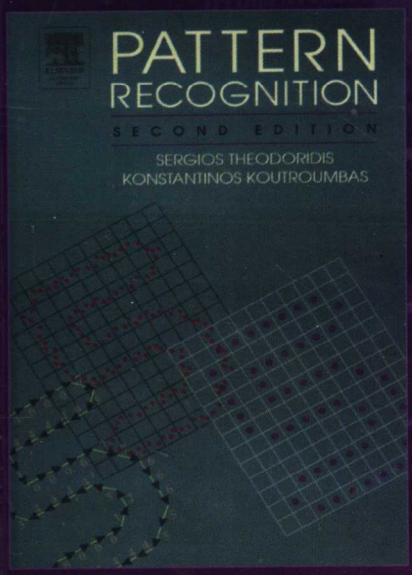
国外计算机科学教材系列

模式识别

（第二版）

Pattern Recognition

Second Edition



[希腊] Sergios Theodoridis
Konstantinos Koutroumbas 著

李晶皎 朱志良 王爱侠 等译

ELSEVIER ACADEMIC PRESS

PHEI

電子工業出版社
Publishing House of Electronics Industry
http://www.phei.com.cn

国外计算机科学教材系列

模式识别

（第二版）

Pattern Recognition

Second Edition

[希腊] Sergios Theodoridis
Konstantinos Koutroumbas 著

李晶皎 朱志良 王爱侠 等译

電子工業出版社
Publishing House of Electronics Industry
北京 · BEIJING

内容简介

本书是在第一版的基础上，由两位有十余年教学经验的资深专家完成的。全书共分16章，主要讲述了特征选择和特征生成，具体有小波、分形和独立成分分析；线性和非线性分类器，具体有贝叶斯分类、多层感知器、决策树和RBF网络；上下文相关分类，具体有动态规划和隐马尔可夫模型技术；新增章节有支持向量机、可变模式匹配和附录的约束最优化等，且包含图像分析、文字识别、医学诊断、语音识别等应用。此外，每章均附有习题。

本书可作为高等院校自动化、计算机、电子和通信等专业研究生和高年级本科生的教材，也可作为计算机信息处理、自动控制等相关领域的工程技术人员的参考用书。

版权贸易合同登记号　图字：01-2003-3951

图书在版编目（CIP）数据

模式识别：第二版/（希）西奥多里蒂斯（Theodoridis, S.）等著；李晶皎等译.
-北京：电子工业出版社，2004.8
（国外计算机科学教材系列）
书名原文：Pattern Recognition, Second Edition
ISBN 7-5053-9924-1

Ⅰ.模...　Ⅱ.①西...　②李...　Ⅲ.模式识别-教材　Ⅳ.TP391.4

中国版本图书馆CIP数据核字（2004）第085025号

责任编辑：谭海平　　特约编辑：李玉龙
印　　刷：北京智力达印刷有限公司
出版发行：电子工业出版社
　　北京市海淀区万寿路173信箱　邮编：100036
经　　销：各地新华书店
开　　本：787 × 1092　1/16　　印张：28.5　　字数：728千字
印　　次：2004年8月第1次印刷
定　　价：45.00元

凡购买电子工业出版社的图书，如有缺损问题，请向购买书店调换；若书店售缺，请与本社发行部联系。联系电话：（010）68279077。质量投诉请发邮件至zlts@phei.com.cn，盗版侵权举报请发邮件至dbqq@phei.com.cn。

出 版 说 明

21世纪初的5至10年是我国国民经济和社会发展的重要时期，也是信息产业快速发展的关键时期。在我国加入WTO后的今天，培养一支适应国际化竞争的一流IT人才队伍是我国高等教育的重要任务之一。信息科学和技术方面人才的优劣与多寡，是我国面对国际竞争时成败的关键因素。

当前，正值我国高等教育特别是信息科学领域的教育调整、变革的重大时期，为使我国教育体制与国际化接轨，有条件的高等院校正在为某些信息学科和技术课程使用国外优秀教材和优秀原版教材，以使我国在计算机教学上尽快赶上国际先进水平。

