

# 一种基于 2D 和 3D SIFT 特征级融合的一般物体识别算法

李新德<sup>1</sup>, 刘苗苗<sup>1</sup>, 徐叶帆<sup>1</sup>, 雒超民<sup>2</sup>

(1. 东南大学自动化学院复杂工程测量与控制教育部重点实验室, 江苏南京 210096;

2. 底特律大学电子与计算机工程系, 美国底特律 48221)

**摘 要:** 如何选择合适的特征表示一般物体类间差异和类内共性至关重要, 因此, 本文在 2D SIFT(Scale Invariant Feature Transform, SIFT)的基础上, 提出了基于点云模型的 3D SIFT 特征描述子, 进而提出一种基于 2D 和 3D SIFT 特征级融合的一般物体识别算法. 分别提取物体 2 维图像和 3 维点云的 2D 和 3D SIFT 特征描述子, 利用“词袋”(Bag of Words, BoW)模型得到物体特征向量, 根据特征级融合将两个特征向量进行融合实现物体描述, 运用有监督分类器支持向量机(Support Vector Machine, SVM)实现分类识别, 给出最终识别结果. 最后, 实验验证了本文提出算法的好处.

**关键词:** 一般物体识别; 点云; 2D SIFT; 3D SIFT; 特征融合; BoW 模型; SVM 分类器

**中图分类号:** TP391.4      **文献标识码:** A      **文章编号:** 0372-2112 (2015)11-2277-07

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2015.11.021

## A Recognition Algorithm of Generic Objects Based on Feature-Level Fusion of 2D and 3D SIFT Descriptors

LI Xin-de<sup>1</sup>, LIU Miao-miao<sup>1</sup>, XU Ye-fan<sup>1</sup>, LUO Chao-min<sup>2</sup>

(1. Ministry of Education Key Laboratory of Measurement and Control, School of Automation, Southeast University, Nanjing, Jiangsu 210096, China;

2. Department of Electrical and Computer Engineering, University of Detroit Mercy, Detroit, Michigan 48221, USA)

**Abstract:** How to choose the appropriate feature to represent differences between classes and the common within class of generic objects is of great importance. So the 3D SIFT(scale invariant feature transform) descriptors of point clouds model based on the 2D SIFT is proposed. Then we propose a new algorithm based on multiple feature fusion of 2D and 3D SIFT descriptors respectively drawn from 2D images and 3D point clouds. The BoW(bag of words) model is used to calculate feature vectors and describe the objects according to the multiple feature fusion. The supervised support vector machine(SVM) classier is used to classify objects. Through several experiments, the advantage of this new algorithm is testified.

**Key words:** generic object recognition; point cloud; 2D SIFT; 3D SIFT; feature fusion; BoW; SVM

### 1 引言

一般物体识别(Generic Object Recognition)是指使得计算机具有和人类一样的,对于任意环境下观察到的任意物体进行检测、分割和识别的能力.不同于特定物体识别(Specific Object Recognition),可以通过海量的训练样本进行训练,仅仅处理某种物体或某类物体;一般物体识别困难的多,因为必须使用物体类间通用的一般特征,而不能为某个特定类别定义特征,并且该特征需要尽可能的表达出类内共性和类间差异,它必须能处理多

类分类及增量学习,在此前提下无法使用给定类别的海量样本进行训练.

目前一般物体识别主要是基于物体图像的局部特征描述子实现的,尤其是尺度不变特征变换 SIFT 在物体识别等领域具有良好的性能.但是基于二维图像识别丢失了实际物体的三维信息,且容易受到光照等外界条件的影响.为了解决这个问题,许多研究者将 SIFT 扩展至高维空间以实现三维模型的描述.另外,也有学者提出了许多基于点云模型的三维特征,R Rusu 等<sup>[1]</sup>针对点云对齐配准提出点特征直方图(Point Feature Histogram,

PFH),并在文献[2]中提出该描述子的快速算法(Fast Point Feature Histogram, FPFH),PFH还有一个扩展实现形式 PFHRGB,该特征描述子在 PFH的基础上加入了颜色特征信息;S Lazebnik 等<sup>[3]</sup>提出旋转不变特征描述子,是 SIFT 特征在点云中的简单化实现,具有旋转不变性. PFHRGB 识别效果较好,但由于是点对点的运算并且特征向量维数大,耗时过长;FPFH 速度上有很大的提高,但是以识别率的降低为代价<sup>[4]</sup>.点云模型仅依赖于物体的几何形状,与物体的亮度和反射等特性无关,不存在阴影或表面投影等问题,所以基于物体点云模型识别物体的过程更为容易.

