Artificial Intelligence and Robotics Research 人工智能与机器人研究, 2019, 8(4), 239-247
Published Online November 2019 in Hans. http://www.hanspub.org/journal/airr
https://doi.org/10.12677/airr.2019.84027

Hans汉斯

Permission Access Control for Multi-User and Multi-Robot Systems Based on Scene Recognition

Xuling Jin[1], Jianyong Feng[2], Yangwei Sun[2], Ben Liu[1], Jian Ye[2]

[1]Beijing University of Civil Engineering and Architecture, Beijing
[2]Institute of Computing Technology, Chinese Academy of Sciences, Beijing
Email: jye@ict.ac.cn

Received: Nov. 1st, 2019; accepted: Nov. 15th, 2019; published: Nov. 22nd, 2019

Abstract

With the development of cloud robots, robots can participate in more and more different scenes. The rapid development of the Internet is followed by the privacy security issues when multiple users use robots in different scenarios. The robot should have different permissions in different scenarios which can maximize the efficiency and protect the user's privacy. To solve this problem, scene recognition is one way to solve this problem. By pre-setting the rights and restrictions owned by each user in different scenarios, the server determines the scene in which the robot is located by the picture taken by the robot, thereby giving each user different permissions, which allows the robot to listen to the user's instructions to protect user's privacy. The application of this method provides a new way of security protection for the multi-user multi-robot system.

Keywords

Robot, Scene Recognition, Authority Management, Multiuser, Robot Cloud Service

基于场景理解的多用户多机器人系统权限访问控制

靳旭玲[1]，冯建勇[2]，孙扬威[2]，刘　犇[1]，叶　剑[2]

[1]北京建筑大学，北京
[2]中国科学院计算技术研究所，北京
Email: jye@ict.ac.cn

收稿日期：2019年11月1日；录用日期：2019年11月15日；发布日期：2019年11月22日

文章引用: 靳旭玲, 冯建勇, 孙扬威, 刘犇, 叶剑. 基于场景理解的多用户多机器人系统权限访问控制[J]. 人工智能与机器人研究, 2019, 8(4): 239-247. DOI: 10.12677/airr.2019.84027

摘　要

伴随着云机器人的发展，机器人能够参与的场景越来越多，互联网的飞速发展，随之而来的是多用户在不同场景下使用机器人构成的隐私威胁。机器人在不同场景下应拥有不同的权限，在保护用户隐私的同时还可以效率最大化地完成任务。为此，场景理解是解决这一问题的基础，通过预先设定不同场景下各用户所拥有的权利与限制，服务器通过机器人拍摄的图片判断机器人所处的场景，从而赋予各用户不同的权限，使机器人听从用户的指令保护隐私。此举的应用为多用户多机器人系统的安全防护方式提供了一种新的思路。

关键词

机器人，场景理解，权限管理，多用户，机器人云服务

1. 引言

人类具有一种卓越的天赋，那就是面对复杂场景可以准确而又迅速地理解。这项能力十分重要，因为它可以帮助人类推断出当前所处的情境并探索周围的环境。而本文中的云机器人在云服务器的帮助下，面对复杂场景可以迅速而准确地获取场景数据，更好地理解周围环境，而不需要在本地进行推理，既提高机器人的反应速度，又节约了单机机器人成本。场景理解的一个重要应用便是机器人系统权限访问控制[1]。在目前的多数云机器人系统中，可以实现多个用户可以操作多个机器人，用户可以命令机器人执行各种项目例如拍照、移动控制等任务。但是在一些采用机器人系统的大型单位中，为了保护隐私，并不是每个用户都有级别利用摄像头进行一些涉密行为(例如拍照)，如何令系统自行判断一个用户能否进行涉密行为是解决问题的重点。普遍的云机器人系统[2]安全防护方法往往采用依照用户 ID 制定好一张权限管理表单，管理用户权限方式相对固定。如果服务器可以理解场景，那么就可以依据当前机器人所处场景来判断用户有无发布任务的权限。本文就基于场景理解的多用户多机器人系统用户权限访问控制进行研究与设计。

