

# 基于位置的增强现实街景留言系统<sup>①</sup>

蒋 浩

(浙江理工大学 信息学院, 杭州 310018)  
通讯作者: 蒋 浩, E-mail: [644549432@qq.com](mailto:644549432@qq.com)



**摘 要:** 人们对移动社交的需求越来越高, 然而陌生人之间想要交流却又有安全等其他问题. 为了解决这些问题, 设计了街景留言系统, 联结同一场景下“共同感受”的人, 区别于以往的社交方式, 更能让人融入这个场景. 留言榜上可以分享自己在这个景点的感想, 形成景点弹幕. 本系统由三大模块组成, 深度学习模块、数据库模块以及显示模块. 深度学习将扫描场景的图像与数据库中已有场景的图像进行图像匹配, 最后通过 AR 方式显示出来. 本系统能提供别样的生活方式, 给人们带来便利.

**关键词:** 社交; 深度学习; AR; 数据库

引用格式: 蒋浩. 基于位置的增强现实街景留言系统. 计算机系统应用, 2019, 28(9): 95-101. <http://www.c-s-a.org.cn/1003-3254/7041.html>

## Location-Based Augmented Reality of Street View Message System

JIANG Hao

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

**Abstract:** The need for social activities is thriving while there does exist security issues during the communication between strangers. In order to address those problems, street view message system is designed to connect someone who "commonly feel" in the same scene, which differ from the previous way of socializing and is better for people to get into this scene. People can share their thoughts of scenic spot on the message board. Those thoughts can form into a barrage attraction. This system consists of three modules: deep learning module, database module and display module. Deep learning matches the images of scanned scenes with those of existing scenes in the database, and finally displays them in AR mode. This system can provide a different way of life and bring convenience to people.

**Key words:** social; deep learning; AR; database

增强现实技术 (Augmented Reality, 简称 AR), 是一种将真实世界信息和虚拟世界信息“无缝”结合的技术, 该技术将虚拟的信息 (视觉信息、声音、味道等) 通过电脑等系统处理, 模拟仿真后叠加在真实世界中, 表现为真实的环境和虚拟的物体实时地叠加到同一个画面或空间, 满足了用户超越现实的感官体验. 深度学习是近年来在大数据环境下兴起的一门机器学习技术, 其实质是通过构建具有多层结构的机器学习模型和海量的训练数据来学习数据中的复杂特性, 用于

之后的分析、预测等应用. 由于其具有海量数据的学习能力和高度的特征抽象能力, 在诸多应用领域都取得了相比传统方法的革命性的成功.

目前经典的 AR 软件包括谷歌星空地图, 用户只需要将摄像头对准天空, 就可以在应用中看到各大恒星、星座的具体位置和准确信息, 同时用户也可以搜索特定的星星. 星空地图反应快速, 使用起来非常流畅, 几乎成了天文爱好者的必装 app 之一. 以及瑞典著名的家具品牌宜家推出的“IKEA Place”, 该应用可以

<sup>①</sup> 基金项目: 国家自然科学基金青年基金 (61501402)

Foundation item: Young Scientists Fund of National Natural Science Foundation of China (61501402)

收稿时间: 2019-02-28; 修改时间: 2019-03-14; 采用时间: 2019-03-18; csa 在线出版时间: 2019-09-05

自动扫描环境,并根据房间尺寸调整3D模型的大小,逼真模拟家具在房间放置的情境,为家具零售业带来一场变革.但是可以看到,这些产品的功能以及其与人的互动方式都较为单一,用户粘性不高,软件使用频率低.

据CNNIC发布的报告,截至2018年6月,我国手机网民规模已达7.88亿,智能手机已经逐渐取代传统PC成为主要的上网终端,人们对移动社交的需求也随之越来越高,然而陌生人之间想要交流却又有安全等其他问题.

