

# 数据驱动的智能车辆目标检测能力测试评价方法<sup>①</sup>

张学显<sup>1</sup>, 刘伟<sup>2</sup>, 余彪<sup>2</sup>, 许铁娟<sup>2</sup>, 周鹏飞<sup>2</sup>

<sup>1</sup>(中国科学技术大学 计算机科学与技术学院, 合肥 230000)

<sup>2</sup>(中国科学院 合肥物质科学研究院 应用技术研究所, 合肥 230000)

**摘要:** 本文针对智能车辆目标检测能力测评存在的指标体系不完整、量化程度和测评实时性低等问题, 聚焦智能车辆目标检测能力中的目标分类和目标识别, 在这两个测评项目上提出了一套量化的评价指标体系, 并用 TOPSIS 方法进行综合的评价. 然后在此指标体系的基础上搭建数据驱动的智能车辆目标检测能力测评平台, 平台可满足对智能车辆目标检测能力测评的实时性要求. 最后采用了若干组车辆检测算法对指标体系进行验证.

**关键词:** 智能车辆; 目标检测; 测评体系

引用格式: 张学显, 刘伟, 余彪, 许铁娟, 周鹏飞. 数据驱动的智能车辆目标检测能力测试评价方法. 计算机系统应用, 2017, 26(11): 249-253. <http://www.c-s-a.org.cn/1003-3254/6043.html>

## Data Driven Test and Evaluation Method for Intelligent Vehicle Object Detection Capability

ZHANG Xue-Xian<sup>1</sup>, LIU Wei<sup>2</sup>, YU Biao<sup>2</sup>, XU Tie-Juan<sup>2</sup>, Zhou Peng-Fei<sup>2</sup>

<sup>1</sup>(School of Computer Science and Technology, University of Science and Technology of China, Hefei 230000, China)

<sup>2</sup>(Institute of Applied Technology, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230000, China)

**Abstract:** To address the problems such as incomplete index system, low degree of quantification and real-time evaluation of test and evaluation for intelligent vehicle object detection capability, we put forward a set of quantitative evaluation index system on object classification and object recognition. Then, we conduct a comprehensive evaluation by TOPSIS. According to the index system, we establish the data driven platform for intelligent vehicle object detection ability evaluation, the platform can also meet the real-time requirements of the evaluation on intelligent vehicle object detection ability. Finally, several groups of vehicle detection algorithms are used to verify the index system.

**Key words:** intelligent vehicle; object detection; test and evaluation system

近些年, 智能车辆领域被广泛地关注, 智能车辆相关核心技术也取得了巨大的发展. 智能车辆是集合了环境感知、决策规划、运动控制、平台执行、人机交互众多关键技术的复杂系统. 作为智能车辆的基础, 智能车辆的环境感知能力也是研究的热点, 目前对环境感知的研究主要集中在目标检测和环境建模的相关领域, 其中环境建模也是以目标检测的结果为输入, 所以目标检测能力对智能车辆的自主性至关重要. 而智能车辆目标检测能力的发展离不开合理的测试评价系统, 因为它不仅明确了目标检测的相关指标, 同时还对智

能车辆目标检测技术的发展方向做出指导.

目前对智能车辆的测评方法主要有实车测试、仿真测试和数据驱动的测试. 在实车测试方面, 美国国防部先进研究项目局 (DARPA) 首先主持了针对沙漠环境的无人驾驶挑战赛和针对城市道路的无人驾驶挑战赛<sup>[1]</sup>. 欧盟委员会也举办了多届地面机器人试验赛. 从 2009 年起, 中国自然科学基金委员会的“视听觉信息的认知计算”重大研究计划持续开展了“中国智能车未来挑战赛”<sup>[2]</sup>. 但是实车测试成本较高、可控性较低, 且实车测试多是任务导向, 无法单独对智能车辆目标检测能力

① 基金项目: 国家自然科学基金 (61503362, 61305111, 91420104); 安徽省自然科学基金 (1508085MF133)

