



中华人民共和国国家标准

GB/T XXXXX—XXXX

机器人智能化视觉评价方法及等级划分

Robot intelligent vision evaluation method and classification

(征求意见稿)

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX - XX - XX 发布

XXXX - XX - XX 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 符号和缩略语	2
5 视觉智能等级划分	2
6 测试要求	4
7 测评方法	8
参考文献	14

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国机械工业联合会提出。

本文件由全国机器人标准化技术委员会（SAC/TC591）归口。

本文件起草单位：上海机器人产业技术研究院有限公司、节卡机器人股份有限公司、美的集团（上海）有限公司、复旦大学、上海电器设备检测所有限公司、上海添唯认证技术有限公司、上海电器科学研究所（集团）有限公司……

本文件主要起草人：

引 言

机器人视觉是机器人智能化发展的重要方向之一，为机器人应用扩大了应用场景及应用能级。近年来机器人视觉的应用呈现爆发式增长，导致技术水平和规模层次不齐，亟需标准化规范行业的持续稳定发展。

本文件为了解决机器人视觉技术的测试评价上受到的瓶颈，满足对机器人视觉科学全面的测试评价的需要而制定。

通过制定本文件，明确机器人智能化视觉评价的等级划分与测评要求，给出具体的测评方法，旨在指导工业机器人、服务机器人、特种机器人的智能化视觉评价活动。

机器人智能化视觉评价方法及等级划分

1 范围

本文件规定了机器人智能化视觉评价的视觉智能等级划分与测试要求、描述了机器人智能化视觉的测评方法。

本文件适用于工业机器人、服务机器人、特种机器人的智能化视觉评价，不适用于公共安全领域使用的机器人的人脸识别检测。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 2423.10 环境试验 第2部分：试验方法 试验Fc：振动（正弦）

GB/T 4208 外壳防护等级（IP代码）

GB/T 17799.1 电磁兼容 通用标准 居住、商业和轻工业环境中的抗扰度

GB/T 17799.2 电磁兼容 通用标准 第2部分：工业环境中的抗扰度标准

GB/T 35273 信息安全技术 个人信息安全规范

GB/T 35678—2017 公共安全 人脸识别应用 图像技术要求

GB/T 38427.1—2019 生物特征识别防伪技术要求 第1部分：人脸识别

GB 17799.3 电磁兼容 通用标准 第3部分：居住环境中设备的发射

GB 17799.4 电磁兼容 通用标准 第4部分：工业环境中的发射

3 术语和定义

下列术语和定义适用于本文件。

3.1

机器人 robot

具有两个或两个以上可编程的轴，以及一定程度的自主能力，可在其环境内运动以执行预期任务的执行机构。

[来源：GB/T 12643—2013，2.6，有修改]

3.2

智能 intelligence

具有人类或类似人类智慧特征的能力。

注：人类或类似人类的智慧特征，表现为在实现某个目的的过程中，总会经历一个或多个的感知、决策、执行的过程或过程循环。并在其中通过不断学习，提高自身实现目的的能力和实现目的的效率与效果；本文件认为，在体现人类或类似人类的智慧特征上，感知、决策、执行和在其中的学习的各项能力和过程具有不可或缺性。

[来源：GB/T 28219—2018，3.1]

3.3

机器人智能化 Robot intelligence

机器人的感知、认知、决策等功能在非结构化或动态环境下自主作业能力的程度。

3.4

人脸识别 face recognition

以人面部特征作为识别个体身份的一种个体生物特征识别方法。其通过分析提取用户人脸图像数字特征产生样本特征序列,并将该样本特征序列与已存储的模板特征序列进行比对,用以识别用户身份。

注1:从应用方式不同,人脸识别可分为人脸验证和人脸辨识。

注2:人脸验证:人脸识别应用之一,将所产生的样本特征序列与按用户标识信息所给定的已存储的用户的模板特征序列进行比对(1:1比对),以确认用户是否为所声明的身份。

注3:人脸辨识:人脸识别应用之一,将所产生的样本特征序列与已存储的指定范围内的所有模板特征序列进行比对(1:N比对),确定用户身份。

[来源:GB/T 38671—2020, 3.1.2, 有修改]

