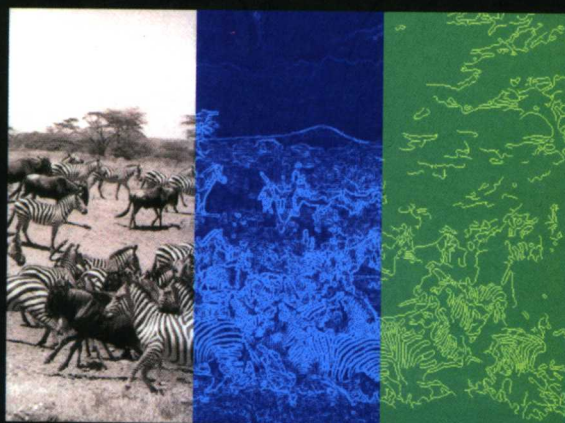


计算机视觉

(美) Linda G. Shapiro George C. Stockman 著 赵清杰 钱芳 蔡利栋 译

COMPUTER VISION



Linda G. Shapiro ■ George C. Stockman

Computer Vision

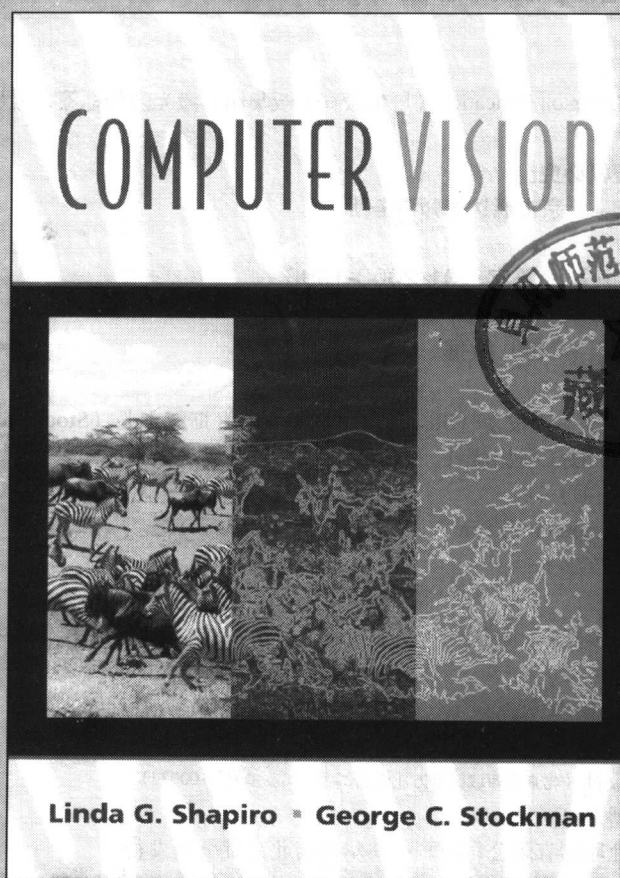


机械工业出版社
China Machine Press

计 算 机 科 学 丛 书

计算机视觉

(美) Linda G. Shapiro George C. Stockman 著 赵清杰 钱芳 蔡利栋 译



Computer Vision



机械工业出版社
China Machine Press

本书系统地介绍了计算机视觉方面的基础知识,详细讨论了从图像自动抽取重要信息的相关理论,内容包括最近出现的研究成果。本书取材新颖精练,重点突出,以解决实际问题为目的。前11章讨论的是2D情况;第12章到第15章从2D情况扩展到3D情况;第16章介绍了利用计算机视觉技术的实际应用系统。书中的大量实例及习题,贴近生活,面向应用,富有情趣。

本书适合作为高等院校计算机及相关专业的高年级本科生和研究生的教材,也可供相关技术人员参考。

Simplified Chinese edition copyright © 2005 by Pearson Education Asia Limited and China Machine Press.

Original English language title: *Computer Vision* by Linda G. Shapiro and George C. Stockman, Copyright 2001.

All rights reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Prentice-Hall.

