

子空间描述的关节式目标跟踪

赵运基^{1†}, 裴海龙²

(1. 河南理工大学 电气工程与自动化学院, 河南 焦作 454000;

2. 华南理工大学 自主系统与网络控制教育部重点实验室, 广东 广州 510640)

摘要: 针对关节式目标变化对子空间描述造成的影响, 本文提出了一种基于增量学习的关节式目标跟踪算法. 该算法通过引入图像分割方法与快速傅里叶变换可有效消除背景像素对目标描述造成的影响以及目标区域前景目标位置对不准造成的误差, 同时应用局部二值模式增加目标描述中像素点间的几何位置信息, 应用基于增量学习的方法实现目标特征的在线更新, 最终为跟踪算法提供较为精确的目标描述. 实验结果表明, 本文提出的关节式目标跟踪算法具有较好的目标跟踪效果.

关键词: 奇异值分解; 子空间描述; 粒子滤波; 局部二值模式

中图分类号: TP391 **文献标识码:** A

Articulate object tracking based on subspace representation

ZHAO Yun-ji^{1†}, PEI Hai-long²

(1. School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo Henan 454000, China;

2. Key Laboratory of Autonomous Systems and Networked Control, Ministry of Education, South China University of Technology, Guangzhou Guangdong 510640, China)

Abstract: Subspace representation of articulate object is subject to its variation. In order to deal with this problem, an articulate object tracking algorithm based on incremental learning is proposed. In this algorithm, graph-cut algorithm is used to reduce the influence of the background pixels on the object representation. Fast Fourier transform (FFT) is used to reduce the error of matching between the foreground locations in object regions. Local binary pattern (LBP) can increase spatial location information of pixels in the object representation. The object features updated online by incremental learning method can provide more accurate object representation. Experimental results demonstrate that the algorithm is able to track articulated objects with higher accuracy.

Key words: singular value decomposition; subspace representation; particle filter; local binary pattern

1 引言(Introduction)

目标跟踪作为计算机视觉中的一个研究领域一直是近年来的研究热点. 目标跟踪可应用于人机接口^[1-2]、运动识别、自动检测、交通监测、车辆导航^[3]和增强现实等领域^[4-5]. 目标跟踪是在后续图像帧中定位选定目标的过程. 尽管针对不同类型的目标提出了相应的目标描述方法和目标跟踪方法, 但是到目前为止仍然没有一种普遍适用的目标跟踪算法. 这主要源于对不同类型的目标随着位置、形状的变化很难得到对目标特征的统一描述, 同时目标的复杂背景等因素也是跟踪算法的一个挑战.

跟踪目标的特征描述直接影响跟踪算法的跟踪效果. 目标的特征描述有多种方法, 例如: 颜色直方图方法^[6-7]、空间颜色分布直方图^[8]、目标轮廓描述^[9]、方向梯度直方图描述^[10]、混合模型的直方图

描述^[11-12]和小波特征描述^[13-15]等. 目标的子空间描述方法因其计算简单、有效等特性在目标跟踪和人脸识别领域得到了广泛应用^[16-17], 其实质是将原始的高维样本数据投影到一个更有利于分类的低维特征子空间中. 文献[16-17]中, 作者提出了可应对目标形变和环境变化的子空间在线更新的目标跟踪算法. 该算法在跟踪过程中通过在线学习更新目标的特征子空间, 同时应用粒子滤波方法代替梯度下降法实现目标运动参数估计. 但是, 文献[16-17]中基于特征子空间的目标描述方法没有考虑各像素点间的位置关系, 同时在特征子空间更新过程中, 目标空间位置对不准带来的位移误差将直接影响目标的特征子空间描述. 局部二值模式^[18](local binary pattern, LBP)描述了像素点与其相邻点的亮度关系. 鉴于该方法在人脸识别领域中的成功应用, 一些研究者提

出了基于局部二值模式的小波特征描述^[19-21].

关节式目标的形状和外观变化较为剧烈, 因此, 如何选择合适的目标描述方法是实现目标稳定跟踪的关键. 早期的研究者们提出了主动轮廓模型^[22]对关节式目标进行描述, 并实现了对关节式目标的跟踪. 但是, 主动轮廓需要离线训练才能够得到, 跟踪过程中如果遇到训练过程中未出现过的目标时, 那么跟踪算法将不能实现目标的稳定跟踪^[23]. 水平集算法也应用到了关节式的目标跟踪中^[24]. 这些方法仅仅应用了目标的轮廓信息, 忽视了外观和丰富的纹理信息, 以及目标区域中各像素点间的几何位置关系. 虽然文献[25]中的目标直方图描述中引入了各像素点间的位置信息, 但是, 此类信息的作用毕竟相对有限.

