

Journal of Image  
and Graphics

# 中国图象图形学报



ISSN1006-8961  
CN11-3758/TB

2013  
Vol.18 No.

1

中国科学院遥感应用研究所  
中国图象图形学学会主办  
北京应用物理与计算数学研究所

# 中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2013年1月 第18卷 第1期(总第201期)

## 目次

### 综述

- 正面人脸图像合成方法综述 ..... 赵林, 高新波, 田春娜(1)  
SAR 图像道路网提取方法综述 ..... 程江华, 高贵, 库锡树, 孙即祥(11)

### 图像处理和编码

- 头部缺失的 JPEG 文件碎片恢复 ..... 徐明, 黄立, 张海平, 徐建, 郑宁(24)  
安全高效的可撤销指纹模板构造 ..... 喻建平, 张鹏, 王瑶, 杨懿竣(36)  
基于残差的图像超分辨率重建 ..... 陈华华, 姜宝林, 刘超, 陈伟强, 陆宇, 张嵩(42)  
旋转的 Wang Tiles 纹理合成算法 ..... 王继东, 庞明勇, 赵瑞斌(49)  
基于圆形约束快速水平集的原生质体细胞分割 ..... 王晓飞, 庞全(55)

### 图像分析和识别

- 采用压缩传感的鲁棒的视频指纹方案 ..... 孙锐, 李超, 蒋飞云(62)  
基于图像显著性的路面裂缝检测 ..... 徐威, 唐振民, 吕建勇(69)  
基于局部熵的主动轮廓模型 ..... 潘改, 高立群, 赵爽(78)  
基于算子的图像分解 ..... 李峰, 曾晓辉, 陈盛霞, 沈玉娟(86)

### 图像理解和计算机视觉

- 中值流辅助在线多示例目标跟踪 ..... 王德建, 张荣, 尹东, 张智瑞(93)  
四叉树直方图的特殊方向关系表达 ..... 张珂, 王小捷, 靳越(101)

### 计算机图形学

- 协同进化的近似规则纹理合成 ..... 王相海, 陶兢喆(107)  
反走样直线的灰度循环生成算法 ..... 牛连强, 张胜男, 钟玲(115)

### 地理信息技术

- 微博客蕴含交通信息的提取 ..... 张恒才, 陆锋, 陈洁(123)

- 
- “计算机视觉前沿论坛”专栏征文通知 ..... (130)

# Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 18 No. 1 January 2013

## Contents

### Review

- Review of frontal face image synthesis methods ..... Zhao Lin, Gao Xinbo, Tian Chunna(1)  
Review of road network extraction from SAR images ..... Cheng Jianguhua, Gao Gui, Ku Xishu, Sun Jixiang(11)

### Image Processing and Coding

- Rrecovery method for JPEG file fragments with missing headers  
..... Xu Ming, Huang Li, Zhang Haiping, Xu Jiang, Zheng Ning(24)  
Secure and efficient scheme to construct a cancelable fingerprint template  
..... Yu Jianping, Zhang Peng, Wang Yao, Yang Yijun(36)  
Image super-resolution reconstruction based on residual error  
..... Chen Huahua, Jiang Baolin, Liu Chao, Chen Weiqiang, Lu Yu, Zhang Song(42)  
Texture synthesis using rotational Wang Tiles ..... Wang Jidong, Pang Mingyong, Zhao Ruibin(49)  
Protoplasm somatic cells segmentation based on circle dependent fast level-set segmentation ..... Wang Xiaofei, Pang Quan(55)

### Image Analysis and Recognition

- Robust video fingerprinting via compressed sensing ..... Sun rui, Li Chao, Jiang Feiyun(62)  
Pavement crack detection based on image saliency ..... Xu Wei, Tang Zhenmin, Lv Jianyong(69)  
Active contour model driven by local entropy energy ..... Pan Gai, Gao Liqun, Zhao Shuang(78)  
Operator-based image decomposition ..... Li Feng, Zeng Xiaohui, Chen Shengxia, Shen Yujuan(86)

### Image Understanding and Computer Vision

- Median flow aided online multi-instance learning visual tracking ..... Wang Dejian, Zhang Rong, Yin Dong, Zhang Zhirui(93)  
Expression of special directional relation based on quadtree histogram ..... Zhang Ke, Wang Xiaojie, Jin Yue(101)

### Computer Graphics

- Fast near-regular texture synthesis based on the concept of co-evolution ..... Wang Xianghai, Tao Jingzhe(107)  
Integral algorithm for generating anti-aliased straight line controlled by gray iteration  
..... Niu Lianqiang, Zhang Shengnan, Zhong Ling(115)

### Geoinformatics

- Extracting traffic information from massive micro-blog messages ..... Zhang Hengcai, Lu Feng, Chen Jie(123)

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2013)01-0093-08

论文引用格式: 王德建, 张荣, 尹东, 张智瑞. 中值流辅助在线多示例目标跟踪[J]. 中国图象图形学报, 2013, 18(1): 93-100.

