# 目录

浅谈计算机视觉技术的发展及应用	2
1.引言	3
2.计算机视觉技术的关键技术发展概况	4
2.1 计算机视觉中的关键技术的发展	4
2.1.1 光源	4
2.1.2 光学镜头	4
2.1.3 摄像机	5
2.1.4 图像采集卡	
2.2 计算机视觉系统的工作原理	5
2.3 计算机视觉的简要发展历程	5
2.4 计算机视觉发展态势	5
3.计算机视觉的应用	6
3.1 计算机视觉在生活方面的应用	6
3.1.1 视觉生成	6
3.2 智能识别,目标检测	6
3.2.1 人脸识别	6
3.2.2 人脸表情识别	7
3.2.3 其他新兴应用	
3.3 计算机视觉技术在工业农业生产中的应用	
3.3.1 分拣技术的应用	
3.3.2 计算机视觉对薄膜厚度的检测	
3.3.3 计算机视觉在汽车领域的应用	
3.4 城市研究中计算机视觉的应用	
4.结束语	
[参考文献]	

# 浅谈计算机视觉技术的发展及应用

姓名: 赵炫皓 学号: 1120200603 班级: 30012003 班

摘要:本文从计算机视觉对人们生活的巨大影响谈起,介绍了计算机视觉实现的 关键硬件条件和完整体系,简述了计算机视觉的起源和发展历程,概括列举了该技术在生活和工农业生产中的应用实例,阐明了计算机视觉广泛应用的实际意义和未 来发展态势,并说明它发展的不足之处,猜测了未来的发展方向。

关键词: 计算机视觉技术,发展历程,发展方向,实际应用

# 1.引言

随着科技的日新月异,计算机技术技术和大规模集成电路的迅猛发展,计算机渐渐成为大众生产生活中不可缺少的一部分。在日常生产生活中,办公室办公需要计算机,信息处理需要计算机,甚至维护社会治安也离不开计算机的帮助……计算机的普及也相应地带动了计算机相关专业的发展,就在近30年来计算机开始真正得到普及,融入到广大人民的生活中,计算机相关的技术也便因此落地生根,成为人们生活当中的一项项重要的工具。

计算机的热潮也相应地点燃了发展人工智能行业的激情。人工智能不再像从前只出现在科幻电影中那样神秘,现在也逐渐被人们了解,开始尝试融入人们的生活当中。

而计算机视觉技术作为人工智能领域

研究的重要领域,由于其应用广泛,社会效益巨大,落地实践较为现实,在近些年持续升温,飞速发展,开始渐渐渗透到人们生活的方方面面:小到出门购物,借助微信和支付宝扫码支付;大到城市规划建设,稳定社会治安,全面实现无人驾驶技术……这些事情都需要依靠计算机视觉技术来完成。这项技术在很大程度上提高了人们生活水平,促进了工业和农业的生产,为人类带来了巨大的便利。

计算机视觉的出现也想人类拥有视觉一样重要。视觉作为人类与生俱来的一种十分重要的感官,帮助人获得了80%的信息<sup>[1]</sup>,在日常的生活中发挥着不可替代的作用。而计算机视觉,顾名思义,就是用计算机去模拟人眼去观察这个世界,获得外界的信息的一项技术。但是计算机毕竟是机器,与人类的血肉之躯有所不同,因此计算机看到的世界与人类所见也大相径庭。

如图 1<sup>[2,3]</sup>所示,普通人眼中一只灰黑色条纹的猫,在计算机看来,似乎只是一串没有任何意义的数字。但实际上,在计算机内部,这只猫的图片被分解成为一个个小的像素点,每一个像素点由 3 个 0 至 255 之间的数字组成矩阵,计算机通过识别矩阵从而看到这张图片。

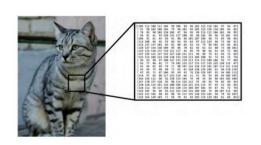


图 1<sup>[3]</sup> 在计算机设备中图像的呈现方法 但计算机视觉绝不仅仅止步于单纯 看到这一点上。计算机所要做的,是模 仿人类看到一件事情,在看到的基础上 进一步获取这张图片的信息,并进行合 理的分析推断,从而真正实现从"看到" 这张图片,迈向像人类一样"看懂"这张图 片的巨大飞跃,这样才能实现拟人化, 在极短的实践内代替人力去做重复和机 械的视觉工作,提高了信息处理和工作



