

计算机视觉目标跟踪的可信估计

李 杨, 李 喆

长春理工大学理学院, 吉林 长春
Email: 1714295405@qq.com, zheli200809@163.com

收稿日期: 2021年6月19日; 录用日期: 2021年7月11日; 发布日期: 2021年7月23日

摘 要

计算机视觉领域, 目标跟踪是一项具有许多现实应用的视觉任务, 对运动的目标进行精准估计成为了一项具有重要意义的工作。本文将区间算法应用到计算机视觉目标跟踪中, 基于频谱滤波器, 利用区间运算实现计算机视觉目标跟踪问题的可信估计。当目标跟踪失败时, 利用颜色识别对每一帧背景进行预处理, 预估目标所在的区域范围, 从而更好的确定目标位置。

关键词

计算机视觉, 目标跟踪, 可信估计

Trusted Estimation of Object Tracking in Computer Vision

Yang Li, Zhe Li

School of Science, Changchun University of Science and Technology, Changchun Jilin
Email: 1714295405@qq.com, zheli200809@163.com

Received: Jun. 19th, 2021; accepted: Jul. 11th, 2021; published: Jul. 23rd, 2021

Abstract

In the field of computer vision, target tracking is a visual task with many practical applications. Accurate estimation of moving targets has become an important work. In this paper, the interval algorithm is applied to computer vision target tracking. Based on the spectral filter, the interval operation is used to realize the trusted estimation of computer vision target tracking problem. When the target tracking fails, the color recognition is used to preprocess the background of each frame to estimate the range of the target area, so as to better determine the target location.

Keywords

Computer Vision, Target Tracking, Trusted Estimation

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

目标跟踪就是对视频图像序列中的运动目标进行检测、提取、识别和跟踪, 获得运动目标的运动参数, 如目标质心位置、速度等, 以及运动轨迹, 从而进行进一步处理与分析, 实现对运动目标的行为理解[1]。目标跟踪作为一门跨学科的前沿技术, 融合了图像处理、模式识别、人工智能、自动控制等多种不同领域的理论知识[2], 并且在视频理解、机器人导航、智能监控等领域都有着广泛的应用, 因此跟踪算法的研究具有重要的意义。

目标跟踪过程中能够影响跟踪结果的因素有很多, 如何提高目标跟踪的稳定性成为了计算机视觉方向一个需要突破的难题。近年来, 众多科学学者致力于此方向的研究, 所提出的算法可以大致分为两类: 基于部分跟踪算法和整体跟踪算法。

Prokhorov [3]等人提出了一个层次化的外观的表示模型用于跟踪, 基于利用多个量化级别的共享信息的图形模型, 通过对像素和超像素的联合分类来找到目标的最可能位置, 在所有级别获得最佳配置。雷震等人[4]提出了一种新的多部件联合在线跟踪分割算法(JOTS), 将跟踪和分割融合为一个统一的能量优化框架来处理视频分割任务。卢湖川等人[5]中提出了一个基于超混合的判别外观模型, 能够解决严重遮挡问题, 并且能够从漂移中恢复跟踪的目标。Malik 等人[6]提出了一种从背景中重复分割图形的跟踪模式, 用于解决长时间跟踪时物体的外观、形状和比例变化对计算结果产生的影响。

不同的是, 整体的跟踪方法似乎更受欢迎, 例如高效且具有良好鲁棒性优点的基于相关滤波器(CF)的模型。在 CF 模型的基础上出现了相关滤波器的一些正则化方法, 此类型的方法从空间抑制背景区域以解决杂乱背景对跟踪器性能的影响。崔振等人[7]提出一种新的方法称为递归目标参与跟踪(RTT), 用来解决估计误差的累积和传播。Danelljan 等人[8]提出了高精度的用于跟踪的空间正则化判别相关滤波器(SRDCF)。与此同时 Danelljan 等人[9]还研究了基于 DCF 的两个跟踪框架(标准 DCF 框架和 SRDCF 框架)中卷积特征的影响, 标准的基于 DCF 的方法依赖于手工制作的特性来进行稳健的图像描述, 提出建议在基于 DCF 的框架中使用卷积特征来进行视觉跟踪。