3.5

目标检测 object detection

确认图像中是否存在指定类别的对象并确定其位置和大小。

3.6

召回率 recall

被正确预测的正样本占全部正样本的比率。

3.7

错误接受率 false acceptance rate

接受不该接受的样本的比率。

[来源:ISO 22300:2021, 3.2.18]

3.8

错误拒绝率 false rejection rate

拒绝不该拒绝的样本的比率。

[来源:ISO 22300:2021, 3.2.19]

4 符号和缩略语

下列缩略语适用于本文件。

FAR: 错误接受率 (False Acceptance Rate)

FRR: 错误拒绝率 (False Rejection Rate)

oks: 对象关键点相似度 (object keypoint similarity);

AP: 平均精度 (Average Precision)

5 视觉智能等级划分

5.1 基本要求

根据机器人应用场景的不同,开展机器人智能化视觉测评活动,得出3类机器人视觉智能等级:

——室内生产环境用机器人视觉智能等级:测试对象如用于3C装配、汽车生产的工业机器人、生产线用自动导引车等;

注:应用场景具有可控光源,光照受天气变化影响小。

- 室内非生产环境用机器人视觉智能等级：测试对象为室内家庭和类似场景、商超、楼宇、办公/餐厅、医院等环境下使用的机器人；
- 室外环境用机器人视觉智能等级：测试对象为室外道路、广场、园区等环境用机器人。

5.2 室内生产环境用机器人视觉智能等级

室内生产环境用机器人视觉智能等级划分与测试指标要求见表1。

表1 室内生产环境用机器人视觉智能等级划分与测试指标要求

测试指标		视觉智能等级				
		L1	L2	L3	L4	L5
人体姿态估计	单人姿态估计	-	-	-	90%~95%	≥95%
	多人姿态估计	-	-	-	90%~95%	≥95%
	人体姿态跟踪	-	-	-	-	≥95%
	3D人体姿态估计	-	-	-	-	≥95%
人脸识别	准确率	-	-	-	-	≥95%
	宏平均准确率	-	-	-	-	≥95%
	微平均准确率	-	-	-	-	≥95%
	FAR	-	-	-	-	≤5%
	FRR	-	-	-	-	≤5%
字符识别	字符识别准确率	-	-	≥90%	≥95%	≥98%
	字符识别召回率	-	-	≥90%	≥95%	≥98%
	平均编辑距离	-	-	≤0.1	≤0.05	≤0.01
物体识别	准确率	-	-	80%~90%	90%~95%	≥95%
	宏平均准确率	-	-	80%~90%	90%~95%	≥95%
	微平均准确率	-	-	80%~90%	90%~95%	≥95%
	平均精度均值	-	-	80%~90%	90%~95%	≥95%
	物体位姿估计	≥1级	≥2级	≥3级	≥4级	≥5级
安全性	≥1级	≥2级	≥3级	≥4级	≥5级	

5.3 室内非生产环境用机器人视觉智能等级

室内非生产环境用机器人视觉智能等级划分与测试指标要求见表2。

表2 室内非生产环境用机器人视觉智能等级划分与测试指标要求

测试指标		视觉智能等级				
		L1	L2	L3	L4	L5
人体姿态估计	单人姿态估计	-	-	-	80%~90%	≥90%
	多人姿态估计	-	-	-	80%~90%	≥90%
	人体姿态跟踪	-	-	-	-	≥90%
	3D人体姿态估计	-	-	-	-	≥90%
人脸识别	准确率	-	-	-	-	≥90%
	宏平均准确率	-	-	-	-	≥90%
	微平均准确率	-	-	-	-	≥90%

