

基于机器视觉的智能导盲眼镜设计

何腾鹏, 张荣芬, 刘超, 房乐楠, 刘宇红

(贵州大学 大数据与信息工程学院, 贵州 贵阳 550025)

摘要: 提出一种基于机器视觉的智能导盲眼镜系统的设计方案。采用三星公司 Cortex-A8 架构的 S5PV210 作为中央处理器, 搭载 Linux 系统, 配备双目采集、GPS 定位、语音播报、GSM 短信、语音通话、无线传输六大核心功能模块搭建智能导盲眼镜系统的硬件平台, 结合深度学习算法在远程云服务器上完成了对目标场景的智能识别, 最后以语音的形式实时对盲人的行走作出准确引导。系统测试结果表明, 该智能导盲眼镜系统在测试环境下不仅能对盲人出行正确导航, 还具有一定的目标识别能力, 能帮助盲人进行简易物品归类。该系统还兼有 GPS 定位、语音通话、GSM 短信等多项辅助功能。

关键词: 机器视觉; Cortex-A8; 双目采集; 智能识别

中图分类号: TN929

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.2017.04.015

中文引用格式: 何腾鹏, 张荣芬, 刘超, 等. 基于机器视觉的智能导盲眼镜设计[J]. 电子技术应用, 2017, 43(4): 58-61.

英文引用格式: He Tengpeng, Zhang Rongfen, Liu Chao, et al. Design of smart seeing glasses based on machine vision[J]. Application of Electronic Technique, 2017, 43(4): 58-61.

Design of smart seeing glasses based on machine vision

He Tengpeng, Zhang Rongfen, Liu Chao, Fang Lenan, Liu Yuhong

(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

Abstract: A system of smart seeing glasses based on machine vision was proposed and designed in this work. Using Samsung Cortex-A8 architecture S5PV210 as the central processor, running on the Linux system, equipping six core modules of binocular acquisition, GPS, voice broadcast, GSM SMS, voice calls and wireless transmission were equipped to build smart seeing glasses systems hardware platform. Then after completing the target scene identification on a remote cloud server through deep learning algorithm, at last, the accurate voice guide for the blind walking in real time was implemented actually. The system test results show that the smart glasses system is not only able to make the right travel guide for the blind, it also has a certain ability to identify simple objects, which can help the blind make a simple items classification. In addition, this system also has GPS positioning, voice calls, GSM SMS and many other auxiliary functions.

Key words: machine vision; Cortex-A8; binocular collection; smart identification

0 引言

据世界卫生组织统计,截至目前全球约有盲人7 800万,其中90%生活在发展中国家,中国现有盲人数量占据世界盲人总数的18%,多达1 400万。盲人作为社会中的一种弱势群体,视力残疾、眼部疾病给其生活带来了诸多不便。另一方面,近年来随着我国老龄化人口的持续增多,视力障碍者数量也在急剧上升,如何保障盲人群体及视障人士安全有效地出行显得尤为重要。为此,本文设计了一种基于机器视觉^[1]的智能导盲眼镜,旨在帮助盲人朋友在行走过程中,安全、有效地避开道路上的目标障碍物,尽最大可能地保障其出行安全。相较于市场上导盲效率不理想的导盲手杖和价格昂贵的导盲犬,基于机器视觉的智能导盲眼镜系统更具竞争力。

1 智能导盲眼镜控制系统的总体设计

本文的智能导盲眼镜控制系统由前端嵌入式采集传输系统和远程云平台服务器两大部分组成。嵌入式采集传输系统以三星 Cortex-A8 架构的 S5PV210 处理器为载体,搭载 Linux 内核,配备双目采集、GPS 定位、语音播报、GSM 短信、语音通话、无线传输等核心功能模块搭建智能导盲眼镜系统的硬件平台,主要完成信息采集传输和智能指令导盲功能。云平台服务器作为智能导盲眼镜的远程数据处理中心,在服务器配置上,选用阿里云作为云端服务器,再融入深度学习、双目测距^[2]等相关算法,实现对智能导盲眼镜前端场景目标的图像识别、距离检测 and 方位判断。此外该服务器还构建了 GPS 卫星数据匹配平台,结合智能导盲眼镜实体,能对眼镜使用者进行实时有效地定位。本系统的总体设计框图如图 1 所示。