电子工业出版社秉承多年来引进国外优秀图书的经验，翻译出版了“国外计算机科学教材系列”丛书，这套教材覆盖学科范围广、领域宽、层次多，既有本科专业课程教材，也有研究生课程教材，以适应不同院系、不同专业、不同层次的师生对教材的需求，广大师生可自由选择和自由组合使用。这些教材涉及的学科方向包括网络与通信、操作系统、计算机组织与结构、算法与数据结构、数据库与信息处理、编程语言、图形图像与多媒体、软件工程等。同时，我们也适当引进了一些优秀英文原版教材，本着翻译版本和英文原版并重的原则，对重点图书既提供英文原版又提供相应的翻译版本。

在图书选题上，我们大都选择国外著名出版公司出版的高校教材，如Pearson Education培生教育出版集团、麦格劳－希尔教育出版集团、麻省理工学院出版社、剑桥大学出版社等。撰写教材的许多作者都是蜚声世界的教授、学者，如道格拉斯·科默（Douglas E. Comer）、威廉·斯托林斯（William Stallings）、哈维·戴特尔（Harvey M. Deitel）、尤利斯·布莱克（Uyless Black）等。

为确保教材的选题质量和翻译质量，我们约请了清华大学、北京大学、北京航空航天大学、复旦大学、上海交通大学、南京大学、浙江大学、哈尔滨工业大学、华中科技大学、西安交通大学、国防科学技术大学、解放军理工大学等著名高校的教授和骨干教师参与了本系列教材的选题、翻译和审校工作。他们中既有讲授同类教材的骨干教师、博士，也有积累了几十年教学经验的老教授和博士生导师。

在该系列教材的选题、翻译和编辑加工过程中，为提高教材质量，我们做了大量细致的工作，包括对所选教材进行全面论证；选择编辑时力求达到专业对口；对排版、印制质量进行严格把关。对于英文教材中出现的错误，我们通过与作者联络和网上下载勘误表等方式，逐一进行了修订。

此外，我们还将与国外著名出版公司合作，提供一些教材的教学支持资料，希望能为授课老师提供帮助。今后，我们将继续加强与各高校教师的密切联系，为广大师生引进更多的国外优秀教材和参考书，为我国计算机科学教学体系与国际教学体系的接轨做出努力。

电子工业出版社

教材出版委员会

译 者 序

模式识别诞生于20世纪20年代，随着20世纪40年代计算机的出现，20世纪50年代人工智能的兴起，模式识别在20世纪60年代初迅速发展成一门学科。模式识别研究的理论和方法在很多地方得到了成功的应用，从最初的光学字符识别（OCR），扩展到笔输入计算机、生物身份认证、DNA序列分析、化学气味识别、药物分子识别、图像理解、人脸辨识、表情识别、手势识别、语音识别、说话人识别、信息检索、数据挖掘和信号处理等。

尽管如此，与生物认知系统相比，模式识别系统的识别能力和鲁棒性还远不能让人满意。模式识别还有许多的基础理论和基本方法等待人们解决，新问题也层出不穷。为此，相关人员很需要一本关于这一领域的高水平学术著作，它既有基础知识的介绍，还有本领域研究现状的介绍，以及未来发展的展望等。本书正是这样一本经典著作。

本书是在第一版的基础上，于2003年由具有25年教学经验的资深专家、希腊雅典大学信息与通信系通信和信号处理专业教授Sergios Theodoridis等两人完成的。本书采用统一的方法讲述各种模式识别方法，以及图像分析、语音处理和通信等应用。

为了适用于电子工程学、计算机工程学、计算机科学和信息、自动控制等专业的研究生，以及高年级本科生等各种不同知识背景的学生，本书内容安排既全面，又相对独立。在各个章节中需要的一些数学工具，如概率、统计和约束优化等知识，在本书的4个附录中做了简单的讲解。本书可以面向大学生和研究生，可以作为一学期或两个学期的课程。本书也可以作为自学教材，或供研究人员和工程技术人员参考。

东北大学李晶皎教授负责译校本书的第1章至第7章，朱志良教授负责译校本书的第11章至第15章，王爱侠负责译校本书的第8章至第10章、第16章以及附录A至附录D。参加初译的有东北大学信息学院计算机专业的硕士研究生甄广启、赵骥、王蓓蕾、唐春强、冯云、何敬禹、左刚、谭光力、赵旸、刘巍等。

