

DOI: 10.3969/j.issn.1006-4729.2016.03.016

计算机视觉原理分析及其应用

王道累, 陈 军, 吴懋亮

(上海电力学院 能源与机械工程学院, 上海 200090)

摘要:介绍了计算机视觉的原理,分析解释了实现原理的步骤以及发展现状,并且举例说明了计算机视觉的应用,解决了在人脸检测和识别、机器人目标定位、导航等领域的应用问题,并简单论述了计算机视觉技术的发展方向.

关键词:计算机视觉; 机器人导航; 物体识别; 立体匹配

中图分类号:TP391

文献标志码:A

文章编号:1006-4729(2016)03-0283-05

Analysis of the Principle of Computer Vision and Its Application

WANG Daolei, CHEN Jun, WU Maoliang

(School of Energy and Mechanical Engineering, Shanghai University of Electric Power, Shanghai 200090, China)

Abstract: The principle of computer vision is introduced, the procedures of carrying out the principle and the development status quo are analytically explained and the application of computer vision technology is exemplified. The problems of human face checking and recognition, the application of object location by robot in navigation are solved. The developing trend of computer vision technology is expounded.

Key words: computer vision; robot navigation; object recognition; stereo matching

计算机视觉是新兴并且迅速发展的一门学科. 计算机视觉是所有从二维图片中获得情景信息的计算机处理方法总称. 由于在工业以及军事实践中,提出了很多新课题,特别是在研制智能机器人、高尖端的武器方面,计算机视觉逐渐受到了人们的关注. 从 20 世纪 70 年代起,科技人员在研究基本理论的同时,还注重研制实用系统.^[1-2]如今计算机视觉理论已广泛应用于神经生物学、人工智能、生物医学、航空航天、模式识别与图像处理等多个领域. 同时,它也是一门由多种学科相互

交叉形成的边缘学科,其研究成果已应用到遍及科学研究、国民经济以及军事部门等各个领域.

1 计算机视觉概述及其基本结构

人类可以通过自身的双眼感知系统,轻松获得周边的三维场景. 比如我们欣赏一盆花时,可以通过叶子的颜色变化,准确预测出这朵花的生长情况;观赏一幅肖像画时,也可以轻松识别出其中的人物,甚至可以从图画呈现出来的面部表情估计出其情感活动变换等. 因为人类视觉系统具有

收稿日期:2015-03-23

通讯作者简介:王道累(1981-),男,博士,讲师,上海人. 主要研究方向为计算机视觉,图像处理, CAD/CAM.

E-mail: alfrdwdl@shiep.edu.cn.

基金项目:国家自然科学基金(61502297);上海高校青年教师培养资助计划项目(ZZsdl15074).

独特功能,可以感知现实三维情景,这促使研究者试图通过传感器和计算机的软硬件去模拟仿真人类视觉系统,再现真实三维场景,比如对三维环境图像的采集、分析、处理和 Learning 能力,并将该能力植入到计算机中,以便让计算机和机器人系统具有智能化的视觉功能。^[3-7]

计算机视觉系统主要由图像采集、图像处理及分析、图像显示输出等组成。因此,计算机视觉理论实现结构可以被划分成图像数据处理层、图像特征描述层及图像知识获取层,具体如图 1 所示。这 3 个层次形成了目前正普遍使用的计算机视觉系统框架。^[8]

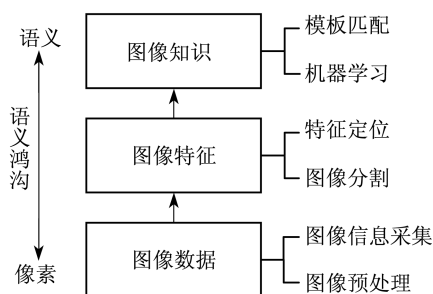


图 1 计算机视觉系统构架

1.1 图像数据处理层

在图像数据处理层中,对要处理的对象即一些像素级的数字信号进行处理与操作,如图像获取、传输、压缩、降噪、转换、存贮、增强和复原等。该层作用是将原始图像转变成为具有所需的某些特性的图像,比如较好的信噪比。它只是图对图的变更,没有一些明显的构造描述。但它又是边界检测的基础。这门技术较成熟、历史长,经常使用的方法有数字滤波以及快速富里埃变换等。^[9]

