

基于蛙眼视觉特性的运动目标模糊化区域理解跟踪方法

王智灵¹ 陈宗海^{1,2} 徐萧萧¹ 吴亮¹

摘要 动态场景下的运动目标检测与跟踪是计算机视觉研究的前沿方向,对场景的背景突变和目标的外观突变的鲁棒性是当前研究的难点所在.针对这种情形,本文提出一种基于蛙眼视觉特性的鲁棒跟踪方法.该方法利用蛙眼视觉认知的生理特性和外部特性,设计了一种与之相应的模糊化区域理解的运动目标跟踪方法.针对实验室环境下的动态序列的实验结果验证了方法的有效性;并进一步将该方法与传统的 Canny 算子理解结果及经典的 Mean shift 算法理解结果进行对比,显示了方法的优越性.

关键词 生物智能,蛙眼视觉特性,运动目标跟踪,局部场景突变,模糊化区域理解,mean shift
中图分类号 TP391

A Fuzzy Region Understanding Tactic for Object Tracking Based on Frog's Vision Characteristic

WANG Zhi-Ling¹ CHEN Zong-Hai^{1,2} XU Xiao-Xiao¹ WU Liang¹

Abstract The detection and tracking of moving objects under dynamic scenes is one of the hot topics in computer vision research. The problem is difficult when local scenes or object's appearance vary saliently. In this paper, we develop a robust intelligent tracking tactic based on the intrinsic and extrinsic features of frog's visual system. It is achieved through a so called "fuzzy region understanding". Experimental results on real lab-video and comparison with two traditional algorithms are reported to demonstrate the validity and robustness of our algorithm.

Key words Biological intelligence, frog's vision characteristic, moving object tracking, scene break, fuzzy region understanding, mean shift

近年来,针对动态混杂环境中特定/非特定运动目标的鲁棒检测与跟踪一直是计算机视觉领域中备受关注的前沿方向之一^[1-3].该问题的难点在于如何在保障实时跟踪的前提下提高对动态场景和动态目标跟踪的鲁棒性与可靠性.

在长期进化过程中所发展出来的各种生物视觉系统为解决该问题提供了最好的实例.一方面,对自然场景的各类运动目标的检测与跟踪关系着生物体的生死存亡,是生物体捕获生存所必须的各种资源、逃离敌害等行为的基本出发点;对于各种动物(如蛙类、鸟类、甚至人类等)而言,认知动态环境下的运

动目标信息的能力对其行为具有显著影响.另一方面,借鉴生物体本身的特性寻找新的跟踪策略是一种可行的途径.利用生物视觉系统的特性,研究新的鲁棒跟踪策略和跟踪方法,对于突破当前计算机视觉研究的困境,实现更为鲁棒的智能视觉系统具有重要的理论意义和现实意义.

1 动态场景下的鲁棒性需求与处理策略

自然条件下的场景丰富多样、背景复杂多变,场景中的运动目标检测与跟踪任务受到众多因素的干扰.不仅受到众多随机不确定因素的影响,比如:拍摄过程中的随机抖动、摄像机的成像质量、光照条件的变化等;还受到各种各样的特殊因素的影响,比如:由于目标形变引入的外观变化、由于目标旋转引入的特征变化、由于背景变化引入的混淆与遮挡、由于多个近似目标引入的干扰与混杂等.其中,动态场景中存在的突变现象(包括场景中的背景突变和运动目标的外观突变)是当前研究的难点所在.

传统的运动分析方法,如光流法^[4],通常认为目标的运动表现为与背景环境不同的独立运动,进而可以通过比较运动目标与背景之间的差异、检测独立运动和补偿全局运动,来达到跟踪的目的.然而,当场景中存在局部背景突变时,该方法容易导致错误的结果.这是因为背景场景的局部突变甚至超过

收稿日期 2008-01-28 收修改稿日期 2008-12-24
Received January 28, 2008; in revised form December 24, 2008
国家高技术研究发展计划(863计划)(2007AA04Z227),多媒体计算与通信教育部微软重点实验室科研基金(05071806),模式识别国家重点实验室开放课题基金(2006-3)资助
Supported by National High Technology Research and Development Program of China (863 Program) (2007AA04Z227), the Opening Fund of Microsoft Key Laboratory of Multimedia Calculation and Communication of Ministry of Education (05071806), and the Opening Task-fund of China National Laboratory of Pattern Recognition (2006-3)
1. 中国科学技术大学自动化系 合肥 230027 2. 中国科学院自动化研究所模式识别国家重点实验室 北京 100190
1. Department of Automation, University of Science and Technology of China, Hefei 230027 2. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190
DOI: 10.3724/SP.J.1004.2009.01048