在识别目标类内差异大、类间相似度高的情况下,基于单一的特征的物体识别算法往往会产生错误的识别结果.为了解决这个问题,很多研究者提出了基于多特征融合的目标识别方法,在人脸识别、飞机目标识别、物体识别等方面有广泛应用<sup>[5~10]</sup>.Bo<sup>[7]</sup>等提取目标物体的梯度、形状和颜色分层核描述子,并将3种特征融合用于物体识别;并且在文献[8]中将核描述子方法扩展到 Depth 图,提取多种深度核特征描述子并将其融合进行目标识别.Janoch<sup>[9]</sup>和 Mirdanias<sup>[10]</sup>等分别提取灰度图和深度图中目标物体的 HOG(Histogram of Gradient)特征和 SIFT、SURF 特征来实现目标识别;该类算法融合了多种特征信息,与直接利用单一特征的识别方法相比识别率有一定的提升.然而,上述方法都没有提出新的三维特征描述子表示物体的 3D 局部几何特性,而是直接利用二维图像的特征提取方法来提取三维深度图中的特征,识别效果有所影响.

综合以上分析,本文提出了一种基于 2D 和 3D SIFT 特征级融合的一般物体识别算法.其创新性在于:(1)将 3D SIFT 特征描述子进行改进,应用于点云模型特征表示中,实现物体点云模型特征提取与表示;(2)将改进 3D SIFT 应用于一般物体识别中,实现一般物体识别功能;(3)将 2D 和 3D SIFT 特征进行特征融合,实现了基于多特征融合的一般物体识别算法,解决了单一特征识别率低下的问题.

## 2 基于 2D 和 3D SIFT 特征融合的一般物体识别算法框架

为了改善一般物体识别性能,本文提出一种基于 2D SIFT(基于图像)和 3D SIFT(基于点云模型)特征融合的一般物体识别算法,主要包括:特征提取与表示、物体 BoW 模型、特征融合和分类器设计.

### 2.1 特征提取与表示

特征提取与表示是物体识别的基础,怎么提取稳定而有效的特征是特征提取研究中的重点和难点,选择好的特征才能在有限的训练样本条件下得到最好的

识别结果.一般物体数量众多,不可能为每一个物体建立模型库,同时,每一类各个物体形状颜色等差异也很大,不同类间也存在相似度很高的情况,所以提取的物体特征必须满足以下条件:(1)使类间差异最大,即最能表征每类物体区别于其他类物体的特征;(2)使类内差异最小,即最能表征每类物体的共性特征.这就需要在一定的语义层次上对每类物体进行抽象和合理表达,用有限的训练物体数量来表征该类物体.本文提出基于点云模型的 3D SIFT 特征,与图像的 2D SIFT 特征一起作为物体特征,实现物体识别.

真实环境中的目标物体存在旋转、缩放、平移、视角变化、局部遮挡等情况,给一般物体识别带来很大的挑战.SIFT 特征<sup>[11]</sup>描述子是图像的局部特征描述子,其对旋转、尺度缩放、亮度变化保持不变性,对视角变化、仿射变换、噪声也保持一定的稳定性,可以在一定程度上解决上述问题.

由于二维图像丢失了重要的三维信息以及容易受到光照等外界条件影响,所以本文将 SIFT 扩展至 3D SIFT,继承了 2D SIFT 的以上特点,同时由于深度信息的加入,使 3D SIFT 描述子能更加准确的描述物体的局部空间关系.很多学者已经将 3D SIFT 应用到许多领域,但对物体空间方位的考虑是不全面的,仅考虑到了方位角和俯仰角,Allaire<sup>[12]</sup>加入倾斜角.本文将其改进应用于点云模型中,结合点云特性,本文将倾斜角替换为点云在关键点处的法向量方向.本文提出的 3D SIFT 特征提取算法主要步骤如下:

#### (1) 关键点检测

物体的点云模型中点坐标可以表示为  $p(x, y, z)$ ,为实现尺度不变性,定义 3D 点云的尺度空间为  $L(x, y, z, \sigma)$ ,由一个变化尺度的高斯核函数  $G(x, y, z, \sigma)$ 与输入点云  $p(x, y, z)$ 卷积得到:

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

其中  $\sigma = \sqrt[3]{2}$  是尺度空间因子,三维高斯核函数为:

$$G(x, y, z, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^3} e^{-(x^2+y^2+z^2)/2\sigma^2} \quad (2)$$