的效率。

a group of zebras standing in a field (<u>Vinyals</u>, CVPR15] a group of zebras grazing on grass (You, CVPR16] a group of zebras grazing in a field (Yao, ICCV17) a group of zebras and a rainbow in the sky (Peter, CVPR18] a group of zebras grazing in a field with a rainbow in the sky (You

图 2[4,5] 计算机的描述图片"天空中的彩虹

例如图 2 所示,在 2015 年至 2018 年"看图说话"中,对于"天空中的彩虹"这幅图在这年来随着计算机视觉技术的进步被不断完善,表述更加全面而具体,越来越接近一个常人肉眼所见后的所思所想。这一经历了从将图片内检测到的实体信息输送到固定的模版上,到对图像进行特征方面的提取再经由循环神经网络(RNN)一个字一个字地进行解析抽取,再到神经网络融合的组合式框架进行关联调用与提炼处理,最终变得愈发精准这样的一个发展的过程<sup>[6]</sup>。因此透过这一幅图可以管中窥豹,看出计算机视觉技术的巨大进步,充满希望的前景以及广阔的应用领域。

计算机视觉的看见,更多的是看见图 片背后无法直接看见的东西,这种抽象 理解能力才是计算机视觉技术得到突破 的关键。

# 2.计算机视觉技术的关键技术发展概况

# 2.1 计算机视觉中的关键技术的发 展

如图 3 可见, 计算机视觉技术拥有多学 科交叉的背景, 融合了各类数字图像,

机械工程等方面的知识,是一项综合性的技术。总体而言,选择一个机器视觉的单个子系统是非常容易的,但是经过恰当的选择,将它们的作用有机结合在一起就有些困难了。其中最关键的技术是光源照明,光学镜头,摄像机和图形采集卡。

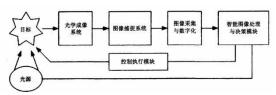


图 3<sup>[7]</sup> 典型工业中机器视觉的应用系统

#### 2.1.1 光源

就目前而言,光源的优劣与否决定了 计算机视觉系统的成败。好的光源能够 突出物体的相对特征,提高明显的对比 度,保证物体的亮度合适,不至于影响 识别,并且能够在清晰照明物体的同时 不影响成像的质量。

其中经过大量的实践<sup>®</sup>证明了 LED 灯 具有发光亮度大,使用寿命长,颜色丰 富,发热少,形状可塑等优良品质,因 而与高频荧光灯一样作为发光设备的首 要选择,在近些年来十分热门。

#### 2.1.2 光学镜头

光学镜头的好坏毋庸置疑直接影响 到最终成像结果的结果。这主要依赖于 镜头的焦距,分辨率,光阑系数,倍率, 接口等问题。不同工业环境所产出的相机质量会有较大差别,价位方面也会出入较大。

在工业视觉中镜头的选择,首先需要了解当前环境下所需相机的类型和传感器尺寸,从而计算焦距来选择合适的镜头<sup>[8]</sup>。

#### 2.1.3 摄像机

当前所应用的摄像机一般是:工业相机与机器视觉系统和嵌入式系统相结合的类型。传感器也分为CCD,CMOS两个类型。CCD由于其体量小,使用周期长,能够成熟物理冲击,画质高等等特点在计算机视觉领域尤其热门。细分它的类型,可以分成面阵式和线阵式两种类型。线阵式只能一次性拍摄从摄像机前直线运动的物体,才能将图像拍得完整。而面阵相机,而面阵式则可以一次成像,可用性强。因此当前的生产中,主要使用CCD摄像机<sup>[7,8]</sup>。

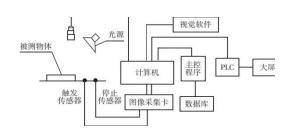
## 2.1.4 图像采集卡

由于图像的传输接口不能够满足图像 信号所需要的快速传输能力,因此需要 图像采集卡的帮忙<sup>[9]</sup>。

图像采集目的就是把物体的信号采集 下来,并将其保存。采集的图像的质量 会因前面几种关键元素的选择不同,从 而产生较大的出入。文献<sup>[10-11]</sup>曾提出过处 理图像边缘的方法。

