

基于无监督在线学习实现视频遮挡边界检测

张世辉^{1,2} 王瑞宇¹ 何欢¹

¹燕山大学信息科学与工程学院, 河北 秦皇岛 066004

²河北省计算机虚拟技术与系统集成重点实验室, 河北 秦皇岛 066004

摘要 为了检测视频序列中的遮挡边界,提出一种新颖的基于无监督在线学习的遮挡边界检测方法。该方法提取视频序列中待测帧的遮挡相关特征并计算其对应的时间长度,利用对冲算法思想并结合时间长度及不同遮挡特征求得待测帧中像素点的遮挡相关信息,利用各特征的遮挡相关信息进行投票,完成当前帧图像的遮挡边界检测。利用 Online Boosting 思想以当前帧的检测结果来估计下一帧的特征投票权重,实现后续帧图像的遮挡边界检测。该方法通过在线学习思想改变不同特征的权重完成遮挡边界检测功能,无需预先获取视频序列的先验知识。实验结果表明,同已有方法相比,该方法具有较高的准确性和较好的通用性。

关键词 机器视觉; 遮挡边界; 无监督学习; 在线学习; 对冲算法; Online Boosting

中图分类号 TP391 **文献标识码** A

doi: 10.3788/AOS201535.1215001

Occlusion Boundary Detection for Video Sequences Based on Unsupervised Online Learning

Zhang Shihui^{1,2} Wang Ruiyu¹ He Huan¹

¹School of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066004, China

²Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, Qinhuangdao, Hebei 066004, China

Abstract In order to detect occlusion boundary in video sequences, a novel occlusion boundary detection approach based on unsupervised online learning is proposed. The occlusion related features of the frame to be detected in video sequences are extracted and the time length corresponding to the frame is calculated. The pixel points' occlusion related information in the frame to be detected is obtained using Hedge algorithm and combining time length with different occlusion features. The occlusion related information of different features is voted to accomplish occlusion boundary detection of current frame. The detection result of current frame is used to estimate the feature weight for next frame based on Online Boosting idea to realize the detection of subsequent frames. The proposed method changes the weight of different features through online learning idea to accomplish occlusion boundary detection, and does not need to obtain priori knowledge of video sequences in advance. Experimental results show that, compared with existing methods, the proposed method has higher accuracy and better generality.

Key words machine vision; occlusion boundary; unsupervised learning; online learning; Hedge algorithm; Online Boosting

OCIS codes 150.0155; 150.4620; 110.4153; 330.1880

1 引言

视频序列中遮挡边界的检测结果将直接影响到视觉测量、场景理解与绘制、形状提取、运动跟踪、舆情监控等诸多视觉应用的效果。与单幅图像相比,视频序列可以为遮挡边界检测提供更多有利的信息,因此,

收稿日期: 2015-05-28; 收到修改稿日期: 2015-07-01

基金项目: 国家自然科学基金(61379065)、河北省自然科学基金(F2014203119)

作者简介: 张世辉(1973-),男,博士,教授,主要从事视觉信息处理、模式识别、智能机器人等方面的研究。

E-mail: sshhzz@ysu.edu.cn

基于视频序列的遮挡边界检测方法逐渐引起了国内外学者的关注^[1-2]。目前,基于视频序列的遮挡边界检测方法主要分为两类:基于运动分析的方法和基于机器学习的方法。基于运动分析的方法^[3-4]单纯利用运动估计确定遮挡边界,但运动矢量场中的任何误差都有可能导致遮挡边界检测结果的不准确,且该类方法的通用性较差。

为了克服基于运动分析方法的不足,一些学者相继提出了基于机器学习的遮挡边界检测方法^[5-10]。最具代表性的是Stein^[6]提出的基于监督学习的遮挡检测方法,此方法以多幅彩色图像为依据,通过结合运动线索和局部边缘线索完成遮挡边界检测。分析可知,基于监督学习思想的遮挡边界检测方法在一定程度上克服了基于运动分析方法存在的问题,然而在很多现实问题中,虽然能较容易地获得大批数据,但对这些数据进行标记却需要耗费大量的人力物力,并且一些情况下无法完成数据的标记,从而导致无法进行监督学习。有标记数据少,未标记数据多的现象在网上应用中更为明显。鉴于此,学者们又提出了基于无监督在线学习的遮挡边界检测方法。具有代表性的是文献[10]以视频序列为依据,通过对冲算法实现遮挡边界检测。此方法虽然不需要对数据进行标记,但并未充分利用视频序列中多帧图像的有效信息,且仅基于单一特征进行遮挡检测,在准确性和适用范围方面仍有待改善。

为了解决已有无监督方法存在的问题,本文提出一种新颖的基于无监督在线学习结合 Online Boosting 思想实现视频遮挡边界检测的方法。实验结果验证了所提方法的可行性和有效性。