利用乘法因子  $k$  得到不同尺度,构建点云高斯金字塔.若每组金字塔组内层数为  $s$ ,则设置  $k^s = 2$ .类似于 2D SIFT,用更高效 DoG 算子代替尺度归一化的高斯拉普拉斯算子  $\sigma^2 \nabla^2 G$ ,进行极值检测. DoG 算子计算公式如下所示:

$$D(x, y, z, k^i\sigma) = L(x, y, z, k^{i+1}\sigma) - L(x, y, z, k^i\sigma) \quad (3)$$

其中,  $i \in [0, s+2]$ .关键点是由 DoG 空间的局部极值点组成的<sup>[4]</sup>,为了寻找 DoG 函数的极值点,每一个点要和它所有的相邻点比较,判断其是否为该邻域范围内

的极大值或极小值. 中间的检测点不仅要和它同尺度的 26 个点比较, 还要和上下相邻尺度对应的  $27 \times 2 = 54$  个点比较, 只有当其大于或者小于所有 80 个邻域点时, 该点被检测为极值点. 同样的, 并不是所有的极值点都是稳定的关键点, 设置阈值  $\tau = 0.5$ , 凡是小于该阈值的极值点都是需要被剔除的低对比度的关键点, 这样检测出来的极值点即为需要的关键点.

### (2) 关键点方向分配

为了使描述子具有旋转不变性, 需要利用点云的局部特征为每一个关键点分配一个基准方向. 在 2D SIFT 中, Lowe<sup>[11]</sup> 利用关键点邻域灰度梯度直方图统计关键点主方向. 为实现旋转不变性, 考虑到点云中邻域点之间的空间关系, 本文统计关键点  $k$  邻域点到邻域中心点角度的二维直方图, 选择中心点而不是关键点的原因是更能表示关键点  $k$  邻域内局部整体特性. 根据式(4)计算关键点  $p$  及其  $k$  邻域点  $(x_i, y_i, z_i)$  ( $i = 1, 2, \dots, k, k+1$ ) 到邻域中心点  $(x_c, y_c, z_c)$  的向量幅值  $m(x, y, z)$ 、角度  $\theta(x, y, z)$  (方位角) 和角度  $\varphi(x, y, z)$  (俯仰角).

$$\begin{cases} m(x, y, z) = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2 + (z_i - z_c)^2} \\ \theta(x, y, z) = \tan^{-1}((y_i - y_c)/(x_i - x_c)) \\ \varphi(x, y, z) = \sin^{-1}((z_i - z_c)/m(x, y, z)) \end{cases} \quad (4)$$

为了确定关键点主方位, 利用直方图统计关键点邻域内的角度  $\theta(x, y, z)$  和  $\varphi(x, y, z)$ , 将幅值  $m(x, y, z)$  作为权值, 对直方图进行高斯加权  $e^{-(2m(x, y, z)/R_{\max})^2}$ ,  $R_{\max}$  表示关键点邻域最大范围, 忽略超过该距离的点. 然后利用高斯滤波器 (Gaussian filter) 对直方图进行平滑滤波, 以消除 (降低) 噪声影响. 直方图的峰值代表了该关键点邻域的方向, 将该方向作为关键点主方向, 为了增强匹配的鲁棒性, 只保留峰值大于主方向峰值 80% 的方向作为该关键点的辅助方向. 定义  $(\theta, \varphi)$  对应主方向为  $(\alpha, \beta)$ . 至此, 检测出的含有位置、尺度和方向的关键点即是该点云的 3D SIFT 特征点.

### (3) 关键点描述

通过以上步骤, 对于每一个关键点  $p$ , 拥有三个信息: 位置  $\mathbf{p} \triangleq [x_p, y_p, z_p]^t$ 、尺度  $\sigma_p$  以及方位  $(\alpha_p, \beta_p)$ . 接下来就是为每个关键点  $p$  建立一个描述子, 用一组向量将这个关键点描述出来, 使其不随各种变化而改变, 比如光照变化、视角变化等. 这个描述子不但包括关键点, 也包含关键点周围对其有贡献的点, 并且描述符应该有较高的独特性, 以便于提高关键点正确匹配的概率. 2D SIFT 描述子是关键点邻域高斯图像梯度统计结果的一种表示, 对于三维点云模型, 本文统计关键点邻域局部空间关系, 通过计算关键点邻域内各角度直方

