

DOI: 10.19783/j.cnki.pspc.190515

变电站二次测试中智能语音控制关键技术研究

徐昱¹, 裘愉涛¹, 侯伟宏¹, 陈嘉宁¹, 蒋逸雯², 王义波³

(1. 国网浙江省电力有限公司杭州供电公司, 浙江 杭州 310000; 2. 华中科技大学, 湖北 武汉 430074;
3. 武汉凯默电气有限公司, 湖北 武汉 430223)

摘要: 为了增强变电站二次测试智能化水平, 提高测试效率, 提出了一种基于智能语音识别的电力二次测试系统。采用一种针对电力专用词库的二次语音识别引擎, 设计了基于离线识别模式的嵌入式测试系统。在实验室与现场, 利用该系统分别实测了智能终端、保护装置、合并单元等设备。实测结果表明, 电力专用词汇语音识别引擎提高了整体识别率和稳定性, 且该测试系统的语音操控能够准确无误地完成整个测试过程, 增强了电力二次测试仪的操控性、灵活性和易用性。

关键词: 智能语音; 离线识别; 麦克风矩阵; 电力专用识别引擎; 嵌入式系统

Key technologies of intelligent speech control in substation secondary test

XU Yu¹, QIU Yutao¹, HOU Weihong¹, CHEN Jianing¹, JIANG Yiwen², WANG Yibo³

(1. Hangzhou Power Supply Company, State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310000, China; 2. Huazhong University of Science and Technology, Wuhan 430074, China; 3. Wuhan Kemov Electric Co., Ltd., Wuhan 430223, China)

Abstract: In order to enhance the intelligent level of substation secondary test and improve the test efficiency, the solution of smart substation testing based on intelligent speech recognition is proposed. A secondary speech recognition engine for power special lexicon is researched and the embedded system design of off-line recognition mode is given. In the laboratory and in the field, the intelligent terminal, protection device, merging unit and other equipment are tested by using the system. The experimental results show that the power special lexicon speech recognition engine significantly improves the overall recognition rate and stability and the speech control of the test system completes the whole test process accurately and correctly, which greatly enhances the maneuverability, flexibility and ease of use.

This work is supported by Science and Technology Project of State Grid Corporation of China (No. 5211HZ1700A).

Key words: intelligent speech; offline recognition; microphone array; power special speech recognition engine; embedded system

0 引言

智能电网(Smart Grid, SG)是电力系统发展的必然趋势, 其核心要义是“智能”, 而人工智能技术取得了突破性进展, 为SG的发展提供了重大机遇和强大支撑^[1]。智能变电站是智能电网的重要组成部分, 随着电网规模的不断扩大, 其建设规模也越来越大。二次设备数量迅速增长, 对设备的管理、运行、维护、检修人员提出了更高的要求, 二次系统运维工作量也日益繁重^[2]。手持式光数字测试仪、继保测试仪、网络测试仪等仪器, 在变电站二次系

统中普遍使用。然而, 这些测试仪器界面复杂、操作流程繁琐, 特别是现场待测的实验设备(保护装置、智能终端、合并单元等)与测试仪有一定距离, 或者多台就地化保护子机需要进行联合测试时, 测试者需要同时操作、检查测试仪与待测设备, 操作繁琐、费时耗力。作为人工智能的一个重要分支——智能语音技术, 极大简化了以传统鼠标、键盘、触摸屏等复杂的输入方式, 使得人机交互模式变得更加简单、快捷、高效。

语音识别方案大多选用云语音识别技术, 例如Google云语音、科大讯飞和百度云语音等等, 而云语音识别效果与网络状况及服务器计算能力相关, 导致语音识别延时大、识别结果情况相关性不强^[3]。

云平台的语音识别方案高度依赖于互联网。然而,依据《电力二次系统安全防护规定》的安全分区划分方案^[4-5],实施起来有一定局限性。待测试的设备如保护、合并单元、智能终端等都部署在电力安全 I 区,基于云平台的二次测试装置,存在着将 I 区与外网连通的风险,导致网络安全分区无法满足要求。即使是目前正在建设的电力专网,也是属于 III 区网络,基于局域网服务器的在线识别同样存在网络安全风险,容易形成外部网络攻击点。相关研究表明,在无标度网络中 5%的关键点受到攻击失败后,就可以使整个网络系统崩溃进而瘫痪^[6]。因此,现阶段电力二次测试优先考虑离线语音识别方案,当然其技术难度也会更高。

文献[7]采用数字信号处理器芯片,将离线语音识别成功地移植到读书郎学习机产品上,读书郎学习机是针对儿童早期启蒙教育类机器,仅对简单的字母、数字具有一定识别能力,限制了电力专用词汇的识别。另一类离线语音识别方案采用 LD3320 芯片^[8-10]。该语音识别芯片,每次最多可以识别 50 项候选句,且无法支持参数变化的数字或字母等,一般应用在稳定性、可靠性较低的智能家居、儿童玩具等领域。一种基于隐马尔可夫模型的电力调度语音识别研究^[11],采用 GPU 服务器训练电网调度语音数据,少量语音命令词下取得了较好识别精度。文献[12]研究了语音识别+鼠标+键盘等多种方式在电力能量管理系统中的应用,认为在控制中心使用语音技术能够减少成本、提高效率,更快地与关键的应用程序进行交互,但并未对语音识别本身有更深入的研究。总体而言,电力系统中的语音识别技术研究尚在起步阶段。

