

遮挡情况下多尺度压缩感知跟踪

张丽娟, 纪峰, 常霞, 李泽仁

(北方民族大学 数学与信息科学学院, 宁夏 银川 750021)

摘要: 针对现有在线学习跟踪算法中目标在线模型更新错误导致跟踪漂移的问题, 提出一种在线模型自适应更新的目标跟踪算法: 首先利用压缩感知技术的高效性, 对多尺度图像特征进行降维, 并提取多尺度样本来实现目标尺度自适应更新, 再由提取的正负样本低维图像特征训练朴素贝叶斯分类器, 利用分类器输出置信度最大处目标样本完成目标跟踪, 并依据当前目标置信度来自适应在线模型更新速率, 减少了遮挡带来的目标错误更新。实验表明: 该方法在尺度变化、局部和全局遮挡、光照变化等情况下均能完成鲁棒跟踪, 平均跟踪成功率较原始压缩感知跟踪算法提高了 20.3%。

关键词: 目标跟踪; 在线学习; 压缩感知

中图分类号: TP391

文献标识码: A

文章编号: 1001-8891(2015)12-1052-06

Multi-scale Compression Perception Tracking under Occlusion

ZHANG Li-juan, JI feng, CHANG Xia, LI Ze-ren

(Northern University for Nationalities, Mathematics and Information Science, Yinchuan 750021, China)

Abstract: In order to deal with the drift problem by updating error in current online learning tracking algorithms, a new adaptive update tracking algorithm is proposed. First of all, based on the efficiency of compressed sensing, the multi-scale image feature space is decreased, and multi-scale samples are exacted to update the target scale. Secondly, a naive Bayes classifier is trained by low dimension image features from positive and negative samples. Experimental results show that the proposed algorithm can complete the robust tracking under the condition of scale changes, partial and full occlusion, illumination changes, etc. Tracking successful rate is improved by 20.3% compared with the original compressive tracking.

Key words: object tracking, online learning, compressive sensing

0 引言

目标跟踪是当前计算机视觉领域内的一个重要研究方向, 在视频监控、运动分析、目标识别等方面得到了广泛应用^[1], 为了完成鲁棒的目标跟踪任务, 目标在线模型受到了广泛关注, 而基于目标在线模型, 当前跟踪算法可分为生成式跟踪算法^[2-4]和判别式跟踪算法^[5-8]。

生成式跟踪算法通过提取目标信息来完成目标表示, 然后利用在线学习到的目标模型在当前帧中基于最小重构误差完成目标定位, 但该类方法忽略了判别性较强的背景信息。判别式跟踪算法将目标跟踪构建为一个二元分类问题, 同时提取目标与背景信息来训练一个分类器, 由分类器决定当前帧目标位置, 因

此也称为基于检测的跟踪算法 (tracking by detection)。Grabner 等^[5]提出了一种在线半监督 boosting 跟踪算法, 利用未标签样本克服了分类器错误更新而导致的跟踪漂移问题; Babenko 等^[6]将多示例学习的思想引入目标跟踪, 利用包含有正负样本的“样本袋”来代替单一样本送入分类器训练, 提高了分类器的判别能力; 近来 Kalal 等^[7]提出了一种结构化约束准则来提升分类器对正负样本的选择, 提升了分类器的抗背景干扰能力。然而这类算法普遍复杂度较高, 为此限制了它们的实时应用。

文献[8]在充分利用上述两类算法的各自优势的基础上, 提出了一种实时跟踪算法——压缩感知跟踪 (Compressive Tracking, CT), 通过建立一个随机投影矩阵对多尺度图像特征进行降维, 并通过朴素贝叶斯

收稿日期: 2015-07-20; 修订日期: 2015-10-05.

