

L₂范数正则化鲁棒编码视觉跟踪

袁广林^{*①} 薛模根^{②③}

^①(解放军陆军军官学院十一系 合肥 230031)

^②(解放军陆军军官学院科研部 合肥 230031)

^③(合肥工业大学计算机与信息学院 合肥 230009)

摘要: 针对基于稀疏表示的视觉跟踪计算效率低和易于产生“模型漂移”的不足, 该文提出一种基于L₂范数正则化鲁棒编码的视觉跟踪方法。该方法利用L₂范数正则化鲁棒编码求解候选目标的编码系数, 以粒子滤波为框架, 利用候选目标的加权重建误差建立似然模型跟踪目标。为了适应目标的变化并克服“模型漂移”问题, 利用L₂范数正则化鲁棒编码估计当前目标的加权矩阵用于遮挡检测, 根据遮挡检测结果实现模型更新。对提出的跟踪方法进行实验的结果表明: 与现有跟踪方法相比, 该方法具有较优的跟踪性能。

关键词: 视觉跟踪; L₂范数正则化; 鲁棒编码; 遮挡检测

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2014)08-1838-06

DOI: 10.3724/SP.J.1146.2013.01416

Robust Coding via L₂-norm Regularization for Visual Tracking

Yuan Guang-lin^① Xue Mo-gen^{②③}

^①(Eleventh Department, Army Officer Academy of PLA, Hefei 230031, China)

^②(Department of Scientific Research, Army Officer Academy of PLA, Hefei 230031, China)

^③(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

Abstract: Sparse representation based visual trackers are very computationally inefficient and prone to model drifting. To deal with these issues, a novel visual tracking method is proposed based on L₂-norm regularized robust coding. The proposed method solves the coding coefficient of candidate objects via robust coding based on L₂-norm regularization, and it achieves visual tracking by taking weighted reconstruction errors of the candidate object as observation likelihood in particle filter framework. In addition, to adapt the changes of object appearance and avoid model drifting, an occlusion detection method for template update is proposed by investigating the weight matrix of current object estimated with L₂-norm regularized robust coding. The experimental results on several challenging sequences show that the proposed method has better performance than that of the state-of-the-art tracker.

Key words: Visual tracking; L₂-norm regularization; Robust coding; Occlusions detection

1 引言

视觉跟踪在智能监控、人机交互、机器人导航和运动分析等方面具有重要的应用价值, 其目的是利用序列图像估计目标的状态。近十年来, 尽管视觉跟踪有了较大的发展, 已经提出了一些方法, 如文献[1-3]提出的产生式跟踪, 文献[4]提出的判别式跟踪, 文献[5]提出的融合产生式和判别式跟踪, 但是实时鲁棒的视觉跟踪尚未解决。目前, 视觉跟踪面临目标遮挡、光照变化、视角变化和背景复杂等难题, 仍然是机器视觉领域的研究热点。

受到稀疏表示在人脸识别中应用^[6]的启发, Mei 等人^[7]首次将稀疏表示应用到视觉跟踪中, 提出了L₁跟踪。L₁跟踪将视觉跟踪等价于候选目标在目标模板和小模板(单位像素基)上稀疏表示系数的求解问题。L₁跟踪提高了对遮挡目标跟踪的鲁棒性, 但是它也存在速度慢和易受噪声模板影响等不足。针对L₁跟踪存在的问题, 近几年提出了一些方法。为了利用高维图像特征提高L₁跟踪的鲁棒性, 文献[8]将动态组稀疏引入到视觉跟踪中, 提出一种基于两阶段稀疏优化的视觉跟踪方法。文献[9]提出一种最小化误差界采样策略, 通过减少求解L₁最小化模型的次数来提高L₁跟踪的速度, 并提出利用小模板系数检测目标遮挡, 据此进行模板更新, 减少

