

一种融合多级稀疏表达和度量学习的目标跟踪方法

彭梦, 蔡自兴, 陈白帆

(中南大学信息科学与工程学院, 长沙 410083)

摘要: 基于稀疏表达的目标跟踪方法通常采用基于固定阈值的模板更新策略, 很难适应不断变化的目标外形; 其次, 稀疏表达缺乏描述目标流行结构的能力, 区分背景和目标的能力差. 针对基于固定阈值的模板更新策略的不足, 提出一种多级分层的模板字典. 为了改善对背景和目标的区分能力, 提出一种融合多级稀疏表达和度量学习的目标跟踪方法. 实验结果表明了所提出的方法能有效提高跟踪的鲁棒性和精度.

关键词: 目标跟踪; 稀疏表达; 度量学习

中图分类号: TP391

文献标志码: A

A target tracking method combining multi-level sparse representation and metric learning

PENG Meng, CAI Zi-xing, CHEN Bai-fan

(School of Information Science and Engineering, Central South University, Changsha 410083, China. Correspondent: PENG Meng, E-mail: pengmeng_pm@csu.edu.cn)

Abstract: Target tracking methods based on the sparse representation mostly apply a template update strategy based on the fixed threshold which is difficult to adapt to the changing shape of target. In addition, sparse representation is inadequate in capturing the manifold structures hidden in target samples. A template update strategy based on the multi-level hierarchical dictionary is proposed according to drawbacks of the template update strategy based on the fixed threshold. A tracking method combining multi-level sparse representation and metric learning is proposed in order to improve the ability to distinguish between background and targets. Experimental results show that the proposed method can improve the tracking accuracy and robustness effectively.

Keywords: target tracking; sparse representation; metric learning

0 引言

基于视觉的目标跟踪在许多领域被广泛应用, 例如智能监控、人机接口、运动分析和车辆导航等. 根据目标外形的学习方法, 视觉跟踪方法可以分为基于生成模型的跟踪方法^[1-3]和基于判决模型的跟踪方法^[4-6]. 基于生成模型的跟踪方法首先根据数据的分布情况来学习一个描述目标的生成模型, 然后通过生成模型寻找与目标相似度函数最大的区域. 生成模型的优点是能够反映同类数据本身的相似度, 缺点是对目标和背景的区分能力不足. 基于判决模型的跟踪方法将跟踪看成一个如何区分目标和背景的二分类问题, 根据目标和背景的决策边界进行跟踪. 判决模型的优点是反映不同类别数据之间的差异, 缺点是不能

反映训练数据本身的特性.

近年来, 人们提出了很多基于稀疏表达理论框架的目标跟踪方法. 基于稀疏表达理论, 目标可以通过同一类别的少数几个目标模板来线性表示, 较好地解决了目标遮挡、图像受损和目标外形的多线性结构低维流行问题.

基于稀疏表达的目标跟踪存在以下两个主要问题. 首先, 由于目标的外观一般都会发生变化, 目标模板需要不断更新^[7]. 1) 如果字典的样本更新过慢, 字典将不能适应目标在图像中的形态变化, 降低跟踪结果的精度. 2) 如果字典的样本更新过快, 容易将跟踪结果的背景部分引入字典, 使得跟踪的“drift”问题变得更加严重. 其次, 因为稀疏表达只用到了少数几个目标

收稿日期: 2014-07-06; 修回日期: 2014-12-25.

基金项目: 国家自然科学基金重大研究计划重点项目(90820302); 国家自然科学基金青年项目(61403423, 61403426).

作者简介: 彭梦(1978-), 男, 博士生, 从事视觉跟踪、多传感器融合的研究; 蔡自兴(1938-), 男, 教授, 博士生导师, 从事人工智能、智能控制、机器人等研究.

样本,不能根据目标样本的分布准确地描述目标外形的流行结构,所以稀疏表达方法对背景和目标的区分能力较弱.当在复杂背景环境中时,容易将背景部分引入目标模板中,使得跟踪结果变得不稳定、可控性差.