在目标的跟踪过程中, 目标跟踪的模板多选择为视频帧中的矩形区域, 因此, 在关节式的目标跟踪中, 由于关节形变使得矩形的目标模板中包含了无效的背景像素. 无效的背景像素会对目标的描述造成一定的影响, 最终对目标跟踪的精度产生影响. 综上所述, 为了得到较为精确的目标描述信息, 本文应用图像分割算法对所选目标区域进行分割, 经过分割的目标区域中所包含的像素即为目标本身. 由于关节式目标的关节形变使得目标本身的像素点在目标矩形区域中存在一定不确定性, 以及二维图像快速傅里叶变换的振幅谱图像的位移不变性, 本文将经过图像分割后的目标图像转化为幅度谱, 对幅度谱图像应用局部二值模式方法进行处理, 得到局部二值模式图像, 该图像已经包含了各像素点与其相邻像素点间的亮度关系. 目标的傅里叶变换表示能够消除因空间位置对不准带来的目标子空间描述误差. 局部二值模式图像将各像素点与相邻像素点的相对亮度信息考虑其中, 增加了目标描述中的像素点的局部信息. 跟踪算法在粒子滤波框架下实现. 各个粒子区域中前景目标的二值模式图像与目标的基于子空间线性描述的线性差值作为粒子权重描述. 同时跟踪过程中实时更新特征子空间向量, 最终实现基于子空间描述的关节式目标跟踪. 通过实验比较证明该方法能够有效实现对关节式目标的有效跟踪.

算法初始时由于目标描述的子空间向量为空, 因此, 初始的 m 帧中目标跟踪采用模板匹配的方式, 各个粒子的权重系数由其与第1帧图像中的目标模板的欧式距离负数的幂函数描述. m 值的大小确定了目标子空间向量的更新速度, 而粒子数的选择对跟踪精度和计算复杂度同时产生影响, 因此, 在跟踪过程中需要做到跟踪精度和计算复杂度的平衡. 跟踪算法总体流程如图1所示.

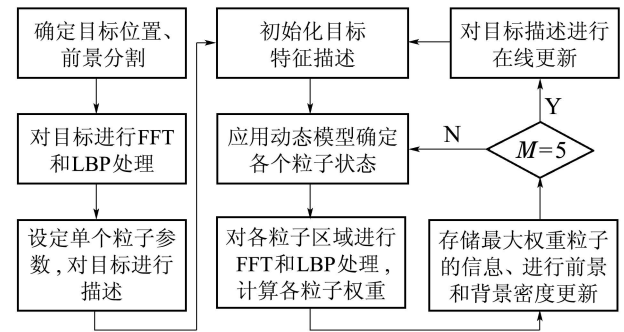
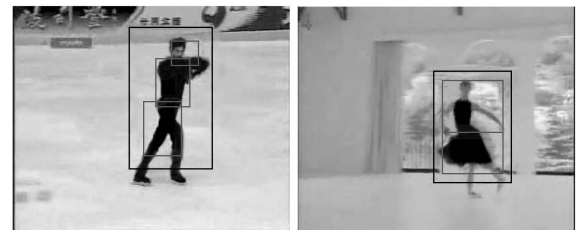


图1 本文算法的整体流程

Fig. 1 Flowchart of the algorithm proposed in this paper

2 前景目标分割(Foreground object segmentation)

关节式目标跟踪过程中, 无论是初始帧或是后续帧的目标区域中都会包含背景像素, 因此, 将包含背景像素的目标矩形作为训练样本进行目标特征子空间描述的基向量计算将存在较大误差, 甚至出现对跟踪目标的子空间描述基向量的漂移, 最终影响目标的跟踪效果. 因此, 为了得到关节式目标的准确描述, 在目标矩形区域应用图割法进行前景目标(只包含关节式目标区域)提取. 假设在 t 时刻得到了最优的跟踪目标矩形区域为 W_t , 视频序列选择为文献[25]中相同的视频序列. 图2所示为当 $t=4$ 时图像帧中跟踪到的目标区域. 当 $t=1$ 表示为初始化跟踪目标.



(a) Male skater

(b) Dancer

图2 $t=4$ 时的目标区域

Fig. 2 Object region at $t=4$

如图2所示, 在图2(a)目标窗口中选择3个较小的矩形窗, 在图2(b)中选择为两个较小的矩形区域, 以此作为前景目标的有效逼近. 此方法与文献[25]中相同. 鉴于积分直方图^[26]方法对于计算多个矩形区域直方图时的优越性, 在计算矩形区域直方图时采用了积分直方图的方法. 目标区域前景直方图 $H_t^{B_i}$ 由目标各矩形区域的直方图加权线性组合得到. $H_t^{B_i}$ 表示第 i 个矩形区域. 权重系数由式(1)得到:

$$\lambda_i = k \frac{\sum_{b=1}^N H_t^{B_i}(b)}{\sum_{p \in W_t} C(p)}, \quad (1)$$

其中: $\sum_{p \in W_t} C(p)$ 表示目标区域 W_t 中像素总数, k 为归一化常数, 以确保前景矩形区域的权重系数之和

为1. 参照文献[27–28], 图割算法即是求代价函数的最小值, 代价函数如式(2)所示:

$$E(C_p) = \mu \sum_{p \in \phi} R_p(C_p) + \sum_{(p,q) \in \beta: C_p \neq C_q} B_{p,q}, \quad (2)$$

