

# 一种基于视觉注意机制的改进粒子滤波跟踪算法

田 峥 徐 成 杨志邦 马翌伦  
(湖南大学信息科学与工程学院 湖南长沙 410082)

**摘 要** 目标跟踪是计算机视觉研究中的一个重要课题,它是目标行为理解的基础,是图像系统连续准确工作的重要部分。针对单一特征跟踪算法识别准确性不高,特别是在遮挡状况下无法对目标特征进行检测和跟踪的问题,考虑到粒子滤波算法在处理非线性、非高斯跟踪问题上的优越性,提出一种融合颜色、纹理和运动信息等多类特征的改进粒子滤波跟踪算法;并参考人眼的视觉注意机制,根据目标物体在不同场景下对人眼刺激的显著性不同,对目标的各个特征按照显著性强弱排序,并以此对散布粒子进行过滤。与单一特征和多特征目标跟踪算法的对比实验表明,所介绍的算法比基于单一特征的目标跟踪算法具有更高的准确性和鲁棒性,且比多特征跟踪算法的实时性更好。

**关键词** 目标跟踪 粒子滤波 多特征融合 视觉注意机制

**中图分类号** TP391 **文献标识码** A

## A VISUAL ATTENTION MECHANISM BASED IMPROVED PARTICLE FILTRATION TRACKING ALGORITHM

Tian Zheng Xu Cheng Yang Zhibang Ma Yilun

(School of Information Science and Engineering, Hunan University, Changsha 410082, Hunan, China)

**Abstract** Object tracking is a significant subject in the research of computer vision; it is not only the foundation for target behavior understandings; but also an important contributor for an image system to work continuously and accurately. Aiming at the inaccuracy at recognition for single feature tracking algorithm, especially when shaded, it is impossible to detect or track object features. Considering the superiority of particle filter algorithm at non-linear and non-Gaussian tracking problem handling, the paper proposes an improved particle filtration tracking algorithm that merges multiple features such as color, texture, and motion etc.; meanwhile, referring to human eyes visual attention mechanism, according to different significances to human eye stimulation by an object in different scenarios, sorts the object's features by significance strength so as to filter distributed particles. Comparative experiments between object tracking algorithms with single feature and multiple features indicate that the proposed algorithm is better at accuracy and robustness than single feature tracking algorithm and better at timeliness than multiple features tracking algorithm.

**Keywords** Object tracking Particle filtration Features fusion Visual attention mechanism

## 0 引 言

视觉跟踪通常指在连续变化的视频序列中对感兴趣区域或者目标对象进行精确定位。如何有效提高目标跟踪算法在复杂环境下的准确性,以及如何增加算法在多变场景下的鲁棒性,比如城市交通道路环境,一直是目标跟踪的两个关键问题,也是两大难题<sup>[1]</sup>。粒子滤波跟踪算法<sup>[2]</sup>因能有效解决非线性非高斯系统的状态估计问题而受到了广泛的重视,成为目前视觉跟踪领域研究的主流。粒子滤波是一种基于蒙特卡罗方法和递推贝叶斯估计的统计滤波方法<sup>[3]</sup>,它利用一组带权值的粒子估计和逼近目标状态的后验分布,来预测目标的当前状态;它能有效处理非线性非高斯的跟踪问题,以及处理密度函数呈多峰分布的状态预测问题,但是需要采用大量的随机样本粒子来估计,运算量很大,此外还非常依赖于相似函数的选择,且面临粒子退化和粒子枯竭的问题<sup>[4]</sup>。

近年来,很多学者围绕粒子滤波算法的不足展开了大量研究,并取得了一定的成果。文献[5]基于粒子滤波的思想,结合 Mean Shift 算法非参数概率密度估计的优点,使用 Mean Shift 方法对粒子进行收敛性分析,很好地克服了粒子退化及粒子耗尽的问题;文献[6]提出一种基于颜色信息的粒子滤波跟踪算法,能够用较少的子空间准确地表达目标的颜色分布,并能快速计算出颜色模型,有效提高了粒子滤波跟踪算法的实时性。然而,这类算法往往仅采用单一的颜色信息作为特征向量,无法持续稳健地胜任复杂多变环境下的目标跟踪问题,利用多个特征进行跟踪是提高跟踪算法鲁棒性的一种有效途径。文献[7]提出了一种多特征融合的目标跟踪算法,将颜色、纹理、运动等特征统一使用直方图模型进行描述,使算法在复杂背景下的跟踪

