

# 基于改进SIFT算法及定位信息的建筑物识别系统

刘涛<sup>1</sup>, 姜丰<sup>2</sup>

<sup>1</sup>西南民族大学电子信息学院, 四川 成都

<sup>2</sup>西南民族大学电气工程学院, 四川 成都

Email: 2268979210@qq.com

收稿日期: 2021年6月4日; 录用日期: 2021年6月17日; 发布日期: 2021年7月9日

---

## 摘要

为了解决小范围、图像数据较少区域内传统建筑物识别系统精度偏低的问题, 本文提出一种基于改进SIFT算法与GPS定位信息相结合的建筑物识别系统, 并在系统中加入对建筑物存在可能性预测的预处理机制, 进一步提高准确性和快速性。实验验证结果表明, 该方法为可行有效的且实际效果良好。

## 关键词

建筑物识别, SIFT特征, 直线检测, Canny边缘检测, 特征匹配

---

# Building Recognition System Based on Improved SIFT Algorithm and Positioning Information

Tao Liu<sup>1</sup>, Feng Jiang<sup>2</sup>

<sup>1</sup>College of Electronic and Information, Southwest Minzu University, Chengdu Sichuan

<sup>2</sup>College of Electrical Engineering, Southwest Minzu University, Chengdu Sichuan

Email: 2268979210@qq.com

Received: Jun. 4<sup>th</sup>, 2021; accepted: Jun. 17<sup>th</sup>, 2021; published: Jul. 9<sup>th</sup>, 2021

---

## Abstract

In order to solve the general problem, that is, the accurate recognition rate is low in a small extent or when the image resources are few and scattered. This article puts forward a building recognition system that combines GPS positioning information with an improved SIFT algorithm, and

adds a pretreatment mechanism for predicting the possibility of the presence of the building in the system, further reducing mismatches and improving response speed. Final verification shows that this research is actually effective.

## Keywords

Building Recognition, SIFT Feature, Straight Line Detection, Canny Edge Detection, Feature Matching

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

图像处理和计算机视觉在人工智能(Artificial-Intelligence, AI)的极速发展下开始迅速兴起, 图像信息的识别是研究热点之一, 且用户对信息的获取方式已不再单纯的停留在文字, 人们更希望可以直接从图像本身获取信息[1]。建筑物识别又属于其中一个热点研究方向。传统的识别方式仅仅对建筑物本身特征进行识别或利用全球卫星定位系统进行定位, 但当在小范围、数据较少区域内, 仅靠单一识别技术, 精确性和快速性很难得到保证。针对以上问题, 本文引入对建筑物存在预判的预处理机制, 并基于改进 SIFT 算法与定位信息提出一种新的建筑物识别系统。

近年来, 我国许多研究者在此方向进行了大量研究, 并提出了许多行之有效的方法。2012年, 李松霖等人在特征线匹配的基础上提出一种建筑物识别系统, 以适应在移动条件下的识别要求, 但精度稍差[2]。2015年, 蔡兴泉等人针对传统的识别方式在响应速度和识别效率上难以兼顾, 设计了一种 GPS 匹配与特征匹配结合的识别系统, 在移动设备上仍具有良好的效果[3] [4]。2019年, 刘芯彤等人提出使用方差生成灰度图对图像显著区域进行检测, 提高准确度, 但仍有误匹配点影响, 效果一般[5]。

## 2. 建筑物识别系统

为了解决在小范围、图像数据较少区域内识别准确度低、效率低等问题, 本文所提出的建筑物识别系统中创新点加入针对建筑物的预筛查操作, 进一步减少运算资源的浪费, 此系统将被主要划分为四个模块, 如图 1 所示, 分别为: 预处理模块、特征提取模块、数据库创建模块、特征匹配模块。图 2 所示为系统的逻辑流程图。

### 2.1. 预处理模块

现阶段许多的建筑物识别系统主要是在内容上的图像检索(CBIR) [6], 在此之中并未针对待识别图像中是否存在建筑物进行预判。与自然物体相比, 建筑物的结构更加具有规则, 在相关的图像中可能有迹可循, 列如多条直线的相交、特定的构造模式等。该模块通过对直线及其消失点的检测对是否存在建筑物进行预判。