2 方法概述

基于无监督在线学习思想实现视频遮挡边界检测,具体步骤如下:1) 导入视频序列中的待测帧,提取其对应的遮挡相关特征并计算时间长度。2) 根据时间间隔导入两帧图像,通过遮挡相关特征计算待测帧的瞬间损失与累计损失,然后利用累计损失计算出各像素点的初步遮挡检测结果。3) 根据各像素点的初步遮挡检测结果与时间长度得到各像素点的遮挡类型。4) 根据各像素点的遮挡类型,利用多特征获得各像素点的遮挡检测结果并以此进行加权投票,进而得到待测帧中各像素点的最终检测结果。5) 利用 Online Boosting 思想结合待测帧中各像素点的最终检测结果更新不同特征的权重,为下一帧的检测做准备。6) 循环上述操作,进行下一帧图像的遮挡检测,最终完成视频序列的遮挡边界检测功能。所提方法的总体流程如图1所示。

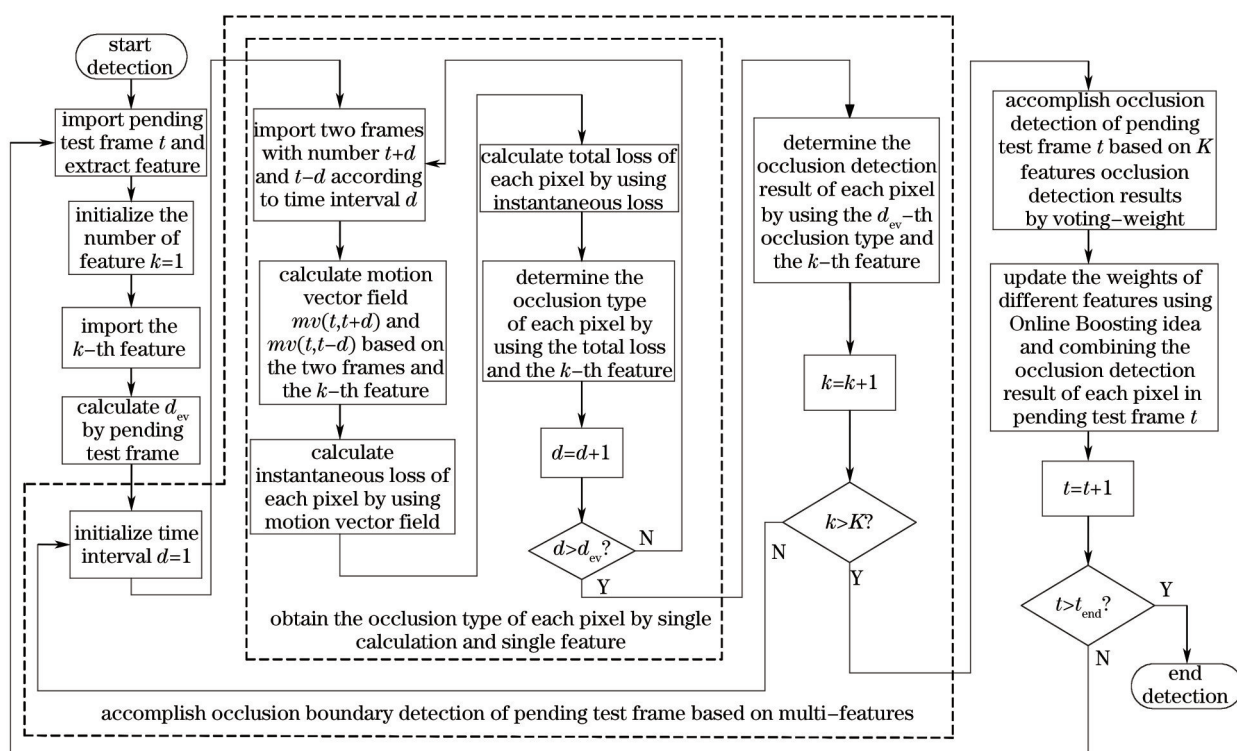


图1 遮挡边界检测方法总体流程

Fig.1 Overall process of occlusion boundary detection approach

3 基于无监督在线学习方法实现视频遮挡边界检测

3.1 遮挡相关特征提取及时间长度的计算

为了实现遮挡检测,需要提取待测帧中各像素点的遮挡相关特征。由于光流计算是很多运动匹配特征提取的前提,故光流算法的好坏对遮挡相关特征的提取至关重要。现有的光流算法有很多,针对光流存在的遮挡、大位移以及运动边界保持等问题,选取分别应对上述光流问题的光流三种算法{Sparse Occlusion, MDP-Flow2, Classic+NL}^[11-13]作为候选光流集,即光流算法 $\alpha \in \{\text{Sparse Occlusion, MDP-Flow2, Classic+NL}\}$,以多光流为基础进行特征的计算与提取。同时,考虑到亮度特征是彩色视频序列中用于对应点匹配的最基本也是最实用的一种特征,而梯度特征则是检测视频图像边界的关键因素,故具体选取亮度一致性特征(PC)^[14]与梯度一致性特征(GC)^[15]组成遮挡特征集{PC,GC}进行遮挡边界检测,即特征 $u \in \{\text{PC,GC}\}$ 。基于光流算法与特征种类,可得特征数 $K=3 \times 2=6$,即特征数 K 由光流算法及特征种类组合决定,并设 k 为特征数 K 的索引。