作者简介: 张丽娟, 女, 河北省沧州人。硕士研究生, 主要研究方向为图形图像处理。E-mail: 372160407@qq.com。

基金项目: 国家自然科学基金项目, 编号: 61440044, 61102008, 61462002, 61163017; 宁夏自然科学基金项目, 编号: NZ13097, 国家民委科研项目, 编号: 14BFZ003。

分类器进行分类，其高效性引起了学者的广泛关注：文献[9]通过过饱和和分块提升了CT的抗局部遮挡能力，但计算复杂度也随之增高；文献[10]利用模板匹配策略提升了分类器的判别能力，文献[11]用互补的多个矩阵代替CT中的投影矩阵。然而上述算法依然存在以下问题：第一，缺乏对目标尺度的自适应更新；第二，缺乏对遮挡的有效判断，进而导致目标在线模型错误更新。

针对以上两个问题，本文在利用压缩感知理论提取图像特征的高效性的基础上，采取多尺度样本来实现目标尺度自适应更新，在由分类器得到当前帧目标位置后，依据当前目标置信度来自适应在线模型更新速率，减少了遮挡带来的目标错误更新，进而能够完成鲁棒的跟踪任务。

1 压缩感知跟踪算法

随机投影矩阵 \mathbf{R} 是压缩感知理论中的重要概念，其作用是将高维图像特征 \mathbf{X} 压缩成低维特征 \mathbf{V} ，如下式：

$$\mathbf{V} = \mathbf{R}\mathbf{X} \quad (1)$$

式中： $\mathbf{R} \in \mathbb{R}^{n \times m}$ ， $\mathbf{X} \in \mathbb{R}^m$ ， $\mathbf{V} \in \mathbb{R}^n$ ，且 $n \ll m$ ，基于最小误差重构，文献[8]选取的随机投影矩阵为：

$$r_{ij} = \sqrt{s} \times \begin{cases} 1 & \text{with probability } 1/2s \\ 0 & \text{with probability } 1-1/s \\ -1 & \text{with probability } 1/2s \end{cases} \quad (2)$$

式中： $r_{ij} \sim N(0,1)$ ， s 在2或3之间随机选择，因此随机投影矩阵中的大量元素为零，大大减少了计算量。

给定样本 $\mathbf{Z} \in \mathbb{R}^{w \times h}$ ，它的高维图像特征 \mathbf{X} 被定义为一列的多尺度矩形滤波器：

$$h_{ij}(x, y) = \begin{cases} 1, & 1 \leq x \leq i, 1 \leq y \leq j \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

式中： i 和 j 是矩形滤波器的宽和高； w 和 h 是样本窗口的宽和高。通过公式(1)高维特征 \mathbf{X} 将被压缩成低维

特征 \mathbf{V} ，文献[8]中采取类Haar特征的加权组合并采取积分图像加速，如图1所示。

给定样本 $\mathbf{Z} \in \mathbb{R}^{w \times h}$ ，低维特征表示为 $\mathbf{V} = (v_1, \dots, v_n)^T$ ，样本标签表示为 $y \in \{1, 0\}$ ，1代表正样本，0代表负样本。文献[8]通过朴素贝叶斯分类器来建模：

$$H(\mathbf{V}) = \lg \left(\frac{\prod_{i=1}^n p(v_i | y=1) p(y=1)}{\prod_{i=1}^n p(v_i | y=0) p(y=0)} \right) = \sum_{i=1}^n \lg \left(\frac{p(v_i | y=1)}{p(v_i | y=0)} \right) \quad (4)$$

其中：

$$p(v_i | y=1) \sim N(\mu_i^1, \sigma_i^1), p(v_i | y=0) \sim N(\mu_i^0, \sigma_i^0) \quad (5)$$

对所有采样样本计算其 $H(\mathbf{V})$ 值，选择最大的候选样本作为最终跟踪结果，再在其周围采样新的正样本，计算其特征值对公式(5)进行更新：

$$\begin{aligned} \mu_i^1 &\leftarrow \lambda \mu_i^1 + (1-\lambda) \mu^1, \\ \sigma_i^1 &\leftarrow \sqrt{\lambda (\sigma_i^1)^2 + (1-\lambda) (\sigma^1)^2 + \lambda(1-\lambda) (\mu_i^1 - \mu^1)^2} \end{aligned} \quad (6)$$

式(6)中： μ_i^1 ， σ_i^1 分别是正样本图像特征的均值和方差； λ 为学习率，原始算法中固定为0.85。

2 本文算法

从以上的分析可以发现CT算法存在两大问题：一是式(3)中多尺度滤波器采用固定大小，一旦目标尺度发生变化将无法提取到正确的特征信息；二是在线模型中采用固定学习速率，当目标受到遮挡或光照变化等因素影响时模型将产生更新错误。为此本文提出改进策略：采取多尺度样本来实现目标尺度自适应更新；依据当前目标置信度来自适应在线模型更新速率。