本书封面贴有Pearson Education(培生教育出版集团)激光防伪标签,无标签者不得销售。

版权所有,侵权必究。

本书法律顾问 北京市展达律师事务所

本书版权登记号:图字:01-2003-1996

图书在版编目(CIP)数据

计算机视觉/(美)夏皮罗(Shapiro, L. G.), (美)斯托克曼(Stockman, G. C.)著;赵清杰等译.-北京:机械工业出版社,2005.3

(计算机科学丛书)

书名原文:Computer Vision

ISBN 7-111-15972-1

I. 计… II. ①夏… ②斯… ③赵… III. 计算机视觉-高等学校-教材 IV. TP302.7

中国版本图书馆CIP数据核字(2005)第000945号

机械工业出版社(北京市西城区百万庄大街22号 邮政编码 100037)

责任编辑:傅志红

北京诚信伟业印刷有限公司印刷·新华书店北京发行所发行

2005年3月第1版第1次印刷

787mm×1092mm 1/16·28.5(彩插0.75印张)印张

印数:0 001-4000册

定价:55.00元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换
本社购书热线:(010) 68326294

出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅肇划了研究的范畴，还揭橥了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短、从业人员较少的现状下，美国等发达国家在其计算机科学发展的几十年间积淀的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章图文信息有限公司较早意识到“出版要为教育服务”。自1998年开始，华章公司就将工作重点放在了遴选、移译国外优秀教材上。经过几年的不懈努力，我们与Prentice Hall, Addison-Wesley, McGraw-Hill, Morgan Kaufmann等世界著名出版公司建立了良好的合作关系，从它们现有的数百种教材中甄选出Tanenbaum, Stroustrup, Kernighan, Jim Gray等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及度藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专诚为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍，为进一步推广与发展打下了坚实的基础。

随着学科建设的初步完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都步入一个新的阶段。为此，华章公司将加大引进教材的力度，在“华章教育”的总规划之下出版三个系列的计算机教材：除“计算机科学丛书”之外，对影印版的教材，则单独开辟出“经典原版书库”；同时，引进全美通行的教学辅导书“Schaum's Outlines”系列组成“全美经典学习指导系列”。为了保证这三套丛书的权威性，同时也为了更好地为学校和老师服务，华章公司聘请了中国科学院、北京大学、清华大学、国防科技大学、复旦大学、上海交通大学、南京大学、浙江大学、中国科技大学、哈尔滨工业大学、西安交通大学、中国人民大学、北京航空航天大学、北京邮电大学、中山大学、解放军理工大学、郑州大学、湖北工学院、中国国家信息安全测评认证中心等国内重点大学和科研机构在计算机的各个领域的著名学者组成“专家指导委员会”，为我们提供选题意见和出版监督。

这三套丛书是响应教育部提出的使用外版教材的号召，为国内高校的计算机及相关专业

的教学度身订造的。其中许多教材均已为M. I. T., Stanford, U.C. Berkeley, C. M. U. 等世界名牌大学所采用。不仅涵盖了程序设计、数据结构、操作系统、计算机体系结构、数据库、编译原理、软件工程、图形学、通信与网络、离散数学等国内大学计算机专业普遍开设的核心课程,而且各具特色——有的出自语言设计者之手、有的历经三十年而不衰、有的已被全世界的几百所高校采用。在这些圆熟通博的名师大作的指引之下,读者必将在计算机科学的宫殿中由登堂而入室。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑,这些因素使我们的图书有了质量的保证,但我们的目标是尽善尽美,而反馈的意见正是我们达到这一终极目标的重要帮助。教材的出版只是我们的后续服务的起点。华章公司欢迎老师和读者对我们的工作提出建议或给予指正,我们的联系方法如下:

电子邮件: hzedu@hzbook.com

联系电话: (010) 68995264

联系地址: 北京市西城区百万庄南街1号

邮政编码: 100037

专家指导委员会

(按姓氏笔画顺序)

尤晋元	王 珊	冯博琴	史忠植	史美林
石教英	吕 建	孙玉芳	吴世忠	吴时霖
张立昂	李伟琴	李师贤	李建中	杨冬青
邵维忠	陆丽娜	陆鑫达	陈向群	周伯生
周立柱	周克定	周傲英	孟小峰	岳丽华
范 明	郑国梁	施伯乐	钟玉琢	唐世渭
袁崇义	高传善	梅 宏	程 旭	程时端
谢希仁	裘宗燕	戴 葵		

