

一种基于标记修正和样本更新的目标跟踪算法

卫保国, 曹慈卓

(西北工业大学 电子信息学院, 西安 710129)

摘要: 针对半监督自训练框架下进行目标跟踪的误差累积问题, 提出一种结合标记修正与区域置信度样本更新的自适应跟踪算法。该算法将视频序列中的目标跟踪视为两类模式即目标与背景的分类问题, 在半监督自训练框架下, 选择SVM分类器分类目标与背景, 结合K近邻和最小距离分类进行标记修正, 并基于区域置信度提取新的样本更新分类器。实验结果显示, 该方法有效改善了由于误差累积导致的漂移问题和目标遮挡后的跟踪失败。

关键词: 目标跟踪; 半监督自训练; 标记修正; 区域置信度

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1001-3695(2012)05-1963-04

doi:10.3969/j.issn.1001-3695.2012.05.096

Object tracking based on label correction and sample updating

WEI Bao-guo, CAO Ci-zhuo

(School of Electronic & Information, Northwestern Polytechnical University, Xi'an 710129, China)

Abstract: To solve the error accumulation problem of object tracking in the framework of semi-supervised self-training, this paper proposed an adaptive tracking algorithm with label correction and sample updating. It treated the object tracking as a binary classification problem between objects and backgrounds, used SVM to classify the image block, corrected the classification result with K-nearest neighbor and minimum distance classification, and extracted new samples based on regional confidence to update the classifier. Experimental results show that the proposed method can effectively avoid the drift problem due to error accumulation and tracking failure since object occlusion.

Key words: object tracking; semi-supervised self-training; label correction; regional confidence

0 引言

目标跟踪是机器视觉领域的一个重要分支, 被广泛用于军事、交通管制、安全监控等领域^[1]。目标跟踪通常被看做是一个搜索与匹配问题^[2], 并通过预测缩小搜索范围, 加快跟踪速度。将目标跟踪视为一个在线分类问题, 是近年来目标跟踪研究的新途径。此类方法寻找能够将目标与背景分离的分类器, 适用于背景凌乱的视频, 解决了目标变化复杂而很难建模的问题, 并且可以实时更新分类器以适应目标外观的变化^[3]。

文献[4]用离线学习的支持向量机(SVM)结合基于光流的方法进行目标跟踪。文献[5]利用基于像素对的随机树分类器对关键点进行分类。这两种方法采用固定的跟踪器, 为了应对目标的变化及复杂的背景需要大量的标记样本, 而手动标记样本的成本巨大, 有时也不现实。为此, 可以考虑利用大量未标记样本来辅助训练分类器以提高其性能, 这类同时利用标记样本与未标记样本的机器学习方法称做半监督学习。在目标跟踪中, 人工标记往往只用于初始的几帧, 跟踪算法必须不断地从跟踪结果中学习目标和背景的变化信息, 因此目标跟踪可以被视做一个半监督学习问题。基于半监督学习的目标跟踪是目前一个新的研究方向, 已有成果显示了其独特的效果。文献[6]提出一种在线特征选择的方法, 提取最能判别目标和背景的颜色特征用于跟踪。文献[7]采用集成学习的思想, 通

过在线训练多个弱分类器, 利用AdaBoost把多个弱分类器合成一个强分类器, 利用强分类器区分出前景和背景。这两种方法的核心是用分类得到的结果来自我更新分类器, 即应用了半监督学习中的self-training思想, 使跟踪性能大大改进。但这种训练方法很容易积累误差, 使得分类器精度下降, 最后导致跟踪失败; 且由于训练多个分类器所需运算时间较长, 影响了算法的跟踪效率。目前基于半监督学习的目标跟踪尚未有研究解决遮挡和偏移问题的相关论文。

为了减少自学习框架下的误差积累, 本文提出了一种结合标记修正与区域置信度样本更新的目标跟踪算法, 算法流程如图1所示。使用SVM分类器分类目标与背景, 通过标记修正技术对所得置信图分类结果进行修正。通过目标局部区域与背景的相似程度判断是否存在遮挡, 基于区域置信度对样本更新以在线地训练分类器来适应目标与环境的变化, 有效地减少了误差积累, 对目标遮挡、干扰、背景变化复杂等情况具有较好的跟踪效果。通过在多个典型的视频中对行人等目标的跟踪, 验证了本文算法的有效性。