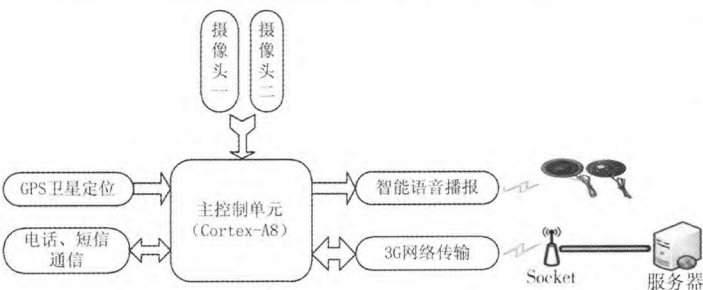


图 1 智能导盲眼镜控制系统结构框图

2 智能导盲眼镜控制系统硬件设计

2.1 双目采集模块设计

双目采集模块选用两个性能参数完全一致的 CMOS 高清摄像头, 用于智能导盲眼镜前方场景信息的采集, 帮助盲人获取前方目标和相应的场景信息。

2.2 GPS 定位模块设计

全球定位系统(GPS)能为全球用户提供低成本、高精度的三维位置,可在全球范围内进行全天候、全方位的实时定位。系统选用瑞士 Ublox 公司的 NEO-6M 模组作为 GPS 模块的核心单元,主要用于实时获取盲人所在地理位置的经纬度坐标。

2.3 无线通信传输模块设计

无线通信传输模块主要由 3G 芯片和对应的外围电路组成,采用 3G 无线技术,实现智能导盲眼镜与远程云平台服务器的双向通信。一方面将双目摄像头模块采集到的图片以及 GPS 模块获取到的地理位置坐标信息通过 3G 网络发到远端云服务器平台,另一方面将服务器的图片识别和地理位置配对结果反送回来传递给智能导盲眼镜,进行语音播报,把结果实时告诉给盲人。同时利用 3G 模块的 GSM 短信功能,智能导盲眼镜系统也会将从云服务器端获取的结果,以短消息的形式及时告知盲人家属。另外,在特殊情况下,盲人也可利用 3G 的电话功能直接与亲属间进行语音通话。图 2 为无线通信传输模块的应用电路图。

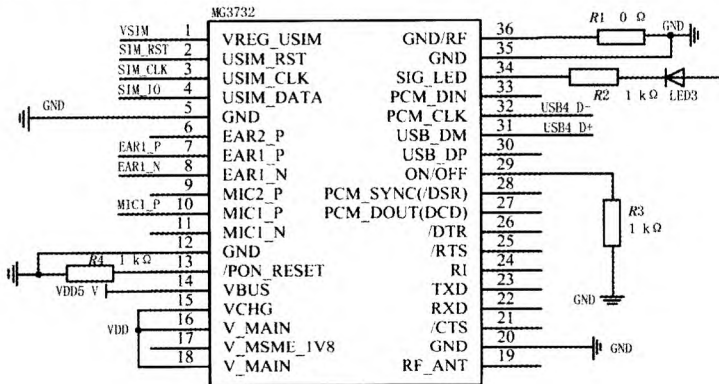


图 2 3G 模块应用电路图

2.4 语音播报模块设计

智能导盲眼镜控制系统的语音播报功能主要用于

《电子技术应用》2017年第43卷第4期

万方数据

将导盲眼镜前方目标的图片识别结果、距离、方位信息以及所处的地理位置通过语音模块播放出来,及时告知盲人所处的周围环境状况。系统选用 SYN6288 中文语音合成芯片作为语音播报模块的主体,实现文本到声音的转换,其外围电路如图 3 所示。