本系统通过创造一种新的社交方式来打破这个格局.而留言墙系统,既能保护发言人的隐私,又能让不认识的人一起交流,作为虚拟社交的一种典型应用,具有受众广、用户粘性大、活跃度高的特点.联结同一场景下“共同感受”的人,区别于以往的社交方式,更能让人融入这个场景,同时留言榜上可以发布寻物启事以及广告,更加丰富了人们的生活方式以及更具有商业价值.特别是基于地理位置服务(LBS)的移动留言系统,因其目标人群定位准确,具有较高的认同度,无论在产品营销、广告宣传还是信息咨询等方面都具有巨大的发展潜力和商业价值.

目前国外也有关于AR留言墙的报道,但是不太人性化,只有单一的普通留言功能.而本系统希望实现的留言系统不仅仅是简单的留言,而是一个聚合平台,商家以及政府机关都可以接入,留言系统可以显示消费者对商家的客观评价、商家也可以自行留言推销产品与促销,以及民众对政府机关的办事效率评价或是政府各个窗口排队的实时人数等.

## 1 系统总体结构

系统总体流程图如图1所示.

1) 构建一个深度学习模型进行图像的匹配,其作用主要是将扫描场景的图像与数据库中已有场景的图像进行图像匹配,筛选出相似度最高的场景图像以待后续操作使用.

2) 根据系统的信息需求、处理需求和数据库的支撑环境,设计出高效、有序和支持多种数据格式的数据库系统.在通过深度学习模型得到相似度最高的场景信息之后,则需要在该数据库中提取出该场景信息对应的留言信息,包括文字、语音、图片、视频.

3) 利用AR技术,将真实场景和虚拟对象进行叠

加,这些虚拟对象包括文字、图像、语音和视频等.其中要利用到虚实结合技术将最后的结果显示到手机上;利用实时交互技术来实现虚拟对象与现实世界之间的交互;利用三维注册技术将虚拟对象准确的叠加到真实环境中去.

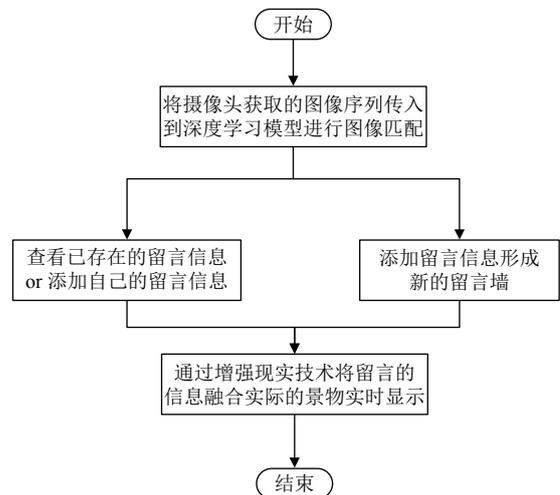


图1 总体流程图

## 2 系统的设计

从基于深度学习的街景对象识别的设计、基于MongDB的街景对象数据库的设计、基于AR技术的留言信息展示的设计三方面进行介绍.采用的三重神经网络,对特定的场景的深度学习比传统方法更加的高效,突破传统识别方式的缺陷,能降低环境变化带来的影响.

多层次图像数据库设计.本系统采用地方数据库的形式,将固定范围内的数据作为一个数据包,结合用户实时的定位信息进行匹配.能够保证数据库大小在一定范围内,便于管理员管理;在进行图片匹配时,能够加快图片匹配速度、提高匹配准确率.

突破了传统AR技术的应用(现实世界相关信息、视频介绍等),本系统发展用户与虚拟信息进行交互功能,交互功能包括用户可以创建留言栏、添加留言、收到回复的留言等等.

### 2.1 基于深度学习的街景对象识别

景物和建筑的识别的目的是精准地调用相应的数据库,识别功能是本系统的基础,只有精准的识别才能保证产品的可靠性.近年来,人脸识别大热,市场上智能手机都开始配备这项功能如iPhone XS、Mate 20

Pro 等. 但是区别于人脸识别, 本系统所需要识别的街景是建筑和景物. 并且由于是在室外, 会受到光照、遮挡和场景动态变化的干扰. 基于以上几点, 我们开始研究适合我们系统的深度学习神经网络结构并对已有的技术进行改进优化.

