

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 **2**
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年2月 第17卷 第2期(总第190期)

目次

综述

内容感知图像缩放技术综述 施美玲,徐丹(157)

图像处理和编码

引入纹理相似性的纺织品图像增强 杨学志,田晓梅,方静,卢洁(169)

保持边缘特征和增强对比度的图像缩放算法 计忠平,方美娥,王毅刚,吴向阳(178)

以图像认证为目的的分形编码数字水印算法 黄晓晴,于盛林(183)

结合运动矢量和像素递归的全局运动估计方法 赵亚湘,刘少强,樊晓平(191)

无缝栅格数据小波金字塔构建 马伯宁,冷志光,汤晓安,匡纲要(197)

图像分析和识别

伪彩色空间完全非结构化道路检测方法 胡晓辉,孙苗强,苏晓许(203)

未知环境下机器人障碍物检测技术 王文格,武凯宾,朱江,段文彬,许重阳(209)

局部 GAC 模型在医学图像分割中的应用 张建伟,方林,陈允杰,詹天明,罗春燕(215)

判别割(Dcut)的图像分割及其快速分割算法 邹小林,陈伟福,冯国灿(222)

改进的遮挡条件下瞳孔检测方法 潘林,魏丽芳,郑炳锜,余轮(229)

木材显微图像的局部水平集分割方法 汪杭军,祁亨年(236)

关键点检测的线要素综合算法 黄志坚,张金芳,徐帆江(241)

GPU 加速的高分辨率 DEM 图像地形特征线提取算法 刘洲俊,胡包钢(249)

改进 GVF 的自动 Snakes 模型 周亚男,程熙,骆剑承,沈占锋,胡晓东(256)

曲波变换用于磨粒图像不变矩的提取 张云强,张培林,任国全,王国德,徐超,李兵(263)

图像理解和计算机视觉

深度图像中基于轮廓曲线和局部区域特征的3维物体识别 吴莉婷,张宇,杨一平,田原(269)

机载 LiDAR 数据的 LZD 航带平差 王丽英,宋伟东(279)

计算机图形学

基于 Z 曲线的瓦片地图服务空间索引 聂云峰,周文生,舒坚,许虎(286)

虚拟现实和增强现实

GPU 加速的台风可视化方法 秦绪佳,张勤锋,陈坚,郑红波,徐晓刚(293)

中国图象图形学报

刊名题字: 宋 健 月刊(1996 年创刊)

第 17 卷 第 2 期

2012 年 2 月 16 日出版

主管单位 中国科学院
主 办 中国科学院遥感应用研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 李小文

编辑出版 《中国图象图形学报》编辑出版委员会

北京 9718 信箱 邮编 100101

电子信箱: jig@irsa.ac.cn

电话: 010-68407995 010-82614429

网 址: www.cjig.cn

印刷装订 北京北林印刷厂

广告经营许可证 京朝工商广字第 0346 号

总 发 行 北京报刊发行局

订 购 全国各地邮局

国外发行 中国国际图书贸易总公司

(中国国际书店)

(北京 399 信箱 邮编 100044)

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing Application,
CAS China Society of Image and Graphics
Institute of Applied Physics and Computational
Mathematics
Chief editor LI Xiaowen
Editor, Publisher Editorial and Publishing Board
of Journal of Image and Graphics
(P. O. Box 9718, Beijing 100101, China)
E-mail: jig@irsa.ac.cn
Distributed by Beijing Bureau for Distribution of Newspapers
and Journals
Domestic All Local Post Offices in China
Foreign China International Book Trading Corporation
(P. O. Box 399, Beijing 100044, China)
Printed by Beijing Beilin Printing House

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 2 February 2012

Contents

Review

Survey on content-aware image resizing techniques Shi Meiling, Xu Dan (157)

Image Processing and Coding

Enhancement of textile image with texture similarity Yang Xuezhi, Tian Xiaomei, Fang Jing, Lu Jie (169)

Edge-preserving and contrast-enhancement image scaling Ji Zhongping, Fang Mei'e, Wang Yigang, Wu Xiangyang (178)

Watermarking method on fractal coding for image authentication Huang Xiaoqing, Yu Shenglin (183)

Global motion estimation method with motion vectors and pixel recursion Zhao Yaxiang, Liu Shaoqiang, Fan Xiaoping (191)

Seamless wavelet pyramid construction for raster data Ma Boning, Leng Zhiguang, Tang Xiao'an, Kuang Gangyao (197)

Image Analysis and Recognition

Completely unstructured road detection for pseudo-color space Hu Xiaohui, Sun Miaoqiang, Su Xiaoxu (203)

Obstacle detection for robot in unknown environment
..... Wang Wenge, Wu Kaibin, Zhu Jiang, Duan Wenbin, Xu Chongyang (209)

Application of local GAC model for medical image segmentation
..... Zhang Jianwei, Fang Lin, Chen Yunjie, Zhan Tianming, Luo Chunyan (215)

Fast image segmentations of Deut Zou Xiaolin, Chen Weifu, Feng Guocan (222)

Improved method for the pupil measurement under occlusion Pan Lin, Wei Lifang, Zheng Bingkun, Yu Lun (229)

Segmentation method of wood microscopic image based on local level set Wang Hangjun, Qi Hengnian (236)

Algorithm of line generalization with key point detection Huang Zhijian, Zhang Jinfang, Xu Fanjiang (241)

High resolution DEM topographic feature line extraction algorithm using GPU Liu Zhoujun, Hu Baogang (249)

Automatic snakes model based on modified GVF Zhou Yanan, Cheng Xi, Luo Jiancheng, Shen Zhanfeng, Hu Xiaodong (256)

Invariant moment extraction by curvelet transform for wear particle images
..... Zhang Yunqiang, Zhang Peilin, Ren Guoquan, Wang Guode, Xu Chao, Li Bing (263)

Image Understanding and Computer Vision

Three-dimensional free-form object recognition based on contour curve and local surface patches in range images
..... Wu Liting, Zhang Yu, Yang Yiping, Tian Yuan (269)

Least Z-difference strip adjustment of airborne LiDAR data Wang Liying, Song Weidong (279)

Computer Graphics

Spatial index for tile map service based on Z curve Nie Yunfeng, Zhou Wensheng, Shu Jian, Xu Hu (286)

Virtual Reality and Augmented Reality

GPU accelerated typhoon visualization method Qin Xujia, Zhang Qinfeng, Chen Jian, Zheng Hongbo, Xu Xiaogang (293)