本文首先针对 L1 算法中模板字典更新策略的不足^[1],提出一种多级分层的目标模板字典,模板更新策略遵循“先低层再高层,先层内再层间”的模板更新原则,使得在每层内部相邻目标模板之间的时间跨度随层次的升高而逐层增大.因此在多级分层的目标模板字典中,低层目标模板的快速更新保证了目标模板字典的实时性,同时高层目标模板的缓慢更新提高了目标模板字典的稳定性、多样性.其次,为了弥补稀疏表达对背景和目标的区分能力比较弱的缺点,本文引入了度量学习理论,提出一种融合多级稀疏表达和度量学习的目标跟踪方法.本文的跟踪方法是基于粒子滤波跟踪框架,融合多级稀疏表达和度量学习方法来计算候选目标的置信度.

1 稀疏表达理论和度量学习理论介绍

1.1 基于稀疏表达理论的相似性计算

假设样本集合存在于一个低维流行中,可以使用过完备的训练样本字典,将每个测试样本通过同一类别的少数几个训练样本线性表示^[1].给定目标模板集合 $T = [t_1, \dots, t_n] \in \mathbf{R}^{d \times N}$,模板集合 T 中包含 N 个目标模板,每个目标模板 $t_i \in \mathbf{R}^d$,一个候选目标 y 的稀疏表达可由下式表示:

$$\min_c \|y - Dc\|_2^2 + \lambda_1 \|c\| + \frac{\lambda_2}{2} \|c\|_2^2. \quad (1)$$

其中训练字典 $D = [T, I, -I] \in \mathbf{R}^{d \times (N+2d)}$,琐碎模板集合 $I = [i_1, \dots, i_d] \in \mathbf{R}^{d \times d}$ 是一个单位矩阵,每一列 $i_i \in \mathbf{R}^d$ 对应一个琐碎模板, $c = \begin{bmatrix} a \\ e \end{bmatrix} \in \mathbf{R}^{N+2d}$ 是非负的稀疏系数向量, $a \in \mathbf{R}^N$ 为目标模板对应的稀疏系数, $e \in \mathbf{R}^{2d}$ 为琐碎模板对应的稀疏系数, λ_1 和 λ_2 为常数.

通过目标模板和琐碎模板作为原子所构建的训练字典,首先可以利用稀疏表达理论得到候选目标的稀疏线性表示;然后使用稀疏表达的重构误差作为候选目标 y 的相似性距离函数,即

$$d(y) = \|y - Ta\|_2. \quad (2)$$

1.2 基于度量学习理论的相似性计算

本文用到的度量学习是有监督的度量学习^[8-9],它利用样本之间的样本对约束,通过在马氏距离意义下最小化同类样本之间的距离和最大化异类样本之间的距离,从而得到一种区分性更好的度量空

间.对于线性度量学习的具体定义就是通过对原数据分析学习得到一种更能代表数据特征的度量矩阵 $M \in \mathbf{R}^{D \times D}$,使其在分类时能够将原数据映射到一个更优分类空间,即

$$D_M(x_i, x_j) = (x_i - x_j)^T M (x_i - x_j). \quad (3)$$

其中: $x_i \in \mathbf{R}^D$ 和 $x_j \in \mathbf{R}^D$ 为样本, M 为半正定矩阵.令 $M = A^T A$,可将 $A \in \mathbf{R}^{d \times D}$ 看成从原始特征空间 \mathbf{R}^D 到新度量空间 \mathbf{R}^d 的线性变换矩阵 $d \ll D$,即

$$D_A(x_i, x_j) = (Ax_i - Ax_j)^T (Ax_i - Ax_j). \quad (4)$$

基于度量学习的跟踪本质是,在一个区分性更好的子空间 \mathbf{R}^d 中去评估候选目标和模板的相似度.通过在线学习不断变化的目标外形和背景,可以得到一个反映样本语义关系的相似性度量函数,从而使得跟踪方法具有很强的区分目标和背景的能力.因此,首先可以通过文献[8]中的方法进行增量的度量学习;然后使用度量矩阵 M 计算候选目标 y 与目标模板 y^o 的马氏距离 $D_M(y, y^o)$,以此作为候选目标的相似性距离函数,即

$$d(y) = D_M(y, y^o). \quad (5)$$

2 多级分层的目标模板的更新策略

稀疏表达字典的更新分为两种:替换模板^[1-2]和字典学习^[10-14].当目标外观的低维流行分布为多线性结构时,字典学习方法使用单一线性模型易使学习出现过拟合,不能很好地保持流行中的局部线性关系.传统的替换模板方法采用基于固定阈值的模板更新策略,不能适应目标外观的变化.因此,为了更合理地对模板字典进行更新,本文改进了替换模板方法,提出一种多级分层的目标模板的更新策略.