本文提出一种基于离线模式的智能语音识别技术,结合二次测试仪器的使用,提出了系统性架构。试验表明,具有智能语音识别的电力二次测试仪器显著提高了测试仪器的操控性、灵活性、易用性,极大地提高了现场测试人员的效率。特别是电力专用词汇语音识别引擎的引入,提升了电力专用词汇识别率与稳定性。本研究也为整个电力二次测试仪器提供了一套智能语音识别技术的系统方案。

1 智能语音技术原理

语音识别是以语音为研究对象,通过语音信号处理和模式识别,使机器自动识别人类口述的语音。经过近 60 年的研究,混合高斯模型-隐马尔可夫(Gaussian mixture model-hidden Markov model, GMM-HMM)声学模型在语音识别领域广泛使用,构成了现代语音识别系统的基础^[13]。随着深度神经

网络(Deep Neural Network, DNN)被应用于机器学习,可结合 HMM 与 DNN 融合搭建 DNN-HMM 混合系统,提高系统训练、解码速度,提高语音的识别率。

典型语音识别系统框图如图 1 所示,包括前端音频模块、语音识别引擎、声学模型、语音数据库、语言识别引擎、语言模型、词组数据、解码搜索、结果输出^[13]。

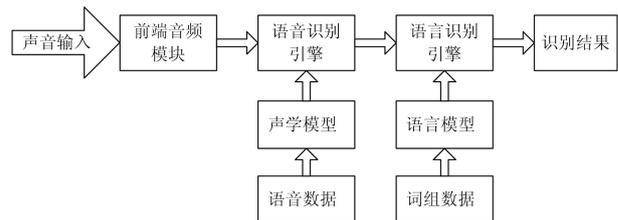


图 1 典型语音识别框图

Fig. 1 Typical block diagram of speech recognition

前端音频模块,将麦克风变化的声音变为信号微弱的电信号,通过消除噪声和信道失真增强语音信号,将信号从时域转化到频域,并为后续的声学模型提取合适的有代表性的特征向量。整个语音识别系统,遵循概率论中的著名贝叶斯模型^[14],其表达式如式(1)、式(2)所示。

$$L = \arg \max P(L|A) \quad (1)$$

$$L = \arg \max P(A|L)P(L) \quad (2)$$

式中: L 表示文字序列概率; A 表示语音输入。式(1)表示语音识别需要在给定的语音输入情况下,找到最有可能的文字序列。根据贝叶斯规则变换,可得式(2)。其中, $P(A|L)$ 表示给定一个字节序列出现该条音频的概率,即语音识别中的声学模型; $P(L)$ 表示出现该文字序列的概率,即语音识别中的语言模型。基于信号统计模型的 GMM-HMM 算法属于双重随机过程,该算法试图找到显性语音输入与隐性文字输出之间的概率。为了获得高识别率, GMM-HMM 算法需要大量的训练样本和存储器,训练过程也耗费过多的时间,多用于连续大词汇语音识别系统。

得益于机器学习的深度神经网络 DNN 的发展,结合 DNN 的 DNN-HMM 声学模型^[15-20]有很强的学习能力和序列化建模能力,在大规模的连续语音识别任务中,其性能远优于传统的 GMM-HMM 系统。

典型的 DNN-HMM 混合型系统如图 2 所示。HMM 用来描述语音信号的动态变化,其模型由不同状态和状态转移概率组成。DNN 的多层神经网络在估计观测概率时考虑了语音的上下文特征,使概率分布更加合理。根据给定的特征向量序列和若干

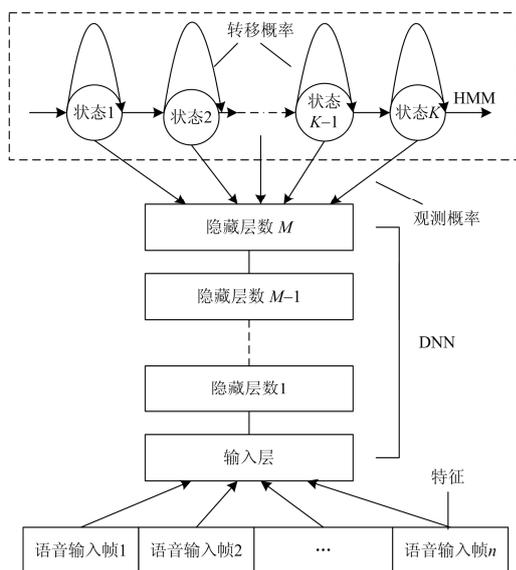


图 2 DNN-HMM 混合型结构

Fig. 2 Hybrid structure of DNN-HMM

假设词序列，可计算出声学模型分数和语言模型分数，总体输出分数最高的词序即为识别结果^[21]。

2 智能语音识别的关键技术

2.1 多麦克风矩阵硬件模块

变电站内的各种声音噪声会大大降低原始声音的信噪比，提高原始声音质量极为重要。一种前端硬件的多麦克风矩阵，可以明显地降低背景声音噪声，加强声源信号强度，改善声音信噪比。