基于图像的局部特征, 可以发现具有重复性的地方, 一定程度上降低了视觉角度和光照方面带来的影响. 识别的时候重点检查这些局部特性上的几何特性和几何排布特性可以进一步提升结果的可靠性. 同时, 全局照片也可以用来检测, 可拓展性更强, 但是在光照和视角变化情况下, 识别率会比较低.

目前已经有一些基于模型或者数据学习的工作, 来设计一个更高效的从变化事物中找到不变特征的方法. 我们研究的方法基于卷积神经网络 (CNN), 因为他们具有更高的学习图像模式的潜力.

本系统采用经过改进的孪生神经网络——三重神经网络来进行特定场景的识别<sup>[1]</sup>.

孪生神经网络简单来说就是“连体的神经网络”, 神经网络的连体通过共享权值来实现, 其中共享权值指的就是左右两个神经网络的权重一模一样, 基本结构如图 2 所示. 孪生神经网络有两个输入, 通过最后输出的 Loss 来衡量两个输入之间的相似程度<sup>[2]</sup>.

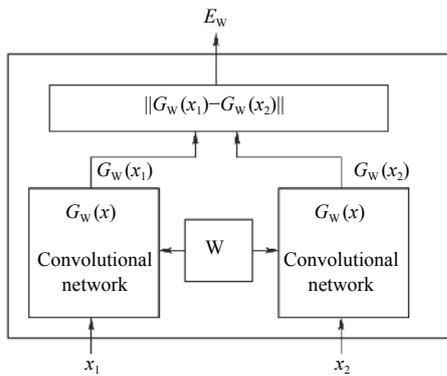


图 2 共享权值

孪生神经网络的目标是寻找两个可比较对象的相似程度 (例如, 签名验证、人脸识别等). 这个网络有两个相同的子网络, 两个子网络有相同的参数和权重.

孪生神经网络在场景图像匹配中的应用如图 3 所示. 第一个子网络的输入是一幅图片, 然后依次送到卷积层、池化层和全连接层, 最后输出一个特征向量. 最后的向量  $f(x^{(1)})$  是对输入  $x^{(1)}$  的编码. 然后, 向第二个子

网络 (与第一个子网络完全相同) 输入图片  $x^{(2)}$ , 我们对它做相同的处理, 并且得到对  $x^{(2)}$  的编码  $f(x^{(2)})$ .

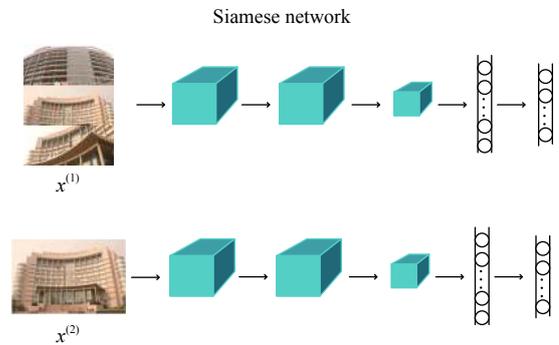


图 3 孪生神经网络的应用

为了比较图片  $x^{(1)}$  和  $x^{(2)}$ , 我们计算了编码结果  $f(x^{(1)})$  和  $f(x^{(2)})$  之间的距离. 如果它比某个阈值 (一个超参数) 小, 则意味着输入的图片与图像数据库中的某张照片为同一场景, 否则, 输入的照片并没有存在于图像数据库中, 即不为同一场景.

$$d(x^{(1)}, x^{(2)}) = \|f(x^{(1)}) - f(x^{(2)})\|_2^2 \quad (1)$$

上式是计算  $x^{(1)}$  和  $x^{(2)}$  的编码之间的距离<sup>[3]</sup>.

系统采用的三重神经网络则是在孪生神经网络基础上进行修改的, 分别在两个神经网络最后的输出增加了一个新的全连接层 (没有激活功能), 具体结构如图 4. 其中基本的神经网络将采用在 Imagenet 上训练的流行的 VGG-16 模型, VGGNet 的主要特性就是利用反复堆叠的  $3 \times 3$  小型卷积核和  $2 \times 2$  的最大池化核, 其目的是在增加网络深度的同时避免各层的输入大小不会随深度的增加而极具减小. 而最后的 loss 则采用欧几里德距离来衡量输入的相似性. 整个工作流程如图 5<sup>[4,5]</sup>.