期望目标的运动变化,这不仅使得传统的全局运动补偿方法失效,甚至会引入错误信息导致对背景对象的误判。

为了增强对动态场景的描述能力, Li 等^[5]构建了一个基于色彩信息的统计模型,该系统能区分静止的和固定的背景的像素,具有对渐进变化的适应能力,还能对由于光源的切断/开启状态所造成的全局性突变鲁棒;但是,该系统仍然缺乏对局部场景突变的理解能力。

目标的外观信息可以有效地削弱背景变化对跟踪结果的影响,构建期望目标的模型描述,通过模型匹配获取目标位置是另一类广为使用的手段。肤色模型和外观模型是最常使用的描述方式。Ramanan 等^[6]通过学习序列中的人的外观模型。该方法处理人运动的跟踪过程中出现的混杂和非刚性变化;Chow 等^[7]建立肤色模型,该方法对环境光照条件鲁棒;但是,该方法没有考虑目标外观突变的问题。

针对这种情形,为了处理存在局部场景突变与目标外观突变的环境下的目标跟踪问题,本文结合青蛙视觉认知的生理特性和外部特性,设计一种与之相应的模糊化区域理解的运动目标跟踪策略。

2 基于蛙眼特性的鲁棒跟踪策略

蛙眼视觉系统以其对运动目标的特殊敏感性而闻名——青蛙看不见(至少是不关注)周围世界中的静态部分细节,它甚至可能在四周环绕静止食物的环境中饥饿致死。目前大多数的研究集中于对青蛙视觉系统的神经生物学机理的研究,以及在此基础上应用计算机信息处理技术对青蛙视觉行为机制的模拟研究。蛙眼认知特性在计算机视觉领域的描述和表达模式还很不完善,蛙眼的许多其他的特殊的视觉特性还没有得到充分的利用。研究如何在运动目标检测与跟踪任务中描述和表达蛙眼的认知特性,研究利用蛙眼的认知特性来提高跟踪策略的鲁棒性与可靠性,仍然是一个很有挑战性的课题。

Lettvian^[8]在其影响深远的论文中研究了蛙眼视网膜的神经结构。在这篇文章中, Lettvian 将青蛙的视神经纤维分为四大类:持续反差检测器、凸边检测器、运动边界检测器和变暗检测器。蛙眼视觉神经对横向穿越感受野的运动凸边具有强烈的敏感性。Lettvian^[8]、Ingle^[9]等对青蛙视网膜神经结构的分析解释了青蛙的这种特殊的视觉行为。此后,研究者们构造出各种类似于蛙眼的视神经纤维功能的检测器,用于在特定环境中监控特定类型的运动目标。

除了上述生理特性外,蛙眼的外部特性也有其突出的特点。本文还将关注蛙眼视觉特性的另外两种外部特征:1)蛙眼的有效分辨距离很短,对青蛙而

言,近视的一个优势就是背景模糊而使得前景相对清晰,这使得青蛙能更准确地捕获猎物;2)蛙眼视觉系统依靠目标大小和运动类型分辨猎物和天敌,它将跳起来捕获任何昆虫或蠕虫大小的运动目标。

借鉴蛙眼的这种外部特性及其神经生理特性,本文提出了一种命名为“模糊化区域理解”的运动目标跟踪策略。与蛙眼跟踪方式类似,在该方法中,我们使用运动目标的外部轮廓作为目标的特征描述,并使其在算法过程中逐步凸显,最终从图像中给出期望的目标。该方法的基本思路就是:通过滤波、保持、合并、滤波的方式,使图像场景中的干扰结构和运动目标的内部纹理逐渐模糊,甚至消失,而保留期望的目标特征——目标外轮廓,并使之逐渐清晰、凸显,结合目标大小的先验信息,获得期望的跟踪位置。

3 跟踪算法设计

3.1 跟踪算法

记输入图像序列为 $\{I_t(\mathbf{x})|t \text{ 为帧计数}\}$,对于二维灰度图像而言, $\mathbf{x} = (x_1, x_2), I_t(\mathbf{x}) \in \mathbf{R}$,以图像空间和图像值空间构成联合特征空间 $Z_t(\mathbf{x}) = (\mathbf{x}, I_t(\mathbf{x})) \in \mathbf{R}^3$ 。首先通过各向异性滤波和 Mean shift 滤波获得对目标图像的模糊化区域理解,并根据区域理解结果获得轮廓边界,最终得到期望目标的跟踪位置。算法设计如下:

步骤 1. 获取序列中的一帧图像 $I_t(\mathbf{x})$,在图像值空间域执行各向异性的自适应平滑滤波;

步骤 2. 在联合特征空间域执行 Mean shift 滤波过程,直到收敛,并记录收敛点;

步骤 3. 根据收敛点进行区域划分,获得初步的模糊化区域理解结果;

步骤 4. 根据预先设定的区域尺度参数对区域划分结果进行融合,获得最终的区域理解结果,并在此基础上获取对跟踪图像序列的轮廓理解结果;