收稿日期:2011-09-01。2011 中国计算机大会论文。国家自然科学基金项目(60973030);湖南大学中央高校基本科研业务专项资金。田峥,博士生,主研领域:计算机视觉。

具有较高的鲁棒性;文献[8]提出一种自适应多特征融合目标跟踪算法,利用相似性度量动态地评价特征对目标与背景的区分能力,并自适应地计算特征融合权重,使算法在复杂背景下具有较高的跟踪精度。但是,这类多特征融合算法通常是以提高计算复杂度为代价,当融入过多的特征时,算法的实时性将难以保证。

自文献[9]提出基于显著性的视觉注意模型之后,视觉注意机制对视觉信息的过滤和选择作用极大地激发了学者的研究热情。文献[10]较早提出利用图像的显著性特征来改进粒子滤波跟踪算法,但是其对图像显著性特征的计算仍较复杂;文献[11,12]主要关注如何利用粒子滤波来对具备显著性特征的目标进行跟踪,但是不适用于对某个特定目标的跟踪问题;文献[13]是目前最新的将显著性特征与粒子滤波结合用于目标跟踪的方法,但是只采用了单一的颜色和亮度作为特征向量。

本文在现有的研究基础上提出了一种改进的粒子滤波跟踪算法,在采用多特征融合机制提高目标跟踪的准确性和鲁棒性的同时,模拟人类视觉注意机制,根据目标物体在不同场景下对人眼刺激的显著性不同,将各种不同特征按其在特定的场景中的显著性进行排序,并以此对散布粒子进行过滤,从而保证算法计算的实时性。

## 1 基于多特征及显著性注意机制的目标跟踪算法

本文在粒子滤波跟踪框架基础上,利用目标的颜色、纹理和运动信息建立多特征融合机制,同时考虑到数量过多的特征的提取会增加计算的复杂度从而影响目标跟踪的实时性,通过对人眼视觉目标跟踪的机制的研究和分析,提出一种改进的基于视觉注意机制的目标跟踪算法,模拟人类视觉对目标的显著性感应的机制,对使用的所有特征按所在场景的显著性进行排序,并以此来对散布粒子进行过滤,直到剩下一个最优的粒子子集,从而实现以最少的计算代价得到较高的准确性和鲁棒性。

### 1.1 多特征提取

本文提出的基于多特征的粒子滤波跟踪算法结合了目标的颜色、纹理以及运动特征,本小节主要讲述这些特征的提取方法,以及目标模型与候选区域相似度的计算方法。

#### 1.1.1 颜色特征

由于HSV色彩空间最接近人类视觉的色彩描述方式,本文采用HSV颜色空间来对颜色信息进行建模,建模的方法为归一化颜色直方图。颜色直方图的基本思想是通过统计图像内的每个像素点进行统计实现构造,通过统计得到图像内各个像素值的像素点数,通过归一化以后直方图具有尺度不变形以及旋转不变形,而且有较好的鲁棒性。同时,本文采用离散巴特查理亚(Battacharyya)距离作为运动目标区域颜色直方图与各个粒子所在的区域的直方图之间相似度的度量。由于图像直方图的值都是离散取值,因此巴氏距离计算公式如下:

$$d(p(n), q(n)) = \sum_{i=1}^n \sqrt{p(i) \cdot q(i)} \quad (1)$$

其中 $n$ 为图像直方图的维度, $p(x)$ 和 $q(x)$ 是需要进行比较的两个归一化的直方图分布函数。

归一化直方图曲线能体现出目标区域内的各个像素点的值

的统计信息,这些统计信息能很好地反映出目标区域内颜色的成分分布信息。但是,它不能反映目标颜色信息的空间特性,因此,当出现另外一个目标颜色成分与目标相似,其他特征完全不一样时,颜色直方图就很难反应出二者的差异了,这种情况下我们需要对目标的纹理特征进行建模分析以弥补颜色直方图信息的不足。