2.1 多尺度更新机制

CT算法式(3)中多尺度滤波器（即图像特征）的位置和大小固定不变，当目标尺度发生变化时将不再适用，而为了自适应目标尺度变化，本文首先将图像特征根据滤波器窗口大小归一化，如下式：

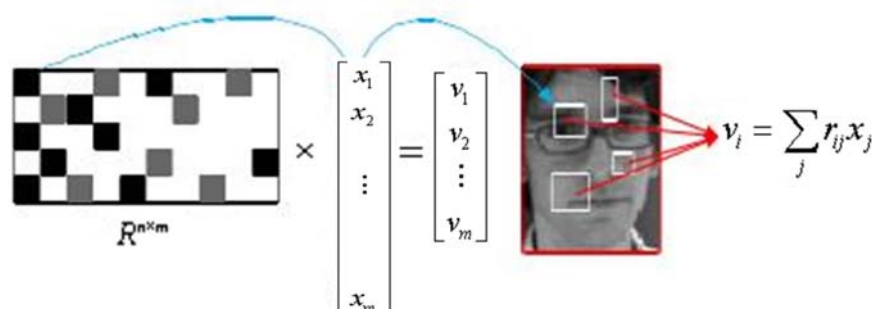


图1 随机投影矩阵提取特征

Fig.1 Feature extraction of random projection matrix

$$h_{ij}(x,y)=\begin{cases} 1/ij, & 1\leq x\leq i, 1\leq y\leq j \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

式中： i 和 j 是矩形滤波器的宽和高；给定初始目标样本 $Z\in\mathbb{R}^{w\times h}$ ， w 和 h 是样本窗口的宽和高，并设定该状态下尺度 $s=1$ ，当目标尺度发生变化时，式(1)中随机投影矩阵也应当随之发生改变，如图2所示。

$R_0^{m\times n}$ 中的每一个元素表示为 $R_0^{m\times n}(r_x, r_y, r_w, r_h)$ ，其中 r_x 、 r_y 、 r_w 、 r_h 分别代表矩形滤波器的位置、宽和高。在尺度 s 下， $R_s^{m\times n}$ 的每一个元素则表示为 $((\text{round}(r_x s), (\text{round}(r_y s), (\text{round}(r_w s), (\text{round}(r_h s)))$ 。

2.2 学习率自适应更新机制

从式(6)可以发现CT算法采用固定学习率更新在线模型，当目标受到遮挡或光照变化等因素影响时模型将产生更新错误，容易导致跟踪漂移。实验表明由分类器 $H(V)$ 得到的每一帧目标置信度值是不一样的，因此可以采用当前帧目标置信度值来决定式(6)中学习率 λ 的大小，如下式：

$$\lambda_t = \begin{cases} 0 & \text{if } H_t(V) < 0 \\ \exp(-(H_t(V) - H_{\max}(V))^2) & \text{else} \end{cases} \quad (7)$$

式中： $H_t(V)$ 代表第 t 帧的目标置信度值； $H_{\max}(V)$ 代表当前所有帧目标置信度值中的最大值； λ_t 代表第 t 帧学习率的大小，当目标被局部遮挡或受光照变化影响时，目标置信度值较小，所以 λ_t 取较小值，而当 $H_t(V) < 0$ 时，则说明样本被分类器判定为负样本，当前目标处于被完全遮挡状态，学习率应当置为0，目标在线模型不再更新。

2.3 目标定位

给定第 t 帧图像以及搜索半径 γ ，以第 $t-1$ 帧目标位置 l_{t-1} 为中心，采样样本集合为 $D^\gamma = \{z \mid \|l(z) - l_{t-1}\| < \gamma\}$ ，得到 D^γ 中每一个样本的低维