1 自训练学习与目标跟踪

基于半监督自训练学习的跟踪方法首先利用已标记的样本来建立初始分类器, 用此分类器对未标记样本进行标记得到置信图, 根据置信图确定目标的新位置; 将置信度高的被标记

收稿日期: 2011-08-25; 修回日期: 2011-10-12

作者简介: 卫保国(1970-), 男, 陕西乾县人, 副教授, 博士, 主要研究方向为图像处理、模式识别与计算机视觉(wbg@nwpu.edu.cn); 曹慈卓(1986-), 女, 陕西凤翔人, 硕士, 主要研究方向为图像处理、模式识别。

样本加入原有标记样本重新训练一个新分类器,再不断重复这个过程,如图 2 所示。其问题是:由于目标跟踪中,目标和环境的不确定因素要比普通分类问题更加复杂,如目标的形变;背景复杂且随摄像机的运动不断变化;目标被遮挡。这些因素容易导致分类误差,随着跟踪帧数的增加会造成错误的积累和扩散^[8]。本文采用结合 K 近邻规则与最小距离分类的标记修正和基于区域置信度的样本更新方法来解决这些问题。

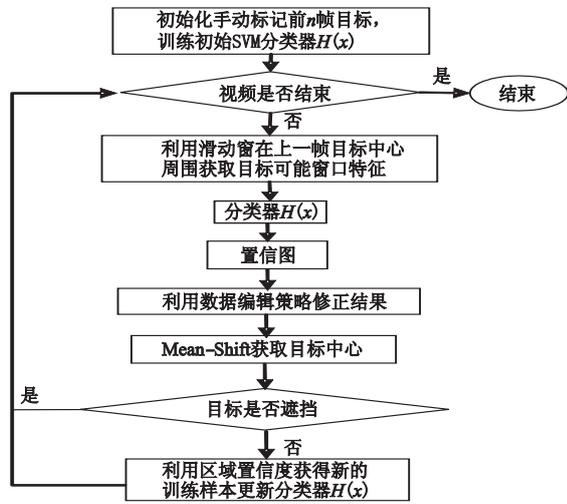


图1 本文跟踪算法的流程

1.1 分类器选择

目标跟踪中,在初始几帧手动标记或通过其他检测方法得到训练样本,样本的数量相对较少,而支持向量机分类器在解决小样本、非线性及高维模式识别中有许多特有的优势^[9]。支持向量机最早由 Vapnik 在 1995 年提出,与传统的分类器相比,支持向量机方法是建立在统计学习理论的 VC 维理论和结构风险最小原理基础上的,根据有限的样本信息在模型的复杂性和学习能力之间寻求最佳折中,解决了神经网络难以控制自身推广能力的问题。

对于两类分类问题,SVM 的目标是寻找一个具有最大间隔的最优分类面,假定 n 个 k 维训练样本 $(x_1, y_1) \dots (x_n, y_n)$, $x_i \in R^k, y_i \in \{-1, 1\}$ 是分类标签。SVM 分类器的决策函数为

$$g(x) = \text{sgn}(K\langle w, x \rangle + b) = \text{sgn}(K\langle \sum_{i=1}^n (a_i y_i x_i), x \rangle + b) = \text{sgn}(\sum_{i=1}^n (a_i y_i) K\langle x_i, x \rangle + b) \quad (1)$$

其中: $w = \sum_{i=1}^n (a_i y_i x_i)$; x 是变量; x_i 是已知的样本; a_i 为拉格朗日乘子,一小部分不等于 0 的拉格朗日乘子对应的样本点 x_i 唯一地确定了分类函数,这些样本就是支持向量(SV); $K(x_i, x)$ 是核函数,接受低维空间的样本输入,将输入值映射到高维特征空间,并在此空间用支持向量构造一个分类超平面。

$g(x)$ 中 w 和 b 的值通过求解下列优化问题得到:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i$$

$$\text{subject to } y_i [(w x_i) + b] \geq 1 - \xi_i$$

$$\xi_i \geq 0 \quad (2)$$

其中: ξ_i 是松弛变量, C 是惩罚因子。由于在绝大多数情况下,数据是线性不可分的,这时通过引入松弛变量实现最少错分样本和最大分类间隔之间的折中,来得到广义最优分类面。其