#### 1.1.2 纹理特征

纹理是一种局部性质,即对局部区域中像素之间的关系进行数学形式上的一种度量。纹理特征描述了颜色空间直方图所不能体现的像素空间信息。本文采用基于图像灰度共生矩阵GLCM(Gray-Level Co-occurrence Matrix)的方法来提取目标区域的纹理特征。GLCM的基本思想是在估计二阶组合条件密度函数基础上进行图像纹理的分析方法,GLCM中记录了包含空间位置关系的像素点值的统计,即统计了图像中灰度为 $h_1$ 的像素点开始到距离它的距离为 $d$ 并且灰度为 $h_2$ 的点出现的概率。

对于纹理特征的相似性度量,我们定义运动目标的纹理特征向量 $T(x)$ ,其中 $Ent(x)$ , $Egy(x)$ , $Con(x)$ , $Cor(x)$ ,以及 $Hom(x)$ 分别表示纹理特征的熵,能量值,对比度,相关性以及均匀度:

$$T(x) = (Ent(x), Egy(x), Con(x), Cor(x), Hom(x)) \quad (2)$$

在纹理特征向量中,不同的场景或者环境下纹理特征向量中的值的敏感性不尽相同,因此可以自定义一个参数矩阵 $M$ 。 $M$ 的取值可以根据具体的情况自己设定,它被用来调整各个特征量在环境中的敏感性:

$$M = (m_{ent}, m_{egy}, m_{con}, m_{cor}, m_{hom})^T \quad (3)$$

由此我们可以得到相似度的判别公式,其中 $T(x_1)$ 和 $T(x_0)$ 分别为粒子目标与原始目标的纹理特征。

$$\Delta T = T(x_1) - T(x_0) \quad (4)$$

$$d = |\Delta T \cdot M| \quad (5)$$

#### 1.1.3 运动特征

前面这两类目标特征都必须使对象在运动的过程中一直暴露在图像中才能被正确地提取。而一旦目标出现被遮挡的情况,由于其主要特征被其他干扰信息覆盖,因此很容易导致错误的跟踪。利用目标的运动特征可以较好地解决目标被完全或者部分遮挡情况下的连续跟踪问题。

运动模型的基本思想是对物体的运动过程进行记录,并基于牛顿运动定律对物体的后续运动进行预测。对于运动目标的运动过程来说,我们感兴趣的是目标运动的位置、速度以及加速度。通过记录这三个运动特性,我们可以根据牛顿力学方程来求出其他相关的运动数据。通常情况下,我们假设目标在相邻两帧所做的运动为匀加速直线运动。由于相邻两帧目标的位移并不是特别大,因此这种假设相当于对物体的运动进行微分,相对符合目标的实际运动过程。假设我们定义一个记录目标当前运动状态的变量 $s$ ,并且有:

$$s = (x, y, v_x, v_y, a_x, a_y) \quad (6)$$

其中 $x$ 和 $y$ 分别表示目标当前所在位置的 $x$ 和 $y$ 坐标, $v_x$ 和 $v_y$ 则表示目标运动速度在 $x$ 和 $y$ 坐标上的分量, $a_x$ 和 $a_y$ 则表示目标运动的加速度在 $x$ 和 $y$ 坐标上的分量,进一步可以建立如下的目标状态转移方程:

$$s_{i+1} = M_i s_i + V_i \quad (7)$$

其中, $M_i$ 为牛顿运动定律所决定的状态转移矩阵, $V_i$ 为观测噪

声。对于观测噪声,我们可以在一定范围内对其进行随机设定,这样根据初试的目标运动状态就能推出下一帧的目标的运动位置,速度以及加速度的信息,反过来,通过下一帧的目标跟踪所得到的位置再对其之前的运动方程的参数进行修正。

## 1.2 视觉注意机制

参考视觉注意机制的计算模型<sup>[14]</sup>,我们对多特征粒子滤波跟踪进行了改进。主要改进思路是在多种特征相融合的基础上,通过模拟人类视觉机制的方法,将各种不同特征按其在特定的场景中的显著性进行排序。接下来在场景中分配粒子,对所有的粒子均先从具有最高显著性的特征开始计算,通过设定阈值排除那些与该显著性特征不匹配粒子。如果这一步完成以后只剩下一个粒子则说明这个粒子即为目标的预测位置,否则说明该显著性不能将目标与所有干扰目标区分,还需要有其他的显著性进行配合,则选取一个次高的显著性作为参照。对剩下的粒子进行同样的筛选工作,直到剩下一个最优的粒子集。同时以上步骤完成以后需要通过反馈的信息对特征的显著性的排序进行更新,以保证在接下来变化的场景中最优的显著性特征排在最前面。