中图法分类号: TP391.4 文献标志码: A 文章编号: 1006-8961(2012)02-0269-10

论文引用格式: 吴莉婷, 张宇, 杨一平, 田原. 深度图像中基于轮廓曲线和局部区域特征的3维物体识别[J]. 中国图象图形学报, 2012, 17(2): 269-278

深度图像中基于轮廓曲线和局部区域特征的3维物体识别

吴莉婷, 张宇, 杨一平, 田原

中国科学院自动化研究所, 北京 100091

摘要: 为了提高3维物体目标识别系统的性能及降低计算复杂度, 提出一种由粗到细的识别方法。该方法利用深度图像所提供的信息, 分两步完成识别过程。首先基于轮廓曲线计算其特征点, 并映射到原有轮廓空间, 以标志点序列表征原由轮廓进行匹配, 在识别初期迅速排除模型库中不相似目标和差异较大的姿态, 生成目标候选列表用于精确匹配, 以提高识别效率。精确匹配采用一种基于局部区域特征的识别方法, 以投票的策略获取最佳结果。局部区域由SIFT算子确定位置和数量, 区域特征主要由表面指数和法向量夹角组成, 具有平移和旋转不变性。为了更进一步提高效率和降低存储空间, 模型库的数据分为轮廓和表面信息两部分, 分别以标志位序列和哈希表的形式存储。实验结果表明, 该方法具有良好的实时性和识别率, 对遮挡和干扰有一定的适应性。

关键词: 3D目标识别; 深度图像; 轮廓曲线; SIFT; 局部区域特征

Three-dimensional free-form object recognition based on contour curve and local surface patches in range images

Wu Liting, Zhang Yu, Yang Yiping, Tian Yuan

Institute of Automation, Chinese Academy of Science, Beijing 100091, China

Abstract: In this paper we propose a two-step approach to recognize free-form objects in range images. First, feature points are calculated based on the contour curve, then mapped to the original shape space. Then a landmark list is determined and used to form a rejection classifier, which quickly eliminates a large number of candidates for an efficient recognition. The remaining free-form objects are then verified using a novel local patch-based matching approach, which is robust to occlusions and noise. The key points are determined based on the scale invariant feature transform (SIFT), and the local surface descriptor is characterized by its two 1D histograms. The two histograms show the frequency of occurrence of the shape index values and the angles between the normal of a key point and that of its neighbors. In order to speed up the retrieval of surface descriptors and save the restore space, the modal data includes shape and surface information. The local surface patched of modal are indexed into a hash table. Verification is performed by running the Iterative Closest Point algorithm. Experimental results with ideal range image are presented to demonstrate the effectiveness, efficiency of the proposed approach. The approach is robust to occlusions and noise.

Key words: 3D object recognition; range images; contour curve; SIFT; local surface patch

收稿日期: 2011-04-19; 修回日期: 2011-06-12

基金项目: 国家自然科学基金项目 (61005067)

第一作者简介: 吴莉婷 (1982—), 女, 中国科学院自动化研究所计算机应用专业博士研究生, 主要研究方向为模式识别与人工智能。

E-mail: iting_zw@gmail.com

0 引言

计算机视觉 (computer vision)、人工智能 (artificial intelligence)、虚拟现实 (virtual reality) 等前沿领域的一个热点和难点^[1]是 3 维物体表面的描述和识别。

识别分为两类:一般物体识别与特定物体识别,它们主要区别在于,后者通过构造高度特化的特征提取,使用海量训练样本进行训练,仅仅处理某种物体,典型如汽车检测及人脸检测^[2]。而一般物体识别则要困难得多,必须使用物体类间通用的一般特性。

在 3 维物体识别中有如下两种最经常使用的传感器:一是灰度传感器(intensity sensor),其获取图像的每个像素点对应一个亮度测量;另一类是深度传感器(range sensor),其像素点对应于传感器到可视的物体表面的距离^[3]。由于深度图像不受光照强度和物体表面特性的影响,直接反映物体的深度信息,能够利用物体表面的结构信息进行识别,因而得到了越来越多的应用。

利用深度图像进行目标识别,已经有许多经典的方法,例如特征图(eigen pictures)^[4]和特征空间(eigen shapes)^[5],但都是属于全局特征,对遮挡非常敏感,而且依赖于初始的分割结果。在识别领域,常用曲面上某点领域内曲面的几何信息或几何统计信息描述该点局部特征。且该局部区域特征一般具有刚体变换恒定的性质,抗遮挡和噪声干扰。常用的局部特征有 spin image^[6], point signature^[7],这些特征的匹配精度高,但是计算复杂。较为简单的局部特征向量有颜色直方图、表面指数(shape index)、曲率等,但是匹配效率依然不高。

针对上述问题,提出一种基于轮廓曲线和局部区域表面特征的 3 维物体识别方法。该方法针对一般形状(free-form)物体,首先以深度图像提供的 2 维信息为基础,提取物体的外轮廓,以标志点序列来简化表示并匹配,作为预处理模块,将相似度较低的部分模型库排除;其次,在后续的匹配中,利用深度图像所包含的物体 3 维结构信息,通过 SIFT 算子获取物体的关键点,并提取关键点的局部区域特征,进行精确匹配,最后需要采用迭代最近点(ICP)算法计算精确位姿。实验证明,本文算法在计算复杂度上有较好性能,对一般物体有良好的适应性,而且对

目标的部分遮挡和噪声干扰有一定的鲁棒性。

1 算法简介

基于轮廓曲线和局部区域特征的 3 维物体识别方法系统模块如图 1 所示。首先获取物体的外轮廓,将其转换为轮廓曲线,提取曲线上的曲率最大点作为特征,投影到原有轮廓空间生成标志点,以标志点序列的形式表征轮廓,与模型库的轮廓信息数据进行相似度匹配,通过排除模型库中相似度较低类别与姿态,得到候选类别编号列表。其次,利用 SIFT 算子提取原图像中物体表面的关键点,并生成该点的局部区域特征,主要包含表面指数和法向量夹角,以直方图的形式进行相似度匹配,以投票的策略来选出目标的类别及最接近的姿态。以此为初始值,通过迭代最近点算法获取物体的精确位姿。