假设将目标模板字典 $T = [t_1, \dots, t_N] \in \mathbf{R}^{d \times N}$ 分为 n 级层次,每层由 m 个目标模板组成 $T = [t_{1,1}, \dots, t_{1,m}, \dots, t_{n,1}, \dots, t_{n,m}] \in \mathbf{R}^{d \times N}$, $N = n \times m$,在每一层内部目标模板按照时间的前后顺序排列.本文目标模板的更新策略遵循以下规则:

1) 更新策略每次使用当前帧的跟踪结果去替换字典中第 1 级层次中最陈旧的目标模板.

2) 当第 $k-1$ 级的目标模板都被循环更新过一次时,选取与第 k 级目标模板差异性最大的第 $k-1$ 级目标模板作为替换模板 \hat{y} ,去替换第 k 级层次中最陈旧的目标模板,其中 $k = [2, 3, \dots, n]$.

3) 为了避免权重最大的目标模板被替换,每次更新字典时保留权重最大的目标模板的一个副本(目标模板的权重计算同文献[1]).本文多级分层的目标模板的更新算法流程如下.

算法 1 模板更新算法.

输入: 新跟踪到的目标 y , 图像时刻 t , 字典的层次数目 n , 每层模板的数目 m ;

输出: 更新后的目标模板字典 T .

1) 保留权重最大的目标模板;

2) For $k = 1, 2, \dots, n$, do

3) If $k = 1$, then

4) $\hat{y} \leftarrow y$;

5) 用 \hat{y} 替换第 k 级中最陈旧的目标模板;

6) break, /*结束更新*/

7) End if

8) If $\text{mod}(t, m^{k-1}) = 0$, then

9) 将第 $k-1$ 级的目标模板看成测试样本

$$Y = [t_{k-1,1}, \dots, t_{k-1,m}] \in \mathbf{R}^{d \times m};$$

10) 将第 k 级目标模板看成训练样本

$$A = [t_{k,1}, \dots, t_{k,m}] \in \mathbf{R}^{d \times m};$$

11) 对每个测试样本 $t_{k-1,i}$ 通过式 (6) 计算对应的稀疏系数 a_i ;

12) 根据式 (7), 计算 $\hat{y} \leftarrow \max_{t_{k-1,i}} \|t_{k-1,i} - Aa_i\|_2$;

13) 用 \hat{y} 替换第 k 级中最陈旧的目标模板;

14) Else

15) break, /*结束更新*/

16) End if

17) End for

本文是根据稀疏表达的重构误差来选取规则 2) 中的替换模板, 首先将第 $k-1$ 级的目标模板看成测试样本 $Y = [t_{k-1,1}, \dots, t_{k-1,m}] \in \mathbf{R}^{d \times m}$, 将第 k 级目标模板看成训练样本 $A = [t_{k,1}, \dots, t_{k,m}] \in \mathbf{R}^{d \times m}$, 训练字典为 $D = [A, I, -I]$, I 的定义同 1.1 节. 首先选用训练样本对测试样本进行稀疏线性表达, 由下式:

$$\min_{c_i} \|t_{k-1,i} - Dc_i\|_2^2 + \lambda_1 \|c_i\| + \frac{\lambda_2}{2} \|c_i\|_2^2 \quad (6)$$

计算 $t_{k-1,i}$ 的稀疏系数向量 $c_i = [a_i; e_i] \in \mathbf{R}^{m+2d}$, 其中 $a_i \in \mathbf{R}^m$ 为 $[t_{k,1}, \dots, t_{k,m}]$ 对应的稀疏系数, $e_i \in \mathbf{R}^{2d}$ 为琐碎模板对应的稀疏系数; 然后根据重构误差越大差异性越大的假设, 由下式:

$$\hat{y} = \max_{t_{k-1,i}} \|t_{k-1,i} - Aa_i\|_2 \quad (7)$$

选取稀疏表达的重构误差最大的测试样本, 作为替换模板 \hat{y} 替换第 k 级层次中最陈旧的目标模板.

