

基于改进的 MOSSE 相关滤波的目标跟踪

纪纲, 高富东, 范加利

(海军航空大学 舰面航空保障与场站管理系, 山东 青岛 266041)

摘要: 为了解决视频目标跟踪过程中遮挡、形变, 光照导致跟踪漂移的问题, 在 MOSSE 相关滤波算法基础上提出了一种多模式的视觉目标跟踪算法, 该算法采用了一个控制模式更新率的策略, 通过对不同模型下的跟踪置信度进行分析, 自适应地对模型进行更新, 有效地解决跟踪过程中目标剧烈外观变化或微弱位移变化, 实现干扰条件下目标的稳定跟踪; 对基准视频数据的仿真实验结果表明, 提出的改进的算法可以明显改善跟踪精度和稳定性, 适合工程应用。

关键词: 目标跟踪; 多模式更新; 相关滤波; 刚性变化; 自适应匹配

Object Tracking Algorithm Based on Improved MOSSE Correlation Filter Algorithm

Ji Gang, Gao Fudong, Fan Jiali

(Department of Carrier Aviation Security and Station Management, Naval Aviation University, Qingdao 266041, China)

Abstract: In order to achieve robust tracking performance for occlusion, noise, and drifts, a multiple model update strategy for visual tracking algorithm is proposed to adapt the target appearance changes. Moreover, for adapting complex appearance change, an update rate strategy has been proposed. The multiple models update strategy for visual tracking algorithm can process both infinitesimally small movements and abrupt changes simultaneously, which can effectively solve the severe appearance changes and weak displacement in the process of object tracking. Simulation experiments results have demonstrated that compared to its single model tracking algorithm, our proposed strategy can improve significantly the tracking stability and accuracy in benchmark datasets, which is suitable for military equipment applications.

Keywords: object tracking; multiple model update; correlation filter; rigidity change; self-adaption matching

0 引言

目标跟踪是计算机视觉研究领域的热点之一, 过去几十年以来, 目标跟踪的研究取得了长足的发展, 和许多机器视觉问题一样, 目标跟踪也是在计算复杂度与跟踪鲁棒性之间权衡。跟踪过程中影响跟踪的每一种情况都会带来不同程度的挑战。为了应对这些挑战, 并获得一个强大而鲁棒的跟踪算法, 需要提出一个普适的复杂算法。在实际目标跟踪过程中, 当目标机动性高且形变较大, 对目标的跟踪非常困难。在此需求条件下, 研究者提出了大量的算法来解决这一问题, 但这些算法大都不具备普适性, 且不能用于实时跟踪。大部分基于生成模型的跟踪算法都是以相关跟踪算法为基础, 它易于实施、跟踪精度高, 但其计算量大、易累积误差, 使其不利于在实际跟踪系统中使用, 伴随着信号处理技术和数字技术的成熟, 视频监控系统在数字化、网络化、智能化和集成化方向飞速发展, 整个视频监控行业已经快速进入了智能监控的时代。经过多年发展, 监控行业已经积累了大量的数据资源, 非常满足大数据为基础的人工智能算法模型训练的要求; 同时, 视频监控行业的事前预防、事中响应、事后追查的特性刚好吻合了人工智能的算法和技术。因此, 在云计算、大数据、芯片、智能算法等技

术的助推下, 视频监控行业也从单一的安全领域向多行业应用、提升生产效率、提高生活智能化方向发展。

视觉跟踪算法的一个重要模块就是适应跟踪目标的外观变化。这可以在跟踪开始时通过在线学习, 或者更新目标模板来实现。文 [1] 中提出, 在跟踪期间可以对前景和背景建模, 跟踪过程对模型进行更新, 其中跟踪过程和模型更新可以同时进行。对模型/模板更新的核心问题之一是目标该何时更新。对于外观变化适应性以及模板更新时机, 若干文献进行了研究^[2-10], 然而这些算法都不适应不同外观变化率。跟踪过程的采用固定的更新率不适合视频目标变换的多样性, 噪声、光照的干扰也会给模板引入干扰。因此, 最近一些研究提出了具有自适应性的高效算法, 如文献 [8] 提到的算法及其基于可控的模板更新算法。另一些方法使用生成模型对目标进行跟踪^[9-10]。在该方法中, 生成模型由不同的姿态的目标构造, 并使用最有可能目标样本进行自适应性训练。多实例学习方法^[11]已成为时下流行的跟踪的方法, 如 MILTrack^[12]就是基于多实例学习方法。鲁棒且跟踪性能良好的跟踪方法例还有 IVT^[13], STRUCK^[14], 和基于 NMF 的跟踪方法^[15]。这些研究提出的方法具有较高的鲁棒性, 但是计算量高于基于相关的跟踪算法。

