

计算机技术

# 一种基于 TLD 改进的视觉跟踪算法

程立英<sup>1,2</sup> 张 丹<sup>1</sup> 赵姝颖<sup>1</sup> 薛定宇<sup>1</sup>

(东北大学信息科学与工程学院<sup>1</sup> 沈阳 110004; 沈阳师范大学物理科学与技术学院<sup>2</sup> 沈阳 110034)

**摘 要** 视觉跟踪是当前计算机视觉的热点问题之一。TLD(Tracking Learning Detecting)算法是一种可以在线学习的新颖视觉跟踪算法。针对算法中跟踪器采用的LK光流法无法捕捉大幅度运动目标的问题,引入图像金字塔模型,提出一种采用金字塔光流法的TLD的改进算法,解决了长时间跟踪中出现运动尺度过大时产生孔径的问题。实验结果表明,算法在复杂场景和大运动条件下,可以长时间准确、快速地实现视觉跟踪,具有较强的适应性和有效性。

**关键词** 视觉跟踪 TLD 算法 金字塔光流法

**中图分类号** TP242.6; **文献标志码** A

视觉跟踪是计算机视觉领域研究的热点问题,根据被跟踪目标的具体性质和场景,跟踪算法也层出不穷<sup>[1-3]</sup>;但是在复杂场景下,长时间的视频由于很可能出现目标快速移动、非刚性物体的形变或目标遮挡等特殊问题,给视觉跟踪带来困难。单纯地跟踪或者单纯地检测算法都无法在长时间跟踪过程中达到理想的效果。而作为当前新颖的TLD<sup>[4]</sup>(Tracking-Learning-Detection,跟踪-学习-检测)视觉跟踪算法将检测和跟踪予以结合,并加入一种改进的在线学习机制,从而使得整体的目标跟踪更加稳定、有效。但是,由于TLD算法的跟踪器采用Lucas-Kanade光流法(LK光流法)进行跟踪,LK光流法是基于运动一致连贯性假设的,算法不适用于大幅度运动目标的跟踪。

针对TLD算法中跟踪器采用的LK光流法无法捕捉大幅度运动目标的问题,引入图像金字塔模型,提出一种采用金字塔光流法的TLD改进算法,解决了在运动尺度过大时产生的孔径问题。

## 1 TLD 视觉跟踪算法及分析

### 1.1 TLD 视觉跟踪算法

TLD算法是Zdenek Kalal在2008年到2010年间提出并不断完善的一种视觉跟踪算法<sup>[4-7]</sup>,此算法主要由三个模块组成:跟踪模块,在线学习模块和检测模块。首先,TLD算法训练一个目标检测器,通过第一帧图像获得运动目标特征,再通过仿射变换这些特征初始化检测器。跟踪器和检测器同时运行,估计目标的位置。在线学习模型不断的更新检测器和跟踪器的数据,同时跟踪器和检测器也将得到的数据反馈给在线学习模型,验证数据的精确性,使得到的目标特征更加精确,其三者相互作用,将得到的数据整合,最终确定下一帧中目标的位置。算法流程示意图如图1所示。

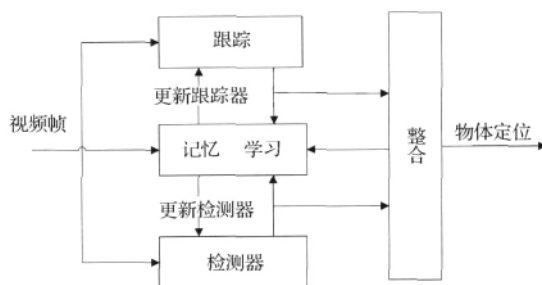


图1 TLD算法流程示意图

2012年11月9日收到

国家自然科学基金(61201378)、

中央高校基础科研基金(N110804005)资助

第一作者简介:程立英(1976—),女,讲师,博士研究生。研究方向:计算机视觉,图像处理。E-mail: clypb@163.com。

## 1.2 TLD 算法存在的问题

TLD 算法中的跟踪器使用 LK 光流法进行跟踪。首先通过在第一帧图像中手动选择跟踪目标,获得目标的特征点。LK 算法在帧与帧之间跟踪这些点,产生矢量光流场,根据所取的特征,使用均值的方式预测下一帧中目标的运动和位移方向,如图 2 所示。这样,在每一帧中都有一系列的新特征点被跟踪,因此使得跟踪具有较强的适应性。