中, $C > 0$ 是一个常数,它控制对错分样本惩罚的程度。

进行 SVM 学习时,如果样本数目不均衡,生成的分类面会偏向样本数量少的类。这是因为式(2)中在求解最优化问题时使用了相同的惩罚系数 C ,分割线偏向样本密度更小的一方,能够使总的错分样本数目减少,从而在间隔尽可能大的情况下降低了结构风险水平。视频帧中采样时,背景样本数量往往多于目标样本数量,标准的支持向量机基于分类精度评价方法显然不合适,例如数据集不均衡,仅有 5% 的正例,那么当分类器把所有测试样本分为负例,便可获得很高的分类精确度(95%)。Veropoulos 等人^[10]提出对不同的类设置不同的惩罚参数 C ,即对正类和负类赋予不同的代价,能够有效地根据不同样本的错分代价进行超平面的优化。本文根据实际正负样本的数量设置惩罚系数,以避免样本数量不均衡情况下的错分,即构造如下二次规划问题:

$$\min \frac{1}{2} \|w\|^2 + C^+ \sum_{\{i|y_i=+1\}} \xi_i + C^- \sum_{\{i|y_i=-1\}} \xi_i$$

$$\text{subject to } y_i [w \cdot x_i + b] \geq 1 - \xi_i$$

$$\xi_i \geq 0 \quad (3)$$

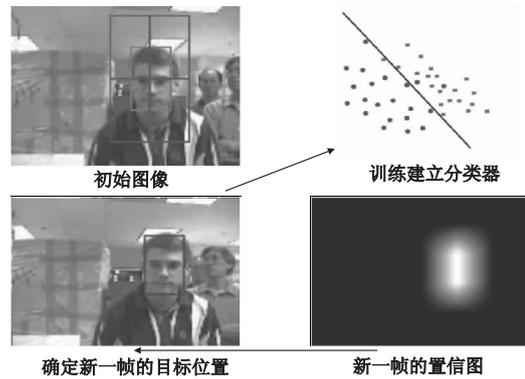


图2 半监督自训练学习的目标跟踪方法

1.2 结合 K 近邻规则与最小距离分类的标记修正

本文采用的特征为像素的颜色信息。分类时,先把属于目标的样本点视为正样本,把属于背景的样本点视为负样本,这样,样本点集合 $\{x_i\}$ 的标记 $\{y_i\}$ 就为 $+1$ 或 -1 。

在 self-training 训练过程中,SVM 分类器误标记是不可避免的,从而导致置信图中包含噪声,影响目标中心点的确定。如果将这些噪声引入到训练集中,随着迭代过程的继续,分类性能将持续降低。因此,找到这些误标记数据并进行修正是改善跟踪结果的有效途径。本文提出一种结合 K 近邻规则^[11]和最小距离分类^[12]的标记修正方法来识别并修正被错误标记的样例。

设在视频第 t 帧中,训练集为 L_t ,利用滑动窗对第 t 帧分类所得数据集为 L_t' ,计算 L_t 中正类与负类的均值向量 $z_i (i = \pm 1)$ 。对新标记的 L_t' 中的每个样例 x 按 K 近邻规则从 $L_t \cup L_t'$ 中选取它的 $k (k$ 为偶数) 个近邻,当与 x 标记相同的数量小于 $k' (k' \leq k/2)$ 时,将该样例标志为可疑的样例。分别计算此可疑样例与正负类的距离

$$d_i = |x - z_i| (i = +1, -1) \quad (4)$$

利用最小距离分类法对可疑样例的标记重新修正,即若 $d_{+1} < d_{-1}$, x 标记为正,否则标记为负。

1.3 Mean-Shift 算法简介

Mean-Shift^[13]算法是一种有效的统计迭代算法,是由

Fukunaga 等人于 1975 年首先提出的。Mean-Shift 算法的基本思想是,通过反复迭代搜索特征空间中样本点最密集的区域,先算出当前点的偏移均值,移动该点到其偏移均值,然后以此为新的起始点继续移动,直到满足一定的条件结束。其显著的优点是算法计算量小、简单易实现。