步骤 5. 结合目标大小的先验信息,获得期望的跟踪位置。

在上述算法过程中,各向异性滤波被用于在保持显著边界的同时平滑图像;Mean shift 滤波可以在保持期望目标结构的前提下滤除目标的内部结构和背景结构;区域融合则进一步凸显目标轮廓;而目标大小的先验信息指引了目标的期望跟踪位置。

3.2 自适应平滑滤波

为了能在平滑图像区域内部的同时仍能够保持显著的区域边缘,避免由于边界模糊化而导致的边缘位置偏移,对所获的序列图像定义其自适应平滑滤波操作。记时刻 t 获得的图像帧为 $I_t^k(x_1, x_2)$,定义 $T_t(x_1, x_2)$ 为图像梯度的幅值函数,即

$$T_t(x_1, x_2) = \sqrt{\left(\frac{\partial I_t(x_1, x_2)}{\partial x_1}\right)^2 + \left(\frac{\partial I_t(x_1, x_2)}{\partial x_2}\right)^2} \quad (1)$$

自适应平滑后的信号 $I_t^{k+1}(x_1, x_2)$ 为

$$I_t^{k+1}(x_1, x_2) = \frac{\sum_{i=-1}^1 \sum_{j=-1}^1 I_t^k(x_1 + i, x_2 + j) w_t(x_1 + i, x_2 + j)}{\sum_{i=-1}^1 \sum_{j=-1}^1 w_t(x_1 + i, x_2 + j)} \quad (2)$$

式中, 权函数 $w_t(x_1, x_2)$ 为

$$w_t(x_1, x_2) = \exp \frac{-(T_t(x_1, x_2))^2}{2\sigma^2}$$

这是一种非线性的滤波方式^[10], 实际上是各向异性滤波^[11]的一种离散化形式; 因此具有与各向异性滤波相似的性质, 它可对二维图像信号执行保持边缘的平滑.

3.3 Mean shift 滤波

Mean shift 算法是由 Fukunaga 和 Hostetler^[12]提出的一种基于核密度估计的自适应梯度算法, 它从任意一点出发, 沿着核密度的梯度上升方向, 以自适应的步长进行搜索, 最终收敛于核密度估计函数的局部极大值处. 自从 Cheng^[13]的论文发表以后, Mean shift 算法在目标跟踪、模式识别与分析等领域开始受到研究者的重视.

为了进一步将期望目标外轮廓从图像中凸显出来, 应用 Mean shift 滤波过程实现保持目标结构的区域模糊化操作.

对于 $I_t^{k+1}(x_1, x_2)$ 中的每一个像素, 其联合特征空间为 $Z_t(\mathbf{x}) = (\mathbf{x}, I_t(\mathbf{x})) \in \mathbf{R}^3$, 记特征矢量为 $\mathbf{z} = Z_t(x_1, x_2)$, 其概率密度函数的核密度估计为

$$\hat{f}(\mathbf{z}) = \frac{1}{n} |H|^{-1/2} \sum_{i=1}^n K \left(\frac{\mathbf{z} - \mathbf{z}_i}{|H|^{-1/2}} \right) = \frac{1}{n} |H|^{-1/2} \sum_{i=1}^n k \left(\left\| \frac{\mathbf{z} - \mathbf{z}_i}{|H|^{-1/2}} \right\|^2 \right) \quad (3)$$

式中, $K(\cdot)$ 是高斯核的核函数, $k(\cdot)$ 是 $K(\cdot)$ 的剖面函数, $K(\cdot) = k(\|\cdot\|^2)$, H 为尺度参数矩阵.

尺度参数是决定 Mean shift 滤波的重要参数, 它不仅决定了参与迭代的采样点数量, 还将影响算法的收敛速度和收敛的准确性. 本文算法对尺度参数进行进一步解耦, 采用图像空间尺度与图像值空间尺度解耦方式^[14], 此时, 核函数化为如下形式

$$K_{h_s, h_c}(\mathbf{z}) = \frac{C}{h_s^2 h_c} k \left(\left\| \frac{\mathbf{z}^s}{h_s} \right\|^2 \right) k \left(\left\| \frac{\mathbf{z}^c}{h_c} \right\|^2 \right) \quad (4)$$

式中, h_s 、 h_c 分别为解耦后的空间尺度参数和值空间尺度参数, C 为归一化常数. 在联合特征空间内, 以空间尺度参数 h_s 、值空间尺度参数 h_c 所决定的超椭球体为 Mean shift 迭代的样本区域. h_s 、 h_c 共同决定了模糊化区域理解的效果. 在实验部分, 将进一步分析尺度参数选择对理解结果的影响.