图, 生成 3D SIFT 特征向量, 表示关键点及其邻域局部特性. 表面法线是几何体表面的重要属性, 法向量分布可以表达物体表面的 3D 几何特征, 所以本文在计算 3D SIFT 特征向量时加入法向量, 更加全面的表达物体的局部空间特征. 关键点描述子生成过程如下: 计算关键点  $p$  的  $k$  邻域点  $p_i$  ( $i = 1, 2, \dots, k$ ), 对应坐标为  $\mathbf{p}_i \triangleq [x_i, y_i, z_i]$ . 为了保证旋转不变性, 将关键点  $p$  的  $k$  邻域点  $p_i$  旋转至关键点主方位, 邻域点坐标变换公式为:

$$\mathbf{p}'_i = \begin{bmatrix} \cos\alpha_p \cos\beta_p & -\sin\alpha_p & -\cos\alpha_p \sin\beta_p \\ \sin\alpha_p \cos\beta_p & \cos\alpha_p & -\sin\alpha_p \sin\beta_p \\ \sin\beta_p & 0 & \cos\beta_p \end{bmatrix} \cdot \mathbf{p}_i \quad (5)$$

其中  $\mathbf{p}'_i$  为  $\mathbf{p}_i$  旋转后对应的点坐标. 然后利用开放点云库 (Point Cloud Library, PCL) 提供的方法计算旋转后关键点  $p$  的  $k$  邻域在点  $p$  处的法向量  $\mathbf{n}$ ; 对于关键点  $p$  邻域内的任意一个点  $p'_i$ , 利用式(4)计算向量  $\mathbf{pp}'_i$  的幅值  $m$  和两个角度  $\theta$  及  $\varphi$ , 同时计算法向量  $\mathbf{n}$  和向量  $\mathbf{pp}'_i$  夹角  $\delta$ .

$$\delta = \cos^{-1} \left( \frac{\mathbf{pp}'_i \cdot \mathbf{n}}{|\mathbf{pp}'_i| \cdot |\mathbf{n}|} \right) \quad (6)$$

对于关键点  $p$  及其邻域点  $p'_i$  可以用得到的 4 元组  $(m, \theta, \varphi, \delta)$  表示, 将关键点邻域分为  $n \times n$  的子区域, 分别统计子区域中所有邻域点在各个方向的角度分布, 得到关键点  $p$  处的特征描述子. 在文献[11]中, Lowe 取  $n = 4$ . 为了减少特征描述子计算时间, 本文将关键点邻域作为一个整体, 即  $n = 1$ . 统计关键点  $p$  邻域内角度分布直方图, 按照  $45^\circ$  一个柱, 分别将  $(\theta, \varphi, \delta)$  分为 8 个、4 个和 4 个子区间, 并统计落在每个子区间的点数目, 同时将幅值  $m$  作为权值, 统计区间点数目时进行高斯加权  $e^{-(2m/R_{\max})^2}$ , 这样就得到了一个  $n \times n \times 4 \times 4 \times 8 = 128$  维的特征向量. 最后对每个关键点生成的特征描述子进行归一化处理.

至此, 完整的 2D 和 3D SIFT 特征描述子生成.

## 2.2 物体 BoW 模型

本文采用物体识别中经典的统计学模型 BoW 模型<sup>[13]</sup> 统计物体特征向量分布, 用一个多维向量实现物体描述, 不同的是本文利用 KMeans++ 聚类算法<sup>[14]</sup> 得到物体视觉单词库. 与 KMeans 聚类算法相比, KMeans++ 聚类算法对初始聚类中心进行了改进, 使得算法无论在聚类结果准确度还是运行时间上都有提升. KMeans++ 聚类算法初始聚类中心的选取的基本思想是: 初始的聚类中心之间的相互距离要尽可能的远. 初始聚类中心选择步骤如下所示:

(1) 聚类中心集合记为  $center$ , 从输入的向量集合  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{n-1}, \mathbf{x}_n\}$  中随机选择一个向量  $\mathbf{x}_i \in \mathbf{X}$  作为第一个聚类中心.

(2) 对于满足  $\{\mathbf{x}_j | \mathbf{x}_j \in \mathcal{D}_{\mathbf{x}center}\}$  的任意向量, 计算其

与最近聚类中心(指已选择的聚类中心)的距离 $D(x_j)^2$ .