收稿时间: 2017-02-20; 修改时间: 2017-03-09; 采用时间: 2017-03-13

进行测评;而仿真测试因输入的是仿真环境模型,数据还原度较低,一般进行智能车辆的决策规划能力测评而不进行感知能力的测评;而数据驱动的测评方式以其高安全低成本、可控可重复等优点受到很多学者的青睐,它的主要思想是把已经采集的数据回放给待测智能车辆,然后对待测车辆的处理结果进行测评,如机器视觉领域的 PASCAL VOC 挑战赛和 ImageNet 数据库、以及目前被智能车辆广泛使用的 KITTI<sup>[3]</sup>,但是 KITTI 等平台也存在很多问题,如评价指标较少(只提供平均精度),对智能车辆要求较高的实时性的测评关注度不够(仅有代码运行时间,且仅精确到 0.1s),而最终的排名也仅以平均精度一项决定,无法合理全面的对智能车辆的目标检测能力进行综合评价。

针对以上问题,本文聚焦数据驱动的智能车辆目标检测能力测评,提出智能车辆目标检测力量化评价指标(包括性能指标和实时性指标),并对部分指标的计算方法做出优化,然后综合以上指标并利用 TOPSIS 对智能车辆目标检测能力进行多维度的综合的评价。最后在此测评指标上建立基于服务器/客户端模式的测评系统,测评系统可对智能车辆的目标检测耗时进行精确的测量,可以对测评指标的合理性进行验证。

## 1 智能车辆目标检测能力测评指标体系

本文聚焦智能车辆目标检测能力测评,主要对目标分类和目标识别两个测评项目提出量化指标,下面分别就这两个项目的测评指标进行阐述:

### 1.1 目标分类能力测试评价指标

目标分类能力测试主要用于测评智能车辆对于特定目标的识别能力。对于给定的分类目标,待测智能车辆需返回当前测试样本是否有此给定的目标。

此测试项目的输入为:测试样本集合(相机图片,可以是不连续帧数据), $I = \{S_1, S_2, \dots, S_n\}$ ,其中:每个样本的数据结构为 $S_i = (ID_i, DATA_i, CLS_i)$ ,其中 $ID_i$ 为每个样本的编号, $DATA_i$ 为具体的样本数据, $CLS_i$ 为对应样本的标注类别,即当前样本中需识别的目标类别(只考虑单分类问题)。

此测试项目的输出为:测试结果集合, $O = \{R_1, R_2, \dots, R_n\}$ ,其中每个测试结果的数据结构为 $R_i = (ID_i, ANS_i)$ ,其中 $ID_i$ 为对应测试样本编号, $ANS_i$ 为每个样本的测试结果(即当前测试样本是否含有需要识别的目标)。

此测试项目的评价指标有:精度、召回率、误检率、漏检率、F-measure、平均耗时和耗时标准差。我们希望精度和召回率都尽可能的高,但是这两者在某些情况下是矛盾的,这样就需要用 F-measure 来综合考虑它们, F-measure 是精度和召回率的加权调和平均,我们采用最常见的 F1-measure,其计算公式如下:

$$F1\text{-measure} = \frac{2 \cdot P \cdot R}{P + R} \quad (1)$$

### 1.2 目标识别能力测试评价指标

目标识别能力测试相对于目标分类能力测试难度更大,对于给定的测试样本集合,智能车辆需返回所有的目标位置与类别。

此测试项目的输入为:测试样本集合(相机图片,可以是不连续帧数据), $I = \{S_1, S_2, \dots, S_n\}$ ,其中:每个样本的数据结构为 $S_i = (ID_i, DATA_i, OBS_i)$ , $ID_i$ 为每个样本的唯一标识, $DATA_i$ 为具体的样本数据, $OBS_i$ 为对应样本标注的目标信息,数据结构为 $OBJ_i = \{ob_{i1}, ob_{i2}, \dots, ob_{in_i}\}$ , $n_i$ 为对应样本标注的目标数量, $ob_{ij}$ 可以表示为 $ob_{ij} = (id_{ij}, cls_{ij}, x_{ij}, y_{ij}, w_{ij}, l_{ij})$ , $id_{ij}$ 为每个目标的编号, $cls_{ij}$ 为每个目标标注的类别, $x_{ij}$ 、 $y_{ij}$ 、 $w_{ij}$ 以及 $l_{ij}$ 分别为对应目标的中心位置、宽度以及高度。要求测试程序采用矩形对检测与识别的对象进行标注。

