

基于 Haar-like 特征的实时 L1-跟踪算法

阎刚^{1,2,3} 屈高超^{2,3} 于明^{1,2,3}

(河北工业大学电子信息工程学院 天津 300401)¹ (河北工业大学计算机科学与软件学院 天津 300401)²
(河北省大数据计算重点实验室 天津 300401)³

摘要 稀疏表示技术已成功应用于视觉跟踪,但是仍然存在跟踪算法效率低的问题。提出一种基于 Haar-like 特征的视频跟踪算法,该算法是基于粒子滤波框架的 L1-跟踪算法,其特点是运用 Haar-like 特征及特征块的思想对完备基进行重新构造。将正负小模板由单个像素改为像素块,降低稀疏表示中过完备基的维数,大幅减少稀疏矩阵的计算量;同时,在保证跟踪质量的前提下适当减少目标模板数量,减少稀疏计算的次数,并控制模板更新频率。实验结果表明,所提算法能大幅提高跟踪的实时性,同时很好地解决了跟踪问题中的短时间遮挡、目标物体的形变以及光照变化等问题。

关键词 L1-跟踪算法,粒子滤波,稀疏表示,目标跟踪,Haar-like 特征

中图分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.03.061

Real-time L1-tracker Based on Haar-like Features

YAN Gang^{1,2,3} QU Gao-chao^{2,3} YU Ming^{1,2,3}

(School of Electronic and Information Engineering, Hebei University of Technology, Tianjin 300401, China)¹

(School of Computer Science and Engineering, Hebei University of Technology, Tianjin 300401, China)²

(Hebei Province Key Laboratory of Big Data Calculation, Tianjin 300401, China)³

Abstract In order to solve the problem of real-time in L1-tracker algorithm under the framework of particle filter, this paper proposed an improved L1-tracker algorithm based on Haar-like features. Firstly, the complete dictionary is reconstructed by the Haar-like features and feature blocks. The single pixels are replaced by the pixel blocks to make up positive and negative trivial templates. Then, the dimensions of over-complete dictionary are reduced by sparse representation and the calculation amount of the sparse matrices is significantly reduced. Secondly, the number of the target template is reduced in order to decrease the calculation of sparse representation. Finally, the updating frequency of the templates is controlled by experiential value. Experimental results demonstrate that the proposed algorithm can significantly improve real-time of L1-tracker algorithm and is effective under short time occlusion, the deformation of target and illumination changes.

Keywords L1-tracker, Particle filter, Sparse representation, Object tracking, Haar-like features

1 引言

稀疏表示的研究在最近几年成为许多研究领域的焦点。Mei 等^[1]将 John Wright^[2]运用于人脸识别的稀疏表示思想引入视觉目标跟踪领域中,将目标跟踪问题看作粒子滤波框架下的稀疏逼近问题,提出了 L1-跟踪算法。L1-跟踪算法通过引入两组小模板(trivial template)来对目标内的所有像素进行处理,使得该算法能够处理遮挡、目标物体的形变以及背景和光照变化等一些目标跟踪领域具有挑战性的难题,该算法表现出很好的鲁棒性以及跟踪准确性。相对传统的粒子滤

波跟踪算法^[3], L1-跟踪算法具有以下特点:1)可靠的目标候选模板。将过完备基下的稀疏表示作为候选模板(观测模型)对目标进行描述,目标在过完备基中的稀疏表示可以很好地克服遮挡、光照等问题。2)目标模板实时更新。采用一种在线增量学习方法进行更新,克服了场景、光照等变化,从而增强了目标描述的鲁棒性。但是, L1-跟踪算法中的核心是 l_1 正则化最小二乘问题, l_1 -正则化最小二乘求解效率低^[4], 小模板的引入导致最小化问题维数增加;同时为了达到更好的跟踪效果,利用粒子滤波跟踪框架时会选用较多的候选粒子,进一步降低了算法的跟踪效率。针对运算量大、实性差这一

到稿日期:2016-06-30 返修日期:2016-12-08 本文受国家自然科学基金资助项目(60302018),天津市科技计划项目(14RCGFGX00846),河北省自然科学基金面上项目(F2015202239),天津市科技计划项目(15ZCZDNC00130),河北省高等学校科学技术研究项目(Z2015044)资助。
阎刚(1977-),男,博士生,副教授,主要研究方向为图像处理、模式识别, E-mail: yangang@scse.hebut.edu.cn; 屈高超(1987-),男,硕士生,主要研究方向为图像处理; 于明(1964-),博士,教授,主要研究方向为图像处理、视觉计算及其应用、图像与视频编码的高效算法。

缺点, Mei 等^[5,20]运用最小误差边界策略使运行速度加快了 4~5 倍,但是正负小模板的维数很大,仍然无法满足实时性的要求; Bao 等^[6]在 Mei 的研究基础上首先运用加速近端梯度计算(Accelerated Proximal Gradient, APG)方法加速最小化运算,同时取消负小模板,进一步减少了运算次数,使改进后的算法能达到 12f/s 左右(在本文实验环境下),但正小模板维数依然很大,影响了其实时性;文献[7-8]在 Mei 提出的框架下增加了局部稀疏模型,将目标区域中的背景与目标分离,增强了跟踪的抗背景干扰能力,但跟踪效率没有得到提高; Wang 等^[9]将主成分分析算法与 L1-跟踪算法结合,减少了样本噪声干扰,但依然使用了小模板,运算量很大,跟踪效率很低。由上面分析可知,影响 L1-跟踪算法效率的主要原因是小模板维数太大。在稀疏模型下,文献[10]运用线性分类器直接对原始图像块进行分类, Babenko 等^[11]运用分类器对 Haar-like 特征^[12]进行分类,提出了新的 Multiple Instance Learning(MIL)算法,提高了算法的实时性。MIL 算法可以有效地减少目标跟踪漂移问题,但是由于受更新的多示例学习外观模型与新的被跟踪的目标噪声结果影响,示例包中的示例不能够有效选择,在算法应用于长时间的视频序列目标跟踪问题中时仍然会发生目标漂移现象。此外,文献[13]将退火粒子滤波应用于 L1-跟踪算法中,用于稀疏表示,取得了很好的跟踪效果,但是文中没有提及算法的效率问题; Wu 等^[14]将拉普拉斯错误分布以及结构相似性引入到粒子滤波框架中,提高了现有 L1-跟踪算法的准确性以及鲁棒性,但处理速度仅能达到 1.9f/s。由此可见,在保证跟踪准确性的基础上, L1-跟踪算法需要着重提高处理效率。