(1) 增强图像 其目标是要改善图像的视觉效果,将需要图像的整体或感兴趣的局部特性强化出来,将不清晰的图像变得清晰,扩大不同物体特征之间的差别,同时抑制不需要的特征,从而使图像质量得到改善,信息量得到丰富,图像判读与识别效果得到加强,满足特殊分析的需要。

(2) 平滑图像 其目的是使图像的宽大区域、主干部分、低频成分或干扰高频成分和抑制图像噪声被突出,这样图像的亮度趋于平缓并渐变,从而减小突变梯度程度,该处理方法能进一步改善图像的质量。

(3) 图像数据编码和传输 图像编码是以较

少的数据量无损或无损地表示原来像素矩阵的技术。数字图像的数据量巨大,如像素级的数字图像,其每个像素为 256 k 字节,如果直接进行传输,非常耗时。因此,要对数字图像数据进行变换、编码和压缩,便于图像的存储以及传输。

1.2 图像特征描述层

(1) 边缘锐化 目的是使图像的轮廓线、边缘以及图像的细节变得清晰,而经过平滑的图像变得模糊的根本原因是图像受到了平均或积分运算,因此可以对其进行逆运算,就可以让图像变得更加清晰。它是早期视觉理论及算法中的基本问题之一,也是中后期视觉算法成败的重要因素之一。

(2) 图像分割 是将图像分成若干个、特定的、具有独特性质的区域,依据灰度值、空间特性、颜色、纹理特性和频谱特性等提取出感兴趣目标的技术和过程。现有的图像分割方法被划分为基于区域的分割方法、基于阈值的分割方法、基于特定理论的分割方法和基于边缘的分割方法等。1998 年以来,国内外学者不断改进原有分割方法,结合其他学科的一些新理论和新方法,提出了很多新的分割方法。已被标示或提取的目标图像区域可被用于医学图片病症确认、图像搜索、图像语义识别等领域。

1.3 图像知识获取层

图像识别是指利用计算机对图像进行处理、分析和理解,以识别各种不同模式的目标和对象的技术。这也是计算机视觉系统必须完成的任务之一。图像识别主要包括图像匹配和机器学习。

国内外有大量研究者对图像匹配工作展开研究,^[10-11]并且取得了较好的成果。图像匹配的研究大致集中在了 3 个方面(即 3 要素):特征空间;相似性度量;搜索策略。图 2 中列举了图像匹配 3 要素以及具体内容。

1996 年,LANGLEY 给出了机器学习的定义,即机器学习是一门人工智能的科学,该领域的主要研究对象是人工智能,特别是如何在经验学习中改善具体算法的性能。机器学习是人工智能的核心技术之一,也是实现计算机具备智能方法的根本途径,在人工智能的各个领域中得到了普遍的应用。

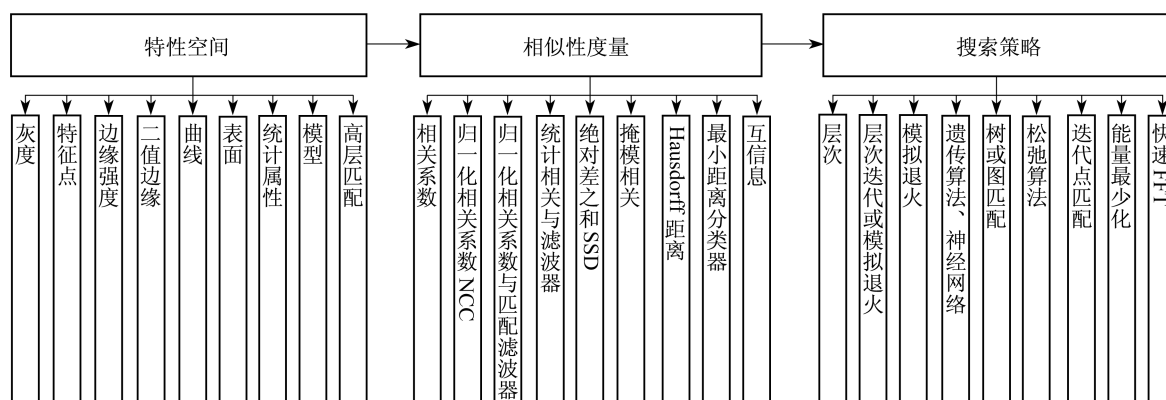


图2 匹配3要素

2 计算机视觉的应用

2.1 计算机视觉的人脸检测与识别

2.1.1 发展和研究现状

19世纪末法国人 GALTON Sir Franis 就开始研究关于人脸识别课题,直到20世纪90年代,人脸检测与识别才开始作为一个独立的学科发展起来.如今东方人脸的图像数据库也已在我国建成,这也是世界上较全面、大规模的数据库.