此测试项目的输出为:测试结果集合, $O = \{R_1, R_2, \dots, R_n\}$ ,其中每个测试结果的数据结构为 $R_i = (ID_i, ANS_i)$ , $ID_i$ 为对应测试样本的编号, $ANS_i$ 为测试结果,其结构表示为 $ANS_i = \{d_{i1}, d_{i2}, \dots, d_{im_i}\}$ , $m_i$ 为第 $i$ 各样本上检测的目标数量( $m_i$ 不一定等于 $n_i$ ), $d_{ij}$ 可以表示为 $d_{ij} = (id_{ij}, cls_{ij}, x_{ij}, y_{ij}, w_{ij}, l_{ij})$ , $id_{ij}$ 为每个目标的编号, $cls_{ij}$ 为每个目标识别的类别, $x_{ij}$ 、 $y_{ij}$ 、 $w_{ij}$ 以及 $l_{ij}$ 分别为对应目标的中心位置、宽度以及高度。

此测试项目的评价指标有:平均检测精度、精度、召回率、F-measure、平均耗时、耗时标准差、漏检率、误检率。平均检测精度计算方法如下,其余指标同上。

平均识别精度 (Average detection precision, ADP):

$$ADP = \frac{1}{\sum_{i=1}^n m_i} \sum_{i=1}^n \sum_{j=1}^{m_i} (w_1 \alpha_{ij} + w_2 \beta_{ij}) \quad (2)$$

其中,

$$\alpha_{ij} = \frac{S_{A \cap B}}{S_{A \cup B}} \quad (3)$$

$$\beta_{ij} = \exp\left(-\frac{(S_B - 1)^2}{2}\right) \quad (4)$$

对于任意一帧测试样本  $S_i, A_{ij} \in OBJ_i, B_{ij} \in ANS_i$ ,  $A$  与  $B$  的匹配集合  $M_i = \{(A_1, B_1), (A_2, B_2), \dots, (A_{imi}, B_{imi})\}$  为使  $ADP$  达到最大的一个匹配,  $\alpha_{ij}$  表示  $A$  与  $B$  的重合度,  $\beta_{ij}$  表示表示  $A$  与  $B$  面积的差距,  $w_1$  和  $w_2$  分别取 0.5.  $S_A$ 、 $S_B$  分别为  $A$  与  $B$  覆盖的面积.

要正确求解  $\alpha_{ij}$  和  $\beta_{ij}$ , 关键在于确定匹配对  $M_i$ , 具体算法如下.

算法 1. 确定检测匹配对  $M$

Step1. 构建关系矩阵:  $R = (r_{ij})_{ni \times mi}$ ,

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1mi} \\ r_{21} & r_{22} & \dots & r_{2mi} \\ \dots & \dots & \dots & \dots \\ r_{ni1} & r_{ni2} & \dots & r_{nmi} \end{bmatrix} \quad (5)$$

Step2. 构建最优化模型: 可以采用 0-1 指派问题进行建模与求解<sup>[4]</sup>,

$$\begin{aligned} \max z &= \sum_{i=1}^{ni} \sum_{j=1}^{mi} r_{ij} x_{ij} \\ s.t. & \begin{cases} \sum_{j=1}^{mi} x_{ij} \leq 1 & i = 1, 2, \dots, ni \\ \sum_{i=1}^{ni} x_{ij} \leq 1 & j = 1, 2, \dots, mi \\ x_{ij} \in \{0, 1\} & i = 1, 2, \dots, ni; j = 1, 2, \dots, mi \end{cases} \end{aligned} \quad (6)$$

通过求解上述优化模型, 可以得到所有  $x_{ij}$  的值, 其中  $x_{ij}=1$  表示对应的  $o_i$  与  $d_j$  为正确匹配.