通过平滑操作以减少图像中噪声的影响, 在此之后引入 Canny 算子完成边缘检测。所获得的边缘图像将用于之后的直线检测, 此处通过霍夫变换得到直线参数并判断其数量。空间坐标中同一条直线上的点映射到另一参数空间中来定义曲线, 他们都会相交于一点。但通常这些交点在此空间中的某一点(x, y)附近。由于这些离群点的存在, 会产生误差直线, 必定会对下一步消失点检测产生影响。本文通过寻找消失点簇, 并通过计算两点之间距离  $d$ , 设定阈值  $T$ , 若  $d \leq T$  则视为相邻点, 相邻点数量大于某一设定值时, 作为消失点。

在此之上对整幅图像中是否存在建筑物进行预判操作, 而不是直接对其进行特征匹配, 这将减少对无关图像的操作, 以此提高本系统的准确度和高效性。

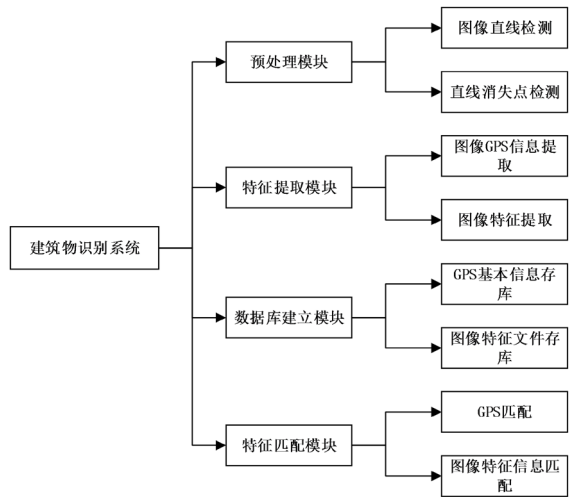


Figure 1. Structure diagram of building recognition system  
图 1. 建筑物识别系统结构图

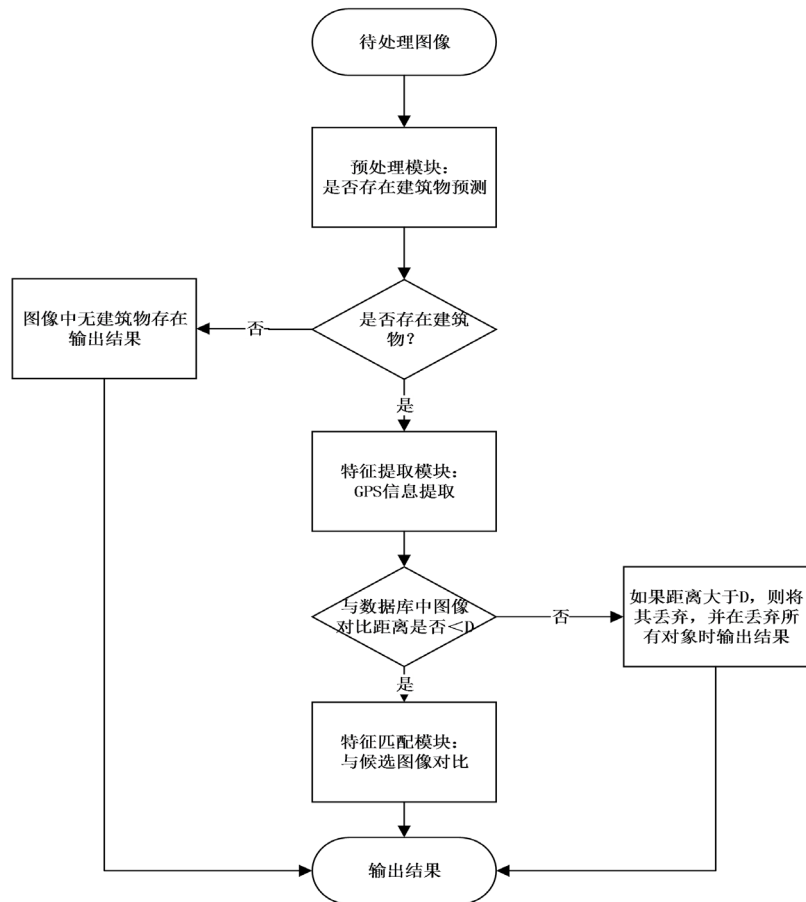


Figure 2. System logic flow chart  
图 2. 系统逻辑流程图

## 2.2. 特征提取模块

由于拍摄设备、拍摄角度等外因影响, 故相应图像并不是每张都为标准尺寸。因此, 先进性待识别图像的尺寸缩小, 本文采用双线性插值法进行操作, 以此减少响应时间。然后对待识别图像进行 GPS 信息的提取, 为后续缩小图像匹配范围; 之后使用改进的 SIFT 算法对待识别图像进行特征提取, 并将提取到的特征描述算子保存为特征文件, 为后续匹配操作使用。

SIFT 作为一种成熟算法, 其基于待识别物体的局部特点加以提取, 对待识别物体的角度、尺寸等要素无关且对噪声有很高的容忍度[7], 但传统 SIFT 在建筑物识别方面存在特征点的误匹配, 增加无关点的存在会影响响应时间和系统精准度, 且高达 128 维的描述算子会使得其具有很高的时间和空间复杂度。本文针对此进行改进, 在不影响原始算法鲁棒性的基础上, 提升对图像中显著区域的特征提取和缩短响应时间。

### 1) 构建高斯差分金字塔

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

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

其中  $L(x, y, \sigma)$  高斯尺度空间模型,  $G(x, y, \sigma)$  为尺度变化高斯函数,  $I(x, y)$  是原图像像素灰度值,  $*$  为卷积运算。