(3)选择一个向量作为新的聚类中心,每个向量被选为聚类中心的概率 $P(x_j)$ 由式(7)计算得到, $P(x_j)$ 最大时对应的向量即新的聚类中心

$$P(x_j) = D(x_j)^2 / \sum_{x_j \in \vartheta_{x,center}} D(x_j)^2 \quad (7)$$

(4)重复(2)和(3)直到 $K$ 个初始聚类中心被选出来,得到 $K$ 个初始聚类中心后执行标准的KMeans算法.通过选取不同 $K$ 值进行试验对比,本文取 $K=300$ .利用KMeans++聚类算法得到的 $K$ 个聚类中心称为视觉单词库,然后根据每个特征描述子到视觉单词库的欧式距离将该描述子映射到距离最短的视觉单词中,统计每个视觉单词在特征描述子中出现的频数,得到2D和3D SIFT对应的BoW特征向量.所以,利用BoW模型,每个物体由两个 $1 \times 300$ 的BoW特征向量描述,分别记为 $BoW_{2D}$ 和 $BoW_{3D}$ .

### 2.3 特征级融合

对于任意样本 $O_\xi \in O$ , $O$ 是样本空间,其对应的2D和3D SIFT特征描述子利用BoW模型得到特征向量为 $BoW_{2D}$ 和 $BoW_{3D}$ ,则 $O_\xi$ 的融合特征为 $BoW_{2D,3D} \triangleq [BoW_{2D}, BoW_{3D}]$ ,利用该特征实现物体描述.

(1)提取和描述待识别物体的2D和3D SIFT特征描述子,其对应的2D和3D SIFT特征描述子集为 $m_{2D} \times 128$ 和 $m_{3D} \times 128$ , $m_{2D}$ 和 $m_{3D}$ 分别为样本2D和3D SIFT

关键点个数,利用BoW模型统计物体特征描述子分布,得到物体特征向量,分别表示为 $BoW_{2D}$ 和 $BoW_{3D}$ .

(2)对两个特征向量进行特征级融合,形成新的特征向量 $BoW_{2D,3D}$ ,实现物体描述.

(3)将新的特征向量输入所有1V1 SVM判别函数得到对应的判别结果,通过投票得到该物体属于第 $i$ 类的概率记为 $P(i)$ , $i \in [1, n]$ , $n$ 表示物体总类数.

(4)通过概率最大的值判定待识别物体对应的类class,数学公式为:

$$class = \arg \max_{1 \leq i \leq n} \{P(i)\} \quad (8)$$

### 2.4 分类器设计

物体描述完成后,利用SVM学习目标类型并实现目标分类.SVM在解决小样本、非线性及高维模式识别问题中表现出许多特有的优势,在许多领域都取得了成功的应用.SVM的关键在于核函数,RBF核是应用最广泛的核函数,但是特征向量维数特别大时,线性核表现出较好的性能.SVM是典型的两类分类器,而更多时候需要实现的是多类分类问题,本文通过训练多个两类分类器的方法构建多类分类器解决上述问题.

### 3 算法流程

基于2D和3D SIFT特征级融合的一般物体识别算法流程,如图1所示.