人脸识别的研究发展过程主要分为3个阶段.^[12]

第一阶段 主要研究人脸识别中所需要的面部特征,主要的识别过程全部依赖于操作人员,是以 ALLEN 和 PARKE 为代表.

第二阶段 人机交互式的识别阶段,用多维特征矢量来表示人脸面部的特征,是以 HARMON 和 LESK 为代表.然而以 KAYA 和 KOBAYASHI 为代表的统计识别,将欧氏距离用于表示人脸的特征.

第三阶段 机器自动识别的阶段,人脸识别技术进入了实用化的阶段,例如 Eyematic 公司研发的人脸识别系统,清华大学“十五”攻关的项目“人脸识别系统”也通过了由公安部主持的专家鉴定.

2.1.2 人脸检测与识别算法

人脸检测与识别系统是通过计算机的“眼睛”(如摄像机、数码相机等)观察“影像”(人脸),从影像中提取有效特征来鉴别身份信息的能力.人脸检测与识别可分为人脸检测、特征提取和识别3个部分.该算法整个流程如图3所示.首

先采集图像,接着检测判断人脸,即对像集中的图像逐幅进行检测,如果人脸存在,则对其进行精确定位,同时通过特征提取进行人脸识别进而获得人脸信息,最后鉴别身份、验证结果.

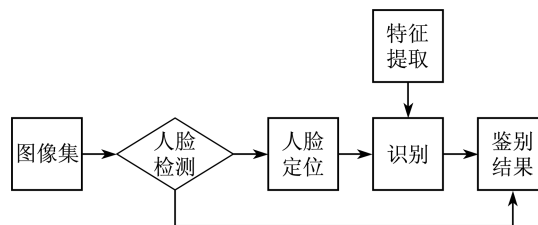


图3 人脸识别流程

2.2 计算机视觉在机器人目标定位中的应用

2.2.1 基于视觉的自主导航定位系统

机器人导航技术是智能机器人领域的一项关键技术,同时也是智能机器人的一个重要的研究热点.根据工作环境的不同要求,可以制定移动机器人导航定位系统的不同方法,比如采用双目立体视觉系统以及三角测量的原理来测量机器人在场景中移动的位置情况.国内外有大量的学者多年专门研究这方面的问题,因此在视觉导航和机器人定位等方面取得了很大的进步.而且在工业领域应用中,移动机器人导航技术也已得到了广泛的应用.^[13-15]

立体视觉系统的视差功能可实现对目标的三维定位,因此可以采用简单的固定式的双目立体视觉系统.该系统采用两个固定的摄像机来实现这种视差,简单且易操作.两个固定的相机就像人的两只眼睛.通过这对相机采集图像对来恢复三维数据点云,进而确定三维目标位置.这种设备要

求具有较高的精度,成本较高.另一方面,目标定位可视区域大小难以把握,并且需要复杂计算量来对两个摄像机进行标定,所以误差较大.而通过将移动末端执行器安置在不同位置来实现视差的手眼式立体视觉系统,只需要一个 CCD 摄像机就可以实现.

2.2.2 基于视觉的手眼目标定位系统

自动化装配领域以及航空航天领域中广泛地应用于手眼系统,该系统也促进了现代工业的飞速发展.在机器人进行装配、搬运等工作中利用视觉系统,识别需要装配的零部件并确定其安装方位,进而引导机器手臂抓取所需的零件,并能准确地放到指定的位置,因此能帮助完成工业生产中分类、搬运、装配等任务.

2.3 基于计算机视觉的机器人导航

机器人导航技术有多种,如基于地图的机器人导航、基于光流的机器人导航、基于地貌的机器人导航等.^[16]

2.3.1 基于地图的机器人导航

基于预定义的地图导航,分为绝对定位和增量定位.首先,通过在摄像机采集图像中获取图像中的不同特征,对这些特征建立相关联关系,同时建立三维坐标系,这一过程都在远程控制下进行.接着,在运动过程中不断地在网格中循环标记跟踪得到的特征.最后,将活动的环境网格化到地图中.