Haar-like 特征最早由 Papageorgiou 等^[12]应用于人脸描述, Viola 等^[15]在此基础上,使用 3 种类型 4 种形式的特征组合成特征模板。特征模板内有白色和黑色两种矩形,依据两种矩形区域像素和计算出来的差值就是 Haar-like 特征的特征值。

为了加快 L1-跟踪算法的速度同时保证目标跟踪的精度,本文以 L1-跟踪算法为框架,通过分析稀疏表示中过完备基的构造,运用 Haar-like 特征及特征块的思想,将完备基中正负小模板由单个像素替换为像素块,适当减少目标模板个数,控制模板更新次数,提出一种高速的 L1-跟踪算法。通过在不同视频序列上的实验,验证了本文方法在保证跟踪精度的前提下能大幅提高 L1-跟踪算法的跟踪效率,同时能克服短时间遮挡等目标跟踪领域中的难题。

同时,考虑到算法的跟踪精度,本文还提出了一种鲁棒的纹理模板更新策略。在目标跟踪初始阶段取初始帧纹理作为纹理模板,由用户手动地在第一帧图像上进行标注。在实际跟踪过程中,人体运动和光照的改变会使纹理不断地变化,因此需要动态更新纹理模板。由于人体的自遮挡,会产生一个新的问题,即人体的某些部位通常会被遮挡一段时间,如果更新纹理模板太频繁,很容易把遮挡后的纹理带入到纹理模板中。另外,跟踪不好的纹理如果被带入到纹理模板中,会使后续帧错误累加,产生漂移问题。为了解决这两个问题,算法必须能够检测出这两种情况,阻止其加入到纹理模板中。本文

利用人体的三维模型来检测遮挡后的纹理,用模板系数来检测跟踪不好的纹理。基于这两种检测器的纹理模板更新策略,可以提高跟踪的准确度。

2 粒子滤波下的 L1-跟踪算法

粒子滤波能对复杂环境中的运动目标进行鲁棒的跟踪,主要用于非线性和非高斯的运动目标跟踪问题。L1-跟踪算法将目标跟踪看作粒子滤波框架下的稀疏最小化问题, APG-L1 利用加速近端梯度计算来加速最小化运算,提高 L1-跟踪算法的速度。

2.1 基于稀疏表示的 L1-跟踪算法

2.1.1 粒子滤波算法框架

粒子滤波^[16](Particle Filter, PF)的思想是基于蒙特卡洛方法(Monte Carlo methods, MC),通过递归抽样逼近复杂的后验概率分布,对非高斯和非线性系统都适用;在视觉跟踪中,粒子滤波算法提供了一种在不知道具体的观测概率密度的情况下估计下一帧目标位置的方法。该方法可以分为两步:预测和更新。令 x_k 表示在第 k 帧时目标的位置和观测值, $z_{1:k-1} = \{z_1, z_2, \dots, z_{k-1}\}$ 表示目标从第一帧到 $k-1$ 帧的观测值。粒子滤波进行预测和更新时运用以下两个概率公式:

$$p(x_k | z_{1:k-1}) = \int p(x_k | x_{k-1}) p(x_{k-1} | z_{1:k-1}) dx_{k-1} \quad (1)$$

$$p(x_k | z_{1:k}) = \frac{p(z_k | x_k) p(x_k | z_{1:k-1})}{p(z_k | z_{1:k-1})} \quad (2)$$

其中, $p(z_k | x_k)$ 表示观测概率。以上两个公式构成了最优贝叶斯,但是在很多情况下无法求解,只能通过近似的方法求解,粒子滤波就是其中一类有效的逼近方法。粒子滤波中后验概率 $p(x_k | z_{1:k})$ 可以由一组 N 个带有权值 w_k^i 的粒子 $\{x_k^i\}_{i=1, \dots, N}$ 近似表示为:

$$p(x_k | z_{1:k}) \approx \sum_{i=1}^N w_k^i \delta(x_k - x_k^i) \quad (3)$$

其中, $\delta(\cdot)$ 为狄拉克函数。粒子的权重 w_k^i 的更新方法为:

$$w_k^i \propto w_{k-1}^i \frac{p(z_k | x_k^i) p(x_k^i | x_{k-1}^i)}{q(x_k^i | x_{k-1}^i, z_{1:k})} \quad (4)$$