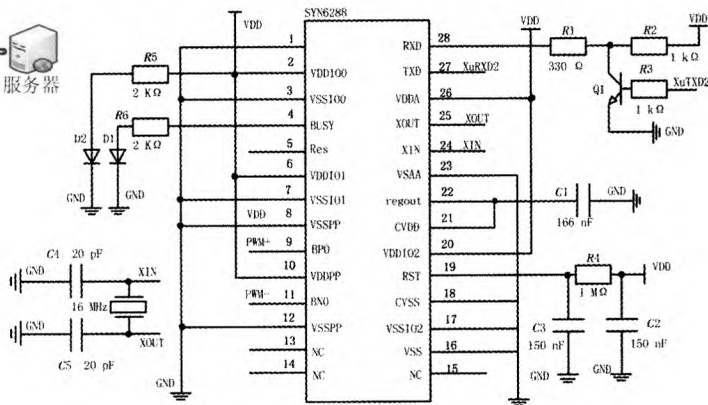


图3 语音播报模块应用电路图

3 智能导盲眼镜控制系统的软件设计

本系统的软件设计分为远端云平台服务器上的程序设计和智能导盲眼镜前端的程序设计两部分。远程云平台服务器上的软件设计主要是采用 C/C++ 等高级编程语言将图片的识别、测距、方位检测等相关算法转换成计算机系统能够识别的程序指令,从而实现智能导盲眼镜远程云服务器的识别、测距、方位检测等功能。另外 GPS 的经纬度解析也在云服务器上通过软件编程实现。导盲眼镜前端软件设计主要包括双目摄像头的图片采集、GPS 地理位置坐标的获取、无线通信传输模块的数据传输与接收、语音播报模块的调度和按键中断的配置等功能模块子程序的编写,其主程序流程图如图 4 所示。

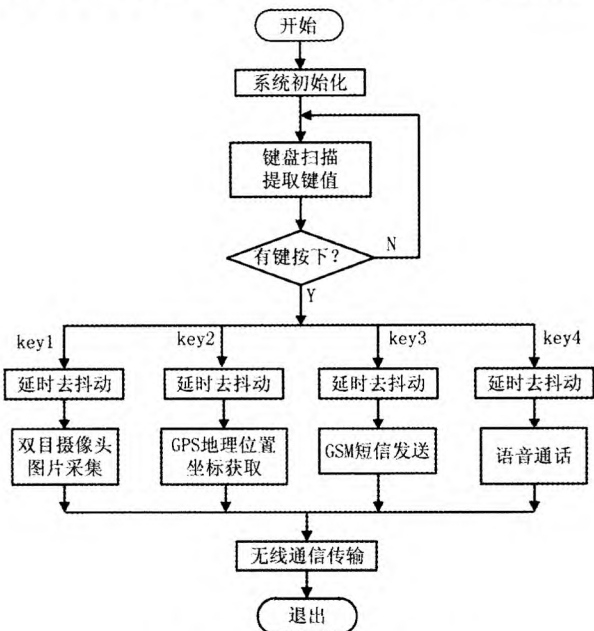


图 4 智能导盲眼镜前端主流程图

4 基于深度信念网络的自然场景识别

深度信念网络(Deep Belief Network, DBN)^[3]作为深度学习中最为广泛应用的一种算法模型,多应用于手写字体识别和自然场景识别这两个方向。在智能导盲眼镜控制系统中,深度信念网络主要用在自然场景中常见的物体识别上,这也是深度学习在机器视觉领域中的一种应用体现。图5是一个典型的深度信念网络的网络结构模型。

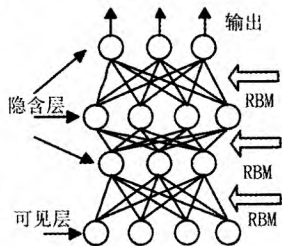


图5 深度信念网络的典型网络结构

从图5可以看出,深度信念网络是由多个受限制玻尔兹曼机(Restricted Boltzmann Machines, RBM)^[4]组成的深层网络,在这个网络中,DBN的训练过程采用逐层训练的方法,每一层RBM都单独训练,参数也是单独调整^[5]。训练完一层后,将该层的训练结果作为下一层RBM的输入,直至每一层RBM都训练完成,这个过程称之为预训练。当深度信念网络中的所有RBM训练完成后,再根据样本的标签值,采用反向传播算法向后微调。

在智能导盲眼镜的图像训练过程中,采用改进型的CIFAR-10自然场景库作为测试训练样本。原始的CIFAR-10数据集有60 000张32×32的彩色图像,分为飞机、汽车、猫、鸟、鹿、狗、青蛙、马、船和卡车共10类。为了增强智能导盲眼镜的实用性,结合盲人的特殊情况,本文在实际的系统图像训练中,加入了餐桌、椅子、人、垃圾桶、树木等生活中常见的一些目标场景,对原始的CIFAR-10自然场景库进行了改进,然后采用改进型的CIFAR-10自然场景库通过图6所示的深度信念网络模型进行训练识别。