秘 书 组

武卫东

温莉芳

刘 江

杨海玲

译者序

本书系统阐述了计算机视觉的相关理论和应用技术基础，内容广泛，深入浅出，列举了大量习题和应用实例，不仅适合作为高年级本科生和研究生的教材，也适合作为相关领域研究人员和工程技术人员的参考资料。

本书内容涉及计算机视觉的各个方面，很多内容参考了近期的研究成果，取材新颖精练，重点突出，并以解决实际问题为目的。书中列出了很多算法，都以函数或者过程的形式给出，读者只需用自己熟悉的编程语言稍加修改，即可实现这些算法。书中的大量实例及习题贴近生活，面向应用，富有情趣。另外，在每章后面列出了大量参考文献，这些文献不仅能够帮助读者巩固所学的内容，而且方便读者在自己感兴趣的方向上进行更深入的研究。

本书由赵清杰、钱芳、蔡利栋共同翻译。其中蔡利栋教授翻译了第1、2章的内容，钱芳博士负责第4、5、7、8、10章的翻译工作，赵清杰博士翻译了其余部分并负责全书的统稿工作。参与本书翻译的还有宋霏、王宗远等，在此对他们的工作表示感谢。

由于译者水平有限，加上时间仓促，译稿中难免有错误和遗漏，谨向读者和原作者表示歉意，并欢迎批评指正：zhaoqingjie@tsinghua.org.cn。

赵清杰

2004年9月于北京

前 言

本书系统地介绍了计算机视觉方面的基础知识,内容适合于从事视觉领域研究的广大读者。书中详细讨论了从图像自动抽取重要信息的理论知识,并列举了很多应用实例,为从事这方面学习和研究的学生及科研工作者提供帮助。该书不仅是专业技术人员的一本实用参考资料,更适合作为高年级本科生和研究生的教材。本书主要介绍基本概念与算法,对当前迅速发展的视觉应用领域也进行了论述。本书的独特之处在于,第8章的图像数据库以及第15章的虚拟现实和增强现实,这两部分是迅速发展的最新应用领域。第16章简单介绍了利用计算机视觉技术的实际应用系统。

随着计算机技术的最新发展,计算机图像已经成为一种经济灵活的技术手段,并渗透到各行各业。图像计算不再只属于科学研究领域,也属于艺术领域、社会科学领域,甚至成为人们的业余爱好。这本书适合有专业背景和正在进行专业学习的相关人员,包括对多媒体、艺术设计、地理信息系统和图像数据库感兴趣的读者,以及传统的自动化、图像科学、医学成像、远程感知和计算机绘图等领域的读者。

要使书的内容面面俱到是不可能的。微积分、物理学和常规计算等方面的内容,已有专门的相关教材。我们希望本书不仅可以作为教材,同时又能对一般读者有所帮助。本书所选内容新颖有趣,相信大多数读者都能够看懂。作为研究生或高年级本科生计算机视觉课程的教材使用时,应把参考文献作为课程的补充材料。每章后面都列出了适当数量的参考文献,但并没有包括全部文献。

前面各章首先介绍底层基本知识,并逐步过渡到数学模型部分。目的是为了在涉及图像特征之前,让大家先有一个直观性的了解。标注“*”的部分需要更多的数学知识或者难度更深,在专业性不强的课程中可以不讲这些内容。为了加强直观性理解,在前面的11章里,我们一直在讨论二维(2D)图像,到了后面几章才开始讨论三维(3D)计算机视觉。有经验的教师可以针对不同课程和教学风格,重新安排各章的讲解顺序。2D图像处理有很多用处,许多概念和算法在2D情况下讲解起来更容易理解。第4章介绍模式识别方面的基本知识,使学生在全面掌握图像特征和匹配之前,对完整的识别系统有所了解。学完第4章之后,读者会对2D图像处理应用有更深入的理解。第5、6和7章是有关灰度、颜色和纹理特征的内容。第8章介绍图像数据库方面的知识,这是一个较新的研究内容。一些同仁建议把这部分内容放在书的末尾,我们把它安排得稍微靠前,目的是为了强化前面几章中的有关概念,以及为学期中间的课程作业提供素材。第10和11章讲的是图像分割与匹配,主要针对的是2D情况,不涉及复杂的3D变换,这样可使基本概念描述起来更加简单。

