

# 基于浅层特征的印刷品分拣识别系统<sup>①</sup>



谢成亮<sup>1,2</sup>, 王鸿亮<sup>1</sup>, 何薇薇<sup>3</sup>, 赵杰<sup>3</sup>, 王帅<sup>4</sup>

<sup>1</sup>(中国科学院 沈阳计算技术研究所, 沈阳 110168)

<sup>2</sup>(中国科学院大学 计算机与控制学院, 北京 101408)

<sup>3</sup>(陆军炮兵防空兵学院 士官学校, 沈阳 110161)

<sup>4</sup>(山东省科学院 情报研究所, 济南 250014)

通讯作者: 王鸿亮, E-mail: [wanghl@sict.ac.com](mailto:wanghl@sict.ac.com)

**摘要:** 针对印刷领域印刷品的识别问题, 设计了一种基于浅层特征的视觉识别系统. 浅层特征适用于对精准特定目标的识别, SIFT 是一种被广泛使用且效果良好的浅层特征. 该系统首先建立一个小型样本图像数据库, 然后采用 SIFT 算法进行特征提取并与样本图像进行匹配识别. 另外, 该系统实现了在线学习的功能. 经过实验验证, 该系统可以实时、准确的实现对印刷品内容的识别.

**关键词:** 浅层特征; 图像检索; 相似度检测; 特征匹配; 在线学习

引用格式: 谢成亮, 王鸿亮, 何薇薇, 赵杰, 王帅. 基于浅层特征的印刷品分拣识别系统. 计算机系统应用, 2019, 28(2): 62-67. <http://www.c-s-a.org.cn/1003-3254/6787.html>

## Print Sorting Recognition System Based on Shallow Features

XIE Cheng-Liang<sup>1,2</sup>, WANG Hong-Liang<sup>1</sup>, HE Wei-Wei<sup>3</sup>, ZHAO Jie<sup>3</sup>, WANG Shuai<sup>4</sup>

<sup>1</sup>(Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China)

<sup>2</sup>(School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing 101408, China)

<sup>3</sup>(Sergeant School, Army Artillery Air Defense Academy, Shenyang 110161, China)

<sup>4</sup>(Information Research Institute, Shandong Academy of Sciences, Jinan 250014, China)

**Abstract:** Aiming at the problem of the recognition of print content in the printing field, a visual identity system based on shallow features is designed. Shallow feature is applicable to the recognition of specific target. SIFT is a widely used and effective shallow feature. The system first establishes a sample image database, uses SIFT algorithm to extract features, finally identifies by matching with sample features. In addition, the system implements online learning functions. After experimental verification, the system can realize the identification of print content in real time and accurately.

**Key words:** shallow features; image retrieval; near-duplicate detection; feature matching; online learning

随着计算机视觉技术的飞速发展, 工业领域对具有识别功能的视觉识别系统的需求不断增加. 图像检索技术作为视觉识别系统的核心技术, 近年来成为了人们研究的热点.

在印刷品分拣的过程中, 传统的方法是工人根据印刷品的内容进行分拣, 但这样会消耗大量的人力资

源; 而需要识别的印刷品 80% 左右是相同的, 即仅需根据印刷品 20% 左右的内容对印刷品进行分拣, 可以采用图像检索技术实现印刷品的分拣. 为了减少人力成本的投入, 同时提高印刷品分拣的效率, 需要实现一种用于印刷品自动分拣的视觉识别系统.

基于内容的图像检索方法所需提取的特征主要包

① 基金项目: 辽宁省重点研发计划 (2018225096); 中国科学院沈阳分院-山东省科学院青年科学家合作伙伴项目

Foundation item: Key Research and Development Program of Liaoning Province (2018225096); Young Scientist Partner Program of Shenyang Branch, Chinese Academy of Sciences and Academy of Sciences of Shandong Province

收稿时间: 2018-08-26; 修改时间: 2018-09-20; 采用时间: 2018-09-29; csa 在线出版时间: 2019-01-28

括两种: 深层特征和浅层特征. 深层特征主要是通过神经网络和全链接层提取, 对于一类物体具有广泛的代表性, 但是对设备的要求高. 浅层特征是基于领域知识通过固定的算法提取特征<sup>[1]</sup>, 适用于对精准特定目标的识别, 硬件上相对容易实现. 对于图像相似度判断, 可采用 PSNR 峰值信噪比方法<sup>[2]</sup>, 一种全参考的图像质量评价指标, 其感知结果会受到许多因素的影响而产生变化; 直方图方法<sup>[3]</sup>, 将图像转换成直方图进行比较, 会受到亮度光照条件的影响; SIFT 方法, 提取图像的局部特征, 其对旋转, 尺度缩放, 亮度变化保持不变性, 对视角变化, 仿射变化, 噪声也保持一定的稳定性.