表2 室内非生产环境用机器人视觉智能等级划分与测试指标要求（续）

测试指标		视觉智能等级				
		L1	L2	L3	L4	L5
人脸识别	FAR	-	-	-	≤15%	≤10%
	FRR	-	-	-	≤15%	≤10%
字符识别	字符识别准确率	-	-	-	≥90%	≥95%
	字符识别召回率	-	-	-	≥90%	≥9%
	平均编辑距离	-	-	-	≤0.05	≤0.01
物体识别	准确率	≥70%	70%~80%	80%~90%	90%~95%	≥95%
	宏平均准确率	≥70%	70%~80%	80%~90%	90%~95%	≥95%
	微平均准确率	≥70%	70%~80%	80%~90%	90%~95%	≥95%
	平均精度均值	≥60%	60%~70%	70%~80%	80%~90%	≥90%
	物体位姿估计	≥1级	≥2级	≥3级	≥4级	≥5级
安全性		≥1级	≥2级	≥3级	≥4级	≥5级

5.4 室外环境用机器人视觉智能等级

室外环境用机器人视觉智能等级划分与测试指标要求见表3。

表3 室外环境用机器人视觉智能等级划分与测试指标要求

测试指标		视觉智能等级				
		L1	L2	L3	L4	L5
人体姿态估计	单人姿态估计	-	-	-	90%~95%	≥95%
	多人姿态估计	-	-	-	90%~95%	≥95%
	人体姿态跟踪	-	-	-	90%~95%	≥95%
	3D人体姿态估计	-	-	-	90%~95%	≥95%
人脸识别	准确率	-	-	-	90%~95%	≥95%
	宏平均准确率	-	-	-	90%~95%	≥95%
	微平均准确率	-	-	-	90%~95%	≥95%
	FAR	-	-	-	≤10%	≤5%
	FRR	-	-	-	≤10%	≤5%
字符识别	字符识别准确率	-	-	-	≥95%	≥98%
	字符识别召回率	-	-	-	≥95%	≥98%
	平均编辑距离	-	-	-	≤0.05	≤0.01
物体识别	准确率	≥75%	75%~85%	85%~90%	90%~95%	≥95%
	宏平均准确率	≥75%	75%~85%	85%~90%	90%~95%	≥95%
	微平均准确率	≥75%	75%~85%	85%~90%	90%~95%	≥95%
	平均精度均值	≥80%	80%~90%	90%~95%	95%~98%	≥98%
	物体位姿估计	≥1级	≥2级	≥3级	≥4级	≥5级
安全性		≥1级	≥2级	≥3级	≥4级	≥5级

6 测试要求

6.1 一般要求

机器人视觉的测试应包括两方面：

- 机器人视觉算法测试；
- 机器人视觉整机测试。

6.2 测试集要求

6.2.1 基本要求

6.2.1.1 测试集的数据安全应符合 GB/T 35273 的要求。

6.2.1.2 针对不同类型的机器人视觉传感器安装位置，应采用特定的测试集进行测试。

示例1：扫地机器人的测试集视角距离地面 5 cm ~10 cm；

示例2：物流机器人用于人体、物体目标检测的测试集视角距离地面 80 cm ~150 cm；

示例3：工业机器人的测试集视角距离地面 20 cm~350 cm。

6.2.1.3 测试集应具有独立性，确保机器人视觉模型在训练时没有使用过测试集中的数据。

6.2.2 数量要求

测试集数量应不小于20000张。

6.2.3 质量要求

6.2.3.1 旋转

以样本的平面视图为基准，进行 0° ~ 360° 的随机旋转，样本数量不少于测试集总数量的10%。

6.2.3.2 视角

样本数量不少于测试集总数量的10%，测试集对视角的要求应包括：

- 室内生产环境用测试集以样本的正视图为基准，从前、后、左、右四个方向随机旋转 0° ~ 90° ；
- 室内非生产环境、室外环境用测试集以样本的正视图为基准，从左、右两个方向随机旋转 0° ~ 30° ，数量不低于测试集总数量的10%。

6.2.3.3 光照

测试集对光照的要求应包括：

- 室内非生产环境用测试集包含不同时间、不同天气光照强度下的数据，包括晴天室内(100 lx~2000 lx)，阴天室内(5 lx~50 lx)，不同光照下的样本数量不低于测试集的5%；
- 室内生产环境用测试集可不考虑光照因素的影响；
- 室外环境用测试集包含不同时间、不同天气光照强度下的数据，包括晴天室外(30000 lx~100000 lx)，晴天背阴处(5000 lx~10000 lx)，阴天室外(50 lx~1000 lx)，夜晚(0.2 lx~10 lx)，不同光照下的样本数量不低于测试集的3%。