此外,考虑到视频序列中相邻帧的间隔非常短,对于运动矢量场来说变化十分有限,为了提高检测结果的准确性,通过引入时间长度 d ,利用多帧图像同时计算像素点的遮挡分类结果。设视频序列第一帧为 t_{begin} ,最后一帧为 t_{end} ,待测帧为 t ,将 $t-t_{\text{begin}}$ 和 $t_{\text{end}}-t$ 中的较小值作为时间长度 d ,即 $d = \min\{t_{\text{end}}-t, t-t_{\text{begin}}\}$ 。从待测帧 t 开始向两侧以 1 帧为时间步长取前后两帧图像,第一次取 $t+1$ 帧与 $t-1$ 帧,第二次取 $t+2$ 帧与 $t-2$ 帧,以此类推,直到 $t+d$ 帧与 $t-d$ 帧,然后分别利用所取的前后两帧及待测帧信息多次计算像素点的遮挡分类结果,完成对该待测帧的遮挡边界检测,具体如图 2 所示。

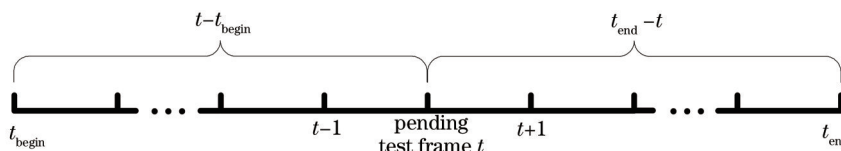


图 2 时间长度 d 的计算

Fig.2 Calculation of time length d

同时,为了便于比较不同方法的效果,按照文献[10]方法将遮挡分为 9 类,即 $\gamma(\theta, \mu)$ (包括 8 种遮挡情况)和无遮挡。其中, $\theta \in \{0, \pi/4, \pi/2, 3\pi/4\}$ 表示遮挡边界像素点以何种角度遮挡背景像素点, $\mu \in \{0, 1\}$ 表示前景物体与背景像素点的遮挡关系。 $\mu = 1$ 表示前景物体即将遮挡背景像素点, $\mu = 0$ 时表示前景物体正在远离背景像素点。 θ 和 μ 的组合为 8 种遮挡情况。对于视频帧中的每一个像素点,分别计算这 9 种遮挡与该像素点的匹配情况,选出其匹配程度最好的遮挡类型作为此像素点的遮挡类型,得到此帧图像像素点的遮挡相关信息。同时,引入像素块概念以辅助估计像素点的遮挡类型,依据 9 种不同的遮挡类型分别以待测像素点为中心建立 9 个像素块,各像素块均由 $\delta \times \delta$ 的像素点组成。各像素块与中心点的位置如图 3 所示。

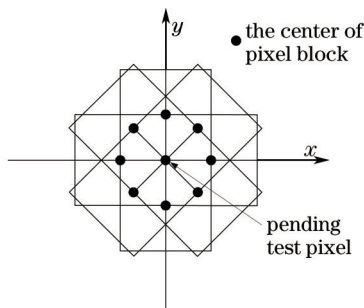


图 3 像素块与像素块中心点示意图

Fig.3 Sketch map of pixel block and its central point

3.2 基于时间长度及多特征多次利用对冲算法计算待测帧的遮挡边界

对冲算法^[16]是在线学习中的一种重要方法,其主要思想是利用预测结果与真实结果的差异衡量预测结果与真实结果的匹配程度,并将匹配程度最好的情形作为分类结果。提出一种多特征对冲算法框架,利用多特征共同决策完成遮挡边界检测功能。此框架首先根据不同特征计算瞬间损失,然后由瞬间损失计算累积损失,在此基础上得到各像素点与遮挡类型的匹配程度,从而进一步确定出待测帧的遮挡边界。

3.2.1 瞬间损失计算

瞬间损失是指像素块中各像素点真实位置与预测位置所对应特征的绝对差值之和。瞬间损失的计算公式定义为

$$l_{i,m}^{i,k}(d) = l_0^k(d) + l_1^k(d) \quad , \quad (1)$$

式中 $l_{i,m}^{i,k}(d)$ 表示使用第 k 个特征计算时,待测帧(第 t 帧)中像素点 m 所对应的瞬间损失, $i \in [1,9]$ 代表不同的遮挡类型。其中,1~8代表不同的遮挡情况,9代表无遮挡, $d \in [1, d_{\max}]$ 代表不同的时间间隔。 $l_0^k(d)$ 、 $l_1^k(d)$ 在遮挡情况下分别表示待测像素点两侧像素块中各像素点的真实位置与预测位置所对应特征的绝对差值之和,在无遮挡情况下表示前后两帧像素块与待测帧像素块中各像素点的真实位置与预测位置所对应特征的绝对差值之和。