大多数的相关跟踪算法都是生成目标模型以便更新目标外观的变化。目标模型的更新速率与目标的外观变化率成正比。如果目标外观变化剧烈, 应增加更新速率以保持跟踪稳定性。另一方面, 使用高更新率会使目标模型对目标环境敏感, 但对目标非常小的运动不敏感 (尤其是亚像素运动)。为了克服这

收稿日期: 2018-03-19; 修回日期: 2018-03-26。

作者简介: 纪纲 (1985-), 男, 江苏省镇江市人, 硕士, 讲师, 主要从事计算机科学与技术 (仿真)、机械工程及其自动化及人工智能方向的研究。

种折衷带来的跟踪不稳现象, 本文在 MOSSE 相关滤波算法基础上提出了一种多模式的视觉目标跟踪算法, 该算法不仅具有相关算法的实时性, 还适应跟踪目标外观大幅度变化情况。对基准视频数据的仿真实验结果表明, 与对应的单模型跟踪算法相比, 本文提出的算法可以明显改善跟踪精度和稳定性。

MOSSE 相关滤波算法在目标跟踪的任务中, 越来越多的方法开始利用深度学习来进行性能的提升, 但是基于 MOSSE 算法思想的传统方法依然占据一席之地, 本文中, 我们提出的跟踪方法是以相关跟踪为基础。由于 MOSSE 跟踪算法^[11]是一种具有鲁棒性的有效方法, 我们在这一章节简要概括 MOSSE 方法的。

假定训练目标样本为 f_i , 它们的对应期望输出为 g_i 。这些目标图像和期望输出的 FFT 都分别以大写字母 F_i 和 G_i 表示。MOSSE 方法的优化目的是在公式 (1) 中找到一个最优滤波器 H_{opt} 。

$$H_{opt} = \operatorname{argmax}_{H^*} \sum_i |F_i \otimes H^* - G_i|_2^2 \quad (1)$$

如文献 [8] 所示, 通过推导, 等式 (1) 可以得到 H_{opt} 的解析解, 其表达式如下:

$$H_{opt}^* = \frac{\sum G_i \otimes F_i^*}{\sum F_i \otimes F_i^*} \quad (2)$$

因此, 基于少量训练样本并使用公式 (2) 对滤波器初始化, 就可以采用下列公式在跟踪时对模板更新:

$$H_i^* = \frac{A_i}{B_i} \quad (3)$$

$$A_i = \mu G_i \otimes F_i + (1 - \mu) A_{i-1} \quad (4)$$

$$B_i = \mu F_i \otimes F_i + (1 - \mu) B_{i-1} \quad (5)$$

可以看出, MOSSE 滤波器是误差最小均方滤波器, 通过找出一个合适的滤波器使其在目标上的响应最大值。在实际过程中, 考虑到目标外观变换以及环境的复杂性, 需要参考目标的多个图像以便增加跟踪的鲁棒性。但是在实际跟踪的过程中我们要考虑到目标的外观变换等因素的影响, 所以需要同时考虑目标的 m 个图像作为参考, 从而提高滤波器模板的鲁棒性, 所以作者就提出了 MOSSE 这个模型, 然而, 如何处理目标的大尺度变化和微弱位移这两种看似矛盾的情况, 是实现目标稳定跟踪的关键问题。因此, 本文在相关滤波的基础上, 结合 MOSSE 滤波器的最优滤波的策略, 提出了一种多模式的自适应外观跟踪算法。

1 本文改进的算法

目标跟踪就是在连续的视频序列中, 建立所要跟踪物体的位置关系, 得到物体完整的运动轨迹。给定图像第一帧的目标坐标位置, 计算在下一帧图像中目标的确切位置。在运动的过程中, 目标可能会呈现一些图像上的变化, 比如姿态或形状的变化、尺度的变化、背景遮挡或光线亮度的变化等。

本文在相关跟踪的基础上提出了一个多模式的自适应外观模型。该模型可以在跟踪期间, 同时处理目标的微量位移变化和迅速形变。因此, 我们提出的自适应跟踪方法可以用来抵消目标不同层次的视觉变化。

为了解决由于目标遮挡以及光线等问题带来的影响, 同时为了解决不同层次的视觉变化, 可以使用一种多模型目标跟踪策略, 如公式 (6) 所示。

$$\begin{aligned} H_i^1 &= (1 - \mu_1) H_{i-1}^1 + (\mu_1) H_{current} \\ H_i^2 &= (1 - \mu_2) H_{i-1}^2 + (\mu_2) H_{current} \\ &\dots \\ H_i^N &= (1 - \mu_N) H_{i-1}^N + (\mu_N) H_{current} \end{aligned} \quad (6)$$