2. 系统分析与设计

系统整体可分为 Android 应用、云平台、机器人三个部分。Android 应用主要负责与用户的交互，通过访问云平台[3]提供的接口展示当前可以执行任务的机器人以及给机器人分配任务。云服务器[4]是整个系统的核心，所有用户的任务借助移动智能设备上传到服务器后，服务器对收到的任务分析处理，所有任务都会解析为一个个包含重要参数的子任务，发布到特定的 URLs，同时一些必要的信息也会发布在平台。与云服务器连接的众多机器人查询到自己当前的任务后，首先进入特定任务界面，获取需要的任务信息并提取重要参数，然后才能执行任务。任务结束后需要返回任务的结束数据到云服务器上，供移动设备端查询当前任务的完成情况。

基于以上的系统框架，本系统将对用户在服务器后台按照不同场景划分权限，服务器对机器人上传的环境图片进行场景理解，根据场景理解结果分用户 ID 进行过滤(访问控制)，系统总体设计架构如图 1 所示。

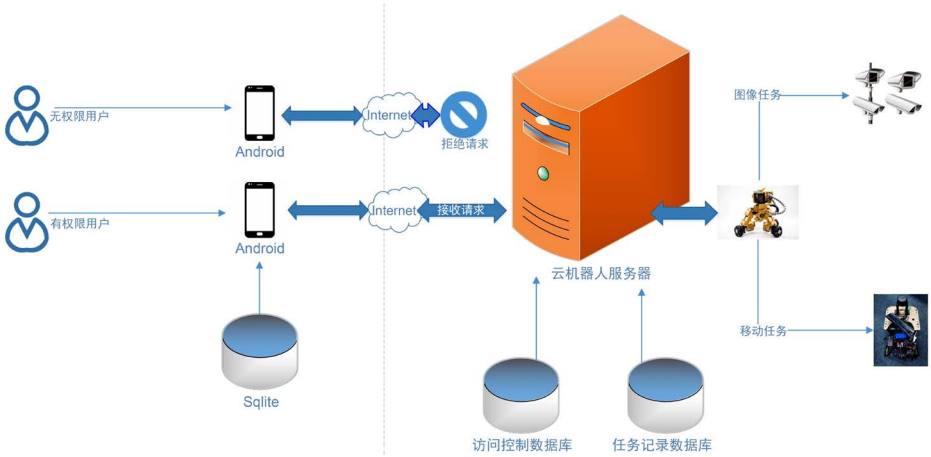

Figure 1. System architecture and working principle
图 1. 系统整体架构与工作原理图

3. 基于场景理解的用户访问控制

3.1. 访问控制的意义

在应用本系统的单位中，并不是所有用户的权限级别都相同。如果在所有场景中的所有用户都有权通过 Android 应用对机器人发布任务，那么单位就会有被泄密的风险。所以依据用户本人对存在的各个场景来划分其权限是十分必要的。

3.2. 注册登录模块

系统将发布命令的控制者和接受命令的机器人都看作用户，提供给用户注册与登录的界面，为避免越权操作，在注册的同时，就将控制者按照 commander_id 划分好了其所在各个场景的权限范围。

3.3. 访问控制模块

本系统服务器在经过将图片识别出具体场景后，根据用户发布任务时的 id 与当前场景进行匹配判断当前场景用户能否发布任务以及发布何种任务，判断通过则发布成功，否则会驳回用户的发布请求。

3.4. 任务发布与任务获取模块

用户通过本系统的 Android 应用登录账号进行发布任务控制机器人行为，发布任务时可选择发布何种任务与命令指定机器人执行任务，任务发布后存储在服务器的数据库中。机器人端通过轮询服务器的任务列表进行任务获取，机器人执行任务与上传任务信息和安卓端的任务发布这三者与服务器之间的消息通信依靠 HTTP 协议[5] [6]。服务器对控制端的 IP 地址 + 端口号 + ID 号的访问，当控制端每发送一个新的请求时，进程会产生新的线程对该请求进行服务[7]。