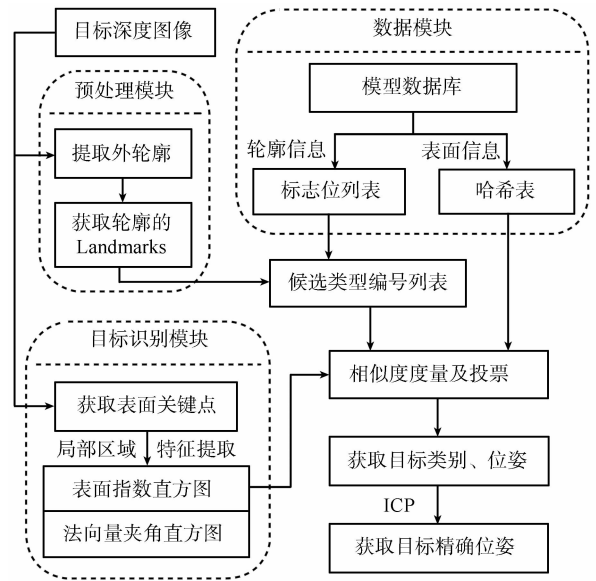


图 1 系统模块图

Fig. 1 System diagram

2 预处理模块

预处理模块的主要目的是对目标进行粗匹配,在模型库中排除与其外形差异较大的类别,获取候选类别和姿态的编号列表。同时,由于在精确匹配中选用的局部区域特征具有旋转不变性(参看目标识别模块小节),所以在匹配时,也不需要单个目标的所有姿态都搜索一遍。而预处理模块不但缩小

的类别搜索范围,也在一定程度上缩小了目标姿态的搜索范围。

预处理模块利用物体的2维信息进行匹配。其中,形状特征是图像的一个重要的视觉特征,它对位移、旋转和尺度的变换具有不变性。利用物体的轮廓特征作为匹配的基础。

为了进一步压缩目标的信息量,使得计算量小,速度快,提取目标轮廓的显著特征用于配准。点特征是在图像配准中最常用最基本的图像特征之一。

但是,直接在轮廓上提取显著特征,易受噪声的干扰,出现特征点分布过于集中的情况。因而提出一种基于轮廓曲线提取轮廓的特征点的方法。基于轮廓曲线计算得到的特征点映射到原来的轮廓空间,获取标志点序列,以此表征原有轮廓。

2.1 基于轮廓曲线的识别方法

在计算机视觉和模式识别中,形状是对目标的二值图像表示,可以看成是目标的轮廓,它是用于目标识别的重要特征^[8]。基于不同的理解,许多形状识别方法已经被提出,形状特征的描述分为2类方法:1)基于边界特征;2)基于区域特征。

直接对轮廓进行识别,存在不足之处:1)轮廓易受噪声干扰;2)对遮挡比较敏感。

标记图(以下称为轮廓曲线)是一种1维函数的边界表达方法^[9]。最简单的方法之一就是将从质心到边界线的距离转化成一个角度函数。在此基础上对标记图进行了修改,横坐标为外轮廓上的点。以下面的函数式表示:

$$r(c_i) = d(c_i, c_{center}) \quad (1)$$

式中, c_{center} 为外轮廓的质心, c_j 为轮廓上的点。 $r(c_i)$ 是单值函数,把2维封闭的轮廓转换成1维曲线,显然这种映射为一一映射,即这条曲线唯一确定了目标轮廓的特性。如图2所示。

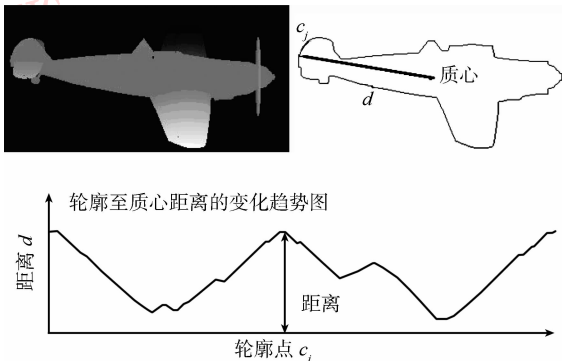


图2 轮廓曲线

Fig.2 Curve of contour

2.2 标志点序列表征轮廓

文献[10]中,提出了利用标志点(landmark)来表征轮廓曲线,用于识别,进一步降低运算效率。标志点作为轮廓曲线的重要特征,决定了图像中目标的形状。该方法利用 Legendre 多项式逼近轮廓曲线,以消除噪声的干扰,在此基础上通过提取曲线上的极值点,作为特征点,并映射到原有的轮廓,以标志点序列的形式表征,简化匹配。

预处理模块的作用在于高效率地剔除在模型库中与匹配目标外形差异较大的候选类别。并不需要精确表示目标。同时轮廓曲线上的特征点与原有轮廓上的显著特征一一对应,因而将标志点定义为轮廓曲线上曲率足够高的点。

为了减少噪声对轮廓曲线的干扰,同时不增加过多计算量,计算 landmark 前,需要对轮廓曲线进行高斯平滑。

在实际应用时,高斯平滑的窗口大小并不好确定,而且平滑窗口过大,会导致轮廓曲线过多偏离原来的形状,获取的特征点偏离原有轮廓的显著位置。同时,即使经过了平滑处理,轮廓曲线仍然可能受到噪声的干扰,导致标志点之间非常靠近。这样得到的表征有些冗余,而且对遮挡很敏感。为了得到较为合理分布的标志点列表,需要对标志点点簇进行聚类处理,将非常靠近的点去除,以点簇的聚类中心为新的标志点。

具体算法如下:

- 1) 对目标进行形态学处理,去除目标细小的部分,获取其大致的外轮廓,同时提高对噪声的抗干扰能力;
- 2) 依据标记图原理和式(1),计算外轮廓的质心边界关系,获取轮廓曲线;
- 3) 对轮廓曲线进行高斯平滑;
- 4) 计算整个轮廓曲线上各个位置的曲率;
- 5) 按照曲率大小排序,选取最高的 N 个点作为候选标志点;
- 6) 利用聚类的方法避免标志点过于集中。遍历所有候选标志点,计算相邻点的距离,小于一定阈值则合并为同一类,得到 M ($M \leq N$) 个类,即 M 个中心点,作为新的标志点;
- 7) 将标志点映射到原有轮廓的图像空间,形成标志点序列表征目标用以匹配。

