

DOI: 10.13382/j.jemi.2017.08.021

# 采用局部线性嵌入的稀疏目标跟踪方法<sup>\*</sup>

孙锐 王旭 张东东 高隽

(合肥工业大学 计算机与信息学院 合肥 230009)

**摘要:**目标跟踪是从复杂的背景中辨认出运动目标,并且对目标进行准确且连续的追踪。如何在遮挡、形变、背景复杂的条件下鲁棒性跟踪目标仍是亟待解决的问题。针对遮挡和形变问题,提出一种局部线性嵌入(LLE)和稀疏表示的算法来有效的学习外观模板。其中LLE是流形学习的一种典型算法。在该算法中每个点的近邻权值在平移、旋转、伸缩变化下是保持不变的,因此可以用来提取目标的本质特征,发现数据的内在规律。算法首先采用局部线性嵌入提取低维特征,提取后的特征作为基向量与琐碎模板组成稀疏原型,稀疏原型用于模板的更新。算法保持了原有稀疏跟踪方法对遮挡处理的优势,同时对目标形变有较好的稳健性。实验结果表明,跟踪算法比其他7个常用的算法在9个视频序列中有较好的鲁棒性能。

**关键词:**外观模型;目标跟踪;局部线性嵌入;稀疏表示;流形学习

中图分类号:TN911.73 文献标识码:A 国家标准学科分类代码:510.4050

## Sparse object tracking method using local linear embedding

Sun Rui Wang Xu Zhang Dongdong Gao Jun

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

**Abstract:** Definition of object tracking is identifying the moving targets from complex background, and it should track the target accurately and continuously. In occlusion, deformation, complex background conditions robust tracking target is still a challenging problem to be solved. A novel online object tracking algorithm is proposed for the occlusion and deformation with sparse prototypes, which exploits local linear embedding (LLE) algorithm with sparse representation scheme for learning effective appearance model. LLE is a classic manifold learning algorithm. In the algorithm, the neighbor points weight of each point remains unchanged in translation, rotation, scale changes. Thus, it can be used to extract the essential characteristics of target and find the inherent law of data. Firstly, the algorithm uses the local linear embedding algorithm to extract low dimensional characteristic. Then the sparse prototype is composed of the base vector which is extracted from the low dimensional characteristic and trivial templates. It can be used to update templates. This algorithm maintains the advantages of the original sparse tracking method to occlusion, and has a good robust tracking effect of deformation object. The experimental results show that the proposed tracking algorithm is better than the other seven commonly used algorithms in the nine video sequences.

**Keywords:** appearance model; object tracking; local linear embedding (LLE); sparse representation; manifold learning

## 0 引言

目标跟踪是计算机视觉研究邻域的一个热门方向,被广泛应用于多个邻域,如智能导航、军事识别、视频安

全监控等。虽然经过十年的发展,已在目标跟踪邻域已经取得了很大的进步,但是遮挡、形变、光照的变化等导致的外观的变化一直是目标跟踪的一个难点。典型的跟踪方法由3个部分组成:1)外观模型<sup>[1]</sup>,这个模型可以用来评估观测图像块的可能性;2)动态模型<sup>[2]</sup>这个动态模

型旨在描述目标随时间的变化状态(卡尔曼滤波器<sup>[3]</sup>粒子滤波<sup>[3-4]</sup>);3)在当前帧找到最优的图像块的搜索策略(均值漂移<sup>[1]</sup>、滑动窗<sup>[4]</sup>)。本文提出了一个可以考虑遮挡和形变影响的具有鲁棒的外观模型。

目前已经提出的跟踪算法中,其外观模型的建模方式主要分为 3 类。1)产生式外观模型,它通过学习目标在视频中各种运动的变化,建立自适应外观模型实现跟踪,依据目标的表征方式不同可以再分为混合产生式学习和子空间产生式学习。近几年,基于稀疏表示的产生式外观模型引起了很高的关注,它将目标通过基函数字典映射成稀疏向量建立外观模型,结合增量学习更新字典子空间在粒子滤波框架下实现鲁棒跟踪。Adam 等人用多碎片设计一个外观模型,该模型对于部分遮挡是强健的。2)判别式外观模型<sup>[5]</sup>学习跟踪方法,它将跟踪问题转换成二分类问题,依据目标和背景的分离边界建立自适应外观模型(如支持向量机<sup>[6-7]</sup>、增强类算法<sup>[8]</sup>)。3)混合产生-判别式外观模型<sup>[9]</sup>学习跟踪方法,它结合了产生式模型能够描述广义变化和判别式模型能够描述差异性变化的优点。动态模型主要描述目标随着时间的推移不断变化,常用的方法有卡尔曼滤波器的算法和粒子滤波算法。粒子滤波的实质是用由粒子及其权重组成的离散随机测度近似相关的概率分布并且根据算法递推更新离散随机测度。由于粒子退化是粒子滤波算法中不可避免的现象,解决的方法主要是选取好的重要性密度函数和粒子重采样,于是有人提出了各种改进的粒子滤波算法,如基于重要密度函数选择的改进粒子滤波算法、基于重采样技术的改进粒子滤波算法和自适应粒子滤波算法。MeanShift<sup>[10]</sup>最早是由 Fukunaga 等人于 1975 年提出的一种常用的搜索策略,通过迭代运算找到目标位置,实现目标跟踪。其显著的优点是算法计算量小,简单易实现,很适合于实时跟踪场合;但是跟踪小目标和快速移动目标时常常失败,而且在全部遮挡情况下不能自我恢复跟踪。自从 Comaniciu 等人将均值漂移应用到跟踪邻域研究分为两大类,一类应用 MeanShift 搜索到目标模板和候选目标相似函数的最大值,另一类将 MeanShift 和粒子滤波结合,应用 MeanShift 搜索密度函数的模式。

遮挡、形变、光照的变化一直是目标跟踪邻域的难点问题。如何处理好遮挡形变的问题更是衡量跟踪算法鲁棒性的关键。流形学习<sup>[11]</sup>是从高维空间中找到低维流形,以实现维数约简。其中的典型算法一局部线性嵌入算法(LLE)<sup>[12-13]</sup>因其算法中有每个点的近邻权值在平移、旋转、伸缩变化下是保持不变的优点,可以用于图像特征的降维和提取,从而找到事物的本质,有助于解决遮挡和形变问题。Mei 等人<sup>[14]</sup>提出一个跟踪算法通过琐碎模板的稀疏表示模拟部分遮挡问题。基于这种算法思