3.5. 多用户多机器人权限控制

每个用户创建之后都会有一张默认的场景任务权限表，一些比较低级的任务(如移动)是被允许的。场景识别任务全部在服务器端完成，这也就意味着，机器人只需要拍摄一张场景图片并上传到服务器，服务器端借助场景图片即可推测出当前机器人所处场景。本文中每个用户的场景任务发布权限由系统管理员控制，不同的用户在不同的场景下可能有不同的权限，用户每次最多只能命令一个机器人执行任务。一个任务从发布到结束，用户与机器人的关系是一对一的，并且机器人通过服务器查询任务时，会对用

户 ID 和机器人 ID 进行比对。多个用户同时发布多个任务时，每个机器人都执行发布命令的用户的相应任务，避免了多用户多机器人情况下系统可能存在的冲突问题。

4. 场景理解

作为本系统重要的基础，场景理解模块的作用是识别机器人上传的图片中的场景，理解机器人所处的场景。机器人所处环境的依据，借助场景理解算法可以根据机器人拍摄的照片分析出机器人所处的场景，当用户上传任务时，服务器会根据用户的 commander_id 查询当前场景下用户所拥有的权限，如没有权限则会将请求驳回。本文中场景理解模块主要基于特征增强的深度学习主干网络，融合目标检测和语义分割，利用识别的多级语义信息，建立视觉模型，充分挖掘图像所表达的场景信息，识别出当前场景。

4.1. 数据集构建

本次模型训练采用的数据集为公共数据集，主要有目标检测数据集和语义分割数据集，其中目标数据集包含 80 类物体，每一类物体含有精确标注的包围框。目标检测和语义分割是两种不同的视觉任务，为了解决数据集之间图像数目、物体类别、数据集特点分布均不一致的问题，本文对数据集进行了筛选与合并，去掉会对模型训练带来干扰的数据，选择特征性更强的数据。经过处理之后的目标数据集相比原来类别数量较少，但各个类别之间 bounding box 和 image 数量相对比较平衡，有利于模型的训练。在测试阶段，根据训练集的类别对测试集进行筛选和过滤，以保持和训练集同样的类别。

4.2. 图像场景理解方法

本文设计了多任务学习的统一端到端识别框架，融合了图像分类、目标检测和语义分割等任务。模型框架采用多分支并行的结构，以特征增强方法的主干网络为特征提取主干网络，对于不同的识别任务用不同的分支进行并行训练。