#### 2.2 计算机视觉系统的工作原理

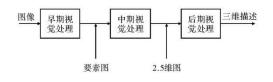
虽然计算机视觉所要处理的内容数目 庞大,具有极大的不确定性和复杂性, 但其内部构造上基本相同,主要以面积, 高度,坐标方位等要素为关键点,从而 进行形式化显示的一个机械过程<sup>[12]</sup>,具体 流程如图 4 所示;



# 2.3 计算机视觉的简要发展历程

机器视觉的发展可以追溯到上个世纪50年代的统计识别模式,历经20年的积蓄,在上世纪70年代开始呈现发展的态势,并涌现出很多科学的理论框架支撑它稳步发展。其中的 Marr 理论是较为完善的一个体系<sup>[9]</sup>。

Marr 将系统分为三个层次: 计算理论 层次, 硬件实现层次和算法层次。而视 觉系统自上而下分为 3 个步骤, 解释了 从原始数据对三维环境表达所经过的三 段过程。如图 5 所示。



#### 2.4 计算机视觉发展态势

由于生物学方面未能破解人类视觉机制之迷,因此计算机视觉体系还有太多漏洞等待修复。在可以遇见未来里,计算机视觉潜力巨大,应用领域广泛,智能摄像机将会占据市场主要的地位,逐渐实现自动化与智能化<sup>[11]</sup>。

# 3.计算机视觉的应用

# 3.1 计算机视觉在生活方面的应 用

#### 3.1.1 视觉生成

计算机视觉在大量深度学习后,在理解大部分图片以后,可以完成对损坏的视频或照片的修复,完成视频或海报的辅助设计<sup>[1]</sup>,在物件分类的基础上可以实现由文本输入图像输出的反向过程,在早期技术还不成熟的时候,所生成的图像辨识度极低,模糊不清。但随着时间的推移和计算机视觉技术的不断完善和改进,使得现如今生成的图象已经非常逼真自然。由图6可见。

画面中的山川,蝴蝶,蘑菇和气泡看似是自然界真实存在的事物,事实上是借助计算机视觉当中的视觉生成,输入描述性的词汇后,系统根据相应的特征检索有关的数据库,从而在海量的照片

中合成这些图片。



## 3.2 智能识别,目标检测

计算机图像识别和分类方向的 Image Net 比赛自 2010 年办赛以来,到 2017 年已经取得了突破性的进展。这几 年间深度学习,即卷积神经网络不断取 得新的进展,其在计算机视觉方面的应 用使得机器对于物体识别的错误率显著 降低。目前人类在其上数据集的精度提 高到了 94.9%左右<sup>[14]</sup>。

#### 3.2.1 人脸识别

人脸识别技术已经广泛引用于日常的生活中:面部解锁,门禁刷脸,扫描面部付款.....

但人脸识别技术涉及到大量个人隐私 问题,需要做到错误率几乎为 0,才能免 于造成不必要的损失。同时,这项技术 存在被破解的风险,因此应用此项技术 的团体或个人应注意个人隐私保护,不 断维护该项技术以免遭受攻击造成严重 后果。

#### 3.2.2 人脸表情识别

心理学家 Mehrabiadu 认为面部表情在 人们的交流中传达了半数以上的有效信息。人脸表情识别正是意识到面部表情 的重要性。主要识别方式就是在不同面 部映射单元上建立肌肉组织与基本表情 的联系,从而得出结论<sup>[16]</sup>。

这一点的识别与上述人脸识别十分相似,虽然实现的复杂程度要明显高于人脸识别,但所试错成本要远远低于人脸识别,几乎不存在隐私问题和重大损失。

面部表情的识别主要分为 Gabor 的面部提取法则,ASM/AAM(主动形状法则),Haar-like 特征提取的方式,以及光流法和特征点跟踪法,在提取识别面部特征算法后,分类特征的算法主要有K-NN 算法, Adaboost 算法以及 SVM 算法<sup>[15]</sup>。