考虑到系统对识别稳定性的要求和成本问题, 针对浅层特征, 采用基于 SIFT 算法的识别技术, 设计和实现了应用于印刷品分拣的视觉识别系统. 下面, 将从系统架构, 关键算法和功能实现等方面进行介绍.

## 1 系统架构

该系统主要由工业摄像头, 工控主机, 显示设备和执行设备组成. 系统架构框图如图 1 所示. 工业摄像头主要用来采集印刷品图像, 并传入到工控主机中进行处理; 工控主机是整个系统的处理核心, 主要实现了两个功能, 样本的检测识别和在线学习. 针对样本检测, 工控主机对从摄像头输入的视频, 按帧提取图像并做大小和灰度处理, 然后提取图像特征与存储的样本进行匹配, 最后将识别结果和处理信息分别传入到显示设备和执行设备; 针对在线学习, 工控主机读取输入样本和样本描述信息, 然后对样本图进行特征提取, 最后存储到系统的样本数据库中. 显示设备主要用来进行人机交互. 执行设备就是根据工控主机传出的信息对相应的印刷品进行操作.

## 2 浅层特征识别算法

针对印刷品识别过程中容易出现亮度变化, 印刷品旋转和尺度缩放等问题, 本系统选择了 SIFT 算法作为本系统的核心算法. SIFT 算法由 David Lowe 在 1999 年提出, 在 2004 年加以完善<sup>[4,5]</sup>. 下面, 将介绍一下本系统的关键算法.

SIFT 特征匹配主要包括两部分: 第一部分是 SIFT 特征的提取和描述特征的特征向量的生成; 第二部分是 SIFT 特征向量的匹配.

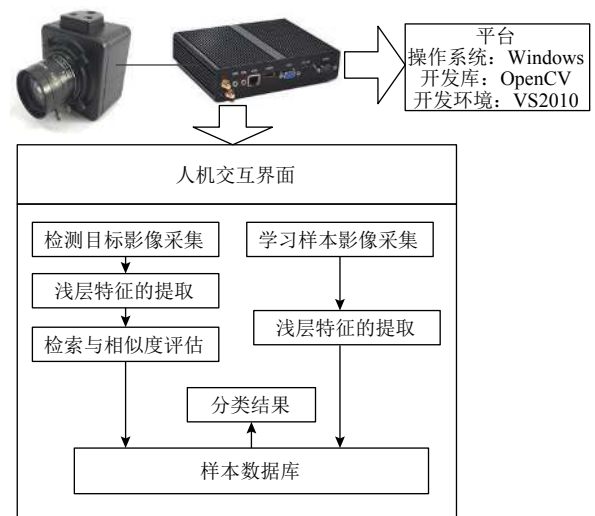


图 1 系统架构框图

### 2.1 特征提取与描述生成

特征提取与描述特征向量的生成一般包括以下几步:

1) 构建尺度空间, 检测极值点, 获取尺度不变性. SIFT 算法是在不同的尺度空间上查找关键点, 而尺度空间的获取需要使用高斯模糊来实现, Lindeberg<sup>[6]</sup>等人已证明高斯卷积核是实现尺度变换的唯一变换核, 并且是唯一的线性核. 尺度空间理论的基本思想是在图像信息处理模型中引入一个被视为尺度的参数, 通过连续变化尺度参数获得多尺度下的尺度空间表示序列, 对这些序列进行尺度空间主轮廓的提取, 并以该主轮廓作为一种特征向量, 实现边缘, 角点检测和不同分辨率上的特征提取等. 尺度空间可表示为:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

其中, \* 表示卷积运算,

$$G(x, y, \sigma) = (1/2\pi\sigma^2) \times e^{-[(x-m/2)^2 + (y-n/2)^2]/2\sigma^2} \quad (2)$$