2013-09-17 收到, 2014-01-08 改回

国家自然科学基金(61175035, 61379105)资助课题

*通信作者: 袁广林 yuanguanglin1008@sina.com

了模型漂移的发生。在文献[9]的基础上,文献[10]提出一种新的 L_1 范数最小化模型及其快速算法,进一步提升了 L_1 跟踪的鲁棒性与速度。根据粒子稀疏表示系数的相似性,文献[11]提出利用多任务稀疏学习模型联合求解粒子的表示系数,提高了 L_1 跟踪的速度。文献[12]提出一种结构化的局部稀疏表现模型,该模型的编码系数包含了目标的局部与空间信息,进一步提高了 L_1 跟踪的精度。与上述方法不同,文献[13]提出融合基于稀疏表示的判别式模型和产生式模型建立目标表现模型,实现基于稀疏性的协作模型跟踪目标。尽管基于稀疏表示的跟踪方法提升了对遮挡目标跟踪的鲁棒性,但是这些跟踪方法仍然存在一些不足。首先, L_1 跟踪潜在的假设是重建误差具有稀疏性,然而在目标严重遮挡时这一假设并不成立;其次,针对 L_1 跟踪的速度问题,虽然已经提出了一些方法^[9-11],但是快速的稀疏编码仍然没有解决,现有方法只能使用低分辨率的图像在跟踪精度与速度之间实现折中;再者,表示系数的稀疏性使得其对模板的依赖更强,一旦模板中引入背景信息会导致跟踪失败。2011年,文献[14,15]对基于稀疏表示的人脸识别与图像分类进行了研究,结果表明:编码系数的稀疏性并不是提高图像分类与人脸识别性能的真正原因,尤其是文献[15]指出:并不是稀疏表示提高了人脸识别的性能,而是多个人脸的协作表示保证了基于稀疏表示的人脸识别性能的提升。

在上述研究的启发下,本文提出一种基于 L_2 范数正则化鲁棒编码的视觉跟踪方法。不同于现有 L_1 跟踪,首先,本文提出的跟踪方法以重建误差的加权 L_2 范数作为编码模型的损失函数项提高了对遮挡目标跟踪的鲁棒性;其次,利用 L_2 范数正则化最小二乘模型求解候选目标的编码系数,提高了目标跟踪的速度;最后,利用权值矩阵检测目标遮挡和实现模板更新,降低了模型漂移的发生概率。

2 L_2 范数正则化鲁棒编码

由文献[14-16]可知,为了使基于多视图模板与线性模型的视觉跟踪对目标遮挡具有鲁棒性,编码系数的稀疏性不是必须的,在模板字典由目标与背景两类样本组成时,对编码系数的 L_2 范数约束也可获得对目标遮挡的鲁棒性。据此,在视觉跟踪中,文献[7]提出的 L_1 最小化模型应改为如式(1)模型:

$$\min_{\mathbf{c}} \|\mathbf{D}\mathbf{c} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{c}\|_2^2 \quad (1)$$

其中, $\mathbf{y} \in R^d$ 是候选目标, $\mathbf{D} = [\mathbf{T}, \mathbf{B}] \in R^{d \times n}$ 称为模板字典, $\mathbf{T} \in R^{d \times n_t}$ 和 $\mathbf{B} \in R^{d \times n_b}$ 分别是目标模板和背景模板, $\mathbf{c} = [\mathbf{c}_T; \mathbf{c}_B] \in R^n$, $\mathbf{c}_T \in R^{n_t}$ 和 $\mathbf{c}_B \in R^{n_b}$

分别是目标模板系数和背景模板系数。

式(1)的本质是 L_2 范数正则化最小二乘估计问题,其前提假设是重建误差 $\mathbf{e} = \mathbf{D}\mathbf{c} - \mathbf{y}$ 符合高斯分布。在视觉跟踪中,由于目标表现变化的复杂性,使得重建误差 \mathbf{e} 的分布不可预知,因此式(1)并不符合视觉跟踪的实际情况。由文献[15]可知,为了描述视觉跟踪中未知分布的重建误差,需要将损失函数项 $\|\mathbf{D}\mathbf{c} - \mathbf{y}\|_2^2$ 改为 $\|\mathbf{W}^{1/2}(\mathbf{D}\mathbf{c} - \mathbf{y})\|_2^2$, 这样式(1)成为 L_2 范数的正则化鲁棒编码模型:

$$\min_{\mathbf{c}} \|\mathbf{W}^{1/2}(\mathbf{D}\mathbf{c} - \mathbf{y})\|_2^2 + \lambda \|\mathbf{c}\|_2^2 \quad (2)$$

