

SIFT 特征在三维物体识别中的应用

熊英 马慧敏

(清华大学电子工程系, 图象图形研究所, 三维图象仿真研究室, 北京, 100084)

摘要 SIFT 算法因其良好的尺度不变性、旋转不变性、抗噪声能力强、受光照变化影响小等优点, 在当今图形图像学的很多领域得到了广泛的应用。该算法所运用的核心特征 (SIFT 特征) 是一种基于局部梯度的向量特征, 能够抵抗图像的伸缩、旋转等多种变化, 非常好地满足了三维物体识别的实际需要。将 SIFT 特征应用在三维物体识别系统中的视点空间划分、背景物体分割、模式特征匹配等模块, 可以有效地提高系统的识别速度与效率, 增强系统的稳定性。本文通过理论分析与实验验证, 讨论了将 SIFT 特征应用于三维物体识别系统中的可行性与优势。

关键词 SIFT 算法 三维物体识别 模式匹配 视点空间划分

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2009)-

Application of SIFT Feature in 3D Object Recognition

Xiong Ying, Ma Hui-min

(Institute of Image and Graphics, Department of Electronic Engineering, Tsinghua University, Beijing, 100084)

Abstract Because of its invariance to image scaling, rotation, noise and illumination changes, SIFT algorithm is nowadays widely applied in many aspects of the image study. The feature used by this algorithm (SIFT feature) is a vector based on local gradient, which can resist a lot of image variation such as extension, compression and rotation, meeting the practical requirements of 3D object recognition. Applying SIFT feature in view space partition, cutting the object from its background and pattern matching can effectively enhance the robustness of the system and improve its speed and efficiency. With theoretical analysis and experimental verification, we prove the feasibility and advantages of applying SIFT feature in 3D object recognition system.

Keywords SIFT algorithm, 3D object recognition, pattern match, view space partition

1. 引言

出于在自动导航、自动检测、机器人抓取物件及装配任务等许多实际需要, 三维物体识别如今已成为一个非常活跃的研究领域。三维物体识别的一

个难点在于克服因视点位置改变而引起的目标姿态变化。这种姿态变化一般可以用一个投影变换来描述。但投影变换的非线性极大地增加了运算的复杂度, 因此在实际中往往采用仿射变换来代替投影变换。于是三维物体识别的一个关键问题就是找到物体的仿射不变性特征, 并以此进行目标的匹配识别。在小角度近似下, 仿射不变性又可以归结为尺度不变性和在平面内的旋转不变性。

SIFT(Scale-Invariant Feature Transform) 算法由

基金项目: 国家自然科学基金项目 (60502013); 国家高技术研究发展计划 “863” 项目 (2006AA01Z115)

收稿日期:2009-; **改回日期:**2009-

第一作者简介:熊英(1987年3月), 男, 现就读于清华大学电子工程系。Email: xiongying0007@gmail.com

David G. Lowe 于 1999 年提出^[1]，并于 2004 年完善总结^[2]。该算法通过抽取图像的局部不变性特征 (SIFT 特征) 进行图像识别匹配。由于具有良好的尺度不变性和旋转不变性、抗噪声能力强等诸多优点，该算法在许多领域得到了广泛的应用，如图像拼接、模式识别等。

目前 SIFT 算法在三维物体识别领域的应用并不多见，但实际上，该算法能够有效抵抗图像伸缩、旋转、噪声干扰及光照变化等优点恰恰能够满足三维物体识别中的一些实际需要。虽然在大角度下 SIFT 特征的仿射不变性会受到一些限制，但通过合理划分视点空间可以保证算法有较高的识别准确率。另外，本文将证明对三维模型做充分的预处理后，系统识别速度也可以大幅提高。

2. 三维物体识别系统结构

三维物体识别在当今计算机视觉中已成为一个非常活跃的领域，现今已提出较多的识别技术与算法^[3]。一个典型的识别系统大致可分为两个部分，即对三维物体建立特征数据库与对目标图像特征提取 (见图 1)。