2.3.2 基于光流的机器人导航

Santos-Victor 等人研发出了一种基于光流的视觉系统 Robee,该系统能模拟出蜜蜂的视觉行为及运动规则.该系统认为昆虫的眼睛长在两侧的优势是基于运动产生的特征来导航蜜蜂行为,而不是获得深度信息.

2.3.3 基于地貌的机器人导航

室外环境导航大多数采用基于地貌的机器人导航,这类导航技术的核心问题为关于数字图像的模式识别,具体地说就是物体纹理、颜色的识别问题.但由于环境色以及光照的影响,在不同的环境下,物体具有相同本质色能呈现出来的完全不同的颜色.由于地貌导航很难预知先验知识,而只能实时处理视野中的对象,无法建立一幅关于周围环境的完整地图.

3 计算机视觉的发展方向

近年来,国内外在机器视觉技术领域进行了积极大胆的思索和研究,如美国卡内基-梅隆大学机器人技术研究所视觉与自主系统中心建立了一套由 49 个经过同步的 CCD 摄像机组成的“3D ROOM”系统,^[17]主要用来对实时变化的动态场景及事件进行三维建模;美国马里兰大学自动控制研究中心的 Keck 实验室使用一套由 64 个同步摄像机组成的视觉运动分析系统,对人体在三维空间中的运动进行捕捉、分析和建模;^[18]美国斯坦福大学计算机图形学实验室设计并实现了一套由 128 个经过同步的 CMOS 摄像机组成的“Light Field”多摄像机阵列,用于对高性能成像技术、高速摄像技术以及被遮挡表面的重建技术进行研究;^[19]美国明德学院(Middlebury College)提供了一套多视点三维重建算法的标准评估平台,^[20]可用于对多视点三维重建算法的精度和完整性提供定量评估,当前已有超过 40 种多视点三维重建算法的精度评估结果及排名;香港科技大学、中国科学院自动化研究所、北京大学三维视觉计算与机器人实验室等诸多研究机构也都在这个领域展开了大量的研究工作.^[21-23]另外,对于基于单棱镜立体视觉系统^[24-26]及多视图立体视觉重构^[27-33]等都有深入的研究.同时,这项技术逐步应用于工业现场,这些应用大多集中在药品检测分装、印刷色彩检测、制药印刷以及矿泉水瓶盖检测等领域.

4 结 语

虽然计算机视觉技术这门学科刚刚兴起,技术还不够成熟,但其应用前景广阔,相信在不久的将来,未来计算机视觉的应用将更加深入到人类现代生活的每个方面.

参考文献:

- [1] FAUGERAS O, LONG Q. The geometry of multiple images [M]. Cambridge, MA, USA: MIT Press, 2001: 50-70.
- [2] HARTLEY R, ZISSERMAN A. Multiple view geometry in computer vision second edition [M]. Cambridge, UK: Cambridge University Press, 2003: 40-110.
- [3] WANG Z X, WU Z Q, ZHEN X J, et al. A two-step calibration method of a large FOV binocular stereovision sensor for onsite measurement [J]. Measurement, 2015, 62(3): 15-24.
- [4] HUANG Z R, XI J T, YU Y G. Accurate projector calibration based on a new point-to-point mapping relationship between