关于3D特征,在第2章做了介绍,在第12章进行了详细讨论。第12章综述了从2D图像恢复3D世界的多方面内容,包括立体视觉的量化模型,由焦距变化恢复深度的薄透镜模型,以及分辨力的概念。第13章介绍3D计算机视觉变换,教学过程中发现这个问题对学生来说难度较大。关于齐次变换的内容安排在这一章内,而没有放到附录中。3D变换是对第11章中2D简单情况的推广。最小二乘拟合也从第11章的2D简单情况推广到第13章的3D情况。本章介绍了

P3P非线性优化方法,并用于进行摄像机标定,包括建立镜头的径向畸变模型。第14章讨论的是3D模型以及模型与3D数据的匹配,这部分难度更大。第15章讨论虚拟现实和增强现实技术,以及计算机视觉在其中扮演的重要角色。

编程语言问题

本书不依赖任何编程语言,而是使用了通用算法符号。用特定语言编写不仅没有必要,而且对许多读者来说也不合适。对会编程的学生来说实现这些算法并不困难,这一点在我们的学生身上已得到证明。在适当和可能的时候,相关例子会公布在WWW上,一方面是为了让学生能够快点儿进行实验,另一方面也使他们能够学习编写代码。

教师和学生可以利用软件工具和程序库,例如Khoros、NIH-Image、XView、gimp和MATLAB等软件工具,也可以从生产视觉硬件设备的公司购买现成的程序包。作者在书中没有用专用软件,因为多数读者使用不同的软件工具,另外工业专用软件具有复杂的数据结构和算法,用这种软件工具进行图像运算达不到预期的学习目的。在简单环境下掌握了算法的基本原理之后,读者在选用专用软件工具时就会得心应手。

如何使用本书

教师和学生可以根据课程的目的和兴趣有选择地学习书中的内容,也可以打乱书中的章节顺序。以下内容仅供参考:

- 第2章简介,第3章作为重点

在数据结构和算法课程中至少需要1~3讲。在第2章的背景知识基础上,第3章内容包括对2D图像阵列、深度优先搜索以及并查数据结构等的应用和编程练习。

- 第1、2、3章和第4、5、6章中的部分内容

大学生做课程设计时,这部分内容可选讲1~3周。要求他们写出简单的学期报告或者设计一个小项目,项目可以是建立2D零件识别系统,利用连通成分和特征向量原型匹配方面的知识。

- 第1~11章的大部分内容

作为地理学、自然资源或微生物学学生的选修课,可以不讲其中带“*”的选学部分。如果作为本科生的图像处理与分析课程,第1~11章的大部分内容都应该讲到。

- 书中的大部分内容

对于相关专业的高年级本科生或者研究生,要用一学期的时间学习计算机视觉这门课。而本书内容用一个学期是讲不完的,部分内容可以不讲或只做简单介绍,后面的习题也可以只做一部分。如果是半个学期的本科生课程,第1~4、6~12和14章内容应该介绍,这些内容可使学生对计算机视觉有比较深的了解。如果是半个学期的研究生课程,第1~4章内容可做简单介绍,第6~14章应该重点讲解,第15章选讲主要内容。如果是研究生课程,应该增加参考文献中的有关内容。

感谢与我们有同样兴趣的同事、教师和学生,他们为本书做了大量的贡献,并把自己的研究成果拿出来与大家分享。许多人为这本书的出版提供了无私的帮助,他们不断鼓励我们,贡献自己的想法、图表和算法等,书中都做了引用说明。几位审稿人和同事提供的宝贵意见对本书的改进帮助很大。特别感谢Mohammad Ghavamzadeh、Nick Dutta、Kevin Bowyer、Adam Clark、Yu-Yu Chou、Habib Abi-Rached和Valentin Razmov,他们对书中的文字做了认真修改。我们对书中的任何错误负责,在以后的版本中将进行改正。