由算法 1 可知, 目标模板字典每层内部的相邻目标模板平均时间跨度为 $m^{n-1}t$, 其中 m 表示层内目标模板的数目, n 为目标模板所在层的级数, t 为图像的采样间隔. 例如, 假设有 3 级目标模板, 每层目标模板的数目为 10, 则第 1 级相邻目标模板平均时间

跨度为 $10^0t = t$, 第 3 级相邻目标模板平均时间跨度为 $10^2t = 100t$.

由于每层内部的相邻目标模板平均时间跨度随层次的升高而增大, 低层相邻模板之间较小的时间跨度保证了模板更新的实时性, 高层相邻模板之间较大的时间跨度提高了目标模板字典的稳定性、多样性. 相比较 L1 算法^[1], 本文的模板更新方法不会因为频繁地引入新样本而造成字典中样本变化过快, 保证了目标模板字典结构的稳定性; 也不会因为过快淘汰旧样本而降低目标模板之间的差异性, 保证了目标模板的多样性; 同时本文方法每次都最新的跟踪结果作为目标模板加入字典中, 所以保证了目标模板字典在时间序列上的实时性.

本文使用文献 [1] 中的稀疏表达算法, 理论上对于每个测试样本计算其稀疏表示的时间复杂度为 $O(d^2 + dm)$, 因此本文使用的目标模板字典更新算法的时间复杂度为 $O(mn(d^2 + dm))$, 其中 d 为目标模板的像素个数 (即模板的向量维数), m 为每层内部的目标模板个数, n 为层的个数.

3 融合多级稀疏表达和度量学习的目标跟踪方法

背景环境呈现多样性和复杂性, 因此通过增加背景模板^[15-16]和提取区分性能高的特征^[16-17]并不能有效提高稀疏表达方法对背景和目标的区分能力. 为了使本文的跟踪方法能适应目标的外观变化和复杂的背景, 引入了度量学习理论, 联合多级稀疏表达和度量学习两种方法来改善跟踪的精度和鲁棒性. 稀疏表达反映了目标模板内在的特性, 即使在多线性结构的低维流行中也能较好地为目标进行线性表示, 因此跟踪精度较高; 度量学习反映了目标与背景之间的差异, 能较好地描述异类样本的分布, 因此区分能力较高. 当在稀疏表达框架中, 原始特征空间中的重构误差能较好地区分目标和背景时, 稀疏表达的跟踪结果更可信; 反之, 当稀疏表达的重构误差不足以区分目标和背景时, 度量学习的跟踪结果更可信.

在时间 t 给定粒子状态 $\{x_t^i\}_{i=1}^N$, 它的观测值为 $\{y_t^i\}_{i=1}^N$, 置信值为 $\omega_t^{i,j} = p_j(y_t^i | x_t^i)$ ($j = 1, 2$), N 为粒子数目. 本文的融合跟踪方法首先利用多级稀疏表达^[1]和度量学习^[8]两种方法 (式 (2) 和式 (5)) 分别计算每个粒子 x_t^i 的两种置信值

$$p_1(y_t^i | x_t^i) \propto \exp(-\|y_t^i - Ta^i\|_2^2), \quad (8)$$

$$p_2(y_t^i | x_t^i) \propto \exp(-D_M(y_t^i, y^o)); \quad (9)$$

然后根据粒子的两种置信值 $\{x_t^i, \omega_t^{i,j}\}_{i=1}^N$ ($j = 1, 2$) 的分布统计信息来计算每种置信值的可信度 π_t^j ($j =$

