

基于计算机视觉的手势识别康复系统研究与应用

陈壮炼, 林晓乐, 陈银菊, 黄秋莹, 李超

(广东轻工职业技术学院 信息技术学院, 广州 510300)

摘要: 针对目前医学上手部康复治疗的康复器械功能单一, 训练过程动作反复且枯燥无趣, 恢复过程比较缓慢等问题, 提出了一种基于计算机视觉的手势识别康复系统; 首先介绍了图像采集、分割、平滑处理、分类、识别等计算机视觉关键技术, 其次, 重点阐述了手势康复系统的实现细节; 通过摄像头采集不同年龄, 性别人群手势样本数据, 建立康复手势数据库, 利用计算机视觉, 卷积神经网络, PyQt 图形界面等技术来构建康复系统, 给出具有趣味性并且高效便利的康复训练方案, 治疗方法弥补传统康复治疗的缺陷; 实验结果表明, 系统运行可靠、准确率达 96%, 不但可以提高患者对康复训练的兴趣, 积极性, 而且价格相对于其他康复器械更加低廉, 应用前景更广阔。

关键词: 计算机视觉; 手势识别; 康复系统; 卷积神经网络; PyQt 图形界面

Research and Application of Gesture Recognition and Rehabilitation System Based on Computer Vision

CHEN Zhuanglian, LIN Xiaole, CHEN Yinju, HUANG Qiuying, LI Chao

(Information Technology College, Guangdong Light Industry Vocational and Technical College, Guangzhou 510300, China)

Abstract: Aiming at the problems of single function of rehabilitation equipment for hand rehabilitation in medicine, repeated and boring movements during training, and slow recovery process, a gesture recognition rehabilitation system based on computer vision is proposed. First, it introduces the key technologies of computer vision such as image acquisition, segmentation, smoothing, classification, and recognition. Second, it focuses on the implementation details of the gesture rehabilitation system. Collect gesture sample data of people of different ages and genders through the camera, establish a rehabilitation gesture database, use computer vision, convolutional neural network, PyQt graphical interface and other technologies to build a rehabilitation system, and provide interesting, efficient and convenient rehabilitation training programs and treatments methods to make up for the shortcomings of traditional rehabilitation therapy. The experimental results show that the system is reliable and get 96% accuracy, not only can increase the patient's interest and enthusiasm for rehabilitation training, but also the price is lower than other rehabilitation equipment, and the application prospect is broader.

Keywords: computer vision; gesture recognition; rehabilitation system; convolutional neural network; PyQt graphical interface

0 引言

在现实生活中, 手部障碍给人们带来了诸多不便, 而且严重影响人们的日常生活。随着科学技术的发展, 手部障碍患者的治疗方式愈趋于智能化, 如基于数据手套的手部远程虚拟康复系统^[1-2]和基于触觉的虚拟康复系统^[3-4]。前者是利用数据手套来控制虚拟场景中中小球的拿捏来完成手部的训练。该虚拟系统虽然一定程度能达到手部康复的目的, 但患者在佩戴时会有负重感, 容易造成二次伤害, 而且数据手套价格较昂贵, 不适合普通家庭的使用。后者是利用触觉系统和虚拟环境结合设计出用户图形, 患者通过举起杯子和挤压球等动作走出“3D 迷宫”, 完成训练。该康复系统虽提高了患者训练的积极性, 但虚拟设备价格较高, 而且只能训练某些特定的动作, 不具备灵活性。

针对上述手部康复系统的不足, 利用计算机视觉技术设计出的手势识别康复系统兼具灵活、高识别准确率、低成本的特点, 并且在训练过程中能与患者进行动态交互, 通过数据库记录下训练过程, 有助于医务人员针对该训练过程给患者制定合理的康复训练疗程。