阵列信号处理技术是指将一组传感器按一定的方式布置在空间的不同位置上，组成传感器阵列来接收空间信号，经过适当的处理提取信号源及信号的属性等信息，这些信息包括信号辐射源的数目、方向、幅值等^[22]。多麦克风矩阵形成的声束，能够加强确定方向上的声源强度，抑制其他方向声束旁瓣能力。图 3 显示声束主瓣指向声源 1，旁瓣指向声源 2 和其他噪声源。由于主瓣指向声源 1，该声源将会被放大、加权增强，而其他声源或噪声将被有效抑制。

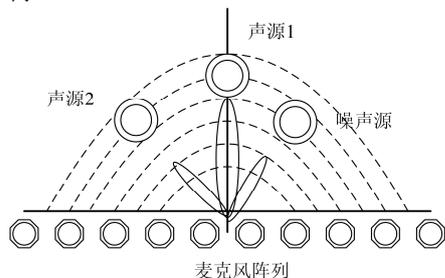


图 3 麦克风阵列声束

Fig. 3 Microphone array sound beam

2.2 通用语音识别模块

通用语音识别模块集成了语音识别处理器和一些外部电路，包括 AD、DA 转换器、麦克风接口、音频输入接口。通用语音识别模块主要由语音唤醒引擎和通用识别引擎组成，系统框图如图 4 所示。

语音唤醒引擎用于唤醒睡眠中的语音识别模块，也可以作为系统唤醒的信号源，让处在睡眠中的系统立即进入工作状态。该功能能够显著地降低整机功耗。

通用识别引擎是一个普适性语音识别引擎，完成了从声音到汉字、字母、数字的转换。麦克风阵列高信噪比的语音信号输出到通用语音识别模块，模块内部的 DSP 处理器提取语音的频谱特性、音素等特征，这些特征与常用汉字、字母、数字比相匹配，输出可能的汉字、字母、数字组合序列，从而实现了识别过程中复杂的 DNN-HMM 算法的硬件化和模块化。

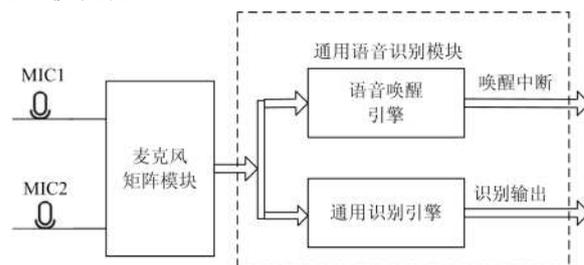


图 4 通用语音识别模块

Fig. 4 Universal speech recognition module

2.3 电力专用识别引擎

电力专用识别引擎以提高电力词库识别率为目的，其过程是求解通用识别引擎输出字符串与电力专用词库的相似度问题。通过一定规则的特征再提取、数字化、判决等方法，进一步提高语音识别准确率与稳定性，系统框图如图 5。

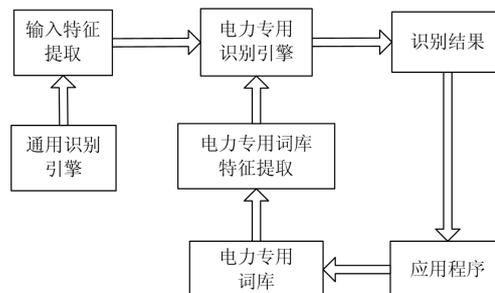


图 5 电力专用的识别引擎框图

Fig. 5 Block diagram of power-specific identification engine

特征提取成为电力词库二次再识别的关键。基于汉字发音特征能够快速、准确地在限定的电力专用词库中搜索到正确的命令词。汉字发音的显著特

征为声母、韵母、辅音和声调，将这些信息进行数字化归类评分，进行相似度计算。将一个汉字的发音按如图 6 所示，拆分为 4 个单元，即声母、韵母、辅音、音调。

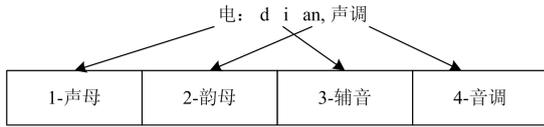


图 6 汉字拼音组合

Fig. 6 Chinese character pinyin combination

将如上发音单元进行数字化，即

$$Sum(V) = W_1 \times v_1 + W_2 \times v_2 + W_3 \times v_3 + W_4 \times v_4 \quad (3)$$

式中： W 为权重； v 为单元的编码。声母权重 W_1 取 0.4，韵母权重 W_2 取 0.4，辅音权重 W_3 取 0.1，音调权重 W_4 取 0.1。

声母与韵母进行数字化编码，分别见表 1 和表 2。

表 1 声母数字化表

Table 1 Initials digitization table

声母	编号	声母	编号	声母	编号	声母	编号
b	1	p	2	m	3	f	4
d	5	t	6	n	7	l	7
g	8	k	9	h	10	j	11
q	12	x	13	zh	14	ch	15
sh	16	r	17	z	14	c	15
s	16	y	18	w	19		