其中, \mathbf{W} 是一个对角阵,称为权重矩阵,其值 $\mathbf{W}_{i,i}$ 是候选目标 \mathbf{y} 中第 i 个像素的权值。

在视觉跟踪中,如果候选目标 \mathbf{y} 未被遮挡,则候选目标可以由目标模板 \mathbf{T} 较好地重建,即 $\mathbf{y} = \mathbf{T}\mathbf{c}_T$, 这时 \mathbf{y} 中各个像素的重建误差 e_i 较小;当日目标遮挡时,如果利用目标模板 \mathbf{T} 重建候选目标 \mathbf{y} , 则遮挡像素的重建误差较大,非遮挡像素的重建误差较小。为了减小遮挡像素对编码系数 \mathbf{c} 的影响,使得目标遮挡时线性关系 $\mathbf{y} = \mathbf{T}\mathbf{c}_T$ 得以较好地近似,在求解编码系数 \mathbf{c} 时,遮挡像素应赋予较小的权值,非遮挡像素应赋予较大的权值。不失一般性,令 $\mathbf{W}_{i,i} \in [0,1]$, 定义 $\mathbf{W}_{i,i}$ 为 logistic 函数^[15], 如式(3)所示。

$$\mathbf{W}_{i,i} = \exp(-\mu e_i^2 + \mu\delta) / (1 + (\exp(-\mu e_i^2 + \mu\delta))) \quad (3)$$

其中, μ 和 δ 均是正的标量。参数 μ 控制 $\mathbf{W}_{i,i}$ 从 1 到 0 的下降率, δ 控制分界点位置。

对于式(2)的求解,本文采用类似于文献[15]提出的算法,具体算法示于表 1。

3 遮挡检测与模板更新

为了适应跟踪过程中目标视觉特征的变化,需要更新目标模板。文献[7]直接利用跟踪结果更新目标模板,在目标遮挡时会使得遮挡信息引入到目标模板中,进而导致跟踪失败。为了解决文献[7]中模

表 1 迭代加权 L_2 范数正则化鲁棒编码算法

算法 1 迭代加权 L_2 范数正则化鲁棒编码算法

输入: 候选目标 \mathbf{y} , 字典 \mathbf{D} , 最大迭代次数 I_{\max} 。

输出: 编码系数 \mathbf{c} 。

令 $\mathbf{c}_1 = [1/m, 1/m, \dots, 1/m]^T$

for $k=1$ to I_{\max}

步骤 1 计算误差 $\mathbf{e}_k = \mathbf{y} - \mathbf{D}\mathbf{c}_k$;

步骤 2 利用式(3)计算权重矩阵 \mathbf{W}_k ;

步骤 3 求解编码系数 $\mathbf{c}_k = (\mathbf{D}^T \mathbf{W}_k \mathbf{D} + \lambda \mathbf{I})^{-1} \mathbf{D}^T \mathbf{W}_k \mathbf{y}$;

end

步骤 4 令 $\mathbf{c} = \mathbf{c}_k$ 。

板更新方法存在的问题,文献[9,10]提出利用小模板系数检测遮挡,但是由于小模板也能表示跟踪目标^[10],这时会导致检测失败。考虑到式(3)权重矩阵 W 中遮挡像素的权值较小,非遮挡像素的权值较大,并且遮挡像素具有空间连续性,据此本文提出根据权值矩阵检测目标遮挡和计算遮挡矩阵,如表2所示。

表2 遮挡矩阵计算与遮挡检测

方法1 遮挡矩阵计算与遮挡检测	
输入:	当前目标 y
输出:	遮挡矩阵 M 和遮挡标示 O_f
步骤1	根据当前目标 y , 利用算法1计算权值矩阵 W ;
步骤2	将权值向量 $[W_{1,1}, W_{2,2}, \dots, W_{m,m}]^T$ 重排为2维图像 $I_w(p)$;
步骤3	对 $I_w(p)$ 进行阈值化分割和闭合操作;
步骤4	计算 $I_w(p)$ 中值为“0”的最大连续区域的面积 A_{\max} 及其位置 p_{\max} ;
步骤5	令 $I_w(p \neq p_{\max}) = 1$ 并将 $I_w(p)$ 重排为1维向量,令 $M = \text{diag}(I_w(p))$;
步骤6	遮挡判决: 如果 $A_{\max} > 0.2 \cdot A_y$ (A_y 为目标的面积), 则 $O_f = 1$, 否则 $O_f = 0$ 。

图1给出了利用方法1检测目标遮挡的一个例子。图1(a), 1(b)和图1(c)分别是目标图像、权值图像和遮挡矩阵图像。在遮挡图中黑色像素值对应的是遮挡像素,白色像素值对应的是非遮挡像素。