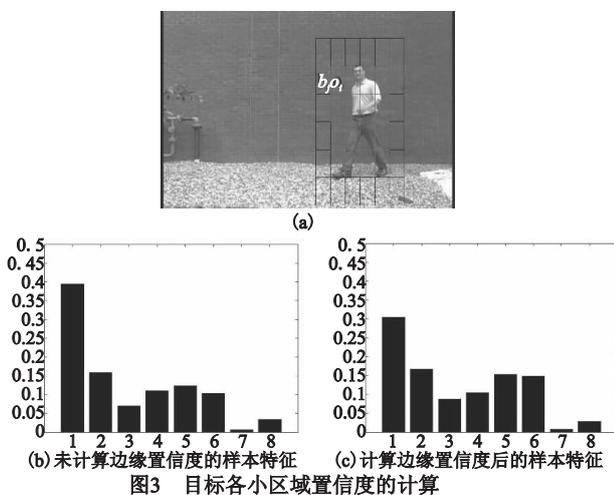
1.4 基于区域置信度样本更新的半监督跟踪算法

在半监督自训练框架下的跟踪过程中,根据前几帧图像中已知的目标与背景计算出正负样本的统计信息建立分类器,在新的一帧中识别非样本像素的归属,再根据自身的预测能力实时学习目标背景的变化信息在线的更新分类器,以适应目标的形态变化和背景的复杂性。因此,训练阶段的样本质量决定着分类的结果,也决定着从分类过程所获得的信息价值。为保证分类结果有较高精度,训练数据必须具有完整性和代表性,即训练样本的统计特征与该类型总体统计特征相接近。当选取的训练样本中含有不合要求的像素(如异类像素)时,需要对训练样本进行纯化。

视频中的目标往往是不规则的形状,跟踪中的目标框边缘区域含有背景像素,即噪声点,这些噪声点降低了训练样本的准确性,加重了分类器的分类误差,随着误差的积累导致漂移。如果在目标遮挡情况下错误地更新样本,会导致分类器丢目标。

为了有效地去除这些噪声点以纯化训练样本,本文根据图像的区域性,即图像中普遍存在着的相邻像素点之间具有某种相似性的特点,通过计算目标框边缘不同区域与背景相似性来定义边缘各区域为目标即正类的置信度,与背景相似度大的区域为正类的可信度较低,在下一步提取新的样本时可将这些区域从目标中排除。

如图 3 所示,(a)中矩形框的目标区域掺杂了背景像素,目标区域内存在与其相邻背景相似性较大的小区域,因此,需要对样区进行纯化处理;(b)(c)分别为样区纯化前后所得到的颜色特征直方图。



将目标框划分为若干大小相等的区域 o_i ,假设其边缘小区域的个数为 m ,将周围背景区域划分为若干个与 o_i 面积相等的区域 b_j ,统计 o_i 与 b_j 的颜色直方图分别为 $his(o_i)$ 与 $his(b_j)$ 。通过下式得到 m 个 o_i 与其紧邻的 b_j 的相似性值:

$$D_i(o_i, b_j) = \| his(o_i) - his(b_j) \| = \sqrt{\sum_{i=1}^{bin} (his(o_i) - his(b_j))^2} \quad (5)$$

其中: $i=1, 2, \dots, m$ 。设置阈值 T_1 ,当 $D_i(o_i, b_j) \leq T_1$ 时, o_i 区域所含背景像素数较多,丢掉这一区域;当 $D_i(o_i, b_j) > T_1$ 时, o_i 所含背景像素数较少,将其作为目标区域,与目标框内其他区域结合计算样本的特征向量。可以表示为

$$o_i = \begin{cases} 1 & \text{inside}(A) \text{ and } D_i(o_i, b_j) > T_1 \\ -1 & \text{otherwise} \end{cases} \quad (6)$$

