

# 基于差分图象的多运动目标的检测与跟踪

王 栓 艾海舟 何克忠

(清华大学计算机科学与技术系, 北京 100084)

(清华大学智能技术与系统国家重点实验室, 北京 100084)

**摘 要** 运动目标的检测与跟踪在许多领域有着广泛的应用,它是应用视觉研究的焦点之一。文中介绍了一种基于差分图象的运动目标检测算法,检测结果是符号化了的图象,其中运动目标由其外接矩形表示,然后根据连续性约束假设,实现了运动目标的跟踪。实验表明这种检测和跟踪方法是快速有效的,与现有其它方法相比它能更好地处理跟踪目标之间的重叠以及目标的暂时消失等情况。

**关键词** 差分图象 运动检测 目标跟踪

## 0 引 言

运动目标的检测与跟踪是应用视觉研究领域的一个重要课题。在现实生活中,大量的有意义的视觉信息包含在运动之中,甚至有些动物的眼睛经过进化,只能看见运动的物体。尽管人类视觉既能看见运动又能看见静止的物体,但是在许多场合,比如说交通流量的监测、重要场所的保安、航空和军用飞行器的制导、汽车的自动驾驶或辅助驾驶等,我们往往对运动的物体更感兴趣。因此研究只对运动目标敏感的检测与跟踪系统是很有意义的。此外,运动目标的研究对象是图象序列,而对图象序列的研究一般要比对单帧图象作静态分析容易。

一般而言,运动检测与跟踪研究的难点在于:

- (1) 运动物体的正确检测与分割;
- (2) 物体之间的遮挡与重叠;
- (3) 初始运动参量的确定。

关于运动目标的检测与跟踪研究大致可分为两类:

- (1) 摄像头随着运动目标移动,始终保持目标在图象的中心附近<sup>[1-3]</sup>;
- (2) 摄像头固定,只对视场内的目标进行跟踪<sup>[4-6]</sup>。

已有的算法包括基于模板匹配的算法<sup>[9]</sup>、基于镶嵌图的算法<sup>[6]</sup>、基于主动轮廓线的算法<sup>[4,8]</sup>等,多

数是对目标模型的直接匹配,实际上是对静态(目标)图象的处理分析,显然计算量大,难以实时应用。另外还有一些针对点目标的跟踪算法,如 R. Mehrotra<sup>[7]</sup>对图象序列中的兴趣点进行匹配以确定运动轨迹的算法:依据运动轨迹尽可能光滑、位移尽可能小的假设在相邻帧中对点进行匹配。该方法的缺点是不能处理点的重叠(即使一个点被多个运动轨迹所匹配)以及点的暂时消失等情况。实际上,这是许多现有的跟踪方法的共同问题。另外需指出的是,把目标作为点来处理过于简单化。我们的工作介于他们之间,主要研究了在固定摄像头的情况下对视场内的目标进行跟踪的问题,首先将运动区域用外接矩形表示(相当于符号化),然后考虑对这些矩形序列进行匹配跟踪,所实现的系统能够处理跟踪目标之间短暂(秒的数量级)重叠或者目标暂时消失情况下的跟踪问题,而且具有准实时性。所谓目标的暂时消失是指在序列的若干帧中某些跟踪目标由于种种原因没有被检测出来,但它实际上并没有运动出视场的情况。所用的核心方法是基于运动预测的最近邻法。

## 1 运动目标的检测算法和实验

在应用视觉系统中,检测运动目标常用差分图象的方法,一般有两种情况:

- (1) 当前图象与固定背景图象之间的差分;
- (2) 当前连续两幅图象(时间间隔  $\Delta t$ )之间的差分。

第1种情况即固定背景法检测物体的优点是位置精确、速度快,因为它只需获取当前的一幅图象。不足之处是受环境光线变化的影响,在非受控环境下需要加入背景图象更新机制,且不适用于摄像头运动或者背景灰度变化很大的情况。