具体的生成过程如图3所示。

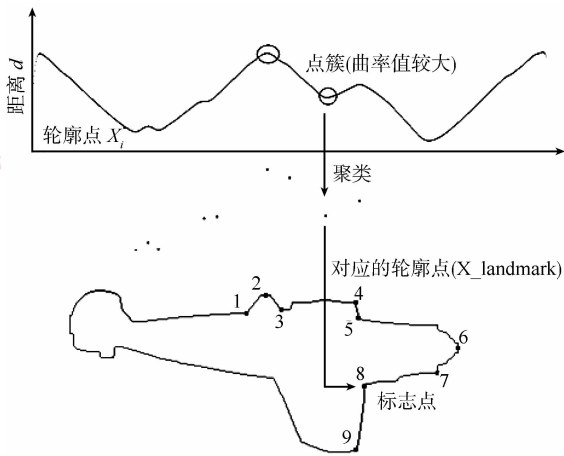


图3 获取标志点序列的过程

Fig. 3 The process for landmarks

由图3可看出,标志点序列并没有完整的表达整个目标的轮廓。实际上,该目标的标志点生成过程中,只选取了前 $N = 45$ 个轮廓曲率最高的点,最后通过聚类得到了 $M = 9$ 个聚类中心,即9个标志点形成的列表,映射到原有轮廓空间,每一格标志点存储2维空间信息 (x, y) ,以 $L\{l_1(x, y), l_2, \dots, l_m\}$ 的形式表征原有轮廓。这并不影响后面的匹配,因为该模块的主要目的只是剔除外形差异较大的候选目标,进行快速匹配,缩小精确匹配的搜索范围。若需要,可增大 N 的大小,获取较为完整的轮廓表达。

相对于直接利用轮廓进行识别,基于标志点序列的识别方法,能够降低模型库的存储空间,提高匹配效率,同时在有效性和准确度上也有一定保证。

2.3 标志点 (landmark) 序列的相似度量

标志点匹配的主要目的是建立两个标志点序列之间的对应关系。

为待匹配标志点序列的某一个标志点,寻找其在模型库中参考序列的对应标志点 (corresponding landmark) 就要依据以下原则:仅考虑2维空间的限制,在一定邻域范围能寻找最相近的标志点,形成相关点对。如果相关标志点对的数量大于一定阈值时,认为两个目标的外形相似。若不考虑遮挡的情况,标志点序列可按顺序匹配,效率能进一步提高。

待匹配标志点序列 $L\{l_1(x, y), l_2, \dots, l_m\}$ 和模型库中的参考序列 $L^c\{l_1(x, y), l_2, \dots, l_n\}$ 相似度量准则:

1) 每一个标志点 l_k 在附近领域(半径为 d_{thre}) 寻找参考序列中最近的标志点,得到相似点对如图4所示并计数;

2) 相似点对的数量为 h , 则相似度量度量为 $S_l = \frac{h}{m} > s_{l_{thre}}$ 。

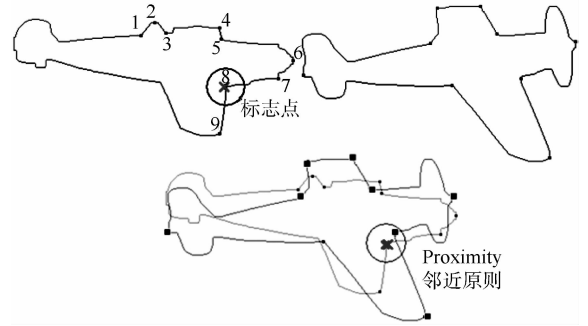


图4 标志点序列的相似度量

Fig. 4 Similarity measure of landmark sequence

事实上,这是较为宽松的相似性度量,匹配时没有对序列间的关系进行约束,但是能够满足预处理模块的需求。若需要增加匹配精度,需要增加匹配条件,相似度量不能仅仅考虑空间的限制,具体算法可参考文献[10]。

3 目标识别模块

通常,我们对目标进行识别,提取的特征大都只考虑其是否具有平移、旋转和尺度不变性。但在实际应用中,由于遮挡、图像噪声、分辨率等因素的影响,当目标出现缺损或局部畸变时,所生成的全局特征就会发生改变,影响结果。

在众多物体识别方法中,基于图像小块(patch)的方法得到研究者们的青睐。简单来说,它在图片中选取一些小块作为单位,在小块上进行一些图像的基本操作,得到描述小块的特征;然后存储这些特征以供识别时使用^[2]。基于图像小块(patch)的方法优势在于:1)它对图片的表示比较灵活,且抗各种变换能力强;2)识别前不需要做分割或者其他任何预先的对于图像的语义解释;3)不需要对数据作统计推断,速度快、稳定性好,没有陷入局部极值的危险。从使用的角度上看,小块方法实现简单方便,综合性能也最好。局部特征(图像小块描述子, local patch descriptor)是指依赖于图像的局部区域而构成的特征,当目标某一部分发生形变时,其他部分的特征仍维持不变。可应用于目标出现遮挡或缺损的识别中。

使用基于局部区域特征的识别方法,需要解

决以下3个问题:1)如何定位局部区域的位置;2)如何表征该区域;3)各局部区域间的相似度量。

基于这些问题,提出一种新的方法,利用SIFT算子提取目标中的关键点,确定图像小块的位置;由形状指数(shape index)和法向量夹角来表征该区域(小块,patch),并以此确定其大小,将其统称为局部区域特征;最后,与模型库进行精确匹配。

3.1 关键点提取

SIFT(scale invariant feature transform)算子由Lowe 1999年提出^[11]并于2004年进行完善^[12]。该算法通过抽取图像的局部不变性特征进行图像识别匹配,具有良好的尺度和旋转不变性,而且抗噪声能力强。SIFT关键是图像在不同尺度下的连续几幅高斯差分图像中非图像边缘附近的极值点。高斯差分图像中的极值点是在图像进行高斯平滑滤波后变换剧烈的像素点,因而他们具有很强的稳定性。

结合深度图像的特点,Rachel等人^[13]提出2.5D SIFT算子来定位关键点,并利用一个2维直方图对关键点附近的局部区域进行描述,完成一个对尺度和旋转不敏感的人脸识别系统。

尺度空间理论最早在计算机视觉领域提出是为了模拟图像数据的多尺度特征。Koenderink^[14]研究证明了高斯卷积核是实现尺度变换的唯一变换核。随后,Lindeberg等人^[15]进一步研究证明了高斯卷积核是唯一的线性核。

2维高斯函数定义为

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (2)$$

σ 表示高斯正态分布的方差,即尺度因子。尺度因子越小表示该图像被平滑的越少。大尺度对应图像的概貌特征,小尺度对应图像的细节特征^[14]。

一幅2维图像 $I(x, y)$ 在高斯尺度空间的表示 $L(x, y, \sigma)$ 可以由图像 $I(x, y)$ 与高斯核 $G(x, y, \sigma)$ 相卷积得到,其定义如下:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3)$$