这本书撰写了四年。感谢Addison Wesley-Longman 的Paul Becker, 他在这个过程中做了许多指导性工作。感谢Prentice Hall的Tom Robbins, 因为他使本书得以顺利出版。感谢Cathy Davison 和Lorraine Evans, 他们一直在对很多案例进行跟踪研究。感谢ICC的Rose Rummel-Eury 和Chanda Wakefield, 他们认真修改了书中的符号和语言, 并推进计划的顺利进行。编写这本书任务繁重, 好在我们有一个训练有素和幽默的团队。

Linda Shapiro

shapiro@cs.washington.edu

George Stockman

stockman@cse.msu.edu



图 6-1

(左图) 老虎在草地上的自然色图像

(右图) 由于颜色的改变，对老虎的识别变得不太可靠，也许是只站在地毯上的家猫？（原图经Corel Stock Photos许可）

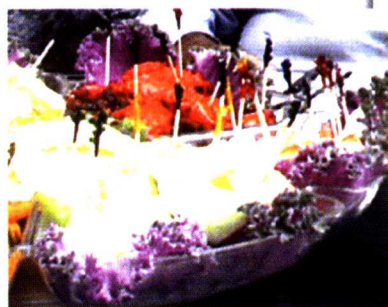
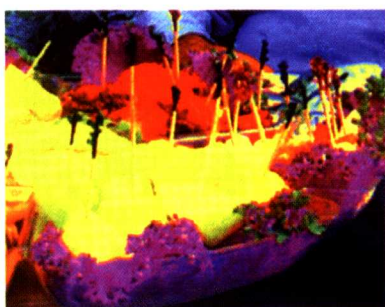


图 6-9

(左图) 输入的RGB图像

(中图) 饱和度S增加40%

(右图) 饱和度S降低20% (Frank Biocca 提供)



图6-10 从左边的彩色图像中分割出白色像素。白色像素的单个连通成分用第3章的颜色算法任意标记 (David Moore提供分析)

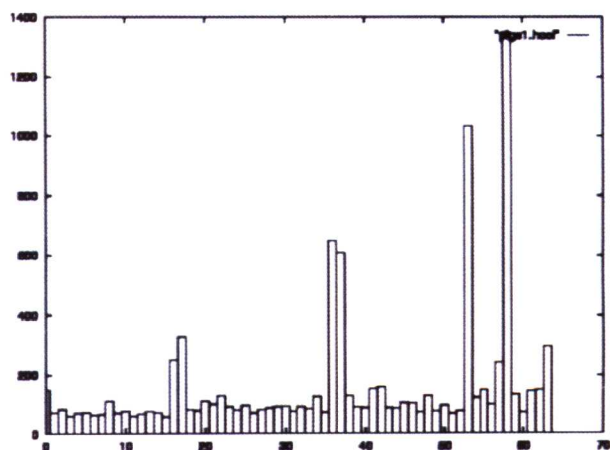
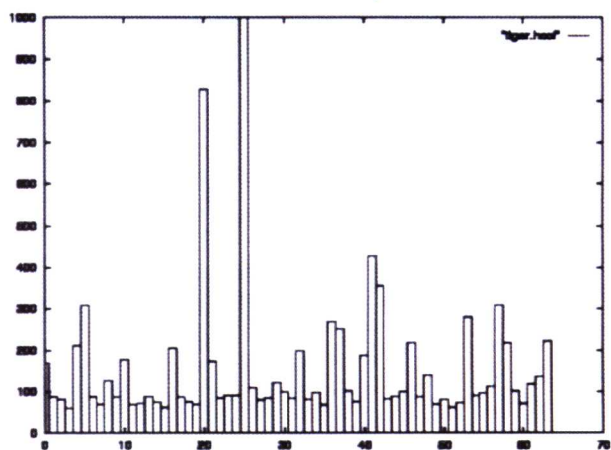