几种机器学习算法之间存在着明确的 区别,又暗合着种种联系。不同算法结 合,得到识别结果的准确率出入较大。 因此在实际应用中需要不断调整找到较 优解。

#### 3.2.3 其他新兴应用[3]

许多电商平台都支持通过拍照来获取相关物品的商品信息。这是计算机视

觉技术能够落地企业的成功范例。这一 实践极大地促进了互联网信息与现实信息的交互,提高了用户的购物体验,而 且具有即时性,可操作性的大优势。

拍照识别商品的技术主要运用了目标检测,分类,深度特征学习及检索等技术,在获得特征后,通过离线特征库的检索,将相似产品进行重排序后反馈给用户。

此外,识别技术还可用于搭配推荐 算法中,在用户完成一次衣物的购物后, 为用户推荐合理的衣物搭配,目前这一 技术已经日趋成熟,其搭配效果堪比专 业的服装设计搭配师,在提高效率降低 成本的同时增强了拓展性。

这一技术也广泛地适用于内容审核中。内容审核可以确保互联网平台内容符合国家的安全规范,对内容进行严格把控。在海量的鉴别不同的信息后,将正常的内容与禁止的内容进行区别与分类,通过大量的学习从而掌握不良内容的基本特征。通过鉴别视频或图片中的违禁物品或暴恐行为等不良内容,从海量的信息中过滤掉不良的内容,减少了这类不良信息在互联网传播的风险,减少了审核工作人员的心理负担。该技术大提高了审核效率,极大地保障了社会稳定和人民幸福安康。

除此以外, 计算机视觉技术还可以应 用与智能安防领域<sup>[22]</sup>。主要通过安插多 台摄像机,可以监视密集场所的人员流动,实现动态视角下目标物体的精确捕获与追踪,有效增加了园区密集场所的 预警和管理。

还可以通过当前的人工智能数据库 分析出疑似的可以人员,通过实时动态 追踪将其步态与人工智能数据库中的步 态相比较,在发现步态一致或特征行为 之后,会发送警告信息告知相关的安全 工作者,并进行实时追踪。

智慧社区服务和智能养老服务作为 安防技术的典型应用对象,前者能够识别出家居者的情绪,并通过情绪与数据 库的映射关系做出相应的处理。同时在 住宅遭受入侵时,可以在第一时间告知 用户,从而是用户的安全得到保障,将 财产和人身安全的损失降低到最小。这一项目的实现得益于多摄像机检测目标的方法和智能化视频联动方式而落地实践的。

智慧养老主要帮助年轻用户检测老人行为,以及健康状况。依据骨骼识别算法快速计算人体骨骼当中 20 个特征点,通过特征点的局部描述分析出老人整体的骨骼状态和状况,最终通多相机运动检测,实现跌倒检测,夜间身体状况检测和分析老人行为的完整过程<sup>[22]</sup>。

智慧安防的实践中,如果全景视频技术得到飞速提升,那么智慧安防的理想将会迅速成为现实,大大降低与之相关

领域从业这的工作量,给用户带来极优 的生活体验。

# 3.3 计算机视觉技术在工业农业生 产中的应用

#### 3.3.1 分拣技术的应用

计算机视觉的发展使得工业过程中出现了轻量化。低耗能,成本低的工业视觉设备,整个智能生产行业的发展都由此推动<sup>[17]</sup>。

晏祖根<sup>[18]</sup>等人研发出能够高效进行食品分拣的智能分拣系统。这套系统在保障食品卫生降,低劳动力强度的同时提高了分拣效率。



(a) 原图像



(b) 二进制图像



(c) Sobel算子边缘检测图像

图 7[19] 关于孟加拉国货币的识别

Kamruzzaman<sup>[19]</sup>等人针对孟加拉纸币的不同类型,提出了纸币分级系统。平均分拣速度约为8张纸币每秒,对正常的纸币,它的分拣失误率几乎为0,大大提高了纸币分拣的效率<sup>[20]</sup>。如图七所示。

#### 3.3.2 计算机视觉对薄膜厚度的检测

薄膜厚度主要由自身生产环境和工艺设备,螺栓位置和薄膜厚度传感器决定。 计算机视觉技术可以对不同位置的螺栓进行位置追踪,确定螺栓位置,实现薄膜的均匀<sup>[25]</sup>。

#### 3.3.3 计算机视觉在汽车领域的应用

虽然在设备生产初期,用延时摄影的方式对车身进行质量检测能够分析生产问题,但由于其研发落地成本高,对精度要求比较高,没能得到大规模实现。

主要思路是用光源照射底面,从而获取被测物的位置信息,对车身不同位置的平整度和大小缝隙进行可持续化分析,保障生产车辆的整体一致性能有效消除误差,实现对于大规模生产质量的控制<sup>[25]</sup>。

