

基于各向异性核函数的均值漂移跟踪算法

齐苏敏^{①②} 黄贤武^① 伊怀峰^①

^①(苏州大学电子信息学院 苏州 215021)

^②(曲阜师范大学计算机科学学院 曲阜 273165)

摘要: 均值漂移算法是一种将迭代轨迹滑向局部邻域内均值的迭代算法, 已应用于目标跟踪领域。传统的均值漂移算法通常采用各向同性核函数进行跟踪, 但视频序列中的跟踪目标的结构随时间而变化, 尤其当目标结构快速变化时, 基于各向同性核函数的均值漂移跟踪算法常常会导致目标的丢失。该文采用各向异性核函数均值漂移算法实现目标跟踪, 由于该核函数的形状、大小、方向能自适应于目标局部结构的变化, 保证了跟踪效果的稳定性和鲁棒性。实验结果证明该算法是有效的。

关键词: 目标跟踪; 均值漂移; 各向异性核函数; 模板匹配

中图分类号: TN919.8

文献标识码: A

文献标识码: 1009-5896(2007)03-0686-04

Object Tracking by Anisotropic Kernel Mean Shift

Qi Su-min^{①②} Huang Xian-wu^① Yi Huai-feng^①

^①(School of Electronics & Information Engineering, Soochow University, Suzhou 215021, China)

^②(School of Computer Science, Qufu Normal University, Qufu 273165, China)

Abstract: Mean shift, an iterative procedure that shifts each data point to the average of data points in its neighborhood, has been applied to object tracking. However, with the changing structure of object in video sequences, traditional mean shift tracker by isotropic kernel often loses the object, especially when object structure varies fast. This paper implements object tracking with anisotropic kernel mean shift in which the shape, scale, and orientation of the kernels adapt to the changing object structure. The algorithm ensures tracking robust and real-time. Experimental results show it is effective.

Key words: Object tracking; Mean shift; Anisotropic kernel; Modal matching

1 引言

随着计算机视觉的发展, 实时的可视跟踪已经成为一个热门的研究领域。运动目标的实时跟踪在民用上和军事上都有广泛的应用, 同时也产生了各种实时目标跟踪算法, 一般来说, 跟踪算法可分为 3 类^[1]:

(1) 相关性跟踪 给定某个时刻的目标跟踪窗口, 根据最大相关性, 在下一个时刻寻找到与其最为匹配的窗口。

(2) 光流跟踪 从光流场近似获得序列图像的运动场, 由运动区域的运动规则指导目标跟踪。

(3) 特征跟踪 自下而上地提取跟踪目标的边缘、角点等特征, 然后用这些特征作为在后继帧中跟踪与匹配的主要原始数据。

均值漂移跟踪算法^[2,3]是从彩色图像中提取目标的表面颜色分布特征, 采用核函数直方图进行跟踪。核函数直方图是一种多元核函数无参数密度估计, 它同时适用于低维空间和高维空间。核函数直方图对部分遮挡具有鲁棒性, 旋转和

尺度不变性, 特别对于非刚体目标的跟踪具有更强的适应性。核函数窗宽^[1,4]是均值漂移算法的关键参数, 决定了目标跟踪的鲁棒性和准确性。本文采用了一种能自适应于跟踪目标尺度的各向异性核函数, 实现目标跟踪的稳健性、有效性和实时性。

2 基于各向同性核函数的均值漂移跟踪算法及其局限性

2.1 均值漂移算法

均值漂移算法^[2,3]是一种有效的无参数统计迭代算法, 它是使每一个点“漂移”到密度函数的局部极大值点。近年来, 均值漂移算法已广泛应用于计算机视觉领域, 如跟踪、图像分割、图像平滑、聚类分析等方面。

假设 X 是 d 维欧氏空间的总体, $\{\mathbf{x}_i, 1 \leq i \leq n\}$ 是来自总体 X 的独立分布样本集, 则 \mathbf{x} 点的多元核函数密度估计是

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

其中 $w(\mathbf{x}_i)$ 是表示 \mathbf{x}_i 的权重, 即 \mathbf{x}_i 发生的先验概率; $k(\bullet)$ 是

文献[3]中定义的轮廓函数; h 为窗宽; c 为归一化常数。Cheng^[2], Comaniciu^[3]等提出如下形式的均值漂移算法:

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^n \mathbf{x}_i k' \left(\left\| \frac{\mathbf{y}_j - \mathbf{x}_i}{h} \right\|^2 \right) w(\mathbf{x}_i)}{\sum_{i=1}^n k' \left(\left\| \frac{\mathbf{y}_j - \mathbf{x}_i}{h} \right\|^2 \right) w(\mathbf{x}_i)} \quad (2)$$

其中 $k'(\bullet)$ 是 $k(\bullet)$ 的导数。Cheng^[2]证明了由式(2)得到的均值漂移序列向着函数值 $\hat{f}(\mathbf{x})$ 升高的方向移动, 经过若干次迭代后, 收敛到稳定状态, 同时“漂移”到局部极大值点处。

由上可知, 均值漂移算法从样本集 $\{\mathbf{x}_i, 1 \leq i \leq n\}$ 出发给出密度函数 $f(\mathbf{x})$ 的非参数估计, 进而得到均值漂移序列, 对先验知识要求较少, 完全依靠样本集进行估计, 而且可用于任意密度函数的估计, 因此得到了广泛的应用。

2.2 基于各向同性核函数的均值漂移跟踪算法的局限性

Comaniciu^[3]等基于式(2)的均值漂移算法实现了目标的实时跟踪, 它由目标建模、相似性度量和模板匹配3部分组成。首先在当前帧中初始化候选目标模板, 计算与目标模板相似度函数; 然后利用均值漂移迭代快速的特点, 得到均值漂移序列, 实现目标的定位。

虽然均值漂移迭代算法收敛且迭代轨迹是光滑的, 但基于式(2)的均值漂移跟踪算法采用各向同性空间核函数, 而且核函数的窗宽在整个跟踪过程中保持不变, 忽略了运动目标的形状、大小、方向等因素在图像序列中的变化。这样, 在复杂的目标运动情况下, 极易出现定位不准确甚至造成目标的丢失。本文采用各向异性核函数均值漂移算法实现目标跟踪, 由于该核函数的形状、大小、方向能自适应于目标局部结构的变化, 保证了跟踪效果的稳定性和鲁棒性, 实现了视频序列中运动目标的有效跟踪。

3 基于各向异性核函数的均值漂移算法

3.1 各向异性核函数密度估计器

定义 形式上, 各向异性核函数多元密度估计器^[5, 6]表示为

$$\hat{f}_K(\mathbf{x}, \Sigma) = \sum_{i=1}^n K_{\Sigma}(\mathbf{x} - \mathbf{x}_i) w(\mathbf{x}_i) \quad (3)$$

$$K_{\Sigma}(\mathbf{x}) = c_{k,d} |\Sigma|^{-\frac{1}{2}} k(g(\mathbf{x}, \Sigma)) \quad (4)$$

其中 $c_{k,d}$ 为归一化系数, $g(\mathbf{x}, \Sigma) = \mathbf{x}^T \Sigma^{-1} \mathbf{x}$ 是 Mahalanobis 距离, \mathbf{x}^T 为 \mathbf{x} 的转置矩阵, Σ 为采样窗口内样本的协方差矩阵。由于 Σ 为矢量集的协方差矩阵的统计量, 反映了样本的局部结构, Σ 也称为窗宽(bandwidth)矩阵。

协方差矩阵 Σ 是对称正定的。若 $\Sigma = h^2 I_d$, 则 $K_{\Sigma}(\mathbf{x})$ 简化为各向同性核函数, $g(\mathbf{x}, \Sigma)$ 等值点轨迹为球体; 反之, $K_{\Sigma}(\mathbf{x})$ 为各向异性核函数, $g(\mathbf{x}, \Sigma)$ 等值点轨迹为椭球体, 该椭球体的主轴平行于协方差矩阵 Σ 的特征矢量方向, 其长度与相应的特征值的方根成正比。本文利用上述椭球体反

映跟踪目标的局部结构。

3.2 均值漂移

根据式(3)和式(4)在 \mathbf{x} 点得到密度估计 $\hat{f}_K(\mathbf{x}, \Sigma)$, 求其关于 \mathbf{x} 的偏导为

$$\frac{\partial \hat{f}_K(\mathbf{x}, \Sigma)}{\partial \mathbf{x}} = 2c_{k,d} \Sigma^{-1} \sum_{i=1}^n |\Sigma|^{-\frac{1}{2}} k'(g(\mathbf{x} - \mathbf{x}_i, \Sigma)) w(\mathbf{x}_i) (\mathbf{x} - \mathbf{x}_i)$$