在建立三维物体特征数据库过程中，需要完成的核心任务有：

(1) 视点空间划分：选择代表视点及抽取代表视点的特征。其中视点空间划分目前较多采用均匀划分后聚类、突变论两种方法，但前者存在划分不够精细准确的问题，而后者需求的海量计算往往无法实现。

(2) 目标分割：在目标图像特征提取过程中需要将输入图像预处理，如将物体从背景中分割，以便有效消除背景噪声干扰，准确地提取物体特征。

(3) 特征匹配：最后将输入图像特征与物体模型特征进行比对，进行分类决策并给出识别结果。

3. SIFT 特征在视点空间划分中的应用

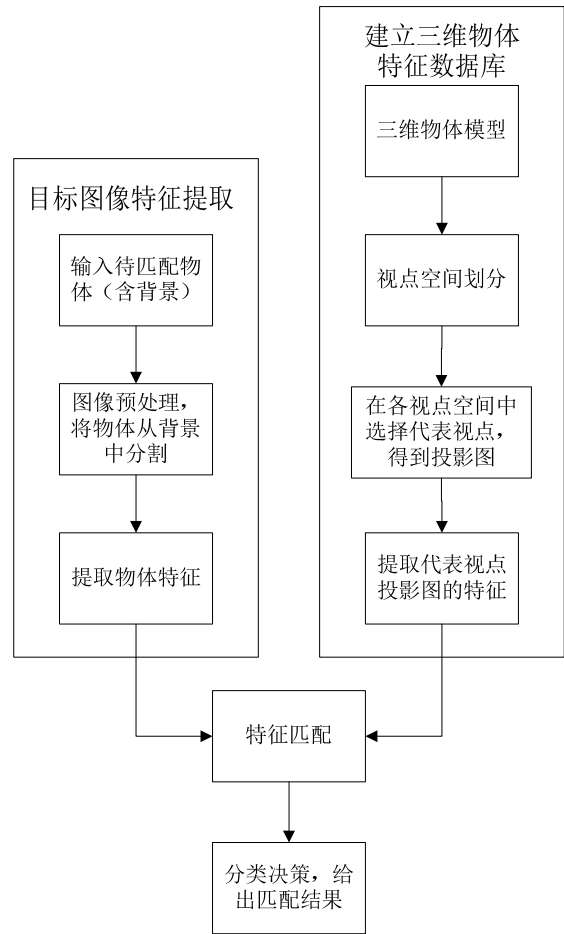


图 1 典型三维物体识别系统结构框图

SIFT 特征可以应用于一种基于聚类思想的视点空间划分。通过采样得到三维物体各个视点下的投影视图，抽取这些视图的 SIFT 特征进行聚类分析。将特征大体相同的视点集聚为一类，用一个代表视点表达该类，本文中我们直接取该类点的公有 SIFT 特征。

首先在高斯球面上均匀且较为密集地采集三维物体的投影视图，计算每个视图的 SIFT 特征。接着对相邻视图应用匹配算法，判断其相似度。根据突变理论，当达到某个临界点 (如 EV 事件或 EEE 事件发生点)，两幅图像的重要结构会发生突变而产生失配。于是可以用这些失配点作为临界点，将临界点之间的部分作为一个聚类。

一个聚类往往有相同或相似的 SIFT 特征，可以将它们提取出来，这样就不需要重新选取代表视点，而且可以通过统计平均方法增大所选取 SIFT 特征的代表性，提高匹配准确度。