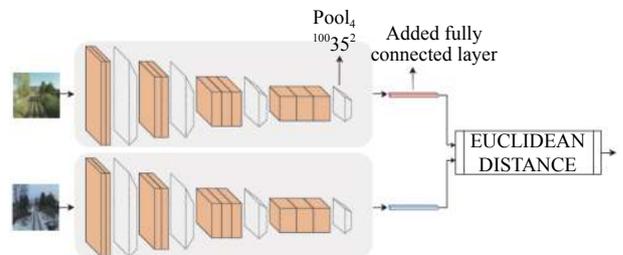


图 4 三重神经网络结构图

2) Triplet loss 三重损失函数定义:

$$L = \text{Max}(d(A, P) - d(A, N) + \alpha, 0) \quad (2)$$

所谓三重损失函数指的是使用三幅图片计算一个损失函数:一副固定影像 A,一副正例图像 P (与固定影像是同一场景),以及一张反例图像 N (与固定影像不是同一场景)三重损失函数处理过程如图 6<sup>[6]</sup>。

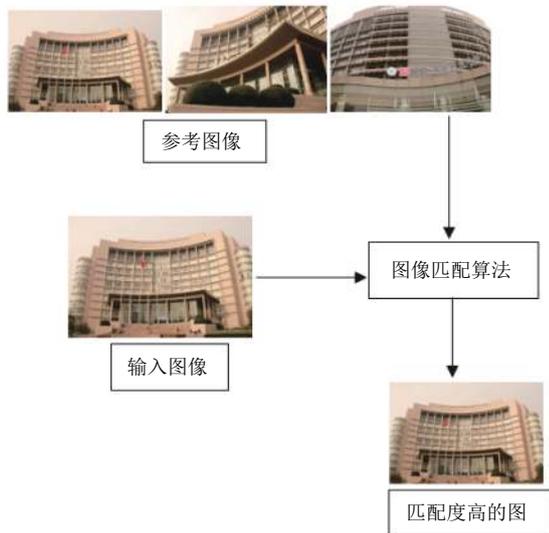


图 5 三重神经网络工作流程

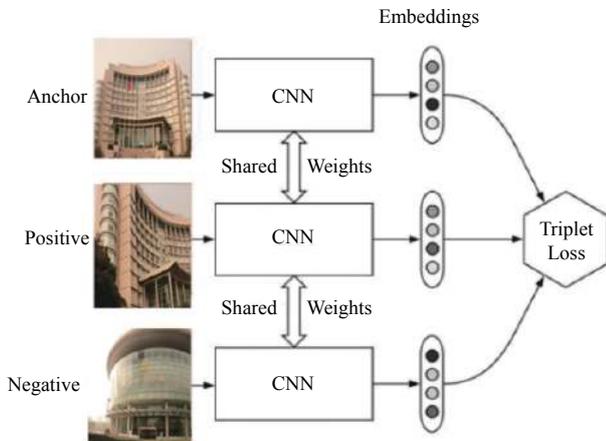


图 6 三重损失函数示意图

表 1 为采用改进的孪生神经网络——三重神经网络识别校园各个景点识别精度结果。

表 1 不同时间不同场景的街景识别精度

场景	晴天 (%)	阴天 (%)	雨天 (%)
图书馆	93.2	92.8	91.5
教学楼 1	90.5	91.2	93.4
教学楼 2	91.3	93.5	91.4
体育馆	90.5	91.5	90.3
钟楼	93.5	91.3	92.3

## 2.2 基于 MongoDB 的街景对象数据库设计

基于位置的增强现实街景留言系统中的用户信息、留言信息、留言墙信息和场景信息都需要数据库技术的支持.而其中涉及存取大量的文字,图片信息,甚至视频信息等,此外考虑到此系统以后的被推广以及大量用户的使用造成数据量的极大上升带来的查询效率等一系列问题.这里我们采用的是基于分布式文件存储的 MongoDB 数据库系统. MongoDB 是一个开源的,基于分布式的,面向文档存储的非关系型数据库,是非关系型数据库中功能最丰富,并且最像关系型数据库的. MongoDB 由 C++ 语言编写,旨在处理海量的数据,并能提供高效的性能和良好的可靠性<sup>[7,8]</sup>.