特征 $V(z)=(v_1, \dots, v_n)^T$ ，后由式(4)计算分类器响应值 $H(V)$ ，根据响应值最大值输出当前帧目标位置 l'_t ，定义尺度空间 $S=\{a^{-n}, \dots, a^0, \dots, a^n\}$ ，对每一个 $s\in S$ ，分别计算分类器响应值，根据响应值最大值输出当前帧目标位置 l_t 。

再以 l_t 为中心，给定搜索半径 α ， β ，和 ζ ，采样正样本集合 $D^\alpha = \{z \mid \|l(z) - l_t\| < \alpha\}$ 和负样本集合 $D^{\zeta, \beta} = \{\zeta < z \mid \|l(z) - l_t\| < \beta\}$ ，提取两个集合中样本的低维特征，由式(7)对分类器参数进行更新。

3 实验与结果分析

实验采用常见的具有3段不同挑战性的视频，并与原始CT算法^[8]以及多示例学习MIL算法^[6]进行对比分析，实验利用OpenCV2.4.9在VS2010上进行调试。视频主要特性如表1所示。

表1 实验视频的不同特性
Table 1 Different characteristics of experimental videos

视频名称	帧数	主要特性
Faceocc2	812	局部遮挡，外观变化
David	761	光照变化，外观变化
Car4	659	光照变化，尺度变化

3.1 参数设置

分类器学习率初始化为 $\lambda_1=0.85$ ，搜索窗口半径 $\gamma=25$ ，正样本采样半径 $\alpha=4$ ，负样本内外采样半径 $\zeta=8$ 和 $\beta=30$ ，尺度更新机制中 $a=1.02$ ， $n=1$ ，也即尺度空间共3个尺度。

3.2 结果分析

采用文献中常见的两个评估标准来对本算法的性能进行测试，一是跟踪成功率(success rate, SR)，

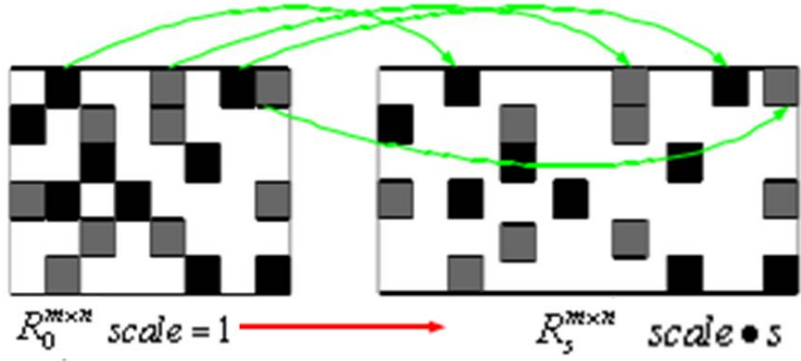


图2 多尺度随机投影矩阵
Fig.2 Multi-scale random projection matrix

定义 $\text{score} = \frac{\text{ROI}_T \cap \text{ROI}_G}{\text{ROI}_T \cup \text{ROI}_G}$ 代表每一帧跟踪结果的得分，其中 ROI_T 代表跟踪结果区域， ROI_G 代表目标真实区域，如果 $\text{score} > 0.5$ 则认为跟踪结果正确；二是中心位置误差 $\sqrt{(x_T - x_G)^2 + (y_T - y_G)^2}$ ，其中 (x_T, y_T) 是跟踪结果的目标位置， (x_G, y_G) 是目标真实位置，表2给出了3种算法在测试视频中的成功率对比，图3给出了不同算法的中心位置误差对比，可见本文算法在3组测试视频中基本保持在20 pixels内。下面结合图4、图5、图6对本文算法鲁棒性进行分析。

图4显示了本文算法的抗遮挡能力，Suv序列在第185帧左右目标逐步被树木遮挡，CT算法由于采用固定学习率不断将错误信息更新到在线模型而逐渐丢失目标，MIL算法在第350帧左右目标被遮挡后也出现了同样的问题，而本文算法一旦判定完全遮挡

则不再更新在线模型，全程完成了鲁棒跟踪任务。

表2 算法跟踪成功率对比

Table 2 Algorithm tracking success rate			
Sequence	Proposed	CT	MIL
Suv	89.1	49.1	40.2
David	84.8	89.1	68.2
Car4	83.2	58.1	56.3
Average SR	85.7	65.4	54.9