LK 光流法最初于 1981 年提出<sup>[8]</sup>,LK 光流法基于以下三个假设:

(1) 亮度恒定 图像场景中目标像素在不同帧间运动时外观上保持不变,对于灰度图像,像素被跟踪时其亮度不发生变化;

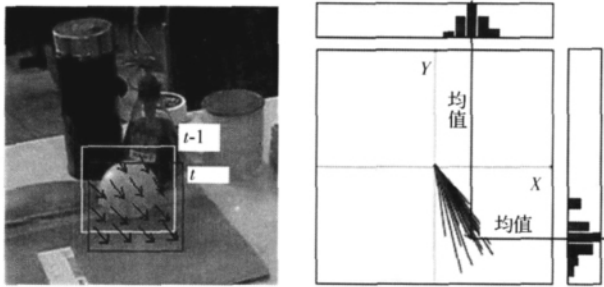


图 2 跟踪算法示意图

(2) 时间连续或者运动是“小运动” 图像的运动随时间变化比较缓慢。实际应用中是指时间变化相对图像中运动比例要足够小,这样目标在不同帧间的运动就比较小;

(3) 空间一致 一个场景中同一表面上邻近点具有相似的运动,在图像上的投影也在邻近区域。

亮度恒定是指被跟踪部分像素不随时间变化:

$$f(x, t) = I(x(t), t) = I[x(t + dt), t + dt] \quad (1)$$

针对第二个假设将运动的变化看成是亮度对时间的导数,考虑隐含的  $x$  为  $t$  的函数,将亮度的定义  $f(x, t)$  用  $I[x(t), t]$  替换,应用偏微分链式规则:

$$\frac{\partial I}{\partial x} \left[ \left( \frac{\partial x}{\partial t} \right) + \frac{\partial I}{\partial t} \right]_{x(t)} = 0 \quad (2)$$

式(2)中,  $I_x$  是图像的偏导数,  $I_t$  是随时间的导数,  $v$  是要求的速度。

$$v = - \frac{I_t}{I_x} \quad (3)$$

前面是 LK 光流法在一维空间上的用法,将其扩展到二维图像上,速度的  $y$  分量为  $v$ ,  $x$  分量为  $u$ ,得到:

$$I_x u + I_y v + I_t = 0 \quad (4)$$

但对于单个像素来说,等式的约束条件太少,不能得到二维运动的定解,只能求得光流方程垂直方向的运动分量。垂直光流由孔径问题产生,即用小孔或小窗口去测量运动,这样,只能观测到物体边缘而观测不到角点,而只依靠边缘是不足以判断整个运动物体的运动朝向的。

利用光流法的第三个假设求解整个运动,建立邻域像素的系统方程来求解中心像素运动。如果利用  $5 \times 5$  邻域的像素亮度值计算此像素的运动,则可以建立 25 个方程。

$$\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_{25}) & I_y(p_{25}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots \\ I_t(p_{25}) \end{bmatrix} \quad (5)$$

为解这个系统方程,需要建立该方程的最小平方法,通过下面方程来求解最小化的  $\|Ad - b\|^2$ :

$$(A^T A) d = A^T b \quad (6)$$

这样,可以得到  $u$  和  $v$  运动分量,表述如下:

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix} \quad (7)$$

当  $A^T A$  可逆时,方程解如下:

$$\begin{bmatrix} u \\ v \end{bmatrix} = (A^T A)^{-1} A^T b \quad (8)$$

当  $A^T A$  满秩,即其有两个较大的特征向量时,  $A^T A$  可逆,在这种情况下跟踪窗口的中心在图像的角点区域。正因为  $(A^T A)$  在角点处有两个大的特征向量,这些角点是可用于跟踪的良好特征点。角点以有限的数据量集中表达了图像中物体关键的形状信息,很大程度上降低了数据冗余度,为实时检测提供了保障。但是,在实际应用中,大而不连贯的运动普遍存在。基于小窗口 LK 算法的跟踪器存在不足,较大的运动会使目标将移出小窗口,从

而无法再找到这些点。因此需要一个大的窗口来捕获较大的运动,而大窗口会违背LK光流法运动连贯的假设。

## 2 TLD 视觉跟踪算法的改进

通过前面的分析研究,可知TLD算法中的跟踪器采用的LK光流法容易产生孔径问题,不能准确跟踪运动尺度过大的目标,只适用小图像灰度模式的运动估计,受灰度不变约束和速度平滑性约束,这样会影响跟踪的效果。但在实际应用的跟踪环境下的目标运动速度较快,难以满足光流估计方法的约束条件,所以为了解决这个问题,现引入金字塔光流法。图像的金字塔结构<sup>[9]</sup>是一种多分辨率表示方法,不同金字塔分辨率层可分析不同尺度的目标。金字塔光流法首先在图像金字塔的最高层计算光流,用得到的运动估计结果作为下一层金字塔的起始点,重复这个过程直到金字塔的最底层。这样就将不满足运动假设的可能性降低到最小,从而实现快速运动目标的跟踪,原理如图3所示。

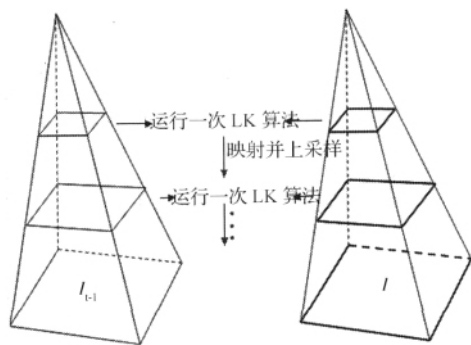


图3 金字塔光流法示意图

根据金字塔光流法原理,将原图像作为金字塔基层  $I = 0$ ,将原图减采样至原尺寸的  $1/2^{2N}$  (一般  $N = 1$ ),获得  $I = 1$  层图像,则该层相邻帧间的目标像素运动距离为  $D/2^N$  ( $D$  为原图中相邻帧间的目标像素运动距离)。当  $I$  达到一定值时(一般  $3 \leq I \leq 5$ ),可满足光流约束条件。算法流程图如图4所示。在最高层  $V_f = 0$ ,图像细节最少, $\Delta V_f$  是  $f$  层的光流计算结果,作为下一层图像的运动估计,并根据运算规则计算此帧的光流,直到运算到图像的最底层。

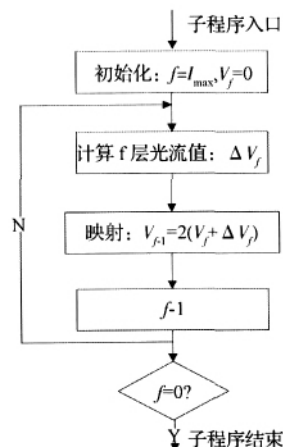


图4 金字塔光流法计算流程图

改进的TLD算法在线学习模块采用原算法中的P-N学习框架进行特征的更新,检测器模块使用的分类器采用随机森林方法。

## 3 算法验证

分别针对背景简单、但快速运动的摩托车运动视频<sup>[4]</sup>和自拍的复杂背景下大幅度运动的盆花视频作实验资源,在实验用计算机为CPU E5500 @ 2.80 GHz 2 GB 内存,在Matlab R2009a环境下运行对本文算法进行验证。