令

$$h(\mathbf{x}) = -k'(\mathbf{x}) \quad (5)$$

$$H_{\Sigma}(\mathbf{x}) = c_{h,d} |\Sigma|^{-1/2} h(g(\mathbf{x}, \Sigma)) \quad (6)$$

由式(3)可知: $\hat{f}_H(\mathbf{x}, \Sigma) = \sum_{i=1}^n H_{\Sigma}(\mathbf{x} - \mathbf{x}_i) w(\mathbf{x}_i)$, 假定 $\hat{f}_H(\mathbf{x}, \Sigma) \neq 0$, 这时有

$$\frac{\partial \hat{f}_K(\mathbf{x}, \Sigma)}{\partial \mathbf{x}} = \frac{2c_{k,d}}{c_{h,d}} \hat{f}_H(\mathbf{x}, \Sigma) \cdot \Sigma^{-1} \left[\frac{\sum_{i=1}^n H_{\Sigma}(\mathbf{x} - \mathbf{x}_i) w(\mathbf{x}_i) \mathbf{x}_i}{\sum_{i=1}^n H_{\Sigma}(\mathbf{x} - \mathbf{x}_i) w(\mathbf{x}_i)} - \mathbf{x} \right] \quad (7)$$

由式(7)可得均值漂移矢量:

$$\Delta \mathbf{x} = \frac{\sum_{i=1}^n H_{\Sigma}(\mathbf{x} - \mathbf{x}_i) w(\mathbf{x}_i) \mathbf{x}_i}{\sum_{i=1}^n H_{\Sigma}(\mathbf{x} - \mathbf{x}_i) w(\mathbf{x}_i)} - \mathbf{x} \quad (8)$$

由式(7)可知, 均值漂移向量 $\Delta \mathbf{x}$ 正比于 $\frac{\partial \hat{f}_K(\mathbf{x}, \Sigma)}{\partial \mathbf{x}} / \hat{f}_H(\mathbf{x}, \Sigma)$, 即均值点 \mathbf{x} 总是沿着 $\hat{f}_K(\mathbf{x})$ 函数值升高的方向移动, 以此保证在其后的跟踪中, 由均值漂移向量确定当前帧的最佳匹配模板位置。

3.3 窗宽选择

根据Lindeberg理论^[1], 将式(3)的密度估计 γ 归一化为

$$\begin{aligned} \hat{f}_{\gamma K}(\mathbf{x}, \Sigma) &= |\Sigma|^{\gamma/2} \hat{f}_K(\mathbf{x}, \Sigma) \\ &= c_{k,d} |\Sigma|^{(\gamma-1)/2} \sum_{i=1}^n k(g(\mathbf{x} - \mathbf{x}_i, \Sigma)) w(\mathbf{x}_i) \end{aligned} \quad (9)$$

假定 $\hat{f}_{\gamma K}(\mathbf{x}, \Sigma) \neq 0$, 对式(9)关于 Σ 求偏导, 并由式(5)和式(9)可得

$$\begin{aligned} \frac{\partial \hat{f}_{\gamma K}(\mathbf{x}, \Sigma)}{\partial \Sigma} &= \frac{(1-\gamma)}{2} f_{\gamma K}(\mathbf{x}, \Sigma) \Sigma^{-1} \left[\frac{2}{1-\gamma} \right. \\ &\quad \left. \times \frac{\sum_{i=1}^n (\mathbf{x} - \mathbf{x}_i)(\mathbf{x} - \mathbf{x}_i)^T h(g(\mathbf{x} - \mathbf{x}_i, \Sigma))}{\sum_{i=1}^n k(g(\mathbf{x} - \mathbf{x}_i, \Sigma)) w(\mathbf{x}_i)} - \Sigma \right] \Sigma^{-1} \end{aligned} \quad (10)$$