想,运用 LLE 算法提取图像特征作为基向量和琐碎模板构成稀疏原型的算法来解决遮挡和形变问题。局部线性嵌入算法,能在保持原始数据内在特性不变的情况下,可以约简数据维数,得到维数低、敏感度高且聚类性好的低维特征。通过 LLE 基向量和琐碎模板的稀疏性更新模板,可以有效的解决遮挡和形变的问题。

## 1 流形学习与稀疏表示

### 1.1 流形学习降维

近几年来,流形学习已经在数据降维方面引起了很大的关注,而其中 LLE<sup>[15]</sup>在保持原始数据性质不变的情况下,约简数据的维数被应用到数据降维中。其核心思想是用局部线性拟合来求得可呈现在全局中的线性结构。具体的解释就是把原空间进行划分,使得各局部中的数据尽可能的可以用线性结构表述,即数据在局部区域呈线性状态,通过对每个小块的线性运算映射,再把所有部分进行整合,结果非常接近于全局的非线性结构。在保证局部信息不变中,要求解带约束的最小二乘问题,最优的解即数据点的重构权,这样不仅能保持区域特征属性不变,而且全部的信息也能通过局部相互叠加来实现,最终能够达到保持整体几何特点的效果。一些实验结果表明 LLE 学习方法应用到目标跟踪可以增加跟踪的稳健性而且还可以有效的处理遮挡和形变等问题。其中 LLE 算法流程如图 1 所示。

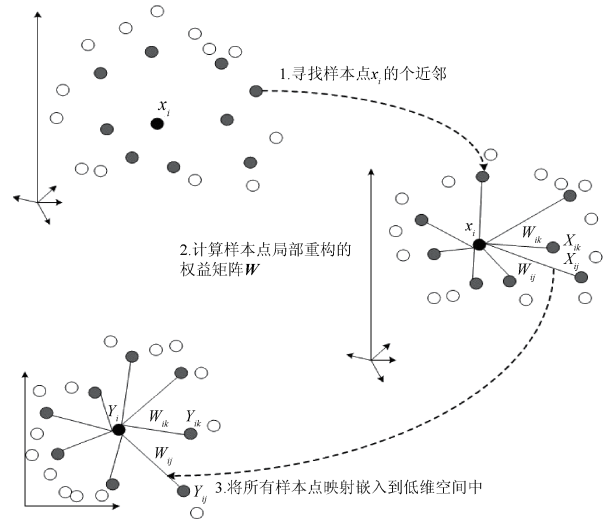


图 1 流形学习(LLE)的过程

Fig. 1 Manifold learning (LLE) process

1) 寻找每个样本点  $x_i$  的  $k$  个  $d_{ij}$  近邻。定义样本集矩阵  $X = \{x_1, x_2, \dots, x_N, \}$  中样本的  $x_i$  点与  $x_j$  的距离为

$$d_{ij} = \left[ \sum_{m=1}^D |x_{mi} - x_{mj}|^p \right]^{1/p} \quad (1)$$

式中:通常取  $p = 2$ ,即采用欧氏距离。根据式(1)计算得到样本点  $x_i$  的  $k$  个距离最近的点作为  $g_i$  的  $k$  个近邻点。

2) 计算样本点局部重构的权值矩阵  $W$ 。定义误差函数为:

$$\min \varepsilon(W) = \sum_{i=1}^N (x_i - \sum_{j=1}^k w_{ij} g_j)^2 \quad (2)$$

式中:  $g_j (j = 1, 2, \dots, k)$  为  $x_i$  的  $k$  个近邻点,  $w_{ij}$  为  $x_i$  与  $g_j$  之间的权值系数,且满足:

$$\min \varphi(Y) = \sum_{i=1}^N (y_i - \sum_{j=1}^k w_{ij} h_j)^2 \quad (3)$$

将式(3)代入式(2),得到:

$$\min \varepsilon(W) = \sum_{i=1}^N [ \sum_{j=1}^k w_{ij} (x_i - g_i) ]^2 = \sum_{i=1}^N w_i^T z_i w_i \quad (4)$$

其中,求解式(4)是一个约束最小二乘问题。

3) 将所有样本点映射嵌入到低维空间中。映射嵌入满足:

$$\min \varphi(Y) = \sum_{i=1}^N (y_i - \sum_{j=1}^k w_{ij} h_j)^2 \quad (5)$$

式中:  $\varphi(Y)$  为损失函数,  $y_i$  为  $x_i$  的输出向量。  $h_j (j = 1, 2, \dots, k)$  为  $k$  的  $y_i$  个近邻点,且满足:

$$\begin{cases} \sum_{i=1}^N y_i = 0 \\ \frac{1}{N} \sum y_i y_i^T = I \end{cases} \quad (6)$$

式中:  $I$  为  $d \times d$  的单位矩阵。

$w_{ij} (i = 1, 2, \dots, N)$  可以存储的稀疏矩阵  $W$  中。当  $x_j$  是  $x_i$  的近邻点时,  $W_{ij} = w_{ij}$ ;  $W_{ij} = 0$ ; 否则  $Y = [y_1, y_2, \dots, y_N]$  用  $W$  表示  $W_i$  矩阵的第  $i$  列其中  $Y$  向量,即:

$$\min \varphi(Y) = \sum_{i=1}^N (YI_i - YW_i)^2 = Y(I - W)(I - W)^T Y^T = YMY^T \quad (7)$$

结合约束条件,利用 Lagrange 乘子有:

$$L(Y) = YMY^T + \lambda (YY^T - NI) \quad (8)$$

为使损失函数  $L(Y)$  最小,则:

$$\frac{\partial L}{\partial Y} = 2MY^T + 2\lambda Y^T = 0 \quad (9)$$

将  $M$  的特征值按从小到大顺序排序  $Y$  取  $M$  的最小非零特征值所对应的特征向量。一般地,  $M$  的第 1 个特征值几乎等于 0。因此取  $2 \sim d + 1$  至个特征值所对应的特征向量作为低维嵌入。