在实验中确实可以发现相邻图像匹配点数曲线在某些地方出现极小值 (见图 2)。

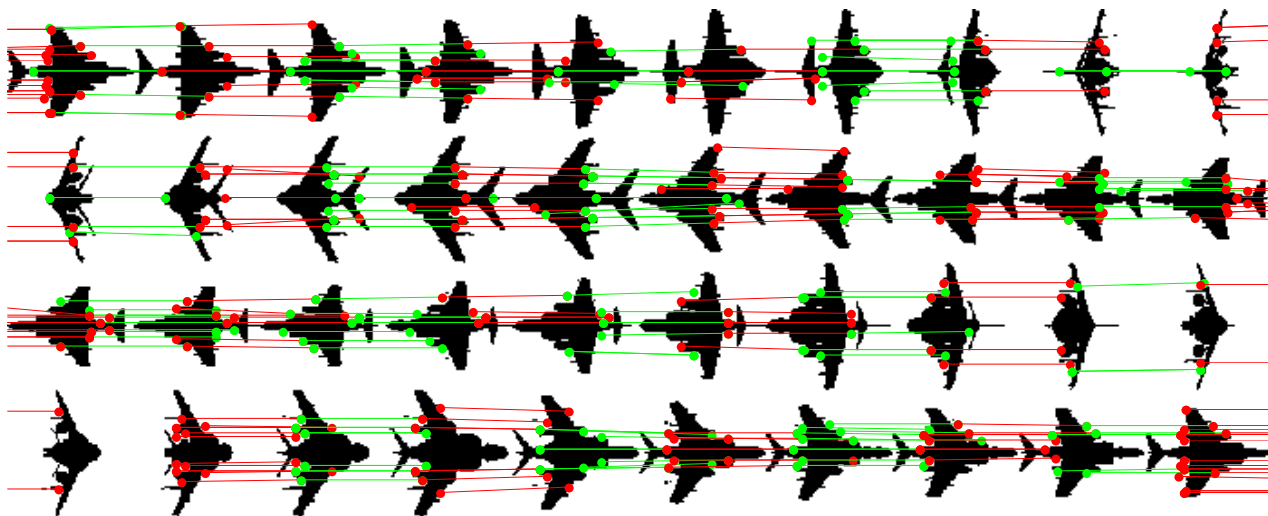


图 2 某飞机的高斯球赤道面上等间隔 40 个投影图的匹配情况

图 3 给出上图飞机在赤道面各角度处具体的匹配情况。

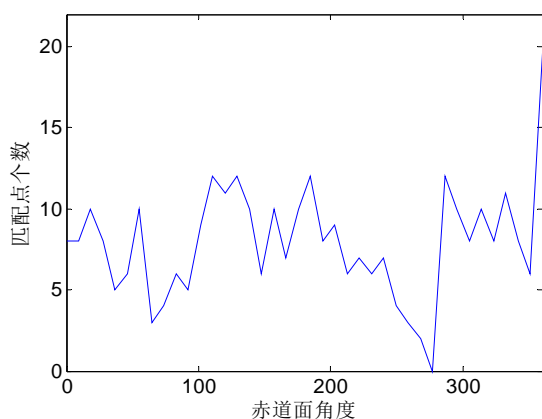


图 3 赤道面各角度处匹配点个数

图 3 中可以看到，在赤道面 280 度左右的地方匹配点数减小至零，对应图 2 第四行开始两个视图适配。经分析可以发现，这个失配点恰是飞机翻过程中的临界点（临界点左侧为飞机的俯视图，右侧为仰视图）。另一个临界点在赤道面 60 度左右的地方，对应图 2 第一行最后两个视图的失配。这样，本例中飞机的赤道面就被分成两段，每段为一个聚类，分别为飞机的俯视图和仰视图。如果需要，还可以更精细地在赤道面上采样，划分出更多（更准确）的聚类。

同样方法在全球面采样处理，可以得到物体在整个高斯球面的视点空间划分。

上述视点空间划分的方法是十分精细准确的，但也存在一个问题，即需要在全高斯球面密集采集视点和投影视图，而且需要对每幅投影视图计算 SIFT 特征，并与其相邻视图进行特征匹配，这个过程需要的计算相对来说是比较烦杂的。但好在上述过程是在建立特征数据库过程中进行的，因此不会影响系统匹配的实时性。事实上，这相当于通过前期大量的后台准备工作以保证输入图像到来时的快速匹配。

