

摘 要

图像是普遍存在于个人主机和网络上的一种包含丰富信息的载体，因此正确的理解图像内容对于图像数据库的检索、个人主机管理、数字图书馆和媒体制作等领域有重要的作用。图像具有直观、内容丰富等特点，所包含的语义内容非常多。用户对图像的理解通常不仅仅是图像中所包含的各个孤立的物体，而且能根据图像的整体抽象出一些概念，比如情景、事件或是情感等。为了能更好的理解图像的语义内容，减少语义鸿沟的影响，本文在图像内容的语义理解方面进行了研究，主要解决的问题是：特定图像的识别、纹理图像分类，场景图像分类。我们提出的解决这些问题的方法结合了具有不变性的局部图像特征和机器学习模型。

特定图像识别可以应用到很多计算机视觉问题中，比如图像拼接、场景图像匹配、图像监控、图像检索等。提出一种新的方法来识别两幅图像是否包含相同的内容。提出的方法是基于局部不变特征的。首先，通过一个三步过程，来决定图像上局部不变区域的位置、尺度和主要方向。然后利用一种新的描述子来表示这个区域。最后计算两幅图像间描述子的对应关系，根据建立对应关系的描述子的个数来判断两幅图像是否含有同样的内容。提出的方法可以处理具有较大几何变形和成像变化情况下的特定图像识别任务。实验结果证明了我们方法是有效的。

提出了一种新的无监督分类方法来分类具有尺度、仿射等变化的纹理图像集合。提出的方法不需要专家来提供标注过的训练集合。给定纹理图像集合后，从每张纹理图像上提取一组不变描述子。所有纹理图像的描述子经过矢量量化形成视觉词汇表。然后把纹理图像表示成“Histogram of VisualWords”矢量。通过类比于文本分类，利用概率隐含语义索引模型(PLSI)和非负矩阵分解模型(NMF)来进行无监督的纹理分类。提出的方法可以从给定的纹理图像集合上自动的发现纹理类别，而不需要提供任何纹理类别的先验知识。提出的方法在UIUC的纹理数据库上进行了一系列的测试，这个数据库包含了非常大的视点和尺度变化。我们还在Brodatz纹理数据库上进行了无监督纹理分类实验。我们还利用从没有标注的纹理集上学到的纹理模型来分类新的图像。实验结果证明我们的方法对视点变化和尺度变化具有鲁棒性，并且无需标注的训练集合就可以得到很好分类精度。

场景图像分类是计算机视觉领域中的一个基本问题。在本文中提出了一个新的模型来进行场景图像的分类。提出的模型对场景图像内容的相关性进行了建模，并且可以描述图像局部特征之间的空间关系。首先从图像上提取视觉单词，

并把图像表示成视觉单词的词频矢量。然后我们利用产生式模型来学习训练集合中包含的主题，和每一幅图像的主题比例矢量。最后用判定式分类器进行多类学习。提出的方法利用logistic正态分布对主题的相关性进行建模，从而在学习过程中减少不相关主题出现的可能，使得学习得到的类别的主题分布更准确。并且在学习过程中不需要对图像内容进行人工标注，从而节省了很多人力，同时避免了人工标注中的主观性偏差。并且通过构造图像的空间金字塔来提取图像的层次型视觉直方图矢量，从而把特征的空间位置关系结合到我们的模型中。我们提出了一种新的局部区域描述方法，它结合了局部区域的梯度信息和彩色信息。在自然景色图像集合和人造图像集合上成功实验了这种方法，它相对于传统方法取得了更好的结果。