# 3.4 城市研究中计算机视觉的应 用

计算机视觉在城市环境的研究中,最显著的一个作用是奎尔恰<sup>[23]</sup>等人用该技术研究城市环境中影响美感的要素。如图 8 所示。



图 5[23] 与美感相关的视觉识别元素

研究组分析了不同街区在美感,安静感,愉悦感方面的数据,发现色彩中发现不同的颜色和肌理对于人心理对于建筑的感受有不同的影响,可以具体到颜色的类别。同时,不同的视觉元素对于也会影响人类的观感体验。例如廉价住房,广告牌等会提高人们对于城市街区的负面评价。麻省理工学院发起的"地点脉冲"所收集的数据领域更加广泛,对研究城市布局和市容风貌,增强城市美感有着参考价值[24]。

# 4.结束语

随着时代的变换和科技的不断创新, 人工智能最终终将会实现,融入每个人 的生活当中,影响着整个社会科技的发 展和生产力的变化。这是科技发展的必 然趋势,每个人都应该顺应时代的浪潮, 去迎接智能化时代的到来,而不是以排 斥的心态去看待这种新的潮流。

在深度学习即卷积神经网络的加持下,计算机视觉技术经过数十年的发展,终于在当前开始落地生根。计算机现如

今可以在极短的时间内有效大规模处理 信息世界的海量信息,并作出相应的基 本识别与判断, 在社会生产生活中发挥 着难以替代的重要作用。但现如今计算 机只能像一个人类儿童一样看出所检测 图片的表层含义,进行简单的分析,无 法真正做到像人类一样, 把图片的全部 信息在短时间内充分挖掘出来。因此计 算机视觉在实际应用当中, 仍存在诸多 不足,例如在图像检测中有近5%的错误 率;在汽车制造生产中的技术,成本过高 且无法保证高精度;在智能安防,智能家 居领域由于运动检测的精度不够, 理论 与实践落地的距依旧相差很远.....由于 现实世界的复杂性和不确定性太多,因 此计算机所面临的挑战就是在不断的实 践中成功对付这些不确定因素,提高识 别的准确程度和降低错误率。

因为计算机视觉的人脸检测等方面涉及隐私与利益等问题,因此在实际应用中需要追求 0 出错率,这样才能保证应用时的精准无误。一旦出现识别失误的话不但会增加巨大的麻烦,同时还将对个人和集体造成不可估量的损失。因此在实践的过程中要等待技术发展成熟以后才能将它缓步地应用于实际之中。

虽然目前中国的计算机视觉技术理论 目前已经较为完备,每年的论文数量和 高质量论文数量在不断地攀升,微信支 付等等计算机科学技术的应用实例在华

夏大地上遍地开花,这得益于计算机发展浪潮时,中国工业飞速发展,赶上信息爆炸的时代,计算机视觉技术也顺着时代的浪潮飞速发展,但这无法掩盖中国的计算机科学技术与世界先进水平以及发达国家仍然存在较大差距,仍然需要奋斗在一线的科技工作者继续努力,使计算机视觉技术在理论和应用层面平衡发展,最终实现计算机视觉的繁荣与先进。

中国在计算机视觉飞速发展的时刻刚刚起步,虽然经过数十年的奋进在理论层面已经有所突破,但在应用层面仍有较大不足,甚至可以说在某些方面有些滞后<sup>[7]</sup>,难以将完备的理论体系进行落地实践,所以说这是我国计算机视觉领域发展的一大重点。

目前无人零售,无人车驾驶,智能制造以及文本语义分割,自然语言处理仍处于一个迅速发展的态势,在提高硬件CPU,视觉传感器的基础上,还需要优化基础算法,使得基础算法与硬件的功能得到极大的提高,同时相互促进,从根本上提升计算机视觉技术的巨大飞升,在该领域中达到国际先进水平。从而实现为实现上述的技术迈出重要的一步,最终实现计算机领域的人工智能的快速发展。

这些创新新的发展都将打破固有僵化 的生活秩序,从而为社会发展和科技进 步注入新的活力,推动生产力的快速发 展从而在实际意义方面深刻地改变人们 的生活方式,最终提高人们的生活质量。