### 1.2 稀疏表示

稀疏表示 (sparse representation), 也称为稀疏编码 (sparse coding) 或者稀疏分解 (sparse decomposition), 最早来源于神经科学领域的有效编码<sup>[16-17]</sup> 即认为生物视觉系统总是用尽量少和有限的资源去表示尽量多的外界信息。研究人员发现,对于自然环境中的刺激,视觉皮层细

胞的响应满足稀疏分布,进而认为,可以根据该机理对自然图像进行编码,这称为稀疏编码。为此研究人员基于图像稀疏特性相继构建有关模型。其中 Olshausen 和 Field 所做的工作较为关键,为后续稀疏表示的深入研究和广泛应用奠定了坚实基础。他们构建了基于自然图像稀疏特性的数学模型,认为每幅图像由过完备字典 (overcomplete dictionary)<sup>[16]</sup> 线性组合而成。基函数可以从自然图像中学习而来,且相对于字典是稀疏的。文献通过模拟实验证明经过学习的基函数与简单细胞的感受野特性相似。稀疏表示的基本思想从过完备字典(可能是通过学习得到)中找到尽量少的更能符合被逼近信号的原子 (atom), 用它们的线性组合来表示原信号。若定义向量  $y \in R^N$  表示待逼近的信号用矩阵  $D \in R^{N \times M}$ , ( $N < M$ ) 表示过完备字典, 向量  $\beta \in R^M$  表示系数 (coefficient) 那么稀疏表示抽象模型如图 2, 稀疏的数学公式如下:

$$\min \|\beta\|_0, \text{ s. t. } y = D\beta \quad (10)$$

式中:  $\|\cdot\|_0$  为  $l_0$  范数,表示向量中非零元素的个数,称为稀疏度  $s_p \in Z^+$ 。  $N < M$  意味着矩阵  $D$  的行数小于列数,体现了过完备性 (overcompleteness); 而中“0”值个数越多,则越能体现稀疏性 (sparsity)。这里,矩阵  $D$  的每一列称为基向量,他们并非一般意义上满足正交性的基, Zhang 等人也称其为原子,这些原子组成了过完备字典。流形学习因其能够在数据降维中很好的保存高维数据的结构,用于特征的提取。其中局部线性嵌入算法因为有许多优点而被广泛的应用到数据降维中,可以提取特征。本文中用局部线性嵌入方法来提取目标的特征作为基向量,用琐碎模板来解决遮挡问题。

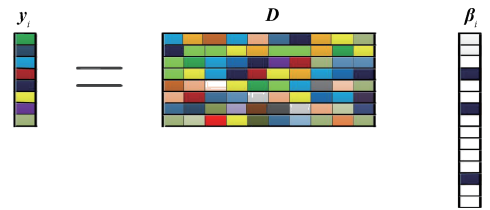


图2 稀疏表示的数学抽象

Fig. 2 Sparse representation of mathematical abstraction

## 2 基于 LLE 算法的稀疏目标跟踪

近年来,稀疏编码已经很成功的应用于目标跟踪中。稀疏目标跟踪主要思想是将每个候选目标表示成字典模板的一种稀疏线性组合。跟踪过程中需要对字典模板进行动态更新以迎合目标外观的变化。由式(11)可知,  $D$  是模板字典,  $D$  里每个列向量均为一个目标模板,稀疏表示实际是在解线性方程的过程中,使解向量尽量稀疏,求

解向量的过程实际上是一个凸优化的问题,可以通过凸松弛算法与贪婪算法计算。本文用的是基于局部直方图特征,开始时跟踪目标的状态是手动设置的。本文提出的是一种典型的基于贝叶斯推理框架下的追踪方法,即将目标跟踪问题看做是用隐藏状态变量在马尔科夫模型下的贝叶斯推理任务。在  $t$  坐标系下,给出一系列观察到的图像,  $Y_t(y_1, y_2, \dots, y_i)$  在隐藏状态变量  $X_t$  的递归:

$$p(x_t | Y_t) \propto p(y_t | x_t) \int p(x_t | x_{t-1}) p(x_{t-1} | Y_{t-1}) dx_{t-1} \quad (11)$$

式中:  $p(x_t | x_{t-1})$  表示在两个连续的状态的动态模型,  $p(y_t | x_t)$  表示的是观察模型在  $X_t$  状态下估计  $Y_t$  的可能性。在  $t$  时刻通过  $N$  个样本的最大后验概率可以获得在坐标系下观察值的最优的目标跟踪的模型。

$$\hat{x}_t = \operatorname{argmax}_p(y_t^i | x_t^i) p(x_t^i | x_{t-1}^i) \quad i = 1, 2, \dots, N \quad (12)$$

式中:  $x_t^i$  表示在第  $i$  个样本中的状态  $x_t$ ,  $y_t^i$  表示由预测的图像块。图 3 所示为跟踪算法的主要步骤。本文的创新点在于跟踪框架后的模板更新。Mei 等人<sup>[14]</sup> 提出一个跟踪算法通过计算模板的稀疏表示来找到最有可能的图像块。这种方法能够通过琐碎模板的稀疏表示模拟部分遮挡问题。基于 Mei 的算法,提出 LLE 算法和稀疏表示算法有效的学习外观模型。LLE 算法的基本思想是在样本点和它的邻域点之间构造局部线性平面,在此基础上建立优化目标函数也就是说,一个样本点可以由它的邻域点线性重构而成,而且重构权值使得样本点与邻域点的线性重构误差达到最小。对一个样本点首先通过计算该样本点的个近邻点,再由近邻点计算出该样本点的局部重建权值矩阵如式(2),最后由重建权值矩阵和近邻点计算出该样本点的输出值。在该算法中每个点的近邻

权值在平移、旋转、伸缩变化下是保持不变的,因此可以用来提取目标的本质,LLE 算法最突出的优点是经过处理后的低维数据均能保持原有的拓扑关系。本文算法首先采用局部线性嵌入提取低维特征,因此该算法可以有效的处理遮挡和形变的问题。提取后的特征作为基向量与琐碎模板组成稀疏原型,稀疏原型用于模板的更新如图 3 所示的算法流程和图 4 所示稀疏原型图。

