

基于 Mean Shift 的目标平移与旋转跟踪

王长军 朱善安

(浙江大学电气工程学院, 杭州 310027)

摘要 针对目前多数实时跟踪算法只能跟踪目标平移运动,不能跟踪旋转运动的问题,提出一种基于 Mean Shift 的快速旋转跟踪算法。该算法以目标区域的梯度方向分布(直方图)为特征,构造了可用 Mean Shift 算法寻优的相似度函数,将旋转跟踪转化为寻优问题,并利用 Mean Shift 寻优过程收敛速度快的特点,有效跟踪目标旋转。又提出交替迭代的方法,将旋转跟踪与 Meer 的平移跟踪算法融合起来,构造了可以同时跟踪目标旋转和平移完整跟踪算法。

关键词 梯度方向分布 Bhattacharyya 相似度 Mean Shift 旋转与平移 目标跟踪

中图分类号: TP391.4 **文献标识码**: A **文章编号**: 1006-8961(2007)08-1367-05

Mean Shift Based Targets' Rotation and Translation Tracking

WANG Chang-jun, ZHU Shan-an

(College of Electrical Engineering, Zhejiang University, Hangzhou 310027)

Abstract The present real-time tracking algorithm can perform well in term of translation tracking but can seldom do so in rotation tracking. A rotation tracking algorithm was proposed, which utilized the gray gradient direction distribution of the target region as the feature and constructed a similarity function that can be optimized by Mean Shift method. Therefore the rotation tracking was transformed into an optimization problem. Due to the fast convergence of the Mean Shift, this algorithm can be run in real-time. Combining the rotation tracking with the translation tracking algorithm proposed by Meer, a whole algorithm was obtained by alternate iteration, which can track both translation and rotation of the targets.

Keywords gradient direction distribution, Bhattacharyya similarity, Mean Shift, rotation and translation, targets tracking

1 引言

目标跟踪是通过对目标某一个或多个特征的跟踪实现的,目标特征主要分为两类,一类为基于边缘的特征,如目标轮廓,另一类为基于区域的特征,如目标像素值或其统计量。一般而言,区域特征比边缘特征包括更大的信息量,因此基于区域特征的跟踪算法有更强的鲁棒性。Meer 提出的基于 Mean Shift 的跟踪算法^[1]就是一种以目标区域像素值的概率分布为特征的跟踪算法。该算法由于采用了统计特征,因此对噪声有很强的鲁棒性,同时该算法构造了一个可以用 Mean Shift 算法进行寻优的相似度函数。由于 Mean Shift 本质上是最陡下降法,因此其寻优过程收敛速度快,使得该算法具有很好的实

时性。但正是由于采用像素值分布为特征,丢失了目标的几何信息,所以该算法只能跟踪目标的平移,无法跟踪目标的旋转。而对于某些应用^[2,3],跟踪目标的旋转是必要的。为此,本文提出一种基于 Mean Shift 的旋转跟踪算法,它通过引入目标区域的梯度方向分布(直方图)这个特征,同样构造出可用 Mean Shift 算法寻优的相似度函数,对目标的旋转进行有效跟踪,并提出交替迭代的方法,将 Meer 的平移跟踪算法融合起来,构造出可以对目标的旋转和平移进行实时跟踪的完整跟踪算法。

2 Mean Shift 及平移跟踪

2.1 Mean Shift 算法

Mean Shift 是一种非参数的密度估计算法,最早

收稿日期:2005-02-08;改回日期:2006-03-20

第一作者简介:王长军(1979~),男,现为浙江大学电气工程学院博士研究生。主要研究方向为模式识别、智能监视等。E-mail: wangzju@126.com

由 Fukunaga 提出,希望应用于模式识别问题。真正将该算法付诸应用的是 Cheng^[4]和 Comaniciu^[5]。

Mean Shift 算法可以表述如下^[5]:

设集合 $\{x_i\}_{i=1,\dots,n}$ 是 d 维空间 \mathbf{R}^d 中的 n 个点,则具有 Epanechinkov 核 $k(\|\cdot\|)$ 和带宽 h 的多变量密度估计

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

其中

$$k(x) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1-x) & \text{如果 } x < 1 \\ 0 & \text{其他} \end{cases}$$

c_d 为 d 维单位球体积。欲求 $\hat{f}(x)$ 极大值时的 x , 可以通过对下式迭代获得^[1]。

$$y_{j+1} = \frac{\sum_{i=1}^n x_i g\left(\frac{\|y_j - x_i\|^2}{h^2}\right)}{\sum_{i=1}^n g\left(\frac{\|y_j - x_i\|^2}{h^2}\right)}, j = 1, 2, \dots$$