应用 SIFT 算法的另一好处在于其数据库占用存储空间非常小，一个三维物体经过预处理后将被转换成几百个 128 维向量。数据库中只需要存储这些 128 维向量和几个代表视点的二维投影图，而其他各个视点的二维投影图则不需要进行存储。

4. SIFT 算法在物体背景分割与特征提取中的作用

在三维物体识别系统中需要将输入图像预处理，以便对其进行特征提取及后续的匹配识别。预处理过程中要解决的一个关键问题是将图像中的物体与背景区分开来，避免背景对识别结果的影响。一般来讲，当不规则物体以任意位置姿态嵌入在背景中时，将其完全从背景中分离出来是比较困难的，至

少需要一定量较复杂的计算,从而会增长处理时间,破坏系统的实时性。

而由于 SIFT 特征描述子的特点使得其在处理物体背景分割与特征提取过程中有一些特别的优势,可以很好地解决上述问题。这些优势主要表现在以下几方面:

首先, SIFT 特征点取自输入图像差分高斯空间的极值点。而对图像的差分高斯运算相当于一个高通滤波器,能够有效滤除背景中的低频分量。也就是说,在纯色背景以及变化缓慢的渐变背景中一般不会出现 SIFT 特征点,进而也就不会影响特征匹配。

另一方面, SIFT 算法在提取特征点时有一个门限参数(Threshold),可以剔除对比度较低的极值点,因为这些极值点往往源自背景噪声。实验证明,通

过适当提高门限参数,可以有效剔除背景中的特征点,而不会影响对物体的特征计算(见图 4)。

最后, SIFT 算法进行的是分层(Octave)计算,大多数背景中的特征点都会出现在尺度层次较低的视图中(对应文章[2]中参数 σ 较小的视图),而物体中的特征点往往出现在高尺度层次的视图中。通过适当地选择计算层次,也可以有效地滤除背景对特征选取的影响。

总之, SIFT 算法的上述一些特点使得其在处理带有背景的输入图像时有相当好的稳定性。在本文的实验部分,我们证实了 SIFT 算法能够较好地解决物体在纯色背景,渐变背景以及白噪声背景中的物体分割及特征提取问题。