在翻译过程中，我们力求忠实、准确地把握原著，同时保留原著的风格。但由于译者水平有限，书中难免有错误和不准确之处，恳请广大读者批评指正。

前　言

本书是作者在20年来给研究生和本科生教学的基础上编写的，该课程面向很多专业的学生，例如电子工程学、计算机工程学、计算机科学和信息以及自动控制等专业的研究生。这些经验使我们得以把本书内容编写得既全面又相对独立，并且适用于各种不同知识背景的学生。读者需要具备的知识包括基础的微积分学、初等线性代数和概率论基础。在各个章节中需要的一些数学工具，如概率、统计和约束优化等知识，在本书的4个附录中做了简单的讲解。本书可以面向大学生和研究生，可以作为一学期或两个学期的课程。本书也可以作为自学教材，或供研究人员和工程技术人员参考。我们编写本书的动力之一是使这本书适合于所有从事模式识别相关研究的人员。

本书采用统一的方法讲述各种模式识别，包括图像分析、语音处理和通信应用。尽管这些领域有很多不同点，但也有共同之处，对它们的研究也有统一的方法。本书的每一章都采用循序渐进的讲解方式，即从基础开始过渡到比较高深的课题，最后对最新技术发表评论。在每章的末尾都提供一些习题和上机练习，读者可以从出版商处购买解答指南。此外，读者可以从本书的网站（www.di.uoa.gr/~stpatrec）得到一些MATLAB示例。

我们定期在网站上增加和更新示例，欢迎读者多提建议。尽管网站上的内容经过多次仔细检查，但有些地方还是不可避免地存在错误，欢迎读者批评指正。

本书的出版离不开广大师生多年来的支持和帮助。特别感谢K. Berberidis教授、E. Kofidis博士、A. Liavas教授、A. Rontogiannis博士、A. Pikrakis博士、Gezerlis博士和K. Georgoulakis博士，I. Kopsinis博士自始至终都给予了莫大的支持和帮助；另外，G. Moustakides教授、V. Digalakis教授、T. Adali教授、M. Zervakis教授、D. Cavouras教授、A. Böhm教授、G. Glentis教授、E. Koutsoupias教授、V. Zissimopoulos教授、A. Likas教授、A. Vassiliou博士、N. Vassilas博士、V. Drakopoulos博士和S. Hatzispyros博士在仔细阅读书稿后提出了许多宝贵建议，大大提高了本书的可读性。同时，也非常感谢广大学生的付出和给予的建设性意见。非常感谢N. Kalouptsidis教授，长期以来我们的合作和友谊是本书灵感的来源。

最后，K. Koutroumbas感谢Sophia的耐心和支持，S. Theodoridis感谢Despina、Eva和Eleni，她们是快乐和动力的源泉。

目　录

第1章 导 论

1.1 模式识别的重要性

模式识别是一门以应用为基础的学科，目的是将对象进行分类。这些对象与应用领域有关，它们可以是图像、信号波形或者任何可测量且需要分类的对象。可以用专用术语“模式”来称呼这些对象。模式识别具有悠久的历史，但在20世纪60年代以前，模式识别主要是统计学领域中的理论研究。同其他事物一样，计算机的出现提高了对模式识别实际应用的需求，而这反过来又对理论发展提出了更高的要求。就像我们的社会从工业化到后工业化阶段一样，工业生产中的自动化以及信息处理和检索的需求变得日益重要，这种趋势把模式识别推向今天的工程应用和研究的高级阶段。在大多数机器智能系统中，模式识别是用于决策的主要部分。

在机器视觉中，模式识别是非常重要的。机器视觉系统通过照相机捕捉图像，然后通过分析生成图像的描述信息。典型的机器视觉系统主要应用在制造业中，作为自动化视觉检验或装配线的自动化。例如，在自动化视觉检验应用中，生产的产品通过传送带移动到检验站，检验站的照相机确定产品是否合格。因此，必须在线分析图像，模式识别系统必须将产品分为“合格”和“不合格”。然后，根据分类结果采取相应的行动，比如丢弃不合格的产品。在装配线上，必须对不同的对象进行定位和识别，也就是说，将对象分类到已知类别的某一类中，如镙丝刀类、德国钥匙类以及任何工具制造单元，然后机器手把这些对象放置在正确的位置。