其中, $g(\cdot) = -k'(\cdot)$ 。

2.2 基于 Mean Shift 的平移跟踪

在当前帧中定位目标的问题可以采用如下方法解决^[1]: 设 z 表示目标的某种特征, 如颜色或纹理等, 其密度分布为 q_z , 同时, 在位置 y 处的候选目标, 其特征 z 的分布为 $p_z(y)$, 那么跟踪问题转化为求位置 y , 使得 $p_z(y)$ 与 q_z 最相似。相似度采用 Bhattacharyya 系数表征。其定义为

$$\rho(y) = \int \sqrt{p_z(y)q_z} dz$$

实现上, 用归一化直方图代替连续的密度分布, 即

$$q_z = \{\hat{q}_u\}_{u=1,\dots,m}$$

$$p_z(y) = \{\hat{p}_u(y)\}_{u=1,\dots,m}$$

其中, $\sum_{u=1}^m \hat{q}_u = 1, \sum_{u=1}^m \hat{p}_u(y) = 1$ 。

现以目标像素的灰度为特征, 并引入 kernel 计算直方图 $\{\hat{q}_u\}$ 和 $\{\hat{p}_u(y)\}$ 如下:

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

$$\hat{p}_u(y) = C_h \sum_{i=1}^n k\left(\frac{\|y - x_i\|^2}{h^2}\right) \delta[b(x_i) - u] \quad (3)$$

其中

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

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

是归一化常数。 x_i^* 为以 0 为中心的目标像素坐标, x_i 为以 y 为中心的候选目标像素坐标。 $b(x)$ 表示 x 点像素在原始直方图 (未经 kernel 加权) 中对应值。

假设目标的初始位置为 \hat{y}_0 , 由 Taylor 公式可知

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

将式(2)代入上式, 得

$$\rho(y) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_0)\hat{q}_u} + \frac{1}{2} \sum_{u=1}^m w_i k\left(\frac{\|y - x_i\|^2}{h^2}\right) \quad (4)$$

其中, $w_i = \delta[b(x_i) - u] \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}}$ 。

由式(3)可知, 第 1 项独立于 y_0 , 因此要使 $\rho[\cdot]$ 极大, 只需使第 2 项极大。而第 2 项具有式(1)的形式, 可用 Mean Shift 算法取得使第 2 项极大时的 y 。迭代过程如下:

$$\hat{y}_{j+1} = \frac{\sum_{i=1}^n x_i w_i g\left(\frac{\|\hat{y}_j - x_i\|^2}{h^2}\right)}{\sum_{i=1}^n g\left(\frac{\|\hat{y}_j - x_i\|^2}{h^2}\right)}, j = 1, 2, \dots \quad (5)$$

3 基于 Mean Shift 的旋转跟踪

Comaniciu 提出的位置跟踪算法, 对噪声有很好的抑制能力, 同时, 即便目标发生旋转, 只要目标像素点在跟踪区域内 (一般为矩形或椭圆), 便不会影响直方图, 因此也能够很好地跟踪目标的中心位置, 但它不能跟踪目标的旋转角度, 本文称之为对目标旋转的“被动处理”。这种“被动处理”在以下两种情况下是不适用的。第一, 只要跟踪区域不是圆形或正方形, 就会由于目标旋转而无法覆盖全部的目标像素点, 同时使一些背景像素点进入跟踪区域, 这时, 直方图发生很大变化, 导致跟踪失败。第二, 某些应用需要跟踪目标的旋转^[2]。

Bradski^[2] 在人脸跟踪算法中提出一种基于目标统计特征的旋转跟踪方法, 它利用目标的水平和垂直二阶矩与目标对称轴角度之间关系跟踪目标的旋转。该方法要求目标形状是对称的, 因此应用受限。本文算法采用目标梯度方向的密度分布为特

征,并构造出可用 Mean Shift 算法寻优的相似度函数,利用 Mean Shift 算法收敛快速的特点,实现了对目标旋转的实时跟踪。该算法不要求目标形状具有对称性,克服了 Bradski 方法的缺点。

设 $\{\theta_u\}_{u=1,\dots,l}$ 为目标像素值梯度幅角离散值,它表示对 $(-\pi, \pi]$ 等分为 l 个角度,对应的密度分布(直方图)为 $B(\theta_u)_{u=1,\dots,l}$ 。构造基于 kernel 加权直方图,

$$\hat{Q}_u = cB(\theta_u)k\left(\frac{\|\theta_u\|^2}{\pi^2}\right) \quad (6)$$

$$\hat{P}_u(\theta) = c_p B(\theta_u)k\left(\frac{\|\theta - \theta_u\|^2}{\pi^2}\right) \quad (7)$$