图6-11 彩色图像及其64箱格的直方图（直方图由A. Vailaya提供，图片经Corel Stock Photos许可）

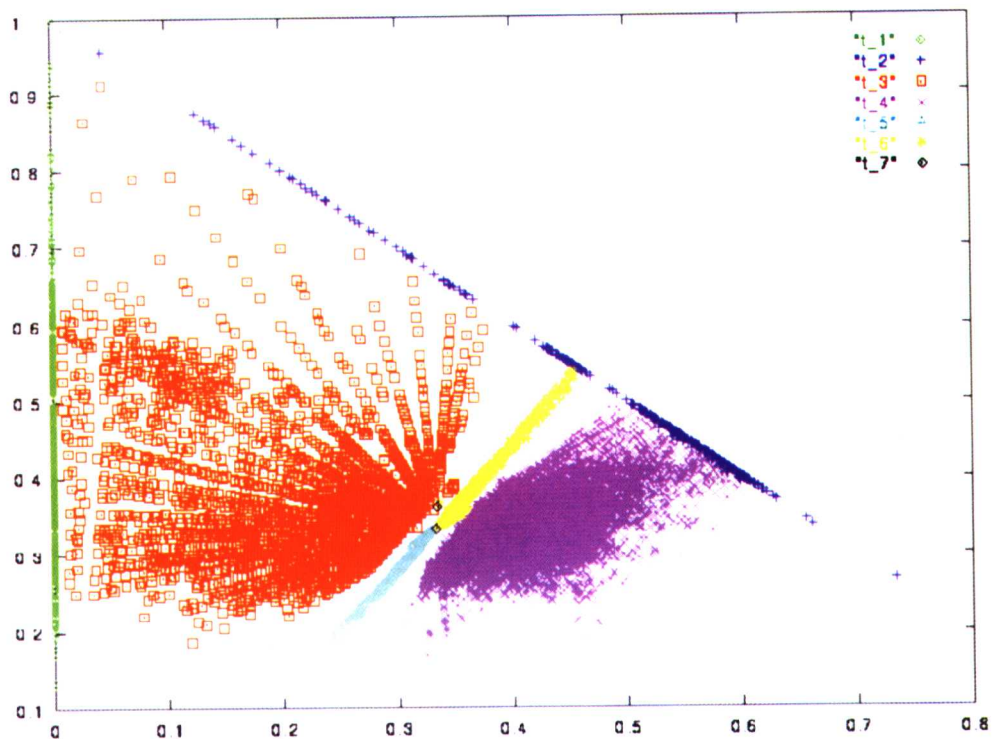
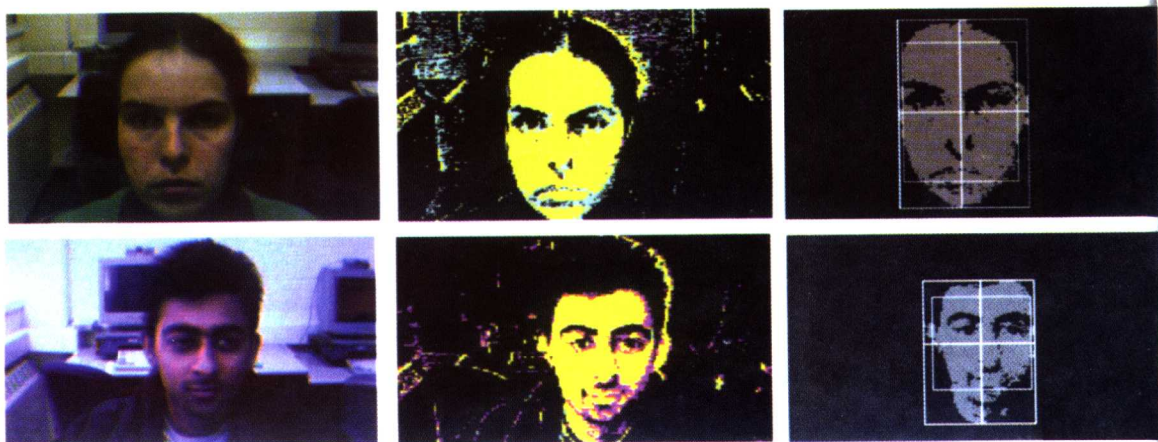