其中: H_i^l 是第 l 个目标模型, μ_l 是对应的目标模型的更新率, 其中 $\mu_l < \mu_{l+1}$ 。使用公式 (6) 中的方程组, 得到每一帧中不同更新速率的不同模型。基于该策略, 使用 MOSSE 算法的评价准则作为跟踪模板评价指标, 那么一个目标模型的跟踪得分就可以视为该目标模型在候选图像上的跟踪质量。这些模型的跟踪质量的评价, 可以通过公式 (7) 进行量化。跟踪质量可以定义为跟踪器的一个置信度量。如果任何跟踪质量超过了预定的阈值, 跟踪算法就会报告该目标的位置。因为我们假定目标外观通常会发生极微小的变化, 选择最小更新率的模型, 并且其跟踪质量超过质量阈值, 用于对目标的定位。

$$Q = \max F^{-1}(F \otimes H) \quad (7)$$

找到目标位置后, 在频域更新模板。更新方法是通过引入学习速率 μ , 如下所示。

$$H_i = (1 - \mu) H_{i-1} + (\mu) H_{current} \quad (8)$$

其中: $H_{current}$ 是当前计算模板, H_i 和 H_{i-1} 分别表示相邻模板。因此, 使用傅里叶逆运算的线性变换, 可以得到如下等式。

$$Q_i(\mu) = \max(1 - \mu) F^{-1}(F \otimes H_{i-1}) + \mu F^{-1}(F \otimes H_{current}) \quad (9)$$

可以看出, 由于 $H_{current}$ 被设定为当前帧的最佳滤波器 F , H_{i-1} 设定为前一帧的先前的滤波器, 因此 $\max(F \otimes H_{current})$ 一般会大于 $\max(F \otimes H_{i-1})$ 。那应该增大该值, 以满足最优匹配效果。因此, 使用我们提出的算法策略, 可以确定较低跟踪质量的更新率。

滤波器 H_i 和 μ_l 都以升序初始化, 以确定多个模型的更新率。在每一帧, 滤波器的响应和相应的质量, 按更新速率 μ_l 升序计算。如果一个跟踪质量超过于阈值, 则假定目标成功地定位, 这个滤波器更新策略开始在定位的位置使用当前计算滤波器 $H_{current}$ 。超过质量阈值的滤波器 H_i 的 μ_l , 被用于更新不会超过质量阈值的较低更新率的滤波器 (H_i^k 使得 $k < l$)。具有较高更新率的滤波器 (H_i^k 使得 $k > l$), 使用相应的更新率来更新。简言之, 我们提出的更新策略暂时增加了每个模型的更新速率, 直至该模型的跟踪质量超过最低质量标准。

2 实验结果及其分析

为了更加全面地验证本文提出的方法的跟踪能力, 本文选择了基准视频外序列进行定性定量的实验分析。所有的序列都是针对单一目标的跟踪, 具有剧烈运动、相似背景、遮挡、旋转、尺度等特点。

2.1 参数设置

为了验证算法的有效性, 本文采取的 4 个滤波器进行分析。整个实验中, 所有参数都是固定的, 更新率被定为 $\mu_1 = 0.005$, $\mu_2 = 0.005$, $\mu_3 = 0.1$, $\mu_4 = 0.25$ 。本文中 MOSSE 的更新率保持在 0.12。此外, 在实验过程中, 本文使用高斯窗抑制背景干扰, 其覆盖范围为超过目标大小 0.3 倍。使用跟踪框和基准框的重叠率作为目标成功跟踪的标准。如果重叠区域为 50%, 那么可以确定该帧被成功跟踪。为了算法公平比较, 本文采用的对比算法都使用作者提供的源代码, 参数及其初始化值为默

认值。

2.2 定性定量分析

图 2 展示了各算法的跟踪性能比较,其中黑色框是目标基准位置,绿色跟踪框为 MOSSE 跟踪结果,红色则是我们提出算法的结果。实验分析发现当具有较低更新率的滤波器 H_1 ,其质量降低至低于最小 PSR 值时,由于滤波器 H_2 或滤波器 H_3 开始在新一帧中使用一个较高的更新率, H_1 的质量开始增加。当超过最小 PSR 值后, H_1 滤波器开始使用对应更新率。这表明本文提出的算法完全适合目标的外观变化。几乎所有的序列中,本文提出算法定位精度都优于 MOSSE 方法。

表 1 不同数据集中 MOSSE 和提出的跟踪算法的跟踪成功率和定位误差

| 基准序列 | MOSSE | Proposed | MOSSE | Proposed |
|-----------|-------|----------|-------|----------|
| | 成功率 | 成功率 | 误差 | 误差 |
| Dudek | 78 | 100 | 12 | 9 |
| Sylvester | 97 | 97 | 5 | 5 |
| Carl1 | 56 | 100 | 4 | 2 |
| Coke | 86 | 86 | 11 | 1 |
| Faceoccl | 7 | 66 | 9 | 4 |
| Girl | 56 | 94 | 12 | 12 |
| Bike | 98 | 99 | 5 | 7 |
| Shaking | 5 | 82 | 16 | 11 |
| Animal | 59 | 100 | 10 | 10 |
| Football | 79 | 78 | 9 | 9 |
| Singer1 | 100 | 100 | 9 | 10 |