本系统研究基于位置的增强现实街景留言系统,该系统通过基于地理位置服务 (Location Based Service, LBS) 获取当前位置附近的热门留言墙,扫描目标景物,然后通过增强现实技术将留言信息叠加到移动终端捕捉到的场景中<sup>[9,10]</sup>.

该数据库主要包括用户信息、留言信息、留言墙信息和场景信息 4 个实体.

其中场景信息存储有手机 GPS 和陀螺仪的信息,其目的是用于目标匹配时的第一轮筛选和推荐附近热门区域.

在当用户对目标景物进行扫描时,可通过手机定位信息来判断是否进入深度学习模型来进行图像匹配,这样可以大大避免运算资源的浪费.

在当用户进入系统界面时,系统可以根据手机的定位信息在地图上标注出附件热门留言墙以及热门留言墙上 TOP3 的简要信息,引导人们去高人气地区.

用户信息中还可以包含有微信帐号关联的信息,然后用户可以查看好友中热门的留言区域,同时也可以跟微信好友进行互动,在看到自己喜爱的内容时进行点赞.

开发基于位置的增强现实街景留言系统.经过可行性分析和初步的需求调查,确定了系统的功能边界,该系统能完成以下的功能:

- 1) 用户注册 (可以微信账号注册).
- 2) 用户登录.
- 3) 用户可以查看自己的相关信息,如在哪些地方留过言,留言的关注度、点赞数等.
- 4) 根据用户当前地理位置信息显示附近留言墙信息.
- 5) 用户对目标物体进行扫描,显示当前标志物的留言墙及留言信息,若无留言墙,可以创建留言墙.

6) 用户可以查看当前打开的留言墙上的信息, 其中留言墙上微信好友的留言会有标识, 也可以在上面以文字、语音、图片、视频等方式在上面留言.

图7反映了基于位置的增强现实街景留言系统与用户的接口, 但未表明数据的加工要求, 需要进一步细化. 根据前面基于位置的增强现实街景留言系统功能边界的确定, 再对基于位置的增强现实街景留言系统顶层数据流图中的处理功能进一步分解, 可分解为用户注册、登录、查询个人信息、查询附近留言墙、扫描目标物体查看留言墙或创建留言墙和留言六个子功能, 这样就得到了基于位置的增强现实街景留言系统的第0层数据流图.

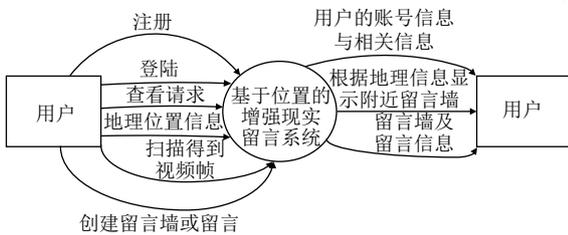


图7 系统顶层数据流图

从图8可以看出, 在此系统的不同功能中, 有必要对其进行更深层次的分析, 图9为扫描目标物体查看留言墙或创建留言墙数据流图<sup>[11,12]</sup>.

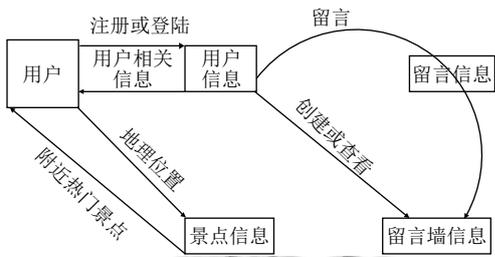


图8 系统第0层数据流图

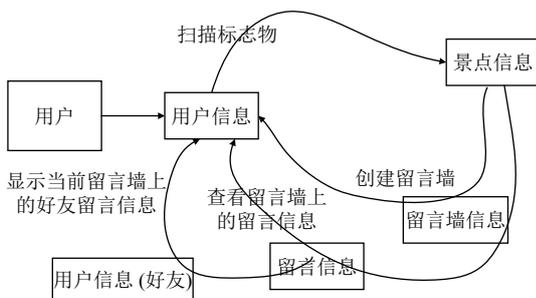


图9 查看留言墙或创建留言墙数据流图

### 2.3 基于 AR 技术的留言信息展示

本系统在 Unity 下基于高通 (Vuforia) 来实现虚拟事物在现实中的显示.