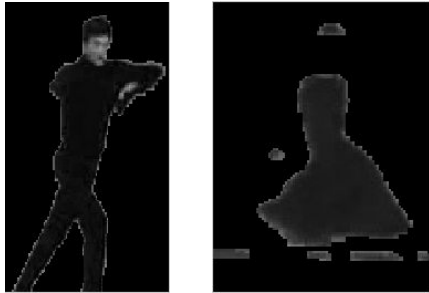
其中: p 表示目标区域 W_i 像素点集 ϕ 中单个像素点, P_B 表示背景密度, 由归一化的背景直方图获得, $p_i (1 \leq i \leq 3)$ 表示前景矩形 B_i 的密度, 由归一化的直方图 $H_i^{B_i}$ 获得, P_i 前景密度由归一化前景直方图 H_i^f 得到; 如果像素点 p 为前景像素点则 $C_p = 1$, 否则 $C_p = 0$. β 表示邻域像素点的集合. $R_p(C_p)$ 由式(3)给出:

$$\begin{cases} R_p(C_p = 0) = -\log P_F(I(p), p), \\ R_p(C_p = 1) = -\log P_B(I(p)). \end{cases} \quad (3)$$

如果 $p \in B_i$, $P_F(I(p)) = P_i(I(p))$, 否则, $P_F(I(p)) = P_i(I(p))$. μ 为权重系数, 与文献[25]采用相同值 $\mu = 0.5$, $B_{p,q}$ 定义如式(4)所示:

$$B_{p,q} \propto \frac{\exp((I(p) - I(q))^2 / 2\delta^2)}{\|p - q\|}, \quad (4)$$

其中 $I(p)$ 表示像素点 p 处的密度值. 整个前景目标分割过程在目标区域中完成, 虽然求取代价函数的最小值的运算过程具有较大的计算复杂度, 但是对整个目标跟踪系统没有造成较大的影响. 与图2相对应的跟踪目标经过图割法处理后的前景目标图像如图3所示.



(a) Male skater (b) Dancer

图3 与图2中相对应的前景目标提取结果

Fig. 3 Foreground object extractions corresponding to Fig.2

在目标前景分割过程中, 随着目标形状的变化, 目标区域的背景直方图和前景直方图也应当做相应的更新. 在新跟踪到的目标区域中根据仿射变换的参数更新目标矩形区域 B_i 的位置, 计算各区域的直方图, 同时依据式(1)可得各区域直方图的权重系数, 进而对前景直方图进行归一化处理得到前景密度 P_i . 计算目标区域的背景直方图并进行归一化处理得到背景密度 P_B . 因此, 经过上述处理过程可实现目标的前景密度和背景密度的更新.

在目标前景提取过程中, 初始帧中目标前景提取是在初始选定的区域中进行的, 在后续帧中为了降低计算负担将目标前景提取在包含所有粒子确定的区域中进行, 如图4所示的黑色区域即为进行前景目标提取的区域, 灰色区域为最终确定的目标区域, 白

色矩形区域为在上一帧目标区域的参数在布朗运动条件下所形成的候选目标区域.

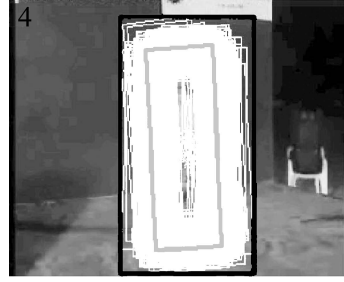


图4 前景分割的区域选择

Fig. 4 Region selection of the foreground segmentation

3 目标特征提取(Object feature extraction)

3.1 前景目标的傅里叶变换(Fast Fourier transform of foreground object)

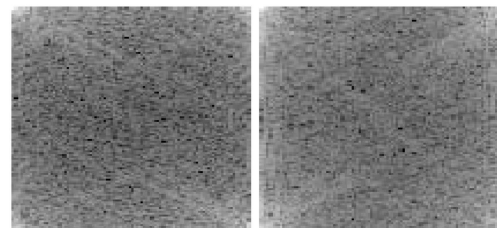
傅里叶变换是信号处理中有效的分析工具, 在图像处理领域也得到了广泛应用. 大小为 $M \times N$ 的二维图像以及经过平移了 (x_0, y_0) 的傅里叶二维变换可以表示为如式(5)所示:

$$\begin{aligned} F(u, v) &= \psi[f(x, y)] = \\ &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{j2\pi(ux/M + vy/N)}. \\ \psi[f(x - x_0, y - y_0)] &= F(u, v) e^{j2\pi(u_0x/M + v_0y/N)}. \end{aligned} \quad (5)$$

傅里叶变换具有如下的平移性质:

$$\begin{aligned} |\psi[f(x - x_0, y - y_0)]| &= \\ |F(u, v) e^{j2\pi(u_0x/M + v_0y/N)}| &= |F(u, v)|. \end{aligned} \quad (6)$$

由式(6)可以得出, 二维图像的傅里叶变换振幅谱具有位移不变性. 因此, 如果将经过前景目标分割所获得的矩形目标区域进行傅里叶变换, 那么关节式目标在矩形目标区域中的位置不确定性将能够得到有效的解决. 与采用矩形目标区域中所有像素点作为目标的基本描述的增量学习的目标跟踪算法^[17]相比较, 采用变换后的振幅图像作为跟踪目标的基本描述的方法具有更高的目标目标描述精度. 此种描述方法对提高目标跟踪算法的跟踪精度也是一种有效的改进. 与图3中前景目标相对应的傅里叶变换幅度谱如图5所示.