在跟踪开始时,根据目标的初始位置 $l_1 = (x_1, y_1)$, 在环形区域 $\{l_T \mid \|l_T - l_1\| < \delta\}$ 内利用高斯扰动采样 n_t 个样本初始化目标模板 T_1 , 在环形区域 $\{l_B \mid \gamma < \|l_B - l_1\| < \eta\}$ 内随机采样 n_b 个样本初始化背景模板 B_1 。在跟踪过程中,利用提出的模板更新方法更新模板字典,该方法包括背景模板更新和目标模板更新。背景模板更新用在背景区域中采样的样本替换背景模板中的所有样本。目标模板更新,首先利用方法1检测遮挡,若检测到目标遮挡,则不更新目标模板,否则利用方法2中的步骤4至步骤6更新目标模板。假设 $D_t = [T_t, B_t]$ 为 t 时刻的模

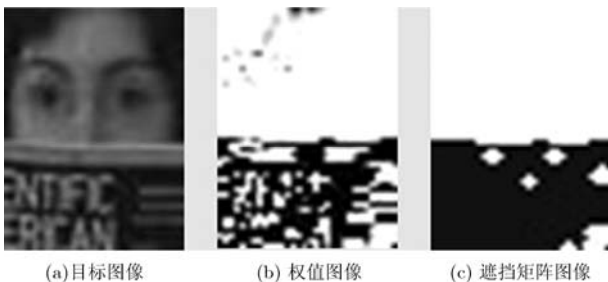


图1 遮挡检测示例

板字典, y_t 和 l_t 分别是 t 时刻的目标图像和目标位置,定义函数 $\text{sim}(\cdot, \cdot) = \cos(\theta)$ 表示两个向量之间的相似性,其中, θ 是两个归一化样本的夹角,则提出的模板更新方法如表3所示。

表3 模板更新

方法2 模板更新	
输入:	$D_t = [T_t, B_t], y_t$
输出:	$D_{t+1} = [T_{t+1}, B_{t+1}]$
步骤1	在背景 $\{l_B \mid \gamma < \ l_B - l_1\ < \eta\}$ 中采样 n_b 个样本替换 B_t 中的所有元素;
步骤2	根据 y_t 利用方法1检测遮挡;
步骤3	如果 $O_f = 1$, 则不更新目标模板, 否则执行步骤4至步骤6;
步骤4	计算 $s_i = \text{sim}(y_t, t_i^t)$, $i = 1, 2, \dots, n_t$;
步骤5	计算 $i_r = \arg \min_i s_i$, $i = 1, 2, \dots, n_t$;
步骤6	目标模板更新: $t_i^t = y_t$ 。

4 目标跟踪

4.1 跟踪方法

本文利用粒子滤波跟踪目标,粒子滤波包括预测和更新两个步骤。假设 $y_{1:t-1} = \{y_1, y_2, \dots, y_{t-1}\}$ 为1到 $t-1$ 时刻的所有可用图像观测,则预测过程为

$$p(x_t \mid y_{1:t-1}) = \int p(x_t \mid x_{t-1})p(x_{t-1} \mid y_{1:t-1})dx_{t-1} \quad (4)$$

其中, x_t 表示 t 时刻的目标状态, $p(x_t \mid x_{t-1})$ 是动态模型。在 t 时刻,当观测 y_t 可用时,则执行如式(5)更新过程:

$$p(x_t \mid y_{1:t}) = \frac{p(y_t \mid x_t)p(x_t \mid y_{1:t-1})}{p(x_t \mid x_{t-1})} \quad (5)$$

其中, $p(y_t \mid x_t)$ 为观测似然模型。在得到所有图像观测 $y_{1:t} = \{y_1, y_2, \dots, y_t\}$ 后,利用最大后验概率准则估计跟踪目标的最优状态 \hat{x}_t , 如式(6)所示。

$$\begin{aligned} \hat{x}_t &= \arg \max_{x_t^i} p(x_t^i \mid y_{1:t}) \\ &= \arg \max_{x_t^i} p(y_t^i \mid x_t^i)p(x_t^i \mid x_{t-1}^i), i = 1, 2, \dots, N \end{aligned} \quad (6)$$

其中, x_{t-1}^i 和 x_t^i 分别是第 i 个粒子在 $t-1$ 和 t 时刻的状态, y_t^i 是 t 时刻第 i 个粒子的图像观测, N 是粒子数。图2给出了本文跟踪方法的主要步骤,其中,目标的初始状态由手工标定得到。