1, 2)^[18], 并且选择可信度高的一种置信值作为粒子的置信值; 最后使用算法 1 更新稀疏表达的字典 D 和使用文献 [8] 中的方法更新度量矩阵 M . 根据置信值的均值 μ_t^j 和方差 σ_t^j 的统计信息

$$\mu_t^j = \frac{1}{N} \sum_{i=1}^N \omega_t^{i,j}, j = 1, 2, \quad (10)$$

$$\sigma_t^j = \frac{1}{N} \sum_{i=1}^N \|\omega_t^{i,j} - \mu_t^j\|_2, j = 1, 2, \quad (11)$$

通过下式:

$$\pi_t^j = \frac{\sigma_t^j}{\sum_{i=1}^2 \sigma_t^i}, j = 1, 2 \quad (12)$$

计算置信值的可信度. 式 (12) 表达了粒子置信值的离散程度, 如果粒子的置信值分布越离散, 粒子的置信值分布越接近于单峰分布, 则这种置信值区分目标和背景的能力越强, 因此它可信度也越高. 本文提出的跟踪算法的流程如下.

算法 2 跟踪算法.

输入: 图像 I_1, I_2, \dots, I_T , 粒子数 N , 目标的初始状态 x_0 ;

输出: 每帧图像中目标的跟踪结果 \hat{x}_t .

1) 根据目标的初始位置初始化目标模板字典 T 、度量矩阵 M 和目标模板 y^o , 初始化粒子状态 $\{x_t^i\}_{i=1}^N$;

2) For $t = 1, \dots, T$, do

3) 根据运动模型 $p(x_t^i|x_{t-1}^i)$ 预测当前的粒子状态 $\{x_t^i\}_{i=1}^N$, 并获取观测值 $\{y_t^i\}_{i=1}^N$;

4) 根据目标模板字典 T 计算 $\{y_t^i\}_{i=1}^N$ 的稀疏表达系数 $\{a^i\}_{i=1}^N$;

5) 根据度量矩阵 M 计算 $\{y_t^i\}_{i=1}^N$ 与目标模板 y^o 的马氏距离 $\{D_M(y_t^i, y^o)\}_{i=1}^N$;

6) 分别根据式 (8) 和 (9) 两种方法计算 $\{x_t^i\}_{i=1}^N$ 的置信值 $\{\omega_t^{i,j}\}_{i=1}^N$ ($j = 1, 2$);

7) 根据置信值的分布, 使用式 (10)~(12) 计算可信度 π_t^j ($j = 1, 2$), 并选择可信度高的置信值作为融合后的置信度 $\{\omega_t^i\}_{i=1}^N$;

8) 取置信值最大的粒子作为跟踪结果 \hat{x}_t , 并更新稀疏表达的字典 T 和度量学习的度量矩阵 M ;

9) 对粒子 $\{x_t^i\}_{i=1}^N$ 进行重采样;

10) End for

4 实验分析

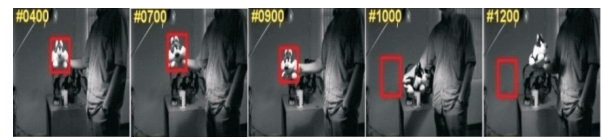
为了评估本文算法的执行结果, 在多个经典的公开测试视频序列上进行了一系列的实验, 并将本文

算法和当前流行的经典跟踪算法 IVT^[3]、LIAPG^[2]、MIML^[8]进行了比较. 这些视频中包含了各种有挑战性的情形, 包括目标遮挡、光照变化、复杂背景、旋转变化和尺度变化等. 本文的 Matlab 源程序源于 <http://pan.baidu.com/s/1bn6GUzmz>, 测试视频数据源于 <http://visual-tracking.net/>.

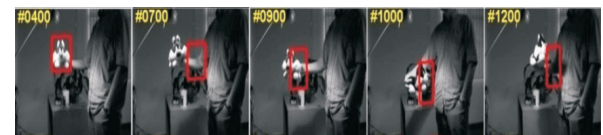
本文中所有算法统一将粒子数定为 200 个. 通过仿射变化将目标区域规则化, IVT 和 LIAPG 算法将目标区域规则化为 32×32 矩阵, MIML 和本文算法将目标区域规则化为 12×15 矩阵. IVT 算法的参数 $f = 0.95$, $m = 5$. LIAPGm 算法的稀疏表达参数 $\lambda_1 = 0.01$, $\lambda_2 = 0.001$, 目标模板字典中的模板个数定为 30 个. 本文算法稀疏表达参数 $\lambda_1 = 0.01$, $\lambda_2 = 0.001$, 目标模板字典中的模板个数定为 30 个, 分为 3 级层次, 每层 10 个模板.