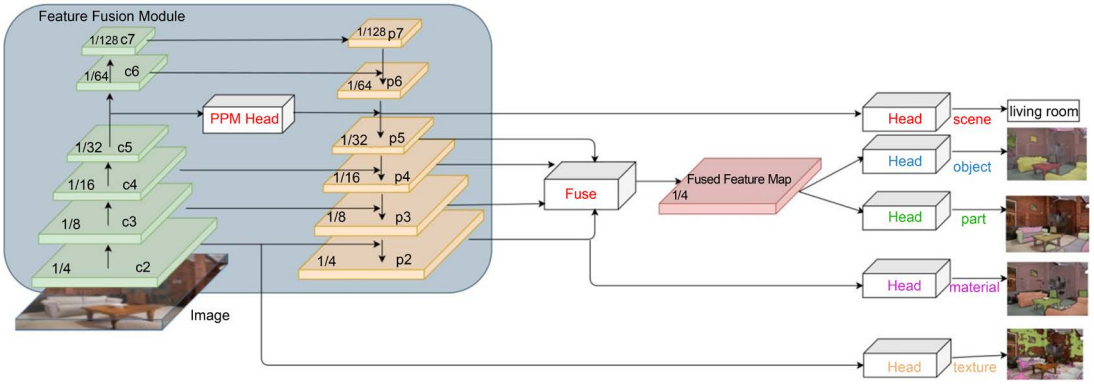

Figure 2. Framework of multi-level semantic recognition
图 2. 基于深度学习的场景理解框架

整体的模型架构如图 2 所示。模型以深度残差网络为主干网络，特征提取分为 $\{C_2, C_3, C_4, C_5, C_6, C_7\}$ 六个阶段，分别对应了卷积步长为 $\{4,8,16,32,64,128\}$。由于语义分割任务对空间信息的要求更高，因此本文在中低层特征图上进行预测，而在所有特征图上进行目标检测任务学习。经过特征融合监督模块之后得到新的特征图 Feature Map $\{P_2, P_3, P_4, P_5, P_6, P_7\}$。对于检测任务，Feature Map 再经过感受野模块连接 FEDet 的 Detection Head，最后输出 Bounding Boxes。

训练与推断设置：训练中采用 SGD 优化算法，对于语义分割任务，其学习率设置为 0.02，权重衰减(Weight-Decay)设置为 0.0001，学习率衰减采用“poly”策略，即当前的学习率等于初始学习率乘以

$\left(1-\frac{iter}{\max\_iter}\right)^{power}$，Power 设置为 0.9；对于检测任务，其学习率设置为 0.001，在第 30 个 Epoch 衰减为 0.0001，整个训练过程经历 40 个 Epoch，每个 Epoch 训练 5000 轮，采用 8 卡同步训练，每轮小批量大小设置为 16，即每张卡为 2 张图片。训练阶段不固定任何批归一化层。考虑到 Texture 分割任务的特殊性，Texture 分支不参与联合训练阶段，待训练结束后再单独进行微调。

损失函数和标签设计：本文设计的框架包含多个任务，因此最终的损失函数由多个子任务损失函数组成。对于 Scene 任务采用交叉熵损失函数，对于 Object、Parts、Material 等任务采用 Sigmoid Cross-Entropy 损失函数，对于检测任务，边框位置回归采用 Smooth-l1 损失函数，类别分类采用交叉熵损失函数。设 Scene 任务为 $Loss_S$，Object 任务为 $Loss_O$，Parts 任务为 $Loss_P$，Material 任务为 $Loss_M$，目标检测任务为 $Loss_{reg}$ 和 $Loss_{cls}$，则整体损失函数为：

$$Loss = \alpha_S \dot{L}oss_S + \alpha_O \dot{L}oss_O + \alpha_P \dot{L}oss_P + \alpha_M \dot{L}oss_M + \alpha_{reg} \dot{L}oss_{reg} + \alpha_{cls} \dot{L}oss_{cls}$$

其中，α 是权重项，目的是平衡不同损失函数的重要性，使得模型可以快速收敛。在本章实验中 $\alpha_S, \alpha_O, \alpha_P, \alpha_M, \alpha_{reg}, \alpha_{cls}$ 取值分别为 1.0，0.5，0.5，1.0，0.1，0.1。

本文采用的数据集全部为公共数据集。其中，目标检测数据集采用 MS COCO [8]数据集，语义分割采用 ADE20K [9]、Pascal-Context [9]等公共数据集。本文的实验采用框架 PyTorch 实现。实验结果中 Top-1 场景的识别正确率达到 68.49%，可以满足需要的识别精度要求。

4.3. 视觉语义知识挖掘

本文设计了一个端到端的多级语义[10]概念识别框架[11] [12]，可以同时得到图像所表达的多级语义信息，除了利用语义分割识别结果来建立语义概念关系，同时也利用了设计的高效目标检测模型的识别结果，进一步区分同类物体之间不同个体，提高语义知识建模的效果。给定一张图像，根据模型的识别结果，可以构建场景的语义知识模型，比如图像场景与物体的关系(Scene-Object Relation)，物体与材料的关系(Object-Material Relation)，物体部分与材料的关系(Parts-Material Relation)，材料与纹理的关系(Material-Texture Relation)等。挖掘这些语义关系信息有助于更好的理解图像场景。对于一张图片，模型可以从中理解出比如“客厅中存在哪些物体？这些物体由哪些部件组成？这些物体由什么材料组成？材料又由什么纹理组成？”这类的知识和概念，从而理解图像语义信息。对场景语义知识的挖掘可以使视觉系统更好的理解机器所处的环境，作为智能机器人的感知模块重要组成部分，视觉语义知识是其他模块进行下一步行动和决策的重要依据。