(a) Male skater (b) Dancer

图5 与图3相对应的目标区域的傅里叶变换幅度谱
Fig. 5 Amplitude spectrum of object region based on fast Fourier transform (FFT) corresponding to Fig.3

3.2 幅度谱的局部二值模式处理(Processing of amplitude spectrum based on LBP)

目标局部的外观信息相对于整体特征, 例如基于主元分析获得的整体特征等具有一定的优越性. 局部的特征信息对跟踪目标局部的光照变化具有一定的鲁棒性. 局部二值模式算子描述了与中心像素点相邻的像素点间的变化情况. 局部二值模式算子最早是由Timo Ojala等^[29]提出的, 其本质是一种有效的纹理描述算子. 在人脸识别领域中的应用最为广泛, 并且在局部二值模式思想的基础上提出了许多类似目标描述方法. 鉴于局部二值模式算子的计算复杂度相对较低同时也能够反映相对像素间的变换关系, 因此本文选择应用局部二值模式对目标区域傅里叶变换的幅度谱进行处理. LBP(local binary pattern)算子处理 3×3 邻域图像的过程如图6所示. 目标区域傅里叶变换后的幅度谱图像经过LBP算子处理后的目标区域图像如图7所示.

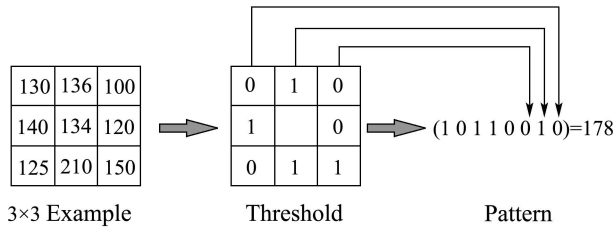


图 6 LBP在 3×3 区域中的计算过程

Fig. 6 LBP operator defined in 3×3 neighborhood

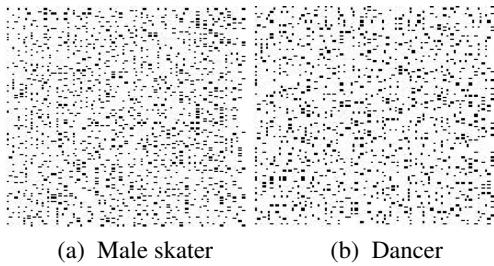


图 7 与图5中相对应的幅度谱的LBP处理结果图像
Fig. 7 The final amplitude spectrum image based on LBP operator corresponding to Fig.5

3.3 基于增量学习的目标特征提取(Object feature extraction by incremental learning)

文献[17]中应用了基于增量学习的目标特征子空间描述方法. 首先应用差值算法或者进行下样对初始跟踪目标区域进行处理, 依据文献[17]中的处理方式将目标区域规范为 32 像素点 \times 32 像素点大小的标准目标区域. 本文中即是类似于如图7所示的经过LBP算法处理的幅度谱图像规范化为 32 像素点 \times 32 像素点大小的目标区域. 此处理过程既可以保证提取的目标子空间基向量能够描述目标, 同时降低计算过程的计算负担. 对图7中各目标的LBP处理

结果的规范化如图8所示. 目标区域的原始大小分别为: Male skater(157 像素点 \times 89 像素点); Dancer(121 像素点 \times 91 像素点), 均为跟踪过程中第4帧图像中的目标区域大小, 本文中所显示的目标区域的大小是经过放大后的图像, 与实际目标区域大小有一定的比例关系. 同时在粒子滤波跟踪过程中每个粒子所确定的目标区域也进行规范化为 32 像素点 \times 32 像素点.

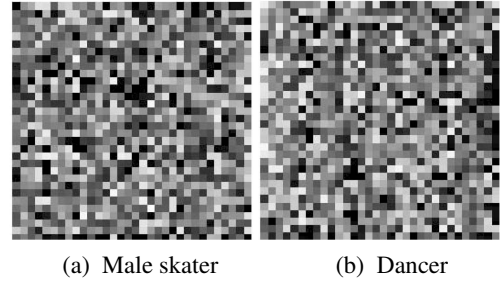


图 8 与图7中相对应的傅里叶变换幅度谱的LBP处理结果图像的规范化结果

Fig. 8 The normalized amplitude spectrum image based on LBP operator corresponding to Fig.7

基于增量学习的目标特征提取的主要依据参考文献[17]实现. 将LBP算法处理的结果图像(32 像素点 \times 32 像素点)转换为 1024×1 的列向量 I_0 作为目标的描述, 粒子滤波框架下的各个粒子确定的候选目标经过前景分割, 傅里叶变换和LBP处理得到相应的 $I_1^i (i = 1, \dots, N)$, N 为粒子滤波所选择的粒子数. 各个粒子的权重由 I_1^i 与 I_0 的欧式距离的指数表示成正比, 同时进行权重系数归一化. 最终确定目标的位置和目标的描述 I_1 , 当得到 T 个目标描述时, 此时进行目标描述的增量学习的更新. 文献[17]给出了增广误差矩阵的分块表示如式(7)所示:

$$[A \ B] = [U \ \tilde{B}] \begin{bmatrix} \Sigma & U^T B \\ 0 & \tilde{B}^T B \end{bmatrix} \begin{bmatrix} V^T & 0 \\ 0 & I \end{bmatrix} \Sigma =$$

$$([U \ \tilde{B}] \tilde{U}) \tilde{\Sigma} (\tilde{V}^T \begin{bmatrix} V^T & 0 \\ 0 & I \end{bmatrix}), \quad (7)$$