将式 (4) 代入式 (3), 得到概率密度函数的新的估计公式

$$\hat{f}(\mathbf{z}) = \frac{c_k}{h_s^2 h_c} \sum_{i=1}^n k \left(\left\| \frac{\mathbf{z}^s - \mathbf{z}_i^s}{h_s} \right\|^2 \right) k \left(\left\| \frac{\mathbf{z}^c - \mathbf{z}_i^c}{h_c} \right\|^2 \right) \quad (5)$$

进一步导出对概率密度函数梯度值的估计

$$\hat{\nabla} f(\mathbf{z}) \approx \nabla \hat{f}(\mathbf{z}) = \begin{pmatrix} \nabla_{\mathbf{z}^s} \hat{f}(\mathbf{z}) \\ \nabla_{\mathbf{z}^c} \hat{f}(\mathbf{z}) \end{pmatrix} = \begin{pmatrix} \frac{2c_k}{h_s^2 h_c h_s^2} \sum_{i=1}^n (\mathbf{z}_i^s - \mathbf{z}^s) \frac{\partial}{\partial \mathbf{z}^s} k \left(\left\| \frac{\mathbf{z}_i^s - \mathbf{z}^s}{h_s} \right\|^2 \right) k \left(\left\| \frac{\mathbf{z}_i^c - \mathbf{z}^c}{h_c} \right\|^2 \right) \\ \frac{2c_k}{h_s^2 h_c h_c^2} \sum_{i=1}^n (\mathbf{z}_i^c - \mathbf{z}^c) k \left(\left\| \frac{\mathbf{z}_i^s - \mathbf{z}^s}{h_s} \right\|^2 \right) \frac{\partial}{\partial \mathbf{z}^c} k \left(\left\| \frac{\mathbf{z}_i^c - \mathbf{z}^c}{h_c} \right\|^2 \right) \end{pmatrix} \quad (6)$$

令

$$g(\mathbf{z}) = -\frac{d}{d\mathbf{z}} k(\mathbf{z}) = \begin{pmatrix} -\frac{\partial}{\partial \mathbf{z}^s} k(\mathbf{z}^s) k(\mathbf{z}^c) \\ -k(\mathbf{z}^s) \frac{\partial}{\partial \mathbf{z}^c} k(\mathbf{z}^c) \end{pmatrix} \quad (7)$$

由式 (6) 和式 (7) 可得

$$\left(\nabla \hat{f}(\mathbf{z}) \right) = \begin{pmatrix} \frac{2c_k}{h_s^2 h_c} \sum_{i=1}^n \frac{g \left(\left\| \frac{\mathbf{z}_i^s - \mathbf{z}^s}{h_s} \right\|^2 \right)}{h_s^2} \left(\frac{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}_i^s - \mathbf{z}^s}{h_s} \right\|^2 \right) \mathbf{z}_i^s}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}_i^s - \mathbf{z}^s}{h_s} \right\|^2 \right)} - \mathbf{z}^s \right) \\ \frac{2c_k}{h_s^2 h_c} \sum_{i=1}^n \frac{g \left(\left\| \frac{\mathbf{z}_i^c - \mathbf{z}^c}{h_c} \right\|^2 \right)}{h_c^2} \left(\frac{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}_i^c - \mathbf{z}^c}{h_c} \right\|^2 \right) \mathbf{z}_i^c}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}_i^c - \mathbf{z}^c}{h_c} \right\|^2 \right)} - \mathbf{z}^c \right) \end{pmatrix} \quad (8)$$

在上式中, 令

$$\nabla \mathbf{m}_{h, g}(\mathbf{z}) = \frac{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}_i - \mathbf{z}}{h} \right\|^2 \right) \mathbf{z}_i}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}_i - \mathbf{z}}{h} \right\|^2 \right)} - \mathbf{z} \quad (9)$$

$\nabla \mathbf{m}_{h, g}$ 即为核函数在该像素处的 Mean shift 增量. 通过迭代, 该过程将收敛到各区域的模式中心

$$y(\mathbf{z}) = \frac{\sum_{j=1}^m n_j \mathbf{z}_j^r g \left(\left\| \frac{\mathbf{z} - \mathbf{z}_j^r}{h} \right\|^2 \right)}{\sum_{j=1}^m g \left(\left\| \frac{\mathbf{z} - \mathbf{z}_j^r}{h} \right\|^2 \right)} \quad (10)$$

式中, $\{z_j^r | r = 1, 2, \dots, m\}$ 是最终形成的区域的聚类中心.

Mean shift 过程具有良好收敛性质, 关于 Mean shift 算法收敛到局部极大值的条件可以参考文献 [14–15].