本文设计的多任务学习框架利用多种不同的标注形式的数据集联合训练，识别图像的多级语义信息。本文采用 ade20k-validation 作为测试数据，该数据集来自于真实世界，符合实际的场景特征，是评估本文框架合适的验证集。本文基于框架的识别结果，构建了如下语义知识关系：

1) Scene-Object

2) Object-Parts

3) Object-Material

4) Material-Texture

上述语义关系除了 Object-Parts 可以直接从标注中提取之外，其他关系都需要对图像进行识别，通过识别结果进行推理统计才能得到。在本文的多任务图像场景理解框架中，目标检测可以识别物体的位置和类别，语义分割可以获得物体的类别和边界，两者都是对 Object 的识别，但不同的是，语义分割得到的是不区分实例(Instance)的 Object 区域分割，而目标检测得到的是 Instance-Level 的 Object 区域检测框(Bounding Box)，两者虽然有所重合，但在识别结果上互为补充。通过对两者识别结果的整理和修正可以得到更精确的 Object 区域分割掩码(Mask)。

5. 场景理解的应用

本系统采用的场景理解模块主要应用在对机器人所处环境的感知上，当机器人执行任务时，伴随着机器人的移动，经常会出现机器人所处环境改变的情况。而就像公司里普通职员没有权限查看公司的财务报表一样，机器人的命令者也有相应的权限制约。例如，办公室场景的命令权限表如表 1 所示，名为 Com1 的命令者在办公室场景下发布了控制机器人移动的命令，当完成此任务时，此时机器人移动到了会议室，同时机器人会自动拍摄并上传一张实时照片供服务器进行分析，此时服务器就掌握了机器人所处的新环境。当名为 Com1 的命令者再发布一个拍照任务时，服务器通过查询 commander_id 在会议室场景下的权限时发现命令者并无拍照权限，就会将该命令者的请求驳回。命令者将无法发布无权限的任务，只能再发布移动命令控制机器人返航。