其中,  $m, n$  表示高斯模板的维度 (由  $(6\sigma + 1) \times (6\sigma + 1)$  确定),  $(x, y)$  代表图像的像素位置.  $\sigma$  是尺度空间因子, 值越小表示图像被平滑的越少, 相应的尺度也就越小. 大尺度对应于图像的概貌特征, 小尺度对应于图像的细节特征. 尺度空间在实现时使用高斯金字塔表示, 高斯金字塔的构建分两部分: 首先, 对图像做不同尺度的高斯模糊; 其次, 对图像做降采样 (隔点采样). Lowe<sup>[4]</sup>使用高效的高斯差分算子代替拉普拉斯算子进行极值检测, 如下:

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

在实际计算时,使用高斯金字塔每组中相邻上下两层图像相减得到高斯差分图像.在空间极值点检测(关键点的初步探查)中,关键点是由DOG空间的局部极值点组成的,关键点的初步探查是通过同一组内各DOG相邻两层图像之间比较完成的.

2) 特征点过滤并精确定位.通过拟合三位二次函数来精确确定关键点的位置和尺度,同时去除低对比度的关键点和不稳定的边缘响应点,以增强匹配稳定性,提高抗噪能力.

3) 为特征点分配方向值.为使描述符具有旋转不变性,需要利用图像的局部特征为每一个关键点分配一个基准方向.使用图像梯度的方法求取局部结构的稳定方向.对于在DOG金字塔中检测出的关键点,采集其所在高斯金字塔图像 $3\sigma$ 邻域窗口内像素的梯度和方向分布特征.梯度的模值和方向如下:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (4)$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1)) / (L(x+1,y) - L(x-1,y))) \quad (5)$$

其中, $L$ 为关键点所在的尺度空间值,梯度的模值 $m(x,y)$ 按 $\sigma = 1.5\sigma_{oct}$ 的高斯分布加成,按尺度采样的 $3\sigma$ 原则,邻域窗口的半径为 $3 \times 1.5\sigma_{oct}$ .在完成关键点的梯度计算后,使用直方图统计邻域内像素的梯度和方向,方向直方图的峰值则代表了该特征点处邻域梯度的方向,以直方图中最大值作为该关键点的主方向.为增强匹配的鲁棒性,只保留峰值大于主方向峰值80%的方向作为该关键点的辅方向.因此,对于同一梯度值的多个峰值的关键点位置,在相同位置和尺度将会有多个关键点被创建但方向不同.仅有15%的关键点被赋予多个方向,但可以明显提高关键点匹配的稳定性.

4) 生成特征描述子.通过以上操作,对于每一个关键点,拥有三个信息:位置,尺度以及方向.接下来就是为每个关键点建立一个描述符,用一组向量将这个关键点描述出来,使其不随各种变化而变化,比如光照变化,视角变化等.这个描述子不但包括关键点,也包括关键点周围对其贡献的像素点,并且描述符应该有较高的独特性,以便于提高特征点正确匹配的概率<sup>[7-9]</sup>.

## 2.2 特征匹配

当图像的SIFT特征向量生成以后,就可以采用关

键点特征向量的欧式距离来作为图像中关键点的相似性判定度量.取一张图像的某个关键点,通过遍历找到另一种图像中的距离最近的两个关键点.在这两个关键点中,如果最近距离除以次近距离小于某个阈值,则判定为一对匹配点.

为了进一步提高匹配精度,我们需要对特征匹配过程中产生的一些误匹配进行消除.消除误匹配主要采用随机样本一致性(RANSAC)方法.RANSAC可以从一组包含“局外点”的数据中,通过迭代的方式训练最优的参数模型,不符合最优参数模型的被定义为“局外点”.消除误匹配的原理是采用RANSAC算法寻找一个最佳单应矩阵 $H$ ,矩阵大小是 $3 \times 3$ .RANSAC的目的是找到最优参数矩阵使得满足该矩阵的数据点个数最多,通常令 $h_{33} = 1$ 来归一化矩阵.RANSAC算法从匹配数据集中随机抽出4个样本并保证这4个样本之间不共线,计算出单应性矩阵,然后利用这个模型测试所有数据,并计算满足这个模型数据点的个数与投影误差(即代价函数),若此模型为最优模型,则对应的代价函数最小.RANSAC消除误匹配点可分为三部分:根据matches将特征点对齐,将坐标转换为float类型;使用求基础矩阵的方法findFundamentalMat()得到RansacStatus;最后,根据RansacStatus来删除误匹配点,即对应RansacStatus的值为0的点<sup>[10-13]</sup>.