4.1 跟踪结果的定性比较

第 1 个测试视频 *sylvester* 包含了光照变化和剧烈的姿态变化. 在第 400、700、900、1000 和 1200 帧各算法的比较结果如图 1 所示, IVT、LIAPG 和 MIML 算法不能适应变化而偏离了实际目标, 本文算法使用了多级分层的目标模板的更新策略, 模板字典具有较好的实时性和稳定性, 所以跟踪结果表现出较好的鲁棒性.



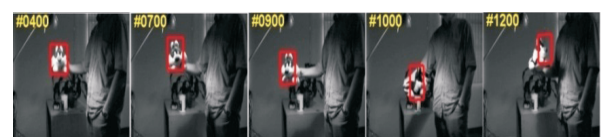
(a) IVT



(b) LIAPG



(c) MIML



(d) 本文算法

图 1 视频 *sylvester* 跟踪结果比较

第 2 个测试视频 *car4* 包含了较大的光照变化和复杂背景. 在第 200、300、400、500 和 600 帧比较各算法结果如图 2 所示, LIAPG 算法不能适应变化而偏离了实际目标, 本文算法由于融合了度量学习方法从而

对目标和背景具有较强的区分能力,对目标能进行稳定而精确的跟踪.

第 3 个测试视频 *dudek* 包含了目标遮挡和尺度、旋转的缓慢变化. 在第 400、600、800、900 和 1000 帧各算法的比较结果如图 3 所示, L1APG 和 MIML 算法不能适应变化而偏离了实际目标, 本文算法使用了多级分层的目标模板的更新策略, 所以跟踪结果表现出较好的稳定性.