字符（字母或数字）识别是模式识别应用的另一个重要领域，主要用于自动化和信息处理。光学字符识别（Optical Character Recognition, OCR）系统已经开始在商业中应用，而且我们或多或少都对其有所了解。OCR系统有一个前端设备，由光源、扫描镜头、文档传送机和检测器组成。在光敏检测器的输出端，光的强度变化转换成数字，并形成图像阵列。然后，一系列的图像处理技术的应用完成了线和字符的分段，模式识别软件完成字符识别的任务，也就是将每一个符号分到相应的“字符、数字、标点符号”类中。存储识别出的文档比存储扫描的图像有两个好处。首先，如果需要，用字处理器使进一步的电子处理更容易；其次，存储ASCII字符比存储文档的图像效率更高。除了印刷体字符识别系统外，现在更多的研究集中于手写体识别。这种系统的典型商业应用是银行支票的机器识别，机器必须能够识别数字的数量和阿拉伯数字，并进行匹配，而且能够检查收款人相应的支出信用是否相符。哪怕只有一半的支票识别正确，这样的机器也可以从枯燥的工作中节省人力。另一个应用是在邮局进行邮政编码识别的自动邮政系统。在线手写体识别系统是具有巨大商业利益的另一应用领域，此系统将用于笔输入计算机。在这种计算机中，数据的输入不是通过键盘而是通过手写，这顺应了开发具有人类技能接口的机器这一发展趋势。

计算机辅助诊断是模式识别的另一个重要的应用，目的是帮助医生做诊断决定，当然最终的诊断由医生来完成。计算机辅助诊断已经应用于实际，主要研究各种医疗数据，如X射线、计算机断层图、超声波图、心电图和脑电图。计算机辅助诊断的需求源于医疗数据较难解释，并且解释结果多依赖于医生的经验这一事实。我们以检查乳腺癌的乳腺X射线照相术为例，尽管乳腺X射线照相术是检测乳腺癌的最好方法，但是10%~30%的患病妇女在乳腺X射线照相术中可能得到相反的

乳腺X射线照片。在这种情况下，大约2/3的放射线医师不能检测出癌变，很明显这是错误的。这可能是因为图像质量不好、放射线医师眼睛疲劳或有疾病等原因造成的。通过另一个放射线医师再次查看照片可以提高正确分类的百分比。因此，可以发明一种模式识别系统通过提出第二种观点来帮助放射线医师进行诊断。基于乳腺X射线照片日益准确的诊断反过来会减少乳腺癌疑似病例的数量，使这些人免于承受外科胸部活组织检查的痛苦。

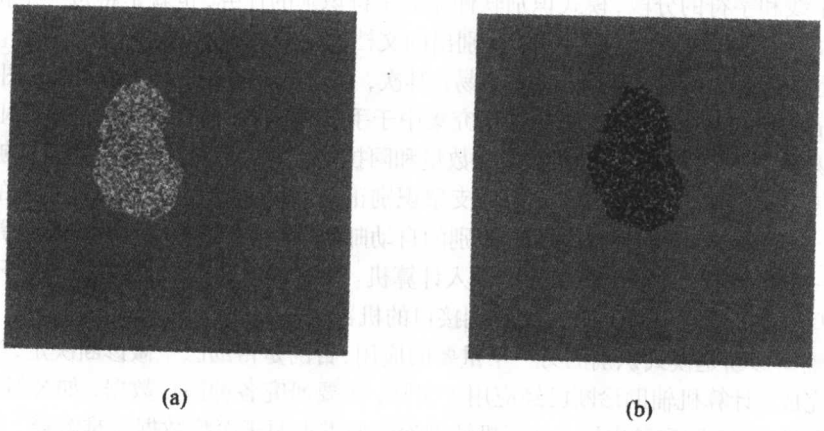
语音识别是模式识别的另一个研究领域，在这个领域中已经进行了大量的研究。语音是人类最自然的沟通和交换信息的方式。因此，长期以来，建立能够识别语音信息的智能机器成为科学家和科幻小说家的目标。这种机器的潜在应用是广泛的。例如，可以用来有效改善制造业的环境，可以远程控制危险环境中的机器，以及通过对话控制机器来帮助残疾人。经过努力，另外一个已经取得一定成功的重要应用是用麦克风向计算机进行语音输入，语音识别系统的软件能够识别语音文本，翻译成ASCII码，并可以显示在显示器上和存储在存储器中。计算机语音输入的速度是熟练打字员输入的两倍，而且有助于增强我们和聋哑人的交流能力。