图5显示了本文算法的抗光照变化能力，David序列中目标从暗处走向亮处，且目标多次出现形变，CT算法和MIL算法由于缺乏对目标信息的有效判断，跟踪框逐渐偏移目标，而本文算法的全程鲁棒跟踪显示了本文算法的抗光照变化能力。

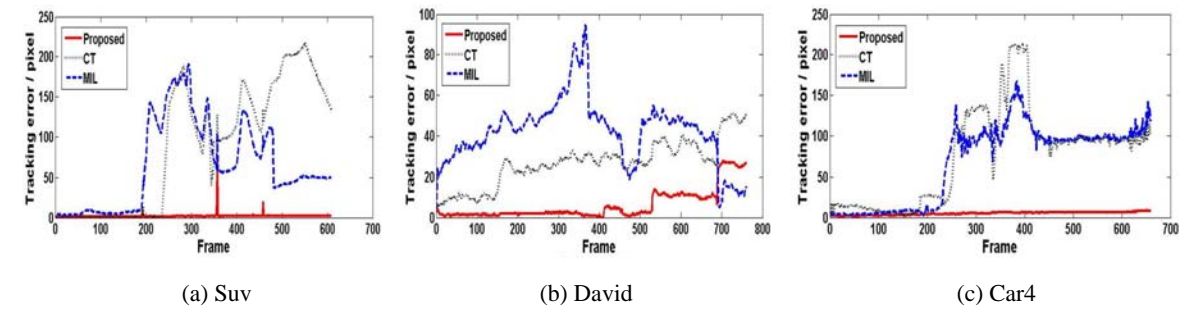


图3 视频序列跟踪误差 Fig.3 Video sequence tracking error

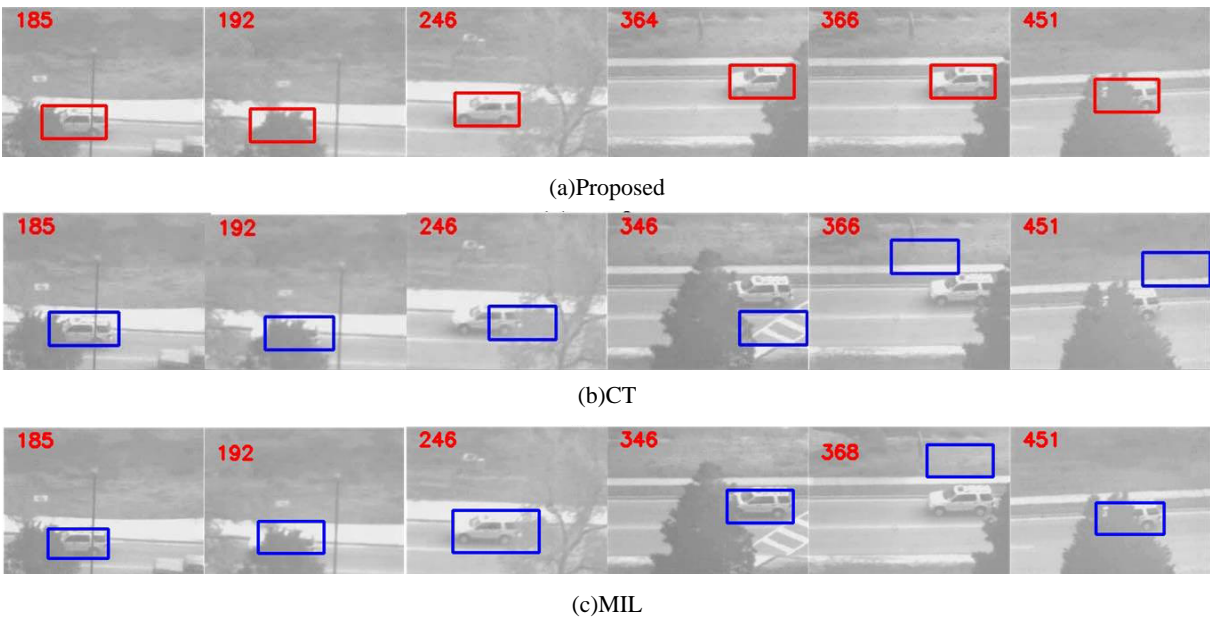


图4 Suv序列跟踪结果 Fig.4 Suv sequence tracking results

图6显示了本文算法的抗尺度变化能力, Car4序列中的目标尺度越来越小, 本文算法的多尺度机制能够不断捕捉到这种尺度变化, 因此跟踪框随着目标减小而减小, 且能对目标进行很好的定位, 而CT算法和MIL算法缺乏尺度更新机制在目标从桥下走出时丢失了目标。