## 中值流辅助在线多示例目标跟踪

王德建, 张荣, 尹东, 张智瑞

中国科学技术大学电子工程与信息科学系, 合肥 230027

**摘要:** 针对机器人演示学习中目标跟踪性能的严格要求, 提出一种可以有效克服快速运动、遮挡和目标漂移的物体跟踪方法。首先计算中值流, 并预测目标的位置偏移, 以此计算高斯权重; 然后修正搜索区域, 并使用在线多示例分类器进行目标搜索, 计算似然度; 最后使用贝叶斯框架对结果进行融合, 使用穷举搜索得到最优的预测位置, 并更新在线分类器。实验结果表明, 与现有方法相比, 该方法对快速运动和目标漂移具有更强的鲁棒性, 而且可以达到实时跟踪。

**关键词:** 服务机器人; 演示学习; 目标跟踪; 在线多示例学习; 中值流

### Median flow aided online multi-instance learning visual tracking

Wang Dejian, Zhang Rong, Yin Dong, Zhang Zhirui

Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China

**Abstract:** To satisfy the stringent requirements of the object tracking performance in the robot's learning-from-demonstration-framework, a new tracking algorithm that can deal with fast motions, occlusions, and drifts, is proposed. First, the Median-Flow method is used to predict the position-shift of the object and the Gaussian weight of each patch. Then, the search-region is modified and the object is located by the online multi-instance learning classifier. Afterwards, the likelihood of each patch is calculated. Finally, the results are combined under the Bayes framework to get the best prediction by exhaustive search and the online classifier is updated. Experiments in several commonly used test videos show that our method outperforms the other state-of-the-art tracking methods, especially for fast motion and drifts. Furthermore, the proposed method can run in real-time.

**Key words:** service robots; learning from demonstration; object tracking; online multi-instance learning; median flow

## 0 引言

随着人口老龄化问题日益严重, 家庭服务机器人将扮演越来越重要的角色。演示学习<sup>[1]</sup> (LD) 是服务机器人自主学习的一种重要方法, 指的是机器人通过观察演示者的动作抽取知识, 完成自主学习的一种学习方式, 是一种“所见即所得”的创新学习方法。与传统的学习方法相比, 演示学习成本低, 操

作简单, 特别适用于家庭服务机器人<sup>[2]</sup>。目前演示学习已经成为了人工智能领域的一个研究热点。但是与工业环境不同, 家庭环境复杂多变, 任务范围广, 因此对于视觉算法的要求更高。家庭环境下机器人的目标跟踪常常遇到的问题主要有遮挡, 快速运动, 光照变化等。如何设计一个鲁棒的目标跟踪算法, 有效地克服以上问题成为演示学习中的一个核心的问题。

目标跟踪在计算机视觉中有非常广泛的应用,

收稿日期: 2012-07-03; 修回日期: 2012-08-29

基金项目: 国家重点基础研究发展计划(973) 基金项目(2010CB327900)

第一作者简介: 王德建(1989—), 男, 中国科学技术大学信号与信息处理专业硕士研究生, 主要研究方向为计算机视觉与模式识别。

E-mail: wdejian@mail.ustc.edu.cn

比如视频监控、机器人视觉、人机交互、车辆跟踪等<sup>[3]</sup>。目标跟踪领域面临着许多的挑战,主要是由于噪声、遮挡、视角变化、光照变化、剧烈运动等因素引起的<sup>[4]</sup>。基于分类的目标跟踪(tracking by classification)是近几年的研究热点。其思想就是将目标跟踪的问题当成分类问题。目前这类跟踪方法已经取得了许多的有效的成果。Avidan<sup>[5]</sup>使用 Adaboost 对每个像素进行分类,能够处理多种视频(灰度、红外等),并且可以处理遮挡对跟踪的影响。Collins 等人<sup>[6]</sup>通过利用在线的特征选择方法进行目标跟踪,可以自适应地解决物体姿态变化的问题。Grabner 等人<sup>[7]</sup>使用 Oza<sup>[8]</sup>提出的在线 boosting 方法,可以较好地处理光照变化、视角变化下的目标跟踪问题。

Babenko<sup>[9-10]</sup>等人认为目标偏移是由于训练样本的潜在的歧义性,使用多示例学习(MIL)可以有效地避开这种歧义,因此设计出基于在线多示例学习的跟踪器,对于部分遮挡和抵抗物体漂移都有很强的鲁棒性,而且可以实时跟踪,是目前这类方法中最有效的方法之一。该方法将应用到机器人演示学习中,由于家庭环境下物体经常发生快速运动,该方法对于这类剧烈运动目标的跟踪效果并不理想。这类基于分类的方法可以很好地处理遮挡、漂移等问题,然而由于其运动模型过于朴素,一般简单地假设帧间物体运动不会超过某个范围,而且出于对计算效率的考虑,一般搜索区域比较小,因此这类方法对于剧烈运动常常没有很强的跟踪能力,在低帧率视频以及目标快速运动的视频中跟踪效果不理想。Kalal<sup>[11]</sup>使用中值流对目标位置和尺度进行估计,对于运动物体的跟踪效果比较好。