4.2 动态模型

本文利用仿射变换对目标在连续两帧图像之间的运动建模。令 $x_t = (x_t, y_t, w_t, h_t, \theta_t)$, 其中, x_t, y_t, w_t, h_t 和 θ_t 分别是 t 时刻目标区域中心的 x 坐标、 y 坐标,目标的宽度,高度和倾斜角。假设状态向量 x_t 中的每一个分量相互独立,则本文采用的动态模型为

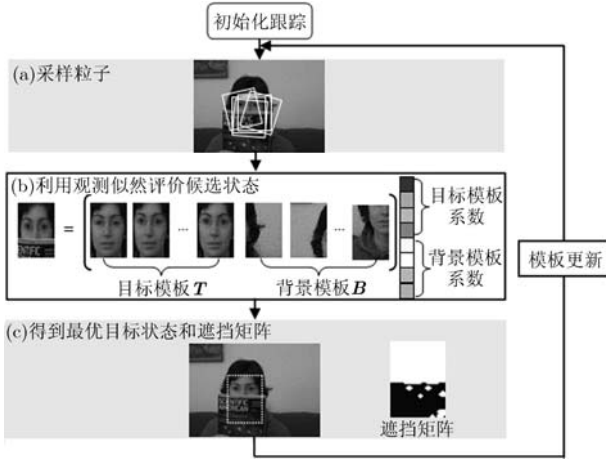


图 2 本文提出的 L_2 正则化鲁棒跟踪方法

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) = N(\mathbf{x}_t; \mathbf{x}_{t-1}, \Psi) \quad (7)$$

其中, Ψ 是对角矩阵, 其对角元素表示相应状态的方差。

4.3 观测模型

对于任一图像观测 \mathbf{y}_t^i , 首先, 求其 L_2 范数正则化编码系数, 模型如式(8)所示。

$$\min_{\mathbf{c}_t^i} \left\| \mathbf{M}_t^{1/2} (\mathbf{D}_t \mathbf{c}_t^i - \mathbf{y}_t^i) \right\|_2^2 + \lambda \left\| \mathbf{c}_t^i \right\|_2^2 \quad (8)$$

其中, \mathbf{M}_t 是由方法 1 得到的 t 时刻的遮挡矩阵, $\mathbf{D}_t = [\mathbf{T}_t, \mathbf{B}_t]$ 是 t 时刻的模板字典, $\mathbf{c}_t^i = [\mathbf{c}_{T,t}^i; \mathbf{c}_{B,t}^i] = \mathbf{P}_t \mathbf{y}_t^i$ 是 \mathbf{y}_t^i 的编码系数, 其中, $\mathbf{P}_t = (\mathbf{D}_t^T \mathbf{M}_t \mathbf{D}_t + \lambda \mathbf{I})^{-1} \mathbf{D}_t^T \mathbf{M}_t$ 称为投影矩阵。在求解图像观测 \mathbf{y}_t^i 的编码系数的基础上, 定义观测似然模型如式(9):

$$p(\mathbf{y}_t^i | \mathbf{x}_t^i) = \frac{1}{\Gamma} \exp \left(-\alpha \left\| \mathbf{M}_t^{1/2} (\mathbf{y}_t^i - \mathbf{T}_t \mathbf{c}_{T,t}^i) \right\|_2^2 \right) \quad (9)$$

其中 α 是高斯核尺度参数, Γ 是归一化常量。

4.4 计算复杂性分析

与本文方法最相近是 IVT 跟踪^[2]和 L_1 跟踪^[7]。假设 $\mathbf{U} \in \mathbb{R}^{d \times k}$ 是 IVT 跟踪使用的特征基, $\mathbf{T} \in \mathbb{R}^{d \times k}$ 是 L_1 跟踪使用的模板, $\mathbf{P} \in \mathbb{R}^{d \times k}$ 是 L_2 范数正则化编码的投影矩阵, 则这 3 个跟踪方法的计算复杂度如表 4 所示。IVT 跟踪最耗时部分是利用特征基计算表示系数, 该计算是矩阵向量乘, 其时间复杂度是 $O(dk)$; L_1 跟踪最耗时部分是利用 PCG (Preconditioned Conjugate Gradients) 算法求解稀疏表示系数, PCG 算法的基本计算是矩阵向量乘, 其时间复杂度是 $O(d^2 + dk)$; 本文方法最耗时部分是用投影矩阵计算表示系数, 该计算是矩阵向量乘, 其时间复杂度是 $O(dk)$ 。由上述分析可知: 这 3 个跟踪方法的计算复杂度级别相同, 都是变量多项式的复杂度。表 4 中第 3 和第 4 列分别给出了在相同软硬件环境下, $d=16 \times 16$ 和 $d=32 \times 32$, $k=16$ 时,