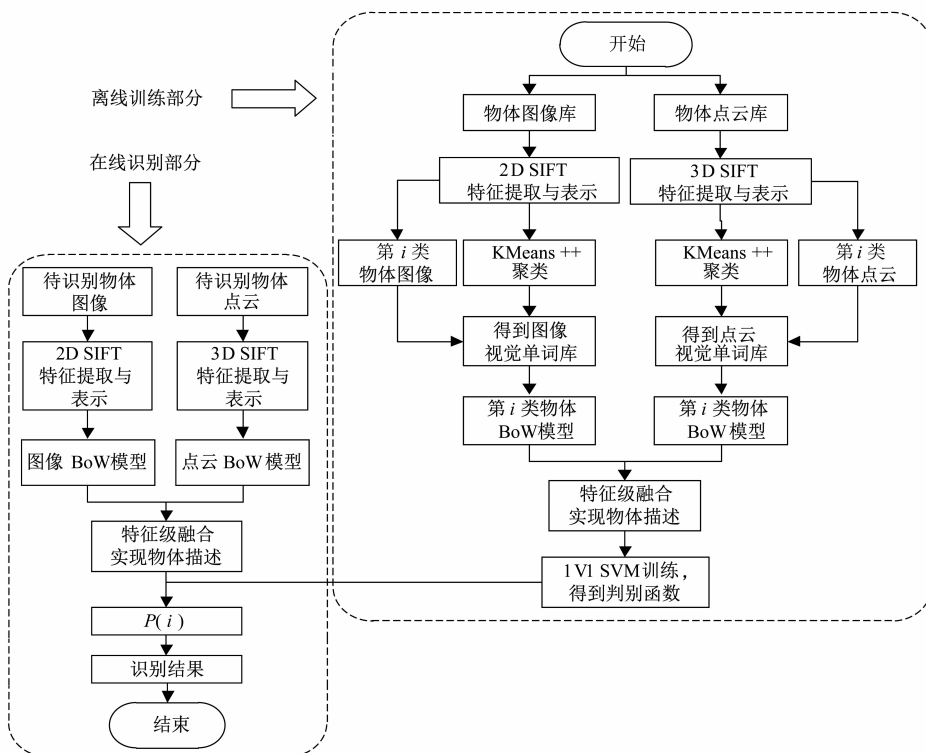


图1 基于2D和3D SIFT特征级融合的一般物体识别算法流程图

## 4 实验结果

本文实验采用的点云模型和 RGB 图像来自于 K lai 等人建立的大型点云数据库<sup>[15]</sup>,该数据库包含 51 类共 300 个物体的点云模型和 RGB 图像,每一个物体点云和图像都包含 3 个视角.实验方法:随机选取每一类中的一个物体作为测试样本,剩余物体则作为训练样本,每类训练样本选择 100 个,测试样本为 60 个,均从数据库中随机抽取.为了评估本文提出算法的性能,该部分进行了多个实验,统计多种情况下正确识别率,正确识别率计算方法:

$$RR = n_r / N \quad (9)$$

其中,  $RR$  表示物体正确识别率 (Recognition Rate,  $RR$ ),  $n_r$  表示测试样本中正确识别个数,  $N$  表示总测试样本个数.

### 实验 1 3D SIFT 正确识别率

本实验选取类内差异明显、类间相似度高的 6 类物体进行实验,分别为 apple, tomato, banana, pitcher, cereal\_box, kleenex. 在本实验中,先对 6 类训练样本进行训练,然后用测试样本进行测试.在已有的众多点云特征中,PFHRGB 和 PFH 是识别率较好的特征<sup>[4]</sup>.为了验证本文提出的 3D SIFT 特征在物体识别中的优势,本文在同样的条件下进行了 3 种特征描述子的对比试验,每种特征描述子都采用 SIFTKeypoint 模块<sup>[17]</sup>检测关键点,然后分别计算关键点处不同特征描述子的特征向量,统计正确识别率,实验结果见表 1.

表 1 各描述子正确识别率

特征描述子	正确识别率 (%)
PFH	81.39
PFHRGB	84.17
3D SIFT	91.11

实验结果分析:PFHRGB 将颜色信息融入到 PFH 特征描述子,由实验结果可知,颜色信息的引入丰富了特征信息,提升了物体正确识别率.由表 1 可知,本文提出的 3D SIFT 特征描述子在识别率方面比 PFH 和 PFHRGB 分别提升了 9.72% 和 6.94%,验证了 3D SIFT 特征描述子在基于点云模型的一般物体识别中的有效性.

### 实验 2 基于 2D 和 3D SIFT 特征级融合的正确识别率

为了克服单一特征在类间相似度高、类内差异明显的情况下往往不能给出正确的识别结果这一问题,本文提出了基于 2D 和 3D SIFT 特征级融合的一般物体识别方法.为了验证本文提出融合算法的优越性,实验 2 在同样的条件下对比了 2D SIFT、3D SIFT 及两者特征级融合的正确识别率,训练样本和测试样本与实验 1 相同,实验结果见表 2.

表 2 特征级融合算法正确识别率

特征描述子	正确识别率 (%)
2D SIFT	88.06
3D SIFT	91.11
2D + 3D SIFT	93.06

实验结果分析:为了表示方便,用 2D + 3D SIFT 表示 2D 和 3D SIFT 特征级融合.由表 2 可知,与 2D SIFT 相比,3D SIFT 识别率提高了 3.05%,可见深度信息的引入有益于实现物体识别.由于物体的多变性,单一特征提供的信息存在不精确、不确定和不完全的问题,使得单一特征算法识别率较低,2D 和 3D SIFT 经过特征级融合后的识别率为 93.06%,比单一特征描述子的识别率有一定提升,说明本文提出的一般物体识别算法在识别率方面有明显的优势.