其中, $q(\cdot)$ 为粒子重要密度函数,作用是防止粒子退化。算法为进一步控制粒子退化现象,利用权值 w_k^i 对粒子进行重采样。

2.1.2 基于稀疏表示的 L1-跟踪算法

稀疏表示是稀疏编码(或压缩感知)中的一个子集。在目标跟踪中,目标所在的线性子空间是由前面若干帧得到的目标构成的模板集张成的。稀疏表示的模型就是计算粒子滤波中样本 x_k 的近似观测概率 $p(z_k | x_k)$ 。在第 k 帧时,模板集 $T_k = \{t_k^1, t_k^2, \dots, t_k^n\} \in \mathbb{R}^{d \times n}$ ($d > n$), 其中 $t_i \in \mathbb{R}^d$ 是由 n 个目标模板归一化后按列生成的一维向量, d 为目标模板像素数; 令 $S_k = \{x_k^1, x_k^2, \dots, x_k^N\}$ 表示样本集, 其中 N 为候选目标个数; $Q_k = \{y_k^1, y_k^2, \dots, y_k^N\}$ 表示候选目标模板集。稀疏模型即可以表示为:

$$y_k^i = T_k a_k^i + I a_k^i, \forall y_k^i \in Q \quad (5)$$

其中, I 是小模板集(正、负小模板), $a_k^i = [a_k^i, a_k^i]$ 为稀疏系数

向量,其中 a_T^i 为目标模板系数向量, a_I^i 为小模板系数向量,均为稀疏向量;另外, a_T^i 加上非零约束后增加了 L1-跟踪算法的鲁棒性。因此,在非零约束下,每个候选模板张量 y_k^i 的稀疏表示可以看成 L1 范数最小化问题:

$$\min_a \frac{1}{2} \|Aa - y_k^i\|_2^2 + \lambda \|a\|_1, a \geq 0 \quad (6)$$

其中, $A = [T_k, I, -I]$ 。因此, x_k^i 的近似观测概率为:

$$p(z_k | x_k^i) = \frac{1}{\Gamma} \exp\{-\alpha \|y_k^i - T_k a_T^i\|_2^2\} \quad (7)$$

其中, α 是 Gaussian 核函数不变控制量, Γ 是一个常系数。 k 帧时最优状态 x_k^* 可以通过式(8)求出:

$$x_k^* = \arg \max_{x_k^i \in S_k} p(z_k | x_k^i) \quad (8)$$

在文献[3]中通过加入最小误差边界减少了 L1 最小化次数,那么样本 x_k^i 的近似观测概率就加了上边界:

$$p(z_k | x_k^i) \leq \frac{1}{\Gamma} \exp\{-\alpha \|T_k \hat{a} - y_k^i\|_2^2\} \cong q(z_k | x_k^i) \quad (9)$$

其中, $\hat{a} = \arg \min_a \|T_k a - y\|_2^2$, $q(z_k | x_k^i)$ 表示样本的概率上边界。增加了上边界就可以使样本 x_k 的概率在 $p(z_k | x_k) < \frac{1}{2N}$ 的情况下不予考虑,从而减少了一部分运算。

2.2 APG 方法改进的 L1-跟踪算法

对于上述算法中的 L1 范数最小化运算量大的问题,文献[6]运用了 APG 方法和取消负小模板来简化运算,式(6)就变为:

$$\min_a \frac{1}{2} \|A'a - y\|_2^2 + \lambda \|a\|_1 + \frac{\mu_k}{2} \|a_I\|_2^2, \text{ s. t. } a_T \geq 0 \quad (10)$$

其中, $A' = [T_k, I]$, $a = [a_T; a_I]$, μ_k 是用来控制正小模板的系数。若要运用 APG 方法,则需将式(10)由限定最小化转换为无约束最小化,可以等价于:

$$\arg \min_a \frac{1}{2} \|A'a - y\|_2^2 + \lambda 1_{\mathbb{R}_+^n} a_T + \frac{\mu_k}{2} \|a_I\|_2^2 + 1_{\mathbb{R}_+^n}(a_T) \quad (11)$$

其中, $1_{\mathbb{R}_+^n} \in \mathbb{R}^n$ 表示单位向量 $1_{\mathbb{R}_+^n}$ 的指数函数,由式(12)给出。

$$1_{\mathbb{R}_+^n}(a) = \begin{cases} 0, & a \geq 0 \\ +\infty, & \text{其他} \end{cases} \quad (12)$$

则式(11)就可以通过式(13)运用 APG 方法计算。

$$F(a) = \frac{1}{2} \|A'a - y\|_2^2 + \lambda 1_{\mathbb{R}_+^n} a_T + \frac{\mu_k}{2} \|a_I\|_2^2 \quad (13)$$

$$G(a) = \|a_I\|_1 + 1_{\mathbb{R}_+^n}(a_T)$$

3 基于 Haar-like 的追踪算法

文献[1,5]引入正负小模板来解决目标形变和遮挡等问题,该方法简单直观,但在粒子滤波框架下每帧图像都会产生数百个粒子。文献[5]中的实验用到了 600 个粒子,每个粒子都需要求解式(6)中 L1 范数最小化问题,并且小模板的维数达到了 $2d+n$ 维,对于 620 帧的视频,需要计算 80000 次 L1-范数最小化,计算量非常巨大。文献[1,5]中过完备基的构造如图 1 所示,文献[6]中的过完备基取消了负小模板并运用 APG 方法简化 L1-范数最小化,该文献指出小模板在跟踪中

的重要性为 30%,但小模板的运算量却占到总运算量的 75%,因此可以通过取消负小模板来减少小模板的总体个数。

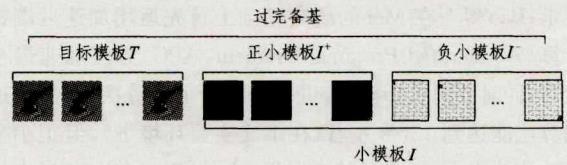


图 1 Mei 等人在 L1-跟踪算法中的过完备基

本文主要工作是对文献[1,5]中 L1-跟踪算法的过完备基运用矩形 Haar-like 特征划分区域的思想 and 分块思想进行重新构造,同时将分块思想融入到过完备基的构造过程中,把正负小模板由单个像素级转换为像素块级,该过程既保留了正负小模板抗遮挡的能力,也保证了正负小模板的稀疏性,同时也大幅减少了模板矩阵的维数,使每帧图像稀疏运算次数大幅减少;然后将改造后的过完备基运用到文献[6]中的 APG L1-跟踪算法中,改变模板更新频率,增强该算法处理遮挡问题的能力。