结合发音规则与计算量考虑，部分声母具有同样的编码值。有些汉字具有辅音或多韵母，该类情况都归类为辅音，采用韵母表数字化。四种声调数字化为 0、1、2、3。便于 CPU 计算，声母、韵母分别采用 5 位二进制，辅音编码采用 4 位二进制，声调采用 2 位二进制编码，总体为 16 位二进制编码。

表 2 韵母数字化表

Table 2 Finals digitization table

韵母	编号	韵母	编号	韵母	编号	韵母	编号
a	1	o	2	e	3	i	4
u	5	ü	6	ai	7	ei	8
ui	8	ao	9	ou	10	iu	11
ie	12	ve	13	er	14	an	15
en	16	in	17	un	18	vn	19
ang	15	eng	16	ing	17	ong	20

定义两个汉字发音相似度 Sim 为

$$Sim(a,b) = 1 - \frac{Sum(V_a) - Sum(V_b)}{Sum(V_b)} \quad (4)$$

假设用户说出“电力”声音，经过通用识别引

擎后，有可能输出为同音的“店里”，其拼音分割为“d”、“i”、“an”、“l”、“i”，再次特征化后，采用式(3)进行数字化，则

$$Sum(V_a) = 0.4 \times 5 + 0.4 \times 15 + 0.1 \times 4 + 0.1 \times 3 + 0.4 \times 7 + 0.4 \times 4 + 0.1 \times 0 + 0.1 \times 0 = 13.1$$

电力专用词库中对应的“电力”进行数字化特征，采用式(3)进行数字化，则

$$Sum(V_b) = 0.4 \times 5 + 0.4 \times 15 + 0.1 \times 4 + 0.1 \times 3 + 0.4 \times 7 + 0.4 \times 4 + 0.1 \times 0 + 0.1 \times 3 = 13.4$$

基于式(4)计算两个词之间的相似度，结果为 97.76%。

通过逐一搜索电力专用词库的数字化值，找到相似度最大的那个词，即认为是最佳语音匹配词。

3 变电站测试仪的智能语音系统硬件设计

变电站测试仪需要考虑实现测试仪的网口接入、FT3 接口、对时接口等常规接口及实现语音识别的硬件电路^[23]。硬件系统的整个方案采用 CPU、FPGA(现场可编程逻辑门阵列)芯片，通用语音识别模块，多阵列麦克风输入模块等组成，框图见图 7。

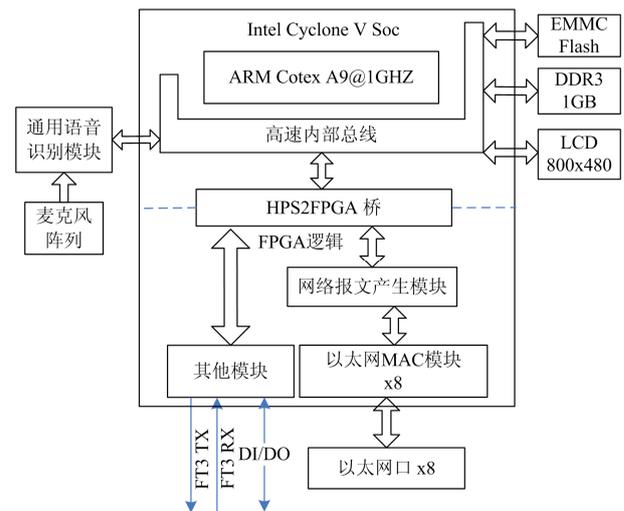


图 7 语音识别硬件系统

Fig. 7 Speech recognition hardware system

CPU 子系统采用 Intel 公司的 Cyclone-V SOC 嵌入式平台，内部集成 ARM Cortex-A9 双核 CPU、主频为 1 GHz，集成 85 k 逻辑单元，作为测试仪的网口、FT3、DI/DO 接口。Cyclone-V SOC 可以外接 4 GB 的大容量 EMMC Flash 作为系统程序空间和电力专用词汇的存储空间，高速 DDR3 存储器运行整个系统。通用语音识别模块采用 USB 接口连接到系统硬件。800X480 分辨率的液晶显示器，便于人机界面交互。多麦克风矩阵和通用语音识别模块，需要考虑到电源噪声对麦克风矩阵电路、语音识别

模块的影响, 建议采用低噪声的线性电源供电。隔离语音识别系统电源与其他硬件电源, 减少数字系统对模拟语音的影响。

4 变电站测试仪的智能语音系统软件设计

4.1 系统软件整体架构

软件系统包括测试应用程序、电力专用识别引擎程序、底层驱动程序, 整个软件框架如图 8。通用语音识别模块采用标准 USB 接口与声卡驱动, 免驱动设计。通用语音识别引擎通过回调函数获得当前语音识别的文字字符串。