1 康复系统结构及原理

1.1 康复系统结构

本文为手部患者设计了较灵活、低成本的康复系统。其系统结构可分为图像采集、图像预处理、图像分类和图像识别 4 个部分。系统结构如图 1 所示。

1.2 系统结构原理

该系统采用的 Python 编辑器是 PyCharm, 该编辑器的好处是能够提升编程的速度, 提高开发效率。主要利用 An-

收稿日期: 2020-11-30; 修回日期: 2021-01-13。

基金项目: 2020 年广东省科技创新战略专项资金立项项目(pdjh2020b0886); 2019 年广东轻工职业技术学院教学改革项目(JG201916)。

作者简介: 陈壮炼(1998-), 男, 广东普宁人, 本科在读, 主要从事卷积神经网络、计算机视觉方向的研究。

通讯作者: 李超(1987-), 男, 湖南长沙人, 硕士, 讲师, 主要从事工业互联网、物联网方向的研究。

引用格式: 陈壮炼, 林晓乐, 陈银菊, 等. 基于计算机视觉的手势识别康复系统研究与应用[J]. 计算机测量与控制, 2021, 29(7): 203-207.

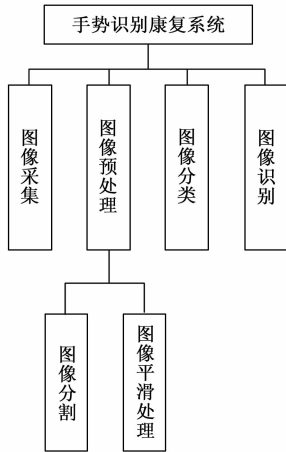


图 1 康复系统结构图

conda 对包进行管理, 利用开源视觉库 OpenCV 和卷积神经网络进行图像处理 and 分类, sklearn 进行模型训练, PyQt 进行 GUI 界面设计和数据可视化。

1.2.1 图像采集

采用 PC 端摄像头对康复手势进行拍摄采集, OpenCV 会将采集的图片会分割成多个像素点, 形成像素图, 再将图像保存到文件夹, 并给文件命名作为图像的标签。

1.2.2 图像预处理

1) 图像分割:

计算机视觉中的摄像头采集到的是不同等级的灰度图像^[5], 而在很多应用中需要的是纯黑或纯白的图像, 如卷积神经网络在进行图像的特征提取时输入只由简单线条、文字组成的图像, 那么在提取的过程就会简单很多, 模型训练的准确率也会提高。本文采用局部二值化图像中的自适应阈值二值化算法进行图像分割^[6]。

将摄像头采集的彩色图片进行灰度化后, 把灰度图像进行二值化处理。即找到一个阈值, 灰度大于这个阈值的区域设置像素为 255, 灰度小于这个阈值的区域设置像素为 0。自适应二值化是针对光影变化较大的图片, 根据图片的灰度直方图, 得到一个适合该图像的二值化阈值。灰度直方图如图 2 所示, 柱状形代表灰度直方图, 设法找到一个合适的混合高斯分布, 两个高斯分布相交的灰度值即要选择的阈值。

本文采用最大类间方差法 (OTSU) 来计算图像二值化的自适应阈值, 阈值将原图像分为前景和背景, 这里令阈值为 T , 图像像素对应的灰度级为 $f(x,y)$, 经阈值分割后的图像 $g(x,y)$ 定义为:

$$g(x,y) = \begin{cases} 1 & f(x,y) \leq T \\ 0 & f(x,y) > T \end{cases} \quad (1)$$

公式 (1) 中, 1 代表前景, 0 代表背景, 即图像二值化。

当选取最佳阈值时, 目标与背景的差别最大, 最大类间方差法 (OTSU) 公式推导如下。

记 T 为分割阈值, W_0 为前景点数占图像比例, U_0 为前景平均灰度, W_1 为后景点数占图像比例, U_1 为后景平均灰

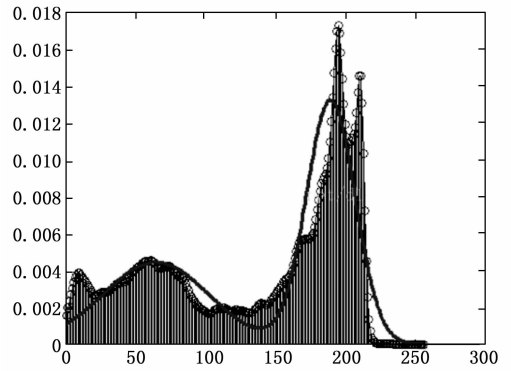


图 2 灰度直方图

度, 则图像的总平均灰度 U 为:

$$U = W_0 * U_0 + W_1 * U_1 \quad (2)$$

前景和后景图像的方差 S 为:

$$S = W_0 * (U_0 - U)^2 + W_1 * (U_1 - U)^2 = W_0 * W_1 * (U_0 - U_1)^2 \quad (3)$$

当方差 S 最大时, 前景和后景的差别最大, 此时的灰度是最佳阈值。

2) 图像平滑处理:

图像平滑处理^[7]是计算机视觉中一种相对简单并且使用频率较高的图像处理方式。常用来减少图像中的噪点和降低图像的分辨率。一张没有经过图像处理的图像中包含有很多的高频信息, 如噪音和边界。如果将未经处理的图像直接进行卷积运算会带来很多不必要的运算量, 使用滤波器可以有效地去除噪音, 达到图像模糊的目的。图像平滑处理一般有图像均值平滑滤波, 高斯滤波和中值滤波, 本文采用的是均值平滑滤波。

均值滤波是一种典型的线性滤波算法, 是指在图像上对目标像素值给出一个滤波模板, 该模板包含目标像素周围的 $M * N$ 个临近像素, 也就是去掉目标像素本身, 再用该模板中目标像素周围的像素平均值来代替原来像素值。其中 $M * N$ 的矩阵称为核, 采用核对原始图像中的像素点进行处理, 从而实现图像去噪, 得到结果图像。均值滤波处理过程如图 3 所示。

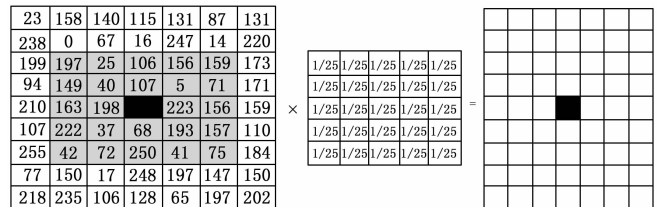


图 3 均值滤波处理过程

均值滤波处理公式为:

$$g(x,y) = \frac{1}{M} \sum_{fes} f(x,y) \quad (4)$$

根据公式 (4) 对图 3 中的区域像素进行均值滤波处理过程为:

$$\begin{aligned} & ((197+25+106+156+159) + (149+40+107+5+ \\ & 71) + (163+198+226+223+156) + (222+37+68+ \\ & 193+157) + (42+72+250+41+75)) / 25 = 125.52 \end{aligned}$$

其中, 5×5 的矩阵为模糊内核。本文使用计算机视觉中的 blur() 方法实现图像的平滑处理。

1.2.3 图像分类

计算机能看到的是一个像素值, 并不能像人类的视觉一样, 可以轻松分辨出图片里面出现的事物, 以及图片的类别, 故计算机需要通过图像分类来区分图片。基于计算机视觉的图像分类方法有朴素贝叶斯分类、K 近邻、支持向量机、决策树、Boosting、随机森林和神经网络。本文采用神经网络中的卷积神经网络图像分类算法^[8-9], 首先通过 PCA 技术提取样本图像特征与待分类的图像特征码, 根据特征码生成输入项, 样本所属标签作为输出项, 将输入与输出项送入神经网络进行训练, 让神经网络学习每个类别图像的特征, 最后将待分类图像的特征码输入神经网络进行测试, 根据神经网络输出项判断其所属类别。

1.2.4 图像识别

计算机和人的眼睛一样, 都是通过颜色的轮廓来区分图像, 而图像的颜色由三原色组成的, 可以通过调色得到各种颜色。在计算机图像识别中, 制定了 RGB (Red, Green, Blue) 通道作为图像识别的基础。OpenCV 中图像通道有单通道、二通道、三通道和四通道, 其中常见通道有单通道和三通道。