第2种情况即相继图象差分法的优点是它只对运动物体敏感,实际上它只检测相对运动的物体,而且因2幅图象的时间间隔较短,差分图象受光线变化影响小,检测有效而稳定;缺点是检测出的物体的位置不精确,其外接矩形在运动方向上被拉伸,这实际上是由相对运动与物体位置并非完全一致引起的。相对运动受物体本身的运动速度及相继图象之间的时间间隔影响。C. Viererl<sup>[4]</sup>提出了一种获取运动物体精确位置的方法,但该方法需要连续3帧图象,并要计算梯度图象,算法的计算量大且有一帧时

间的滞后性。

运动目标的检测过程如图1所示,其结果由运动目标的外接矩形表示。图2给出了一组在固定背景情况下的检测实验图片,图片大小是  $128 \times 128$ 。

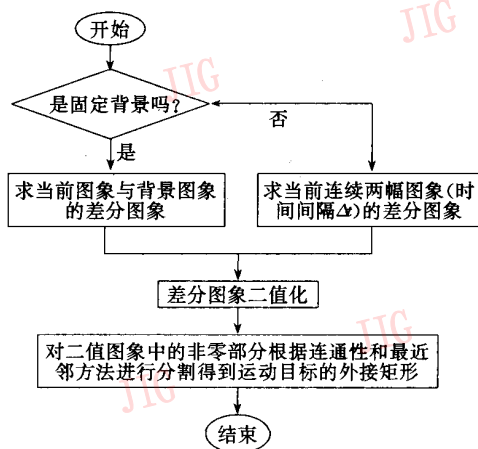


图1 多运动目标的检测算法流程

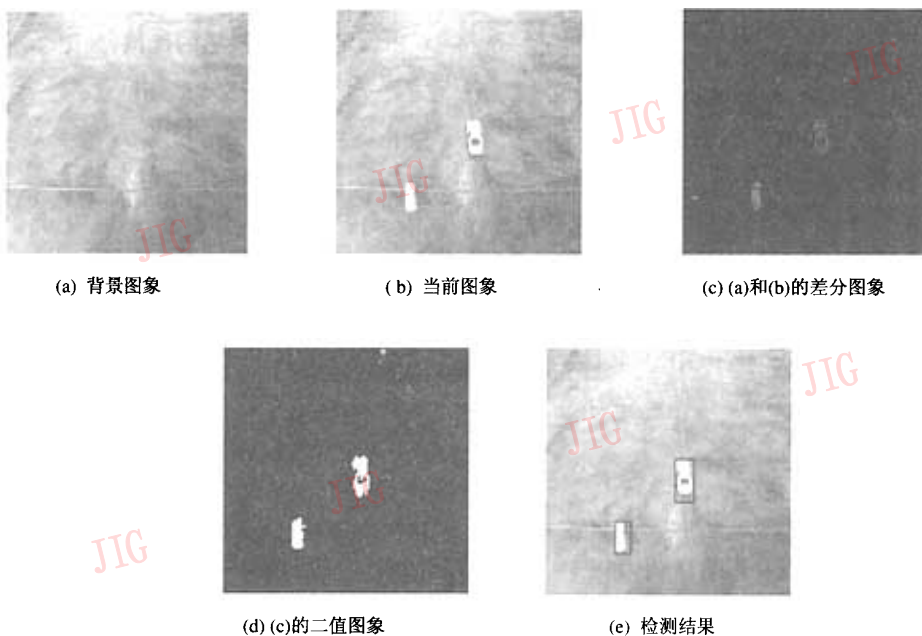


图2 运动目标检测

## 2 运动目标跟踪的算法和实验

运动目标的跟踪是确定同一物体在不同帧中位置的过程,当运动物体被正确检测出来时,它就是对相邻帧中检测出的物体进行匹配的问题,匹配的依据主要是物体的大小和位置关系。假设在一帧的时