实验1采用文献[4]提供的摩托车运动视频,跟踪目标是摩托车,由2 917帧组成(2: 33),在这段视频中包含了目标的遮挡、消失、快速运动、特征的突然变化等。跟踪效果如图5所示,可以看出,在出现部分遮挡之后,目标丢失,但是在下一帧时,改进的TLD算法依然能检测到运动的目标,原因就是在前两帧中运动目标由于运动速度非常快,位置发生了很大的变化,而跟踪对象的搜索范围较小产生的,但是在下一帧时算法便克服其误差,进行大范围的搜索,得到了运动目标。

实验2采用自拍的盆花运动视频作实验资源,分辨率为  $320 \times 240$  帧率为  $30 \text{ f/s}$ 。跟踪目标是花,由4 740帧组成(2: 36)。在视频中盆花被人为的向右、向左、向上做大幅度的运动,跟踪效果如图6所示,可以看出跟踪算法可以长时间在复杂背景和大运动条件下对目标进行跟踪。

为了衡量算法的准确性,采用文献[4]中的评

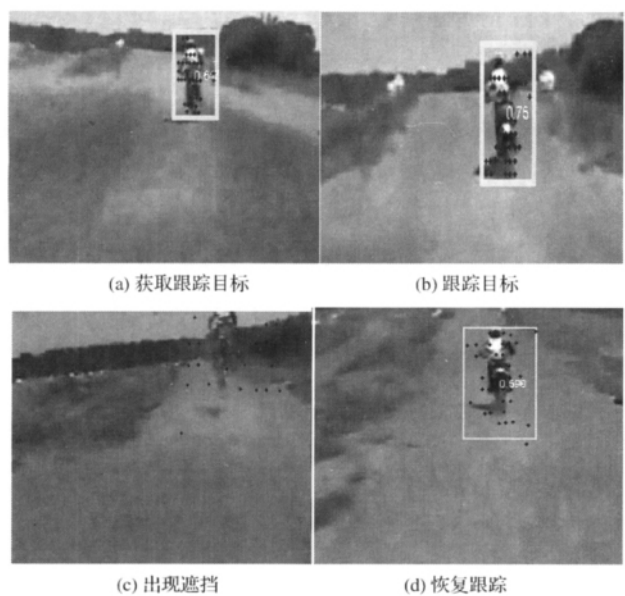


图5 简单背景下跟踪效果图



图6 复杂背景下跟踪效果图

估标准 将跟踪过程中算法选择的所有目标窗口数(样本)定义为  $ST$ , 手动选择的目标窗口(样本)个数定义为  $GT$ , 如果跟踪获得的窗口和手动选择的窗口的重合率超过 70%, 就认为获得是正样本, 用  $TP$  表示, 否则为  $FP$ 。准确率用  $P$  表示, 定义为:

$$P = |TP| / |ST| \tag{9}$$

用  $R$  来表示正样本和手动选择目标窗口的比, 定义为:

$$R = |TP| / |GT| \tag{10}$$

使用  $P$  和  $R$  这两个参数评估跟踪方法的质量。将本文方法与 Mean-Shift 算法和原 TLD 算法进行对比, 跟踪质量对比如表 1 所示。

表1 跟踪质量对比表

评价的方法	摩托车视频		盆花视频	
	$P$	$R$	$P$	$R$
Mean-Shift	0.85	0.43	0.84	0.57
原 TLD 算法	0.93	0.51	0.94	0.73
本文方法	0.94	0.54	0.95	0.78

从表 1 中可以看出, 对于长时间跟踪的两个实验, 原 TLD 算法的准确率明显高于 Mean-Shift 算法。由于跟踪器的改进, 说明现算法无论在简单背景或复杂背景下对有大幅度运动的目标跟踪效果都优于原 TLD 算法。由此可见, 本文算法提高了原 TLD 算法的准确性和适应性。

4 结 语

针对 TLD 算法中跟踪器采用的 LK 光流法无法捕捉大幅度运动目标的问题, 引入图像金字塔模型, 提出一种采用金字塔光流法的 TLD 的改进算法, 实验验证了该算法提高了 TLD 算法在长时间跟踪中, 出现大尺度运动下跟踪目标的准确性和适应性, 该算法针对单运动目标的跟踪效果已经很理想, 今后将考虑如何在此基础上, 实现复杂背景下多目标的理想跟踪。