6.2.3.4 分辨率

样本数量不少于测试集总数量的10%，测试集对分辨率的要求应包括：

- 室内非生产环境用测试集分辨率不低于 256×360 像素；
- 室内生产环境用测试集分辨率不低于 512×512 像素；

——室外环境用测试集分辨率不低于 1028×1028 像素。

6.2.3.5 遮挡

样本数量不少于测试集总数量的10%，测试集对遮挡率的要求应包括：

- 室外环境、室内非生产环境用测试集的遮挡率为 0%~80%；
- 室内生产环境用测试集的遮挡率为 0%~30%。

6.2.3.6 椒盐噪声

样本数量不少于测试集总数量的10%，测试集对椒盐噪声的要求包括：

- 室内环境用测试集椒盐噪声比例为 0.1%~1%；
- 室外环境用测试集椒盐噪声比例为 1%~5%。

6.2.3.7 多个物体

样本要求包括：

- 多个不同类物体同时出现在同一画面中的样本；
- 多个同类物体同时出现在同一画面中的样本；
- 数量不低于测试集总数量的10%。

6.2.4 均衡性、数据集标注

6.2.4.1 数据集均衡性应考虑各种类别的样本数量一致程度和数据集样本分布的偏差程度。

示例：人的数据集包括年龄、性别、种族、表情、体型等。

6.2.4.2 数据集标注信息应完备且准确无误。

6.3 算法测评要求

6.3.1 测试设备要求包括：

- 中央处理器：基础频率不低于 3.6 GHz，最大加速频率不低于 5.0 GHz；
- 内存：内存不低于 16 GB，对于某些大规模数据集和复杂任务，内存不低于 32 GB；
- 图形处理器：显存大小不低于 8 GB，主频大于 1.5 GHz；
- 系统：可采用 Ubuntu22.04/Windows11；
- 硬盘：128 GB SSD 及以上。

6.3.2 被测视觉模型接口要求包括：

- 作为 ROS 的独立节点，和 ROS 操作系统连接时，标准化要求；
- 测试集读入、测试结果输出标准化要求。

6.4 整机测试要求

6.4.1 一般要求

应根据应用场景，按照6.4.2~6.4.4中规定的内容进行测试，对应的测试结果应不低于机器人视觉算法测试结果。

注：跨场景应用的机器人宜分别在室内、室外环境下进行整机测试。

6.4.2 室内生产环境

6.4.2.1 电磁兼容性

电磁发射应符合GB 17799.4的要求，电磁抗扰度应符合GB/T 17799.2的要求。

6.4.2.2 气候环境适应性

测试环境温度应为20℃~25℃。

6.4.2.3 机械环境适应性

应在GB/T 2423.10中规定的扫频振动测试要求如表4。

表4 扫频振动试验要求

频率范围 Hz	位移幅值 mm	每一轴线上的 扫频循环次数	要求
10~30~10	0.75	5	样品应按工作位置在三个互相
30~55~30	0.15	5	垂直的轴线上依次振动

6.4.2.4 防护等级

防护等级应满足GB/T 4208规定的IP40的要求。

6.4.3 室内非生产环境

6.4.3.1 电磁兼容性

电磁发射应符合GB 17799.3的要求，电磁抗扰度应符合GB/T 17799.1的要求。

6.4.3.2 气候环境适应性

测试环境温度应为制造商宣称的温度限值。

6.4.3.3 机械环境适应性

应在GB/T 2423.10中规定的扫频振动测试要求如表5。

表5 扫频振动试验要求

频率范围 Hz	位移幅值 mm	每一轴线上的 扫频循环次数	要求
10~30~10	0.75	5	样品应按工作位置在三个互相
30~55~30	0.15	5	垂直的轴线上依次振动