Step3. 确定识别检测匹配对是否满足目标分类要求:

```

for i = 1 to |M|
    if  $M_i.ob.cls \neq M_i.d.cls$ 
         $M = M \setminus M_i$ 
    end if
end for
    
```

1.3 综合评价方法——TOPSIS

对于以上测试得出的测评数据, 采用 TOPSIS 方法对其进行综合评价, TOPSIS 在多目标决策分析中是一种非常有效的方法<sup>[5]</sup>, 利用 TOPSIS 我们可以对进行测评的智能车辆在每个测试项目上进行排序, 也可以综合所有项目的测评指标, 对其进行综合排序. 其步骤如下.

(1) 构建规范化矩阵

利用测评结果构建初始矩阵  $S$ , 对于成本型指标, 需对其进行取倒数的处理, 并根据公式 (7) 构建规范化矩阵  $S'$ .

$$x'_{ij} = x_{ij} / \sqrt{\sum_{k=1}^n x_{ik}^2}, i = 1, 2, 3 \dots m; j = 1, 2, 3 \dots n \quad (7)$$

$$S = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & \dots & \dots & x_{mn} \end{bmatrix} \rightarrow S' = \begin{bmatrix} x'_{11} & x'_{12} & \dots & x'_{1n} \\ x'_{21} & x'_{22} & \dots & x'_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x'_{m1} & \dots & \dots & x'_{mn} \end{bmatrix} \quad (8)$$

(2) 根据各指标的权重构建权重规范化矩阵

利用权重矩阵  $W$  获得加权决策矩阵  $Z$ :

$$Z = S' \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & w_n \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1n} \\ z_{21} & z_{22} & \dots & z_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ z_{m1} & \dots & \dots & z_{mn} \end{bmatrix} \quad (9)$$

(3) 确定理想解  $Z^+$  与反理想解  $Z^-$

$$Z^+ = (Z_{i1}^+, Z_{i2}^+ \dots Z_{in}^+), Z_{ij}^+ = \max(Z_{1j}, Z_{2j} \dots Z_{mj}) \quad (10)$$

$$Z^- = (Z_{i1}^-, Z_{i2}^- \dots Z_{in}^-), Z_{ij}^- = \min(Z_{1j}, Z_{2j} \dots Z_{mj}) \quad (11)$$

(4) 计算与理想解和反理想解的距离

分别计算每个测评结果与理想解和反理想解的欧氏距离  $D^+$  和  $D^-$ :

$$D_i^+ = \sqrt{\sum_{j=1}^n (Z_{ij} - Z_{ij}^+)^2}; D_i^- = \sqrt{\sum_{j=1}^n (Z_{ij} - Z_{ij}^-)^2} \quad (12)$$

(5) 计算理想解的贴近程度  $C_i^+$ , 根据贴近程度进行排序, 贴近度越小, 即距离最优解距离越小, 排名越高, 贴近度的计算公式如下:

$$C_i^+ = D_i^- / (D_i^+ + D_i^-), i = 1, 2, \dots, m. \quad (13)$$

本文中, 初始化矩阵  $S$  中  $x_{ij}$  为各智能车辆在各指标上的得分. 权重矩阵中  $w_j$  都取 1, 即各指标权重相等, 在不计算时间指标的排名中, 将权重矩阵中平均耗时和耗时标准差两个指标对应的  $w_j$  置为 0 即可. 最后计算出的  $C_i^+$  就是各智能车辆与当前最优解的距离. 距离越近 (即  $C_i^+$  越小) 排名越靠前.