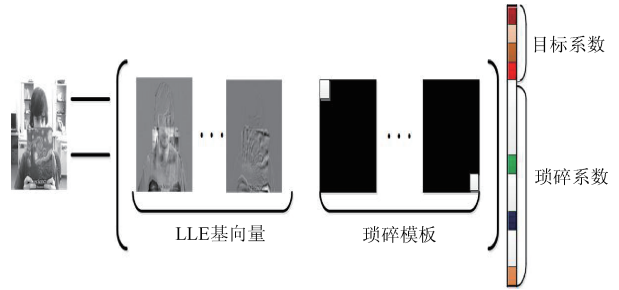


图 4 稀疏原型示意图

Fig. 4 Schematic diagram of sparse prototype

### 2.1 动态模型

在本文用一种仿射图像来模拟连续的两帧目标的运动。仿射变化的 6 个参数被用来跟踪目标的  $p(x_t | x_{t-1})$  的建模。在  $t$  时刻的状态为  $x^t = (x_t, y_t, \theta_t, s_t, \alpha, \phi_t)$ , 其中参数分别对应  $x, y$  水平位移和垂直位移、旋转角度、尺度、纵横比、倾斜度,  $p(x_t | x_{t-1}) = N(x_t; x_{t-1}, \psi)$  其中  $\psi$  代表协方差对角矩阵。

### 2.2 观察模型

如果没有遮挡发生,一个图像观察值  $y_t$  可以假定由跟踪目标子空间的中心  $u$  产生。然而为了增加跟踪的效果需要在外观模型中加入部分遮挡的处理。认为跟踪目标的一个中心图像观察值  $\bar{y}_t(\bar{y}_t = y_t - U)$  可以由 LLE 基向量  $U$  和少量元素的单位矩阵  $I$  (如,简单矩阵) 线性组合表示,  $\bar{y}_t = Uz_t + e_t$ 。注意到  $U$  是由少量的基向量组成  $Z_t$  通常大量的元素组成。另一方面是用  $e_t$  来说明噪音和遮挡。如果没有遮挡发生,通过 LLE 的基向量就可以很高效的表示最优的图像块,而此时的简单模板的系数趋于 0。另一方面,如果候选图像块没有对于图像的真实位置经常导致一个复杂的表示如果发生部分遮挡,最优的图像块可以有 LLE 基向量和简单模板线性组合表示所述,通过用简单模板的可以更加精确的找到目标的位置。对每个观察图像块的重构错误可以测量观测的可能性:

$$p(\bar{y}^i | x^i) = \exp(-\|\bar{y}^i - Uz^i\|^2) \quad (13)$$

式中:  $i$  表示状态  $x$  的第  $i$  样本。然而考虑到遮挡问题,用一种方法来区分遮挡部分和非遮挡部分。

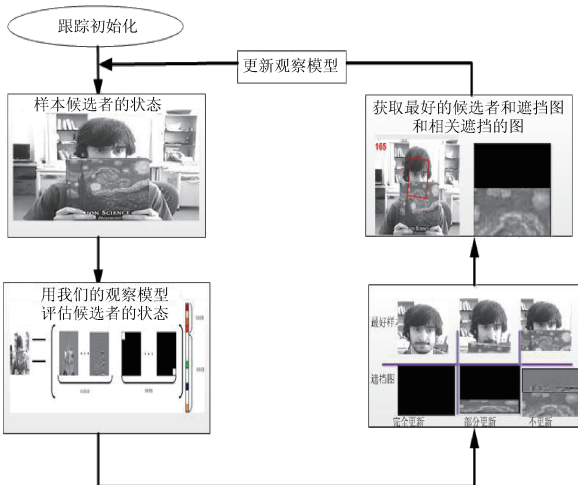


图 3 算法流程

Fig. 3 Flow chart of algorithm

$$p(\bar{y}^i | x^i) = \exp[-(\|w^i \Theta(\bar{y}^i - U z^i)\|_2^2 + \beta \sum (1 - w^j))] \quad (14)$$

式中:  $w^i = [w_1^i, w_2^i, \dots, w_d^i]^T$  是一个向量, 这个向量明  $e^i$  中的零元素。  $\Theta$  是序列分析  $\beta$  是处罚项。 如果  $e^i$  第  $j$  元素是 0, 那么  $w_j^i = 1$ , 或者  $w_j^i = 0$  指数的第 1 部分说明跟踪目标的没有遮挡部分重构错误的比例第 2 部分说明了被遮挡部分。

### 2.3 观察模板的更新

处理视觉跟踪中目标跟踪外观变化时是有必要更新观察模板<sup>[18]</sup>。 如果用一些不准确的样本用来更新时, 模板将退化从而引起跟踪漂移。 相反, 利用简单模板的系数来检测遮挡, 因为它的模板有一定的抗噪性能。 首先, 由于反向光栅扫描的图像块形成了每个简单模板的系数向量。 这个图像的非零元素表明像素是遮挡的, 其次计算非零像素和遮挡图的像素数目的比例。 用两个阈值  $tr1$  和  $tr2$  来描述遮挡程度(本文  $tr1 = 0.1$  和  $tr2 = 0.6$ ) 最后, 依据遮挡比例  $\eta$ , 形成了 3 种不同的类型: 全部更新、部分更新、不更新。 如果  $\eta < tr1$  直接用样本更新模板。 如果  $tr1 < \eta < tr2$  这表明目标是部分遮挡, 用它的相应部分的平均观察值  $u$  来替代遮挡部分像素。 用这种修复的样本更新模板。 另外如果  $\eta > tr2$  这意味跟踪目标的这一部分被显著遮挡。 应舍弃这个样本不用更新模板。 当累计了足够的样本后, 就用 LLE 基向量  $U$  和平均向量  $u$  来更新观察模板。