一般黑白图像是用灰度来表示色彩, 称为单通道, 例如某个像素值为 0 则表示黑色, 255 则表示白色, 灰度的范围为 $0 \sim 255$ 。彩色图像则采用 RGB 通道来表示, 称为三通道, 例如 $(255, 255, 255)$ 表示白色, $(0, 0, 0)$ 表示黑色, 灰度也是彩色的一种。二通道的图像采用 RGB555 和 RGB565 表示, 一个通道为实数, 另一个为虚数。四通道理则是采用 RGBA 来表示, 与三通道相比多了一个 A 通道, A 通道表示透明度。本文图像识别采用的是三通道。

2 康复系统设计与实现

2.1 康复系统的工作流程

康复系统设计主要分为几个步骤: 1) 样本采集: 首先在界面的绿色识别方框内做出不同的手势, 摄像头识别到手势并将手势进行拍摄保存到文件夹中, 再将文件夹命名为对应手势的名称作为手势标签。2) 图片预处理: 将采集到的手势进行图像分割, 形成灰度图像, 以免复杂背景影响训练的准确度, 再将图像进行均值滤波处理, 减少图像的噪声。3) 神经网络训练: 将处理过后的图像通过卷积、池化和全连接层进行特征提取和分类。4) 模型测试: 在相应位置做出对应手势, 观察屏幕是否显示该手势的标签。5) 康复训练: 患者通过打开系统进行手部康复训练, 并将对应训练过程保存到数据库中。6) 康复历史查询: 通过点击查询按钮, 调用数据库查询语句, 将训练过程显示到屏幕上。

2.1.1 样本采集

为了实现手部康复的训练, 定义并采集了 6 种手势,

依次是悬垂、二指捏、击打、柱状抓握、勾拉、侧捏, 建立康复手势数据库。通过调用 cv2.VideoCapture(0) 方法打开 PC 端摄像头, 输入采集手势的种类和训练周期, 将采集的手势保存到 geskind 目录下, 然后将同种类的手势放到一个文件夹并为其取标签名, 方便后续对这些手势图像进行图像处理。采集的分别是 50 岁以上、10 到 50 岁、10 岁以下 3 个年龄段的人群, 一共拍摄 70 组数据, 共 490 张手势图片, 采集的图片大小都是 306×306 , 采集条件是室内正常光照, 采集手势对象的男女比例均衡。图像采集流程如图 4 所示。

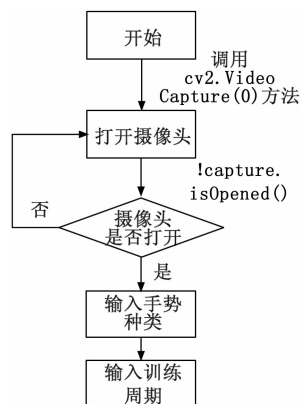


图 4 图像采集

以下是进行图像捕获关键代码:

```

global gestname, sample_nums, path, saveimg
if(self.is_camera_opened == False):
    QMessageBox.warning(self, '警告', '未打开摄像头')
else:
    if self.hand_name.text() == "":
        gestname = self.hand_example.currentText()
    else:
        gestname = self.hand_name.text()
    sample_nums = self.hand_num.value()
    print(gestname + str(sample_nums))
    try:
        os.makedirs('./geskind/' + gestname)
    except OSError as e:
        print(gestname + '文件夹已创建')
    path = './geskind/' + gestname + "/"
    if gestname == "":
        print("请先输入一个存放文件夹的名字")
    else:
        saveimg = True
        # 循环捕获图片
        ret, self.frame = self.camera.read()
        self.frame = cv2.flip(self.frame, 3)
        roi = self.binaryMask(self.frame, x0, y0, width, height)
        framecount = framecount + 1
  