2 数据驱动的智能车辆目标检测能力测评平台

本文所介绍的数据驱动的智能车辆目标检测能力

测试评价平台的设计思想是:利用采集完毕的数据和构建的虚拟传感器对待测智能车辆的目标检测能力进行可控可重复的测评.平台采用服务器/客户端的架构模式.

## 2.1 服务器端

服务器端即测试系统端,主要负责构建虚拟传感器,模拟真实传感器的数据发送,同时对待测智能车辆的感知能力做出评价;服务器端为每种类型的数据类型构建一个虚拟传感器,每个虚拟传感器包括:数据读取线程,用于读取测试数据,并将测试数据格式化为定义的测试样本数据格式;消息发送队列,用于存储数据读取线程格式化的测试样本数据;消息发送线程,用于根据客户端的订阅情况并发送给相应的客户端.

## 2.2 客户端

客户端即待测智能车辆端,用于辅助待测智能车辆进行测评.每个客户端包括:多个数据接收线程,用于接收来自服务器端发来的对应类型的测试样本数据;消息接收队列,用于存储数据接收线程接收的测试样本数据.同时客户端提供消息订阅、数据读取和测试结果回送服务.在智能车辆读取数据时记录时间  $T_1$ ,回送测试结果时记录时间  $T_2$ ,最后由测试系统将  $T=T_2-T_1$  添加到测试结果中,这样就可以准确的统计待测智能车辆在每帧测试样本上的处理时间,以便服务器端对待测智能车辆的平均耗时和耗时标准差进行评价.

## 2.3 平台实现

基于 vs2012 开发数据驱动的智能车辆目标检测能力测评平台,如图 1 所示.

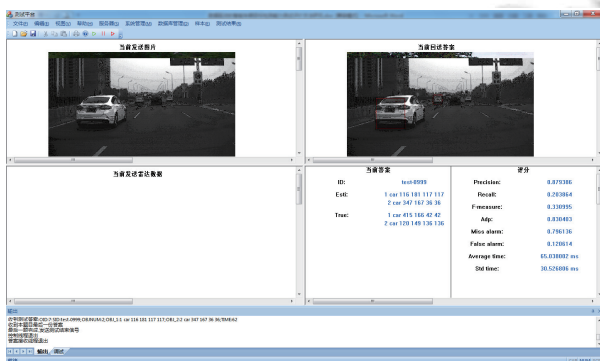


图 1 智能车辆目标检测能力测评平台

## 3 实验和分析

基于数据驱动的智能车辆目标检测能力测评平台进行实验,实验所运行的环境为:Windows7/32位系统,

CPU@3.40 GHz, 内存 2 G, 本次实验所采取的样本为 500 幅车辆样本(正样本)+500 幅行人样本(负样本),如图 2 所示.检测的目标为车辆,待测智能车辆目标检测算法采用 adaboost,使用 5 级强分类器,因算法有限,调整算法的扫描倍率(即本次扫描范围相对于上一次扫描范围的倍率)分别用 1.1 倍, 1.3 倍, 1.5 倍, 1.8 倍的扫描倍率当作四组不同的实验<sup>[6]</sup>,较小的扫描倍率应该有更好的性能,但是更耗时.实验结果如下(漏检率和误检率可以通过精度和召回率计算,故未给出).

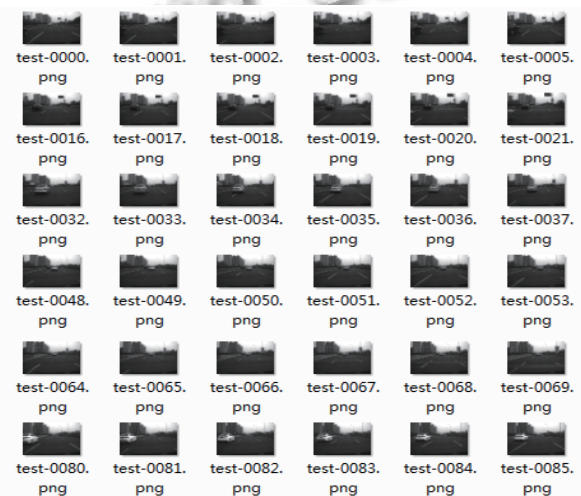


图 2 部分测试样本

### 3.1 目标分类实验结果和分析

目标分类实验结果如表 1 所示.