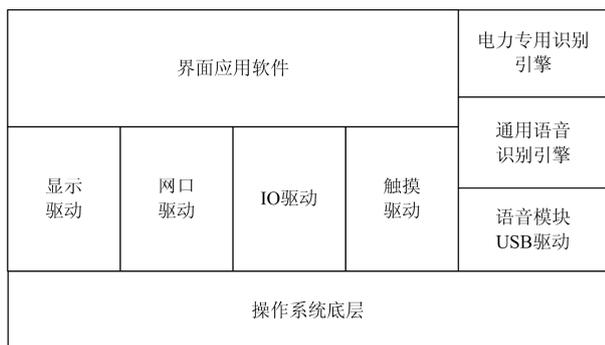


图 8 系统软件框架

Fig. 8 System software framework

系统开机启动后, 首先进入外设初始化, 液晶屏、触摸控制等。设置好这些参数后, 软件系统将进入主界面, 在主界面等待用户触摸输入或等待语音关键词唤醒。

4.2 通用语音识别处理流程

当通用语音识别模块侦听到特定唤醒词后, 系统将从睡眠中唤醒, 开启通用语音识别模块。通用语音识别处理程序将不断轮训该引擎输出, 确认是否有识别字符串输出。如果查询到有字符串输出, 则对该识别结果字符串进行数字化特征提取。否则超过 10 s 后, 系统再次进入睡眠, 等待再次唤醒。整个操作流程见图 9。

4.3 电力专用语音识别引擎处理流程

电力专用语音识别引擎处理通用语音识别引擎输出字符串的特征码与电力专用词汇特征码的相似度。该识别引擎计算并搜索电力词库中的词汇特征码, 找到相似度最高的词汇, 将该词汇输出到应用程序。首先, 将常用电力专用词汇按照前述 2.3 节的方法整理其特征码, 各自对应关系整理成表 3(仅列出部分词汇作为例子)。考虑到计算量与实时性要求及应用程序界面词汇的排版, 再将词汇按照页面分类, 保证一个操作界面的词汇在 20~30 个。既满

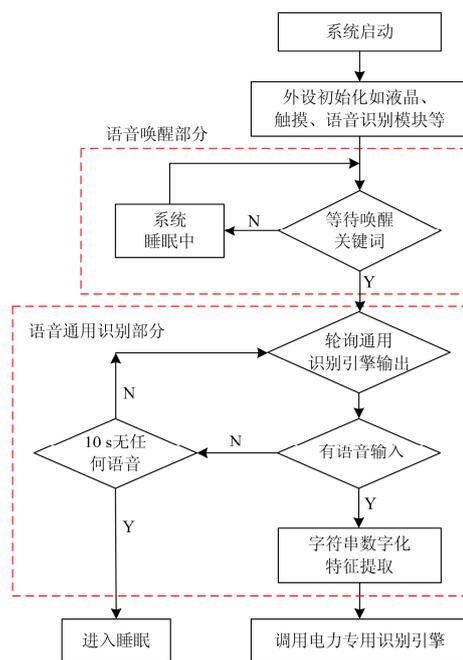


图 9 唤醒与通用语音识别引擎流程

Fig. 9 Wakeup flow and the recognition process of universal speech

表 3 部分电力专用词汇与特征码

Table 3 Partial power-specific vocabulary and signature

电力专用词汇	页面编码	电力专用词汇特征码
保护设备	1	010902-100503-
		160303-010803
合并单元	1	100300-011703-
		051550-181552
智能终端	1	140400-071601-
		142000-051550
就地化 线路保护	2	111100-050400-100153
		131543-070503-
就地化 主变保护	2	010902-100503
		111100-050400-100153
就地化 母线保护	2	140502-0115043-
		010902-100503
选择 SCD 文件	3	111100-050400-100153
		030502-131543-
添加 SV	3	010902-100503
		131550-140302-
添加 GOOSE	3	160000-160803-050403-
		191601-111543
电压电流 额定发送	3	061540-110410-
		160000-190800
置运行	3	061540-110410-
		080500-14002
置运行	3	011543-180100-
		011543-011101-
置运行	3	000300-0516030-
		040100-162003
置运行	3	140003-181803-131701

足界面词汇的使用,也不至于系统计算量过于庞大,以保证系统操作的流畅性。电力专用词汇中,字母按照相近的汉字发音处理。

在通用语音识别引擎有结果输出时,应用软件将调用电力专用语音识别引擎程序。先将当前页面词汇对应的特征码载入到内存中;再逐一计算电力专用词汇表中的特征码与实时输入特征码的相似度;最后搜索出相似度最高的词汇,输出到应用程序。电力专用语音识别流程见图 10。调整相似度阈值,以增强设备的安全性。只有超过 80%相似度的词汇才可以输出到应用程序;否则,提示用户再次语音输入,以避免误操作。

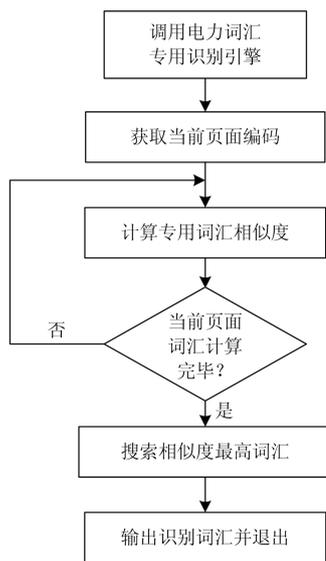


图 10 电力专用语音识别流程

Fig. 10 Recognition process of power-specific speech

5 测试与验证

测试与验证目的分为五点。第一,验证具有智能语音识别功能的测试仪是否具备提高现场测试合并单元、智能终端、保护装置等设备的操控性、便捷性,并正确输出测试向量;第二,判断智能语音整体识别率是否满足现场使用情况;第三,计算智能语音识别引擎在不特定人中识别的稳定性;第四,测试电力专用词库的命令词字数长度对识别率和识别时间的影响;第五,变电站现场实际使用语音输入,测试保护装置,以验证是否满足现场工况环境。