4 结论

针对现有在线学习跟踪算法中存在的跟踪漂移问题, 本文在压缩感知跟踪算法的基础上, 采用在线模型自适应更新机制以及目标尺度自适应更新机制, 提出了一种简单且高效的目标跟踪算法。三组不同挑战性的实验表明: 本文方法在目标被长时或短时遮挡、光照变化及尺度变化等情况下均能完成鲁棒跟踪, 跟踪成功率较原始压缩感知算法提高了20.3%。

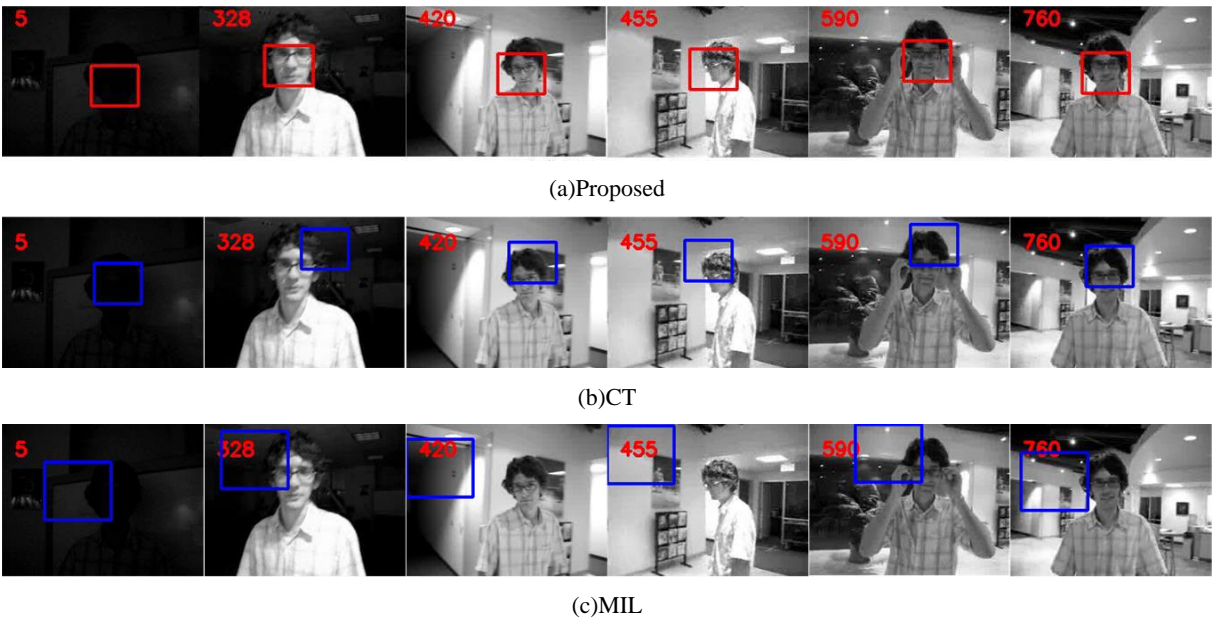


图5 David序列跟踪结果

Fig.5 David sequence tracking results

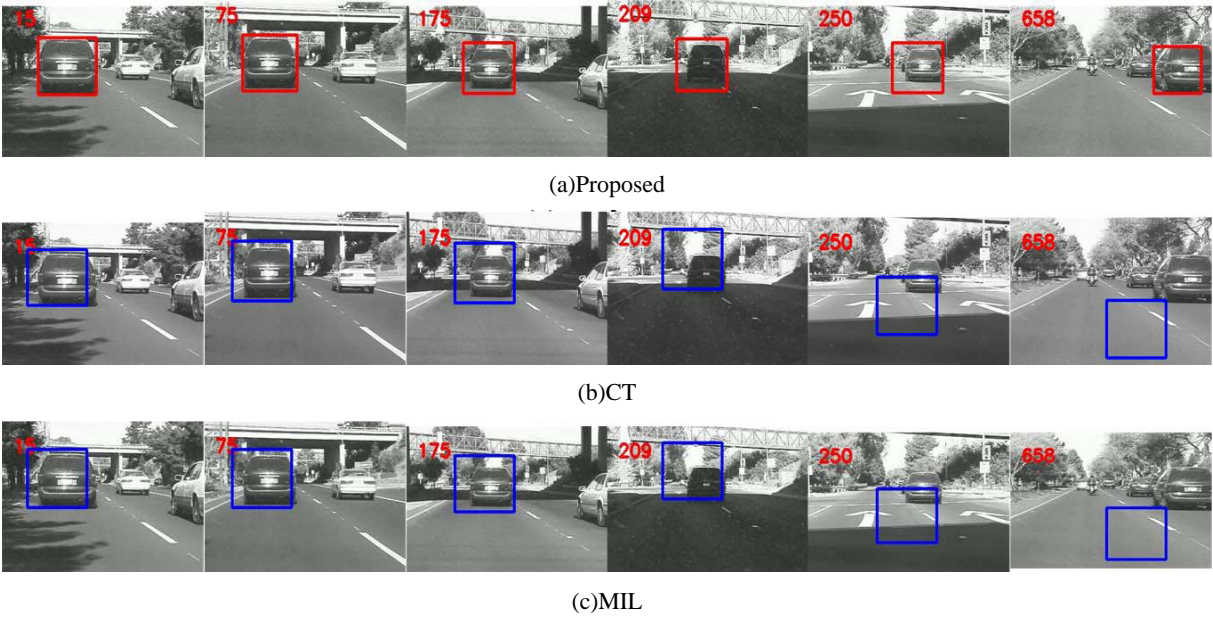


图6 Car4序列跟踪结果

Fig.6 Car4 sequence tracking results

参考文献:

[1] Wu Y, Lim J, Yang M H. Online object tracking: a benchmark[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2013: 2411-2418(doi: 10.1109/CVPR.2013.312).

[2] Ross D, Lim J, Lin R, et al. Incremental learning for robust visual tracking[C]//International Conference on Computer Vision, 2008, 77(1-3): 1056

- 125-141.
- [3] Mei X, Ling H. Robust visual tracking and vehicle classification via sparse representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(11): 2259-2272.
- [4] Li H, Shen C, Shi Q. Real-time visual tracking using compressive sensing[C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2011: 1305-1312 (doi: <http://doi.ieeecomputersociety.org/10.1109/CVPR.2011.5995483>).