其中

$$c = \frac{1}{\sum_{u=1}^l k\left(\frac{\|\theta_u\|^2}{\pi^2}\right)}$$

$$c_p = \frac{1}{\sum_{u=1}^l k\left(\frac{\|\theta - \theta_u\|^2}{\pi^2}\right)}$$

为归一化系数。 \hat{Q}_u 表示目标的密度分布, $\hat{P}_u(\theta)$ 表示候选目标的密度分布。定义相似度

$$\rho(\theta) = \sum_{u=1}^l \sqrt{\hat{P}_u(\theta)\hat{Q}_u}$$

图 1(a)为图 5 序列中目标梯度方向的初始分布,图 1(b)为第 2 帧中目标的梯度方向分布,图 2 为由这两个分布计算得到的 Bhattacharyya 系数,其最大值处对应于目标的旋转角度。

在初始位置估计 $\hat{\theta}_0$ 处进行 Taylor 展开,忽略高阶项,得

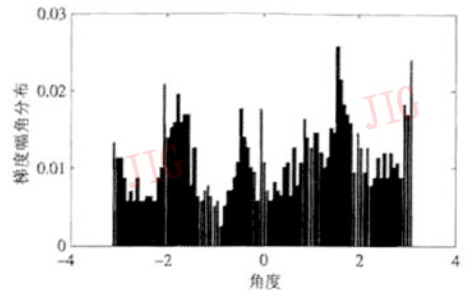
$$\rho(\theta) \approx \frac{1}{2} \sum_{u=1}^l \sqrt{\hat{P}_u(\hat{\theta}_0)\hat{Q}_u} + \frac{1}{2} \sum_{u=1}^l w_u k\left(\frac{\|\theta - \theta_u\|^2}{\pi^2}\right) \quad (8)$$

其中, $w_u = \frac{\hat{Q}_u}{\sqrt{\hat{P}_u(\hat{\theta}_0)}}$

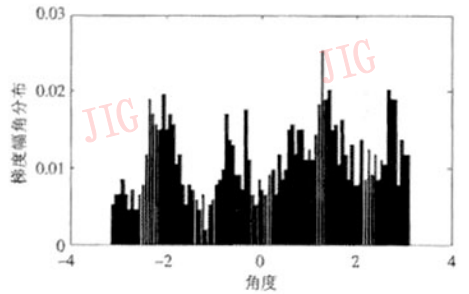
式(8)的第 2 项具有 kernel 密度估计的形式,因此可以用 Mean Shift 算法求出使 $\rho(\theta)$ 极大时 θ 的估计 $\hat{\theta}_j$,其迭代过程如下:

$$\hat{\theta}_{j+1} = \frac{\sum_{u=1}^l \theta_u g\left(\frac{\|\hat{\theta}_j - \theta_u\|^2}{\pi^2}\right)}{\sum_{u=1}^l g\left(\frac{\|\hat{\theta}_j - \theta_u\|^2}{\pi^2}\right)}, j = 1, 2, \dots, \quad (9)$$

本算法实际上是一个基于 Mean Shift 寻优技术的一维跟踪算法。



(a) 初始的分布



(b) 逆时针旋转后的分布

图 1 梯度幅角分布

Fig. 1 Distribution of gradient phase

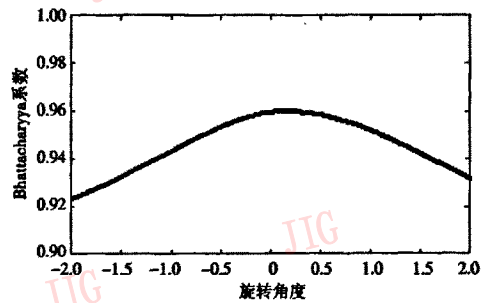


图 2 Bhattacharyya 系数

Fig. 2 Bhattacharyya coefficient

4 旋转与平移联合跟踪算法

上述平移和旋转跟踪算法虽然都是基于 Mean Shift 算法的,但由于采用不同的目标特征,因此其迭代过程是不同的。本文提出利用交替迭代平移和旋转跟踪算法的方法将两个跟踪算法融合起来。整个算法的具体实现步骤如下:

(1) 设定目标初始区域(矩形或椭圆),亦即设定了目标位置和方向角度的初始值 \hat{y}_0 和 $\hat{\theta}_0$ 。依

式(2)计算两个 \hat{q}_n 。

(2) 计算当前位置的 $\hat{p}_n(\hat{y}_j)$ ，并按式(5)迭代一次，得到 \hat{y}_{j+1} 。

(3) 计算 $p_n(\hat{\theta}_j)$ ，并按式(9)迭代一次，得到 $\hat{\theta}_{j+1}$ 。

(4) 计算

$$\Delta \hat{y}_{j+1} = \|\hat{y}_{j+1} - \hat{y}_j\| \text{ 与 } \Delta \hat{\theta}_{j+1} = \|\hat{\theta}_{j+1} - \hat{\theta}_j\|$$

如果 $\Delta \hat{y}_{j+1} < \varepsilon_y$ ，且 $\Delta \hat{\theta}_{j+1} < \varepsilon_\theta$ ，则结束，否则跳转至步骤 2。

这样的交替迭代使两具算法同步地收敛，从而实现了平移和旋转的同时跟踪。

5 实验

本文采用人脸和手的跟踪实验来验证算法有效

性。平移跟踪算法中，考虑到脸和手均为肤色目标，采用 HSV 空间^[6]中的 H 分量代替灰度分量。肤色在 H 分量上表现出一定的聚类性，一定程度上反映了像素的肤色概率，因此用此特征比较合适，当然也可以采用更好的肤色概率模型^[6]，以 H 分量为特征已经显示出很好的跟踪效果。在旋转跟踪算法中，采用 Sobel 算子计算目标灰度的水平和垂直方向梯度分量，然后用反正切计算幅角分量。

图 3 显示了手的平移跟踪效果，由于算法无法跟踪目标旋转，导致目标区域无法覆盖目标的全部像素点，最后算法收敛至背景中 H 分量相近的地方，跟踪失败。图 4 和图 5 显示了加入了旋转跟踪算法后的联合跟踪效果。由于加入了旋转跟踪算法，使目标区域可以覆盖几乎全部的目标像素点，从而保证的直方图的稳定性，达到了很好的跟踪效果。

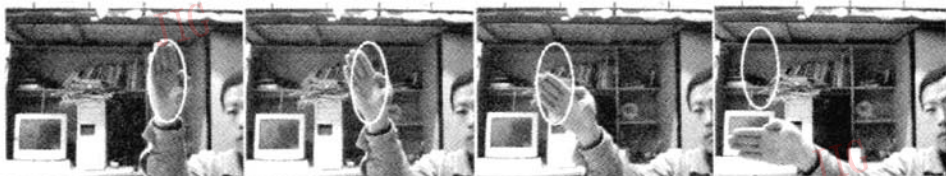


图 3 人手的平移跟踪

Fig. 3 Translation tracking of a face



图 4 人脸的旋转与平移联合跟踪

Fig. 4 Translation and rotation tracking of a face



图 5 人手的旋转与平移联合跟踪

Fig. 5 Translation and rotation tracking of a hand

6 结论

本文针对目前多数实时跟踪算法只能跟踪目标

平移运动，不能跟踪旋转运动的问题，提出一种基于 Mean Shift 的快速旋转跟踪算法。该算法以目标区域的梯度方向分布（直方图）为特征，构造了可用 Mean Shift 算法寻优的相似度函数，将旋转跟踪转化

为寻优问题,借助 Mean Shift 寻优过程收敛速度快的优点,本算法也有很快地处理速度。旋转跟踪的引入,不但满足了某些需要跟踪目标旋转应用的需求,而且消除了平移跟踪算法被动处理目标旋转的问题,提高了跟踪算法的稳定性。最后用交替迭代的方法将两个跟踪算法融合起来,构成了完整的平移旋转跟踪算法。

值得指出,由于本算法利用梯度方向分布特征跟踪目标的旋转,因此应用范围局限于灰度变化显著的目标和梯度方向分布不均匀的目标,这也是本算法的局限性。

参考文献 (References)

- 1 Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using Mean Shift[A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], South Carolina, USA, 2000: 142 ~ 149.
- 2 Bradski Gary R. Computer vision face tracking for use in a perceptual user interface[J]. Intel Technology Journal, 1998, 2(2): 1 ~ 15.
- 3 Gregory D, Belhumeur Hager Peter N. Efficient region tracking with parametric models of geometry and illumination [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(10): 1025 ~ 1039.
- 4 Cheng Yi-zong. Mean Shift, mode seeking, and clustering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(8): 790 ~ 799.
- 5 Comaniciu D, Meer P. Mean Shift: a robust approach toward feature space analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 24(5): 603 ~ 619.
- 6 Benjamin D, Boaz J, et al. Comparison of Five Color Models in Skin Pixel Classification[EB/OL]. <http://www.cs.uic.edu/~super/ratfg1999.pdf>