## 3 检测识别和在线学习

### 3.1 检测识别

本系统主要实现了两个功能,对摄像头获取的印刷品图像进行检测识别和在线增加系统识别种类的学习功能.实现检测识别功能的两个关键步骤是小型样本图像数据库的建立和对输入图像进行实时检测.

对输入图像的检测,一般的方法是,在系统运行过程中,对输入图像和选取的样本图像进行特征提取,然后采用遍历的方式,将输入图像和样本依次进行相似度检测,判断输入图像属于哪个样本图像代表的类型,最后依据检测的结果输出结论,如图2(a)所示.但是,在系统实际运行时,如果系统同时对输入图像和样本图像进行特征提取,再进行相似度检测,会产生非常大的运算量,对系统实时性产生很大的影响.为了提升系统执行的速度,我们可以做出优化的地方有两点,第一是减少系统在特征提取时的运算量;第二是减少系统在进行相似度检测时输入图像和样本图像比较的次数.



针对第二点, 样本图像是以集合的形式存储, 不存在顺序性, 所以在进行图像相似度检测时只能采用循环遍历的形式. 针对第一点, 样本图像是事先提供的, 具有相对的不变性. 因此, 可以提前对要进行检测的图像样本提取特征, 然后保存下来, 如图 2(b) 所示. 当系统运行需要用到样本特征进行相似度检测时, 在读取到内存, 同提取的输入图像的特征进行检测. 由于提前提取了样本图像的特征, 在进行样本和输入图像相似度检测时, 只需要提取输入图像的特征和已提取特征的样本进行相似度比较, 从而避免了对样本特征的重复提取, 大大减少了 CPU 的运算量, 给系统的实时性带来了很大的提升.

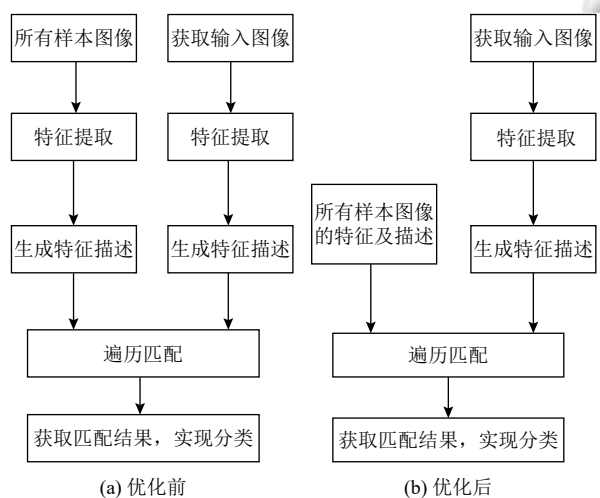


图 2 识别流程图

为了存储提前提取的样本图像的特征和描述信息, 需要建立一个小型的样本数据库. 由于针对的样本类型有限, 因此可以不用选择使用已有的数据库, 只需自己实现一个具有存储数据功能的小型数据库即可. 在创建小型数据库中需要存储的信息主要包括: 样本图像, 样本图像特征, 样本图像特征描述以及对样本信息的描述. 对于样本信息对应问题的解决方法是, 以文件名为纽带将每个样本的图像, 特征, 特征描述以及样本信息关联起来. 小型数据库建立主要包括两步, 1) 在系统中创建一个文件夹, 命名为 `sampledb`, 作为本数据库的根文件. 然后在该文件下分别创建四个文件夹, 用来存储样本的各个信息, 并根据所存的信息为文件命名. 2) 根据提供的样本图像, 分别对它们提取特征, 特征描述并生成样本的描述信息, 存储在相应的文件夹下. 图像特征和特征描述采用 `xml` 文件存储, 实现 `xml` 文件

的读和写采用的是 OpenCV 提供的两个方法:

```
FileStorage fs("name.xml", FileStorage::READ);
```

```
FileStorage fs("name.xml", FileStorage::WRITE);
```

样本信息采用 `txt` 文件存储. 接着将这些信息统一命名, 然后存储在各自对应的文件夹下. 当系统需要获取相应的样本图像信息时, 只要在系统中写定相应的路径, 然后遍历获取对应文件名的文件, 读取该文件中的信息即可. 最后, 需要在根文件夹下创建一个 `txt` 文件, 用于存储数据库中的样本的总体信息, 该信息主要用于实现本系统的在线学习功能.