Table 1. Command table for office scene
表 1. 办公室场景的命令表

用户 ID	拍照	控制移动	控制机械臂
Com1	T	T	T
Com2	T	F	T
Com3	F	T	T

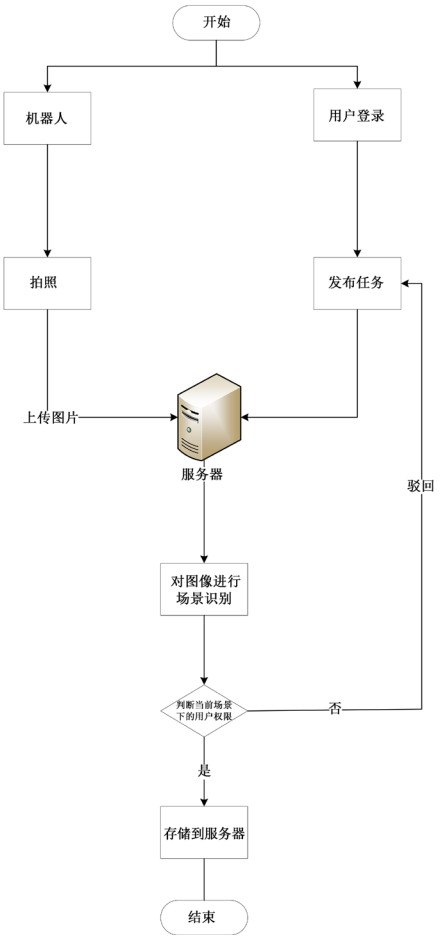

Figure 3. System flow chart
图 3. 系统流程图

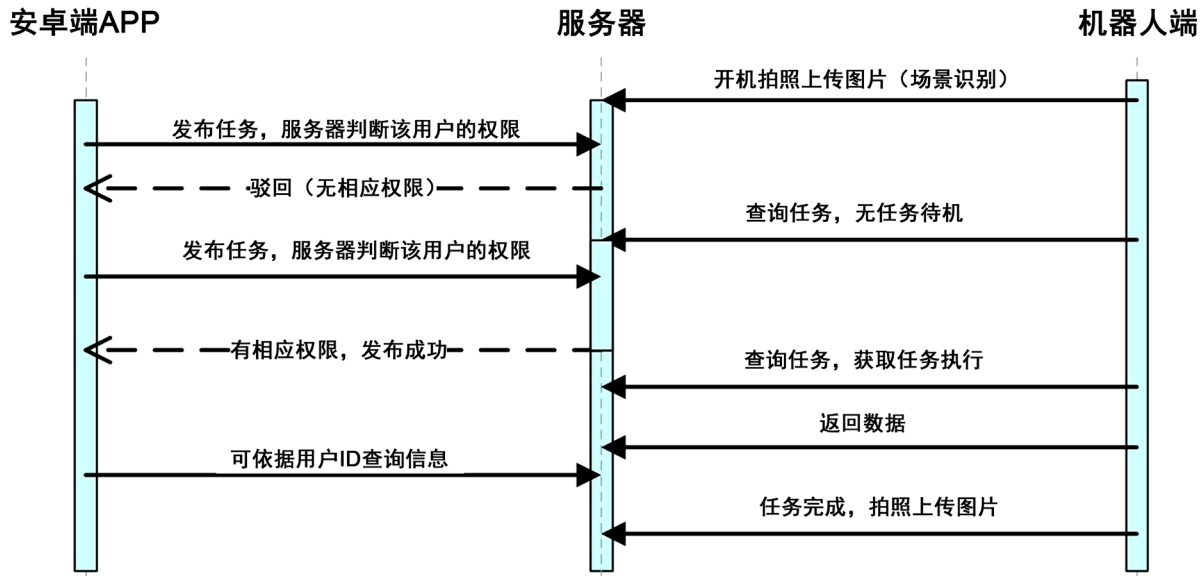

Figure 4. System operation timing diagram
图 4. 系统操作时序图

系统运行流程图如图 3 所示，系统设置了机器人上电启动就会拍照上传至服务器的机制，以便服务器在机器人开机后就能通过对图片进行场景理解获取机器人的初始场景。用户通过 Android 应用可发布任务并选择不同机器人执行任务，服务器会根据当前场景判断此用户有无发布此任务的权限，如果用户拥有发布此任务的权限，任务就会存储到服务器的数据库中，否则，会将任务驳回。当任务发布到服务器上后，机器人通过轮询任务列表，接收任务并执行。任务完成后数据上传到服务器的数据库。当任务完成后，机器人将再次拍照上传服务器，供服务器进行场景理解，获取到当前场景分类，以便服务器了解机器人完成任务时所处的场景。操作时序图如图 4 所示。

6. 实验与实现效果

在用户登录后，填写任务信息、确定发布任务的同时，服务器会获取用户的 ID，通过遍历设定好的表单查询用户有无发布该任务的权限。如果查询结果为有，任务会继续执行发布，服务器分析上传的任务，向指定的机器人提供任务执行数据；如查询不到结果，则会返回 bad_request 页面并显示“User has no permission!”，即为驳回用户请求，如图 5 所示。该任务会被服务器驳回，不记录在系统数据库中。Permission

```
Task create:
Task create:
    user = request.user
    user.id
permission_str = []
    for item in list1:
permussin_str.add(item.can_do)
    if "TaskManagerViewSet.create" not in permission_str:
        return Response('user has no permission!',
status=status.HTTP_400_BAD_REQUEST)
```

判断表：

ID	User_ID	Can_do
1	1	"TaskManagerViewSet.create"
2	1	"****"
3	2	"****"