### 2) 关键点定位

关键点由 DOG 空间中的局部极值点组成。每一个检测点都将来自于来自本层和上下两相邻图层中的 26 个像素点进行比较, 其中 8 个像素点来自本层, 相邻层各 9 个。以确保在尺度空间和二维像素空间都能得到极值点, 此时获得的极值点处在离散空间中。由于某些极值点响应较弱, 所以得到的并不都是稳健的特征点, 需要加入拟合三维二次函数来定位关键点的位置及尺度。

### 3) 特征点主方向的计算

为了消除由图像角度改变等操作对匹配产生的影响, 以保证描述算子的旋转不变性。需要计算梯度值和梯度方向, 并在构建领域梯度直方图后为关键点分配一个主方向。

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3)$$

$$\theta(x, y) = \tan^{-1}((L(x+1, y) - L(x-1, y)) / (L(x, y+1) - L(x, y-1))) \quad (4)$$

梯度值和梯度方向由公式(3), (4)求得, 其中  $L$  表示关键点所在的尺度空间值。

### 4) 生成特征描述子

在进行了上述操作后, 接下来需要利用以获取生成特征描述算子, 本文在实验基础上, 拟定将特征点领域划分为  $3 \times 3$  个子区域, 每个子区域看做一个种子点, 并保留 8 个方向的梯度信息, 此时每个特征点都会有  $3 \times 3 \times 8 = 72$  个数据, 这些数据就是构成了在保证算法性能的基础上, 降维后的 72 维 SIFT 特征向量。

原始 SIFT 算法是在 RGB 模式下直接对图形进行灰度处理。本文将在 Lab 模式下进行操作, 相较于传统 RGB 模式, 本模式在处理速度近似相等, 但可定义色彩更多且与光线、设备等无关。该模式不同于 RGB 由三原色通道组成, 而是由一个明亮通道  $L$  和两个色彩通道组成。其中,  $a$  即表示包括颜色从深绿色变化到粉红色的通道;  $b$  则表示从深蓝色变化到黄色的通道, 相比于 RGB 模式, 此模式表示的颜色范围更广。

之后, 计算各通道的期望与方差, 针对色彩对比度不明显的图像, 本文引入参数  $R$  来调节灰度图, 公式(5)为灰度计算方法。

$$S^2(j) = (P_j(L) - E(L))^2 + (P_j(a) - E(a))^2 + (P_j(b) - E(b))^2 \quad (5)$$

$$Gray(j) = \{S^2(j)\}^R \quad (6)$$

其中,  $Gray(j)$  表示第  $j$  个点的像素点的灰度值,  $R$  定义为影响因子, 通过  $R$  的取值来改变灰度值, 以此调节区分度, 使得保留的有效建筑物信息为相对最佳状态。通过实验表明,  $R$  的取值在  $\pm 1.5$  时效果显著, 也可根据需要在  $0.8 \sim 1.8$  之间进行调节, 区分度越大, 图像中相应建筑物越明显。