由式(10)可得窗宽变化矩阵:

$$\Delta \Sigma = \frac{2}{1-\gamma} \times \frac{\sum_{i=1}^n (\mathbf{x} - \mathbf{x}_i)(\mathbf{x} - \mathbf{x}_i)^T h(g(\mathbf{x} - \mathbf{x}_i, \Sigma))}{\sum_{i=1}^n k(g(\mathbf{x} - \mathbf{x}_i, \Sigma)) w(\mathbf{x}_i)} - \Sigma \quad (11)$$

由式(10)可知,窗宽变化矩阵正比于 $[\partial \hat{f}_{\gamma, K}(\mathbf{x}, \Sigma) / \partial \Sigma] / \hat{f}_{\gamma, K}(\mathbf{x}, \Sigma)$,即窗宽变化矩阵沿着 $\hat{f}_K(\mathbf{x})$ 函数值升高的方向变化,以此保证在其后的跟踪中,由窗宽变化矩阵确定当前帧中的匹配模板的最佳尺度。

4 基于各向异性核函数的均值漂移跟踪算法

4.1 目标模板与候选目标模板

与文献[3]相同,我们用核函数直方图^[5]表示跟踪目标特征,并将彩色特征空间量化为 M 级。假定目标模板区域已被确定(手工或自动检测),区域的中心点为 \mathbf{x}_0 ,目标区域窗口的宽度与高度分别为 h_x , h_y ,那么,初始化窗宽矩阵

$$\text{为 } \Sigma_q = \begin{bmatrix} h_x^2 & 0 \\ 0 & h_y^2 \end{bmatrix}。$$

从初始化区域中提取目标颜色特征——核函数直方图为

$$\mathbf{q} = \{q_u\}_{u=1}^M, \text{ 满足 } \sum_u q_u = 1;$$

$$q_u = c_q \sum_{i=1}^{N_{\Sigma_q}} K_{\Sigma_q}(\mathbf{x}_0 - \mathbf{x}_i) \delta[b(\mathbf{x}_i) - u] \quad (12)$$

其中 c_q 为归一化常数。

目标跟踪算法的目的就是在任一包含跟踪目标的视频序列中找到最佳匹配。假设候选模板区域中心点为 \mathbf{y} ,窗宽矩阵为 Σ ,则候选模板核函数直方图为

$$\mathbf{p}(\mathbf{y}, \Sigma) = \{p_u(\mathbf{y}, \Sigma)\}_{u=1}^M \text{ 且满足 } \sum_u p_u = 1;$$

$$p_u(\mathbf{y}, \Sigma) = c_p \sum_{i=1}^{N_{\Sigma}} K_{\Sigma}(\mathbf{y} - \mathbf{x}_i) \delta[b(\mathbf{x}_i) - u] \quad (13)$$

其中 $b(\bullet)$ 为映射函数, $b(\mathbf{x}_i)$ 等于 \mathbf{x}_i 点的量化特征值, $\delta(\bullet)$ 为Kronecker delta函数, c_p 为归一化常数。

4.2 相似性测度与目标定位和窗宽选择

由于 q 和 $\mathbf{p}(\mathbf{y}, \Sigma)$ 分布密度的相似性程度可用这两个分布函数之间的距离^[7]来度量,于是,我们选择Bhattacharyya距离^[7]作为概率距离度量,即

$$J_B = -\ln \int [q_u p_u(\mathbf{y}, \Sigma)]^{1/2} d\mathbf{v} \quad (14)$$

令 $\rho(\mathbf{y}, \Sigma) = \int [q_u p_u(\mathbf{y}, \Sigma)]^{1/2} d\mathbf{u}$,则对 J_B 求取最小值,等价于求 $\rho(\mathbf{y}, \Sigma)$ 的最大值。对于离散图像的颜色分布密度,有

$$\rho(\mathbf{y}, \Sigma) = \sum_{u=1}^M [q_u p_u(\mathbf{y}, \Sigma)]^{1/2} \quad (15)$$

在 (\mathbf{y}, Σ) 处对 $\rho(\mathbf{y}, \Sigma)$ 进行一阶泰勒展开^[5, 6],则得

$$\rho(\mathbf{y}, \Sigma) \approx c_1 + c_2 \sum_{i=1}^{N_{\Sigma}} w_i K_{\Sigma}(\mathbf{y} - \mathbf{x}_i) \quad (16)$$

$$w_i = \sum_{u=1}^M \left[\frac{q_u}{p_u(\mathbf{y}, \Sigma)} \right]^{1/2} \delta[b(\mathbf{x}_i) - u] \quad (17)$$