### 实验 3 多种特征融合算法正确识别率

本实验给出了多种融合算法识别结果,选取类内差异明显、类间相似度高的 10 类物体,对 2-10 类进行识别实验,分别为 apple, tomato, banana, pitcher, cereal\_box, kleenex, camera, coffee\_mug, calculator, cell\_phone. 共对比了 4 种融合算法的正确识别率,分别为:特征级融合及决策级融合中的平均加权融合、DSmT 理论<sup>[17,18]</sup>和 Murphy 规则<sup>[19]</sup>.实验结果见图 2 所示,ave 表示平均加权融合,横坐标表示类别数,比如,“6”表示该次实验共包括 6 类物体,统计这 6 类物体的正确识别率.由图 2 可知,物体类别数增加时,①特征融合算法比单一特征算法有更高的正确识别率和更强的鲁棒性.4 种融合算法中,平均加权融合和 DSmT 理论融合结果比较接近,低于另外两种融合方法;与特征级融合相比,将 2D 和 3D SIFT 及两者特征级融合的结果共 3 个证据源按照 Murphy 规则融合后的结果并没有提高,所以本文采用特征级融合的方法完成一般物体识别的任务;②本文提出的 3D SIFT 特征描述子相对 PFHRGB 和 2D SIFT 特

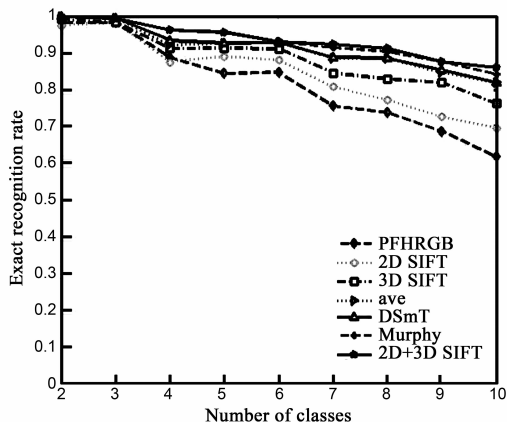


图 2 特征融合方法正确识别率

征描述子有更好的识别效果;③每种识别算法识别率都有所下降,部分原因在于分类器的设计,本文采用的多类分类器是由多个 1V1 SVM 分类器构造的,每一个分类器的误差都会累积到最终投票结果中.随着物体类别数的增加,1V1 SVM 分类器个数迅速增加,如 10 类物体共有 45 个分类器,45 个分类器的判断失误累加到最终投票结果中将会造成很大程度上的识别错误.

#### 实验 4 时间复杂度

在 i7-3770@3.4GHz CPU,64 位 Win7 操作系统的实验平台下,本实验统计了利用不同特征描述子完成识别过程耗费的时间,与实验 1 测试样本相同,平均计算一个物体识别所需时间,实验结果见表 3,与 PFHRGB 相比,本文提出的 3D SIFT 识别算法及融合 2D 和 3D SIFT 的识别算法平均每个测试样本耗时减少了 34.75% 和 22.01%,提高了基于点云模型的识别算法性能.

表 3 不同特征描述子时间比较

特征描述子	360 个测试样本时间 (s)	平均每个测试样本时间 (s)
PFHRGB	3404.628	9.4573
3D SIFT	2221.608	6.1711
2D + 3D SIFT	2653.272	7.3702

## 5 结束语

一般物体种类繁多、变化巨大,给一般物体识别带来很大的困难,物体特征的提取与表示是一般物体识别中的重点和难点.本文针对这一问题,提出了一种基于 2D 和 3D SIFT 特征级融合的一般物体识别方法,取得较好结果.但是本文仍然有许多方面可以改进:(1)提出的算法目前实时性不够,主要是特征提取算法部分耗时较多,需要在保证识别率的基础上,缩短识别时间;(2)分类器的改进对识别率的提升也有一定的帮助.这将是我们的下一步的研究任务.

#### 参考文献

[1] Rusu R B, Blodow N, Marton Z C, Beetz M. Aligning point cloud views using persistent feature histograms[A]. Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems[C]. Nice, France: IEEE, 2008. 3384 – 3391.

[2] Rusu R B, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration[A]. Proceedings of IEEE International Conference on Robotics and Automation[C]. Kobe, Japan: IEEE, 2009. 3212 – 3217.

[3] Lazebnik S, Schmid C, Ponce J. A sparse texture representation using local affine regions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(8): 1265 – 1278.