表 4 计算复杂度与计算时间

方法	计算复杂度	计算时间 (16×16) (ms)	计算时间 (32×32) (ms)
IVT 跟踪 ^[2]	$O(dk)$	0.034	0.085
L_1 跟踪 ^[7]	$O(d^2 + dk)$	17.500	362.000
本文方法	$O(dk)$	0.120	0.912

求解一个样本表示系数的计算时间, 可以看出: 本文跟踪方法的计算速度比 IVT 跟踪慢, 但是快于 L_1 跟踪。

5 实验结果与分析

以 Matlab R2011b 为开发工具实现了本文提出的跟踪方法, 并在 Pentium(R) Dual-Core 2.80 GHz CPU, 2 G 内存的台式机上调试通过。采用 “shop” 序列, “faceocc-woman” 序列, “david” 序列和 “sylv” 序列^[17]对提出的跟踪方法进行了实验验证, 并与 IVT 跟踪方法^[2], MIL 跟踪方法^[4], VTD 跟踪方法^[3]和 APG_L1 跟踪方法^[10]等 4 种跟踪方法进行了比较。实验中, 5 种跟踪方法的粒子数均为 600。对于本文跟踪方法的参数设置说明如下: 参数 δ 用于初始化目标模板, 其值不能过大, 过大会导致目标模板中包含背景样本, 从而降低跟踪的精度或者跟踪失败, 本文 $\delta = 2$ 。参数 γ 和 η 用于背景模板初始化和更新, 其值不能过小, 过小会导致背景模板中包含目标样本, 从而导致跟踪精度降低或者跟踪失败。本文采用与文献[4]相同的方法设置参数 γ 和 η , 即 $\gamma = \max(w/2, h/2)$, $\eta = 2\gamma$, 其中, w 和 h 分别是目标的宽和高。目标模板数 n_t 和背景模板数 n_b 对跟踪的精度与实时性具有影响。由文献[6]可知, 增加 n_t 和 n_b 有助于丰富目标和背景子空间, 从而提高目标跟踪的精度, 但是也会降低跟踪速度。为了保证跟踪的精度和效率, 本文设置 $n_t = 10$, $n_b = 100$ 。模板字典更新频率取决于目标表观变化的快慢, 如果目标表观变化较快, 则需要较大的模板字典更新频率, 反之亦然。本文每 10 帧进行 1 次模板字典更新。动态模型参数取决于目标状态变化的速率, 目标状态变化越快, 动态模型参数越大, 反之亦然。本文对于所有测试序列, 动态模型参数均为(5,5, 0.01, 0.01, 0.002)。

图 3 给出了 IVT 跟踪方法, MIL 跟踪方法, VTD 跟踪方法, APG_L1 跟踪方法和本文跟踪方法对这 4 组图像序列的跟踪结果。“shop” 序列和 “faceocc-woman” 序列存在较大的目标遮挡, 从图 3(a)和图 3(b)可以看出: 本文跟踪方法的跟踪精度优于 IVT 跟踪方法, MIL 跟踪方法, VTD 跟踪