参 考 文 献

- 1 孙 巧. 遮挡情况下基于粒子滤波的视觉跟踪算法研究. 科学技术与工程 2012; 12( 32): 57—59
- 2 袁 霄, 王丽萍. 基于 MeanShift 算法的运动人体跟踪. 计算机工程与科学 2008; 30( 4): 46—49
- 3 刘晓辉, 陈小平. 基于扩展卡尔曼滤波的主动视觉跟踪技术. 计算机辅助工程 2007; 16( 2): 32—37
- 4 Kalal Z, Matas J, Mikolajczyk K. Online learning of robust object detectors during unstable tracking. [http://info.ee.surrey.ac.uk/Personal/Z.Kalal/Publications/2009\\_olcv.pdf](http://info.ee.surrey.ac.uk/Personal/Z.Kalal/Publications/2009_olcv.pdf). 2009
- 5 Kalal Z, Mikolajczyk K, Matas J. Face-TLD: tracking-learning-detection applied to faces. [http://info.ee.surrey.ac.uk/Personal/Z.Kalal/Publications/2010\\_icip.pdf](http://info.ee.surrey.ac.uk/Personal/Z.Kalal/Publications/2010_icip.pdf). 2010
- 6 Kalal Z, Mikolajczyk K, Matas J. Forward-backward error: Automatic detection of tracking failures. <http://info.ee.surrey.ac.uk/Per->

- sonal/Z. Kalal/Publications/2010\_icpr. pdf. 2010
- 7 Kalal Z , Matas J , Mikolajczyk K. P-N learning: bootstrapping binary classifiers by structural constraints. [http://info.ee.surrey.ac.uk/Personal/Z. Kalal/Publications/2010\\_cvpr. pdf](http://info.ee.surrey.ac.uk/Personal/Z.Kalal/Publications/2010_cvpr.pdf). 2010
- 8 Lucas B , Kanade T. An iterative image registration technique with an application to stereo vision. IJCAI ,1981; 81:674—679
- 9 Bouguet J Y. Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm. [http://robots.stanford.edu/cs223b04/algo\\_tracking. pdf](http://robots.stanford.edu/cs223b04/algo_tracking.pdf). 2000

## An Algorithm of Visual Tracking Based on Tracking-learning-detection

CHENG Li-ying<sup>1 2</sup> ZHANG Dan<sup>1</sup> ZHAO Shu-ying XUE Ding-yu<sup>1</sup>

( College of Information Science and Engineering , Northeastern University<sup>1</sup> , Shenyang 110004 , P. R. China;

College of Physics Science and Technology , Shenyang Normal University<sup>2</sup> , Shenyang 110034 , P. R. China)

**[Abstract]** Visual tracking is a research hotspot in computer vision. TLD( Tracking-Learning-Detecting) algorithm is a novel visual tracking algorithm which can online learning. Aiming at the problem that LK flow method can not capture a large scale movement of TLD algorithm , an improved method is presented which used pyramid optical flow as the tracker , to get more corner information to overcome the aperture problem. The experimental results show the presented algorithm is efficient and robust to the large scale movement during long term tracking.

**[Key words]** visual tracking TLD algorithm pyramid optical flow

( 上接第 2370 页)

## Inverter-fault Detection Methods Based on Hilbert-Huang Transform

CHEN Tao ,XIAO Hai-hong

( School of Electrical Information Engineering ,Henan Institute of Engineering Zhengzhou 451191 ,P. R. China)

**[Abstract]** A study on inverter faults of variable frequency speed control system used in electric vehicles is presented , including single power device open circuit fault and one phase missing fault , based on DC side current and Hilbert-Huang Transform ( HHT) . Simulation on both stable state and dynamic operation condition has been carried out. The results show that this method can effectively extract the fault indications from DC side current for both stable state and dynamic operation condition , and it is suitable for car conditions.

**[Key words]** DC side current Hilbert-Huang inverter fault detection