为了解决演示学习中的目标跟踪存在的问题,提出一种中值流预测下的在线多示例目标跟踪方法,主要思想是在线多示例学习目标跟踪方法中引入中值流作为运动模型,对物体的位置先大致地预测,然后再进行分类和检测,最后使用贝叶斯融合方法将两种方法融合起来,并且进行整体模型的调整和优化,从而使得算法不仅能够处理物体遮挡、漂移和姿态变化,对剧烈运动还具有一定的鲁棒性,可以满足机器人演示学习对目标跟踪算法的要求。

## 1 中值流目标跟踪

Kalal 等人<sup>[11]</sup>提出的中值流跟踪本质上是一种点跟踪的方法。Kalal 认为一个有效的跟踪与时间

流的走向无关,即正向跟踪与反向跟踪效果应该一致。例如,给定  $t$  时刻的初始点  $P_t$ ,可以得到  $t$  到  $t+k$  时刻正向跟踪点序列  $\{P_t, P_{t+1}, \dots, P_{t+k}\}$ ,若使用  $t+k$  时刻的  $P_{t+k}$  作为初始点,得到相应的  $t+k$  到  $t$  的反向跟踪点序列  $\{\hat{P}_{t+k}, \hat{P}_{t+k-1}, \dots, \hat{P}_t\}$ ,如果轨迹间的差异很大就说明跟踪失败,否则说明跟踪有效。假设前向后向两次跟踪轨迹为  $T_{fw}^k = \{P_t, P_{t+1}, \dots, P_{t+k}\}$  和  $T_{bw}^k = \{\hat{P}_t, \hat{P}_{t+1}, \dots, \hat{P}_{t+k}\}$ ,则前向-后向误差定义为

$$FBE(T_{fw}^k) = d(T_{fw}^k, T_{bw}^k) \quad (1)$$

式中,  $d(T_{fw}^k, T_{bw}^k) = \|P_t - \hat{P}_t\|$ 。

基于中值流的跟踪过程如图 1 所示。给定  $t$  时刻的跟踪目标位置,首先初始化一些跟踪点,然后使用点跟踪方法得到  $t+1$  时刻这些跟踪点的预测位置并且计算相应的前向-后向误差(FBE)。使用 Lucas-Kanade 光流<sup>[12]</sup>方法,它是一种非常优秀的点跟踪方法,在计算机视觉中得到广泛的应用。然后去除 50% 误差值高的跟踪点。余下的 50% 跟踪点被认为是好的跟踪点,被用于计算  $t+1$  时刻的目标的位置及尺度信息。目标位置偏移  $[dx, dy]^T$  取余下的 50% 好的跟踪点中 FBE 中值对应的偏移,而对于尺度的估计,每个好的跟踪点两两组合成一个点对,计算距离,并且计算出与  $t$  时刻相应点对距离的比值作为当前点对的估计尺度  $\tilde{s}_{ij} = (\|\hat{P}_{t+1,i} - \hat{P}_{t+1,j}\|_2) / (\|P_{t,i} - P_{t,j}\|_2), i, j = 1, 2, \dots, n; i \neq j$ ,将  $\{\tilde{s}_{ij}\}$  排序,选择其中值作为预测的尺度。这样的跟踪过程称为中值流跟踪(median flow)。相应地跟踪器称为中值流跟踪器(median flow tracker)。

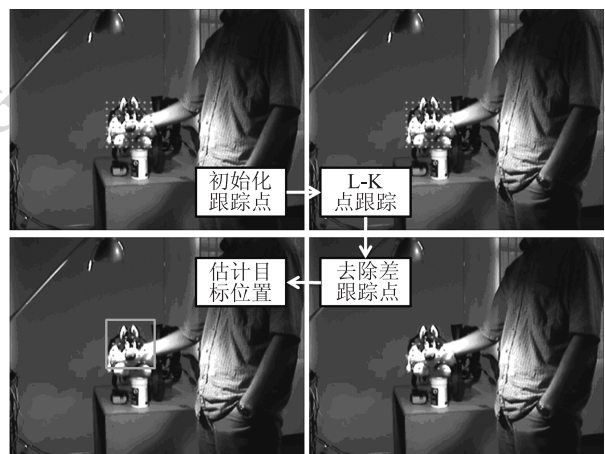


图 1 中值流跟踪示意图

Fig. 1 Chart of median flow tracker

## 2 在线多示例学习目标跟踪

### 2.1 多示例学习

传统的分类器训练方法需要一个训练集  $\{(x_1, y_1), \dots, (x_n, y_n)\}$ , 式中  $x_i$  是一个示例,  $y_i \in \{0, 1\}$  是样本标签, 而在多示例学习中, 训练集变为  $\{(X_1, y_1), \dots, (X_n, y_n)\}$ , 式中  $X_i = \{x_{i1}, \dots, x_{im}\}$  是由示例组成的一个包,  $y_i$  是包标签, 即  $y_i = \max_j (y_{ij})$ , 式中  $y_{ij}$  是示例  $x_{ij}$  对应的标签。