其中: $R = \tilde{U} \tilde{\Sigma} \tilde{V}^T$, \tilde{B} 为 B 中与 U 正交的向量, I 为单位矩阵, 令 $R = \begin{bmatrix} \Sigma & U^T B \\ 0 & \tilde{B}^T B \end{bmatrix}$, 依据文献[17]中的更新策略 \tilde{B} 和 \hat{B} 即可获得, 如式(8)–(9)所示:

$$\hat{B} = [(I_{T+t+1} - \bar{I}_B) \cdots (I_{T+t+m} - \bar{I}_B)$$

$$\sqrt{\frac{nm}{n+m}} (\bar{I}_B - \bar{I})], \quad (8)$$

$$\tilde{B} = \text{orth}(\hat{B} - UU^T \hat{B}), \quad (9)$$

则 $R = \begin{bmatrix} f \Sigma & U^T \hat{B} \\ 0 & \tilde{B} (\hat{B} - UU^T \hat{B}) \end{bmatrix}$, f 为遗忘因子, 遗忘因子的引入可以调整原始数据在特征向量计算中

所占的比重. 对 R 进行奇异值分解, 依据式(7)可得增广中心误差矩阵的奇异值分解后的特征矩阵和对应的特征向量分别为: $U' = [U \ B]\tilde{U}$, $\Sigma' = \tilde{\Sigma}$. 获得增广中心误差矩阵的特征值和特征向量后, 选择特征值较大的 L 个特征值对应的特征向量作为目标描述的基向量, 每个目标向量都可以由此 L 个基向量近似的线性表示, 在粒子滤波过程中, 各粒子所确定的目标向量与均值的误差为 $I_D^i = I_t^i - \bar{I}$, 该粒子的权重系数为 $w_i = \exp(-(I_D^i - H_i U_L'))$, 其中: H_i 为应用 U' 中最大的 L 个特征值所对应的特征向量对 I_D^i 进行线性近似的系数矩阵, U_L' 为对应较大的 L 个特征值所对应的特征向量所组成的矩阵. 选择权重系数较大的粒子所确定的目标即为所要跟踪的目标, 同时对所有粒子的权重进行归一化处理. 将以上过程循环处理即可实现目标特征向量的更新, 也即是基于增量学习的目标特征提取.

4 实验结果与分析(Experimental results and analysis)

整个算法在MATLAB环境下实现, 其中包含了一些经过优化的Mex C++子程序. 为了与文献[25]中提出的方法进行比较, 本文采用了文献[25]所提供的实验视频. 同时本文实现的目标跟踪还将与文献[17]中的跟踪结果进行比较. 所有实验均在2.0 GHz的PC上实现.

实验1采用文献[25]中应用的Male skalter视频进行关节式目标跟踪结果比较. 本实验选取400帧图像进行跟踪实验. 在实验过程中, 需要初始化的参数变量有: 中值误差矩阵的特征向量, 也即是目标描述的基向量更新的速度参数 $m = 5$, 表示每经过5帧更新一次目标描述的基向量; 动态模型中, 状态变量的每个参数相对于前一时刻的状态符合布朗运动规律, 各个仿射变换参数变化的方差设定为: $\delta_x^2 = 9$, $\delta_y^2 = 9$, $\delta_\theta^2 = 0.05$, $\delta_s^2 = 0.05$, $\delta_\alpha^2 = 0.005$, $\delta_\varphi^2 = 0.001$; 遗忘因子在所有的实验过程中全部设定为 $f = 0.95$; 粒子滤波跟踪过程中粒子数设定为 $n = 400$. 应用的3种不同的算法对关节式目标跟踪的结果如图9所示. 3种不同算法的关节式目标跟踪结果的目标中心位置与目标实际的中心位置(目标的中心位置选择为包含目标躯干部分的最小矩形的中心)的欧氏距离作为衡量目标跟踪效果的度量标准, 由图10可以直观的看出: 本文提出的关节式目标跟踪算法的误差曲线基本位于其他两种算法以下, 结合图9与图10可以得出文献[17]中提出的跟踪算法在第350帧左右时, 跟踪误差已经超过50个像素点, 跟踪目标已经丢失. 与图10中(a)-(c)对应于的3种不同算法的关节式目标跟踪结果误差的均值和方差分别为: $\mu_a = 28.8$, $\delta_a = 20.2$, $\mu_b = 17.8$, $\delta_b = 11.4$, $\mu_c = 11.1$, δ_c

$= 9.5$. 由跟踪误差均值和方差的比较可得出, 文献[25]的方法优于文献[17], 本文提出的关节式目标跟踪算法在本视频中的跟踪效果优于其他两种算法.