在8种遮挡情况下, $l_0^k(d)$ 、 $l_1^k(d)$ 计算公式定义为

$$l_0^k(d) = \sum_{a=-\lfloor \frac{\delta}{2} \rfloor}^{\lfloor \frac{\delta}{2} \rfloor} \sum_{b=-\lfloor \frac{\delta}{2} \rfloor}^{\lfloor \frac{\delta}{2} \rfloor} \left| f_i^u(x_0 + a, y_0 + b) - f_{i+d}^u(x_0 + a + u_0^\alpha, y_0 + b + v_0^\alpha) \right| \quad , \quad (2)$$

$$l_1^k(d) = \sum_{a=-\lfloor \frac{\delta}{2} \rfloor}^{\lfloor \frac{\delta}{2} \rfloor} \sum_{b=-\lfloor \frac{\delta}{2} \rfloor}^{\lfloor \frac{\delta}{2} \rfloor} \left| f_i^u(x_1 + a, y_1 + b) - f_{i+d}^u(x_1 + a + u_1^\alpha, y_1 + b + v_1^\alpha) \right| \quad , \quad (3)$$

式中 f_i^u 、 f_{i+d}^u 为特征函数, u 为特征, α 为光流算法, (x_0, y_0) 、 (x_1, y_1) 为待测像素点两侧像素块中心点的坐标, $u_0^\alpha = (u_0^\alpha, v_0^\alpha)$ 、 $u_1^\alpha = (u_1^\alpha, v_1^\alpha)$ 为两侧像素块中像素点的光流值。

无遮挡情况下, $l_0^k(d)$ 、 $l_1^k(d)$ 的计算公式分别定义为

$$l_0^k(d) = \sum_{a=-\lfloor \frac{\delta}{2} \rfloor}^{\lfloor \frac{\delta}{2} \rfloor} \sum_{b=-\lfloor \frac{\delta}{2} \rfloor}^{\lfloor \frac{\delta}{2} \rfloor} \left| f_i^u[x(m) + a, y(m) + b] - f_{i+d}^u[x(m) + a + u_2^\alpha, y(m) + b + v_2^\alpha] \right| \quad , \quad (4)$$

$$l_1^k(d) = \sum_{a=-\lfloor \frac{\delta}{2} \rfloor}^{\lfloor \frac{\delta}{2} \rfloor} \sum_{b=-\lfloor \frac{\delta}{2} \rfloor}^{\lfloor \frac{\delta}{2} \rfloor} \left| f_i^u[x(m) + a, y(m) + b] - f_{i-d}^u[x(m) + a + u_3^\alpha, y(m) + b + v_3^\alpha] \right| \quad , \quad (5)$$

式中 f_{i-d}^u 、 f_i^u 、 f_{i+d}^u 为特征函数, u 为特征, α 为光流算法, $u_2^\alpha = (u_2^\alpha, v_2^\alpha)$ 、 $u_3^\alpha = (u_3^\alpha, v_3^\alpha)$ 为前后两帧像素块中像素点的光流值。

为了进一步说明上述公式的含义,此处以遮挡类型 $\gamma(\pi/2, 0)$ 与无遮挡为例来说明瞬时损失的计算过程。在遮挡类型为 $\gamma(\pi/2, 0)$ 时,设待测像素点的坐标为 $[x(m), y(m)]$,取左右两个像素块,根据前文叙述可知,左右两侧像素块中心点的坐标分别为 $(x_0, y_0) = \{x(m) - \lceil \delta/2 \rceil, y(m)\}$ 和 $(x_1, y_1) = \{x(m) + \lceil \delta/2 \rceil, y(m)\}$,如图4所示。

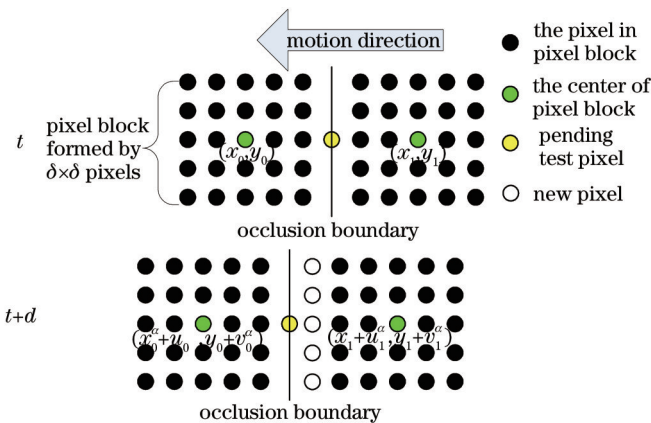


图4 遮挡类型为 $\gamma(\pi/2, 0)$ 时像素点的瞬间损失计算

Fig.4 Instantaneous loss calculation of pixel point with occlusion type $\gamma(\pi/2, 0)$