Viola 等人<sup>[13]</sup>提出了 MILBoost 算法, 通过使用梯度 boosting 的方法最大化包的似然率来训练一个示例分类器。这样就可以估计出示例的似然率  $p(y|x)$ , 然后使用 Noisy-OR 准则得到包的似然率

$$p(y_i | X_i) = 1 - \prod_j (1 - p(y_i | x_{ij})) \quad (2)$$

Noisy-OR 准则使得当包中存在一个似然度很高的示例时, 包的似然度也很高。

### 2.2 在线多示例学习

Babenko 综合了多示例学习<sup>[13]</sup>和在线 boosting 方法<sup>[7]</sup>, 提出了在线多示例学习, 用于解决目标跟踪中样本的歧义性问题, 对于跟踪过程中的目标漂移和遮挡都获得很好的效果。

基于在线多示例学习的跟踪首先在  $t$  时刻目标附近  $r$  像素范围截取正样本, 在  $r$  到  $\beta$  像素范围内截取一定数量的负样本 ( $r < \beta$ ), 分别打成正包和负包, 利用在线多示例学习方法更新分类器。当新的一帧来临时, 用分类器对  $t$  时刻目标位置附近  $s$  像素范围内的图像块进行评价, 得到似然度, 选择似然度最高的作为目标位置, 然后更新分类器。

在视频的每一帧中, 将得到一组训练数据  $\{(X_1, y_1), \dots, (X_n, y_n)\}$ , 示例的似然度通过  $p(y|x) = \sigma(H(x))$  计算, 式中  $\sigma(x) = 1/(1 + e^{-x})$ ,  $H(x) = \sum_{k=1}^K \alpha_k h_k(x)$ ,  $\alpha_k$  是弱分类器  $h_k(x)$  对应的权重。包的似然度可以通过 Noisy-OR 准则 (式(2)) 获得, 建立一个含有  $M$  个弱分类器的分类器池, 从中选出最优的  $K$  个弱分类器进行组合, 并用包的标签来近似示例的标签, 从而得到

$$h_k = \arg \max_{h \in \{h_1, \dots, h_M\}} L(H_{k-1} + h) \quad (3)$$

式中,  $L = \sum_i (y_i \log(p_i) + (1 - y_i) \log(1 - p_i))$  是强分类似然度函数,  $p_i$  可以由式(2)计算,  $\{h_1, \dots,$

$h_M\}$  是弱分类器池,  $h_k$  是第  $k$  轮通过贪心法选出的弱分类器。

## 3 基于中值流的多示例目标跟踪

一个典型的目标跟踪系统主要由 3 个部分组成, 如图 2 所示。

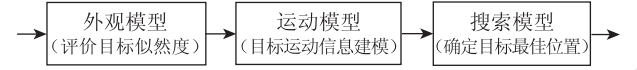


图2 典型物体跟踪系统组成

Fig. 2 Typical tracking system components

外观模型主要对物体的外观建模, 用于评价目标的外观似然度, 运动模型主要对物体的运动信息建模, 用于预测物体的运动并指导搜索模型搜索目标。搜索模型主要负责确定目标的最佳位置。使用在线多示例分类器作为外观模型, 中值流作为运动模型, 并利用中值流的预测信息指导目标的搜索。一般来说搜索模型可以使用 Mean-Shift 等, 但是本文研究的侧重点在于前两个模型的融合的有效性, 所以仅使用简单的穷举搜索以提高搜索性能。

### 3.1 贝叶斯融合

贝叶斯融合框架广泛运用于目标跟踪领域, 用于外观模型和运动模型的融合。使用此方法将多示例学习与中值流进行融合, 即在给定  $t$  及  $t$  时刻以前的观测值的条件下最大化状态量的后验估计

$$p(X_t | Y_{1:t}) \propto p(Y_t | X_t) \cdot \int p(X_t | X_{t-1}) \cdot p(X_{t-1} | Y_{1:t-1}) dX_{t-1} \quad (4)$$

式中,  $X_t$  是  $t$  时刻的状态量,  $Y_{1:t}$  是  $t$  及  $t$  时刻之前的所有观测值。  $p(X_t | X_{t-1})$  为运动模型,  $p(Y_t | X_t)$  为外观模型。在本文中  $X_t = [x, y]^T$ , 即预测的物体中心位置。

### 3.2 外观模型

外观模型  $p(Y_t | X_t)$  由在线多示例分类器给出, 即

$$p(Y_t | X_t) = \sigma(H(x)) \quad (5)$$

式中,  $H(x)$  为训练出的强分类器,  $x$  为  $Y_t$  对应的特征, 为随机化的类 haar 特征<sup>[14]</sup>, 它是一种简单的矩形特征, 由于积分图的引入, 每个特征的计算只需进行简单的加法运算就可以得到, 所以是一种高效的特征。一般来说弱分类器选择比较随意, 在离线算

法中决策树桩的效果最好,但是在目标跟踪里要求弱分类器能够在线更新,因此采用在线的朴素贝叶斯分类器,每一个弱分类器对应一个特征  $f_k$  和 4 个参数  $(\mu_0, \sigma_0, \mu_1, \sigma_1)$ , 这 4 个参数可以在线更新。分类器定义为

$$h_k(x) = \log \left[ \frac{p_t(y=1|f_k(x))}{p_t(y=0|f_k(x))} \right] \quad (6)$$