```

2.1.2 图片预处理

为了降低手势在训练过程过拟合和减少图片处理时的

计算量, 需要进行图片预处理。利用 Opencv 中的 cv2.imread() 方法将图片加载进来, 再通过 Numpy 的 array() 方法将图片进行归一化, 使其能提高模型的收敛速度和精度, 然后使用 LabelBinarizer() 进行二值化处理, 将彩色图像处理成黑白图像, 即二值化灰度图像, 再将二值化灰度图像用 blur() 方法进行平滑处理, 降低图像噪声, 最后使用 pickle.dump(lb, f) 将标签数据保存为 dat 文件, 方便模型训练和手势识别, 具体流程如图 5 所示。

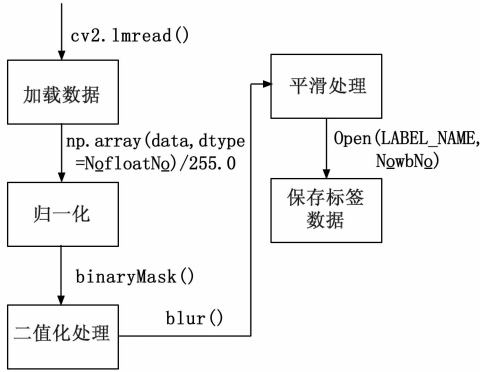


图 5 图像预处理流程

2.1.3 神经网络训练

该手部康复系统的神经网络结构包含 5 个卷积层, 3 个池化层和两个全连接层。其中 5 个卷积层的卷积核尺寸大小都一样, 卷积核个数分别是 32, 64, 64, 128, 128, 而且都是采用全零填充和 Relu 激活函数。3 个池化层采用的是 2 * 2 大小, 步长为 2 的最大池化操作。在全连接层之前该系统采用了 Dropout^[10] 随机丢弃 50% 神经元参数, 让这些参数不参与当前训练而达到了又一次的防止过拟合。最后在输出层采用 Softmax 函数对结果进行分类输出。关键代码如下:

```

# 神经网络构建
model = Sequential()
# 第一层卷积和第一层池化
model.add(Conv2D(32,(3,3),padding="same",input_shape=(100,100,3),activation="relu"))
model.add(MaxPooling2D(pool_size=(2,2),strides=(2,2)))
# 第二层卷积
model.add(Conv2D(64,(3,3),padding="same",activation="relu"))
# 第三层卷积和第三层池化
model.add(Conv2D(64,(3,3),padding="same",activation="relu"))
model.add(MaxPooling2D(pool_size=(2,2),strides=(2,2)))
# 第四层卷积
model.add(Conv2D(128,(3,3),padding="same",activation="relu"))
# 第五层卷积和第三层池化
model.add(Conv2D(128,(3,3),padding="same",activation="relu"))
  
```

```

model.add(MaxPooling2D(pool_size=(2,2),strides=(2,2)))
# 全连接层
model.add(Flatten())
model.add(Dense(64,activation="relu"))
# 防止过拟合
model.add(Dropout(0.5))
# 输出层(用的是多分类的对数损失函数 softmax,如果只有两个手势,用 sigmoid)
model.add(Dense(class_num,activation="softmax"))
  
```

2.1.4 模型测试

为了保证模型的准确度, 需要对模型的准确度进行测试。首先需要使用 pickle.load() 加载标签数据, 调用 load_model() 加载神经网络, 使用 inverse_transform() 将标准化的数据转为原始数据, 再调用 cv2.VideoCapture() 方法打开 Pc 端摄像头, 在界面的绿色区域内做出采集过的手势, 调用 cv2.putText() 将识别到的手势类型名称显示在屏幕上, 根据在绿色识别区域作出的手势与屏幕上显示的手势类型名称是否一致判断模型的准确度。重复进行模型测试, 当模型测试达到高准确度和高稳定性时, 就可以把该模型运用到手部康复训练中。关键代码如下:

```

# 加载标签
try:
with open(LABEL_NAME,"rb") as f:
lb = pickle.load(f)
# 加载神经网络
model=load_model(MODEL_NAME)
except:
print("没有'+LABEL_NAME+'和'+MODEL_NAME'")
gesture=lb.inverse_transform(prediction)[0]
  