以上所描述的算法的实现过程实际上就是一个模拟人眼识别机制的过程。人类视觉机制是动态变化的,首先图像进入眼睛刺激视觉神经,然后那些显著性高的目标首先引起大脑的反应,并且部分特征明显的目标记录在人脑将形成存储信息,当下次出现类似的信息时马上对新图像分区搜索其最显著的特征。类似的,当获得一副图像时,算法首先将图像分区域搜索信息,也就是使用粒子在不同区域的分布获得信息,然后调用存储的显著性最强的特征进行搜索,并依次按照显著性特征的排序对散布的粒子集进行过滤,直到剩下一个最优的粒子集;最后这些搜索匹配的过程所获得的信息又会反馈给系统,以更新特征的显著性顺序。这种显著性的排序就是在不断地获得图像的过程中形成记忆而积累下来的,根据已经被跟踪过的场景和新的场景共同对特征的显著性进行打分。而以上基于显著性特征的目标跟踪过程实际上与人类的视觉机制极为相似。根据上述模型,可以得到算法的实现过程的伪代码表示如下:

```

1. ObjectTracking() {
2.   SelectObject( getImage( &image ) );           //手动选择跟踪目标
3.   InitialParticle();                             //初始化粒子
4.   While( getImage( &image ) ) {
5.     ProcessingImage( image );                   //预处理,过滤图像噪声
6.     for( particle in Particles ) {              //更新粒子
7.       UpdateParticle() {
8.         Gauss2Location();
9.         Gauss2Size();
10.      }
11.    }
12.    do {                                         //根据特征过滤粒子
13.      Filter( particle, image, feature[ index ++ ] );
14.    } while ( sizeof( particle ) != 1 )
15.    for( feature in Features ) {                //更新特征显著性
16.      UpdateSalient( feature[ ] );
17.    }
18.    Output( particle );                          //输出最终预测结果
19.  }

```

通过以上的处理以及反馈机制,我们将保证显著性特征排序表中的特征始终在图像所在的场景或者环境保持显著性最优,从而能通过它们最快地排除所有干扰的粒子,实现目标的正确跟踪。后续实验将证明,本文的算法在目标跟踪的准确性、鲁棒性和实时性方面都具有很好的性能。

## 2 实验与分析

本文提出的算法均在 2.31GHz 的双核 CPU, 2GB 内存配置系统下的虚拟机中完成。虚拟机使用的是 VMware Workstation 7.0.1, 虚拟机的设置为: 单核 CPU, 1G 内存, 操作系统为 Ubuntu 10.04。为了验证提出的算法, 我们首先在上面介绍的平台上实现了自己的算法, 然后分别实现了基于单一特征的 Mean Shift 颜色直方图目标跟踪算法以及文献[7]提出的多特征的粒子滤波目标跟踪算法, 并且采用了同样的几段视频运行在三种算法上。

我们一共选择了四段视频来进行测试, 视频一和视频二为场景较为复杂的体育比赛场景, 如图 1 第一行所示, 第二行为单一特征的目标跟踪对应第一行原始图像的跟踪效果, 第三行为本文提出的算法对所给视频进行跟踪的效果; 第一个场景为冰球赛场的一段视频, 视频中我们跟踪了两支队伍中的一位运动员; 第二个场景为足球训练的一段视频, 视频中我们跟踪了一位穿蓝色球衣的运动员。