## 3 实验

本文提出的算法是通过 MATLAB 2015 平台实现的, 并且在性能测试实验在 Intel ( R ) Core ( TM ) CPU 3.10 GHz, 64 位 Windows 7 系统, 16 G 内存的计算机平台上运行的。 对于每个视频序列, 目标物体的初始位置是通过第一帧中图像手动标记得到的。 在所有的实验中, 为了 LLE 表示每一图像观察值是  $32 \times 32$  像素, 16 维特征向量。 此外用了 1 024 个简单模板。 对于(4)而言,  $z$  和  $e$  是维数分别是 16 和 1024。 本文跟踪用 600 个粒子, 每 5 帧图片跟新一次模板。

为了验证本文算法的性能, 本文在几种不同的视频序列中进行了追踪, 其中包括 Occlusion1、Occlusion2、Car11、Car4、DavidOutdoor、Singer1、Stone、Deer。 这些视频分别包括了部分遮挡、形变、快速移动等目标跟踪中常见的难点。 为了进一步对比本文算法的性能, 通过和其他几种最先进的跟踪算法比较包括 IVT、L1T、PN、VTD、MIL、Frag、OurTracker。 评价其他的跟踪算法性能时都采用原作者提供的原始数据, 并且给予适当的参数。

### 3.1 定量分析

本文采用两种评价准则量化不同跟踪算法的性能。 第一种是跟踪结果重叠率, 对于给定的跟踪结果区域  $R_T$  和目标真实区域  $R_C$ , 通过使用式(15) 来估计重叠率, 当计算结果大于 0.5 认为跟踪成功。

$$score = \frac{area(R_T \cap R_C)}{(R_T \cup R_C)} \quad (15)$$

表 1 给出了 6 种当前比较先进的跟踪算法在不同视频序列下跟踪结果的平均重叠率, 其中黑色代表最优, 其他代表次优。 图 5 所示为整个视频中所有帧的重叠率, 这些视频序列包含了遮挡、形变、光照变化等不同的问题。 可以看出本文的算法具有明显的优越性, 跟踪结果不是最优的就是次优的。 本文算法跟踪的优势主要归功于 LLE 算法可以很好表达数据的内在流形结构, 能够保留数据的本质特征。 用局部线性嵌入算法提取特征可以得到维数低, 敏感度高且聚类性好的低维特征从而使得跟踪结果更加准确。

表 1 7 种跟踪算法在 7 个视频序列行的平均重叠率比较  
Table 1 Average overlap rates comparison of 7 tracking algorithms in 7 lines of video sequence (%)

Sequence	IVT	L1T	PN	VTD	MIL	Frag	Ours
Car4	29	41	19	35	22	43	<b>32</b>
Car11	22	33	<b>25</b>	26	50	64	20
Singer1	85	44	34	153	180	22	<b>42</b>
Stone	111	44	20	89	166	68	<b>33</b>
Occlusion1	92	65	175	111	254	<b>56</b>	47
Occlusion2	<b>102</b>	111	172	141	155	104	49
David Outdoor	66	76	97	15	77	93	<b>51</b>
Girl	90	60	30	85	160	60	<b>50</b>
Deer	80	50	<b>40</b>	96	180	65	39

第 2 种评价准则是平均中心误差, 是通过计算真实状态和测量状态间的像素距离的平均值得到的, 得到的值越小代表误差越小, 追踪越准确。 表 2 给出了不同算法的平均中心误差, 图 6 所示为视频序列每帧的中心误差。 可以看出本文算法特别是对遮挡比较严重的视频序列, 追踪过程中的平均中心误差明显很小, 追踪效果很好, 在所有视频序列中平均中心误差不是最优的就是次优的, 所以可以看出本算法对物体遮挡问题的追踪的优秀性能。

### 3.2 定性分析

如图 7 所示序列 Occlusion2 中被跟踪的人脸经历了严重的遮挡和平面内旋转变化的。 本文的跟踪算法在第 184 帧、337 帧、501 帧中表现良好。 可以看出跟踪算法在处理严重遮挡和目标旋转有明显的优势。 而 IVT 算法因为对遮挡敏感, 所以该跟踪框一直在变小以躲避遮挡区域。 Frag 因采用固定的分块方式, 从而无法精确地定位人脸的位置和尺度。 尽管 MIL 算法能够很成功的跟踪