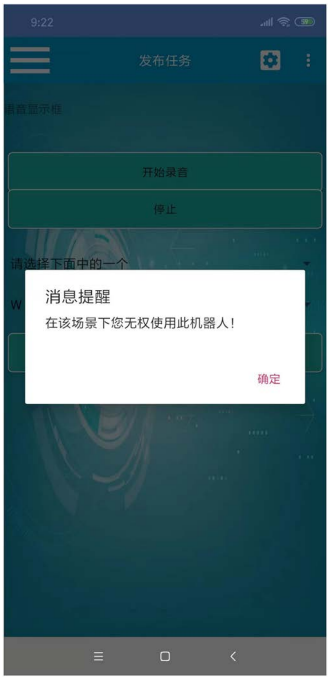

Figure 5. Android application display interface
图 5. Android 应用显示界面

当机器人执行完成一个完整任务时，也会拍摄一张照片上传供服务器分析，进行场景理解，以备下一个用户上线时进行访问控制。

7. 结束语

本文基于多用户多机器人系统的运行状态和用户行为提出一种对用户进行访问控制的方法，并结合场景理解算法，划分不同场景实现了对用户的访问控制工作。测试结果分为同一场景下有权限和无权限用户分别发布任务两个部分。从测试结果来看，同一场景下无权限用户对于特定任务无法发布成功，而有权限用户可以成功发布任务。表明本文提出的方法切实有效，满足了对于不同级别的用户进行访问控制的需求。为多用户多机器人参与的系统安全保密的需求提供了一种新的思路。

基金项目

本课题得到国家重点研发计划课题(2017YFB1302400)、江苏省科技计划产业前瞻与共性关键技术竞争项目(BE2018084)、2018 北京高等学校高水平人才交叉培养“实培计划”面向云机器人的服务平台设计与实现(29041618014)、2018 年工业互联网创新发展工程项目资助。

参考文献

[1] 李牧, 李成群. 基于机器人发展趋势的研究[J]. 山东工业技术, 2018(24): 230.

[2] 黄天龙. 基于云平台的移动机器人自主控制系统研究[D]: []. 哈尔滨工程大学, 2018.

[3] 周风余, 尹磊, 宋锐, 田天, 陈宏兴. 一种机器人云平台服务构建与调度新方法[J]. 机器人, 2017, 39(1): 89-98.

[4] 周斐, 李锦芝. 基于人工智能的广州市政府智能服务机器人云平台设计及应用[J]. 数字技术与应用, 2018, 36(12): 120-122.

[5] 陈宏兴, 周风余, 田天, 姜志飞, 陈竹敏. 服务机器人云计算平台 SOA 接口层模型设计[J]. 山东大学学报(工学版), 2015, 45(4): 31-39.

[6] 龙慧, 朱定局, 田娟. 深度学习在智能机器人中的应用研究综述[J]. 计算机科学, 2018, 45(S2): 43-47+52.

[7] 邓畅. 基于 ROS 的云机器人系统设计与实现[J]. 上海工程大学报, 2018, 32(4): 319-323.

[8] Lin, T.-Y., et al. (2014) Microsoft Coco: Common Objects in Context. *European Conference on Computer Vision*, Springer, Cham, 740-755. https://doi.org/10.1007/978-3-319-10602-1_48

[9] Zhou, B.L., et al. (2017) Scene Parsing through ade20k Dataset. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Honolulu, 21-26 July 2017, 5122-5230. https://doi.org/10.1109/CVPR.2017.544

[10] Mottaghi, Roozbeh, et al. (2014) The Role of Context for Object Detection and Semantic Segmentation in the Wild. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 891-898. https://doi.org/10.1109/CVPR.2014.119

[11] 张继鑫, 武延军. 基于 ROS 的服务机器人云端协同计算框架[J]. 计算机系统用, 2016, 25(9): 85-91.

[12] 陈贤, 武延军. 基于 ROS 的云机器人服务框架[J]. 计算机系统应用, 2016, 25(10): 73-80.