3.1 过完备基的构造

3.1.1 Haar-like 特征分块构造方法

Haar-like 特征也可简称为 Haar 特征,是计算机视觉领域中常用特征之一,它最早由 Viola and Jone 提出。目前常用的 Haar-like 特征有以下 4 类:线性特征、边缘特征、中心特征、对角线特征。本文中的完备基构造利用了 Haar-like 特征中的后 3 种特征中的矩形划分区域思想,构造了 Haar-like 过完备基的像素块的划分。图 2 是 Haar-like 特征中的边缘特征(见图 2(a)、图 2(b))、中心特征(见图 2(c))和对角线特征(见图 2(d))。

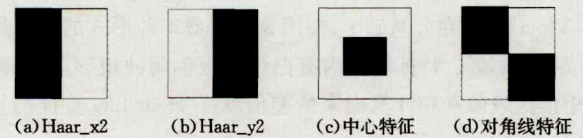


图 2 Haar-like 的 4 种特征

将图 2(a)、图 2(b)、图 2(d) 3 种矩形合并形成图 3 中的边角(见图 3(a)~图 3(d)) and 中心(见图 3(e)) 两类,构成 Haar-like 过完备基中正负小模板的分块思想,如图 3 所示。

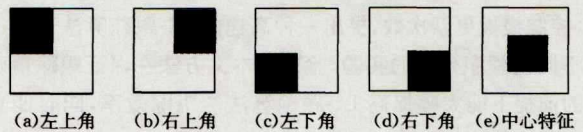


图 3 矩形划分思想

3.1.2 Haar-like 过完备基的构造

本文的 Haar-like 过完备基如图 4 所示。

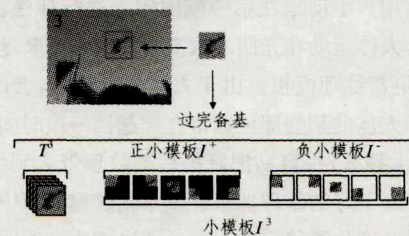


图 4 Haar-like 过完备基

Haar-like 过完备基 D 满足:模板集 $T_k = [t_k^+, t_k^-, \dots, t_k^-] \in \mathbb{R}^{d \times n}$, $d > n$. 使用零向量填充,使式(6)中 $A \in \mathbb{R}^{d \times s}$, 其中, $s \gg d$, 满足稀疏模型过完备基条件,小模板的维数由线性级 $2d+n$ 维缩小到常量级 n . 因此,新的稀疏过程即可由图 5 表示.

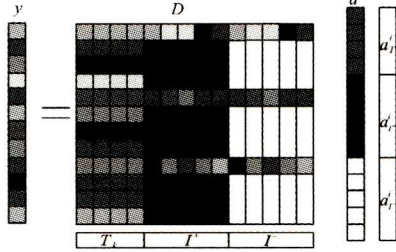


图 5 新的稀疏表示

图 5 中稀疏系数矩阵 a 满足非负,因此式(5)即可写成:

$$y = Aa = [T_k^+, I^+, I^-] \begin{bmatrix} a_i^+ \\ a_i^- \\ a_i^- \end{bmatrix}, \text{ s. t. } a \geq 0 \quad (14)$$

3.2 融入 Haar-like 过完备基的 APG-L1 跟踪

将 Haar-like 过完备基应用到式(10)中可以等价为:

$$\min_a \frac{1}{2} \|Aa - y\|_2^2 + \lambda \|a\|_1 + \frac{\mu_k}{2} \|a_{T^+}\|_2^2 + \frac{\nu_k}{2} \|a_{I^-}\|_2^2$$

s. t. $a_T \geq 0$ (15)

其中, μ_k, ν_k 分别控制正、负小模板的系数. 运用 APG 方法, 式(15)由限定最小化转换成无约束最小化, 可以等价于:

$$\arg \min_a \frac{1}{2} \|Aa - y\|_2^2 + \lambda 1_T^T a_T + \frac{\mu_k}{2} \|a_{T^+}\|_2^2 + \frac{\nu_k}{2} \|a_{I^-}\|_2^2 + 1_{\mathbb{R}_+^s}(a_T) \quad (16)$$

运用 APG 方法, 式(16)可以转化为:

$$F(a) = \frac{1}{2} \|Aa - y\|_2^2 + \lambda 1_T^T a_T + \frac{\mu_k}{2} \|a_{T^+}\|_2^2 + \frac{\nu_k}{2} \|a_{I^-}\|_2^2$$

$$G(a) = \|a_{T^+}\|_1 + \|a_{I^-}\|_1 + 1_{\mathbb{R}_+^s}(a_T) \quad (17)$$

其中, $F(a)$ 是可微且具有李普希茨连续的梯度的凸函数, $G(a)$ 是非光滑的凸函数, 用 APG 方法解决式(18)的无约束最小化问题:

$$\min F(a) + G(a) \quad (18)$$

运用 APG 方法求解 $F(a)$ 的最小值^[17-18]: 令 $\alpha_0 = \alpha_{-1} = 0 \in \mathbb{R}^N$, $t_0 = t_{-1} = 1$, i 为收敛时的迭代次数, L 为李普希茨常量, $\{\alpha_i\}$ 是迭代过程中产生的序列, 则有

$$\begin{cases} \beta_{i+1} = \alpha_i + \frac{t_{i-1}-1}{t_i}(\alpha_i - \alpha_{i-1}) \\ \alpha_{i+1} = \arg \min_a \frac{L}{2} \|a - \beta_{i+1} + \frac{\nabla F(\beta_{i+1})}{L}\|_2^2 + G(a) \\ t_{i+1} = \frac{1 + \sqrt{1 + 4t_i^2}}{2} \end{cases} \quad (19)$$

当满足迭代次数约束条件 $i = O(\sqrt{\frac{L}{\epsilon}})$ 及 $\|\alpha_i - \alpha^*\| < \epsilon$ 时, α^* 即为满足式(17)的最小值.