- the camera and projector images[J]. *Applied optics*, 2015, 54(3):347-356.
- [5] SINHA S, POLLEFEYS M, MCMILLAN L. Camera network calibration from dynamic silhouettes [C] // *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004:195-202.
- [6] GOESELE M, CURLESS B, SEITZ S M. Multi-view stereo revisited [C] // *IEEE CVPR*. New York: IEEE, 2006: 2402-2409.
- [7] CAMPBELL N D, VOGIATZIS G, HERNANDEZ C, *et al.* Using multiple hypotheses to improve depth-maps for multi-view stereo [C] // *ECCV*. Berlin: Springer-Verlag, 2008: 766-779.
- [8] 许志杰,王晶,刘颖,等. 计算机视觉核心技术现状与展望[J]. *西安邮电学院学报*, 2012, 17(6):1-8.
- [9] 蔡愉祖. 计算机视觉概述[J]. *湖北轻工业大学学报*, 2013(7):12-17.
- [10] 熊凌. 计算机视觉中的图像匹配综述[J]. *湖北工业大学学报*, 2006, 21(3):171-173.
- [11] WANG Dao Lei, LIM Kah Bin. Obtaining depth map from segment-based stereo matching using graph cuts [J]. *Journal of Visual Communication and Image Representation*, 2011, 22(4):325-331.
- [12] 谢丽欣,牟会,王欢,等. 基于计算机视觉的人脸检测与识别综述[J]. *计算机安全*, 2010(1):60-63.
- [13] 王俊修,孔斌. 计算机视觉在机器人目标定位中的应用[J]. *微机发展*, 2003(12):7-10.
- [14] 吴福朝,李华,胡占义. 基于主动视觉的摄像机自标定方法研究[J]. *自动化学报*, 2011, 27(6):736-746.
- [15] 雷成,吴福朝,胡占义. 一种新的基于主动视觉系统的摄像机自标定方法[J]. *计算机学报*, 2010, 23(11):1130-1139.
- [16] 吴晓明. 基于计算机视觉的机器人导航综述[J]. *实验科学与技术*, 2007, 5(5):25-28.
- [17] KANADE T, SAITO H, VEDULA S. The 3D room: digitizing time-varying 3D events by synchronized multiple video streams[R]. Pittsburgh, Pennsylvania: Robotics Institute Carnegie Mellon University, 1998.
- [18] SUNDARESAN A, CHOWDHURY A Roy, CHELLAPPA R. 3D modelling of human motion using kinematic chains and multiple cameras for tracking [C]. *International Symposium on the 3D Analysis of Human Movement*. Tampa, USA, 2008:781-785.
- [19] WILBURN B, SMULSKI M, LEE H K, *et al.* The light field video camera [C] // *Proceedings of SPIE Media Processors*, 2012:29-36.
- [20] Middlebury, Inc. [EB/OL]. [2015-07-28]. <http://vision.middlebury.edu/mview/>.
- [21] LHUILLIER M, QUAN L. Surface reconstruction by integrating 3D and 2D data of multiple Views [C] // *IEEE Ninth International Conference on Computer Vision (ICCV)*, 2003: 1313-1320.
- [22] WU F C, DUAN F Q, HU Z Y. An affine invariant of parallelograms and its application to camera calibration and 3D reconstruction [C] // *The 9th European Conference on Computer Vision (ECCV)*. Australia, 2011: 1103-1107.
- [23] LI X, ZHA H B. Realistic human head modeling with multi-view hairstyle reconstruction [C] // *The 5th International Conference on 3D Imaging and Modeling (3DIM)*. Ottawa, Ontario, Canada, 2005:432-438.
- [24] WANG Daolei, LIM Kah Bin. Geometrical approach for rectification on single-lens stereovision using a triprism [J]. *Machine Vision and Applications*, 2013, 24(4):821-833.
- [25] LIM Kah Bin, WANG Daolei, KEE Wei loon. Virtual cameras rectification with geometrical approach on single-lens stereovision using a biprism [J]. *Journal of Electronic Imaging*, 2012, 21(2):023003(1-9).
- [26] LIM Kah Bin, KEE Wei Loon, WANG Daolei. Virtual camera calibration and stereo correspondence of single-lens biprism stereovision system using geometrical approach [J]. *Signal Processing-image Communication*, 2013, 28(9):1059-1071.
- [27] CIPOLLA R, WONG K K. Reconstruction of sculpture from its profiles with unknown camera positions [J]. *IEEE Transactions on Image Processing (IP)*, 2004(13):381-389.
- [28] ZENG G, PARIS S, QUAN L, *et al.* Progressive surface reconstruction from images using a local prior [C] // *International Conference on Computer Vision (ICCV)*, 2005:1230-1237.
- [29] SORMANN M, ZACH C, BAUER J, *et al.* Watertight multi-view reconstruction based on volumetric graph-cuts [C] // *15th Scandinavian Conference on Image Analysis*, 2007:393-402.
- [30] ZAHARESCU A, BOYER E, HORAUD R. Transformesh: a topology-adaptive mesh deformation for surface evolution, morphing, and multi-view reconstruction [J]. *Pattern Analysis and Machine Intelligence, IEEE*, 2011, 33(4):823-837.
- [31] BRADLEY D, BOUBEKEUR T, HEIDRICH W. Accurate multi-view reconstruction using robust binocular stereo and surface meshing [C] // *IEEE CVPR*, Anchorage, 2008:1-8.
- [32] WU Taipang, YEUNG S K, JIA Jiaya, *et al.* Quasi-dense 3D reconstruction using tensor-based multi-view stereo [C] // *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, 2010:1482-1489.

(编辑 桂金星)