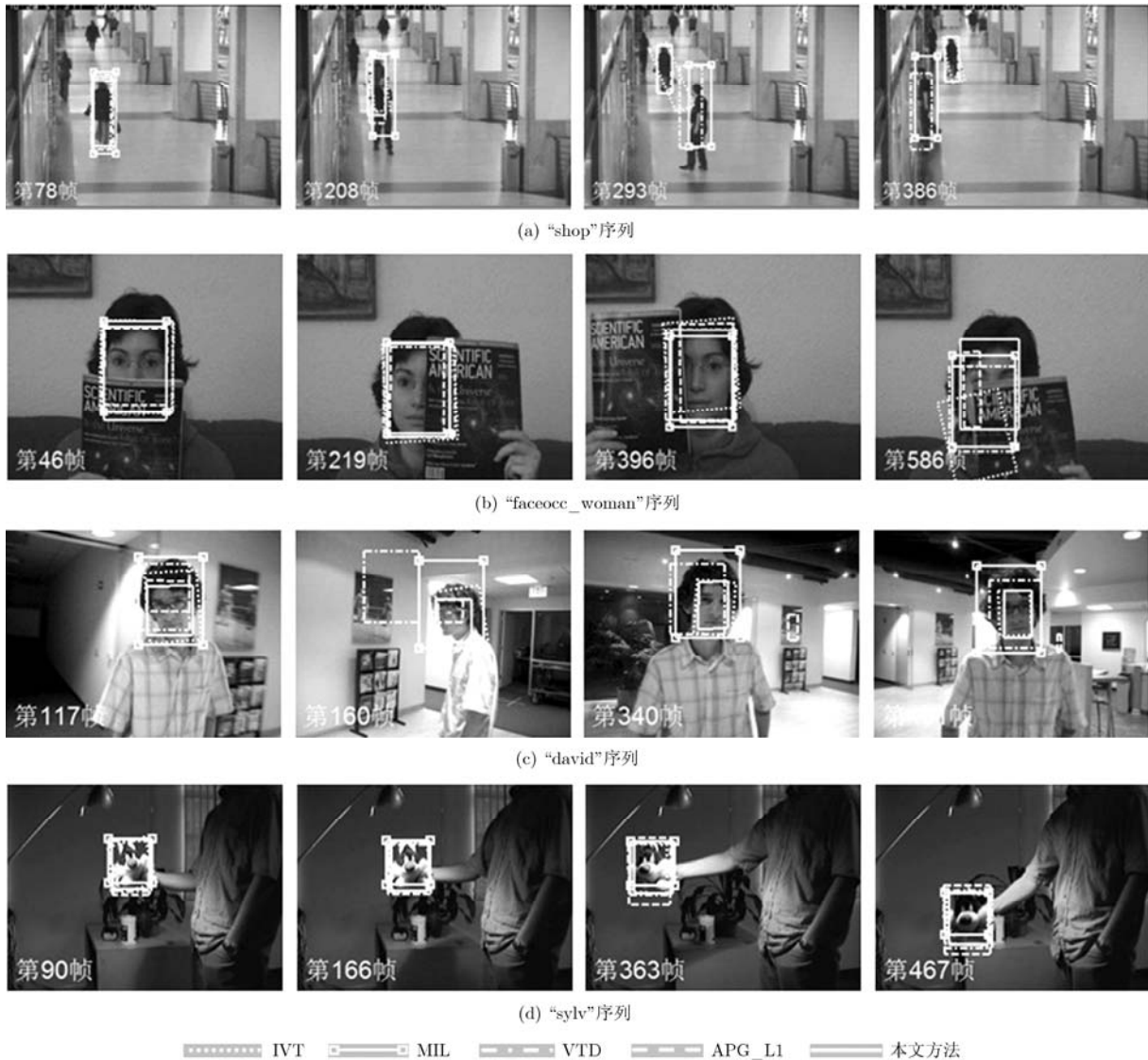


图3 跟踪结果比较

方法, APG_L1 跟踪方法。在“david”序列和“sylv”序列中存在较大的光照变化和姿态变化, 从图 3(c)和图 3(d)可以看出: 本文跟踪方法的跟踪精度与 IVT 跟踪方法相当, 且优于 MIL 跟踪方法, VTD 跟踪方法和 APG_L1 跟踪方法。

通过手工标定得到目标的真实状态参数, 利用相对位置误差 $e = \epsilon/d$ 定量评价跟踪方法的性能, 其中, ϵ 是跟踪结果目标中心位置与目标真实中心位置的距离(单位为像素), d 是目标矩形的对角长度(单位为像素)。表 5 给出了 5 种跟踪算法在这 4 组图像序列上的相对位置误差的均值和方差, 其中黑体为最优指标。由表 5 可以看出: 对于存在目标遮挡的“shop”和“faceocc-woman”图像序列, 本文跟踪方法的跟踪精度优于 IVT 跟踪方法, MIL 跟踪方法, VTD 跟踪方法, APG_L1 跟踪方法。

从表 5 中还可以看出: 对于存在较大光照和姿态变化的“david”和“sylv”图像序列, 本文跟踪方法的跟踪精度最优。

6 结论

本文针对基于稀疏表示的视觉跟踪计算效率低和易于产生“模型漂移”的不足, 提出一种基于 L_2 范数正则化鲁棒编码的视觉跟踪方法。一方面, 该方法利用 L_2 范数正则化鲁棒编码建立目标表观模型, 加快了目标跟踪的速度, 同时提升了对遮挡目标跟踪的鲁棒性; 另一方面, 该方法利用当前目标的权值矩阵检测目标遮挡, 在此基础上进行模板字典的更新, 适应了目标表观的变化, 同时缓解了“模型漂移”。在上述两个方面的基础上, 以粒子滤波为框架建立了基于 L_2 范数正则化鲁棒编码的视觉跟踪方法。利用多个具有挑战性的图像序列对提出的