上述过程中 α_{i+1} 需要解决一个最优化问题:

$$\alpha_{i+1} = \arg \min_a \frac{L}{2} \|a - \beta_{i+1} + \frac{\nabla F(\beta_{i+1})}{L}\|_2^2 + G(a) \quad (20)$$

式(17)中的 G 函数不能直接求解, 需要通过式(19)进行分析, 然后求解.

将式(17)中定义的 $F(a)$ 和 $G(a)$ 代入式(20)中有如下解:

$$\begin{aligned} \alpha_{i+1}|_T &= \max(0, g_{i+1}|_T) \\ \alpha_{i+1}|_{I^+} &= \zeta_T^+(g_{i+1}|_{I^+}) \\ \alpha_{i+1}|_{I^-} &= \zeta_T^-(g_{i+1}|_{I^-}) \end{aligned} \quad (21)$$

其中, $g_{i+1} = \beta_{i+1} - \frac{\nabla F(\beta_{i+1})}{L}$, ζ 是控制正负模板的软阈值算子: $\zeta(x) = \text{sign}(x) \max(|x| - \lambda, 0)$.

根据文献[10]中的 APG 方法, 式(15)可以迭代求解 a , 直至收敛:

$$\begin{cases} \alpha_{i+1}|_T = \max(0, g_{i+1}|_T) \\ \alpha_{i+1}|_{I^+} = \zeta_T^+(g_{i+1}|_{I^+}) \\ \alpha_{i+1}|_{I^-} = \zeta_T^-(g_{i+1}|_{I^-}) \\ \beta_{i+1} = \alpha_i + \frac{t_{i-1}-1}{t_i}(\alpha_i - \alpha_{i-1}) \\ g_{i+1}|_T = \beta_{i+1}|_T - \frac{(A^T(A\beta_{i+1} - y))|_T}{L} - \lambda 1_T \\ g_{i+1}|_{I^+} = \beta_{i+1}|_{I^+} - \frac{(A^T(A\beta_{i+1} - y))|_{I^+}}{L} - \frac{\mu \beta_{i+1}|_{I^+}}{L} \\ g_{i+1}|_{I^-} = \beta_{i+1}|_{I^-} - \frac{(A^T(A\beta_{i+1} - y))|_{I^-}}{L} - \frac{\nu \beta_{i+1}|_{I^-}}{L} \\ t_{i+1} = \frac{1 + \sqrt{1 + 4t_i^2}}{2} \end{cases} \quad (22)$$

3.3 模板更新策略

如果目标模型没有更新, 那么在动态场景中当目标的外观不断发生变化时(比如光照变化、姿态变化等), 跟踪很可能因模型的适应能力比较弱而导致失败. 然而, 如果对模型进行频繁而简单的更新, 每次更新都将引入小的误差, 这些小误差的积累将导致跟踪器与目标发生偏移, 即产生漂移问题.

本文中目标模板更新采用两个策略: 减少模板数量和控制模板更新频率. 由于采用 Haar-like 的特征块代替单个像素来进行过完备基的构造, 正负小模板由 Haar-like 特征的像素块构成, 而不是 L1 算法中的单个像素, 模板的数量减少; 此外, 根据 L1-范数最小化过程中得到的稀疏系数大小来决定模板的更新, 通过稀疏系数粒子的权值来控制粒子滤波框架中的重采样, 分别对正负小模板设置不同的控制系数 μ_k 和 ν_k , 这样既能有效地防止粒子退化, 又能保证样本模板的正确性. 当视频场景中目标被部分遮挡或者发生形变时, 本文方法采用人工选取视频第一帧中的目标作为目标模板, 当目标被完全遮挡时也采用这个模板来跟踪, 此方法能达到较好跟踪效果, 且在跟踪过程中不会因模板的频繁更新而导致目标漂移, 也节省了模板更新时间, 本文的模板更新频率约为文献[6]的 2/3. 图 6 所示的跟踪效果说明了本文方法的有效性.