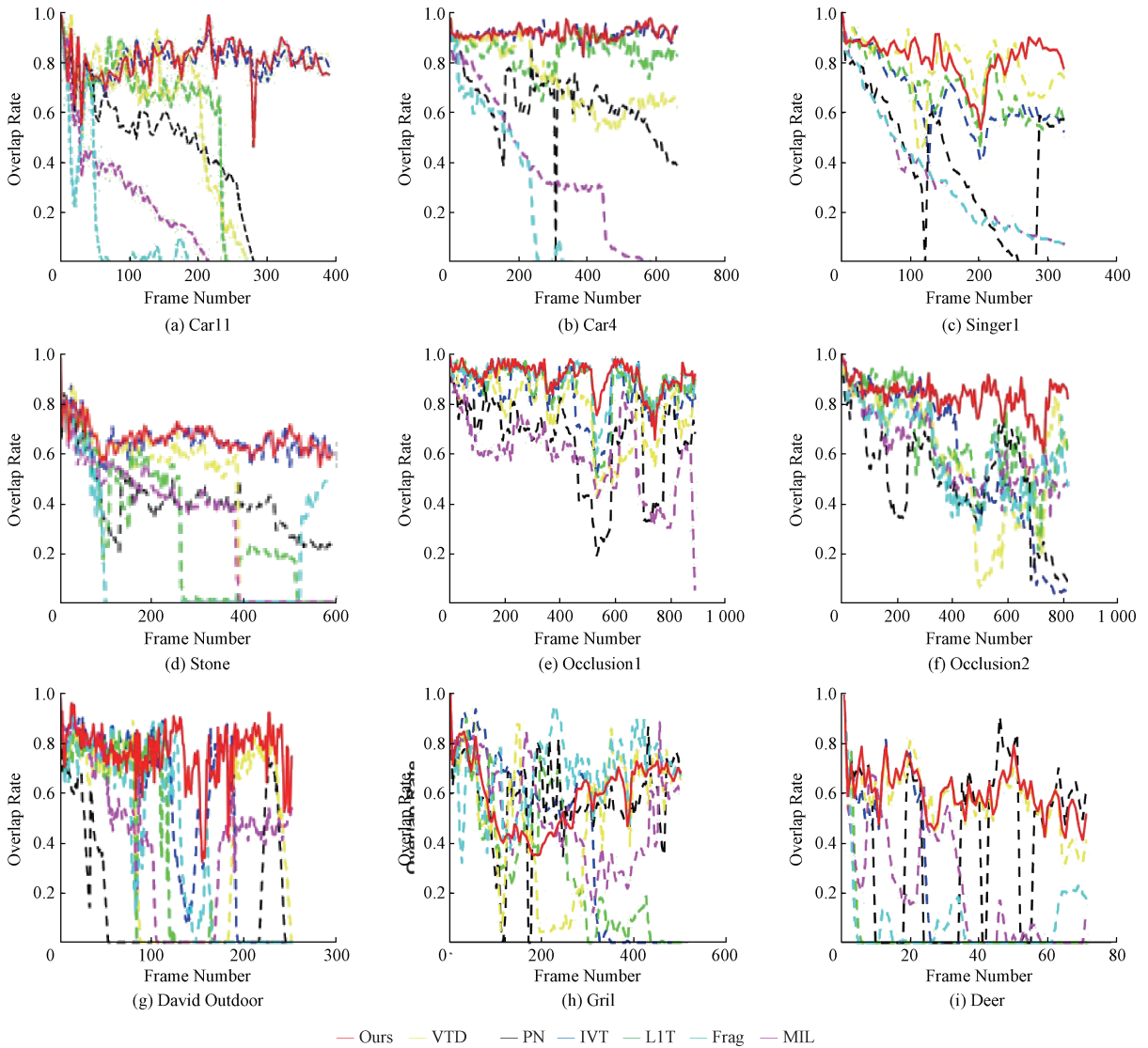


图 5 本文算法在 9 种不同视频帧下的跟踪结果重叠率,并且将本文算法同 6 种先进的跟踪算法比较

Fig. 5 Compared with 6 improved tracking algorithms, the overlap rate of this algorithms in 9 different lines of video sequence

表 2 7 种跟踪算法在 9 个视频序列行的平均中心误差比较  
Table 2 Average center errors comparison of 7 tracking algorithms in 7 lines of video sequence (pixels)

Sequence	IVT	LIT	PN	VTD	MIL	Frag	Ours
Car4	29	41	19	35	22	43	32
Car11	22	33	<b>25</b>	26	50	64	20
Singer1	85	44	34	153	180	<b>22</b>	<b>42</b>
Stone	111	44	20	89	166	68	<b>33</b>
Occlusion1	92	65	175	111	254	<b>56</b>	47
Occlusion2	<b>102</b>	111	172	141	155	104	49
David Outdoor	66	76	97	15	77	93	<b>51</b>
Girl	90	60	30	85	160	60	<b>50</b>
Deer	80	50	<b>40</b>	96	180	65	39

人脸,但是无法捕捉到目标的旋转角度,因此对旋转的跟踪效果也不是很好。在图 8 序列 Occlusion1 中因本算法、Frag 算法和 LIT 算法把遮挡等因素考虑进去了所以表现都比较好。Frag 算法采用局部直方图特征来处理部分遮挡。LIT 采用稀疏表示加琐碎模板的方法处理部分遮挡。同样在图 9 所示的序列 David Outdoor 中的第 80 帧、182 帧、188 帧和图 10 所示的序列 Girl 中的第 423 帧、432 帧、435 帧中被跟踪的人的脸部都经历了严重的遮挡,但是本文跟踪算法依然有较好的鲁棒性能。

如图 11 所示序列 Car4 和图 12 所示序列 Singer1 对目标的跟踪有明显的形变和光照的剧烈变化情况。在图 11 序列中在第 1 帧和第 294 帧中,汽车的行驶过程中经