间间隔内跟踪目标的运动轨迹是平滑的,即它的运动参量(速度和加速度)的改变应该尽可能的小。

对  $n$  帧图象(即对应的矩形数组)序列进行跟踪的步骤如下,其细节将随后分别描述。

- (1) 置运动序列为空;
- (2) 从第1帧到第  $n$  帧图象做:① 为每个未置出界标志的运动序列在当前帧中寻找最佳匹配;②

如果当前帧中有未与任何运动序列匹配的物体,并且它不是其它物体的碎片,那么为它初始化一个新的运动序列。

(3) 结束。

### 2.1 确定跟踪目标的运动参量

假设当前帧  $C$  中物体  $M$  已被确定属于某运动序列,该运动序列中最后一个物体是第  $L$  帧中的  $N$ ,那么  $M$  的速度  $V_{C,M}$  和加速度  $a_{C,M}$  定义为:

$$V_{C,M} = (P_{C,M} - P_{L,N})/\Delta t$$

$$a_{C,M} = (V_{C,M} - V_{L,N})/\Delta t$$

其中  $P_{X,Y}$  表示第  $X$  帧中  $Y$  物体的位置,  $\Delta t$  为  $C$  和  $L$  之间的时间间隔。

由于运动参量在跟踪过程中起十分重要的作用,使得初始运动参量的估计特别重要。我们假设新出现的目标在此后最初的几帧中不连续发生重叠和消失。这个要求是合理的,符合实际情况。对于新出现的目标,在随后最初的几帧中找匹配时,因为无法估计运动参数,我们暂时不考虑其运动参数,直接根据它与待匹配物体在空间上的近邻性以及大小差异的程度,在整个帧中找最佳匹配,若匹配程度超过一选定阈值,则该匹配作为新运动序列的候选且该帧作为新目标出现的另 1 帧图象,直到找到新目标出现的 3 帧图象时,再根据如下的匹配平滑性度量进行重新匹配具体确定新目标的最初跟踪序列。

设新目标出现的 3 帧图象是  $f_{i-2}, f_{i-1}, f_i$ , 假设在  $f_{i-1}$  中与新目标匹配的物体是  $M$ , 在  $f_i$  中与新目标匹配的物体是  $N$ , 那么它的匹配平滑性度量定义为:

$$P'_{M,N} = 1 - \frac{|V_{f_{i-1},M} - V_{f_i,N}|}{\max_{x \in f_{i-1}, y \in f_i} (|V_{f_{i-1},x} - V_{f_i,y}|)}$$

在图象  $f_{i-1}, f_i$  中找出使  $P'_{M,N}$  最大者,如果该最大值大于阈值,则更新运动序列及运动参量,完成了运动序列的初始化过程,否则在以后的帧中按上述方法继续寻找最佳匹配。

### 2.2 确定最佳匹配

运动序列经过初始化过程之后,对某个特定运动序列  $S$  在当前帧  $C$  中进行最佳匹配的过程如下:

(1) 确定匹配候选区域。根据  $S$  中最后一个物体  $N$  (非重叠) 的位置和运动参量,预测出它在  $C$  中的位置,由此确定一个候选区域  $R$ , 下面的匹配在  $R$  中进行。

(2) 求匹配区域中物体的匹配概率。对在  $R$  中

的运动物体  $M$ , 计算它们与  $S$  相匹配的概率  $p_{C,M,S}$ 。  
 $p_{C,M,S}$  定义为两个因素的加权和,这两个因素分别是:

① 根据与预测位置接近程度而得出的平滑匹配概率  $p_{C,M,S}^d$ 。

② 根据它们的相对大小而得出的大小匹配概率  $p_{C,M,S}^s$ 。

于是  $p_{C,M,S}$  可表示为

$$p_{C,M,S} = w_d p_{C,M,S}^d + w_s p_{C,M,S}^s$$

其中  $w_d$  和  $w_s$  分别为平滑匹配概率和大小匹配概率的权重。

(3) 找出匹配概率最大者。在  $R$  的所有物体中找出  $p_{C,M,S}$  最大者  $F$ , 如果  $p_{C,M,S}$  大于阈值且  $F$  未与其它运动序列匹配过,那么可将  $F$  加入  $S$  中,否则跟踪目标可能出现重叠、暂时消失或者运动出视场的现象。

下面对以上三步作详细说明:

(1) 确定候选区域,就是根据  $N$  的运动参量,在  $C$  中确定一个椭圆形的区域,如图 3 所示。图中点  $P_1$  是  $L$  帧中  $N$  的位置,点  $P_3$  是根据  $N$  的运动参量

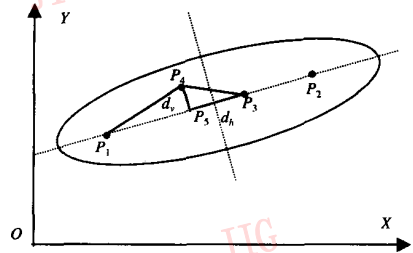


图3 候选区域的确定

预测出的该物体在  $C$  中的位置,假使坐标原点是

$$O, \text{那么 } OP_3 = OP_1 + V_{L,N} \Delta t + \frac{1}{2} a_{L,N} \Delta t^2$$

$P_2$  点在  $P_1$  和  $P_3$  的延长线上且

$$|P_1 P_2| = 1.5 |P_1 P_3|$$

$P_1$  和  $P_2$  是椭圆的焦点,再根据实际情况确定椭圆的焦距就可定出候选区域,以后的匹配就在该区域中进行。 $P_4$  是该物体在  $C$  中的实际位置。

(2) 物体的匹配概率,如前所述,它应为平滑匹配概率和大小匹配概率的加权和。

**平滑匹配概率** 它用来表示跟踪目标运动参量的连续性或目标运动的可预测程度。如果目标运动的连续性越好,即它运动的可预测性越好,则我们根

据它以前的运动参量预测出它在当前帧中的位置越精确。鉴于此,我们用当前帧  $R$  中物体的实际位置与预测位置的偏差程度来确定平滑匹配概率

$$P_{C,M,S}^d = 1 - d_{C,M,S}/D$$

$$d_{C,M,S} = ad_h + bd_v$$

$$D = \max_{X \in R}(d_{C,X,S})$$

其中  $d_h$ ,  $d_v$  的含义见图 4,它们分别表示物体相对于预测位置在运动方向上的偏差以及垂直于运动方向上的偏差,由于这两个方向上的偏差对运动连续性的影响不一样,因此有必要对它们乘上不同的系数,  $a$ 、 $b$  是加权系数。

**大小匹配概率** 大小匹配概率主要由运动目标  $N$  和待匹配物体的外接矩形的宽和高决定。

(3) 找出匹配概率最大者  $F$ 。在这里我们要求如果发生重叠,那么  $F$  外接矩形的大小应该大于  $N$  外接矩形的大小,且在发生重叠的情况下,不对物体的运动参量进行更新。而判断是否出界的依据是  $F$  的匹配概率小于阈值且根据  $S$  运动参量预测的位

置在视场之外,判断暂时消失的依据是  $F$  的匹配概率小于阈值但预测位置在视场之内,但对于连续两帧都消失的跟踪目标作出界处理。

在实验中,我们用玩具车作为被跟踪的物体,图 4 是运动序列中的 3 帧图象。

无论是用固定背景的方法还是用相继图象差的方法检测出的运动物体,都是对应于一系列的矩形框,即图象已经被符号化了,图 5 给出了两辆车的运动序列的矩形框,及其匹配跟踪的结果,跟踪的结果通过每次只回放一个运动序列的方式显示出来。从图中可以看出运动序列 1 和 2 在第 4、5 帧发生重叠,即它们被同一个矩形框所代表。

图 6 是三辆车在视场中运动的情况,在图 6(b) 中所示的第 1 个运动序列由于开始几帧存在噪音干扰致使检测出的前几帧存在丢帧的现象,但随后整个匹配趋于正常。在图 6(c) 所示的第 2 个运动序列中,小车在第 4 帧暂时消失,而在第 6 和第 7 帧和第