5.1 智能语音识别操控性验证

该试验模拟了采用语音输入方式给保护装置施加测试向量的情形。一台具有智能语音识别的光数字测试仪 A,连接光网口到另一台检测设备 B(该设备模拟一台待测保护装置)。

开启光数字测试 A 上的麦克风阵列与通用语音识别模块电源,此时整个语音识别系统上电与初始化完成,等待用户说出唤醒词“凯默你好”;用户说出唤醒词“凯默你好”开始进入实时语音识别;用户说出“电压电流额定值发送”,光数字测试仪 A 识别到该命令词后,找到默认配置的 SMV 开始发送该报文,另一台检测设备 B 准确地显示了该 SMV 数据流信息,如图 11 所示。



图 11 基于语音操控的光数字测试仪

Fig. 11 Optical digital tester based on speech control

该测试表明试验者在连接完线缆后,整个测试完全不需要手动操作任何界面的菜单、按钮即可方便、快捷、准确地执行全部测试流程,也让试验者更加专注于待测就地化保护装置、智能终端、合并单元等设备的自身动作。

5.2 电力专用语音识别引擎识别率测试

测试选取了 100 个电力专用词汇,随机挑选了 20 名测试者,每名测试者分别在只有通用语音识别引擎和增加电力专用识别引擎情况下,依次说出选取的 100 个电力词汇,记录光数字测试 A 的响应是否正确。能够正确响应语音动作认为识别正确,不能正确响应或无响应都认为识别失败。图 12 显示了通用语音识别引擎与电力专用识别引擎的识别率。结果显示,通用语音识别率平均为 69.76%,增加电力专用词汇识别引擎后整体平均识别率为 85.32%,平均识别率提高了 15.56%。增加电力专用词汇语音识别引擎,显著提高了整体识别率。

5.3 电力专用语音识别引擎增加识别稳定度

进一步分析识别率测试结果,得出不同人有明显的识别差异性,即识别率波动较大,识别率最好、最差、最大振幅如表 4。根据该表得出通用识别引擎识别率振幅为 23.27%,增加电力专用语音引擎后,不同人之间的识别率差异明显减小,振幅为 15.97%。电力专用语音识别引擎的使用可以显著地提高 7.30%的语音识别稳定性。

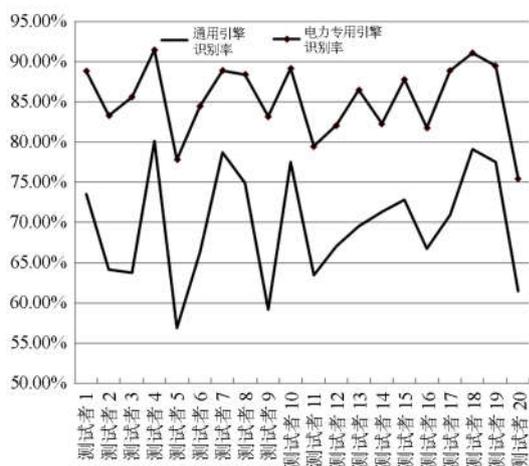


图 12 电力词汇识别率

Fig. 12 Recognition rate of power vocabulary

表 4 通用与专用识别稳定度

Table 4 Power special speech and universal speech stability

识别引擎	识别率最大值	识别率最小值	振幅
通用引擎	80.13%	56.86%	23.27%
电力专用引擎	91.47%	75.50%	15.97%

5.4 电力专用命令词字数长度与识别率及识别时间关系

按照电力专用命令词字数长度归类出 2 个字、4 个字、6 个字、8 个字以及超过 10 个字的命令词列表,对每一类字数列表读出 10 遍,测试出不同字数命令的识别率与识别时间。整个测试数据整理结果如图 13 所示。从图 13 中可以看出,理想命令词字数长度在 4 至 8 个字。字数太少的命令词,误识别率较高;字数太长导致计算量较大,有识别失败或超时的可能性。

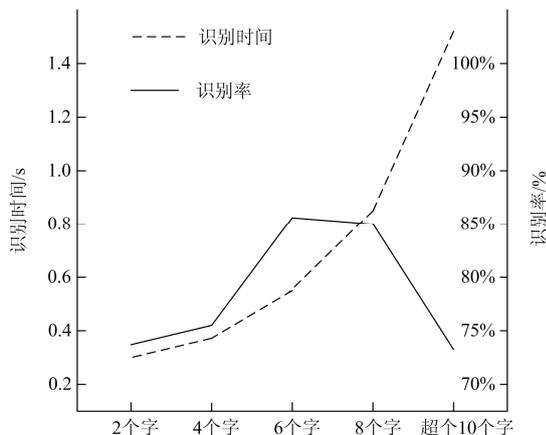


图 13 命令词长度与识别率、识别时间关系

Fig. 13 Relation between the length of command words and recognition rate and recognition time