跟踪结果的准确率会随着外部环境的变化而大幅度的下降, 同时一些传统的目标跟踪滤波器自身也会存在很多缺点。针对如何提高目标跟踪结果准确率的问题, 学者们给出了不同的答案。朱明敏等人[10]提出一种可以实时高效准确的跟踪快速运动目标的长时核相关滤波(LKCF)跟踪算法, 此算法引入了 CN 颜色特征及基于光流法的再检测机制, 当目标跟踪失败时可以对失踪的目标进行重新获得。徐烨超[11]提出了 IFCT (基于独立特征通道的跟踪), 用来解决基于相关滤波器(CF)的视觉目标跟踪算法中非充分学习、非置信结合和非弹性更新在内的三个缺陷, 并且提出了 RCF (基于独立特征通道的可逆向学习相关滤波器)以解决场景导致的模型漂移和算法导致的模型漂移等问题。张雷[12]等人提出基于 KCF 的多尺度目标跟踪的方法, 解决了目标跟踪问题中有关目标尺度估计的困难。

与基于 CF 的方法的整体回归不同, 崔振等[13]人提出了对每个像素的多个局部区域的跟踪模型进行

回归的频谱滤波器(SFT)跟踪方法。本文将 Rump 区间算法应用到频谱滤波器(SFT)的跟踪模型中, 实现目标跟踪问题的可信估计, 同时利用颜色识别提高跟踪结果的准确性。

2. 目标跟踪的可信估计

定理 1: 区间矩阵 $\mathbf{A} \in \mathbb{IR}^{m \times n}$ ($m \leq n$), 和区间向量 $\mathbf{b} \in \mathbb{IR}^m$ 满足 $0 \notin \mathbf{b}$, 如果 verifylss 函数对下列区间线性系统成功运行

$$\begin{pmatrix} A^T & -I_n \\ O_{m,n} & A \end{pmatrix} \begin{pmatrix} w \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$$

则区间矩阵 \mathbf{A} 是满秩的。

频谱滤波器

在给定视频帧的情况下, 由于连续视频帧中目标的运动的变化不容易被察觉, 首先在前一帧的边界框周围确定一个小的候选区域。为了减少局部外观变化对跟踪的影响, 将候选区域建模为具有旋转不变性和平移不变性的像素网格图。

将像素空间网格结构建模为无向加权图, 加权图 $G = \{V, \varepsilon, W\}$ 由一组顶点 $V = (|V| = N)$ 和一组边 ε , 邻域矩阵 W 为连接的边分配的正值。此外, 每个顶点都与信号相关, 即从其坐标位置提取多通道特征向量。形式上特征提取函数 $f: v \rightarrow R^d$ 定义了顶点信号, 其中 d 是特征维数。

谱图理论中图 Laplacian 算子 L 是一个关键的算子, 定义为 $L = D - W$, 其中 $D \in \mathbb{R}^{N \times N}$ 是对角矩阵 $D_{ii} = \sum_j W_{ij}$ 。一种常用的方法是归一化图 Laplacian, 即将每个权重 W_{ij} 与一个因子 $\frac{1}{\sqrt{D_{ii}D_{jj}}}$ 相乘, 即

$$L^{norm} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}} = I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}}, \quad (1)$$

其中 L 为实对称矩阵, 其具有完整的正交特征向量, 其特征向量 $\{u_l\}$ 满足 $Lu_l = \lambda_l u_l$ ($l = 1, 2, \dots, N$), 其中 $\{\lambda_l\}$ 是非负实特征值。假设所有的特征值都是有序的则 $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_N = \lambda_{max}$ 。