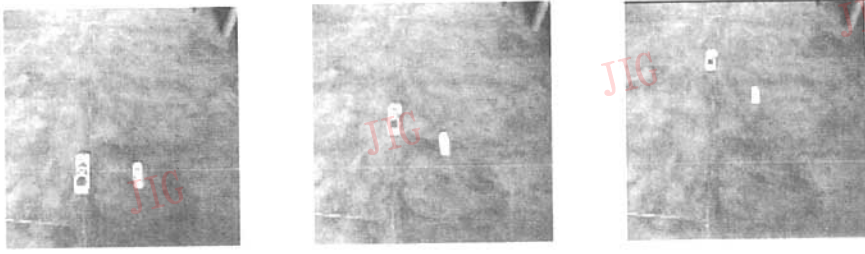
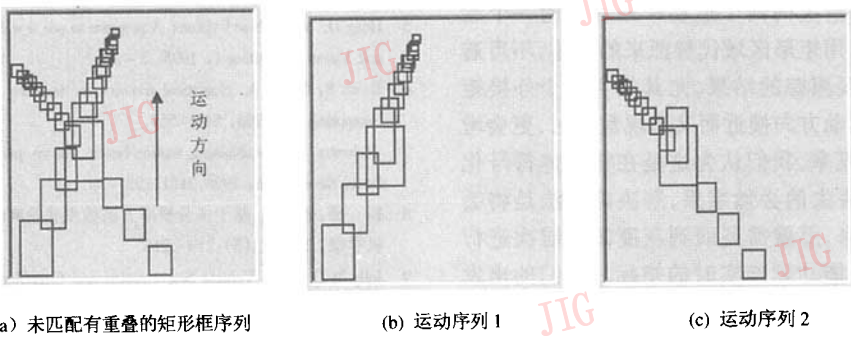


图 4 两辆玩具车从视场中平行驶过



(a) 未匹配有重叠的矩形框序列 (b) 运动序列 1 (c) 运动序列 2

图 5 两辆车有重叠时的跟踪结果

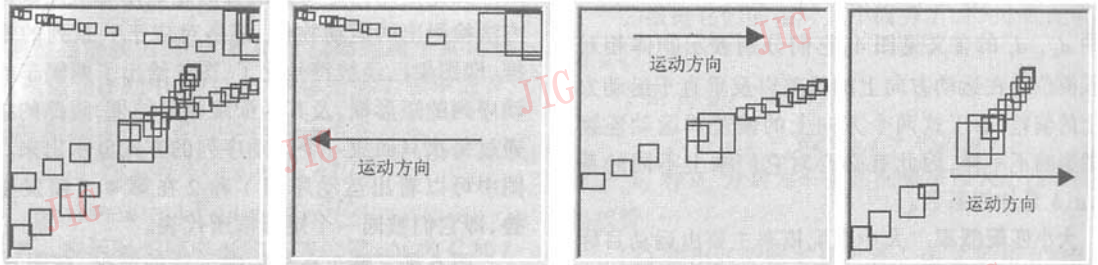
3 个运动序列中的小车发生重叠,同样在图 6(d) 所示的第 3 个运动序列中,小车在第 5 帧时消失,随后在第 6 和第 7 帧和运动序列 2 中的小车发生重叠,

图示检测出的运动结果是令人满意的。

此外,我们做了多组实验,包括仿真及实际采集的数据,对于仿真数据(人工生成的矩形数组框序列)跟

踪完全成功,对于实际采集的实验数据,如果满足文中的假设条件,那么算法也是有效的、鲁棒的,算法复杂度是多项式的,能够实现准实时的检测和跟踪,程序的处理速度与帧中运动物体的数目有关,当运动物体数目为3个左右时,速度为8~9帧/秒。本文中的图片都

是实际采集的实验。实验环境是PC机(Pentium-133MHz, 32M内存),图象技术公司IC-PCI图像采集卡,算法是用MS-VC++5.0编制实现的,本文中的图片都是实际采集的实验。



(a) 未匹配的矩形框序列

(b) 运动序列 1

(c) 运动序列 2

(d) 运动序列 3

图6 三辆车有重叠和暂时消失情况的跟踪结果

### 3 结 论

本文提出了一种基于差分图象的运动目标检测与跟踪算法,跟踪算法的优点是:不在图象层次匹配,而只在矩形框层次(符号)匹配,计算量很小,可以在运动检测过程中同步进行。我们的算法能够处理目标暂时消失,重叠和被跟踪目标运动出视场的情况,另外还能够在跟踪过程中对新出现的物体进行跟踪等。