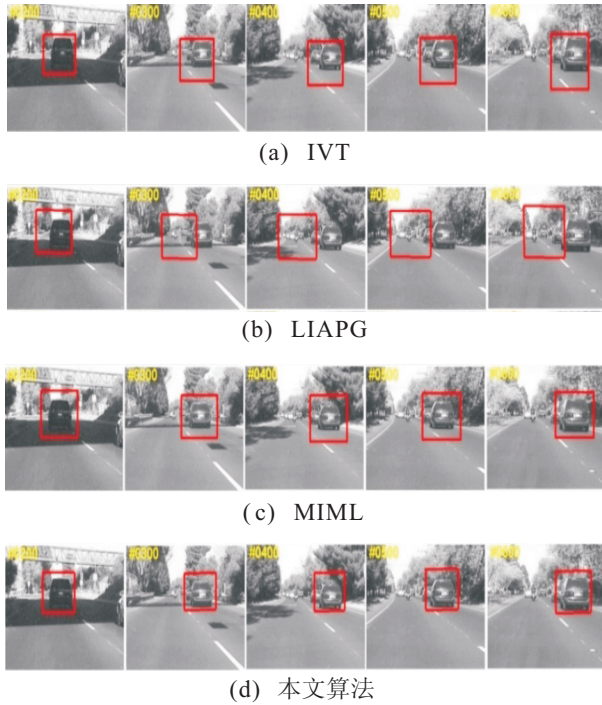


图 2 视频 *car4* 跟踪结果比较

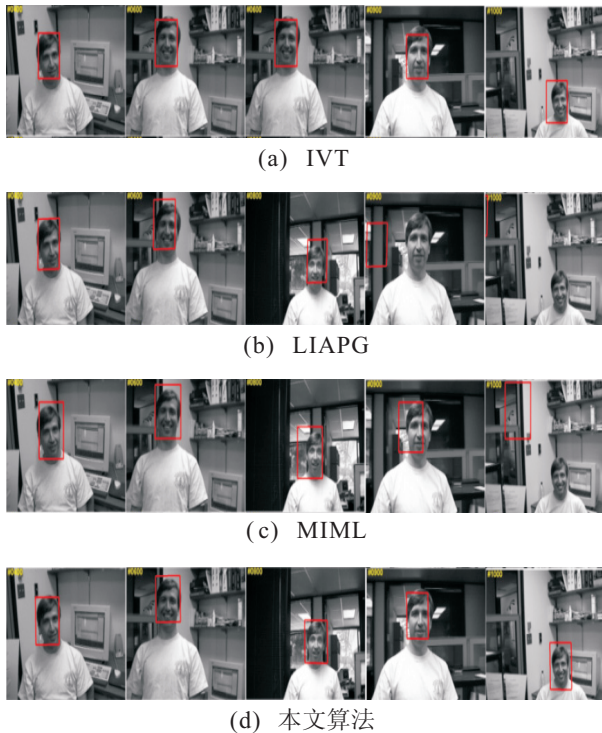


图 3 视频 *dudek* 跟踪结果比较

4.2 跟踪误差的定量比较与分析

本节对跟踪误差进行定量的分析, 使用了平均中心位置误差和平均重叠率作为定量分析的统计量. 中心位置误差是指图像中跟踪结果的中心位置与实际值的中心位置之间的欧式距离, 单位是像素. 重叠率指的是跟踪结果的区域和目标真实值的区域之间的重叠部分所占的比率. 表 1 比较了不同算法在 6 个经典视频中的跟踪结果的平均中心位置误差, 表 2 比较了不同算法在 6 个经典视频中的跟踪结果的平均重叠率, 可以看出本文算法在大部分视频中保持了较低的平均中心位置误差和较高的平均重叠率, 表现出好于其他算法的跟踪精度, 具有较好的跟踪鲁棒性.

表 1 比较不同算法在 6 个测试视频的平均中心位置误差

测试视频	IVT 算法	L1APG 算法	MIML 算法	本文算法
<i>sylvester</i>	34.2	23.5	56.8	7.7
<i>car4</i>	16.9	63.8	5.2	7.4
<i>dudek</i>	14.1	112.8	68.5	11.4
<i>walking2</i>	2.5	6.6	62.3	3.9
<i>skating1</i>	230.8	63.6	15.0	15.2
<i>football</i>	13.5	40.4	15.8	12.3

表 2 比较不同算法在 6 个测试视频的平均重叠率 %

测试视频	IVT 算法	L1APG 算法	MIML 算法	本文算法
<i>sylvester</i>	0.52	0.48	0.21	0.67
<i>car4</i>	0.51	0.25	0.46	0.68
<i>dudek</i>	0.73	0.56	0.62	0.74
<i>walking2</i>	0.79	0.72	0.29	0.60
<i>skating1</i>	0.08	0.20	0.31	0.45
<i>football</i>	0.59	0.48	0.54	0.60

IVT 算法通过单一的线性子空间模型计算候选目标的置信度, 并且采用固定频率更新模型, 所以当目标外观的低维流行分布为多线性结构或者目标外观发生较大变化时, 不能及时修正子空间模型. 因此在测试视频 *sylvester* 中, 由于剧烈的光照变化和目标姿态变化, 跟踪误差曲线出现了发散.

L1APG 算法采用基于固定阈值的模板更新策略, 字典的样本更新过慢, 不能适应目标外观的变化, 在所有测试视频中, 跟踪误差曲线最后都出现了发散.

MIML 算法通过单一的线性映射将高维空间中的样本特征投影到低维空间进行特征提取, 并且采用基于固定频率的度量空间更新策略, 与 IVT 算法同理, 当目标外观的低维流行分布为多线性结构时, 很难构建一个统一的线性的度量空间. 当发生剧烈的目标姿态变化或者遮挡时, 目标模板与候选目标在低维流行空间上的距离已经不能用两者在度量空间中的距离进行衡量, 因此跟踪误差曲线出现了发散.

本文算法由于使用了多级分层的目标模板的更新策略,模板字典具有较好的实时性和稳定性.同时,本文算法融合了度量学习方法,从而对目标和背景具有较强的区分能力.因此,本文的算法对目标能进行鲁棒而精确的跟踪.