表 1 表示不同数据集中 MOSSE^[8] 和我们提出的跟踪算法的跟踪成功率和定位误差。除了足球序列之外,在所有的序列中,我们提出的方法都比 MOSSE 方法的性能更好。这可能是因为 MOSSE 方法的更新率完全适合“足球”序列中目标的外观变化率。在几乎所有的序列里,我们提出的方法定位目标都优于 MOSSE 方法。



图 1 各算法的跟踪性能比较,左图粗线框是目标基准位置,细线跟踪框为 MOSSE 跟踪结果,右图粗线框是本文提出算法的结果。

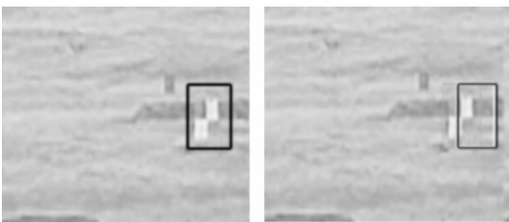


图 2 亚像素级微弱平移的跟踪结果(左右图相差 500 帧,白色跟踪框为 MOSSE 跟踪结果,黑色则是我们提出算法的结果)

图 2 表示,当目标存在一个极小运动时,本文算法的性能比较。由于这种情况的更新率相对较高,MOSSE 方法^[8]不能

处理这种情况,但是由于逐步外观变化使具有最低更新率的滤波器活跃,我们提出的方法可以跟踪这个微弱位移的目标。

3 结论

运动目标跟踪是计算机视觉领域的研究热点,在制导导航,人工智能等领域应用非常广泛,本文针对各种复杂场景及不同目标,如何设计实现效率高、鲁棒性好、实时性强的目标跟踪识别技术仍然是当今业界研究的热点及难点。本文通过分析更新率和跟踪稳定性的相关性,提出了基于改进 MOSSE 的跟踪更新率方法。仿真实验结果表明,该方法优于现有基于多模型框架的跟踪方法,可以处理在目标微小运动时较高的外观变化率的情形,并保持适当的计算复杂度和实时跟踪。所提出的方法是可扩展并应用到任何具有更新策略的视觉跟踪方法,下一步工作主要还是针对目标跟踪的精度上进一步深入研究。

参考文献:

- [1] Chockalingam P, Pradeep N, Birchfield S. Adaptive fragments—based tracking of non-rigid objects using level sets [A]. Computer Vision 2009 [C]. 2009; 1530–1537.
- [2] Balan A O, Black M J. An adaptive appearance model approach for model-based articulated object tracking, in Computer Vision and Pattern Recognition [A]. IEEE Computer Society Conference on Computer Vision & Pattern Recognition, 2006, 1: 758–765.
- [3] Matthews I, Ishikawa T, Baker S. The template update problem [J]. IEEE Trans. Pattern Anal. Mach. Intell., 2004, 26 (6): 810–815.
- [4] 吴宝成. 粒子滤波重采样算法研究及其应用 [D]. 哈尔滨: 哈尔滨工业大学, 2011.
- [5] Kishore M S, Rao K V. Robust correlation tracker [J]. Sadhana, 2001, 26 (3): 227–236.
- [6] 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.
- [7] 姚剑敏. 粒子滤波跟踪方法研究 [D]. 长春: 中国科学院长春光学精密机械与物理研究所, 2014: 17–19.
- [8] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters [A]. Computer Vision and Pattern Recognition (CVPR) [C]. 2010: 2544–2550.
- [9] Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 99: 1–1.
- [10] 连峰, 韩崇昭, 刘伟, 等. 多模型概率假设密度平滑器 [J]. 自动化学报, 2010, 36 (7): 939–950.
- [11] Leistner C, Saffari A, Bischof Horst. Miforests: Multiple-instance learning with randomized trees [A]. Proceedings of the 11th European Conference on Computer Vision: Part VI [C]. Berlin, Heidelberg, 2010, ECCV' 10, 2010: 29–42.
- [12] Babenko B, Yang M H, Belongie S. Visual tracking with online multiple instance learning [A]. CVPR 2009. IEEE Conference on Computer Vision and Pattern Recognition [C]. 2009: 983–990.
- [13] Ross D A, Lim J, Lin R S, et al. Incremental learning for robust visual tracking [J]. International Journal of Computer Vision, 2008, 77: 125–141.
- [14] Hare S, Saffari A, Torr P H S. Structured output tracking with kernels [A]. 2011 IEEE International Conference on Computer Vision (ICCV) [C]. 2011, 263–270.