图6-12 通过训练得到的皮肤颜色类别。水平轴是 R_{norm} ，垂直轴是 G_{norm} 。t_4类是主要的人脸颜色，t_5和t_6是次要的人脸类，它们与人脸上的阴影和胡须区域有关（V. Bakic提供）

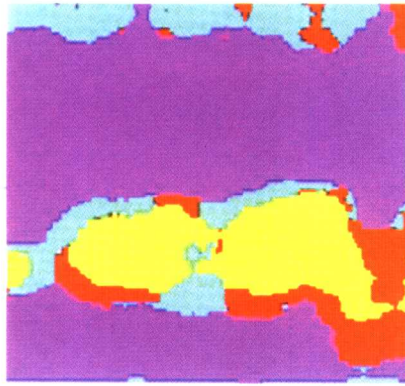
图6-13 人脸抽取实例（图像由V. Bakic提供）

（左图）输入图像 （中图）标记图像 （右图）抽取的人脸区域的边界

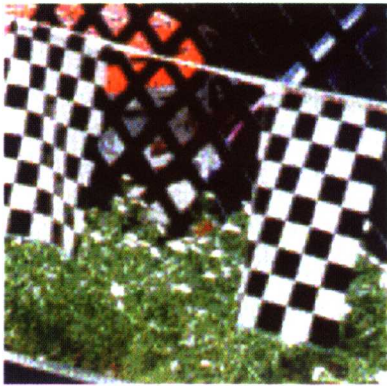




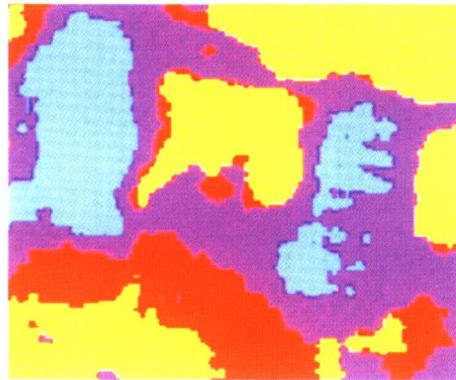
a) 原图



b) 分割成4个类别



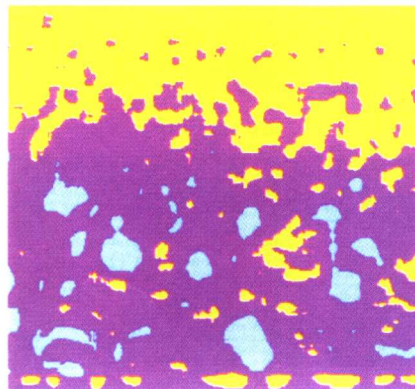
c) 原图



d) 分割成4个类别



e) 原图



f) 分割成3个类别

图7-8 利用Laws纹理能量测度分割图像（原图来自Corel Stock Photos和MIT媒体实验室VisTex数据库）



a) 雷诺阿的绘画



b) 紫水晶图像

图8-1 数字图像示例（皮埃尔·奥古斯特·雷诺阿的绘画，Beaulieu的风景，1893，经旧金山精品艺术博物馆许可，Mildred Anna Williams 收藏，1944.9。紫水晶图像经Smithsonian学院许可，1992）

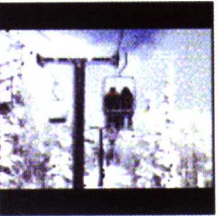
Images 1-8 out of 41			
			
view full size	view full size	view full size	view full size
			
view full size	view full size	view full size	view full size
Columns: Rows:			