3.4 区域融合与跟踪位置获取

除了空间尺度参数和值空间尺度参数外, 本方法引入另一类型的尺度参数—区域尺度参数 h_r , 对 Mean shift 收敛结果做进一步处理

$$h_r = \inf(\sharp(\text{Region})) \quad (11)$$

式中, $\sharp(\text{Region})$ 为区域所包含的像素数目, h_r 是所允许的最小区域面积. 如果获得的收敛区域的实际面积小于 h_r , 则将该收敛区域同与之相邻的区域合并, 简化区域理解结果中的图像轮廓. 区域尺度参数需要根据目标大小的先验信息事先设定.

根据区域分割的结果, 按照式 (12) (见本页下方) 快速获取对模糊化区域分割的轮廓理解 $L_t(x_1, x_2)$, 式中, $S(x_1, x_2)$ 为 (x_1, x_2) 处的区域聚类标号.

在此基础上, 利用目标大小在跟踪序列中不发生显著变化的先验信息, 以期目标收敛中心为中心点, 做该区域轮廓的外接矩形, 提取运动目标的期望位置.

4 实验与分析

为验证方法的性能, 首先针对实验室场景的单帧图像进行分析, 给出模糊化区域理解结果并分析 Mean shift 滤波过程的尺度参数选择; 然后针对不同动态图像序列进行实验, 分析本文方法对场景背景变化、目标外观变化、环境光照变化等情况的适应能力; 最后, 将本文方法与两种经典的方法进行比较.

4.1 单帧图像理解与尺度参数分析

针对实验室场景下的真实环境对该方法进行实验. 首先给出本方法在单帧图像上的处理结果, 如图 1 所示. 图 1(a) 为动态场景中的原始图像帧, 图 1(b) 为基于本文算法所获得模糊化区域理解结果, 图 1(c) 为所提取的区域轮廓, 图 1(d) 给出了基于目标大小先验信息的跟踪位置, 图 1(e) 展示了最终的目标跟踪结果.

运动目标的初始跟踪位置是在模糊化区域理解结果的基础上获得的; 运动目标大小可以通过先验

信息预先设定, 在获取了有效的理解结果后, 以预设定的目标大小进行搜索, 获得目标区域的期望位置.

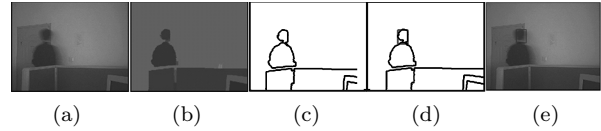


图 1 基于本文方法的单帧图像理解过程 ((a) 原图; (b) 区域理解结果; (c) 轮廓理解结果; (d) 跟踪位置; (e) 跟踪结果)

Fig. 1 Processing of a single frame based on the proposed method ((a) Original image; (b) Region understanding image; (c) Contour image; (d) Tracking location; (e) Tracking result)

尺度参数的选取对模糊化区域理解的性能具有重要影响. 从算法设计中可以看出, 区域尺度参数 h_r 不参与 Mean shift 迭代, 该参数主要用于融合模糊化区域理解结果中过于细碎的图像块. 在算法实现过程中, 可将该参数设置为一个略小于目标大小先验信息的值. 而空间尺度参数 h_s 、值空间尺度参数 h_c 共同决定了参与 Mean shift 迭代的样本的选取, 直接影响到区域理解的效果. 如图 2 所示.

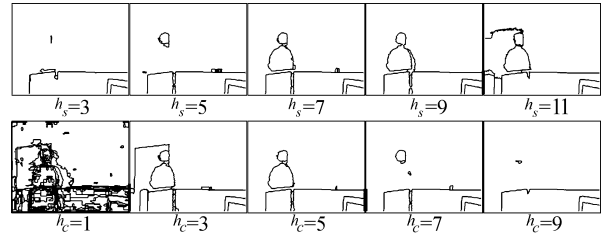


图 2 不同尺度参数对理解结果的影响

Fig. 2 The influences of scale-parameters

在图 2 中, 第一行为固定 h_c 、 h_r , 逐渐增大 h_s 所获得的理解结果, 第二行为固定 h_s 、 h_r , 逐渐增大 h_c 所获得的理解结果. 从图 2 可以看出, h_s 过小时, 无法形成对目标的整体认识; h_s 过大时, 有可能在理解结果中引入额外的对象. h_c 过小, 将导致区域理解的结果支离破碎, 期望目标轮廓将被淹没; h_c 过大, 将导致得到的区域理解结果过于宽泛, 而使得期望目标轮廓在模糊化的过程中被吞噬.