该算法的一个缺点是将连续多帧消失的运动序列拆成两个运动序列。处理这个缺点的一个方法就是当匹配完了之后,再对所有的运动序列做一次处理,将属于这种情况的两个运动序列合并为一个运动序列。另外,用矩形区域代替原来的物体,不可避免地会带来错误跟踪的结果,尤其是当两个外接矩形大小相当,运动方向接近而又出现重叠时,更会增大错误跟踪的概率,我们认为这是在将物体符号化的过程中信息丢失的必然结果,解决的办法是将丢失的信息找回来,那就需要回到灰度图象层次进行匹配,这就不可能达到准实时的指标。我们的出发点是在现有技术条件下,从应用视觉的角度,以求获

得最大限度的速度和可靠性。实际上,个别的问题可以通过好的人机界面由人工干预解决。

### 参 考 文 献

- 1 Piccioli G, Micheli E D, Parodi P, *et al.* Robust method for road sign detection and recognition. *Image and Vision Computing* 14, 1996, 209 ~ 233.
- 2 P. Nordlund, T. Uhlin, Closing the loop: detection and pursuit of a moving object by a moving observer. *Image and Vision Computing* 14, 1996, 265 ~ 275.
- 3 Clarke J C, Zisserman A. Detection and tracking of independent motion. *Image and Vision Computing* 14, 1996, 565 ~ 572.
- 4 Vieren C, Gabestaing F, Postaire J. Catching moving objects with snakes for motion tracking. *Pattern Recognition Letters* 16, 1995, 679 ~ 685.
- 5 Hogg D. Model-based vision: A program to see a walking person. *Image and Vision Computing* 1, 1995, 5 ~ 20.
- 6 Rowe S, Blake A. Statistical mosaics for tracking. *Image and Vision Computing* 14, 1995, 549 ~ 564.
- 7 Mehrotra R. Establishing motion-based feature point correspondence. *Pattern Recognition*, 1998, 31(1): 23 ~ 30.
- 8 杨 杨, 张田文. 基于多分辨率方法的主动轮廓线跟踪算法. *计算机学报*, 1998, 21(3): 210 ~ 216.
- 9 Jolly M D *et al.* Vehicle Segmentation and Classification Using Deformable Templates. *IEEE TRANS. PAMI*, 1996, 18(3).



王 栓 1996年获清华大学现代应用物理系学士学位,现为清华大学计算机系研究生。研究方向为计算机视觉。



艾海舟 1985年、1988年、1991年分别获得清华大学计算机应用专业学士、硕士、博士学位,毕业后留校任教。1994年9月~1996年8月为比利时布鲁塞尔自由大学博士后,现为清华大学计算机系副教授。主要从事计算机视觉,模式识别,移动机器人技术的研究。发表论文20余篇。



何克忠 1962年毕业于清华大学自动控制系,现为清华大学计算机系教授。主要从事计算机控制,智能移动机器人技术的研究。先后主持研制了智能移动机器人 THMR-I、THMR-II、THMR-III,多次获部委科技进步奖,出版专著3部,发表论文40余篇。

## Difference-Image-Based Multiple Motion Targets Detection and Tracking

Wang Shuan, Ai Haizhou and He Kezhong

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

(State Key Laboratory of Intelligent Technology and Systems, Tsinghua University, Beijing 100084)

**Abstract** Motion target detection and tracking is one of the most important issues in applied vision research due to its wide applications in many fields. In this paper, a difference-image-based motion detection algorithm, which results in the symbolic representations of motion targets via their outer rectangular bounds, is presented. Then according to the continuity constraint, motion targets are tracked through image sequence. Experiments show that this method is fast and feasible which can deal with the overlapping and timely missing problems of targets in tracking more properly comparing with other existing methods.

**Keywords** Difference image, Motion detection, Target tracking