式中,  $p_t(y=1|f_k(x)) \sim N(\mu_1, \sigma_1)$ , 即高斯分布。当新的训练数据来时,使用

$$\mu_1 = \gamma \mu_1 + (1 - \gamma) \frac{1}{n} \sum_{i|y_i=1} f_k(x_i) \quad (7)$$

$$\sigma_1 = \gamma \sigma_1 + (1 - \gamma) \sqrt{\frac{1}{n} \sum_{i|y_i=1} (f_k(x_i) - \mu_1)^2} \quad (8)$$

来更新分类器,式中  $0 \leq \gamma \leq 1$  是更新率。负类的更新完全类似,不再赘述。

在线多示例目标搜索区域使用中值流预测的结果进行修正,区域中心修正为  $\mathbf{X}_{t-1} + [dx, dy]^T$ , 搜索区域大小为  $s$ 。这样就可以更好地利用中值流跟踪快速运动物体的能力。

### 3.3 运动模型

运动模型  $p(\mathbf{X}_t | \mathbf{X}_{t-1})$  由中值流跟踪器给出。采用中值流预测结果的高斯加权形式,即

$$p(\mathbf{X}_t | \mathbf{X}_{t-1}) = \exp \left( - \frac{\| \mathbf{X}_t - \tilde{\mathbf{X}}_{t-1} \|^2}{\tau \cdot s^2} \right) \quad (9)$$

式中,  $[dx, dy]^T$  是中值流跟踪器预测的物体中心位移,  $\tilde{\mathbf{X}}_{t-1} = \mathbf{X}_{t-1} + [dx, dy]^T$ 。  $s$  是在线多示例跟踪中的搜索范围(实验中  $s = 25$ ),  $\tau$  为可调参数,表示系统中值流预测位置的信赖程度,若  $\tau \rightarrow +\infty$  说明只利用中值流对目标的大致的位置进行预测,然后完全使用多示例跟踪器检测目标。若  $\tau \rightarrow 0_+$  说明只考虑使用中值流进行目标跟踪,不进行多示例分类器的检测。通过多组  $\tau$  值试验,最终选择  $\tau = 4$ , 这样就融合了中值流跟踪器及多示例跟踪器的结果,既可以充分利用到中值流跟踪器可以跟踪剧烈运动的特点,又可以发挥多示例学习抗遮挡和跟踪漂移的优势。

### 3.4 算法描述

基于中值流的在线多示例目标跟踪算法结构如图 3 所示。

对于  $t$  时刻视频帧,本文算法先使用中值流方法预测目标位置,得到偏移值  $[dx, dy]^T$  并且由此更新多示例学习的搜索区域。在新的搜索区域内取一

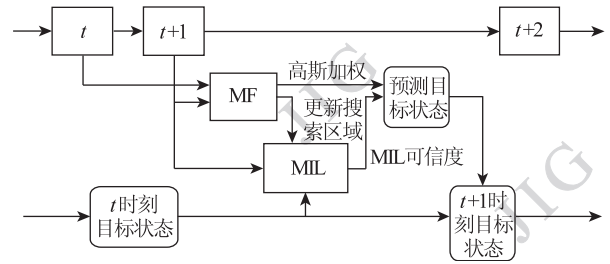


图 3 MIL-MF 算法结构图

Fig. 3 Chart of the MIL-MF algorithm

定数量的待测图像块,对每一图像块分别使用多示例学习方法由式(5)计算  $p(\mathbf{Y}_t | \mathbf{X}_t)$  和中值流式(9)计算高斯权重  $p(\mathbf{X}_t | \mathbf{X}_{t-1})$ , 最后由式(4)得到最终的似然度。取似然度最高的块作为目标的预测。并更新多示例分类器,选择随机化的类 haar 特征作为图像的表示,每个特征由 2~5 个随机生成的矩形组成,每个矩形都有一个随机的权重。通过计算每个矩形的像素和,再加权组合,得到一个特征值。其中矩形像素和可以通过积分图<sup>[14]</sup>快速获得。具体的算法步骤如下:

输入待测图像块  $\mathbf{Y}_{(t,i)}, i = 1, \dots, N$ ;

- 1) 进行中值流预测,得到偏移值  $[dx, dy]^T$ ;
- 2) 利用式(9)计算运动模型权值  $p(\mathbf{X}_{(t,i)} | \mathbf{X}_{t-1})$ ;
- 3) 利用式(5)计算外观模型权值  $p(\mathbf{Y}_{(t,i)} | \mathbf{X}_{(t,i)})$ ;
- 4) 利用式(4)计算每个块似然度  $p(\mathbf{X}_{(t,i)} | \mathbf{Y}_{(t,i)})$ ;
- 5) 取  $\hat{\mathbf{X}} = \max_{\mathbf{X}_{(t,i)}} p(\mathbf{X}_{(t,i)} | \mathbf{Y}_{(t,i)})$ , 即为目标位置;
- 6) 更新外观模型,返回步骤 1)。