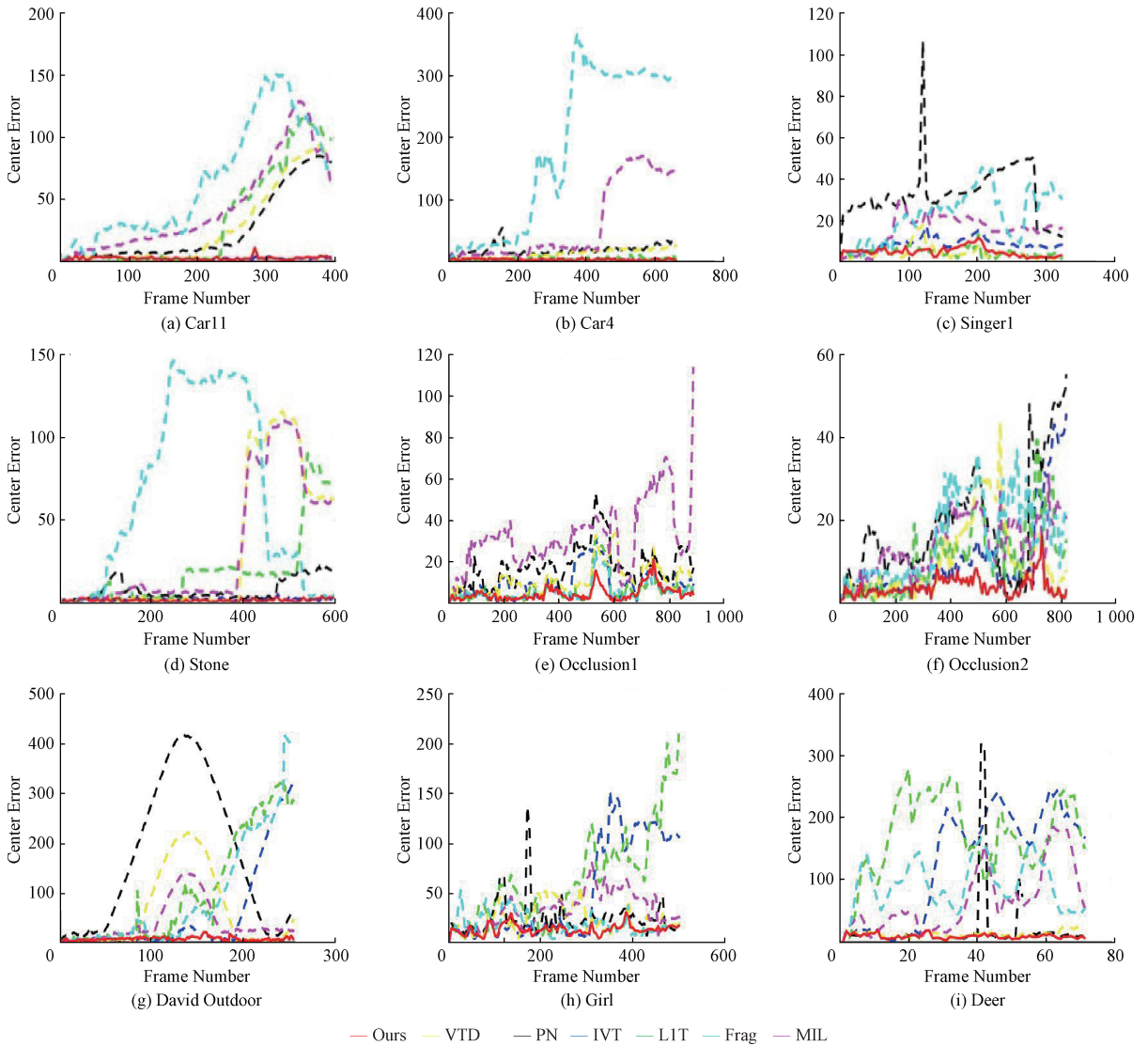


图6 跟踪结果每帧的中心误差

Fig.6 The center error of each frame in tracking result

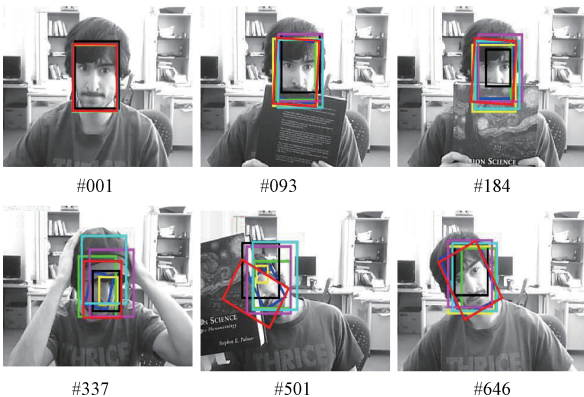


图7 视频序列 Occlusion2 跟踪结果

Fig.7 Tracking results of video sequence Occlusion2

历了尺度变换,必须要求跟踪算法中每一帧中能够精确的定位目标的位置和尺度,否则很容易漂移到其他目标上导致跟踪失败。同时第1帧、第190帧、294帧中虽然背景的光照变化剧烈,但本文算法依然表现很好。在图12所示的 Singer1 序列中在第1帧、第85帧、第321帧中跟踪目标发生了形变,本文算法没有产生漂移,跟踪效果很好。同样在第1帧和第179帧中光照变化剧烈,本文算法表现依旧很好。由于本文算法首先采用局部线性嵌入提取低维特征,提取后的特征作为基向量与琐碎模板组成稀疏原型,算法保持了原有稀疏跟踪方法对遮挡处理的优势,同时对目标形变有较好的稳健性,故而跟踪效果很好。