6.4.3.4 防护等级

防护等级应满足GB/T 4208规定的IP40的要求。

6.4.4 室外环境

6.4.4.1 电磁兼容性

电磁发射应符合GB 17799.3的要求，电磁抗扰度应符合GB/T 17799.1的要求。

6.4.4.2 气候环境适应性

测试环境温度应为制造商宣称的温度限值。

6.4.4.3 机械环境适应性

应在GB/T 2423.10中规定的扫频振动测试要求如表6。

表6 扫频振动试验要求

频率范围 Hz	位移幅值 mm	每一轴线上的 扫频循环次数	要求
10~30~10	0.75	5	样品应按工作位置在三个互相
30~55~30	0.15	5	垂直的轴线上依次振动

6.4.4.4 防护等级

防护等级应满足GB/T 4208规定的IP65的要求。

7 测评方法

7.1 混淆矩阵

7.1.1 混淆矩阵:在二分类任务中,混淆矩阵是通过样本的标签和模型分类结果组成的 2×2 矩阵,其中,左上角为真正样本数(真实值和模型预测值同时为正, TP),左下角为假正样本数(真实值为负,而模型预测值为正, FP),右上角为假负样本数(真实值为正,而模型预测值为负, FN),右下角为真负样本数(真实值和模型预测值同时为负, TN),见表 7。

表7 二分类混淆矩阵

真实值	预测值	
	Positive	Negative
Positive	真正 TP	假负 FN
Negative	假正 FP	真负 TN

7.1.2 在多分类任务中,混淆矩阵的每一列代表了预测类别,每一列的总数等于预测为该类别数据的数目;每一行代表了数据的真实归属类别,每一行的数据总数等于该类别数据样本的数目;第 i 行第 i 列的数值表示第 i 类数据被预测为第 i 类的数目,见表 8。

表8 多分类混淆矩阵

真实值	预测值			
	第1类	第2类	...	第n类
第1类	a_{11}	a_{12}	...	a_{1n}
第2类	a_{21}	a_{22}	...	a_{2n}
...
第n类	a_{n1}	a_{n2}	...	a_{nn}

7.2 准确率

准确率为对于给定的数据集,正确分类的样本数占全部样本数的比率,见公式(1)。

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \dots\dots\dots (1)$$

式中：

accuracy——准确率；
 TP——真正样本数；
 TN——真负样本数；
 FP——假正样本数；
 FN——假负样本数。

7.3 宏平均准确率、微平均准确率

7.3.1 多分类任务下的指标计算：在 n 分类任务中，根据表 5 的多分类混淆矩阵得到了所有种类的预测值分布情况，并基于这些数据计算多分类任务的准确率、精度、召回率、F1 测度等指标，在此基础上，进一步采用宏平均和微平均来评价整个多分类任务的功能有效性。

7.3.2 宏平均准确率指对每一个类别分别计算指标值，即把每个类别视作二分类情况进行统计，然后再对所有类的结果取算术平均值，见公式（2）：

$$acro Acc = \frac{1}{n} \sum_{i=1}^n \frac{TP_i + TN_i}{TP_i + FP_i + TN_i + FN_i} \times 100\% \dots\dots\dots (2)$$

式中：

Macro Acc——宏平均准确率；
 TP_i ——i类任务的真正样本数；
 TN_i ——i类任务的真负样本数；
 FP_i ——i类任务的假正样本数；
 FN_i ——i类任务的假负样本数。

7.3.3 微平均准确率指计算每个类别的真正、真负、假正、假负平均值，然后计算类别预测的统计指标，见公式（3）：

$$Micro Acc = \frac{\overline{TP} + \overline{TN}}{\overline{TP} + \overline{FP} + \overline{TN} + \overline{FN}} \times 100\% \dots\dots\dots (3)$$

式中：

Micro Acc——微平均准确率；
 \overline{TP} ——真正样本数的平均值；
 \overline{TN} ——真负样本数的平均值；
 \overline{FP} ——假正样本数的平均值；
 \overline{FN} ——假负样本数的平均值。

7.4 交并比

交并比用来评价目标框和预测框之间的重合度，见公式（4）：

$$IoU = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})} \dots\dots\dots (4)$$