对于预测为无遮挡的类型,由于此时需要更多的信息预测像素点的状态,故取前后两帧同时进行计

算。该方法是以像素点 $[x(m), y(m)]$ 为中心取像素块, 像素点的瞬间损失通过待测帧与前后两帧所取像素块中各像素点的真实位置与预测位置对应特征的绝对差值之和计算。具体过程如图 5 所示。

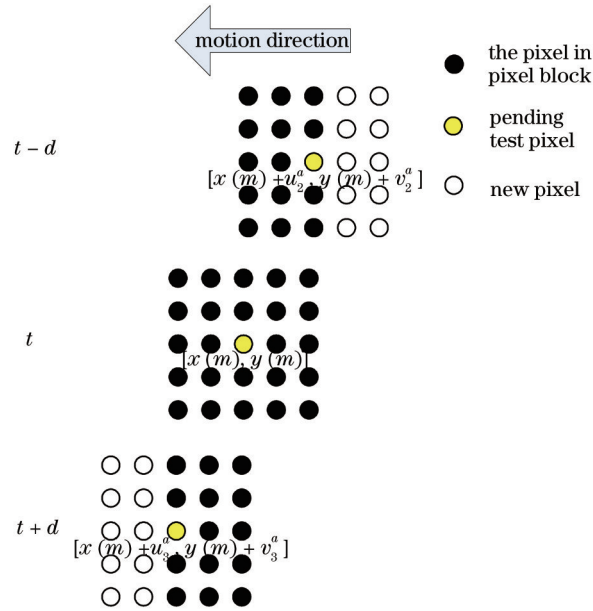


图 5 无遮挡时像素点的瞬间损失计算

Fig.5 Instantaneous loss calculation of pixel point without occlusion

由此可以计算出每一像素点在不同遮挡类型下的瞬间损失, 在得到瞬间损失的基础上, 即可计算累计损失。

3.2.2 累计损失计算及遮挡类型判断

累计损失为瞬间损失的迭代。累计损失的计算公式定义为

$$L_{i,m}^{i,k}(d) = (1 - \mu)L_{i,m}^{i-1,k}(d) + L_{i,m}^{i,k}(d) \quad (6)$$

式中 $L_{i,m}^{i,k}(d)$ 表示待测帧 t 中的像素点 m 基于第 k 个特征在遮挡类型 i 下的第 d 次累计损失, $\mu \in \{0, 1\}$ 表示遮挡物体与背景像素点的遮挡关系, $L_{i,m}^{i-1,k}(d)$ 表示从第一帧 t_{begin} 到 $t-1$ 帧所有瞬间损失的迭代。对于第一帧 t_{begin} , 其累计损失即为其瞬间损失。对于待测帧 t 中任一像素点, 当 d 取不同值时, 由(6)式可得到该像素点基于第 k 个特征的 d_{ev} 个累计损失。

在得到累计损失的情况下, 为了更好地由累计损失判定各像素点的遮挡类型, 将累计损失最小的遮挡类型 i 作为像素点 m 的遮挡类型, 即定义像素点遮挡类型的判定公式为

$$D_m^{i,k}(d) = \arg \min_i [L_{i,m}^{i,k}(d)] \quad (7)$$

式中 $D_m^{i,k}(d)$ 为待测帧 t 中像素点 m 基于特征 k 和时间间隔 d 的遮挡类型判定结果。

3.2.3 基于不同时间间隔及多特征实现当前帧的遮挡边界检测

由于单次计算的不准确性, 为了减少误差, 综合像素点 m 的 d_{ev} 次遮挡类型判定结果 $D_m^{i,k}(d)$ 求得像素点 m 基于特征 k 的遮挡检测结果 $C_k^i(m)$ 。求解过程中, 若 $D_m^{i,k}(d)$ 为某一类遮挡类型则设定其为 1, 若 $D_m^{i,k}(d)$ 为无遮挡类型则设定其为 0, 则 $C_k^i(m)$ 计算公式定义为

$$C_k^i(m) = \frac{\sum_{d=1}^{d_{\text{ev}}} D_m^{i,k}(d)}{d_{\text{ev}}} \quad (8)$$

根据(8)式, 若像素点的 $C_k^i(m) > \frac{1}{2}$ 则判定其为遮挡点, 否则判定为无遮挡点。至此, 得到了待测帧 t 中各像素点基于特征 k 的初步遮挡检测结果。为了将所有 K 类不同特征的检测结果相结合得到待测帧的最终遮挡检测结果, 需要对这些结果进行投票。考虑到一开始不知道每一个特征的优劣程度, 所以默认其初始

权重相同。也就是说,对于某一像素点,若有 $\lceil K/2 \rceil$ 个特征认定其为遮挡像素点,则判定此像素点为遮挡点,反之则判定其为无遮挡,至此可得到待测帧 t 基于不同时间间隔及多特征的最终遮挡边界检测结果。