通常, h_s 、 h_c 使用的是固定的全局性的尺度参数, 该参数可以通过经典的全局性尺度估计方法^[16]进行估计, 也可以针对不同的图像序列, 对参数进行手动调整. 但是, 在实际图像中, 由于各部分的边缘强度的大小是不一致的, 并且各部分的噪声的尺度往往也不相同, 为获得更好的区域理解效果, 我们设

$$L_t(x_1, x_2) = \begin{cases} 1, & \sum_{i=-1}^1 \sum_{j=-1}^1 \|S(x_1 + i, x_2 + j) - S(x_1 - i, x_2 - j)\| > 0 \\ 0, & \text{其他} \end{cases} \quad (12)$$

计一种自适应的局部尺度参数的选择策略: 对于图像中的每一个像素, 以全局性的尺度参数作为 h_s 、 h_c 的初始估计值, 考察由 h_s 、 h_c 所决定的邻域内的样本灰度差的直方图分布, 若:

1) 该邻域内的样本灰度差为单峰分布, 则保持 h_c 不变而增大 h_s ;

2) 该邻域内的样本灰度差为多峰分布, 则保持 h_s 不变而选择 h_c 为灰度差直方图中绝对值最小的谷点.

从而使 h_s 、 h_c 可以根据各点的局部结构自适应的变化, 在模糊化局部背景和保持目标内部结构的同时, 保持图像中重要的边缘结构.

4.2 图像序列理解与目标跟踪

为考察该方法对运动目标跟踪的效果, 研究该方法对背景变化和外观变化的适应能力, 分别针对静态和动态背景的实验室场景序列进行实验.

首先给出静态环境下的跟踪结果, 如图 3 所示. 在图 3 中, 第一行为原始图像序列, 从左到右依次为动态序列中的第 10、20、30、40 帧, 第二行为本文方法最终获得的跟踪位置. 该序列中背景场景固定, 光照条件明亮, 以运动中的人物头部为跟踪目标. 序列中的主要干扰来自运动人物的转身动作, 使得头部的背面/右侧面/正面/左侧面交替出现, 目标的外观特征在运动过程中存在显著变化.

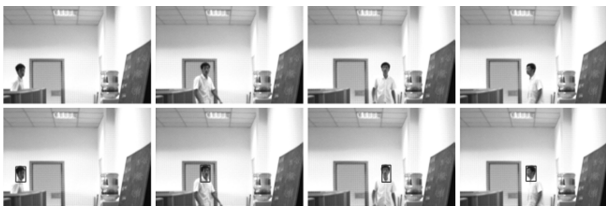


图 3 静态场景中人的头部跟踪

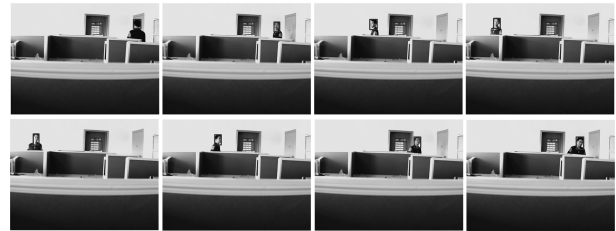
Fig. 3 Tracking under lab environment

为进一步研究该方法对人的旋转过程造成的外观变化和光照导致的外观变化的适应能力, 给出跟踪实验如图 4 所示.

图 4 (a) 给出了目标对象在场景中心顺时针旋转一圈的跟踪过程; 图 4 (b) 给出了目标对象从场景右侧走至左侧并转身走向右侧的跟踪过程. 在图 4 给出的两个跟踪过程中, 都存在由于人的旋转导致的目标外观变化. 在图 4 (a) 中, 人的旋转动作是主要的考察对象; 在图 4 (b) 中, 除了人的转身动作引起的外观变化外, 还受到环境光照渐变的影响, 从场景最右端到场景最左端, 光照逐渐增强. 跟踪结果表明, 本文方法对这两种原因导致的目标外观变化均有很好的适应能力.



(a) 场景 1
(a) Scene 1



(b) 场景 2
(b) Scene 2

图 4 背景不变时的目标跟踪

Fig. 4 Tracking under silent background

接下来考察本文方法在动态场景中的跟踪效果, 如图 5 所示. 在图 5 中, 第一行为原始图像序列, 从左到右依次为动态序列中的第 10、11、30、50 帧, 第二行为本文方法最终获得的跟踪位置.



图 5 动态场景中人的头部跟踪

Fig. 5 Tracking under dynamic lab environment

图 5 所示的环境, 光照略显昏暗, 同样以运动中的人物头部作为跟踪目标. 在运动过程中, 人的正面/侧面/背面交替出现, 目标的外观特征变化仍然是序列跟踪的主要干扰之一. 另一方面, 突变的动态背景是该序列跟踪的难点所在. 由于门的开/关动作使得背景中的局部场景发生了突变, 这种突变即使在相邻两帧中也存在 (第 10、11 帧). 两者都使得传统的基于运动分析和模板匹配的跟踪方法难以适用.

图 5 针对的是背景中的局部场景发生突变时的跟踪效果, 而当背景场景随着目标运动而整体变化时的跟踪效果如图 6 (见下页) 所示.