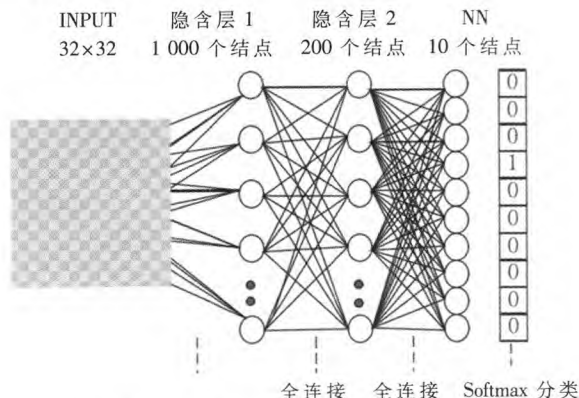


图6 智能导盲眼镜深度信念网络训练模型

在上述训练模型中,改进型CIFAR-10自然场景库中每张图均为32×32的彩色图片,因此输入层大小为3 072个结点(3 072=32×32×3),两个隐含层节点数分别为1 000和200,在经过面向多元分类的Softmax分类器^[6]

后输出层为10个单元,系统最终模型结构为3072-1000-200-10。在智能导盲眼镜系统实际的图像训练阶段,训练模型中的两层RBM的训练迭代次数都设置为200,学习率都设置成0.1。训练完成后,将系统学习到的权重用于初始化神经网络,对网络参数进行参数微调,并用Sigmoid函数^[7]激活神经网络。系统训练过程中,由于样本繁多,数据繁杂,隐含层需要相对较多的节点数才能学习到较好的特征,加之图片本身信息量大,需要较多次的迭代,整个训练过程平均历时10个小时,这与深度学习中的卷积神经网络^[8]以及自动编码模型^[9]相比,训练时间大幅度缩短,并且还具有较为理想的识别率,这也是本系统选用深度信念网络作为识别训练模型的主要原因。

5 系统测试结果与分析

采用上述的深度信念网络训练模型先对改进型CIFAR-10库中的10 000张测试图片随机进行了示范性测试,表1所示为各类样本的识别率和平均识别率。

表1 各类样本的识别率

种类	汽车	摩托车	餐桌	椅子	垃圾桶	树	人	石头	猫	狗
样本数	1 100	940	1 060	990	1 140	1 010	1 070	910	930	850
识别个数	920	792	879	839	933	839	908	742	760	690
识别率/%	83.6	84.3	82.9	84.7	81.8	83.1	84.9	81.5	80.8	81.2
平均识别率/%	82.9									

从表1可以看出改进型CIFAR-10库中的10类测试样本通过导盲眼镜系统的深度信念网络训练模型后,其平均识别率为82.9%,与基于支持向量机^[10]训练识别模型的识别率相比,其平均识别率超出了支持向量机模型的近10%,为进一步的智能导盲眼镜整体系统测试奠定了基础。最后,针对盲人的生活需要,结合智能导盲眼镜的其他功能,对智能导盲眼镜控制系统在实际场景中进行了系统联调,其中双目摄像头的采集帧率设置为3帧/s,语音导航频率设定为每两秒钟导盲一次,远程服务器端测试结果如图7、图8所示。通过图7可以看出智能导盲眼镜的GPS功能能实时、准确地获取佩戴导盲眼镜用户的经纬度,并通过无线通信传输模块送至远程



图7 智能导盲眼镜GPS服务器测试图

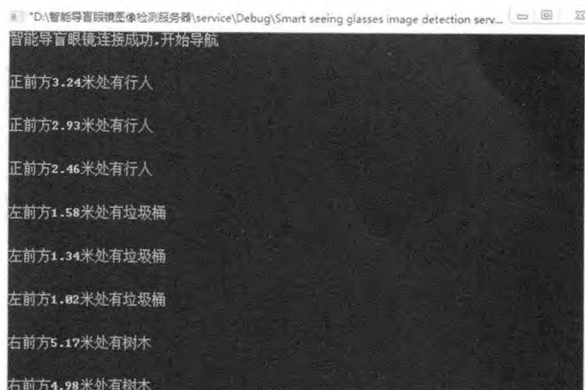


图8 智能导盲眼镜图像检测服务器测试图

服务器,进行地理位置的准确匹配。图8显示该智能导盲眼镜在实际场景中,不仅能较准确识别出前方目标的类别,还能测出目标物体到导盲眼镜的距离,正确表示出障碍物的方位特征,并在眼镜端实时以语音导盲的形式帮助盲人及时有效地避开障碍物,从而保障了盲人的安全出行。