前面提到的只是众多可能应用中的四个例子。典型的应用包括指纹识别、签名认证、文本检索、表情和手势识别。目前的应用研究吸引了很多研究者，其目的是使人机互动更简单，并增强计算机在办公自动化等环境中的作用。为了激发想像力，值得一提的是MPEG-7标准，它包含对数字图书馆中录像带的视频信息检索：在数字图书馆中查找所有显示某人“X”微笑的视频场景。当然，要在所有这些应用中达到最终目标，模式识别还依赖于其他一些学科的发展，如语言学、计算机图形学和计算机视觉等。

为了唤起读者对模式识别的好奇心，下面简要介绍基本结构和方法，其中包括已经研究出来的各种各样的模式识别方法。

1.2 特征、特征向量和分类器

首先模拟一个简化的例子“mimicking”，这是一个医疗图像分类任务。图1.1给出了两个图像，每个图像中有一块突出区域，两个区域彼此有明显的不同之处。我们可以认为图1.1(a)所示的图像是良性的，属于A类；图1.1(b)所示的图像是恶性的（癌），属于B类。进一步假设有效样本（图像）不止这些，我们可以访问图像数据库，那里有一系列样本，其中一些样本是A类，一些是B类。



(a) (b)

图1.1 图像兴趣区举例：(a)A类；(b) B类

第一步是确定可测量值来区别两个图像区域。图1.2显示了每一个区域中的强度均值和其标准偏差的关系图。每一点代表着已知数据库中一个不同的图像，这表明A类样本和B类样本分布于不

同的区域，中间的直线正好将这两类分开。假设现在有一幅新的图像，不知道它属于哪一类，我们计算兴趣区的均值强度和标准偏差，并画出相应的点，在图1.2中用*号表示，从而可以判定未知类型的样本更接近于A类。

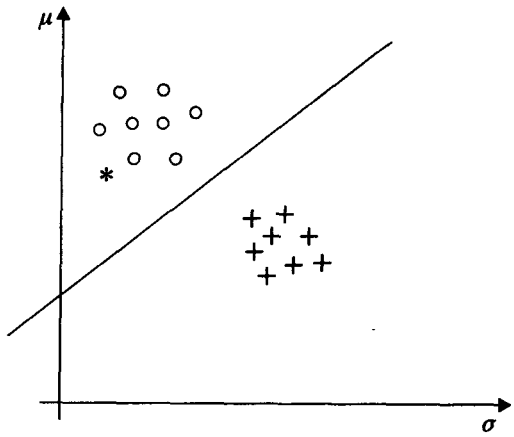


图1.2 来自于A类（o）和B类(+)的一系列图像的相对于标准偏差的均值，这种情况下用一条直线将两类分开

前面所描述的人工分类的任务过程概述了大部分模式识别问题的基本原理。在这个例子中，用来分类的测量方法——均值和标准偏差称为特征值。在一般的情况下使用l个特征$x_i, i = 1, 2, \cdots, l$组成特征向量$\boldsymbol{x} = [x_1, x_2, \cdots, x_l]^T$，其中$T$表示转置。每一个特征向量表示一个样本（对象）。在本书中，特征和特征向量分别看做是随机变量和向量，因为来源于不同样本的测量值是随机的数据。一方面是因为测量仪器的测量噪声，另一方面是因为每一种模式有截然不同的特点。例如，由于不同个体之间生理的不同，所以X射线成像有很大的不同，这也是图1.1中每一类的特征点发散的原因。

图1.2中的直线称为决策线，由它决定的分类器将特征空间划分为不同的类空间。如果对应于一个未知类别的样本特征向量$\boldsymbol{x}$落在A类区域，则划为A类，否则划为B类。但这并不意味着决策是正确的。如果不正确，则出现了一个错误的分类。为了在图1.2中画出这条直线，我们需要知道图中每一个点的类别标签（A类或B类），即用来设计分类器的样本（特征向量）的所属类是已知的，这些样本称为训练样本（训练特征向量）。

上面给出了定义和基本原理，下面给出分类任务中的基本问题。

- 特征提取：在前面的例子中，用均值和标准偏差作为特征，因为我们知道应该从图像中提取这些特征。但在实际问题中，特征不是显而易见的。这是分类系统设计的特征提取阶段的任务，它完成已知样本的识别。
- 特征数l为多少最好？这也是一个很重要的问题，它在分类系统设计的特征选择阶段完成。在实际问题中，总是产生大量的特征供选择，在其中选择最好的使用。
- 对指定的任务选择了合适的特征后，怎样设计分类器？在前面的例子中，只是为了观察方便，根据经验画了一条直线。在实际问题中不可能这样，必须按照最优准则将线画在最优的位置。哪一个线性分类器（直线或l维空间的超平面）具有可接受性能没有固定的判定规则。一般来说，不同类别的区域划分是非线性的。在l维特征空间中，采用什么样的非线性分类器以及采用什么样的优化准则？这些问题在分类器设计阶段解决。