表 5 跟踪算法性能指标

跟踪方法	相对位置误差均值				相对位置误差方差			
	shop	faceocc_woman	david	sylv	shop	faceocc-woman	david	sylv
IVT	0.092	0.117	0.078	0.057	0.006	0.006	0.004	0.001
MIL	0.436	0.077	0.216	0.087	0.153	0.005	0.006	0.001
VTD	0.667	0.075	0.425	0.060	0.373	0.003	0.261	0.001
APG_L1	0.040	0.084	0.410	0.073	0.001	0.003	0.107	0.001
本文方法	0.030	0.032	0.039	0.034	0	0.001	0.002	0

跟踪方法进行了实验验证, 并与现有跟踪方法进行了比较, 实验结果表明: 提出的跟踪方法对目标遮挡、光照变化和视角变化具有较强的鲁棒性。

参 考 文 献

- [1] Comaniciu D, Ramesh V, and Meer P. Kernel-based object tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5): 564-577.
- [2] Ross D, Lim J, Lin R S, *et al.*. Incremental learning for robust visual tracking[J]. *International Journal of Computer Vision*, 2008, 77(1-3): 125-141.
- [3] Kwon J and Lee K M. Visual tracking decomposition[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, 2010: 1269-1276.
- [4] Babenko B, Yang M H, and Belongie S. Robust object tracking with online multiple instance learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(8): 1619-1632.
- [5] Ang Li, Tang Feng, Guo Yan-wen, *et al.*. Discriminative nonorthogonal binary subspace tracking[C]. Proceedings of Europe Conference on Computer Vision, Crete, 2010: 258-271.
- [6] Wright J, Yang A Y, Ganesh A, *et al.*. Robust face recognition via sparse representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(2): 210-227.
- [7] Mei X and Ling H B. Robust visual tracking using l_1 minimization[C]. Proceedings of IEEE International Conference on Computer Vision, Kyoto, 2009: 1436-1443.
- [8] Liu Bai-yang, Yang Lin, Huang Jun-zhou, *et al.*. Robust and fast collaborative tracking with two stage sparse optimization[C]. Proceedings of Europe Conference on Computer Vision, Crete, 2010, Part IV: 624-637.
- [9] Mei X, Ling H B, Wu Y, *et al.*. Minimum error bounded efficient l_1 tracker with occlusion detection[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Colorado, 2011: 1257-1264.
- [10] Bao Cheng-long, Wu Yi, Ling Hai-bin, *et al.*. Real time robust L_1 tracker using accelerated proximal gradient approach[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Rhode Island, 2012: 1830-1837.
- [11] Zhang Tian-zhu, Bernard Ghanem, Liu Si, *et al.*. Robust visual tracking via multi-task sparse learning[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Rhode Island, 2012: 2042-2049.
- [12] Jia Xu, Lu Hu-chuan, and Yang Ming-hsuan. Visual tracking via adaptive structural local sparse appearance model[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Rhode Island, 2012: 1822-1829.
- [13] Zhong Wei, Lu Hu-chuan, and Yang Ming-hsuan. Robust object tracking via sparsity-based collaborative model[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Rhode Island, 2012: 1838-1845.
- [14] Shi Qin-feng, Anders Eriksson, Anton van den Hengel, *et al.*. Is face recognition really a compressive sensing problem? [C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Colorado, 2011: 553-560.
- [15] Zhang Lei, Yang Meng, and Feng Xiang-chu. Sparse representation or collaborative representation: which helps face recognition?[C]. Proceedings of IEEE International Conference on Computer Vision, Colorado, 2011: 471-487.
- [16] Yang Meng, Zhang Lei, Yang Jian, *et al.*. Robust sparse coding for face recognition[J]. *IEEE Transactions on Image Processing*, 2013, 22(5): 1753-1766.
- [17] http://cvlab.hanyang.ac.kr/tracker_benchmark_v10.html.

袁广林: 男, 1973年生, 博士, 讲师, 主要研究方向为图像处理、计算机视觉、机器学习及其应用等。

薛模根: 男, 1964年生, 博士, 教授, 博士生导师, 主要研究方向为图像处理、计算机视觉、光电防御等。