式中：

IoU——交并比；
 B_p ——预测的矩形框区域；
 B_{gt} ——标注的矩形框区域。

7.5 平均重叠率

平均重叠率（AOR）计算每一帧真实标注和算法输出的交并比，对所有帧取平均值。

7.6 错误接受率

错误接受率（FAR）应按照GB/T 38427.1—2019中3.4的要求进行计算。

7.7 错误拒绝率

错误拒绝率（FRR）应按照GB/T 38427.1—2019中3.5的要求进行计算。

7.8 字符识别准确率

字符识别准确率指对于给定的数据集，被正确识别的字符占有所有识别出的字符数的比率，见公式（5）：

$$precision = \frac{TP}{TP+FP} \times 100\% \dots\dots\dots (5)$$

式中：

- precision*——字符识别准确率；
- TP——真正样本数；
- FP——假正样本数。

7.9 字符识别召回率

字符识别召回率指对于给定的数据集，被正确识别的字符占实际字符数的比率，见公式（6）：

$$recall = \frac{TP}{TP+FN} \times 100\% \dots\dots\dots (6)$$

式中：

- recall*——字符识别召回率；
- TP——真正样本数；
- FN——假负样本数。

7.10 平均编辑距离

编辑距离为字符串A到字符串B最少需要的操作次数。操作次数为更改一个字符，删除一个字符或增加一个字符。对整篇文档所有行的编辑距离取平均值，即可求得到平均编辑距离，见公式（7）：

$$AED = \frac{1}{n} \sum_{i=1}^n edit_i \times 100\% \dots\dots\dots (7)$$

式中：

- AED——平均编辑距离；
- edit_i*——字符识别中每一行的编辑距离。

7.11 平均精度均值

对于每一个类别，首先按照置信度（总体参数值落在样本统计值某一区间内的概率）把每个预测结果进行排序，再取不同置信度阈值，把每个预测结果分为真正、假正、真负和假负类，从而获得在该阈值下的精度和召回率值。画出该类别以精度为纵轴，召回率为横轴所绘制的 P-R曲线，该类别的平均准确率就是此P-R曲线下的面积。以上过程遍历所有的类别之后，对所有类别的平均准确率求平均，即可得到模型的平均精度均值（mAP）。

7.12 实时性

视觉实时性是指算法能够在实时或接近实时的速度下处理输入的图像或视频流的能力，计算一秒内处理的图像帧数（FPS），见公式（8）：

$$FPS = \frac{1}{(time_{end} - time_{start})} \dots\dots\dots (8)$$

式中：

FPS——在一秒内处理的图像帧数；

time_{end}——程序结束时间戳；

time_{start}——程序开始时间戳。

7.13 人体姿态估计

7.13.1 人体姿态估计测试指标包括单人姿态估计、多人姿态估计。

7.13.2 对象关键点相似度(oks)通过计算真值和预测人体关键点的相似度，测评人体骨骼关键点检测算法，见公式（9）：

$$OKS = \frac{\sum_i \exp(-d_i^2/2s^2 \cdot v_i)}{\sum_i v_i} \dots\dots\dots (9)$$

式中：

d_i——预测关键点i和真实关键点i之间的欧氏距离；

s——尺度因子，通常设置为关键点标注的最大边缘长度；

v_i——二值指示函数，如果关键点（i）是可见的，则（v_i = 1），否则（v_i = 0）；

∑_iv_i——所有可见关键点的总数。

7.13.3 单人姿态估计通过以任一单张测试集图片中的人进行关键点检测后获得的一组关键点，计算出人与关键点的相似度作为标量，并人为给定阈值T，通过所有图片的 oks 计算平均精度（AP），见公式（10）：

$$AP = \frac{\sum \delta(oks > T)}{\sum 1} \dots\dots\dots (10)$$

式中：

AP——平均精度；

oks——对象关键点相似度；

T——指定的阈值。

7.13.4 多人姿态估计即同时检测和估计多个人体的关键点位置和姿势信息，对复杂场景中多人人体姿态计算平均精度（AP），见公式（11）：

$$AP = \frac{\sum_m \sum_p \delta(oks_{mp} > T)}{\sum_m \sum_p 1} \dots\dots\dots (11)$$