为了在尺度空间中更有效的检测到稳定而独特的特征点,Lowe提出高斯差分尺度空间DoG(difference-of-Gaussian scale space)的概念。DoG算子是两个不同尺度的高斯差分,其计算简单。

$$G(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (4)$$

关键的提取分为两步:1)检测尺度空间的极

值;2)确定关键点的位置。具体的算法可参见文献^[12]。

3.2 局部区域特征

关键点附近的局部特征主要由两个1维的直方图来表征:1)表面指数(shape index),特征由Dorai等人提出^[16],可表示每一点的凹凸程度,能表征立体曲面的8种类型(不包括平面),与坐标系独立,且与刚体运动无关。2)法向量夹角,局部区域内各位置与关键点法向量的夹角。目标表面上各个位置的shape index是由它的最大、最小主曲率 $k_1(p)$ 和 $k_2(p)$ 得到的,其中, $k_1(p) > k_2(p)$ 。

$$S_i(p) = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \quad (5)$$

表面指数的取值范围为 $[0, 1]$,表征的曲面类型由凹面向凸面渐变。 $k_1(p) = k_2(p) = 0$ 时,表征的曲面类型为平面,以曲面类型5来表示。

在目标表面上像素 (x, y) 处的曲面法向量 (N_x, N_y, N_z) 为

$$[N_x, N_y, N_z] = \frac{[-f_x, -f_y, 1]}{\sqrt{(1 + f_x^2 + f_y^2)}} \quad (6)$$

式中, f_x 和 f_y 是在像素 (x, y) 处的一阶高斯导数。

局部区域的大小确定原则为:1)以关键点为中心的一小块领域;2)该区域里各位置的方向法向量差别不大,即与关键点位置的法向量夹角小于一定阈值;3)该区域的各像素的表面指数相近。依据以上原则,局部区域 N 的各个像素应该满足

$$N = \begin{cases} \|N - P\| \leq \varepsilon \\ \arccos(\mathbf{n}_n \cdot \mathbf{n}_p) \leq \theta \\ |S_n - S_p| \leq s \end{cases} \quad (7)$$

式中, P 为关键点, \mathbf{n}_p 为关键点处的表面法向量, \mathbf{n}_n 为局部区域的表面法向量。 ε 限制了局部区域的大小, ε 越大,该区域包含的信息越多,识别的准确率就越高,但降低了识别的效率。 θ 为各像素法向量方向与关键点法向量方向的夹角。 θ 角越大,表示区域内像素的法向量方向与关键点的差别越大。 s 为表面指数,限制了该区域的曲面类型,不希望在这个局部区域(local pathh)中,出现较多曲面类型。 ε 、 θ 和 s 这三个阈值确定了一个合理局部区域的大小,给出即包含了足够多的特征信息用于识别,又确保不会过于冗余,影响识别效率。

图5中标出了在目标中的3个不同的关键点,并由局部区域方程(7)确定相应的局部区域

(local patch) 大小,分别为不同的 3 曲面类型。其中,从区域中提取的表面指数和法向量夹角特征分别以 1 维直方图形式表征。左边为法向量夹角直方图,右边为表面指数直方图,均只显示了非零的部分。

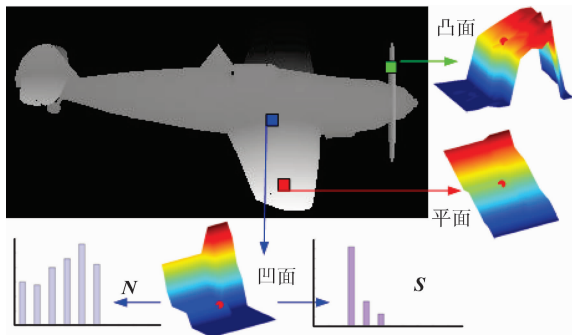


图 5 关键点的曲面类型及其直方图

Fig. 5 Surface types and histograms of concave

4 数据模块

为了降低模型库的规模,并依据识别过程中所需要的信息,将模型库的数据类型分为两类:1) 目标 2 维轮廓信息,以其标志点序列的形式存储;2) 目标 3 维表面信息,以哈希表(hash table)的形式存储。

4.1 模型库中轮廓信息的存储形式

模型库中的轮廓信息主要以其标志点序列形式存储,每个样本包括的主要信息有:1) 所属的类别、姿态号;2) 序列的个数;3) 各个标志点的 2 维位置信息。

4.2 模型库中表面信息的存储形式

哈希表^[17]主要应于查找,实现的是一种从图像空间到参数空间的映射关系。利用哈希表建立关于目标局部信息的模型库,能够简化识别算法,提高识别的速度

实际上,不同类别/姿态需要提取不同数量的局部区域,当模型库中类别数量增加时,识别时所需要处理的数据量越来越大,且往往含有大量相似或异常的信息,处理这样冗余而庞杂的数据集往往导致识别算法效率及性能低下,可扩展性差。为了解决这个问题,提出了用投票机制来获取最佳匹配的算法(参见 4.3 节),并用哈希表存储局部区域信息。一方面有效的降低数据量、提高识别算法的时空效率;另一方面,提取有效的图像小块、去除某些噪声,

提升算法性能。哈希表的结构和存储的信息如图 6 所示。

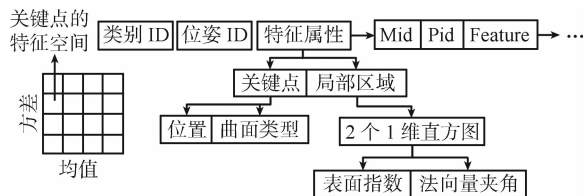


图 6 哈希表结构

Fig. 6 Structure of the hash table

哈希表存储不同候选类别中不同的局部区域,即 3 维表面信息。这些信息包括类别号、位姿号还有该区域的特征属性。其中,特征属性包含该区域的关键点和局部区域描述(特征)信息。具体如图 6 所示。