## 4 实验结果及讨论

### 4.1 算法验证

使用本文算法对 12 组常用测试视频<sup>[9]</sup>进行跟踪。每组视频的特点见表 1。

表 1 测试视频的特点

Table 1 Features of test videos

视频名称	特点
david, sylv	光照, 尺度, 姿态
faceocc, faceocc2, girl	遮挡
tiger1, tiger2, coke11	剧烈运动, 旋转
dollar	外观变化
surfer	运动, 视角变化
cliffbar	相似背景, 尺度
twinings	外观变化, 尺度

对这 12 组视频分别使用基于集成学习的 EnT<sup>[5]</sup>、基于在线 Adaboost 的 OAB1<sup>[7]</sup> 及其变种 OAB5<sup>[7]</sup>、基于分块的 FragT<sup>[15]</sup>、基于中值流的 MF<sup>[11]</sup>、基于在线多示例学习的 MIL 和本文介绍的基于中值流的在线多示例学习 MIL-MF 进行结果的比较。其中 EnT、OAB1、OAB5 均使用默认参数, FragT 使用文献[15]提到的最好的一组参数, MIL 和 MIL-MF 中的多示例分类器的学习率  $\gamma = 0.85$ ,  $M = 250$ ,  $K = 50$ 。因为 OAB1、OAB5、MIL 和 MIL-MF 均采用随机的特征和对应的弱分类器,为了结果更加一般化,对每个算法在每组视频上均进行 10 次测试,取这 10 次结果的平均作为这个算法的结果。跟踪结果见表 2。

表 2 平均中心偏移像素  
Table 2 Average center location error

视频	EnT	OAB1	OAB5	FragT	MF	MIL	MIL-MF
sylv	45.6	28.1	77.7	<b>12.0</b>	21.6	13.9	<b>12.9</b>
tiger1	63.6	34.1	62.7	63.6	51.7	<b>19.1</b>	<b>21.1</b>
tiger2	53.0	28.0	38.4	53.0	23.9	<b>17.6</b>	<b>11.4</b>
surfer	186.8	17.0	83.2	139.9	<b>3.2</b>	8.2	<b>7.1</b>
david	127.0	41.7	72.6	69.6	<b>14.0</b>	23.4	<b>22.2</b>
coke11	84.6	25.3	43.9	69.5	32.6	<b>20.9</b>	<b>18.5</b>
girl	121.6	42.9	66.4	<b>26.5</b>	79.0	31.5	30.8
dollar	46.4	18.8	52.6	56.6	80.8	<b>21.6</b>	<b>21.0</b>
faceocc	61.4	43.5	104.1	<b>6.3</b>	49.1	29.4	<b>28.5</b>
faceocc2	57.5	22.1	93.7	45.5	34.5	<b>18.9</b>	<b>16.1</b>
twinings	74.0	16.2	82.1	13.5	17.6	<b>10.0</b>	<b>10.1</b>
cliffbar	131.6	14.6	46.1	34.1	36.8	<b>11.4</b>	<b>11.1</b>
平均	87.8	27.7	68.6	49.2	<b>37.1</b>	<b>18.8</b>	<b>17.6</b>

注:加粗下划线表示最好的结果,黑体加粗表示次好的结果。

可以看出 12 组视频中有 5 组 MIL-MF 均是最好的。其余的 7 组为次优。其他的算法有局限性,比如 FragT 算法对于遮挡有非常好的处理能力,所以在 faceocc、girl 上表现非常好,然而由于其使用静态的模板,在跟踪过程无法适应外观的变化,所以在其他的视频上表现不好。MIL 由于运动模型过于简

单,所以对于剧烈的运动(如 tiger2 和 coke11)上表现不如 MIL-MF。MF 方法对不同的视频处理结果波动较大,当跟踪目标较平稳的时候效果较好(比如 surfer 和 david),但是对于遮挡的处理能力较弱,而且一旦出现遮挡跟踪性能急剧下降。综合来看 MIL-MF 对于剧烈运动、遮挡、光照变化、视角变化等均具有非常好的处理能力。表 2 给出了平均的中心偏移误差,可以看出 MIL-MF 算法在平均意义下具有最好的性能。图 4 为目标中心像素偏离值曲线图,表示跟踪结果与真实目标的中心位置的欧氏距离,距离越小表示跟踪结果越有效。从图 4 中可以看出, MIL-MF 具有很低的中心偏移误差(对应图中的红色实线)。

图 5 给出了部分视频的跟踪结果及对应的帧号,可以看出 MIL-MF 对于遮挡、光照、剧烈运动的鲁棒性。值得一提的是对于以上的大小为  $320 \times 240$  的视频,在主频为 2.7 G 的普通 PC 上 MIL-MF 的帧率可以达到 14 帧/s 以上,具有较好的实时性。可以广泛应用在家庭服务机器人等领域,具有很大的实际意义。