图 6 (a) 给出了环境光照强度保持不变时, 运动背景下的目标跟踪效果; 图 6 (b) 给出了环境光照强度变化时, 运动背景下的目标跟踪效果. 在图 6 给出的两个跟踪过程中, 都存在两种主要的干扰: 1) 因背景跟随目标运动而变化导致的干扰; 2) 因目标人物的行为动作引起外观变化而导致的干扰. 由图 6

可以看出, 由于采用了模糊化区域理解的方法, 背景运动对目标跟踪并不产生严重影响; 而且, 与静态环境下的跟踪结果类似, 本文方法可以很好地处理由目标的行为动作引起的外观变化. 在图 6 (b) 中还可以看出, 在该序列中, 当被跟踪从实验室大门 (右侧) 走向窗户 (左侧) 的时候, 不仅背景区域因运动而不同, 而且, 随着目标接近窗户, 环境光照强度逐渐增加; 而当目标从窗户 (左侧) 返回大门 (右侧) 时, 环境光照强度逐渐减弱.

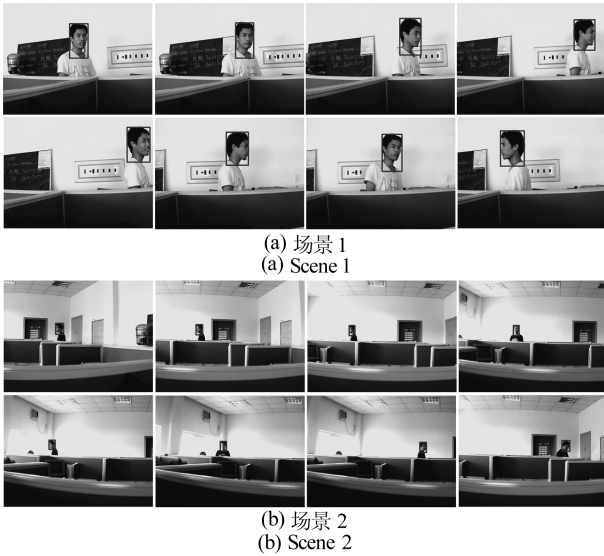


图 6 背景运动时的目标跟踪

Fig. 6 Tracking under dynamic background

从图 3~6 的视频序列跟踪结果可以看出, 本文方法在场景中存在背景变化和目标外观变化时均能给出有效的理解和跟踪.

4.3 比较实验

最后, 针对图 5 所给出的动态场景, 将本文所提出的模糊化区域理解结果与传统的 Canny 边缘检测结果及经典的 Mean shift 算法结果^[14] 进行对比, 如图 7 所示. 图 7 (a) 从上到下依次为动态场景序列中的各帧图像; 图 7 (b) 给出了传统的 Canny 算子对该序列的检测结果; 图 7 (c) 给出了按照经典 Mean shift 算法的检测结果; 图 7 (d) 列给出了基于本文方法的模糊化区域理解结果.

从图 7 可以看出, 传统的 Canny 算子检测结果给出了几乎全部的有意义的和无意义的图像边界, 其结果难以进一步理解; 经典的 Mean shift 算法关注图像当中较为细致的色彩特征变化, 倾向于提取图像灰度的细节变化; 而本文结果提取了运动中的人物对象的轮廓, 理解结果简单明了, 基本上不受目标内部结构的影响, 不受突变的背景局部和头部姿态的影响, 体现了很好的鲁棒性.

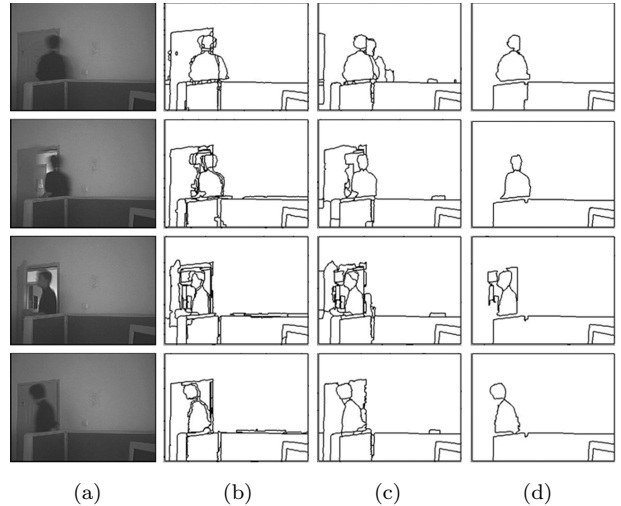


图 7 本文算法与两种经典算法的比较 ((a) 实际场景; (b) 传统 Canny 算子结果; (c) 传统 Mean-Shift 结果; (d) 本文方法结果)

Fig. 7 Comparison of our algorithm with two classical algorithms ((a) Original scene; (b) Results by Canny operator; (c) Results by traditional mean shift; (d) Results by our algorithm)

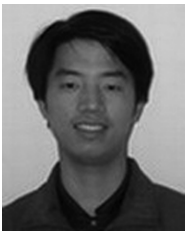
5 结语

本文提出了一种基于蛙眼视觉特性的鲁棒跟踪策略, 该方法的主要贡献有: 1) 该方法借鉴蛙眼认知的外部特性和生理特性, 提出了一种与之相适应的鲁棒跟踪策略; 2) 该方法通过“模糊化区域理解”的方式, 给出了这种跟踪策略的一种实现模式; 3) 该方法对存在背景变化及目标外观变化的动态场景中的运动目标跟踪有很好的鲁棒性. 跟踪实验以及与两种经典方法的对比实验验证了方法的有效性.