通过计算局部区域内各个位置的表面指数,获取该区域均值 μ 和方差 σ^2 作为哈希表的索引,生成一个 2 维的哈希表。

$$\mu = \frac{1}{N} \sum_{n=1}^N S_i(p_n) \quad (8)$$

$$\sigma^2 = \frac{1}{N-1} \sum_{n=1}^N (S_i(p_n) - \mu)^2 \quad (9)$$

式中, N 为该局部区域的大小, $S_i(p_n)$ 为像素 p_n 处的表面指数。

实际上,每一个索引里并不是对应着唯一的一个局部区域,而是拥有相同表面指数均值和方差的局部区域集合。这个集合的大小依赖于索引的细化程度。而哈希表的效率依赖于查找过程中所进行的比较次数。因而其索引细化得越多,每次比较的个数就越少,速度更快。但是过于细化,会导致错误匹配的概率越高,因为表面指数只是特征属性的一种,不能完全表征目标的局部区域。

4.3 表面信息的相似度度量策略

待匹配目标由 SIFT 算子确定关键点的位置和个数后,根据式(5)~(7)获取局部区域的特征属性,并基于预处理模块输出的候选目标列表,进行精确匹配。

相似度度量策略如图 7 所示。

根据表面指数的均值方差,找到模型库中表面信息哈希表的样本子集,而后进行:

1) 关键点匹配,对局部区域的空间位置进行约束。一方面,虽然选取的局部区域特征都具有旋转不变性,但在实际识别时,姿态的差别会影响区域特

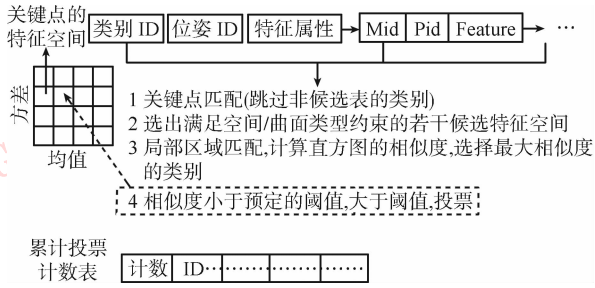


图7 相似度量策略

Fig. 7 Strategy of similarity measure

征的计算。即同一目标的同一关键点的局部区域,在不同姿态下,由于目标的自遮挡,获取的两个直方图并不是完全一致的。另一方面,对关键点位置的限制,是希望获取的识别结果不仅要保证类别正确,而且输出的姿态尽量接近。以此初始化ICP算法,计算效率更高。

2) 曲面类型匹配,对局部区域的曲面类型进行约束。例如,曲面类型有9类,1—4为凹面,5为平面,而6—9为凸面。若待匹配区域的曲面类型为1,而样本的为9,则应该跳出比较,对该集合中的下一个样本进行匹配。

3) 局部区域匹配,待匹配区域与对应哈希表的子集中每一个区域样本进行比较。通过比较两个1维直方图的相似度,主要利用直方图相交法:

$$P(Q, D) = \frac{\sum_{k=0}^{L-1} \min(H_Q(k), H_D(k))}{\sum_{k=0}^{L-1} H_Q(k)} \quad (10)$$

式中, $H_Q(k)$, $H_D(k)$ 分别为待匹配局部区域和子集中样本的表面指数(或法向量夹角)直方图。两个直方图的相似度以加权的形式输出:

$$P = \alpha P_s + \beta P_n \quad (11)$$

式中, P_s , P_n 为表面指数和法向量夹角的相似度, α 和 β 为相应得权重。该相似度需要大于一定阈值时,才能作为候选相似区域。当与该集合所有的样本进行比较后,得到候选相似区域列表。

4) 若候选相似区域列表为空,则放宽索引值,重复步骤1);否则,选取相似度最大的前 N 个样本,根据其类别和姿态ID进行投票。简单的直方图匹配还是较为粗糙的识别方法。因而,还需利用投票计算策略来获取最优结果。

5 实验结果及分析

实验采用了IPVR 3维自由形状(free-form)物体数据库^[18]。该数据库中有42种物体的高分辨率(400×400)的深度图像。训练集包含了5个外形相似的目标(外形差异较大的,在预处理模块就会被排除),每个目标66个姿态(23°~26°视角差)。模型库的数据存储分为两种:1)标志点序列,一共330个;2)表面信息,一共12 264个图像小块(patch),以哈希表的格式存放。在实验中,只分为了10个表,以 μ 为指数。

测试集包含与训练集相对应的5种目标(如图8所示),其中每个目标有258个姿态(与训练集的姿态不同)。



图8 5种自由形状的物体

Fig. 8 Five free-form objects

5.1 在理想的深度图像中测试

算法需要预先设置标志点序列在匹配中的相似度阈值 $s_{L_{thre}}$ 和搜索半径 d_{thre} 。通过设定不同的阈值 $s_{L_{thre}}$ 和 d_{thre} ,可以决定候选目标列表的数目。阈值选取需要在识别精度和效率之间平衡,取较大的阈值 $s_{L_{thre}}$,可以排除较多的类别和姿态型号,提高识别速度,但是会增加错误匹配的概率,更多的情况是,输出的类别正确,但姿态差别大。而取较大的 d_{thre} ,降低了目标间对外形的约束,如图9的识别结果中的候选图像3(图中第2行第3列图),可以避免排除正确类别,但是会导致整个识别效率不高。

考虑以上因素,通过实验选取 $s_{L_{thre}} = 0.5$ 和 $d_{thre} = 30$ (考虑自遮挡和噪声干扰的情况)。预处理后,一般能排除模型库中90%以上的候选(包括目标类别和姿态)