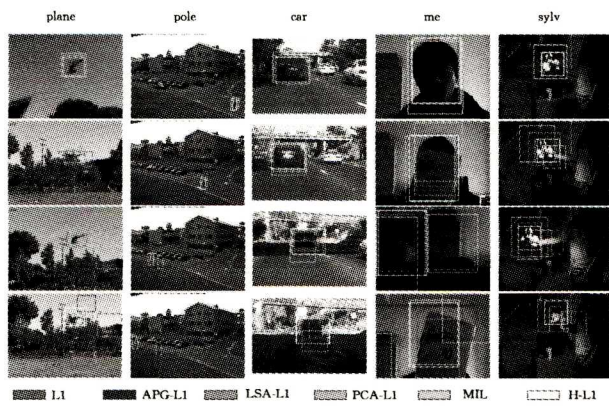


图6 不同算法的跟踪结果

3.4 本文的跟踪方法

基于上面的讨论,给出所提的基于 Haar-like 特征的 L1-跟踪算法的伪代码。

输入:当前第 k 帧视频 F_k ,上一帧目标位置 x_{k-1} ,目标模板集 T^N ,粒

子集 $S_{k-1}=[x_k^1, x_k^2, \dots, x_k^N]$

1. 初始化:设置各粒子的权值为 $w = \frac{1}{N}$;
2. for $i=1;N$
 - 2.1. 获取每个候选目标位置 x_k^i 对应的候选区域,归一化并拉伸为向量 y_k^i ;
 - 2.2. while $j < N \ \&\& \ w$ 大于阈值
 - 2.2.1. 通过式(22)求解式(15)中的最小化问题;
 - 2.2.2. 通过式(7)计算观测概率 p_i ;
 - 2.2.3. 调节粒子权值阈值;
 - 2.2.4. $i++$;
 - end
 - 2.3. 求解式(9)得到近似观测概率 a_k^i (稀疏系数),即当前粒子权值 w_k^i ;
 - end
3. 通过 w_k^i 对粒子进行降序排列;
4. 改变式(15)中的 μ_k 和 ν_k ,更新正负小模板;
5. 更新目标模板 T ;

6. 更新粒子集;

输出:每帧跟踪到的目标位置 x_k

4 实验结果与分析

为了测试本文所提出的视频跟踪算法的跟踪效率、精确度、处理遮挡和目标形变的能力,选择了4个富有代表性的图像测试序列(car, plane, pole, sylv)和1个制作的主要用于测试全遮挡和目标形变问题的图像序列(me),并与L1-跟踪算法(L1)^[5]、APG-L1跟踪算法(APG-L1)^[6]、LSA-L1跟踪算法(LSA-L1)^[7]、PCA-L1跟踪算法(PCA-L1)^[9]和MIL算法(MIL)^[11]5种目前最新的跟踪算法进行比较。

实验中各测试图像序列粒子数 $N=600$,本文算法中的样本个数选用 $n=5$,正负小模板各5个,式(15)中正负小模板控制系数初始化为 $\mu_k = 5, \nu_k = 5$;实验环境为 Intel i5-2430M(2.4GHz),平台为 Matlab 2012b。

4.1 跟踪实时性

本文算法(HL-L1)与L1, APG-L1, LSA-L1, PCA-L1及MIL 5种跟踪算法进行比较,其中,L1, APG-L1, LSA-L1, PCA-L1和本文算法的粒子数均取为600。MIL算法是一种运用 Haar-like 特征且比较有创新性的跟踪算法,所以本文将其作为一个对比算法。表1列出了本文算法与以上5种算法的跟踪效率对比。

表1 本文算法与其他算法的实时性对比/f/s

	L1	APG-L1	LSA-L1	PCA-L1	MIL	HL-L1
plane	0.4	9.7	2.7	2.2	4.3	27.2
pole	0.3	12.0	2.4	1.9	3.6	24.6
car	0.3	12.1	2.5	2.1	4.1	24.5
sylv	0.3	8.9	2.6	2.2	4.2	23.1
me	0.3	9.2	2.9	1.7	3.8	21.9

由表1可以看出,本文的算法在普通电脑上平均能达到21f/s,对各个图像序列都能进行实时处理,能满足实际应用中实时性的要求^[19]。表2列出了在不同图像序列中的详细对比结果。

表2 各算法在不同图像序列中的用时比较(保留到毫秒)

	plane	pole	car	sylv	me	
总帧数/f	785	412	659	1344	900	
图像尺寸/像素	320 * 240	768 * 576	360 * 240	320 * 240	1200 * 900	
目标大小/像素	65 * 60	25 * 58	105 * 85	92 * 96	305 * 413	
L1	3081.253	1785.372	2873.879	6171.211	3865.973	
APG-L1	80.673	44.403	54.515	151.208	98.107	
LSA-L1	288.362	169.827	266.728	512.064	300.892	
PCA-L1	365.543	214.492	317.638	603.456	531.658	
MIL	181.652	115.392	159.076	321.315	236.331	
HL-L1	28.838	16.701	26.749	58.182	41.031	
L1	3.920	4.333	4.361	4.592	4.296	
平均	APG-L1	0.103	0.108	0.083	0.112	0.109
每帧	LSA-L1	0.367	0.412	0.405	0.381	0.334
用时	PCA-L1	0.465	0.521	0.482	0.449	0.591
/s/f	MIL	0.231	0.280	0.242	0.239	0.263
	HL-L1	0.036	0.040	0.040	0.043	0.045

通过表1和表2可知,pole和me图像序列的图像尺寸比较大,本文所提的跟踪算法不会增加处理时间,该算法的效率与图像序列尺寸没有关系;car和me图像序列中的目标尺寸比较大,本文所提算法也不会增加处理时间,说明该算法与目

标大小无关;sylv和me图像序列具有多帧,本文所提算法能高速地处理每一帧,不会出现随着时间增加处理效率降低的情况;plane和sylv图像序列背景比较简单,me图像序列的背景相对复杂,本文算法处理每帧的时间会有限增加,但增加的

时间完全可以控制在毫秒级。因此,本文算法是一种高速的 L1-跟踪算法。

4.2 跟踪精确度与质量

为了验证本文所提算法的跟踪质量,测试图像的选择情况和具体测试问题如表 3 所列。

表 3 图像序列测试用途

序列名称	分类	测试问题
plane	目标与镜头同时动	短时部分遮挡、背景干扰和目标形变
car	目标与镜头同时动	光照、场景变化和非目标运动物体干扰
pole	目标动,镜头静止	小部分遮挡与非目标运动物体干扰
sylv	目标动,镜头静止	光照、物体形变、背景干扰
me	目标与镜头同时动	全遮挡、部分遮挡和目标形变