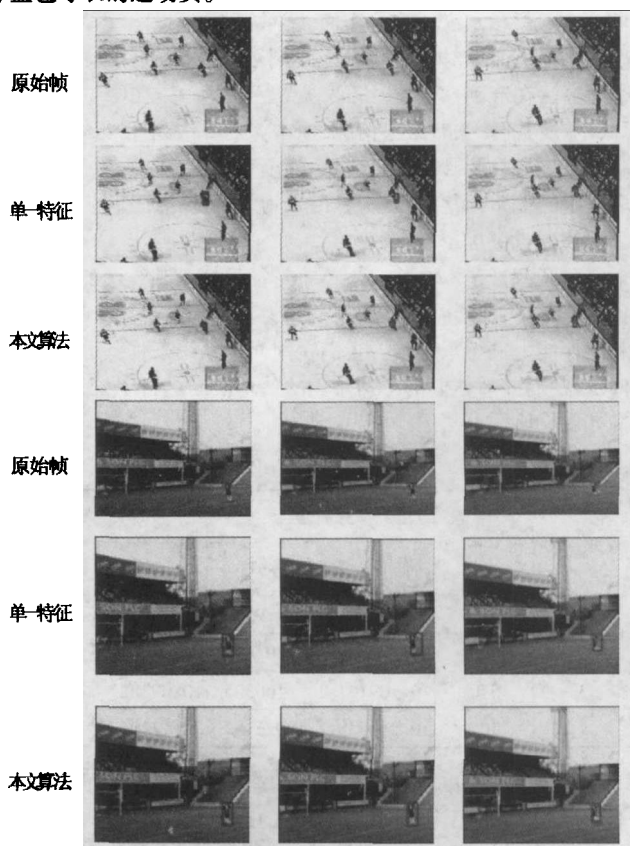


图 1 行人跟踪效果比较

第三个场景为一段公共测试序列, 内容是一段高速公路上的视频, 我们跟踪一辆行驶中的汽车; 场景四是在移动的汽车上自行拍摄的一段麓山南路上的视频, 在视频中我们跟踪了一辆行驶中的出租车, 如图 2 所示:

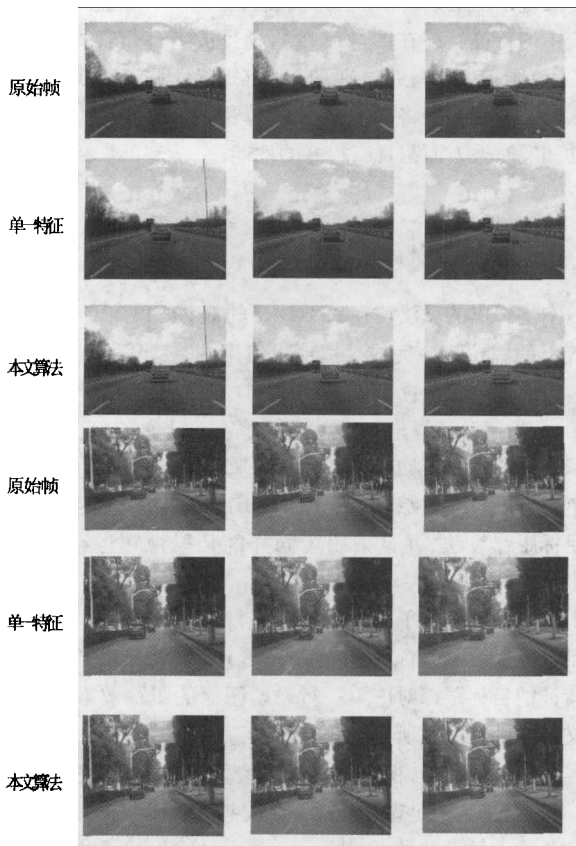


图2 车辆跟踪效果比较

从以上四组对比实验当中我们可以看出,单一特征的目标跟踪算法在准确性上不及本文提出的算法,如在图1中的第二列,单一特征算法有跟丢情况,而本文的算法则能准确地进行跟踪,第二段视频测试表明当目标逐渐变小时,单一特征算法将出现一定的偏差,而在图2所示的较为复杂的路面情况下,本文的算法表现也更优越。

除了测试了本文算法与单一特征目标跟踪算法的准确性及鲁棒性效果外,实验还与非自适应多特征粒子滤波跟踪算法在时间性能上进行了比较。我们在同样的系统条件下分别运行了四段视频,并对他们的平均每帧的处理时间进行了统计,统计结果如表1所示。

表1 跟踪时间对比

视频序号	视频帧数	文献7		本文算法	
		时间(s/f)	帧数(f/s)	时间(s/f)	帧数(f/s)
1	540	0.05787	17.028	0.0575	17.391304
2	1139	0.31216	3.203488	0.18626	5.368843
3	540	0.109167	9.160305	0.092222	10.843373
4	704	0.248438	4.025157	0.173047	5.778781

