matrix mean shift algorithm. Proceedings of the 2004 International Conference on Signal Processing (ICSP'04)[C]. Beijing, 2004, 2: 918-921.
- 沈悦: 男, 1982年生, 硕士生, 研究方向为图像与计算机视觉跟踪.
- 吴小俊: 男, 1967年生, 博士, 教授, 博士生导师, 主要研究方向模式识别、计算机视觉与计算智能.