3.3 基于多特征利用 Online Boosting 思想实现视频序列的遮挡边界检测

基于多特征对冲算法框架完成待测帧 t 的遮挡检测后,需要进行下一帧的遮挡检测,以便完成整个视频序列的遮挡边界检测。在无监督在线学习时,存在一个非常棘手的难题,即分类器的每一次更新都会引入一定的错误,错误累积到一定程度,就会产生漂移,导致检测失败。为了减少此类错误,引入具有在线性且可以在预测未知数据的过程中进行学习的 Online Boosting 思想^[17-19]完成视频序列的遮挡检测,但由于 Online Boosting 思想并不掌握有关样本类别的先验知识,所以如何获取样本的类别是最大的难点。为了解决此问题,创造性地将对冲算法与 Online Boosting 方法相结合实现无监督在线学习,将各特征对应的检测结果作为 Online Boosting 方法中的输入,即根据特征决定分类器的个数,再利用 Online Boosting 思想将不同特征的分类结果进行投票,从而得到最终的遮挡检测结果。后续帧遮挡检测方法的基本思想如下。

首先,以待测帧 t 的最终遮挡边界结果为依据,分别将各特征对应的检测结果与最终结果进行比较,对于分类错误的特征,则减少其对应的权重;对于正确分类的特征,则增加其权重,这样可以区分出不同特征对于图像中各像素点分类的优劣程度。然后,在得到不同特征的权重后,再对下一帧进行检测,因为相邻两帧变化很小,所以将已得到的特征权重估计为下一帧的特征权重,并将下一帧特征得到的初步遮挡检测结果与该权重结合求得下一帧的最终遮挡检测结果。依此类推,即可完成整个视频序列的遮挡边界检测。

基于上述基本思想,将各特征对应的检测结果作为 Online Boosting 方法中的输入,每个特征的初步遮挡检测结果用 $h_k(m)$ 表示。为了衡量每个特征的分类效果,记 λ_k^{corr} 和 λ_k^{wrong} 分别为特征 k 分类正确和错误的像素点个数, $O(m)$ 为最终检测结果。分别将每一特征的分类结果与最终检测结果进行比较,若 $h_k(m) = O(m)$, 则 $\lambda_k^{\text{corr}} = \lambda_k^{\text{corr}} + 1$, 否则 $\lambda_k^{\text{wrong}} = \lambda_k^{\text{wrong}} + 1$ 。也就是说,通过遍历比较两帧图像中的所有像素点,可得到各特征对应的 λ_k^{corr} 和 λ_k^{wrong} 。

在求得各特征对于各像素点的分类情况后,可进一步求得各特征的累计错误率 ε_k 。 ε_k 的计算公式定义为

$$\varepsilon_k = \frac{\lambda_k^{\text{wrong}}}{\lambda_k^{\text{wrong}} + \lambda_k^{\text{corr}}} \quad (9)$$

得到不同特征的累计错误率后可进一步求得不同特征的权重 a_k 。 a_k 的计算公式定义为

$$a_k = \ln \frac{1}{\varepsilon_k} \quad (10)$$

将 a_k 归一化,则 $a_k = \frac{a_k}{a_1 + a_2 + \dots + a_K}$ 。以上一帧的特征权重结合此帧中不同特征得到的分类结果可得到此帧图像的最终遮挡边界检测结果 $C(m)$ 。 $C(m)$ 计算公式定义为

$$C(m) = \begin{cases} 1 & , \quad a_1 \times C_1(m) + a_2 \times C_2(m) + \dots + a_K \times C_K(m) \geq \frac{1}{2} \\ 0 & , \quad a_1 \times C_1(m) + a_2 \times C_2(m) + \dots + a_K \times C_K(m) < \frac{1}{2} \end{cases} \quad (11)$$

式中 $C_1(m), C_2(m), \dots, C_K(m)$ 分别为 K 类特征的分类结果。基于此思想继续进行后续帧的检测,可最终完成视频序列的遮挡边界检测功能。

4 实验及分析

4.1 实验环境及数据集

常用的视频序列遮挡边界检测数据集有 CMU 遮挡数据集^[6]、BVSD^[7]等,其中,CMU 遮挡数据集具有较强的代表性,故采用 CMU 遮挡数据集进行实验。CMU 遮挡数据集包括 30 个真实场景的视频序列,每个视频序列长约 5~30 帧不等。针对每个视频序列,由于 CMU 遮挡数据集中只给出了中间帧的 Ground Truth,所以选择该帧进行实验分析。实验硬件环境为 CPU Intel(R) Core(TM) i7-3770 3.40 GHz,内存 8.0 G。遮挡边界检测程序采用 Matlab 软件编程实现。实验过程中,不失一般性,设定像素块参数 $\delta = 7$ 。