5 结 论

基于稀疏表达的跟踪方法大都是采用固定阈值的模板更新策略,很难适应复杂环境中不断变化外形的目标.为此,本文首先提出了一种多级分层的目标模板字典,既保证了模板更新的实时性又提高了目标模板字典的稳定性、多样性;然后,为了弥补稀疏表达对背景和目标的区分能力较弱的缺点,本文引入了度量学习理论,提出了一种融合多级稀疏表达和度量学习的目标跟踪方法,同时兼顾了跟踪结果的精度和鲁棒性.实验结果定性和定量地显示出本文的方法能有效提高跟踪的鲁棒性和跟踪的精度.

参考文献(References)

- [1] Mei X, Ling H B. Robust visual tracking and vehicle classification via sparse representation[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2011, 33(11): 2259-2272.
- [2] Bao C L, Wu Y, Ling H B, et al. Real time robust L1 tracker using accelerated proximal gradient approach[C]. *IEEE Conf on Computer Vision and Pattern Recognition*. Singapore: IEEE Press, 2012: 1830-1837.
- [3] Ross D A, Lim J W, Lin R S, et al. Incremental learning for robust visual tracking[J]. *Int J of Computer Vision*, 2008, 77(1/2/3): 125-141.
- [4] Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2011, 33(8): 1619-1632.
- [5] Grabner H, Bischof H. On-line boosting and vision[C]. *Proc of the IEEE Computer Society Conf on Computer Vision and Pattern Recognition*. New York: IEEE Press, 2006: 260-267.
- [6] Avidan S. Ensemble tracking[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2007, 29(2): 261-271.
- [7] Matthews L, Ishikawa T, Baker S. The template update problem[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2004, 26(6): 810-815.
- [8] Yang M, Zhang C X, Wu Y W, et al. Robust object tracking via online multiple instance metric learning[C]. *Electronic Proc of the 2013 IEEE Int Conf on Multimedia and Expo Workshops*. San Jose: IEEE Press, 2013: 1-4.
- [9] Jiang N, Liu W Y, Wu Y. Learning adaptive metric for robust visual tracking[J]. *IEEE Trans on Image Processing*, 2011, 20(8): 2288-2300.
- [10] Li H X, Shen C H, Shi Q F. Real-time visual tracking using compressive sensing[C]. *Proc of the IEEE Computer Society Conf on Computer Vision and Pattern Recognition*. Colorado Springs: IEEE Press, 2011: 1305-1312.
- [11] Chen F, Wang Q, Wang S, et al. Object tracking via appearance modeling and sparse representation[J]. *Image and Vision Computing*, 2011, 29(11): 787-796.
- [12] Wang D, Lu H C, Yang M H. Online object tracking with sparse prototypes[J]. *IEEE Trans on Image Processing*, 2013, 22(1): 314-325.
- [13] Zhang Q, Li B X. Discriminative K -SVD for dictionary learning in face recognition[C]. *Proc of the IEEE Computer Society Conf on Computer Vision and Pattern Recognition*. San Francisco: IEEE Press, 2010: 2691-2698.
- [14] Mairal J L, Bach F, Ponce J, et al. Online dictionary learning for sparse coding[C]. *Proc of the 26th Int Conf on Machine Learning*. Montreal: IEEE Press, 2009: 689-696.
- [15] Zhang S P, Yao H X, Zhou H Y, et al. Robust visual tracking based on online learning sparse representation[J]. *Neurocomputing*, 2013, 100(1): 31-40.
- [16] Zhong W, Lu H C, Yang M H. Robust object tracking via sparsity-based collaborative model[C]. *Proc of the IEEE Computer Society Conf on Computer Vision and Pattern Recognition*. Providence: IEEE Press, 2012: 1838-1845.
- [17] Liu B Y, Yang L, Huang J Z, et al. Robust and fast collaborative tracking with two stage sparse optimization[C]. *Proc of 11th European Conference on Computer Vision*. Heraklion: IEEE Press, 2010: 624-637.
- [18] Wang Y R, Tang X T, Cui Q. Dynamic appearance model for particle filter based visual tracking[J]. *Pattern Recognition*, 2012, 45(12): 4510-4523.

(责任编辑: 孙艺红)