References

- 1 Heikkila M, Pietikainen M. A texture-based method for modeling the background and detecting moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(4): 657-662
- 2 Thirde D, Borg M, Aguilera J, Wildenauer H, Ferryman J, Kampel M. Robust real-time tracking for visual surveillance. *Eurasip Journal on Advances in Signal Processing*, 2007, **2007**(1): 1-13
- 3 Veeraraghavan H, Schrater P, Papanikolopoulos N. Robust target detection and tracking through integration of motion, color, and geometry. *Computer Vision and Image Understanding*, 2006, **103**(2): 121-138
- 4 Zheng J, Li B, Zhou B, Li W. Fast motion detection based on accumulative optical flow and double background model. In: *Proceedings of International Conference on Computation*. Xi'an, China: Springer, 2005. 291-296
- 5 Li L Y, Huang W M, Gu I Y H, Tian Q. Statistical modeling of complex backgrounds for foreground object detection. *IEEE Transactions on Image Processing*, 2004, **13**(11): 1459-1472

- 6 Ramanan D, Forsyth D A, Zisserman A. Tracking people by learning their appearance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(1): 65–81
- 7 Chow T Y, Lam K M. Mean-shift based mixture model for face detection in color image. In: *Proceedings of International Conference on Image Processing*, Washington D. C., USA: IEEE, 2004. 601–604
- 8 Lettvin J Y, Maturana H R, McCulloch W S, Pitts W H. What the frog's eye tells the frog's brain. *Proceedings of the IRE*, 1959, **47**(11): 1940–1951
- 9 Ingle D. Disinhibition of tectal neurons by pretectal lesions in the frog. *Science*, 1973, **180**(84): 422–424
- 10 Saint-Marc P, Jer-Sen C, Medioni G. Adaptive smoothing: a general tool for early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, **13**(6): 514–529
- 11 Barash D. A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(6): 844–847
- 12 Fukunaga K, Hostetler L D. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Informations Theory*, 1975, **21**(1): 32–40
- 13 Cheng Y Z. Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, **17**(8): 790–799
- 14 Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 603–619
- 15 Wen Zhi-Qiang, Cai Zi-Xing. Convergence analysis of mean shift algorithm. *Journal of Software*, 2007, **18**(2): 205–212 (文志强, 蔡自兴. Mean Shift 算法的收敛性分析. *软件学报*, 2007, **18**(2): 205–212)
- 16 Sheather S J, Jones M C. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society*, **53**(3): 683–690



王智灵 中国科学技术大学自动化系博士研究生. 2003 年获中国科学技术大学工学学士学位. 主要研究方向为计算机视觉、鲁棒估计、模式识别和生物认知.
E-mail: zlwang3@mail.ustc.edu.cn
(**WANG Zhi-Ling** Ph.D. candidate in the Department of Automation,

University of Science and Technology of China (USTC). He received his bachelor degree from USTC in 2003. His research interest covers computer vision, robust estimation, pattern recognition, and biological intelligence.)



陈宗海 教授. 主要研究方向为复杂系统的建模、仿真与控制、智能机器人和量子信息控制. 本文通信作者.

E-mail: chenzh@mail.ustc.edu

(**CHEN Zong-Hai** Professor in the Department of Automation, University of Science and Technology of China.

His research interest covers modeling, simulation and control of complex system, intelligent robot, and quantum control. Corresponding author of this paper.)



徐萧萧 中国科学技术大学自动化系博士研究生. 2005 年获中国科学技术大学工学学士学位. 主要研究方向为计算机视觉和模式识别.

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

(**XU Xiao-Xiao** Ph.D. candidate in the Department of Automation, USTC.

He received his bachelor degree from USTC in 2005. His research interest covers computer vision and pattern recognition.)



吴亮 中国科学技术大学自动化系硕士研究生. 2005 年获南京理工大学工学学士学位. 主要研究方向为计算机视觉, 视频运动分割和特征识别.

E-mail: wuliang3@mail.ustc.edu.cn

(**WU Liang** Master student in the Department of Automation, USTC. He

received his bachelor degree from Nanjing University of Science and Technology in 2006. His research interest covers computer vision, motion detection in video, and character recognition.)