在矩阵表达式中 $L = U \Lambda U^T$, 其中 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ 。

步骤 1 滤波器函数的构造

x 的傅里叶变换 $x = U^T \hat{x}$ 以及逆变换 $x = U \hat{x}$, 其中 \hat{x} 是产生的频率信号。

设滤波器函数 $g(\cdot)$, 输入信号 x 上的频率滤波定义为 $\hat{z}(\lambda_l) = \hat{x}(\lambda_l) \hat{g}(\lambda_l)$, 或为逆图傅里叶变换

$$z(i) = \sum_{l=1}^N \hat{x}(\lambda_l) \hat{g}(\lambda_l) \hat{u}_l(i), \quad (2)$$

其中 $\hat{z}(\lambda_l), \hat{x}(\lambda_l), \hat{g}(\lambda_l)$ 是对应频谱 λ_l 的傅里叶系数。通过使用矩阵表示法, 信号 x 被过滤为

$$z = \hat{g}(L)x = U \begin{bmatrix} \hat{g}(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \hat{g}(\lambda_N) \end{bmatrix} U^T x. \quad (3)$$

步骤 2 函数的求解

给定的输入 x 和输出 z , 通过特征值分解求解可求得(3)中的 $g(\cdot)$, 同时利用 K 阶的 Chebyshev 展开式来逼近频域中的 $g(\cdot)$ 降低计算量。

Chebyshev 展开式由递归关系定义为

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x),$$

其中 $T_0 = 1, T_1 = x$ 。

在适当的 Sobolev 空间中, Chebyshev 多项式的集合构成了一个基, 因此 $[-1, 1]$ 中任何一个函数都可以通过展开来表示 $f(x) = \sum_{k=0}^{\infty} a_k T_k(x)$ 。

为使得 Laplacian L 的特征值 $\{\lambda_l\}$ 落在 $[-1, 1]$ 中, 将其缩放并转换为 $\tilde{\lambda}_l = \frac{2}{\lambda_{\max}} \lambda_l - 1$, 并在 $\{\tilde{\lambda}_l\}$ 上使用 Chebyshev 多项式, 考虑多项式分量的线性组合, 则 K 阶滤波器

$$\hat{g}(\lambda_l) = \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\lambda}_l), \quad (4)$$

其中 $\theta \in \mathbb{R}^K$ 是多项式系数的参数向量, K 是多项式的阶。由(3)(4)有

$$\begin{aligned} z &= U \begin{bmatrix} \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\lambda}_l) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\lambda}_l) \end{bmatrix} U^T x \\ &= \sum_{k=0}^{K-1} \theta_k U \begin{bmatrix} T_k(\tilde{\lambda}_l) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & T_k(\tilde{\lambda}_l) \end{bmatrix} U^T x \end{aligned} \quad (5)$$

$$= \sum_{k=0}^{K-1} \theta_k T_k \left(U \begin{bmatrix} \tilde{\lambda}_l & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \tilde{\lambda}_l \end{bmatrix} U^T \right) x \quad (6)$$

$$= \sum_{k=0}^{K-1} \theta_k T_k \left(\frac{2}{\lambda_{\max}} L - I \right) x \quad (7)$$

在(6), (7)上使用 Laplacian 矩阵的谱分解 $L = U \text{diag}([\lambda_1, \dots, \lambda_N]) U^T$ 。通过(5)和(6)利用滤波函数的基本计算即

$$L^k = U \text{diag}([\lambda_1^k, \dots, \lambda_N^k]) U^T = (U \text{diag}[\lambda_1, \dots, \lambda_N] U^T)^k \quad (8)$$

在(7)中的 K 阶多项式是拉普拉斯图上的一个精确的 K 局部化滤波函数。为了得到图上的局部滤波响应, 只需对拉普拉斯矩阵 L 进行运算, 即多项式的每一项都可以作为滤波基, θ 是需要求解的参数。