4.2 实验比较及分析

为了验证所提遮挡边界检测方法的有效性,将本文方法与代表无监督在线学习遮挡边界检测方法最新进展的文献[10]方法在 CMU 遮挡数据集上进行了实验对比。此外,为了验证时间长度及多特征对遮挡检测效果的影响,增加了一组基于本文方法思想但去除 Online Boosting 环节的特征平均投票方法(FMV)的实验对比。图 6 给出了部分实验结果。其中,图 6(a)为视频序列中间帧对应的原始图像,图 6(b)为 Ground Truth,图 6(c)为文献[10]方法的检测结果,图 6(d)为 FMV 方法的检测结果,图 6(e)为本文方法的检测结果。

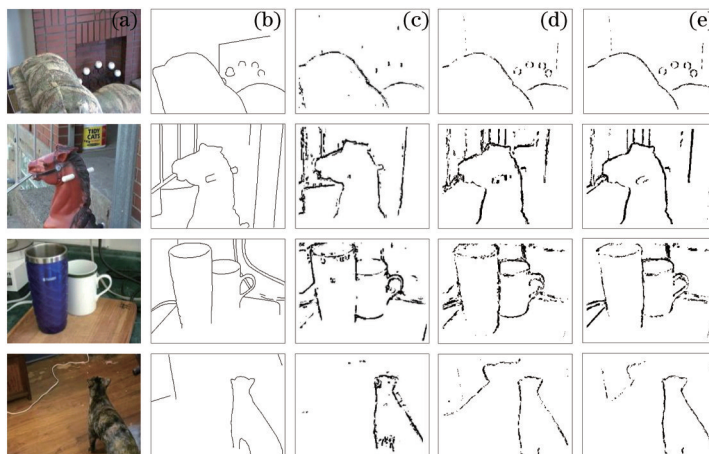


图 6 不同遮挡边界检测方法的实验结果比较。(a)原始图像; (b) Ground Truth; (c) 文献[10]方法; (d) FMV 方法; (e) 本文方法
Fig.6 Experimental result comparison of different occlusion boundary detection methods. (a) Original image; (b) ground truth; (c) Ref.[10] method; (d) FMV method; (e) proposed method

结合图 6 不同方法的实验结果并对照图 6(b)的 Ground Truth 进行分析可知,第一行中文献[10]方法仅检测出沙发的遮挡边界,漏检了壁炉处圆球的遮挡边界,FMV 方法虽然检测出了壁炉处圆球的遮挡边界但将墙壁纹路误检为遮挡边界,而本文方法在检测出沙发遮挡边界的同时也检测出圆球的遮挡边界且误检较少。第二行中文献[10]方法没有检测出马的左侧扶手的遮挡边界且马身的遮挡边界连续性较差,FMV 方法在马的左侧扶手的遮挡边界检测上有了一定地改善但马身的遮挡边界仍然比较粗糙,而本文方法在马的左侧扶手和马身的遮挡边界检测效果上明显高于前两种方法。第三行中文献[10]方法将图中机器下侧边缘及左侧杯子上的标签边缘误检为遮挡边界,FMV 方法虽然避免了对左侧杯子上标签的误检,但对机器下侧边缘仍存在误检现象,而本文方法则实现了正确检测。第四行中文献[10]方法将猫身体上的部分花纹误检为遮挡边界且遮挡边界比较粗糙,FMV 方法中猫身体上的遮挡边界存在明显的不连续现象,而本文方法则减少了此类错误。

由上述分析可知:文献[10]方法虽然能够较为准确地检测出图像中比较明显的遮挡边界,但是检测结果中存在一些较明显的误检和漏检,遮挡边界的准确性和连续性有待提高。与文献[10]方法相比,FMV 方法虽然减少了误检和漏检现象,但仍有待改善。比较而言,本文方法的检测结果明显优于文献[10]方法与 FMV 方法。

为了更加全面合理地评估各遮挡检测方法的准确性,基于 F-score 对文献[10]方法、FMV 方法和本文方法进行了量化评估。F-score 的计算公式为 $2PR/(P+R)$,其中, P 为准确率, R 为召回率。表 1 给出了各方法对 CMU 数据集中全部 30 组测试样本进行遮挡边界检测时整体 F-score 的比较结果。由表 1 可知,本文方法在 30 个视频序列上的平均 F-score 明显高于文献[10]方法和 FMV 方法。

表 1 不同方法基于 CMU 数据集的平均 F-score 比较

Table 1 Average F-score comparison for different methods based on CMU dataset

Detection method	F-score
Ref.[10] method	43.86%
FMV method	46.19%
Proposed method	50.62%

综合上述分析可知,无论从定性还是从定量角度看,FMV方法比文献[10]方法的准确性高,而本文方法比FMV方法和文献[10]方法更精确。FMV方法比文献[10]方法准确性高的主要原因是:与仅基于相邻帧及单特征的文献[10]方法相比,FMV方法考虑了时间长度且使用了多特征,通过综合考虑多帧图像和多种特征计算像素点的遮挡情况,提高了遮挡边界检测的准确性。同样,本文方法与FMV方法相比,虽然FMV方法也考虑了时间长度且使用了多特征,但FMV方法在决策时使用的是平均投票法,而本文方法则结合了 Online Boosting 思想动态调整不同特征的权重,将准确性好的特征动态赋予更高的权重,充分体现了在线学习的思想,所以本文方法又比FMV方法的准确性高。总体而言,本文方法通过综合多帧图像多特征的加权投票结果更有利于检测出多种情况下的遮挡边界,从而使得本文方法检测结果的准确性更高。