5.5 变电站应用测试

在浙江某 220 kV 变电站现场,运维人员采用具有智能语音的光数字测试仪,选取了 24 条常用语音命令词,以语音输入的方式,分别对智能终端(型号: PCS-222)、合并单元(型号: PCS-221)、线路保护(型号: PCS-931)、母线保护(型号: PCS-915)这四台设备进行了现场使用测试,测试结果如表 5 所示。该结果显示:一次语音输入情况下,操作完成率为 87.50%;一次语音未正常操作的命令词,在第二次输入的情况下都能正常操作。其使用测试结果表明了离线语音识别能够适应变电站电磁噪声与声音噪声,准确地完成测试操作。

表 5 变电站现场语音测试结果

Table 5 Speech test results of substation

测试语音命令	被测试设备	一次语音操作完成	两次语音操作完成
开始智能终端测试		完成操作	
查看 goose 实时值	智能终端 型号:PCS-222	完成操作	
查看 goose 报文监测		完成操作	
查看 goose 间隔监测		完成操作	
查看 goose 报文异常		完成操作	
查看 SV 有效值	合并单元 型号:PCS-221	完成操作	
查看 SV 谐波		完成操作	
查看 SV 报文监测		完成操作	
查看 SV 报文统计		完成操作	
查看 SV MU 延时		完成操作	
开始距离保护测试	线路保护 型号:PCS-931	完成操作	
接地距离一段定值		未完成操作	完成操作
接地距离一段时间		未完成操作	完成操作
相间距离一段定值		完成操作	
相间距离一段时间		未完成操作	完成操作
开始网络压力测试	母线保护 型号:PCS-915	完成操作	
开始报文记录		完成操作	
查看光功率		完成操作	
置检修状态		完成操作	
置运行状态		完成操作	
查看语音关键字	所有设备	完成操作	
对话框确认		完成操作	
对话框取消		完成操作	
返回主界面		完成操作	

6 结论

当前变电站测试和运维中面临大量的试验与测试等工作,费时耗力,提出了一种基于智能语音控

制的变电站二次测试系统, 研究面向电力专用语库的语音识别引擎, 具体研究结论如下:

1) 采用本文的方案, 试验者在连接完线缆后, 整个测试完全不需要手动操作任何界面的菜单、按钮即可方便、快捷、准确地执行全部测试流程, 增强了电力二次测试仪操控性、灵活性、易用性。

2) 本文所提出的电力专用识别引擎显著提高了语音识别率与稳定度, 识别率从通用语音识别引擎的 69.76% 提高到 85.32%, 平均识别率提高了 15.56%。能够准确地执行测试者的测试目的, 也能够适应不特定人的语音差异。

3) 实际测试了语音命令词字数长度可能造成的语音识别率变差与识别时间加长的问題, 建议合适的命令词在 4 至 8 个字之间, 对语音命令词字数设计具有实际参考价值。

4) 本文研究了在网络安全要求较高的离线模式下, 智能语音技术应用到变电站二次测试中的可行性, 为变电站其他测试设备提供了一套智能语音识别的参考方案。

基于离线语音的识别方案相对于云语音识别方案, 在识别率、鲁棒性、口音等都有一定差距。因此, 还需对电力专用语音识别引擎进一步改善, 例如使用测试仪的特定人群、部分词汇的再学习与训练等。另外, 正在推广与建设的泛在电力物联网也使电力行业部署私有云成为可能^[24], 也为智能语音识别提供了兼顾安全与可靠、识别率更高的方案。