图8-3 基于颜色分布相似性的QBIC检索结果。查询图像是位于左上角的图像（Egames提供）

图8-4 基于颜色百分比的QBIC检索结果。查询定义为40%的红色、30%的黄色和10%的黑色（Egames提供）

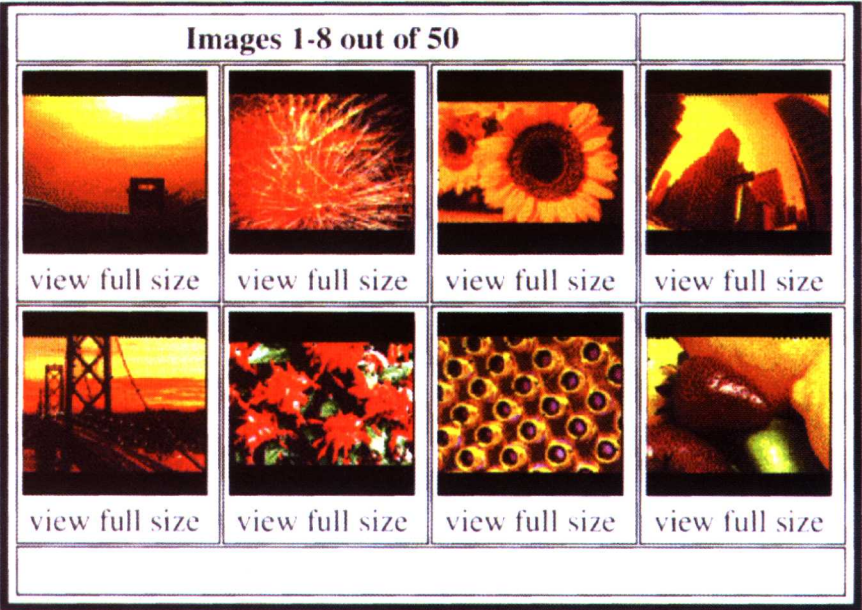


图8-5 图像库检索结果，其中查询图像是涂色的栅格（图像来自MIT媒体实验室的VisTex数据库:<http://vismod.www.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>）

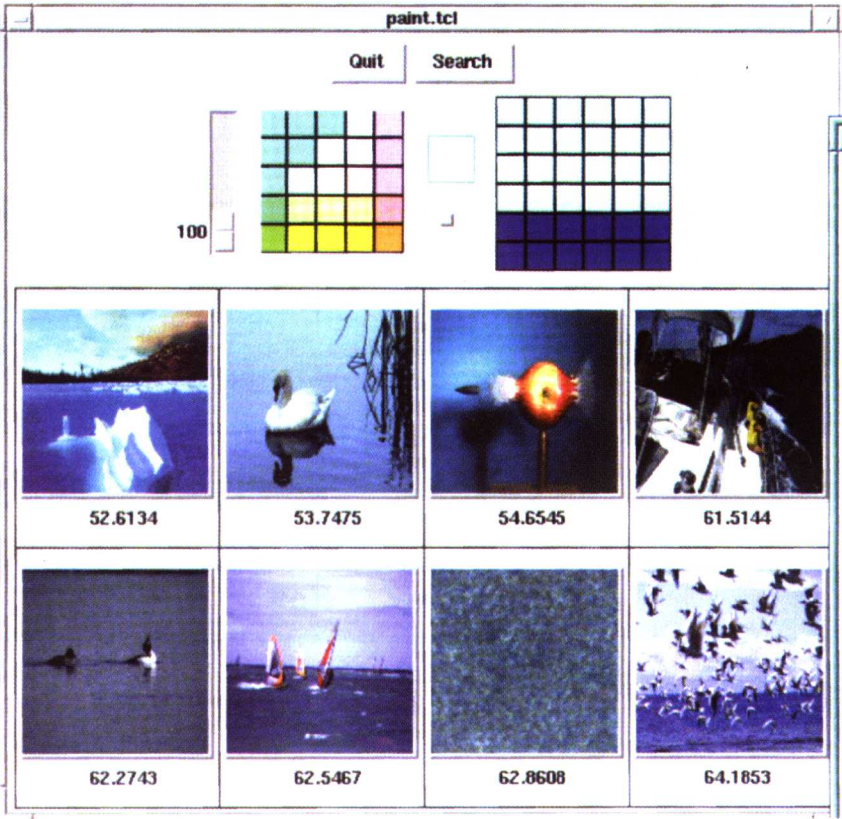


图8-6 基于纹理相似性的图像库检索结果 (来自MIT媒体实验室的VisTex数据库:
<http://vismod.www.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>)