计算机视觉研究正出于一个向上飞速 发展的大阶段中,出于目前阶段,我国相关行业的从业人员应该密切关注国际 发展态势,稳步提高其理论水平的同时,积极提高计算机视觉技术在现实层面的 应用,包括但不限于工农业生产,日常社会生活,军事交通等各个领域,从而达到将计算机视觉技术的优越性充分利用的效果,将人力无法简单实现的无意义的机械劳动快速实现,有效对数据的 质量进行甄别与分类,致力与实现提供少量样本即可快速实现进行完备的信息 标注和信息处理的模式。不难看出,计算机视觉的发展还有着极为广阔的领域。

#### [参考文献]

[1]Janusz S , Pawel H . Workspace supervising system for material handling devices with machine vision assistance [J] . Journal of KONBiN , 2009 (11 /12):7- 16.

[2]JOHNSON J , YEUNG S , LI F F. CS231n: convolutional neural network for visual recognition [EB/OL] . (2019)

- 02- 04) [2019- 11- 18] . http://vision.stanford.edu/teaching/cs231n/slides/2019/cs231n\_2019\_lecture01.pdf. [3]余京蕾.浅谈计算机视觉技术进展及其新 兴 应 用 [J]. 北 京 联 合 大 学 学报,2020,34(01):63-69.

[4] ANDERSON P, HE X, BUEHLE R C, et al. Bottom- up and top-down attention for image captioning and visual question

answering [C] / 2018 IEEE Computer Society Conference on Computer Vision and Pattern R ecognition (CVP R 2018), 19-

21st June, 2018, Salt Lake City, Utah, USA. IEEE, 2018: 6077- 6086.

[5]YAO T, PAN Y, LI Y, et al. Exploring visual relationship for image captioning [C] // 2018 European Conference on Computer Vision(ECCV 2018), 8- 14 th September, 2018, Munich, Germany. 2018: 711- 727.

[6]魏忠钰,范智昊,王瑞泽,承怡菁,赵王榕,黄萱菁.从视觉到文本:图像描述生成的研究进展综述[J].中文信息学报,2020,34(07):19-29.

[7]章炜·机器视觉技术发展及其工业应用 []].红外,2006(02):11-17.

[8]何新宇,赵时璐,张震,赵天宇.机器视觉的研究及应用发展趋势[].机械设计与制

造,2020(10):281-283+287.

[9]刘焕军,王耀南,段峰.机器视觉中的图像 采集技术 [J]. 电脑与信息技术,2003(i):15 — 21.

[10]延渊渊, 范文慧, 冯袁.汽车连杆生产 线的建模与仿真 [J] .系统仿真学 报, 2018, 30 (9): 3360-3365.

(Yan Yuan-yuan, Fan Wen-hui, Feng Yuan.Modeling and simulation of connecting rod production line [J] Journal of System Simulation, 2018,

30 (9) : 3360-3365.)

[11]柴富荣, 王红军.基于 Witness 的凸轮轴生产线生产节拍的仿真与优 化 [J].制造业自动化, 2014, 36 (9): 39-42.

( Chai Fu-rong , Wang Hong-jun.Simulation and optimization of production rhythm of camshaft production line based on witness [ J ] .Manufacturing Automation , 2014, 36 (9) : 39-42.)

[12]曹国斌,刘雪娇,王华. 图像和机器视觉技术概述 [J]. 电子工业专用设备,2008(8):27 - 31.
[13]B R OCK A, DONAHUE J, SIMONYAN K. Large scale GAN training for high fidelity natural image synthesis [Z/OL]. (2019

- 02- 25) [2019- 11- 18] . https://arxiv.org/abs/1809.11096.

[14]HE Kaiming, ZHANG Xiangyu, R
EN Shaoqing, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNetclassification [C] //Proceedings of 2015

IEEE International Conference on Computer Vision . Santiago , Chile , 2015: 1026- 1034 .

[15]王信,汪友生.基于深度学习与传统机器学习的人脸表情识别综述[J.应用科技,2018,45(01):65-72.

[16]EKMAN P, FR IESEN W V, HAGERJ C. Facial action coding system(FACS)[M]

. Manual, Salt Lake City (USA): A Human Face, 2002.

[17]罗尤春, 孙容磊. 机器视觉理论及应用 [J]. 软件导刊, 2014, 13