之后对使用欧式距离的原始 SIFT 算法进行进一步改进, 原文作者思路是每个特征点的描述符都具有 128 维信息, 通过欧式距离完成对比[3]。本文改为采用马氏距离作为特征点之间紧邻的度量标准。马氏距离表示的是数据的协方差距离, 相比于欧式距离, 更适合处理高维数据, 并且独立于测量尺度。以此进一步提高了数据库中图像的特征匹配速度。

若向量样本为  $X_1 \sim X_n$ ,  $S$  为对应的协方差矩阵,  $u$  为均值, 任意  $X$  到  $u$  的马氏距离可以由公式(7)求得。

$$D(X) = \sqrt{(X-u)^T S^{-1} (X-u)} \quad (7)$$

两样本  $X_i, X_j$  之间的马氏距离公式(8)所示:

$$D(X_i, X_j) = \sqrt{(X_i - X_j)^T S^{-1} (X_i - X_j)} \quad (8)$$

相较于原始算法对有效的建筑物区域特征提取得到了改善(与环境相比), 但存在大量相似结构时仍会出现一部分误匹配点, 此处引入 RANSAC 算法加以改进。该算法通过迭代训练的方法从一组包含“局外点”观测数据中训练最佳参数模型。RANSAC 的目的即为寻找最优的参数矩阵使得满足该矩阵的数据点最多[8]。

$$S \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (9)$$

在公式(9)中,  $(x, y)$ ,  $(x', y')$  分别表示比较和输入图像的角点位置,  $S$  为尺度参数。算法从匹配数据集中随机抽取四个不共线的样本, 然后计算其单应性矩阵, 然后利用和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, 则对应值为最小值。代价函数可由公式(10)求得。

$$\sum_{i=1}^n \left( x'_i \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 + \left( y'_i \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 \quad (10)$$

### 2.3. 数据库建立模块

在进行识别操作时, 首先要创建相应的图像特征数据库。本文针对小范围区域; 视野、信号较差; 图像资源较少区域设计, 故数据库建立是搜索有针对性图像资源。笔者以西南民族大学武侯校区为例, 以不同设备, 在不同角度、时间、光线下拍摄了带有 GPS 信息的学校中建筑物照片, 供创建数据库使用。对以上叙述中所获得的图像进行 GPS 信息提取, 并将此信息与图像的景点信息存入数据库中; 之后对相关图像进行尺寸变换并提取相应的特征信息, 最终存放到 XML 特征文件中。

### 2.4. 特征匹配模块

通过预处理模块进行判断后, 若图像中存在建筑物, 将会进入特征匹配模块进行最终的对比。此模块内包含 GPS 信息匹配和特征匹配两方面。



首先, 待识别图像的 GPS 信息与前文已建立的数据库中信息进行匹配操作。我国常用的 WGS-1984 坐标系中, 经度  $1'' \approx 23.6$  米; 纬度  $1'' \approx 30.9$  米。待识别图像定位信息与数据库中所有定位信息计算距离  $D$  (distance), 若之间距离小于 30 米, 即在秒级变化时, 库中相应图片作为候选, 反之则舍弃, 最终减少特征匹配时的计算量。最后, 待识别图像与上一步操作后所形成的候选图像进行特征匹配, 此处采用改进的 SIFT 算法, 若匹配成功, 将对应的图像和相关信息输出给操作界面。