- 当分类器设计完毕后，如何评估分类器的性能？也就是说，分类误差率是多少？这是系统评估阶段的任务。

图1.3给出了分类系统设计的各个阶段，从这些反馈箭头可以看出，每一步都不是独立的。相反，它们相互关联，相互依赖。为了提高整体性能，每一阶段都有可能返回到前一阶段重新设计。而且有一些阶段可以合并，例如，特征选择和分类器设计阶段处于同一优化任务中。

虽然已经向读者描述了分类系统设计核心的一些基本问题，但是还有一些问题必须提到。

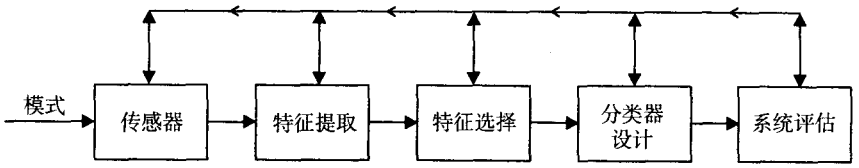


图1.3 分类系统设计的基本步骤

1.3 有监督和无监督模式识别

在图1.1的例子中，假设有一个可用的训练数据集，并通过挖掘先验已知信息来设计分类器，这称为有监督模式识别（Supervised Pattern Recognition）。但是，并不总是出现这种情况，另外一种模式识别没有已知类别标签的训练数据可用。在这种情况下，给定一组特征向量$\boldsymbol{x}$来揭示潜在的相似性，并且将相似的特征向量分为一组，这就是无监督模式识别（Unsupervised Pattern Recognition）或聚类（Clustering）。在社会科学和工程中会出现这种情况，例如遥感、图像分段、图像和语音编码。下面来看两个这样的问题。

在多光谱遥感中，放在人造卫星、航天飞机或太空工作站的灵敏扫描器测得从地球表面发射的电磁能量。这个能量可能是反射的太阳能（被动），或者反射从媒介发射的部分能量，目的是为了探测地球表面的情况。扫描器对电磁辐射的部分波段敏感，地球表面情况的特征不同，对波段反射的能量就不同。例如，在可见红外波段内，矿物质、潮湿的土壤、水域和潮湿的植被都是反射能量的主要贡献者。在热红外区，主要反映地球表面和地表下的热容量和热特性。因此，每一个波段测量地球表面同一块地方不同的特性，用这种方法可以根据不同波段的反射能量分布来生成地球表面的图像。研究这些信息的目的是识别各种地面类型，如公路、农田、森林、火烧地面、水和患病的农作物等。最后，生成了地球表面每一个单元的特征向量$\boldsymbol{x}$，向量中的元素$x_i, i = 1, 2, \cdots, l$对应于各种光谱波段中像素的强度。实际上，光谱波段的数量是变化的。

聚类算法可以用来完成对l维特征向量的分组。对应于相同地面类型的点，如水，将其聚类在一起形成一组。一旦这样分组以后，分析人员就可以通过每一组中的样本点和地面数据的参考信息（地图或观察结果）相联系来识别地面类型，图1.4说明了这个过程。

聚类也广泛地应用于社会科学，进行研究、调查、统计数据以及得到一些有用的结论来引导正确的行为。再来看一个简单的例子，假定我们既要研究一个国家的国民生产总值和人的文盲水平是否有关，又要研究国民生产总值和儿童的死亡率是否有关。在这个例子中，每个国家都用一个三维特征向量来表示，且特征向量的每一项与之对应。聚类算法将揭示低国民生产总值、高文盲和高儿童死亡率（以人口百分比表示）的这些国家的聚类相似性。

无监督模式识别主要用于确定两个特征向量之间的“相似度”以及合适的测度，并选择一个算法方案，基于选定的相似性测度对向量进行聚类（分组）。通常，不同的算法方案可能导致不同的结果，这一点必须由专家进行解释。