其 c_1 , c_2 为常数。

求 $\rho(\mathbf{y}, \Sigma)$ 的最大值即为求式(16)中的核密度估计的最大值。由式(8)和式(11)我们得到

$$\mathbf{y}' = \frac{\sum_{i=1}^n H_{\Sigma}(\mathbf{y} - \mathbf{x}_i) w_i \mathbf{x}_i}{\sum_{i=1}^n H_{\Sigma}(\mathbf{y} - \mathbf{x}_i) w_i} \quad (18)$$

$$\Sigma' = \frac{2}{1-\gamma} \times \frac{\sum_{i=1}^n (\mathbf{y} - \mathbf{x}_i)(\mathbf{y} - \mathbf{x}_i)^T h(g(\mathbf{y} - \mathbf{x}_i, \Sigma)) w_i}{\sum_{i=1}^n k(g(\mathbf{y} - \mathbf{x}_i, \Sigma)) w_i} \quad (19)$$

式(18)与式(19)即为我们进行目标跟踪的两个基本迭代公式。

4.3 利用窗宽矩阵估算目标团块的形状

利用窗宽矩阵对目标团块进行形状估算。其估计过程如下:

设候选模板区域由 (\mathbf{y}, Σ) 确定,其中 $\Sigma = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$;令 $l_1 = a - y_1^2$; $l_2 = 2(b - y_1 y_2)$; $l_3 = c - y_2^2$,

则得椭圆长轴 l ,短轴 w 和长轴方向 θ ^[8]分别为

$$l = \sqrt{\frac{l_1 + l_3 + \sqrt{l_2^2 + (l_1 - l_3)^2}}{2}} \quad (20)$$

$$w = \sqrt{\frac{l_1 + l_3 - \sqrt{l_2^2 + (l_1 - l_3)^2}}{2}} \quad (21)$$

$$\theta = \frac{\arctan(l_2 / (l_1 - l_3))}{2} \quad (22)$$

由式(20)到式(22)可求得椭圆区域,同时将其外接矩形作为下次迭代的搜索区域。

4.4 算法实现

由上述过程可知,基于各向异性核函数的均值漂移算法的目标跟踪可归纳为如下步骤:

假定已知目标模板 $\mathbf{q} = \{q_u\}_{u=1}^M$,

(1) 在当前帧中初始化候选模板 $\mathbf{p}(\mathbf{y}_0, \Sigma_0) = \{p_u(\mathbf{y}_0, \Sigma_0)\}_{u=1}^M$,将 (\mathbf{y}_0, Σ_0) 确定的椭圆区域标记为 O_0 。

(2) 由式(17)求得图像加权矢量。

(3) 由式(18)和式(19)计算新的目标模板中心点 \mathbf{y}_1 及窗宽矩阵 Σ_1 。

(4) 由式(20)到及式(22)式计算二维椭圆区域 O_1 。

(5) 将(4)中得到椭圆 O_1 与 O_0 比较,判断是否为相同的区域,若是,则结束迭代;否则令 $\mathbf{y}_0 \leftarrow \mathbf{y}_1$, $\Sigma_0 \leftarrow \Sigma_1$, $O_0 \leftarrow O_1$,执行(2)。

注:在算法具体实现时,前一帧的最后迭代结果作为当前帧的初始化值。

5 实验结果

采用不同的各向异性核函数对不同的视频序列中的运动目标进行跟踪实验,均取得了很好的跟踪结果。若采用高

斯核函数, 取 $\gamma = 1/7$; 若采用 Epanechnikov 核函数, 取 $\gamma = 1/6$ 。

图 1 所示的 hand 序列用来验证算法的有效性, 序列中手的位置和形状快速变化。我们对彩色空间 HSV 中 H 和 S 信息进行 16×16 级量化, 采用文献[3]的跟踪算法和我们的跟踪算法进行实验, 其结果分别如图 1(a)~图 1(c), 图 1(d) ~图 1(f)所示。两种算法的迭代次数如图 2 所示。前一种方法平均迭代次数(图 2(a)所示)为 2.5863, 后一种方法平均迭代次数(图 2(b)所示)为 1.835。实验结果表明, 本文算法比文献[3]的算法迭代次数少近 1 次, 而且本文算法能够根据图像序列中运动目标的形状自适应地调节跟踪尺度和方向。