其中,plane 图像序列背景由简单到复杂,主要用来测试跟踪算法处理短时遮挡、背景干扰和目标形变问题的能力;car 图像序列中既有小的光照变化下道路边树的阴影,又有大的光照变化下的天桥阴影,它主要用来测试光照和场景变化问题;pole 图像序列中有目标遮挡以及行人和移动车辆干扰,主要用来测试运动物体干扰和小部分遮挡问题;sylv 图像序列在台灯下存在光照变化,同时伴随目标形变与背景颜色的干扰,主要测试光照与物体形变等问题。L1,APG-L1,LSA-L1,PCA-L1,MIL 和本文算法(HL-L1)的实验结果见图 6。

图 6 利用不同亮度深浅的矩形框对各种算法的跟踪效果进行标注。其中,颜色最亮的白色矩形框代表本文算法 HL-L1 跟踪算法的结果,其它矩形框按照亮度由浅到深依次代表 MIL,PCA-L1,LSA-L1,L1 以及 APG-L1 等跟踪算法的跟踪结果。

本文以欧氏距离(ED)作为跟踪精度衡量标准,定量比较各种算法的跟踪精度,人工手动标定目标真实位置中心,跟踪位置误差(欧氏距离)定义如下:

$$ED(t) = \sqrt{(x_t - \hat{x}_t)^2 + (y_t - \hat{y}_t)^2} \quad (23)$$

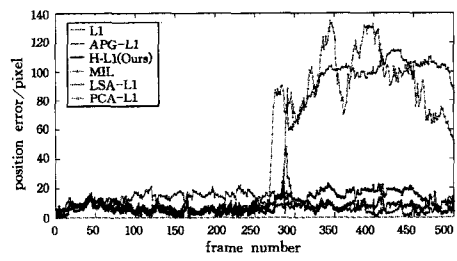
其中,(x,y)为目标真实位置中心,(\hat{x}, \hat{y})为算法预测目标位置中心。

各图像序列的误差曲线如图 7 所示。

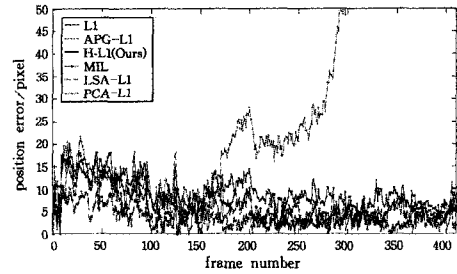
由图 6 和图 7 可知,本文算法提高了 L1-跟踪算法和 APG-L1 跟踪算法的鲁棒性,在处理目标形变与光照变化等问题上表现出了优越的处理能力。

对于 plane 和 pole 两个图像序列,由图 7(a)和图 7(b)可知,本文算法拥有合理的模板更新规则,能与 L1-跟踪算法和 APG-L1 跟踪算法一样准确地跟踪目标,克服复杂背景干扰与短时大部分遮挡,LSA-L1 跟踪算法将背景与目标分离,对背景干扰不敏感;由于 plane 图像序列中第 272 帧存在背景干扰,MIL 与 PCA-L1 对该帧图像之后的序列无法准确跟踪目标,说明 MIL 与 PCA-L1 无法有效地解决背景干扰。

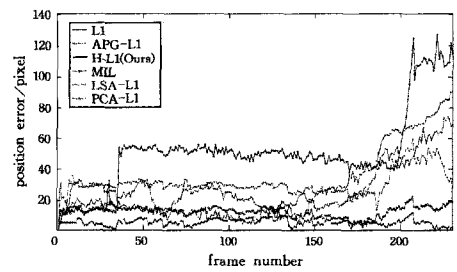
对于 car 和 sylv 图像序列,再由图 6 与图 7(c)和图 7(e)可知,由于存在局部和全部光照变化,L1,APG-L1 跟踪和 MIL 算法的跟踪精度都存在一定的偏差;对于 sylv 图像序列,PCA-L1 跟踪算法在 110 帧时由于目标在光照变化时出现了形变,因此导致跟踪失败;本文算法由于选用了像素块的思想,在处理光照和目标形变问题表现得特别突出,跟踪精确度明显优于其他算法。



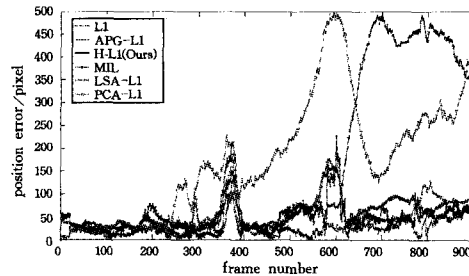
(a) plane



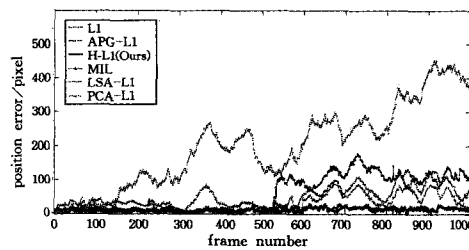
(b) pole



(c) car



(d) me



(e) sylv

图 7 定量比较跟踪位置误差曲线图(像素)

me 图像序列尺寸和目标尺寸都比较大,由图 6 和图 7 (d)可知,LSA-L1 跟踪算法具有很好的抗背景干扰能力,但在 me 图像序列中第 620 帧跟踪失败,主要原因是 LSA-L1 跟踪算法在目标变化或移动较小的前提下候选模板集由前一帧得到的目标构成,对剧烈形变的目标的处理能力较弱,本文保留了 me 在 L1-跟踪算法中的目标坐标放射思想,跟踪区域根据目标形变进行小幅的变化来克服目标形变与瞬间较大目标