参考文献

- [1] 鞠平, 周孝信, 陈维江, 等. “智能电网+”研究综述[J]. 电力自动化设备, 2018, 38(5): 2-11.
JU Ping, ZHOU Xiaoxin, CHEN Weijiang, et al. “Smart Grid Plus” research overview[J]. Electric Power Automation Equipment, 2018, 38(5): 2-11.
- [2] 秦红霞, 武芳瑛, 彭世宽, 等. 智能电网二次设备运维新技术研讨[J]. 电力系统保护与控制, 2015, 43(22): 35-40.
QIN Hongxia, WU Fangying, PENG Shikuan, et al. New technology research on secondary equipment operation maintenance for smart grid[J]. Power System Protection and Control, 2015, 43(22): 35-40.
- [3] 刘敏. 嵌入式平台云语音识别技术的研究[D]. 哈尔滨: 哈尔滨工业大学, 2013.
LIU Min. Research on speech recognition technology in embedded platform[D]. Harbin: Harbin Institute of Technology, 2013.
- [4] 电监会 5 号令. 电力二次系统安全防护规定[S]. 2005. Order No.5 of the electric regulator commission. Regulations on the safety protection of secondary power systems[S]. 2005.
- [5] 刘国军, 郑晓崑, 杨会峰, 等. 电力通信网安全分区研究[J]. 电力信息与通信技术, 2016, 14(8): 28-31.
LIU Guojun, ZHENG Xiaokun, YANG Huifeng, et al. Research on the security partition of electric power communication network[J]. Electric Power Information and Communication Technology, 2016, 14(8): 28-31.
- [6] 王汪兵, 王先陪, 尤泽樟, 等. 电力通信网关键节点辨识方法研究[J]. 电力系统保护与控制, 2018, 46(1): 44-49.
WANG Wangbing, WANG Xianpei, YOU Zezhang, et al. Research on key node identification method in electric power communication network[J]. Power System Protection and Control, 2018, 46(1): 44-49.
- [7] 赵欢. 面向嵌入式计算平台的自动语音识别关键技术研究[D]. 长沙: 湖南大学, 2010.
ZHAO Huan. The key techniques of automatic speech recognition for the embedded computing platform[D]. Changsha: Hunan University, 2010.
- [8] ICRout Ltd. LD3320 数据手册[R]. 2011.
ICRoute Ltd. LD3320 datasheet[R]. 2011.
- [9] 罗贵舟, 熊晓东. 基于 LD3320 的智能 LED 照明系统研究[J]. 电子世界, 2017(2): 182-183.
LUO Guizhou, XIONG Xiaodong. Research on the system of intelligent LED lighting base on LD3320[J]. Electronics World, 2017(2): 182-183.
- [10] 孙红军, 周万根. 基于 LD3320 语音识别的智能家居控制系统设计[J]. 电子测量, 2017(18): 18-20.
SUN Hongjun, ZHOU Wangen. Design of smart home control system based on LD3320 speech recognition[J]. Electronic Test, 2017(18): 18-20.
- [11] 鄢发齐, 王春明, 窦建中, 等. 基于隐马尔可夫模型的电力调度语音识别研究[J]. 武汉大学学报: 工学版, 2018, 51(10): 920-923.
YAN Faqi, WANG Chunming, DOU Jianzhong, et al. A study of speech recognition in power dispatching based on hidden Markov model[J]. Engineering Journal of Wuhan University, 2018, 51(10): 920-923.
- [12] 朱全胜, 刘婉, 李卫东. EMS 人机界面及其交互技术研究[J]. 电力自动化设备, 2007, 27(8): 77-81.

- ZHU Quansheng LIU Rao, LI Weidong. HCI and interaction technologies of EMS[J]. Electric Power Automation Equipment, 2007, 27(8): 77-81.
- [13] RABINER L R. A tutorial on hidden Markov models and selected applications in speech recognition[J]. Proceedings of the IEEE, 1989, 77(2): 257-286.
- [14] 俞动, 邓力. 语音识别实践[M]. 俞凯, 钱彦旻, 等译. 北京: 电子工业出版社, 2016.
- [15] 刘加. 汉语大词汇量连续语音识别系统研究进展[J]. 电子学报, 2000, 28(1): 103-128.
- LIU Jia. Research on large vocabulary mandarin Chinese continuous speech recognition system[J]. ACTA ELECTRONICA SINICA, 2000, 28(1): 103-128.
- [16] DAHL G E, YU D, DENG L, et al. Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition[J]. IEEE Transactions on Audio, Speech and Language Processing, 2012, 20(1): 30-42.
- [17] MOHAMED A R, DAHL G, HINTON G. Deep belief networks for phone recognition[C] // Conference on Neural Information Processing Systems, 2010: 1-9.
- [18] SAINATH T N, KINGSBURY B, RAMABHADRAN B, et al. Making deep belief networks effective for large vocabulary continuous speech recognition[C] // IEEE Workshop on Automatic Speech Recognition & Understanding, December 11-15, 2011, Waikoloa, HI, USA.
- [19] MOHAMED A, DAHL G E, HINTON G. Acoustic modeling using deep belief networks[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2012, 20(1): 14-22.
- [20] HINTON G, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups[J]. IEEE Signal Processing Magazine, 2012, 29(6): 82-97.
- [21] CHO K, VAN M B, BAHDANAU D, et al. On the properties of neural machine translation: encoder-decoder approaches[J]. Computer Science, 2014: 1-9.
- [22] 宋雷鸣. 基于相控麦克风阵列的逆向噪声源识别原理与技术研究[D]. 北京: 北京交通大学, 2010.
- SONG Leiming. Reverse Noise-identify mechanism and technology based on phased microphone array[D]. Beijing: Beijing Jiaotong University, 2010.
- [23] 李志勇, 孙发恩, 翟晓宏. 智能变电站综合测试仪的研究与实现[J]. 电力系统保护与控制, 2018, 46(15): 149-150.
- LI Zhiyong, SUN Fa'en, QU Xiaohong. Research and implementation of integrated tester for smart substation[J]. Power System Protection and Control, 2018, 46(15): 149-150.
- [24] 陈皓勇, 陈永波, 王晓娟, 等. 基于 LPWAN 的泛在电力物联网[J]. 电力系统保护与控制, 2019, 47(8): 1-8.
- CHEN Haoyong, CHEN Yongbo, WANG Xiaojuan, et al. Ubiquitous power internet of things based on LPWAN[J]. Power System Protection and Control, 2019, 47(8): 1-8.

收稿日期: 2019-05-09; 修回日期: 2019-06-27

作者简介:

徐 昱(1975—), 男, 高级工程师, 从事电网生产营销管理工作;

王义波(1981—), 男, 通信作者, 本科, 高级工程师, 从事电力系统继电保护相关技术工作。E-mail: wangyibo@kemov.com

(编辑 张爱琴)