5 结 论

提出一种利用无监督在线学习思想实现视频遮挡边界检测的方法。该方法以不同特征为依据进行对冲算法计算,得到不同的遮挡相关信息,综合这些信息进行投票得到待测帧的遮挡边界检测结果,再以此结果为依据更新特征权重,进行下一帧图像的检测,最终完成视频序列的遮挡边界检测功能。该方法的贡献在于:1) 定义了时间长度这一概念,并基于此概念提出了多次求解待测帧遮挡边界的方法,综合多帧图像信息计算像素点的遮挡情况,提高了检测结果的准确性。2) 提出了一种多特征无监督学习遮挡检测框架,利用该框架不但可以结合实际问题需要灵活更换特征的种类和数量,而且可以依据不同特征的性能差异自动更新其对应的权重。3) 首次将 Online Boosting 思想与无监督在线学习思想相结合实现了视频序列中遮挡边界的检测,有效地改善了无监督在线学习在检测准确性方面的不足并扩展了 Online Boosting 思想的应用范围。实验结果表明,与现有代表最新进展的遮挡边界检测方法相比,该方法提高了遮挡边界检测的准确性。

参 考 文 献

- 1 A S Ogale, C Fermüller, Y Aloimonos. Motion segmentation using occlusions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(6): 988-992.
- 2 D Feldman, D Weinshall. Motion segmentation and depth ordering using an occlusion detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(7): 1171-1185.
- 3 Wang Wei, Yu Zhijing, Guo Yin, *et al.*. Study on the occlusion problem in dynamic space intersection measurement with multi-camera systems[J]. Acta Optica Sinica, 2014, 34(4): 0415003.
王 伟, 于之靖, 郭 寅, 等. 多摄像机空间交会动态测量中遮挡问题的研究[J]. 光学学报, 2014, 34(4): 0415003.
- 4 A N Stein, M Hebert. Local detection of occlusion boundaries in video[J]. Image and Vision Computing, 2009, 27(5): 514-522.
- 5 N Apostoloff, A Fitzgibbon. Learning spatiotemporal T-junctions for occlusion detection[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2005: 11553-11559.
- 6 A N Stein, M Hebert. Occlusion boundaries from motion: Low-level detection and mid-level reasoning[J]. International Journal of Computer Vision, 2009, 82(3): 325-357.
- 7 M E Sargin, L Bertelli, B S Manjunath, *et al.*. Probabilistic occlusion boundary detection on spatio-temporal lattices[C]. Proceedings of the IEEE International Conference on Computer Vision. IEEE, 2009: 560-657.
- 8 X He, A Yuille. Occlusion boundary detection using pseudo-depth[C]. Proceedings of the European Conference on Computer Vision. Springer, 2010: 539-552.
- 9 D Chen, Z Yuan, G Zhang, *et al.*. Detecting occlusion boundaries via saliency network[C]. Proceedings of International Conference on Pattern Recognition. IEEE, 2012: 2569-2572.
- 10 N Jacobson, Y Freund, T Q Nguyen. An online learning approach to occlusion boundary detection[J]. IEEE Transactions on Image Processing, 2012, 21(1): 252-261.
- 11 A Ayyaci, M Raptis, S Soatto. Sparse occlusion detection with optical flow[J]. International Journal of Computer Vision, 2012, 97(3): 322-338.
- 12 L Xu, J Jia, Y Matsushita. Motion detail preserving optical flow estimation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(9): 1744-1757.
- 13 D Sun, S Roth, M Black. Secrets of optical flow estimation and their principles[C]. Proceedings of the IEEE Computer Society Conference

- on Computer Vision and Pattern Recognition, IEEE, 2010: 2432–2439.
- 14 B K P Horn, B G Schunck. Determining optical flow[J]. Artificial Intelligence, 1981, 17(1–3): 185–203.
- 15 A Bruhn, J Weickert. Towards ultimate motion estimation: Combining highest accuracy with real-time performance[C]. Proceedings of the IEEE International Conference on Computer Vision, IEEE, 2005: 749–755.
- 16 Y Freund, R E Schapire, Y Singer, *et al.*. Using and combining predictors that specialize[C]. Proceedings of the Annual ACM Symposium on Theory of Computing. ACM, 1997: 334–343.
- 17 N C Oza. Online bagging and boosting[C]. IEEE International Conference on Systems, Man and Cybernetics. IEEE, 2005: 2340–2345.
- 18 H Grabner, H Bischof. On-line boosting and vision[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2006: 260–267.
- 19 P Sundberg, T Brox, M Maire, *et al.*. Occlusion boundary detection and figure/ground assignment from optical flow[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2011: 2233–2240.

栏目编辑: 苏 岑