如图10为基于 Unity3D 的增强现实系统的工作流程, 也即是增强现实实现原理, 大体上包括图像获取、图像匹配以及虚实结合等部分内容<sup>[13,14]</sup>.

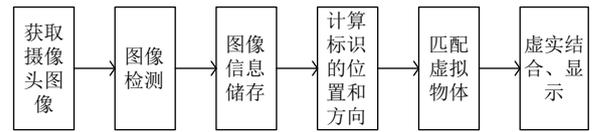


图10 AR 工作流程

系统通过摄像头获取到识别对象, 通过识别对象上的特征点将它与系统中本已上传的识别图进行匹配达到虚拟与现实结合的目的. 而实现这一功能, 我们采用 Vuforia SDK. Vuforia 调用计算机摄像头, 利用视觉技术实时识别和捕捉来控制虚拟物体与识别图位置关系.

Vuforia 对被识别的模型有着相应的要求:

1) 对象被假定为静态的, 在被检测到后, 将不得不在环境中保持静态. 用户则可以在对象周围移动, 但不移动对象本身. 如果需要支持移动或者替换对象, 在对象移动后停止并重新开始追踪;

2) 相对来说, 单一统一颜色的物体更加难以追踪, 尽管他们通常可以被可靠地检测到. 对于这些对象, 往往需要通过表面外观的一些变化来区分, 因此, 纯白背景上的纯白色或者纯黑背景上的纯黑色物体区分效果并不明显; 相比较于几何复杂的对象, 诸如立方体这类的简单形状或者非常细长的简单形状更加容易被混淆;

3) 而很重要的一点, 对象应当具有与模型完全相同的形状, 虽然模型目标跟踪器允许它们各自的几何形状有一些偏差, 但肯定不能是柔性的物体.

4) 在实例中, 用户使用手机扫描相应的实物或者实景, 通过手机摄像头来捕捉图像, 将图像上传至手机后台数据库中进行图像匹配.

在进行识别图读取定位后, 计算机已经拥有了虚拟物体的三维坐标, 下一步即时将它与真实世界进行结合. 这个过程包括3个部分: 三维注册、物体位置、光照条纹. 三者协调工作, 当虚拟物与现实场景出现失真情况时, 三维注册进行实时定位跟踪, 不断调取特征值进行匹配, 具有自我纠错能力<sup>[15]</sup>. 三维注册技术是实

现移动增强现实应用的基础技术,也是决定移动增强现实应用系统性能优劣的关键,因此三维注册技术一直是移动增强现实系统研究的重点和难点.本系统主要是实时检测出摄像头相对于真实场景的位姿状态,确定所需要叠加的虚拟信息在投影平面中的位置,并将这些虚拟信息实时显示在屏幕中的正确位置,完成三维注册<sup>[16,17]</sup>.

### 3 系统展示

首先弹出一个登陆界面,用户需要登陆自己的账号如图 11,如果没有要先进行注册.



图 11 登陆界面

登陆后,软件需要取得用户手机位置获取的权限获取用户当前的地理坐标,下载相关的场景包,配合手机后置摄像头,弹出一个新界面.然后用户对准所要识别的景物进行扫描如图 12 所示.

扫描成功之后会弹出这个景物所属的留言墙如图 13 所示,你可以查看别人在留言墙上发布的信息,也可以自己发布信息,甚至还能看到有关扫描的景物的有关信息,附上一个范例留言墙如图 14 所示.

为了使这个系统更加的完善,带来更好的体验感,加入了最近识别点,热门景点和足迹三个子功能,点击最近的识别点能出现离你距离最近的一系列留言墙,热门景点则是根据事实的留言数选进附近的几个热门景点,用户可以前去扫描参与留言.同时还会记录下你

扫描过的街景,方便你想看看自己留下的痕迹与留言之后与其他用户的互动.这些功能如图 15 所示.