同理,在识别模块中SIFT算子和局部区域特征的相似度量策略需要确定的参数,也需要考虑效率和精确的平衡问题。

SIFT算子的参数选取原则,主要考虑关键点个数。因为个数越多,每个目标生成的局部区域越多,有可能出现区域重叠,导致更多的冗余,影响识别效

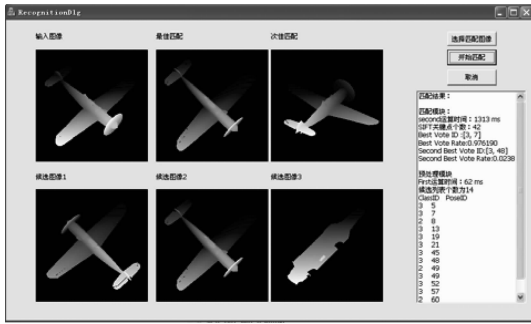


图 9 理想情况下识别结果

Fig. 9 Result of recognition

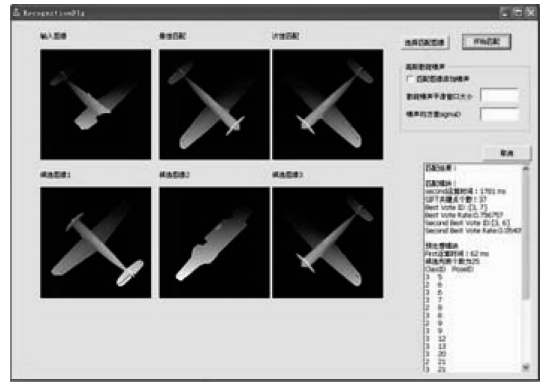


图 10 目标被遮挡的识别结果

Fig. 10 Result of recognition with occluded object

率。一般关键点个数在 100 以内,效果较好,设置的参数以此为标准。

局部区域由式(7)确定。并根据 ε 、 θ 和 s 这三个阈值确定了一个合理局部区域的大小。 ε 确定一个初步搜索的范围,而 θ 和 s 确定最终的区域大小。在实验中, $\varepsilon = 11$,即以关键点为中心,在半径为 11 的领域内,搜索与其曲面类型相似 ($s = 0.25$) 以及法向量夹角不超过 θ ($\theta = 30^\circ$) 的连续区域作为局部区域,并提取局部特征。阈值的选取原则参见 3.2 小节。

图 9 为飞机某姿态下的识别结果,左上角为待匹配图像,第 1 行后两个图像分别为识别模块输出的最佳匹配和次佳匹配(投票率最高的前两个)。第 2 行为预处理模块输出的候选类别列表,随机地显示 3 个目标。在理想状态下,对目标类别的识别率在 90% 以上,差别在于最终输出的姿态。

5.2 目标存在遮挡的情况

当目标出现遮挡时,主要对预处理模块产生影响,而识别模块的局部区域特征(区别于全局特征)对遮挡有一定的适应性^[19]。

当目标出现遮挡时,目标的外轮廓出现了形变,在相似度阈值不变的情况下,候选目标列表的数目将增加,导致了识别模块需要匹配的数目增加,效率降低。若遮挡较小(30% 以下),对识别结果的类别基本没有影响,但获取的姿态并不一定是最优的。若遮挡较大(50% 以上),识别率会降低到 50% 以下。

图 10 为飞机头部和机翼出现遮挡的情况下的识别结果。当出现遮挡时,预处理模块输出的候选列表中类别/姿态号增多,影响识别效率,同时导致最佳匹配结果的投票率降低,但对识别结果影响不大。

5.3 目标存在散斑噪声的情况

当目标存在噪声干扰时,对预处理模块的影响不大,因为在获取轮廓时进行了形态学处理,同时相对于轮廓,轮廓曲线对噪声的敏感度要低很多。但是在识别模块,局部区域的特征对噪声比较敏感。

目标加入了散斑噪声^[20]的干扰。实验中对不同散斑噪声的方差进行测试。在识别前,需要对图像进行平滑处理。实验中,选取的平滑窗口为 5。

图 11 显示了加入方差为 0.5 的散斑噪声,对识别结果的影响。散斑噪声的加入,影响了 SIFT 算子,增加了关键点的个数,导致识别效率降低。同时,每一个局部区域也受到了干扰。

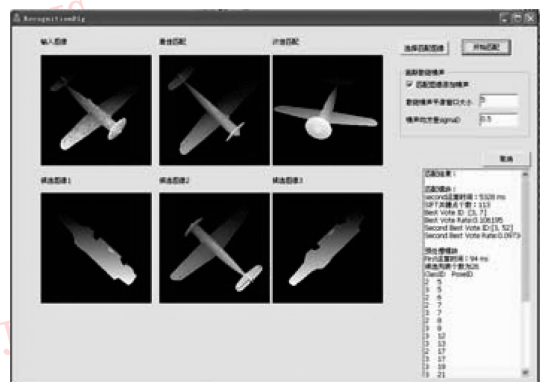


图 11 目标加入散斑噪声的识别结果,噪声方差为 0.5

Fig. 11 Result of recognition with sparkle noise

当进一步增加噪声干扰时,一些关键点的局部区域在模型库中都找不到相似区域,导致投票率大大降低。如图 12 所示,输出的识别结果并不是最佳的。

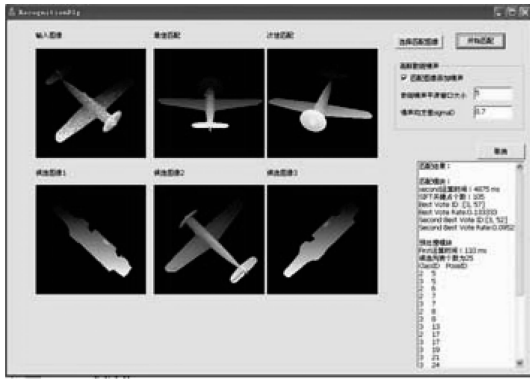


图12 目标加入散斑噪声的识别结果,噪声方差为0.7

Fig. 12 Result of recognition with sparkle noise

5.4 系统模块运行时间

系统运行的硬件平台 Pentium® Dual-Core CPU, 2.60 GHz, 内存为 1.96 GB; 软件平台为 C++ 编写的测试程序。

预处理模块在各种情况下, 运算时间的差别不大。而精确识别模块的耗时则根据情况差异较大, 在模型库规模一定的情况下, 主要两个因素有关:

1) 候选列表中类别/姿态的个数; 2) 关键点的个数。其中, 每个关键的局部区域与模型库的匹配时间与候选列表有关, 与候选的类别或姿态个数呈线性关系, 如图 13 所示, 数据采集于两种情况下: 理想和噪声干扰。

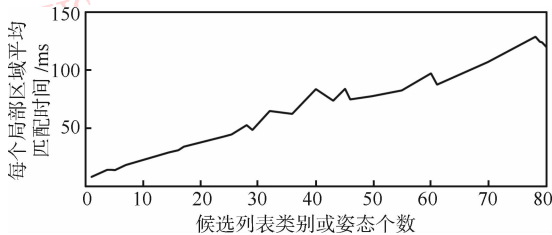


图13 每个局部区域匹配所需要的时间

Fig. 13 Matching time of local patch

由图 13 可以看出, 预处理模块排除的类别/姿态越多, 精确匹配需要的时间就越少, 显著的提高运销效率。

表 1 列出了在理想和散斑噪声干扰(散斑噪声方差为 0.5, 平滑窗口大小为 5 个像素)情况下模块的平均运行时间和识别率。

目前大多数 3 维识别算法研究主要集中在精确匹配上, 区别在于选择不同的特征组合或者不同的匹配策略, 往往忽略了识别效率。本文算法通过增加预处理模块, 识别所消耗的平均时间为 2 152 ms,