在系统开始执行后, 从数据库中将每个样本的特征信息都入到内存中. 接着从摄像头输入的视频中按照一定的时间间隔获取视频的帧, 保存到 `Mat` 类型的变量中, 对获取的图像进行大小和灰度处理, 使输入的图像成为固定大小的灰度图像, 调用 SIFT 算法提取图像特征并生成图像特征描述. 接着利用获取的输入图像的特征及描述与读入的样本分别进行相似度检测, 判断获取的图像与哪一类的样本相匹配. 相似度检测的实现, 在判断两种图像是否匹配时, 首先以输入图像的特征描述为基础与样本图像的特征相匹配, 接着进行消除误匹配的操作, 得到消除误匹配后两张图像的特征点的个数, 通过验证当特征点匹配的个数大于样本图像的特征点的个数的 40% 时, 便可以判断两种图像是匹配的. 针对特殊情况, 当出现多个满足上述验证过程的样本时, 根据消除误匹配后特征点依旧匹配的个数进行排序, 选取个数最大的样本匹配. 如果依旧无法判断, 那么视这种情况为输入图像没有匹配的样本. 现在, 便完成了输入图像类型的判断, 然后从样本数据库中调出相应的样本信息提交给系统即可. 如此, 便实现了系统进行检测识别的功能.

### 3.2 在线学习

对于本系统实现的在线学习功能, 首先要与机器学习中的学习做出区别. 在机器学习中, 学习的含义是从数据中学得模型的过程, 这个过程通常执行某个学习算法来完成, 学习过程就是为了找出或逼近关于数据的某种潜在的规律<sup>[14]</sup>. 在本系统中, 学习的含义是可以向系统中添加新的样本图像, 使得系统具备识别与输入样本图像相匹配的图像的功能. 本系统实现的学习功能本质上是, 设计出一个人机交互的界面, 通过此界面操作者可以向系统的样本数据库中添加新的样本, 使得系统增加识别的类型.

系统学习功能的实现, 是通过在系统主界面上添加一部分, 在该部分上可以选择需要加入的样本图像和添加对样本信息的描述, 操作人员可以通过该部分添加新的样本和其信息的描述, 确认后系统可以自动调用 SIFT 算法对样本图像进行特征提取, 然后存储到样本数据库中, 文件命名的方式则是根据数据库根文件夹下存放的样本的整体信息文件确定, 从而使样本具有唯一性. 如此, 系统便实现了在线学习的功能.

## 4 实验分析

### 4.1 系统搭建

本系统硬件包括: 工业摄像头、工控主机和显示设备, 如图 3 所示. 在软件方面, 工控主机安装的是 windows 系统. 本系统采用的开发环境是 vs2010+MFC+openCV, 系统开发完成后直接发布到工控主机上.



(a) 工业摄像头

(b) 工控主机



(c) 系统整体硬件结构

图 3 实验阶段系统的硬件图

### 4.2 性能测试

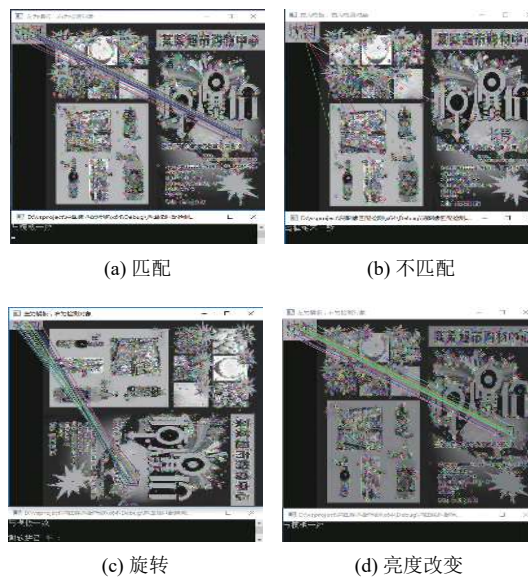
本系统主要关注的实验数据有两部分: 系统在检测识别的时间效率和准确率. 系统的时间消耗主要体现在输入图像的特征提取及描述生成的时间消耗, 以及输入图像与样本匹配的时间消耗. 通过对上述两点的改进和优化, 本系统基本满足印刷品识别中对系统实时性的要求, 具体的实验数据如表 1 所示. 针对检测识别的准确率, 我们通过比较各种图像相似度比较方

法, 最终选择了效果很好的 SIFT 算法, 同时在识别的过程中加入了消除误匹配方法. 通过大量实验对图像匹配标准的调整, 本系统对环境光照变化、输入图像的旋转变化的鲁棒性基本满足实验对环境和输入的要求, 图 4 是系统识别方法在输入图像正常、发生旋转和光照变化情况下的匹配效果.