步骤 3 跟踪目标模型的建立

在视觉跟踪中, 从多通道特征 $X \in \mathbb{R}^{N \times d}$ 中回归峰值映射 $y \in \mathbb{R}^{N \times 1}$, 其中 N 候选区域内的像素数(即顶点), 每一行 X 对应一个顶点的信号。

$\tilde{L} = \frac{2}{\lambda_{\max}} L - I$, 将(7)中定义的滤波器基变成 $\{T_0(\tilde{L}), T_1(\tilde{L}), \dots, T_{K-1}(\tilde{L})\}$ 。

给定一个信号和 x 和滤波器参数 $\theta = \{\theta_0, \theta_2, \dots, \theta_{K-1}\}$, 采用 K 个滤波器基的线性组合获得局部滤波响应 z 。使用 K 个滤波器基对图进行滤波, 将滤波参数和特征投影函数的学习结合到一个最小二乘回归模型中

$$\arg \min_w \|F(X)w - y\|^2 + \gamma \|w\|^2, \quad (9)$$

其中 γ 是平衡参数, $F(X)$ 在特征维数上连接 K 滤波器基的响应,

$$F(X) = [T_0(\tilde{L})X, T_1(\tilde{L})X, \dots, T_{K-1}(\tilde{L})X]. \quad (10)$$

因此, 根据定理 1, 跟踪模型中的系数 w 可以利用 Rump 区间算法估计求解为

$$\begin{pmatrix} (F(X))^T & -I_n \\ O_{m,n} & F(X) \end{pmatrix} \begin{pmatrix} w \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}. \quad (11)$$

步骤 4 跟踪目标的坐标的计算

初始化拉普拉斯矩阵 L , 从候选区域中提取特征 X , 计算(10)方程中的 K 阶响应 $F(X)$, 利用(11)获取得分

$$\tilde{y} = F(X)w$$

通过找到得分最大的目标中心获取被跟踪目标的最终坐标。

3. 颜色识别

跟踪过程中由于外部环境的影响会导致跟踪失败, 为了提高模型对目标捕获的准确性, 在基于区间算法的频谱滤波器跟踪模型中增加颜色识别部分。首先从目标图像中抽取颜色特征, 彩色图像得格式有很多, 本节以 RGB 的格式来提取颜色特征。

3.1. 提取颜色特征

当物体的呈蓝色时, 物体的蓝色分量值大于红色、绿色分量值, 则蓝色特征的提取公式

$$D(x, y) = 2B(x, y) - R(x, y) - G(x, y).$$

当物体呈黄色时, 物体的红色分量和绿色分量值大于蓝色分量, 并且红色和绿色分量之间的差距较小时, 则黄色特征的提取公式

$$D(x, y) = R(x, y) + G(x, y) - 2B(x, y) - k \times \text{ABS}(R(x, y) - G(x, y)),$$

其中, 系数 k 的选择取决于黄色的个体差异, 当背景颜色不是纯黄色而是偏于绿色或红色的时候, 系数 k 的值可以适当减少。

当物体呈黑色时, 物体的红色、绿色、蓝色三个分量的值都比较小, 同时三个分量之间的差异也比较小, 则黑色特征的提取公式

$$D(x, y) = 255 \times 3 - (R(x, y) + G(x, y) + B(x, y)) - k \times (\text{ABS}(R(x, y) - G(x, y)) + \text{ABS}(R(x, y) - B(x, y)) \text{ABS}(B(x, y) - G(x, y)))$$

其中, 系数 k 的选择取决于黑色的纯正程度。

当物体的呈白色时, 物体的红色、绿色、蓝色三个分量的值都比较大, 但是三个分量之间的差异比较小, 则白色特征的提取公式

$$D(x, y) = R(x, y) + G(x, y) + B(x, y) - k \times (\text{ABS}(R(x, y) - G(x, y)) + \text{ABS}(R(x, y) - B(x, y)) \text{ABS}(B(x, y) - G(x, y)))$$