- [5] Grabner H, Leistner C, Bischof H. Semi-supervised on-line boosting for robust tracking[C]//*European Conference on Computer Vision*, 2008, Part I: 234-247.
- [6] Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**: 1619-1632.
- [7] Kalal Z, Matas J, Mikolajczyk K. Tracking-learning-detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**: 1409-1422.
- [8] Zhang K H, Zhang L, Yang M H. Real time compressive tracking [C]//*European Conference on Computer Vision*, 2012, **7574**: 864 - 877.
- [9] Zhu Qiuping, Yan Jia, Deng Dexiang. Compressive tracking via over-saturated sub-region classifiers[J]. *IET Computer Vision*, 2013, **17**(6): 448-455.
- [10] Jing Jing, Xu Guangzhu, Lei Bangjun. Using appearance rematching to improve compressive tracking[C]//*Proceedings of the fifth International Conference on Internet Multimedia Computing and Service*, 2013: 28-33(doi: 10.1145/2499788.2499825).
- [11] 朱秋平, 颜佳, 张虎, 等. 基于压缩感知的多特征实时跟踪[J]. *光学精密工程*, 2013, **21**(2): 437-444.
- Zhu Qiuping, Yan Jia, Zhang Hu, et al. Real-time tracking using multiple features based on compressive sensing[J]. *Optics and Precision Engineering*, 2013, **21**(2): 437-444.