图 8 视频序列 Occlusion1 跟踪结果

Fig. 8 Tracking results of video sequence Occlusion1

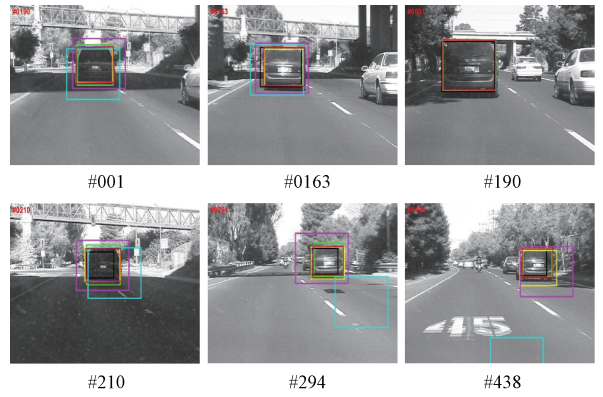


图 11 视频序列 Car4 跟踪结果

Fig. 11 Tracking results of video sequence Car4

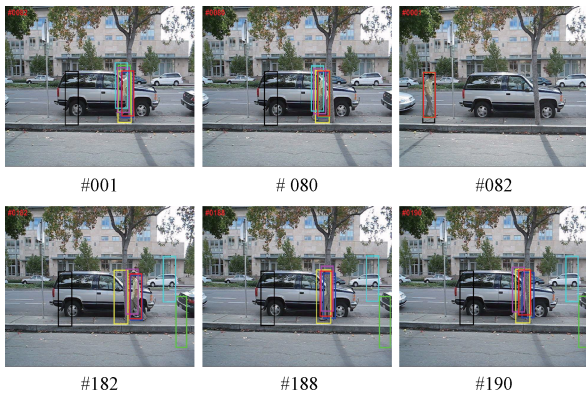


图 9 视频序列 David Outdoor 跟踪结果

Fig. 9 Tracking results of video sequence David Outdoor

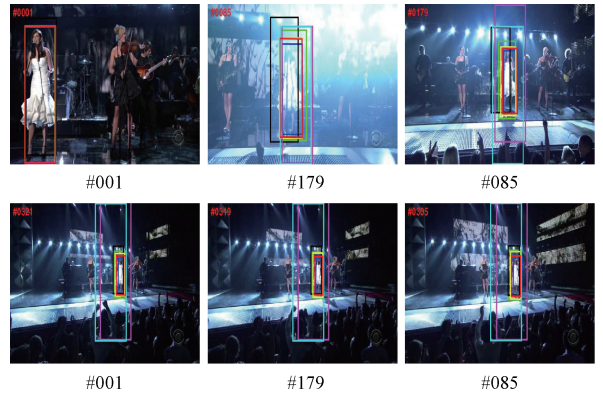


图 12 视频序列 Singer1 跟踪结果

Fig. 12 Tracking results of video sequence Singer1

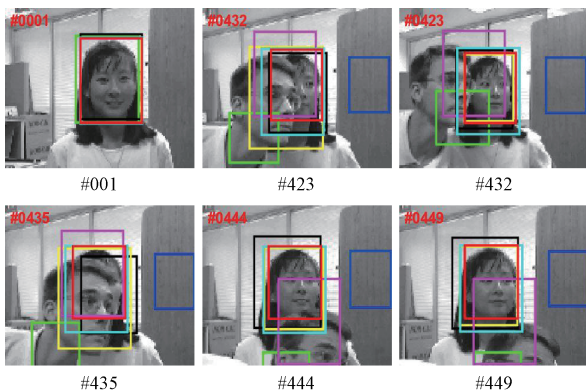


图 10 视频序列 Girl 跟踪结果

Fig. 10 Tracking results of video sequence Girl

地把部分遮挡和形变等因素考虑在外观更新和目标跟踪中。在经过一系列有挑战性的图像序列实验后,本文跟踪算法表现良好。实验结果也可以看出,所提算法很好地解决了部分遮挡和形变的问题,使得跟踪结果非常准确。未来关注点在于把稀疏表示计划应用到机器视觉其他领域中,如目标识别等。

### 参考文献

[ 1 ] LI ZH Y, WANG D M, NAI K, et. al. Robust object tracking via weight-based local sparse appearance model [ C ]. International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD) ,2016:560-565.

[ 2 ] JI ZH G, WANG W Q, LU K. Robustly tracking objects via multi-task kernel dynamic sparse model [ C ]. International Conference on Image Processing (ICIP) , 2015:266-270.

[ 3 ] 李奇越,吴忠,黎洁,等. 基于改进卡尔曼滤波的 NLOS 误差消除算法 [ J ]. 电子测量与仪器学报, 2015, 29(10) :1513-1519.

## 4 结 论

本文通过提出一种稀疏原型表示来展示了一种鲁棒性的跟踪算法。通过 LLE 算法和稀疏表示的优势明确



- LI Q Y, W ZH, LI J, et al. NLOS error elimination algorithm based on modified Kalman filtering [J]. Journal of Electronic Measurement and Instrumentation, 2015, 29(10):1513-1519.
- [4] 任航. 基于拟蒙特卡洛滤波的改进式粒子滤波目标跟踪算法[J]. 电子测量与仪器学报, 2015, 29(2): 289-295.  
RENG H. Improved particle filter target tracking algorithm based on Quasi Monte Carlo filtering [J]. Journal of Electronic Measurement and Instrumentation, 2015, 29(2): 289-295.
- [5] ZHU G, WANG J, LU H. Object tracking with part-based discriminative context models [C]. IEEE International Conference on Image Processing (ICIP), 2014: 4932-4936.
- [6] 吴哲君, 赵忠华, 唐雷. 基于 SVM 的行人步态实时分类方法[J]. 电子测量技术, 2015, 38(7):41-44.  
WU ZH J, ZHAO ZH H, TANG L. SVM-based method for real time classification of pedestrian gait [J]. Electronic Measurement Technology, 2015, 38(7): 41-44.
- [7] 罗子安, 毛征, 张辉. 基于特征点集群目标跟踪与检测[J]. 国外电子测量技术, 2016, 35(7):34-42.  
LUO Z N, MAO ZH, ZHANG H. Base on cluster feature tracking and detection [J]. Foreign Electronic Measurement Technology, 2016, 35(7):34-42.
- [8] 徐超, 高梦珠, 查宇锋, 等. 基于 HOG 和 SVM 的公交乘客人流量统计算法[J]. 仪器仪表学报, 2015, 36(2):446-452.  
XU CH, GAO M ZH, ZHA Y F, CAO L M. Bus passenger flow calculation algorithm based on HOG and SVM [J]. Chinese Journal of Scientific Instrument, 2015, 36(2):446-452.
- [9] THANG B D, MEDIONI G. Co-training framework of generative and discriminative trackers with partial occlusion handling [C]. WACV, 2011:642-649.
- [10] BAO X, DU K, TIAN H. Object tracking based on improved mean shift and SURF [C]. Consumer Electronics, Communications and Networks (CECNet), 2012:2716-2719.
- [11] SANCHES J M, NASCIMENTO J C. Manifold learning for object tracking with multiple nonlinear models [J]. IEEE Transactions on Image Processing, 2014, 23(4): 1593-605.
- [12] LI M A, LUO X Y. Feature extraction and visualization of MI-EEG by LLE algorithm [C]. IEEE International Conference on Mechatronics and Automation, 2016: 1989-1994.
- [13] TANG Y F, ZOU D Q. LLE coordinates for image blending [C]. International Conference on Virtual Reality and Visualization (ICVRV), 2014:278-283.
- [14] MEI X, LING H. Robust visual tracking using LI minimization [C]. IEEE International Conference on Computer Vision, 2009:1436-1444.
- [15] WANG D, LU H CH. Online object tracking with sparse prototypes [J]. IEEE Transactions On Image Processing, 2013, 22(1):314-325.
- [16] WANG W, YAN Y, ZHANG L M, et al. Collaborative sparse coding for multiview action recognition [J]. IEEE Multimedia, 2016, 23(4):80-87.
- [17] HU W R, YANG Y H, ZHANG W SH, et al. Moving object detection using tensor-based low-rank and saliently fused-spare decomposition [J]. IEEE Transactions on Image Processing, 2017, 26(2):724-737.
- [18] GEORGE A, SRINATH S A, PILLAI V Y. Modulation identification for intercept sonar based on over complete dictionary [C]. IEEE International Conference on Engineering and Technology (ICETECH), 2016: 1155-1159.

### 作者简介



孙锐, 1976 年出生, 现任合肥大学教授, 目前主要的研究方向为智能信息处理等。

E-mail: sunrui@hfut.edu.cn

**Sun Rui** was born in 1976. And he is a professor in Hefei University of Technology now. His present research interests include intelligent information processing and so on.



王旭, 1989 年出生, 现为合肥工业大学研究生, 目前主要研究方向为计算机视觉、目标跟踪。

E-mail: 913609584@qq.com

**Wang Xu** was born in 1989. And he is a M. Sc. candidate in Hefei University of Technology now. His present research interest includes computer vision and target tracking.