其中, 系数 k 的选择取决于白色的纯正程度。

3.2. 选取阈值

经过颜色特征的提取之后, 原彩色图像变成了几个灰度图像。在对每个灰度图像进行搜索时, 一般需要进行二值化的处理, 此时阈值的选择成为了首要问题。

在这里采用依据背景和物体之间的类间方差最大, 而背景之间以及物体之间的类内方差最小的大津方法。

设一块图像区域的灰度值为 $f(i, j)(i = 1, \dots, m; j = 1, \dots, n)$, 则阈值计算公式

$$sum = \sum_{i=1}^m \sum_{j=1}^n f_{ij} \quad dots = m \times n, \quad (10)$$

D_k 为灰度直方图 ($k = 0 \sim 255$),

$$a = \sum_{k=1}^g k \times D_k \quad b = \sum_{k=1}^g D_k, \quad (11)$$

$$t_1 = b \times dots - b \times b \quad t_2 = sum \times b - dots \times a, \quad (12)$$

$$p = \frac{t_2 \times t_2}{t_1 \times dots \times dots}, \quad (13)$$

计算 g 从 1 到 255 时, p 的值, 找出最大的 p 就是所求的阈值。

4. 实验结果及分析



Figure 1. Experimental results 1
图 1. 实验结果 1



Figure 2. Experimental results 2
图 2. 实验结果 2

如图 1、图 2 所示, 明确的给出了本文改进前算法的部分跟踪结果, 从跟踪结果中可以清楚地看出, 当目标与背景颜色对比鲜明、图中物体颜色反差较为明显, 目标物体未发生形变的状态下跟踪效果较为良好, 可以很清晰的看出物体的颜色, 物体的轮廓等。



Figure 3. Experimental results 3
图 3. 实验结果 3

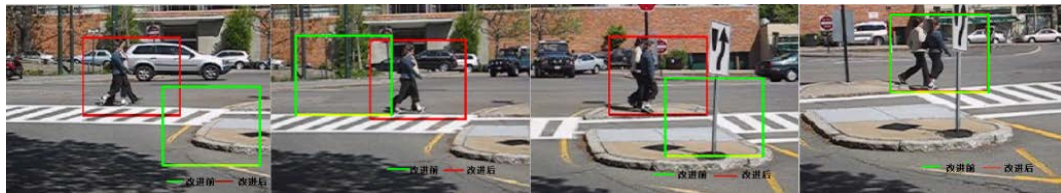


Figure 4. Experimental results 4
图 4. 实验结果 4

如图 3、图 4 所示, 改进前的方法无法清晰地给出跟踪结果, 当跟踪环境未能达到理想的状态下, 且目标物体发生较大形变时, 容易导致目标物体的跟踪失败, 此时采用改进后的跟踪方法能够对目标重新准确的跟踪。

通过以上的跟踪实验可以得出, 针对跟踪失败的案例, 在对目标物体进行跟踪识别前, 利用颜色识别对目标背景进行处理, 以此获取目标物体颜色的 RGB 阈值。再根据选定的 RGB 阈值范围, 在视频帧中选定在 RGB 阈值范围内的区域进行调试, 对目标进行重新获取, 再利用频谱滤波器跟踪模型确定重新捕获的目标在背景中的位置。

从实验中可以清晰地看出, 与原有的计算方法相比下, 改进后的计算方法确实能够更好更明确的对目标物体进行捕捉跟踪, 但该算法也存在着一些弊端的短板, 例如计算的过程较为复杂、计算效率比较低等, 所以跟踪效率方面依旧存在着很大的提升空间。在未来该算法需要通过进一步的优化和研究, 从而取得更为准确的计算结果, 实现更高效率的跟踪。