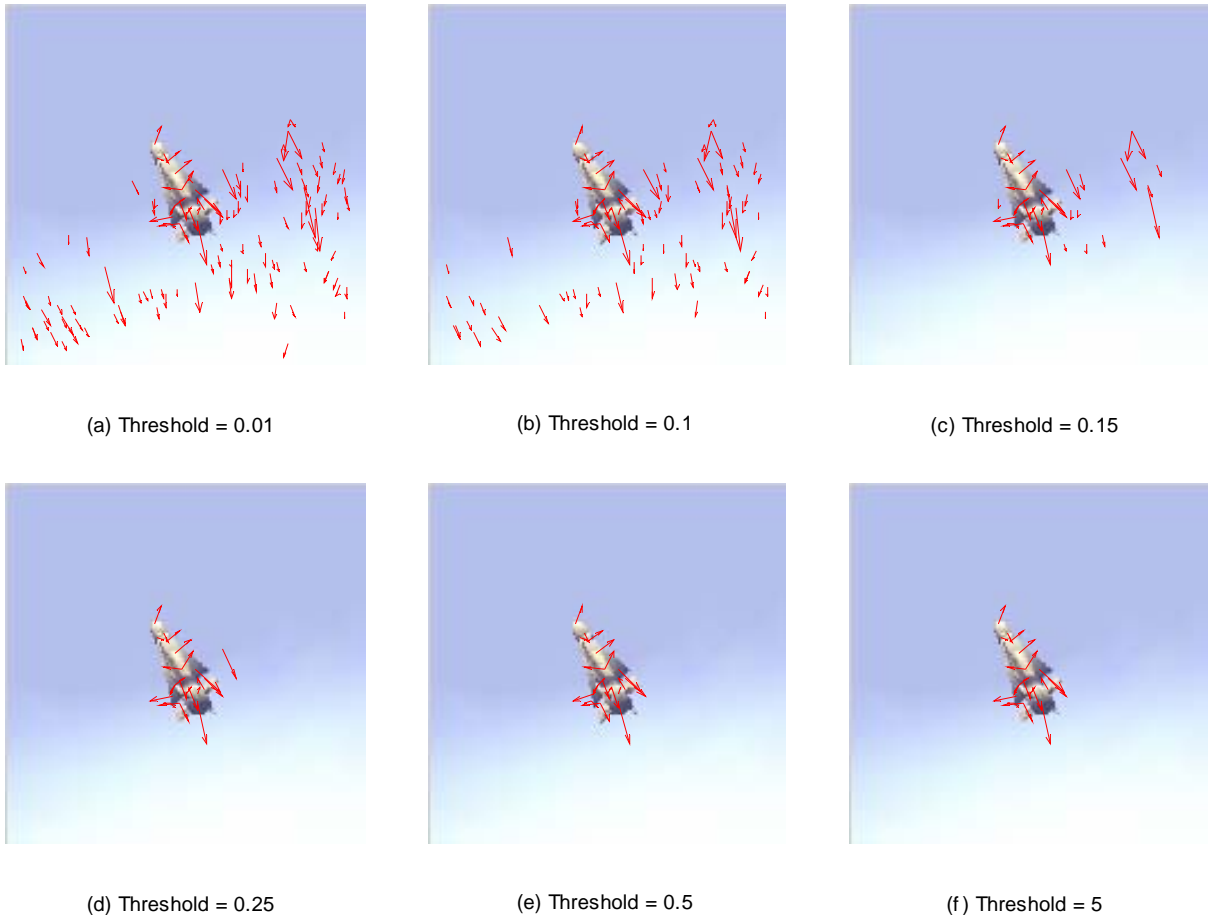


图 4 调节门限参数剔除 SIFT 背景特征点 (红色箭头为特征点及其主方向)