从表1中可以看出统一算法的运算时间在不同的视频中是不一样的,这是因为不同的视频中的跟踪目标不一样,目标的大小对我们所散布的粒子运算直接相关联。由于汽车的目标相对较大,因此视频三和视频四的运算时间比视频一多。从两种算法的运行时间对比可以看出,本文的自适应算法由于对多余的粒子计算进行过滤,使得一些明显可以淘汰的粒子不再重复计算,因此减少了计算的负载,在四段视频所代表的不同场景中,本文算法都取得了较好的效果。

综合本算法与单一特征算法在准确性上的对比以及与多特征粒子滤波的目标跟踪之间的对比,本文的算法结合了多特征识别准确性和鲁棒性高的特点,同时计算复杂度比普通的多特征粒子跟踪大大减低。

### 3 结语

本文在粒子滤波目标跟踪算法的基础上,针对单一特征目标跟踪对复杂场景下目标跟踪准确性不足的缺点以及多特征目标跟踪实时性难以满足的特点,提出了一种模拟人类视觉机制的多特征目标跟踪算法。通过模拟人类视觉对目标的显著性感应的机制,对使用的所有特征按所在场景的显著性进行排序,并以此来对散布粒子进行过滤,直到剩下一个最优的粒子子集。由于不同的显著性特征在不同场景对粒子的筛选能力不同,因此多类特征的组合构成了目标跟踪的自适应性。通过实验证明,本文所提算法得到了比单一特征跟踪的准确性和鲁棒性更高的效果,同时利用人眼视觉的模式大大降低了重复粒子的运算。

但是,算法仍然存在对大尺寸目标跟踪效率低的问题,下一步我们将对特征建模方法和特征显著性判断机制方面进行进一步的优化。

### 参 考 文 献

- [1] 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.
- [2] Arulampalam M S, Maskell S, Gordon N, et al. A tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking[J]. IEEE Trans on Signal Processing, 2002, 50(2): 174-188.
- [3] 朱志宇. 粒子滤波算法及其应用[M]. 北京: 科学出版社, 2010.
- [4] 王江涛. 基于视频的目标检测、跟踪及其行为识别研究[D]. 南京: 南京理工大学, 2008.
- [5] 云廷进, 郭永彩, 高潮. 基于粒子 Mean Shift 迁移的红外人体目标跟踪算法[J]. 计算机学报, 2009, 32(6): 1222-1228.
- [6] 李培华. 一种新颖的基于颜色信息的粒子滤波器跟踪算法[J]. 计算机学报, 2009, 32(12): 2454-2463.
- [7] 王欢, 王江涛, 任明武, 等. 一种鲁棒的多特征融合目标跟踪新算法[J]. 中国图象图形学报, 2009, 14(3): 489-498.
- [8] 袁广林, 薛模根, 谢恺, 等. 基于核函数粒子滤波和多特征自适应融合的目标跟踪[J]. 计算机辅助设计与图形学学报, 2009, 21(12): 1774-1784.
- [9] Itti L, Koch C, Niebur E. Model A of Saliency-Based Visual Attention for Rapid Scene Analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.
- [10] Zhang Longfei, Cao Yuanda, Zhang Mingjie, et al. Object Tracking Based on Visual Attention Model and Particle Filter[J]. International Journal of Information Technology, 2005, 11(9): 109-118.
- [11] 张焱, 张志龙, 沈振康, 等. 基于动态显著性特征的粒子滤波[J]. 电子学报, 2008, 36(12): 2306-2311.
- [12] Zhang G, Yuan Z, Zheng N, et al. Visual Saliency Based Object Tracking[C]//Proceedings of ACCV (2). 2009: 193-203.
- [13] Sidibe D, Fofi D, Meriaudeau F. Using Visual Saliency for Object Tracking with Particle Filters[C]//18th European Signal Processing Conference (EUSIPCO), Aalborg, Denmark, August 2010.
- [14] Itti L, Koch C. Computational modeling of visual attention[J]. Nature Reviews Neuroscience, 2001, 2(3): 194-230.