从整体来看,不考虑时间因素,扫描倍率越小性能越好,小的扫描倍率更加耗时,所以在加上时间因素综合排名,大的扫描倍率综合排名较好,综合来看 1.5 倍扫描倍率最好.从单个指标来看,随着扫描倍率的增加,召回率呈明显下降趋势,虽然精度出现波动,但是这是以召回率下降作为代价的,所以能综合评价精度和召回率的 F-Measure 也呈下降趋势,符合实验预期.

### 3.2 目标识别实验结果和分析

目标识别实验结果如表 2 所示.

从整体来看,不考虑时间因素,仍然是扫描倍率越小性能越好,但是大的扫描倍率综合排名较好,综合来看 1.5 倍扫描倍率最好.从单个指标来看,随着扫描倍率的增加,召回率呈明显下降趋势,精度也出现波动,但是综合的评价指标 F-Measure 和 ADP 基本呈下降趋势,基本符合实验预期.

表1 目标分类实验结果

扫描倍率	精度	召回率	F-Measure	平均耗时(ms)	耗时标准差(ms)	排名	不计时排名
1.1倍	0.972458	0.976596	0.974522	313.791	135.012	4	1
1.3倍	0.997475	0.840426	0.91224	112.239	51.965	3	2
1.5倍	0.997151	0.744681	0.852619	55.963	27.400	1	3
1.8倍	0.993174	0.619149	0.762778	61.389	30.266	2	4

表2 目标识别实验结果

扫描倍率	精度	召回率	F-Measure	ADP	平均耗时(ms)	耗时标准差(ms)	排名	不计时排名
1.1倍	0.840595	0.402135	0.544016	0.868659	332.458008	139.870422	4	1
1.3倍	0.919708	0.320285	0.475113	0.869279	119.196999	47.992977	3	2
1.5倍	0.952128	0.273005	0.424338	0.851685	63.937	27.886332	1	3
1.8倍	0.879386	0.203864	0.330995	0.830403	66.666	31.295443	2	4

#### 4 结语

智能车辆的目标检测能力作为其自主性的基础,重要性不言而喻,而其相关关键技术的发展离不开合理的测评体系. 本文就目标分类和目标识别两个测评项目给出了量化的测评体系,并在此基础上构建了智能车辆目标检测能力测评平台,最后进行若干组实验对测评体系加以验证,实验证明,此测评体系可以满足对智能车辆目标检测能力进行综合评价的要求.

#### 参考文献

- Mason R, Radford J, Kumar D, *et al.* The golem group/university of California at Los Angeles autonomous ground vehicle in the DARPA grand challenge. *Journal of Field Robotics*, 2006, 23(8): 527–553. [doi: 10.1002/(ISSN)1556-4967]
- 郝盛, 张红卫. “智能车挑战赛”车辆参赛资格测试规范的研究. *专用汽车*, 2012, (6): 95–97.
- Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite. *Proc. of 2012 IEEE Conference on Computer Vision and Pattern Recognition*. Providence, RI, USA. 2012. 3354–3361.
- 张惠珍, 魏欣, 马良. 求解 0-1 线性整数规划问题的有界单纯形法. *运筹学学报*, 2014, 18(3): 71–78.
- Shih HS, Shyur HJ, Lee ES. An extension of TOPSIS for group decision making. *Mathematical and Computer Modelling: An International Journal*, 2007, 45(7-8): 801–813. [doi: 10.1016/j.mcm.2006.03.023]
- Lampe A, Chatila R. Performance measure for the evaluation of mobile robot autonomy. *Proc. of 2006 IEEE International Conference on Robotics and Automation*. Orlando, FL, USA. 2006. 4057–4062.