位移;在 me 图像序列中的 282-301 帧和 760-788 帧,跟踪目标被完全遮挡,而本文 HL-L1 跟踪算法依然能很准确地跟踪目标。

综上所述,本文改进的算法将算法完备基中正负小模板由单个像素替换为像素块,减少了目标模板个数,从而大幅提高了 L1-跟踪算法的效率;同时,利用目标模板实时更新策略解决了目标跟踪中遮挡、相似物体干扰等问题,保证了算法的鲁棒性。

结束语 本文在 L1 跟踪算法框架的基础上对过完备基构造和模板更新两方面进行了改进,提出的 HL-L1 算法大幅提高了 L1-跟踪算法的跟踪效率,能满足实际应用中的实时性要求,同时对目标物遮挡、形变、环境场景、光照变化等问题具有很强的处理能力。如果目标很小(几个像素),那么本文算法与 APG-L1 跟踪算法的效率基本上差不多。本文所提算法存在一个待改进的问题:在背景颜色与目标颜色基本相似的情况下,本文算法处理能力相对较弱,在适当增加样本模板的情况下其能得到相对改善,这也是我们下一步研究的方向。

参考文献

- [1] MEI X, LING H B. Robust visual tracking using ℓ_1 minimization [C]// 2009 IEEE 12th International Conference on Computer Vision. IEEE, 2009: 1436-1443.
- [2] WRIGHT J, YANG A Y, GANESH A, et al. Robust Face Recognition via Sparse Representation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2009, 31(2): 210-227.
- [3] ARULAMPALAM M S, MASKELL S, GORDON N, et al. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking [J]. IEEE Transactions on Signal Processing, 2002, 50(2): 174-188.
- [4] CANDÈS E J, WAKIN M B. An Introduction To Compressive Sampling [J]. IEEE Signal Processing Magazine, 2008, 25(2): 21-30.
- [5] MIE X, LING H B, WU Y, et al. Minimum error bounded efficient ℓ_1 tracker with occlusion detection [C]// 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2011: 1257-1264.
- [6] BAO C L, WU Y, LING H B, et al. Real time robust L1 tracker using accelerated proximal gradient approach [C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2012: 1830-1837.
- [7] JIA X, LU H C, YANG M H. Visual tracking via adaptive structural local sparse appearance model [C]// IEEE Conference on Computer Vision & Pattern Recognition. 2012: 1822-1829.
- [8] ZHONG W, LU H C, YANG M H. Robust object tracking via sparsity-based collaborative model [C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2012: 1838-1845.
- [9] WANG D, LU H C, YANG M H. Online object tracking with sparse prototypes [J]. IEEE Transactions on Image Processing, 2013, 22(1): 314-25.
- [10] WANG Q, CHEN F, XU W L, et al. Online discriminative object tracking with local sparse representation [C]// IEEE Workshop on the Applications of Computer Vision. IEEE Computer Society, 2012: 425-432.
- [11] BABENKO B, YANG M H, BELONGIE S. Robust Object Tracking with Online Multiple Instance Learning [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(8): 1619-1632.
- [12] PAPAGEORGIOU C P, OREN M, POGGIO T. A General framework for object detection [C]// International Conference on Computer Vision. 1998: 555-562.
- [13] WU G X, ZHAO C X, LU W J, et al. Efficient structured ℓ_1 tracker based on laplacian error distribution [J]. International Journal of Machine Learning and Cybernetics August, 2015, 6(4): 581-595.
- [14] WANG X Y, WANG Y, WAN W G, et al. Object tracking with sparse representation and annealed particle filter [J]. Signal Image and Video Processing, 2014, 8(6): 1053-1068.
- [15] VIOLA P, JONES M. Rapid Object Detection using a Boosted Cascade of Simple Features [C]// CVPR 2001. 2001: 1-511-1-518.
- [16] ZHOU F, JIANG W, LI S Q, et al. Moving Target Localization and Tracking Algorithms: A Particle Filter Based Method [J]. Journal of Software, 2014, 24(9): 2196-2213. (in Chinese)
周帆, 江维, 李树全, 等. 基于粒子滤波的移动物体定位和跟踪算法 [J]. 软件学报, 2013, 24(9): 2196-2213.
- [17] TSENG P. Approximation accuracy, gradient methods, and error bound for structured convex optimization [J]. Mathematical Programming, 2010, 125(2): 263-295.
- [18] ZHANG K, ZHANG L, YANG M H. Real-Time Compressive Tracking [C]// European Conference on Computer Vision. Springer-Verlag, 2012: 864-877.
- [19] YILMAZ A, JAVED O, SHAH M. Object tracking: A survey [J]. ACM Computing Surveys, 2006, 38(4): 81-93.
- [20] MEI X, LING H B, WU Y, et al. Efficient minimum error bounded particle resampling L1 tracker with occlusion detection [J]. IEEE Transactions on Image Processing, 2013, 22(7): 2661-2675.
- [34] BAO Y K, RU H Y, JIN S J. A New Algorithm of Knowledge Mining in Factor Space [J]. Journal of Liaoning Technical University (Natural Science), 2014, 33(8): 1141-1144. (in Chinese)
包研科, 茹慧英, 金圣军. 因素空间中知识挖掘的一种新算法 [J]. 辽宁工程技术大学学报(自然科学版), 2014, 33(8): 1141-1144.
- [35] WANG P Z, GUO S C, BAO Y K, et al. Factorial analysis in factor space [J]. Journal of Liaoning Technical University (Natural Science), 2014(7): 865-870. (in Chinese)
汪培庄, 郭嗣琮, 包研科, 等. 因素空间中的因素分析 [J]. 辽宁工程技术大学学报(自然科学版), 2014(7): 865-870.

(上接第 273 页)