式中：

AP——平均精度；

oks_{mp}——第 m 个图像中的第 p 个人的对象关键点相似度；

T——指定的阈值。

7.14 物体位姿估计

7.14.1 工业场景下机器人智能化视觉物体位姿估计等级划分见表 9。

表9 工业场景下机器人智能化视觉物体位姿估计等级划分

1 级	2 级	3 级	4 级	5 级
具备 2D 视觉识别功能； 可输出位置信息。	具备 2D 视觉识别功能； 可输出位置、姿态信息。	具备 3D 视觉识别功能。 可输出复杂环境中的物体位置、姿态和运动方向估计。 位置误差：≤5 mm； 姿态误差：≤5°； 路径预测误差：≤5 mm； 识别时间：≤30 ms。	高精度 3D 视觉系统，能够在动态环境中实时追踪多个物体，准确估计其位置、姿态和运动方向。 位置误差：≤1 mm； 姿态误差：≤1°； 路径预测误差：≤1 mm； 识别时间：≤20 ms。	具备智能化学习与适应能力，可在未知环境中自主识别和适应新物体及其变化，实现高度自动化和智能化的视觉感知。 位置误差：≤0.5 mm； 姿态误差：≤0.5°； 路径预测误差：≤0.5 mm； 识别时间：≤10 ms。

7.14.2 非工业场景下机器人智能化视觉物体位姿估计等级划分见表 10。

表10 非工业场景下机器人智能化视觉动态物体位姿估计等级划分

1 级	2 级	3 级	4 级	5 级
能够识别和定位简单的物体。	能够识别物体的位置、大小和形状。	具备 3D 视觉识别能力，输出物体的位置、姿态和运动方向。 位置误差：≤10 mm； 姿态误差：≤10°； 路径预测误差：≤10 mm； 识别时间：≤50 ms。	具备 3D 视觉识别能力，在未知环境输出物体的位置、姿态和运动方向。 位置误差：≤5 mm； 姿态误差：≤5°； 路径预测误差：≤5 mm； 识别时间：≤30 ms。	具备智能化学习与适应能力，可在未知环境中自主识别和适应新物体及其变化并，实现高度自动化和智能化的视觉感知； 其它性能指标不低于 4 级。

7.15 安全性

安全性主要包括两方面：视觉系统参与路径规划、导航、避障任务引入的安全风险；人为欺诈引入的安全风险。机器人智能化视觉安全性等级划分见表11。

表11 机器人智能化视觉安全性等级划分

1 级	2 级	3 级	4 级	5 级
基本的视觉系统功能,无视觉系统故障报警功能。	能够检测到视觉系统的故障并发出基本的警报信号,警报的响应速度 ≤ 1 s,无明显误报。	视觉系统故障时,能发出警报信号,警报的响应速度 ≤ 0.3 s。	<ol style="list-style-type: none"> 视觉系统故障时,能发出警报信号,警报的响应速度≤ 0.2 s。 具备应对措施,如切换到备用视觉系统并发出详细的故障报告。 能够有效防止欺骗攻击。 	<ol style="list-style-type: none"> 视觉系统故障时,警报时间≤ 0.017 s; 具备切换到备用视觉系统并生成详细故障报告的功能; 能够预测并主动预防潜在故障,系统能自我修复或通过智能算法调整工作状态,确保任务的连续性和安全性。 能够有效防止欺骗攻击。

参 考 文 献

- [1] GB/T 12643—2013 机器人与机器人装备 词汇
 - [2] GB/T 28219—2018 智能家用电器通用技术要求
 - [3] GB/T 38671—2020 信息安全技术 远程人脸识别系统技术要求
 - [4] GA/T 1755—2020 安全防范 人脸识别应用 人证核验设备通用技术要求
 - [5] T/CESA 1036—2019 信息技术 人工智能 机器学习模型及系统的质量要素和测试方法
 - [6] T/SAITA 001—2021 人工智能 计算机视觉系统测评规范
 - [7] ISO 22300: 2021 Security and resilience — Vocabulary
 - [8] Goodfellow I J, Shlens J, Szegedy C. Explaining and harnessing adversarial examples[C], ICLR. 201
 - [9] Du M, Liu N, Hu X. Techniques for interpretable machine learning[J]. Communications of the ACM, 2019, 63(1):68-77
-