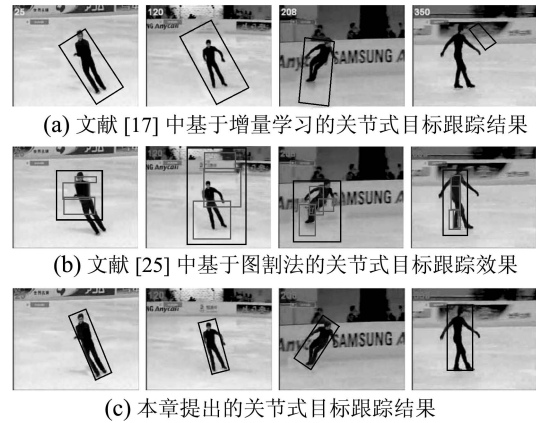


图9 在视频Male skalter中本文提出的跟踪算法与文献[17,25]的跟踪结果

Fig. 9 Tracking effects of the methods [17, 25] and the method proposed in this article in Male skater sequence

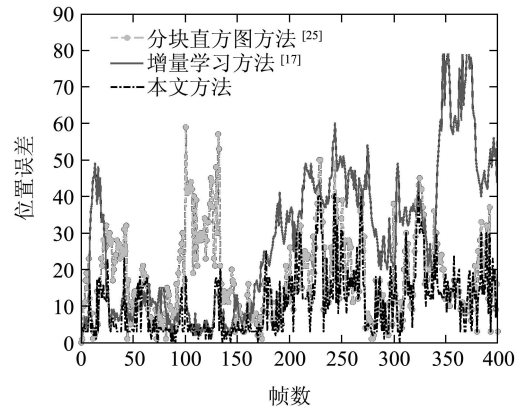
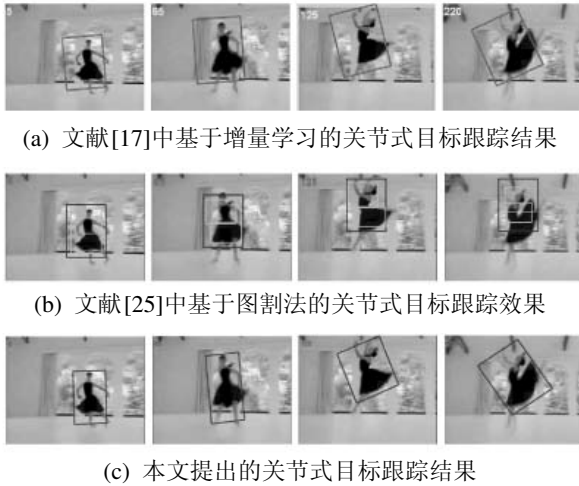


图10 在视频Male skalter中本文提出的跟踪算法与文献[17,25]的跟踪算法的位置误差曲线

Fig. 10 Distance error of the tracking methods [17, 25] and the tracking algorithm proposed in this article in Male skater sequence

实验2采用文献[25]中应用的Dancer视频进行关节式目标跟踪结果比较, 此视频序列共225帧. 跟踪过程的参数设定与实验一中的参数均相同. 目标跟踪结果如图11所示, 跟踪结果的中心位置误差曲线如图12所示. 图11表明, 3种跟踪算法基本都能够实现对目标的跟踪, 但是通过目标区域的背景像素点数的比较可以很直观的看出, 本文方法和文献[25]中的方法的目标矩形区域中包含了相对较少的背景像素点数, 但是文献[25]中算法的跟踪结果目标矩形区域之外存在较多的前景像素点数. 在图12中, 相对应与图11中的(a)-(c)3种不同跟踪算法的误差曲线的均值和方差分别为: $\mu_a = 25.7$, $\delta_a = 7.5$, $\mu_b = 25.2$, $\delta_b = 9.2$, $\mu_c = 16.7$, $\delta_c = 7.3$. 通过跟踪误差

的均值和方差的比较可以得出: 文献[17]和文献[25]两种算法的跟踪误差曲线的均值基本相似, 但是文献[17]的跟踪误差波动相对较小, 因此, 相对于文献[25]具有一定的优越性, 本文提出算法的跟踪误差曲线的均值和方差均小于文献[17]和文献[25]中的算法, 因此在跟踪过程中具有相对的优越性能。



(a) 文献[17]中基于增量学习的关节式目标跟踪结果
(b) 文献[25]中基于图割法的关节式目标跟踪效果
(c) 本文提出的关节式目标跟踪结果

图 11 在视频Dancer中本文提出的跟踪算法与文献[17,25]的跟踪算法的跟踪结果
Fig. 11 Tracking effects of the methods [17,25] and the method proposed in this article in Dancer sequence