#### 4.2 演示学习中目标跟踪实验

设计了一个演示学习案例,通过向机器人示范拿取红瓶子和蓝瓶子的不同动作,一共两组演示,并给出对演示的评价,机器人通过推理学习就可以获得正确的知识,进而学会执行正确的操作。在这个示例下关键是正确地跟踪和检测红瓶子和蓝瓶子(图 6)。图 6 显示的是实验室服务机器人采集的视频数据中的部分图,从图中可以看出,实验环境十分接近家庭环境,背景较复杂,光照条件也比较一般。特别需要机器人采集的视频数据的帧率较低(约为 5 帧/s),因此相当于物体的运动非常快速。在上面的实验中可以看出 MIL-MF 和 MIL 效果最好,这里只对这两种方法进行对比。图 6 中黄色虚线框为使用 MIL 的结果,绿色实线框为使用 MIL-MF 的部分跟踪结果,图的左上角的为帧号。可以看出 MIL-MF 算法性能比 MIL 好得多,这是因为低帧率下的视频图像相当于目标进行了快速运动,原始的 MIL 方法运动模型较为朴素, MIL-MF 采用了中值流作为 MIL 中的运动模型,对于快速运动具有非常强的适应性。

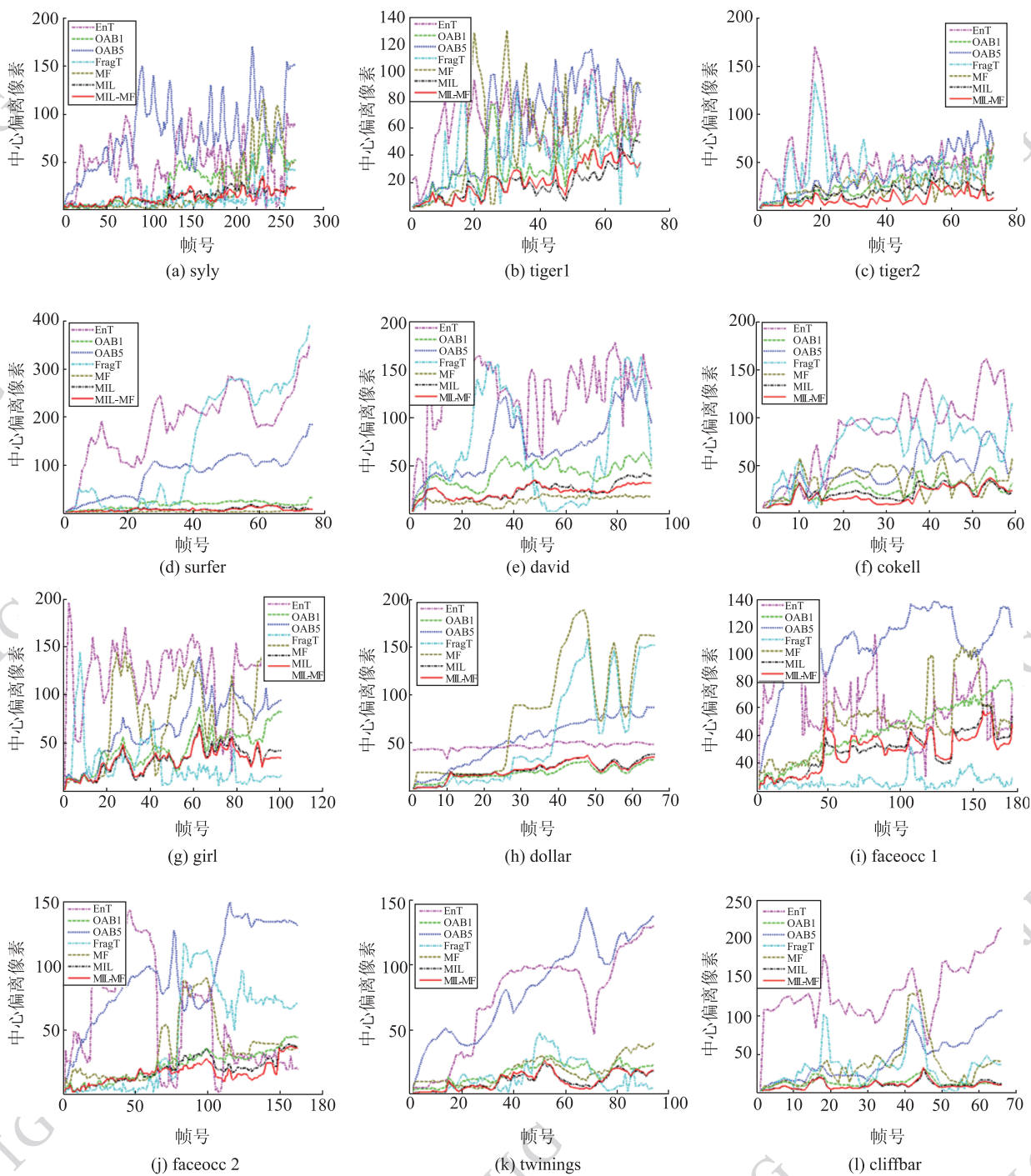


图 4 中心偏离误差曲线  
 Fig. 4 Location error plots

## 5 结 论

使用中值流辅助在线多示例学习跟踪,并对多组视频进行测试,结果表明提出的 MIL-MF 具有很好的跟踪性能,对遮挡、剧烈运动、漂移、光照等均具有很好的效果。并且针对家庭环境下的演示学习场景进行了真