假设 S 为边缘小区域判定为背景的个数,即

$$S = \sum_{i=1}^m (\delta(o_i + 1)) \quad (7)$$

设置阈值 T_2 ,当 $S \geq T_2$ 时,不提取新样本;当 $S < T_2$ 时,提取新样本,进行分类器更新。

如图 4 所示,改进前由于正样本特征向量中附带大量背景信息,容易使分类器误将背景判定为目标,改进后置信图会变得紧凑,使重心位置更接近目标中心。

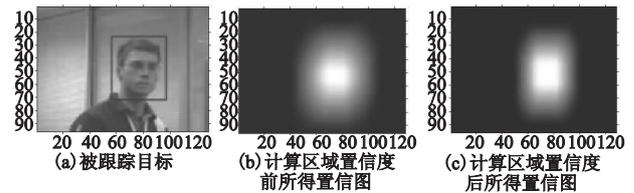


图4 计算区域置信度前后的置信图对比

上述过程中将颜色直方图量化为 8 个 bin。用量化的方法计算小区域的颜色直方图产生的问题是,由于小区域的直方图是稀疏的,当光照变化或噪声引起像素颜色发生微小变化时,其分布极易从其颜色区间落入相邻的颜色区间,从而造成两幅图像的颜色直方图形状几乎相同,只是互相错开了几个 bin,这时采用欧氏距离计算两者的相似度,会得到很小的相似度值。为了克服这个缺陷,可以对颜色直方图事先进行平滑过滤^[14],让每个 bin 中的像素对于相邻的几个 bin 有所贡献。

1.5 算法流程

输入:视频帧 I_1, I_2, \dots , 前 n 帧的目标位置和范围。

输出:每帧的目标位置和范围 A_{n+1}, A_{n+2}, \dots 。

1)初始化 在初始图像中选择包含目标的矩形框,根据其中数据提取目标特征,标记为 1,将目标框之外更大一些的区域设为背景,提取背景特征,标记为 -1;用标定好的前 n 帧图像的样本训练支持向量机分类器 $H(x)$ 。

2)在线跟踪 从 $n+1$ 帧到视频结束:

a)在第 $n+l$ 帧图像中,将上一帧目标框 A_{n+l-1} 周围区域中的每个位置作为当前帧的目标备选位置。

b)提取每个备选位置的颜色特征对其进行分类,得到整个备选区域的目标位置置信分布图。

c)结合 K 近邻规则和最小距离分类器对所得置信图进行修正。

d)在修正后的置信图上将 A_{n+l-1} 中心作为目标中心的初始位置,运用 Mean-Shift 算法确定第 $n+l$ 帧的目标中心,得到 A_{n+l} 。

e)在目标框 A_{n+l} 中分别根据式(5)(7)计算 $D_i(o_i, b_j)$ 与 S ,判断目标是否遮挡。如果 $S \geq T_2$,转步骤 a);否则,将提取的新的正负样本加入原始样本中,并去掉从最早一帧中提取的正负样本重新训练分类器。

f)在新的一帧中重复 a) ~ e)。

2 实验结果及分析

为了验证本文算法的性能,在多个典型的视频序列上进行大量实验。实验平台为 MATLAB 2008a, Pentium® Dual-Core 处理器,2 GB 内存,操作系统为 Windows XP。

实验时,手动标记视频序列的前 5 帧,计算正负样本训练 SVM 分类器。滑动窗搜索范围取 4 倍跟踪框大小。每帧中正负样本的比例数为 1:4,支持向量机正负类别惩罚系数比设为 4:1。 T_1 根据经验设为 0.5, T_2 设为边缘像素数的 50%。

图 5 是在序列 seq_cubicle 中的目标跟踪,图像分辨率为 97×129 。视频序列中目标发生了严重的遮挡,即人的头部被隔板遮挡。从结果可以看出,目标被遮挡后会出现一些误差,但是遮挡结束后,能够继续捕捉到目标。

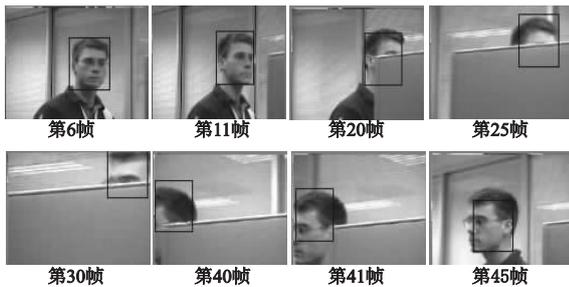


图5 本文算法在序列seq_cubicle中的跟踪结果

图 6 是在序列 seq_villains 中的目标跟踪,将视频中人的头部作为目标,图像分辨率为 97×123 。视频序列中,一个人在房间行走,随着目标的运动背景不断变化,且整个过程都有其他人的干扰。图 6(a)显示未进行标记修正与根据区域置信度更新样本的跟踪结果,由于背景复杂且视频中目标发生较大的变化,出现了跟踪误差,并随着误差的积累导致跟踪结果错误;(b)中未出现(a)的错误,并且本文算法有效地避免了其他相似目标的干扰。算法改进前后跟踪轨迹比较如图 7 所示。在图 7 中,横轴表示图像帧数,纵轴分别表示 X、Y 方向的运动距离。本文将跟踪的目标中心到实际目标中心的距离作为误差,前 45 帧未改进算法未跟丢目标时的跟踪误差均值为 4.59,改进后算法的误差均值为 2.63,可见改进后的算法跟踪结果精度有所提高,并且不易跟丢目标。