```

2.2 康复训练

手部康复患者首先打开康复训练程序, 若是首次使用的患者可以根据需要录入各种手势。非首次使用可以跳过录入手势的步骤, 设定好康复训练的组数作为一个疗程进行康复训练, 在康复训练中患者需要根据屏幕上提示的文字信息在界面的绿色区域中做出对应的手势, 完成疗程后结束康复训练。训练完毕会生成数据录入数据库中, 患者可点击历史查询按钮查看历史康复训练数据的反馈, 根据系统记录判断患者的训练完成程度对训练计划做出相应的调整, 患者可设定较低组数作为手部的适应性训练, 当患者适应训练强度后可逐步增加训练的难度。循序渐进的手部康复训练, 有利于提高患者对康复训练的信心和加快患者手部的恢复速度。

2.3 实验结果分析

通过摄像头拍摄的各种手势图像一般都是 RGB 图像, 该系统使用了 OpenCV 技术把输入的 RGB 图像数据经过预处理转化为灰度图像, 同时也采用归一化提升手势识别康复系统模型的收敛速度和精度。该手部康复系统的神经网络结构包含 5 个卷积层, 3 个池化层和两个全连接层。经过 5 个仅取图像数据局部特征的卷积层, 有效地减少了图片的

长度和宽度, 减少了特征权值的数量。3 个一样的池化层不断地降低了维数, 通过降采样的方式, 在不影响我们的特征图像质量的情况之下压缩图片, 从而达到了减少网络中的参数和计算次数, 缩短了训练时间等目的, 也很好地控制了过度拟合情况的发生。最后在全连接层之前该系统也采用了 Dropout 随机丢弃 50% 神经元参数, 让这些参数不参与当前训练而达到了又一次的防止过拟合。其中第一个全连接层也使用了拉直层。由于手势识别康复过程中需要训练的手势往往是两个以上的, 所以输出层用的是输出值概率总和始终为 1 的多分类的对数损失函数 Softmax 函数。该系统的训练过程中是用总样本的 20% 为测试数据, 剩下的 80% 为训练数据。最终的训练的准确率可以达到 96% 以上。

该系统不仅使用了计算机视觉技术, 卷积神经网络也应用在其中, 使训练模型的准确度不断提高。整个康复系统也采用 PyQt5 设计 GUI 界面, 实现了可视化。以历史查询界面为例, 用户通过时间选择范围, 查看在各阶段的手势恢复数据反馈, 如训练时间, 训练手势种类, 训练周期等数据可以帮助医生以及患者查看患者在各个阶段的手部恢复情况, 便于医生对恢复程度进行针对性的康复训练调整。由于计算机视觉对收集手势的环境背景要求比较高, 该康复系统在使用时, 最好保证摄像头开启时的环境背景足够单调, 这样更有利于降低周围环境对识别手势的干扰, 准确识别到手势, 提高训练使用准确率。

3 结束语

本文设计并实现了基于计算机视觉的手部康复系统, 给手部障碍患者提供一个良好的训练环境并进行有效的康复训练, 提高患者训练的积极性。系统具有以下点: 采用 PyQt 设计了 GUI 界面, 呈现效果美观; 利用计算机视觉技术与患者进行交互, 识别速度和识别准确率较高; 患者的训练过程会保存到数据库中, 可以随时查看训练过程并做

出相应的训练调整。

手势康复训练是近年来人工智能研究邻域的热点, 对此我们对该系统如下展望: 提高在复杂背景下识别的准确率; 通过跟相关医院合作, 根据更专业的医护知识完善训练手势的种类、训练过程设计等; 进一步研究智能决策系统, 通过主观评估结果结合智能决策法, 产生康复训练的初步方案, 并将该方法运用到游戏训练的过程中, 通过监控当前游戏训练中的参数的变化趋势, 以此作为后续训练方案制定的依据。

参考文献:

- [1] 胡艳玲. 基于手势识别的手部虚拟康复系统研究 [D]. 秦皇岛: 燕山大学, 2017.
- [2] 郭路. 手指康复训练机器人系统设计 [D]. 南京: 东南大学, 2019.
- [3] 刘方全. 基于虚拟现实与触觉感知的多功能上肢康复系统研究 [D]. 济南: 山东大学, 2019.
- [4] 张晶晶. 面向上肢功能康复的视触觉增强现实系统研究 [D]. 南京: 南京信息工程大学, 2019.
- [5] 郭子雷. 基于计算机视觉的手势识别系统的设计与实现 [D]. 武汉: 华中科技大学, 2016.
- [6] 黄鹏, 郑淇, 梁超. 图像分割方法综述 [J]. 武汉大学学报 (理学版), 2020 (6): 519-531.
- [7] 魏挺. 基于平滑滤波处理的图像插值通用方法 [J]. 舰船电子工程, 2019, 39 (12): 133-136.
- [8] 罗斌, 韩纪东. 基于主成分分析和 CNN 的图像分类算法研究 [J]. 科学技术创新, 2020 (30): 99-101.
- [9] 王艳, 孙薇, 周小平. 基于深度学习的中草药植物图像识别方法研究 [J]. 中医药信息, 2020, 37 (6): 21-25.
- [10] 高雷鸣, 肖满生, 向华政. 基于深度学习的肺部肿瘤图像识别方法 [J]. 计算机测量与控制, 2020, 28 (10): 160-164.
- [11] 王亚刚, 邵惠鹤. 基于相位裕度和幅值裕度的 PI/PID 参数自适应新方法 [J]. 电气自动化, 2000 (5): 13-15.
- [12] 王繁珍, 陈增强, 姚向峰, 等. 基于频域的 PID 广义预测控制器的鲁棒性分析 [J]. 中国工程科学, 2006, 8 (10): 66-70.
- [13] HO W K, HANG C C, CAO L S. Tuning of PID controllers based on gain and phase margin specifications [J]. Sydney, Australia: 1993: 267-270.
- [14] HO W K, HANG C C, CAO L S. Tuning of PID controllers based on gain and phase margin specification [J]. Automatica, 1995, 31 (3): 497-502.
- [15] 王曦, 曾庆福. 频域不确定性系统加权混合灵敏度函数频域整形 [J]. 航空学报, 1999, 20 (4): 358-361.
- [16] 陈光炯. 基于频域法的直升机旋转桨叶模型分布动载荷识别 [D]. 南京: 南京航空航天大学, 2014.
- [17] 董磊, 宋汉文, 郑铁生. 频响函数曲线拟合与模态分析精细化 [J]. 振动与冲击, 2016, 35 (2): 69-75.
- [18] 吴吉平. 多输入多输出频域正交多项式模态参数识别方法 [D]. 长沙: 中南大学, 2005.
- [19] 唐钰婷. 航空发动机 PID 控制频域设计方法研究 [D]. 南京: 南京航空航天大学, 2012.
- [20] 张勇. 面向整车 NVH 分析的频响函数子结构混合建模方法研究 [D]. 北京: 清华大学, 2016.
- [21] 李伟华, 庄奕琪, 包军林, 等. 基于虚拟仪器的频响分析系统 [J]. 仪器仪表学报, 2004 (1): 354-356.
- [22] 张峰, 徐坚, 王雷. 多通道频响测试系统设计与开发 [J]. 测控技术, 2018 (02): 122-124.
- [23] 赵力海. 三轴转台液压控制系统低速和频响性能的理论分析与工程实现 [D]. 上海: 上海交通大学, 2014.
- [24] 赵军. 转台频域性能测试软件开发及控制技术研究 [D]. 南京: 南京航空航天大学, 2007.
- [25] 叶明. 基于 LabVIEW 的角振动台频响测试系统实现 [J]. 航空精密制造技术, 2011 (4): 60-62.
- [26] 孙鑫晖, 张令弥, 王彤. 基于奇异值分解的频响函数降噪方法 [J]. 振动、测试与诊断, 2009, 29 (3): 325-328.