[4] Alexandre L A. 3D descriptors for object and category recogni-

tion: a comparative evaluation[A]. Proceedings of IEEE International Conference on Intelligent Robotic Systems (Vol Workshop on Color-Depth Camera Fusion in Robotics)[C]. Vilamoura, Portugal: IEEE, 2012. 1 – 6.

[5] 陈北京,王蔚,宋加涛,任小波.一种融合二值边缘特征和灰度特征的人脸识别方法[J].电子学报,2009,37(6): 1180 – 1184.

Chen Bei-jing, Wang Wei, Song Jia-tao, Ren Xiao-bo. A novel face recognition method based on the fusion of binary edge and grayscale features [J]. Acta Electronica Sinica, 2009, 37(6): 1180 – 1184. (in Chinese)

[6] 李新德,潘锦东,Dezert Jean.一种基于 DSMT 和 HMM 的序列飞机目标识别算法[J].自动化学报,2014,40(12): 2862 – 2876.

Li Xin-de, Pan Jin-dong, Dezert Jean. A target recognition algorithm for sequential aircraft based on DSMT and HMM[J]. Acta Automatica Sinica, 2014, 40(12): 2862 – 2876. (in Chinese)

[7] Bo L, Lai K, Ren X, Fox D. Object recognition with hierarchical kernel descriptors[A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition[C]. CO, USA: Colorado Springs, 2011. 1729 – 1736.

[8] Bo L, Ren X, Fox D. Depth kernel descriptors for object recognition[A]. Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems[C]. San Francisco, CA, USA: IEEE, 2011. 821 – 826.

[9] Janoch A, Karayev S, Jia Y, Barron J T, Fritz M, Saenko K, Darrell T. A category-level 3-D object dataset: putting the kinect to work[A]. 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops) [C]. Barcelona, Spanish: IEEE, 2011. 1168 – 1174.

[10] Mirdanias M, Prihatmanto A S, Rijanto E. Object recognition system in remote controlled weapon station using SIFT and SURF methods[J]. Mechatronics, Electrical Power, and Vehicular Technology, 2013, 4(2): 99 – 108.

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

[12] Allaire S, Kim J J, Breen S L, Jaffray D A, Pekar V. Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis[A]. Proceedings of 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops[C]. Anchorage, AK: IEEE, 2008. 1 – 8.

[13] Sivic J, Zisserman A. Video Google: a text retrieval approach to object matching in videos[A]. Proceedings of 9th IEEE International Conference on Computer Vision[C]. Nice, France: IEEE, 2003. 1470 – 1477.

[14] Arthur D, Vassilvitskii S. k-means + + : The advantages of

careful seeding [A]. Proceedings of the 18th annual ACM-SIAM symposium on discrete algorithms [C]. PA, USA: Society for Industrial and Applied Mathematics, 2007. 1027 - 1035.

- [15] Lai K, Bo L, Ren X, Fox D. RGB-D dataset [DB]. <http://rgbd-dataset.cs.washington.edu/dataset.html>, 2011-03-05.
- [16] Rusu R B, Cousins S. 3D is here: point cloud library (pcl) [A]. Proceedings of 2011 IEEE International Conference on Robotics and Automation [C]. Shanghai, China; IEEE, 2011. 1 - 4.
- [17] Dezert J. Foundations for a new theory of plausible and paradoxical reasoning [J]. Information and Security, 2002, 9: 13 - 57.
- [18] 李新德, Dezert Jean, 黄心汉, 孟正大, 吴雪建. 一种快速分层递阶 DS<sub>m</sub>T 近似推理融合方法 (A) [J]. 电子学报, 2010, 38(11): 2566 - 2572.  
Li Xin-de, Dezert Jean, Huang Xin-han, Meng Zheng-da, Wu Xue-jian. A fast approximate reasoning method in hierarchical DS<sub>m</sub>T(A) [J]. Acta Electronica Sinica, 2010, 38(11): 2566 - 2572. (in Chinese)

- [19] Murphy C K. Combining belief functions when evidence conflicts [J]. Decision Support Systems, 2000, 29(1): 1 - 9.

### 作者简介



李新德(通信作者) 男, 1975年9月出生, 山东海阳人. 东南大学自动化学院副教授、博士生导师. 主要研究方向为智能机器人、人机交互、机器感知、信息融合、不确定推理和机器视觉.  
E-mail: xindeli@seu.edu.cn



刘苗苗 女, 1990年2月出生, 河南安阳人. 东南大学自动化学院硕士研究生. 主要研究方向为模式识别与分类.  
E-mail: 374561475@qq.com