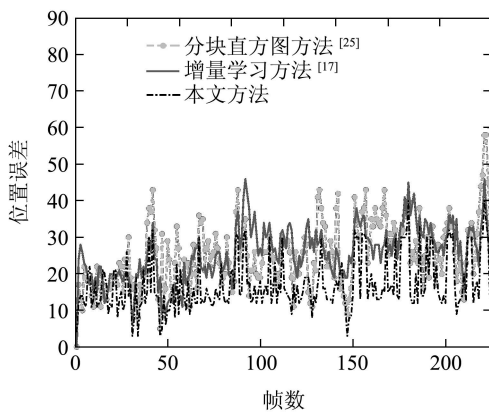


图 12 在视频Dancer中本文提出的跟踪算法与文献[17,25]的跟踪算法的位置误差曲线
Fig. 12 Distance error of the tracking methods [17,25] and the tracking algorithm proposed in this article in Dancer sequence

通过上述两个实验表明, 本文提出的关节式目标跟踪算法相对于文献[17]、文献[25]能够提供更为精确的目标位置. 同时通过视频序列变化的比较表明, 视频序列Male skater中目标的外观变化速度最快, 因此在应用3种不同算法进行目标跟踪过程中, 应用文献[17]中的方法时, 选择参数 $m = 5$ 作为向量子空间描述更新的参数时, 在跟踪过程中由于目标的子空间描述更新较慢直接导致跟踪目标丢失. 更新参数设定较小时, 在跟踪过程中的奇异值分解计算次数

相应的增加, 最终影响跟踪效率. 3种算法的计算复杂度可以通过相同的硬件和软件环境条件下的帧率表明. 3种算法在本文所选取的视频序列中的跟踪时间消耗与跟踪过程中的图像处理效率如表1所示, 其中时间单位为s, 帧率的单位为帧/s. 由帧率的比较结果可以得出, 本文所提出的算法在计算复杂度方面低于文献[25]中的方法, 但是高于文献[17](源于在基于增量学习的基础上增加了目标的前景提取与快速傅里叶变换与LBP处理等)的方法. 结合实验结果与计算复杂度的分析可得出, 本文提出的算法能够实现关节式目标的有效跟踪, 同时相对于文献[25]中的方法在跟踪效率上上了一定的改进, 虽然相对于文献[17]中的方法计算复杂度较大, 但是在关节式目标的跟踪过程中具有更好的稳定性.

表 1 跟踪过程中的时间消耗与跟踪算法处理效率
Table 1 Treatment efficiency and time consumption of the tracking algorithm in tracking process

	文献[17]算法		文献[25]算法		本文算法	
	用时	帧率	用时	帧率	用时	帧率
Male skater	63.109	6.3	207.242	1.9	172.824	2.3
Dancer	33.534	6.7	95.735	2.4	55.741	4.0

本文算法的影响跟踪实时性的主要因素有: 粒子数目、目标前景提取过程、LBP的计算以及增量学习的更新速度. 在降低计算复杂度方面应该考虑如图8所示的目标可能出现的区域进行LBP计算, 此过程可以避免重叠的候选目标区域的计算消耗. 粒子数目的选择会对跟踪精度产生一定的影响, 同时较大的粒子数目明显增加计算复杂度, 因此在粒子数目选择问题上应该考虑跟踪精度与计算复杂度的平衡问题. 增量学习的更新速度直接影响目标描述的准确性, 较大的更新速度增加计算复杂度, 较小的更新速度不能够应对较快的目标变换, 因此在后续的研究中应该考虑加入自适应的更新速度设置.

5 总结(Conclusions)

本文利用文献[17]的工作, 将文献[17]中基于增量学习的目标跟踪算法扩展到了关节式目标跟踪领域, 同时为了增加目标描述的准确性引入了图像分割算法和傅里叶变换, 此方法可有效降低背景像素对目标描述的影响以及目标空间位置对不准所产生的影响. 局部二值模式的引入为基于向量的目标描述增加了目标像素点间的空间位置信息, 提高了目标瞄的精确度. 跟踪实验结果表明, 本文的算法能够有效的实现关节式目标的有效跟踪. 本文的不足之处在于, 目标前景分割算法对整个算法的跟踪实时性影响较大, 增量学习的更新参数 T 确定了增量学习的更新速度, 固定的更新速度不能准确及时的描

述变化的目标. 后续研究中将考虑增强跟踪算法的实时性, 同时考虑将更新参数 T 进行在线更新.

致谢 感谢评审专家对本文提出的具有建设性的意见和建议, 同时感谢文献[17, 25]作者所提供的目标跟踪源程序, 为本文的软件编程实现提供了参考和跟踪结果比较的实现提供可能.

参考文献(References):