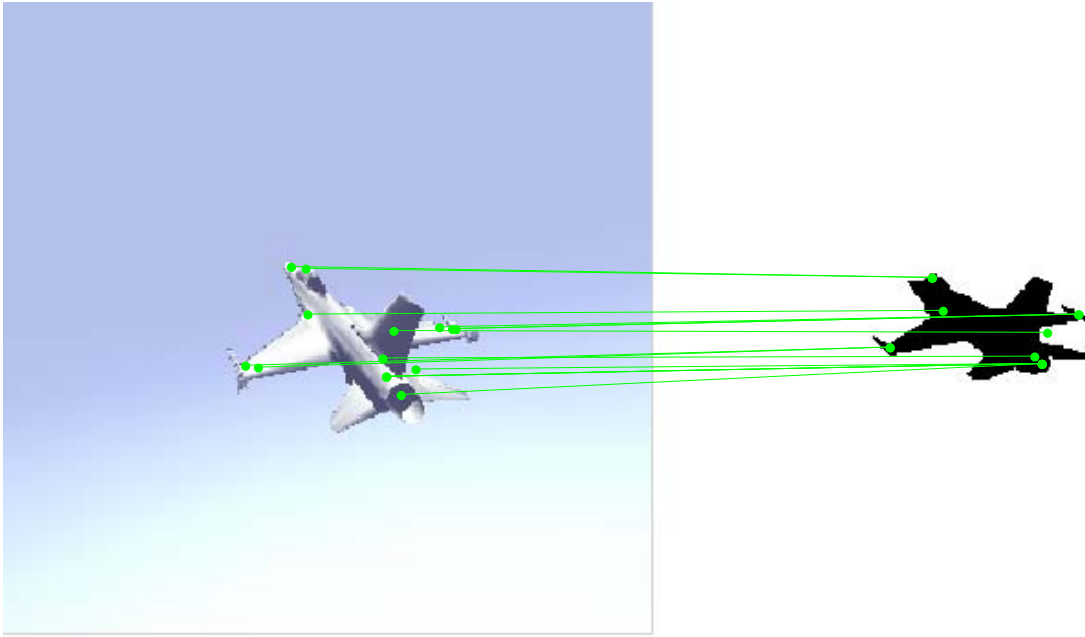


图 5 SIFT 算法在彩色图像和二值图像之间的匹配

5. SIFT 特征匹配

得到待匹配物体的特征后，就可以与数据库中的特征进行比对，通过分类决策得到匹配结果。为了加快匹配速度，还可以对特征数据库建立索引或哈希。由于 SIFT 特征描述子非常规整(128 维向量)，因此建立索引或哈希也相对容易。

SIFT 特征的另一好处在于它是基于局部梯度特征的，因此对尺度、旋转和光照强度都有非常好的不变性。只要局部形态一致，甚至可以进行彩色图像与二值图像之间的匹配（见图 5）。一方面，这可以简化三维物体建模；另一方面，也可以优化输入图像的预处理过程，省去二值化过程。（事实上，图 5 左边图像二值化后将会导致物体和背景融合，无法分离。）

6. 结论

本文分析了目前三维物体识别技术中的主要目标，即通过仿射不变性特征进行识别匹配，讨论了 SIFT 算法在图像模式匹配中的特点，并论述了将其应用在三维物体识别中的可行性与优势。

使用 SIFT 特征作为模式匹配的特征，可以保证识别系统有非常好的尺度不变性与旋转不变性，经过合理的视点空间聚类划分后可以克服仿射不变性限制，达到较高的准确率。在处理将物体从背景分割问题中，通过合理设定门限参数滤除背景中的特征点，就可以直接处理输入图像而不受背景噪声的影响。最后，SIFT 特征规整，占用存储空间小，且易于建立索引和哈希，这些都有助于提高匹配速度。

参考文献

- [1] Lowe, D.G. Object recognition from local scale-invariant features [A]. In: International Conference on Computer Vision[C], Corfu, Greece, 1999: 1150-1157.
- [2] Lowe, D.G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [3] Li Qing, ZHOU Man-li, Liu Jian. A review on 3D

objects recognition [J]. Journal of Image and Graphics, 2000, **5A**(12): 985-993

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, **5A**(12): 985-993

[4] Sun Jie, Li Feng-ting. Research on 3D object recognition based on affine invariance [D]. Beijing: Tsinghua University, 2008

孙洁,李凤亭. 基于仿射不变性的三维目标识别研究, 北京:清华大学, 2008

[5] Lu Jing, Su Xian-yu ,Wang Hai-xia. Three-dimensional object recognition with rotation-invariant by structure light [J]. Journal of Optoelectronics •Laser, 2004 **15**(12): 1492-1497

吕静, 苏显渝, 王海霞.旋转不变的三维物体识别[J]. 光电子·激光, 2004 **15**(12): 1492-1497

[6] Sun Yi-gang, Yang Li-yong, Sun Cheng-qi. Study on approach of 3D object recognition across multiple views in affine coordinate system [J]. Journal of civil aviation university of China, 2005, **23**(6): 53-56

孙毅刚, 杨立勇, 孙承琦. 仿射坐标系下多视点的三维物体识别方法研究[J]. 中国民航学院学报, 2005, **23**(6): 53-56

[7] Silvio Savarese, Li Fei-fei. 3D generic object categorization, localization and pose estimation [A]. In: International Conference on Computer Vision [C], Rio de Janeiro, Brazil, 2007: 1-8