实测试,表明了在实际场景下本文算法可以有效地完成低帧率下目标跟踪,并且可以达到实时跟踪,因此可以满足演示学习中对目标跟踪算法的特殊要求,可以应用在家庭服务机器人等领域。但是由于在机器人等领域常常是多目标跟踪,而将本文算法直接扩展至多目标跟踪将得不到实时的跟踪结果,因此如何使用本文算法进行多目标跟踪是今后的一个重要的工作。



图 5 不同跟踪算法比较

Fig. 5 Tracking results of different algorithms

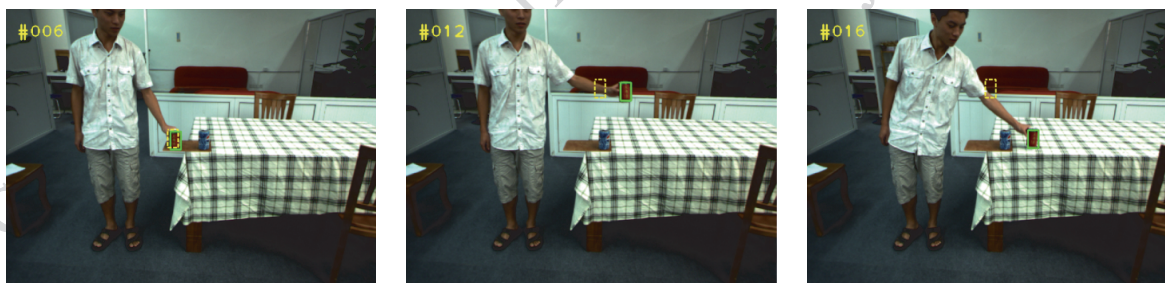


图 6 演示学习环境下的目标跟踪结果

Fig. 6 Tracking results of the robot's learning from demonstration scenario

参考文献 (References)

[ 1 ] Argall B D, Chernova S, Veloso M, et al. A survey of robot learning from demonstration [J]. Robotics and Autonomous Systems, 2009, 57(5): 469-483.

[ 2 ] Zhang R, Wang S F, Chen X P, et al. Towards robot incremen-

tal learning constraints from comparative demonstration (demonstration) [C] // Proceedings of the 10th International Conference on Autonomous Agents and Multi-agent Systems. Taipei, Taiwan, China; ACM, 2011: 1301-1302.

[ 3 ] Zhou Q H. Tracking of moving object based on multiple instance learning [D]. DaLian; Dalian University of Technology, 2010 [周秋红. 基于多示例学习的运动目标跟踪算法研究 [D].

- 大连:大连理工大学,2010.]
- [ 4 ] Hou Z Q, Han C Z. A Survey of visual tracking [J]. *Acta Automatica Sinica*, 2006, 32(4):603-617. [侯志强,韩崇昭. 视觉跟踪技术综述[J]. *自动化学报*, 2006, 32(4):603-617.]
- [ 5 ] Avidan S. Ensemble tracking [J]. *IEEE Trans. on PAMI*, 2007, 29(2):261-271.
- [ 6 ] Collins R, Liu Y, Leordeanu M. Online selection of discriminative tracking features [J]. *IEEE Trans. on PAMI*, 2005, 27(10):1631-1643.
- [ 7 ] Grabner H, Bischof H. On-line boosting and vision [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. New York, USA:IEEE, 2006:260-267.
- [ 8 ] Oza N C. Online ensemble learning [D]. California: University of California, 2001.
- [ 9 ] Babenko B, Ming-Hsuan Y, Belongie S. Visual tracking with online multiple instance learning [C] // *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Miami, USA:IEEE, 2009:983-990.
- [ 10 ] Babenko B, Ming-Hsuan Y, Serge B. Robust object tracking with online multiple instance learning [J]. *IEEE Trans. on PAMI*, 2011, 33(8):1619-1632.
- [ 11 ] Kalal Z, Mikolajczyk K, Matas J, et al. Forward-backward error: automatic detection of tracking failures [C]// *Proceedings of the 20th International Conference on Pattern Recognition*. Istanbul, Turkey:IEEE, 2010:2756-2759.
- [ 12 ] Baker S, Matthews I. Lucas-Kanade 20 years on: A unifying framework [J]. *International Journal of Computer Vision*, 2004, 56(3):221-255.
- [ 13 ] Viola P, Platt J, Zhang C. Multiple instance boosting for object detection [C] // *Proceedings of Neural Information Processing System*. Vancouver, Canada:ACM, 2006:1419-1426.
- [ 14 ] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features [C] // *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Hawaii, USA:IEEE, 2001(1):511-518.
- [ 15 ] Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram [C] // *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. New York, USA:IEEE, 2006:798-805.