对于系统在检测识别的准确率验证, 选取 10 个样本, 在相同的光照条件、不同的旋转角度下, 对多种不同类型的印刷品进行识别验证. 通过大量实验, 得出系统在检测识别的准确率达到 98%, 满足印刷品识别中对准确率的要求.

表 1 系统单张输入图像检测识别消耗时间

匹配样本所存位置 (样本总数 10)	生成特征及 描述 (s)	与样本 匹配 (s)	单图检测 识别 (s)
1	0.61	0.02	0.65
5	0.67	0.09	0.81
10	0.60	0.20	0.91
无	0.63	0.25	0.94
平均值	0.65	0.18	0.87



(a) 匹配

(b) 不匹配

(c) 旋转

(d) 亮度改变

图 4 匹配鲁棒性测试

### 4.3 原型系统实现

通过研究与开发, 最终实现了本系统. 系统基于 MFC 的开发基本上满足了与操作人员交互的功能. 操作人员可以通过显示设备操作整个系统、检测系统的运行情况以及可以向系统的样本数据库中添加或删除所要识别的样本类型. 系统的效果图如图 5 所示.



(a) 系统初始化界面



(b) 系统运行时界面

图5 系统成果展示

## 5 总结

为解决印刷品自动分拣的识别问题, 本系统采用 SIFT 算法, 提出并实现了基于 XML 格式浅层特征保存的在线学习和基于线性表检索的图像检测方法. 使用 OpenCV 库、MFC 框架, 在 VS2010 平台上设计并实现了对印刷品检测识别和在线学习功能. 通过发布在工控主机上进行实验验证, 系统实现了对印刷品的实时检测, 并取得了很好的效果. 本系统的实现将会在很大程度上节约人力, 提高印刷品分拣的效率.

### 参考文献

1 于俊清, 吴泽斌, 吴飞, 等. 多媒体工程: 2016——图像检索

研究进展与发展趋势. 中国图象图形学报, 2017, 22(11): 1467-1485.

2 佟雨兵, 张基善, 祁云平. 基于 PSNR 与 SSIM 联合的图像质量评价模型. 中国图象图形学报, 2006, 11(12): 1758-1763. [doi: 10.11834/jig.2006012307]

3 邹承明, 薛栋, 郭双双, 等. 一种改进的图像相似度算法. 计算机科学, 2016, 43(6): 72-76.

4 Lowe D G. Object recognition from local scale-invariant features. Proceedings of the 7th IEEE International Conference on Computer Vision. Kerkyra, Greece. 1999. 1150-1157.

5 Lowe D G. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91-110.

6 Lindeberg T. Scale-space theory: A basic tool for analyzing structures at different scales. Journal of Applied Statistics, 1994, 21(1-2): 225-270. [doi: 10.1080/757582976]

7 周楠钊. 基于 SIFT 特征的物体识别系统的设计与实现[硕士学位论文]. 兰州: 兰州大学, 2016.

8 孙光民, 王晨阳. 一种基于改进 SIFT 的图像检索算法. 国外电子测量技术, 2016, 38(8): 32-37. [doi: 10.3969/j.issn.1002-8978.2016.08.008]

9 杜浩翠, 范建德. 基于 SIFT 局部特征的图像检索算法研究. 信息与电脑, 2017, (13): 53-55. [doi: 10.3969/j.issn.1003-9767.2017.13.022]

10 张岩, 孙世宇, 胡永江, 等. 基于特征距离与内点的随机抽样一致性算法. 电子与信息学报, 2018, 40(4): 928-935.

11 段嘉旭, 罗林, 高晓蓉, 等. 基于抽样随机一致性和图像特征的图像配准. 信息技术, 2017, (11): 100-104.

12 赵焯, 蒋建国, 洪日昌. 基于 RANSAC 的 SIFT 匹配优化. 光电工程, 2014, 41(8): 58-65. [doi: 10.3969/j.issn.1003-501X.2014.08.010]

13 刘川熙, 赵汝进, 刘恩海, 等. 基于 RANSAC 的 SIFT 匹配阈值自适应估计. 计算机科学, 2017, 44(S1): 157-160.

14 周志华. 机器学习. 北京: 清华大学出版社, 2016.