6 结语

本系统以 S5PV210 为主控制器搭建基于机器视觉的智能导盲眼镜,通过搭载 Linux 内核,配备双目采集、GPS 定位、语音播报、GSM 短信、语音通话、无线传输六大核心功能模块组成系统,完成了系统的硬件电路设计和软件设计。通过系统测试,该智能导盲眼镜不仅能对盲人的独立出行进行实时语音导航,而且在特殊情况下,盲人可以通过智能眼镜上的触发按键,使用导盲眼镜的 GPS、GSM 短信、语音通话等功能,及时向亲朋好友获取帮助。另外由于智能导盲眼镜还具有图像识别能力,能帮助盲人进行简易物品归类,在一定程度上也使盲人的生活自理能力得以提升,这对于盲人基数较大的中国而言,显得尤为重要。

参考文献

- [1] Milan Sonka, Vaclav Hlavac, Roger Boyle, 等. 图像处理、分析与机器视觉[M]. 北京: 清华大学出版社, 2016.
- [2] 岳荣刚, 王少萍, 李凯, 等. 基于相似原理的新型双目测距法[J]. 光电工程, 2008, 35(4): 64-68.
- [3] 陈翠平. 基于深度信念网络的文本分类算法[J]. 计算机系统应用, 2015, 24(2): 121-126.
- [4] 张春霞, 姬楠楠, 王冠伟. 受限波尔兹曼机简介[J]. 工程数学学报, 2013(2): 159-173.
- [5] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. Computer Science, 2012, 3(4): 212-223.
- [6] 王爽, 马文萍, 谢慧明, 等. 一种基于栈式编码和 softmax 的极化 SAR 图像分类方法[P]. CN104156728A, 2014.
- [7] 张雪伟, 王焱. 基于 Sigmoid 函数参数调整的双隐层 BP 神经网络的板形预测[J]. 化工自动化及仪表, 2010, 37(4): 42-44.
- [8] 陈先昌. 基于卷积神经网络的深度学习算法与应用研究[D]. 杭州: 浙江工商大学, 2013.
- [9] 吴海燕. 基于自动编码器的半监督表示学习与分类学习研究[D]. 重庆: 重庆大学, 2015.
- [10] 崔鹏宇. 基于支持向量机的分类器训练研究[J]. 数字技术与应用, 2016(6): 58-58.

(收稿日期: 2016-10-14)

作者简介:

何腾鹏(1990-), 男, 硕士, 主要研究方向: 电路与系统。

张荣芬(1977-), 女, 教授, 硕士生导师, 主要研究方向: 嵌入式系统与视频图像处理、大数据应用。

刘宇红(1963-), 通信作者, 男, 教授, 硕士生导师, 主要研究方向: 嵌入式通信系统、大数据应用, E-mail: liuyuhongyx@sina.com。

研究与实现[D]. 杭州: 浙江大学, 2014.

- [10] 陈铖颖, 杨丽琼, 王统. CMOS 模拟集成电路设计与仿真实例: 基于 Cadence ADE[J]. 北京: 电子工业出版社, 2013.

(收稿日期: 2016-10-17)

作者简介:

王帅(1990-), 男, 硕士研究生, 主要研究方向: 专用集成电路设计。

黄海生(1964-), 男, 教授, 硕士, 主要研究方向: 专用集成电路设计与系统研究。

(上接第 57 页)

- [6] LI X Q, WEI Q, QIAO F, et al. Balanced switching schemes for gradient-error compensation in current steering DACs[J]. IEICE Trans Elec, 2012, 95(11): 1790-1798.
- [7] RAMASAMY S, VENKATARAMANI B, RAJKUMAR C K, et al. The design of an area efficient segmented DAC. IEEE, 2010.
- [8] Cong Yonghua, GEIGER R L. Formulation of INL and DNL Yield Estimation in Current-Steering D/A Converter[C]// ISCAS 2002 IEEE international symposium. 2002, 3: 1959.
- [9] 薛晓博. 高速高精度电流舵数模转换器关键设计技术的