本文的跟踪算法对真实场景中的机动车辆的跟踪可同样达到实时跟踪的效果, 如图 3 所示, 平均迭代次数为 2.5。从实验结果可以看出, 跟踪器能够根据图像序列中运动目标的结构自适应地调节尺度。

6 结束语

本文简洁地给出了基于各向异性核函数的均值漂移算法的理论证明, 同时总结了基于各向异性核函数的均值漂移算法用于目标跟踪的步骤。实验结果证明该算法能对快速运动目标进行准确、自适应的定位, 同时能很好地估算运动目标的形状。理论分析和实验结果充分说明基于各向异性核函数的均值漂移跟踪算法对视频序列中的运动目标的跟踪具有很好的准确性、自适应性、实时性和鲁棒性。

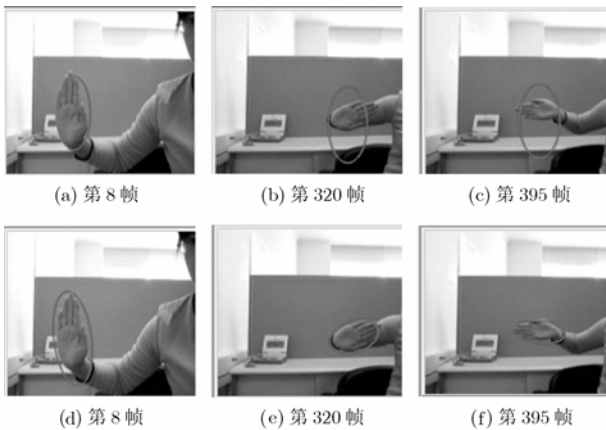


图 1 Hand 序列跟踪效果比较

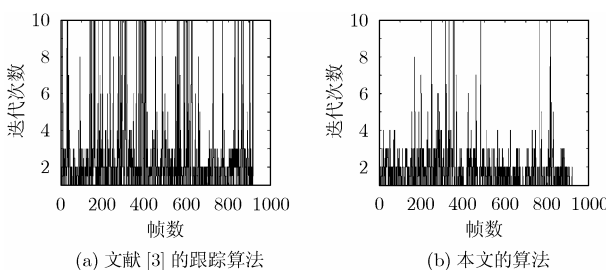


图 2 Hand 序列迭代次数比较



图 3 Car 序列实时跟踪效果

参考文献

- [1] Bretzner L and Lindeberg T. Feature tracking with automatic selection of spatial scales[J]. *Computer Vision and Image Understanding*, 1998, 71(3): 385–392.
- [2] Cheng Yizong. 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, and Meer P. Kernel-based object tracking[J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2003, 25(5): 564–577.
- [4] Zivkovic Z and Krose B. An EM-like algorithm for color-histogram- based object tracking[C]. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington DC, USA, 2004, vol.1: 798–803.
- [5] Chen H F and Meer P. Robust computer vision through kernel density estimation[C]. *Computer Vision - ECCV 2002. 7th European Conference on Computer Vision Proceedings*, Copenhagen, Denmark, Part I (Lecture Notes in Computer Science Vol.2350), 2002: 236–250.
- [6] Wang J, Bo T, Xu Y Q, and Cohen M. Image and video segmentation by anisotropic kernel mean shift[C]. *Computer Vision - ECCV 2004. 8th European Conference on Computer Vision, Proceedings (Lecture Notes in Comput. Sci.Vol.3022)*, Prague, Czech Republic, 2004, Vol.2: 2638–249.
- [7] 边肇祺, 张学工. 模式识别[M]. 北京: 清华大学出版社, 2000: 180–181.
- [8] Bradski G R. Computer vision face tracking as a component of a perceptual user interface[C]. *Proceedings Fourth IEEE Workshop Applications of Computer Vision*, Berlin, Germany, Oct. 1998: 214–219.

齐苏敏: 女, 1976年生, 博士生, 讲师, 研究方向为视频跟踪、机器视觉。
 黄贤武: 男, 1941年生, 博士生导师, 教授, 主要研究方向为数字图像处理, 模式识别。
 伊怀峰: 男, 1980年生, 硕士, 研究方向为视频分割, 模式识别。