- [1] LEONID S, ALEXANDRU O B, MICHAEL J. Humaneva: synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion [J]. *International Journal of Computer Vision*, 2010, 87(1/2): 4 – 27.
- [2] BULLING A, GELLERSEN H. Toward mobile eye-based human-computer interaction [J]. *IEEE Pervasive Computing*, 2010, 9(4): 8 – 12.
- [3] YILMAZ A, JAVED O, SHAH M. Object tracking: a survey [J]. *ACM Computer Surveys*, 2006, 38(4): 229 – 240.
- [4] HENDERSON S, FEINER S. Exploring the benefits of augmented reality documentation for maintenance and repair [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2011, 17(10): 1355 – 1368.
- [5] ROBERT O C, GEORG K, DAVID W M. Wide-area augmented reality using camera tracking and mapping in multiple regions [J]. *Computer Vision and Image Understanding*, 2011, 115(6): 854 – 867.
- [6] BEYAN C, TEMIZEL A. Adaptive mean-shift for automated multi object tracking [J]. *IET Computer Vision*, 2012, 6(1): 1 – 12.
- [7] LI P. An adaptive binning color model for mean shift tracking [J]. *IEEE Transaction on Circuits System and Video Technology*, 2008, 18(9): 1293 – 1299.
- [8] BIRCHFIELD S T, SRIRAM RANGARAJAN. Spatiograms versus histograms for region-based tracking[C] // *Computer Vision and Pattern Recognition*. San Diego, CA, USA: IEEE, 2005: 1158 – 1163.
- [9] THI Y, VASWANI N, TANNENBAUM A, et al. Tracking deforming objects using particle filtering for geometric active contours [J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2007, 29(8): 1470 – 1475.
- [10] HAN Z J, YE Q X, JIAO J B. Combined feature evaluation for adaptive visual object tracking [J]. *Computer Vision and Image Understanding*, 2011, 115(1): 69 – 81.
- [11] WANG Y, TAN Y H, TIAN J W. Adaptive hybrid likelihood model for visual tracking based on Gaussian particle filter [J]. *Optical Engineering*, 2010, 47(9): 1 – 8.
- [12] ZHANG B, TIAN W F, JIN Z H. Efficient hybrid appearance model for object tracking with occlusion handling [J]. *Optical Engineering*, 2007, 46(8): 1 – 11.
- [13] XU Y, YANG X, LING H B, et al. A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid [C] // *IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco: IEEE Computer Society, 2010: 161 – 168.
- [14] NACEREDDINE N, TABBONE S, ZIOU D, et al. Shape-based image retrieval using a new descriptor based on the radon and wavelet transforms [C] // *International Conference on Pattern Recognition*. Istanbul, Turkey: IEEE, 2010: 1997 – 2000.
- [15] SU Y, SHAN S G, CHEN X L, et al. Hierarchical ensemble of global and local classifiers for face recognition [J]. *IEEE Transactions on Image Processing*, 2009, 18(8): 1885 – 1896.
- [16] LIM J, ROSS D, LIN R, et al. Incremental learning for visual tracking [M] // SAUL L, WEISS Y, BOTTOU L. *Advance in Neural Information Processing Systems*. Cambridge: MIT Press, 2005: 793 – 800.
- [17] ROSS D, LIM J, LIN R, et al. Incremental learning for robust visual tracking [J]. *International Journal of Computer Vision*, 2008, 77(1): 125 – 141.
- [18] AHONEN T, HADID A, PIETIKAINEN M. Face decription with local binary patterns: Application to face recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(12): 2037 – 2041.
- [19] LIU L, SUN Y F, YIN B C, et al. Local gabor binary pattern random subspace method for eyeglasses-face recognition [C] // *International Congress on Image and Signal Processing*. Yantai, China: IEEE, 2010: 1892 – 1896.
- [20] ZHANG W, SHAN S, CHEN X, et al. Are Gabor phases really useless for face recognition [C] // *International Conference on Pattern Recognition*. Hong Kong: Springerlink, 2006: 606 – 609.
- [21] TAN X Y, TRIGGS B. Enhanced local texture feature sets for face recognition under difficult lighting conditions [J]. *IEEE Transaction on Image Processing*, 2010, 19(6): 1635 – 1650.
- [22] KASS M, WITKIN A, TERZOPOULOS D. Snakes: active contour models [J]. *International Journal of Computer Vision*, 1988, 1(4): 321 – 331.
- [23] TOYAMA K, BLAKE A. Probabilistic tracking in a metric space [C] // *Proceedings of the International Conference on Computer Vision*. Vancouver, Canada: IEEE, 2001: 50 – 59.
- [24] FREEDMAN D, ZHANG T. Active contours for tracking distributions and shape prior [J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 518 – 526.
- [25] SHALED NEJHUM S M, JEFFREY H, YANG M H. Online visual tracking with histogram and articulating blocks [J]. *Computer Vision and Image Understanding*, 2010, 114(8): 901 – 914.
- [26] PORIKLI F. Integral histogram: a fast way to extract histograms in Cartesian spaces [C] // *International Conference on Computer Vision and Pattern Recognition*. San Diego: IEEE, 2005: 829 – 836.
- [27] FREEDMAN D, ZHANG T. Interactive graph cut based segmentation with shape priors [C] // *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. San Diego: IEEE, 2005: 755 – 762.
- [28] BOYKOV Y, JOLLY M P. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images [C] // *Proceedings of the International Conference on Computer Vision*. Vancouver, Canada: IEEE, 2001, 1: 105 – 112.
- [29] TIMO O, MATTI P, DAVID H. A comparative study of texture measures with classification based on featured distributions [J]. *Pattern Recognition*, 1996, 29(1): 51 – 59.

作者简介:

赵运基 (1980–), 男, 博士研究生, 主要从事目标跟踪与视觉导航的研究, E-mail: auyjz@hpu.edu.cn;

裴海龙 (1965–), 男, 教授, 主要从事机器人控制、嵌入式系统、数控系统和神经网络的研究, E-mail: auhlpei@scut.edu.cn.