图 12 扫描街景界面



图 13 识别成功后弹出界面



图 14 其他范例留言墙



图 15 其他辅助功能

## 4 结束语

这是一款基于位置的移动增强现实街景留言系统, 用户可以将该系统安装于手机上. 这个留言系统不仅仅是简单的留言, 而是一个聚合平台, 后期商家, 政府机关都可以接入, 留言系统可以显示消费者对商家的客观评价, 商家也可以自行留言推销产品与促销, 民众对政府机关的办事效率的评价.

### 参考文献

- 1 Yu TS, Wang RS. Enhancing scene parsing by transferring structures via efficient low-rank graph matching. Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. Burlingame, CA, USA. 2016. 22.
- 2 王坤. 基于深度学习的自然图像分类方法的研究[硕士学位论文]. 南昌: 东华理工大学, 2017.
- 3 张安琪. 基于孪生卷积神经网络与三元组损失函数的图像识别模型. 电子制作, 2018, (21): 49-50, 20. [doi: 10.3969/j.issn.1006-5059.2018.21.017]
- 4 沈雁, 王环, 戴瑜兴. 基于改进深度孪生网络的分类器及其应用. 计算机工程与应用, 2018, 54(10): 19-25. [doi: 10.3778/j.issn.1002-8331.1801-0269]
- 5 吴汉钊. 基于孪生卷积神经网络的人脸追踪. 计算机工程与应用, 2018, 54(14): 175-179. [doi: 10.3778/j.issn.1002-8331.1801-0369]
- 6 Olid D, Fácil JM, Civera J. Single-view place recognition under seasonal changes. arXiv: 1808.06516, 2018.
- 7 Boicea A, Radulescu F, Agapin LI. MongoDB vs oracle -- database comparison. Proceedings of the Third International Conference on Emerging Intelligent Data and Web Technologies. Bucharest, Romania. 2012. 330-335.
- 8 Plugge E, Hawkins T, Membrey P. The Definitive Guide to MongoDB: The NoSQL Database for Cloud and Desktop Computing. Apress Berkely, CA, USA: Springer, 2010.
- 9 徐旭平, 李小勇. 基于 MongoDB 的元数据管理研究. 信息技术, 2018, (8): 87-93.
- 10 张天宇, 贺金鑫, 王阳, 等. 基于 NoSQL 数据库的地质大数据高效存储方法. 吉林大学学报(信息科学版), 2013, 31(6): 604-608. [doi: 10.3969/j.issn.1671-5896.2013.06.008]
- 11 杨磊. 基于 NoSQL 数据库的结构化存储设计与应用. 科技风, 2011, (18): 99. [doi: 10.3969/j.issn.1671-7341.2011.18.084]
- 12 刘愉, 王立军. 基于 MongoDB 的 EHR 存储方案研究与设计. 中国数字医学, 2013, 8(6): 20-24. [doi: 10.3969/j.issn.1673-7571.2013.06.006]
- 13 朵雯娟. 虚拟现实图像增强现实 AR 的使用价值与技术研究. 现代电子技术, 2018, 41(19): 75-78.
- 14 雷卓石. 虚拟与增强现实技术. 科技创新导报, 2018, 15(3): 150, 152.
- 15 周见光, 石刚, 马小虎. 增强现实系统中的虚拟交互方法. 计算机工程, 2012, 38(1): 251-252, 255. [doi: 10.3969/j.issn.1000-3428.2012.01.082]
- 16 王坚. 虚拟现实和增强现实的人机交互技术. 计算机技术在工程建设中的应用——第十二届全国工程建设计算机应用学术会议演讲辑录. 北京, 2004: 18. [doi: 10.12120/bjutsxb20040418]
- 17 高宇, 邓宝松, 杨冰, 等. 基于增强现实的虚拟实景空间的研究与实现. 小型微型计算机系统, 2006, 27(1): 146-150. [doi: 10.3969/j.issn.1000-1220.2006.01.033]