图6 在序列seq_villains中的跟踪结果

3 结束语

本文介绍了一种在半监督 self-training 框架下,结合标记修正与区域置信度在线自适应更新分类器的目标跟踪算法。这种方法先把跟踪问题转换为一个两类分类问题,采用 SVM 作为分类器,通过合理的参数设置训练具有较高精度的 SVM 分类器,利用这个分类器来对新一帧中的样本进行分类得到置

信图,并对所得结果进行修正,运用均值漂移找到置信图的重心,也就是目标的新位置。为了与目标和背景的变化相适应,结合区域置信度在每一帧中提取新的样本训练一个新的分类器。实验结果证明,该方法有效避免了误差累积导致的漂移问题和目标遮挡导致的跟踪失败,能在各种场景下实现稳定有效的跟踪。然而本文算法中使用的跟踪框大小是固定的,不能随着目标的形变而改变大小和角度,因此下一步的工作是着重解决跟踪框的缩放和旋转,以及多目标跟踪问题。

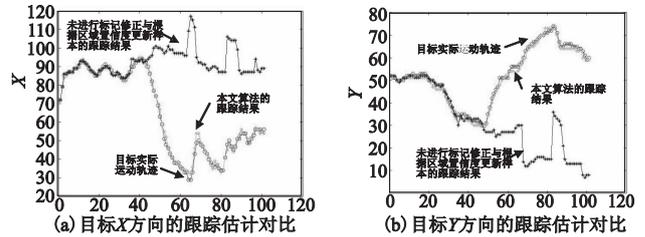


图7 目标位置的跟踪轨迹

参考文献:

- [1] 夏瑜. 视觉跟踪技术的研究与应用[D]. 无锡:江南大学,2010.
- [2] 徐杨,吴成东,陈东岳. 基于视频图像的交通事件自动检测算法综述[J]. 计算机应用研究,2011,28(4):1206-1210.
- [3] TANG Feng, BRENNAN S, ZHAO Qi, et al. Co-tracking using semi-supervised support vector machines[C]//Proc of the 11th IEEE International Conference on Computer Vision. 2007:1-8.
- [4] AVIDAN S. Support vector tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence,2004,26(8):1064-1073.
- [5] LEPETIT V, LAGGER P, FUA P. Randomized trees for real-time keypoint recognition[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society,2005:775-781.
- [6] COLLINS R T, LIU Yan-xi. On-line selection of discriminative tracking features[J]. IEEE Trans on Pattern Analysis and Machine Intelligence,2005,27(10):1631-1643.
- [7] AVIDAN S. Ensemble tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence,2007,29(2):261-271.
- [8] CHAPELLE O, SCHOLKOPF B, ZIEN A. Semi-supervised learning[M]. London: MIT Press,2010.
- [9] VLADIMIR N V. The nature of statistical learning theory[M]. Berlin: Springer-Verlag,2000.
- [10] VEROPOULOS K, CAMBELL C, CRISTIANINI N. Controlling the sensitivity of support vector machines[C]//Proc of the International Joint Conference on AI. 1999:55-60.
- [11] DENG Chao, GUO Mao-zu. Tri-training and data editing based semi-supervised clustering algorithm[C]//Proc of the 5th Mexican International Conference on Artificial Intelligence. Berlin: Springer-Verlag, 2006:641-651.
- [12] JAIN A K, DUIN R P W, MAO Jiang-chang. Statistical pattern recognition: a review[J]. IEEE Trans on Pattern Analysis and Machine Intelligence,2000,22(1):4-37.
- [13] CHENG Yi-zong. Mean shift, mode seeking, and clustering[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1995,17(8):790-799.
- [14] 何明一,卫保国. 数字图像处理[M]. 北京:科学出版社,2008:125-127.
- [15] 张学工. 模式识别[M]. 北京:清华大学出版社,2010.