5. 总结

在目标物体跟踪的过程中, 会存在着很多直接影响跟踪计算结果的准确性的因素, 例如跟踪目标物体的形状、尺寸的大小、位置的变化, 和跟踪目标背景环境以及光照的条件变化, 以及目标物体被遮挡以及部分被遮挡等, 都会导致的目标物体跟踪出现失败。针对这些难点, 本文提出基于区间算法的频谱滤波器跟踪方法, 即通过利用区间运算来计算跟踪模型系数值的区间, 从而实现目标物体跟踪的可信估计。当目标物体跟踪失败, 利用目标物体颜色, 估计目标物体的位置区域范围, 实现目标的重新捕获跟踪。通过实验可以证明, 与原有的跟踪方法相比, 改进后的跟踪计算方法能更准确的对目标进行跟踪。

参考文献

- [1] 卢晓鹏. 视频序列中目标跟踪技术研究[D]: [博士学位论文]. 北京: 中国科学院电子研究所, 2007.
- [2] 张娟, 毛晓波, 陈铁军. 运动目标跟踪算法研究综述[J]. 计算机应用研究, 2009, 26(12): 4407-4410.
- [3] Hong, Z., Wang, C., Mei, X., Prokhorov, D. and Tao, D. (2014) Tracking Using Multilevel Quantizations. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds., *Computer Vision-ECCV 2014, ECCV 2014, Lecture Notes in Computer Science*, 8694, 155-171. https://doi.org/10.1007/978-3-319-10599-4_11
- [4] Wen, L., Du, D., Zhen, L., Li, S.Z. and Yang, M.H. (2015) Jots: Joint Online Tracking and Segmentation. 2015 *IEEE*

-
- Conference on Computer Vision and Pattern Recognition (CVPR)*, **6**, 2226-2234. <https://doi.org/10.1109/CVPR.2015.7298835>
- [5] Wang, S., Lu, H., Yang, F. and Yang, M.H. (2011) Superpixel Tracking. 2011 *International Conference on Computer Vision*, **24**, 1323-1330.
- [6] Ren, X. and Malik, J. (2007) Tracking as Repeated Figure/Ground Segmentation. 2007 *IEEE Conference on Computer Vision and Pattern Recognition*, **1**, 1-8. <https://doi.org/10.1109/CVPR.2007.383177>
- [7] Cui, Z., Xiao, S., Feng, J. and Yan, S. (2016) Recurrently Target-Attending Tracking. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, **1**, 1449-1458. <https://doi.org/10.1109/CVPR.2016.161>
- [8] Danelljan, M., Häger, F., Khan, S. and Felsberg, M. (2015) Learning Spatially Regularized Correlation Filters for Visual Tracking. 2015 *IEEE International Conference on Computer Vision (ICCV)*, **1**, 4310-4318. <https://doi.org/10.1109/ICCV.2015.490>
- [9] Danelljan, M., Häger, F., Khan, S. and Felsberg, M. (2015) Convolutional Features for Correlation Filter Based Visual Tracking. 2015 *IEEE International Conference on Computer Vision Workshop (ICCVW)*, **1**, 58-66. <https://doi.org/10.1109/ICCVW.2015.84>
- [10] 朱明敏, 胡茂海. 基于相关滤波器的长时视觉目标跟踪方法[J]. 计算机应用, 2017, 37(5): 1466-1470.
- [11] 徐焯超. 基于相关滤波器的视觉目标跟踪方法研究[D]: [硕士学位论文]. 南京: 南京邮电大学, 2019.
- [12] 张雷, 王延杰, 刘艳滢. 基于相关滤波器的视觉目标跟踪方法[J]. 光电子·激光, 2015(7): 1349-1357.
- [13] Cui, Z., Cai, Y.Y. and Zheng, W.M. (2019) Spectral Filter Tracking. 2019 *IEEE Transactions on Image Processing*, **28**, 2479-2489. <https://doi.org/10.1109/TIP.2018.2886788>