在不考虑地球表面实际地形差异的影响下, 可将地球看做是平均半径  $R = 6371.004$  km 的正球体。此时根据经纬度信息即可得到两点之间距离。若存在  $A, B$  两点, 分别用  $(LonA, LatA), (LonB, LatB)$  表示经纬度。以 0 度经线为基准, 东经设为正值, 西经为负, 在本研究的实际情况下, 纬度暂不考虑正负。由公式(11), (12)可根据经纬度信息求得距离, 其中  $NLonA, NLonB$  表示经过正负处理后的经度信息。

$$C = \sin(LatA)\sin(LatB) + \cos(LatA)\cos(LatB)\cos(NLonA - NLonB) \quad (11)$$

$$Distance = \frac{R \times \arccos(C) \times \pi}{180} \quad (12)$$

### 3. 实验及总结

由图 3 对比可以看出, 原算法对特征点的提取较为分散且对图像中环境干扰的处理能力较弱, 而改进后的算法对建筑物区域有更好的识别效果, 特征点较为集中于目标区域。对比图 4 的实验结果可以发现, 原算法在匹配时存在较多误匹配, 且对于遮挡现象没有很好的处理, 但改进后的算法所展示的结果更为理想, 对局部图像的特征匹配效果更为优秀。



Figure 3. Comparison of extracted feature points  
图 3. 提取的特征点对比

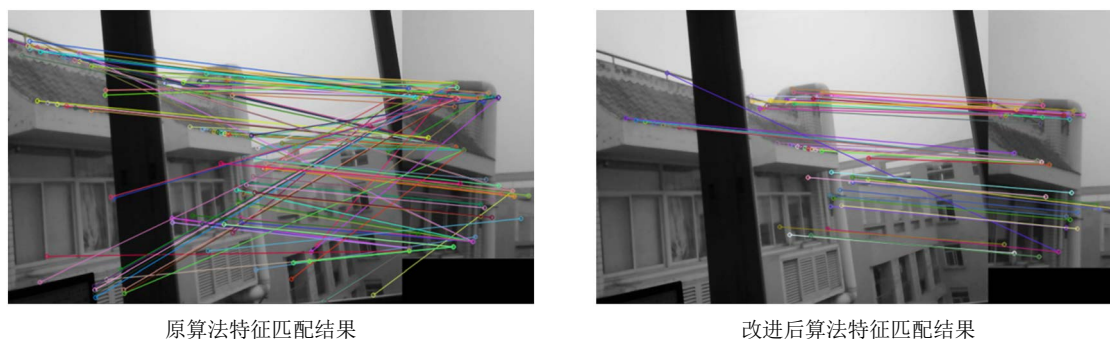


Figure 4. Algorithm improvement comparison  
图 4. 算法改进对比

针对小范围、视野、信号较差区域中传统建筑物识别方式精度不足等问题, 本文提出的识别系统及基于传统 SIFT 算法的改进方法, 可以有效减少数据库中的匹配数据数量, 并减少误匹配点所造成的影响, 增强系统的动态性能和准确性。通过实际操作, 表明本文所提出的建筑物识别系统具有良好的实验效果, 但在预处理方面还有待改进, 避免无关的物体的影响。

## 基金项目

西南民族大学研究生创新型科研项目(项目编号: CX2021SP117)。

## 参考文献

- [1] 王泽泓, 刘厚泉. 基于迁移学习与自适应特征融合的建筑物识别[J]. 计算机技术与发展, 2019, 29(12): 40-43.
- [2] 李松霖, 范海生, 陈秀万. 基于特征线匹配的城市建筑物识别方法研究[J]. 遥感技术与应用, 2012, 27(2): 190-196.
- [3] 柳静华. 面向移动应用的建筑物图像识别技术研究[D]:[硕士学位论文]. 北京: 北方工业大学, 2015.
- [4] 蔡兴泉, 柳静华. 建筑物图像识别系统设计与实现[J]. 现代计算机(专业版), 2015(14): 18-20.
- [5] 刘芯彤, 张辉, 邓志强. 基于 SIFT 算法的建筑物图像识别[J]. 科学与财富, 2019(20): 350.
- [6] 颜廷法. 基于图像序列的特定建筑物识别算法[J]. 泰山学院学报, 2018, 40(3): 68-72.
- [7] Lowe, D.G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, **60**, 91-110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [8] Xu, X.B., et al. (2021) LiDAR-Camera Calibration Method Based on Ranging Statistical Characteristics and Improved RANSAC Algorithm. *Robotics and Autonomous Systems*, **141**, Article No. 103776. <https://doi.org/10.1016/j.robot.2021.103776>