而仅用精确匹配, 所需要的平均时间为 10 912 ms (每个关键点的平均匹配时间为 167 ms)。本文算法在保证识别率较好(92.7%)的情况下, 有效地提高了识别效率。

表 1 实验结果

Table 1 Experiment results

	平均数目		模块平均运行时间/ms		识别率 /%
	候选类别	关键点	预处理	识别	
理想	25	59	41	2 111	92.7
噪声	28	113	89	5 208	80.1

实际上, 模型库中的哈希表只做了简单的分块, 若进一步细化, 能够非常有效的提高精确识别的效率。

5.5 识别率分析

Hetzel 等人基于 IPVR 物体数据库^[19]进行了实验, 特征选用了图像深度、表面法向量和表面指数的不同组合。根据不同的特征组合, 识别率达到 94.25% ~ 97.26%。

针对该数据库, 仅采用精确识别模块, 识别率达到 95% 以上。

本文的方法在识别率上略低于以上提到的算法。产生这种现象的原因, 在于预处理模块在轮廓匹配时错误地将一些正确的模型排除。因为该模块为了提高匹配效率, 仅仅考虑了 2 维轮廓信息。

6 结论

给出一种基于轮廓曲线和局部区域特征匹配的由粗到细的一般物体识别方法。第 1 步以提取轮廓的特征点做相似匹配, 作为排除算法, 将模型库中外形或姿态差异较大的排除; 第 2 步对剩余的模型库进行精确匹配, 完成识别。实验结果证明, 本文的方法可以保证较好识别精度的情况下, 有效提高识别速度, 而且对部分遮挡和噪声干扰有一定的适应性。提出的方法有如下特点:

1) 基于物体局部区域特征的匹配, 计算复杂, 耗时长, 难以满足实时性要求。结合简单、稳定的标志点序列匹配算法, 在精确匹配前快速排除部分类别和姿态, 提高了识别速度;

2) 利用轮廓曲线来计算轮廓的标志点, 降低了噪声对轮廓干扰的影响;

3) 局部区域特征采用两个 1 维直方图的形式,

在模型库中只存储非零的部分,简化了模型数据,提高匹配速度;

4)模型库采用两种形式的存储方式,降低了占用的存储空间,尤其物体表面信息利用哈希表的方式存储,有效地降低了匹配过程中比较的次数,大大提高了识别效率。

参考文献 (References)

- [1] Liu B F, Xiao J, Ping X J, et al. The representation and comparability determination of three-dimensional surface [J]. Journal of Image and Graphics, 2001, 6(2): 199-203. [柳葆芳,肖菁,平西建,等. 三维表面的曲面描述与相似性判别 [J]. 中国图象图形学报, 2001, 6(2): 199-203.]
- [2] Tian Y D. Relevant Algorithms for General Object Recognition Based on Feature Combination [D]. Shanghai Jiaotong University, 2007. [田渊栋. 基于特征组合的一般物体识别相关算法研究 [D]. 上海交通大学, 2007.]
- [3] Li Q, Zhou M L, Liu J. A review on 3D objects recognition [J]. Journal of Image and Graphics, 2000, 5(12): 986-993. [李庆,周曼丽,柳健. 三维物体识别研究进展 [J]. 中国图象图形学报, 2000, 5(12): 986-993.]
- [4] Murase H, Nayar S. Visual learning and recognition of 3D objects from appearance [J]. International Journal of Computer Vision, 1995, 14: 5-24.
- [5] Campbell R, Flynn P. Eigenshapes for 3D object recognition in range data [C] // The Proceedings of International Conference on Computer Vision and Pattern Recognition. Collins, USA: IEEE Computer Society, 1999, 505-510.
- [6] Johnson A E, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes [J]. IEEE Transactions on Pattern Recognition and Machine Intelligence, 1999, 21(5): 433-449.
- [7] Chua C S, Han F. 3D human face recognition using point signature [C] // Proceedings of International Conference on Automatic Face and Gesture Recognition. Grenoble, France: IEEE Computer Society, 2000: 233-238.
- [8] Zhou Z J. Object recognition based on shape features [D]. Chang Sha: Natinal University of Defense Technology, 2007. [周正杰. 基于形状的目标识别方法研究 [D]. 长沙: 国防科学技术大学, 2007.]
- [9] Gonzalez R C, Woods R E. Digital Image Processing [M]. 2ed, Beijing: Publishing House of Electronics Industry, 2006: 525-527.
- [10] Zhang J, Zhang X. Object representation and recognition in shape spaces [J]. The Journal of Pattern Recognition Society, 2003, 36: 1143-1154.
- [11] Lowe D. Object recognition from local scale-invariant features [C] // Proceeding of the Seventh IEEE International Conference on Computer Vision. Collins, USA: IEEE Computer Society, 1999, 1152: 1150-1157.
- [12] Lowe D. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [13] Rachel T V, Paul J S. Local feature extraction and matching on range image: 2. 5D SIFT [J]. Computer Vision and Image Understanding, 2009, 113(12): 1235-1250.
- [14] Koenderink. The structure of images [J]. Biological Cybernetics, 1984, 50(5): 363-370.
- [15] Lindeberg. Scale-space for discrete signals [J]. Pattern Analysis and Machine Intelligence, 1990, 12(3): 234-254.
- [16] Dorai C, Jain A K. COSMOS-A representation scheme for 3D free-form objects [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(10): 1115-1130.
- [17] Belkasim S, Shridhar M. Pattern recognition with moment invariants: a comparative study and new results [J]. Pattern Recognition, 1991, 24(12): 1117-1138.
- [18] Hetzel G, Leibe B. 3D object recognition from range images using local feature histograms [C] // Proceedings of International Conference on Computer Vision and Pattern Recognition, Kauai, USA: IEEE Computer Society, 2001, 2: 394-399.
- [19] Tan F. Object recognition based on features extraction [D]. Chengdu: University of Electronic Science and Technology of China, 2010. [谭芳. 基于特征信息提取的目标识别算法研究 [D]. 成都: 电子科技大学, 2010.]
- [20] Li Z Q, Li Q, Wang Q. Noise characteristic in active laser imaging system by statistic analysis [J]. Chinese Journal of Laser, 2004, 31(9): 1081-1085. [李自勤, 李琦, 王骥. 由统计特性分析激光主动成像系统图像的噪声性质 [J]. 中国激光, 2004, 31(9): 1081-1085.]