

视频分析算法60讲

谢剑斌 闫 玮 刘 通 编著
李沛秦 Yan Shuicheng



科学出版社

视频分析算法 60 讲

谢剑斌 闫 玮 刘 通 编著
李沛秦 Yan Shuicheng

国防科学技术大学学术著作出版资助专项经费资助

科 学 出 版 社

北 京

内 容 简

为使读者全面了解视频分析算法的原理，本书详尽地介绍了 60 多种有关视频分析的算子、描述子、变换、方法的基本理论，深入地阐述了视频分析算法的改进措施和系统地总结了其优缺

点，并提供配套的实验仿真源代码和象库。相关资料详见

www.kedachang.com。

本书特别重视如何将视频分析算法的解决视觉分析领域中的诸多基础问题，可和实验仿真有机结合，

生物特征识别和智能视频监控等领域。器视觉、大数据分析、

本书可作为计算机、自动化、电子与通信等专业高年级本科生和研究生的教材，也可作为从事视频分析与智能识别领域研发人员的参考资料。

图书在版编目(CIP)数据

视频分析算法 60 讲 / 谢剑斌等编著. —北京: 科学出版社, 2014.11
ISBN 978-7-03-124236-6

I. ①视… II. ①谢… III. ①视频系统—系统分析 IV. ①TN94

中国版本图书馆 CIP 数据核字 (2014) 第 257073 号

策划编辑: 陈 静 / 责任编辑: 陈 静 邢宝钦 / 责任校对: 郭瑞芝
责任印制: 张 倩 / 封面设计: 迷底书装

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

文林印务有限公司 印刷

科学出版社发行 各地新华书店经销

*

2014 年 11 月第 一 版 开本: 720×1 000 1/16

2014 年 11 月第一次印刷 印张: 17 1/4

字数: 330 000

定价: 79.00 元

(如有印装质量问题, 我社负责调换)

作者简介



谢剑斌，博士，教授，国防科学技术大学电子科学与工程学院研究生导师，中国生物特征识别国家标准组委员，中国数字电视国家标准组委员，中国图象图形学会高级会员，计算机学会高级会员，《电光与控制》编委会委员。长期从事海量视频分析与生物特征识别方面的研究工作，作为项目负责人主持国家级项目 8 项、部委级项目 27 项、横向课题 37 项。在国内外知名期刊发表学术论文 80 多篇，出版专著 3 部，获国家发明专利授权 20 项、实用新型专利 31 项。荣获国际发明展金奖 2 项，湖南省科技进步二等奖 1 项，公安部技术革新特别项目奖 2 项和全国发明展金奖 1 项、铜奖 1 项。



闫 玮，博士，讲师，任职于国防科学技术大学电子科学与工程学院。长期从事海量视频分析与生物特征识别方面的研究工作，作为技术骨干参与国家级项目 4 项、省部级项目 17 项。在国内外知名期刊发表学术论文 23 篇，参与出版专著 3 部，获国家发明专利授权 12 项、实用新型专利 25 项。荣获国际发明展金奖 2 项，湖南省科技进步二等奖 1 项，公安部技术革新特别项目奖 2 项和全国发明展金奖 1 项、铜奖 1 项。



刘 通，博士，讲师，任职于国防科学技术大学电子科学与工程学院。长期从事海量视频分析与生物特征识别方面的研究工作，作为技术骨干参与国家级项目 4 项、省部级项目 18 项。在国内外知名期刊发表学术论文 29 篇，参与出版专著 3 部，获国家发明专利授权 12 项、实用新型专利 25 项。荣获国际发明展金奖 2 项，湖南省科技进步二等奖 1 项，公安部技术革新特别项目奖 2 项和全国发明展金奖 1 项、铜奖 1 项。



李沛秦，博士，讲师，任职于国防科学技术大学电子科学与工程学院。长期从事海量视频分析与生物特征识别方面的研究工作，作为技术骨干参与国家级项目 3 项、省部级项目 15 项。在国内外知名期刊发表学术论文 25 篇，参与出版专著 3 部，获国家发明专利授权 12 项、实用新型专利 25 项。荣获国际发明展金奖 2 项，湖南省科技进步二等奖 1 项，公安部技术革新特别项目奖 2 项和全国发明展金奖 1 项、铜奖 1 项。



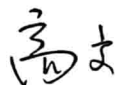
Dr. Yan Shuicheng is currently an Associate Professor in the Department of Electrical and Computer Engineering at National University of Singapore, and the founding lead of the Learning and Vision Research Group (<http://www.lv-nus.org>). Dr. Yan's research areas include computer vision, multimedia and machine learning, and he has authored/co-authored over 300 technical papers over a wide range of research topics, with Google Scholar citation >12570 times and H-index-50. He is an associate editor of IEEE Transactions on Circuits and Systems for Video Technology (IEEE TCSVT) and ACM Transactions on Intelligent Systems and Technology (ACM TIST). He received the Best Paper/Demo Awards from ACM MM'13, ACM MM'12, PCM'11, ACM MM'10, ICME'10 and ICIMCS'09, the winner prizes of the classification task in PASCAL VOC 2010-2012, the winner prize of the segmentation task in PASCAL VOC 2012, the honourable mention prize of the detection task in PASCAL VOC'10, 2010 TCSVT Best Associate Editor (BAE) Award, 2010 Young Faculty Research Award, 2011 Singapore Young Scientist Award, and 2012 NUS Young Researcher Award.

序

能从浩如烟海的视频数据中快速准确地找到有价值的信息，是大数据时代解决多媒体疑难问题的决定性手段。而视频特征提取与内容分析技术，是海量视频分析与理解的基石。《视频分析算法 60 讲》，系统地介绍了视频分析领域的经典理论和算法，以及由此产生的最新成果。

该书主要介绍视频特征提取与内容分析领域经典理论和最新成果的起源、发展、改进、实现方法和应用特点。该书分为 5 篇。第 1 篇系统介绍了包括 Harris、SUSAN 和 FAST 等在内的 8 种特征提取算子；第 2 篇系统介绍了包括 HOG、LBP 和 SURF 等在内的 8 种特征描述子；第 3 篇系统介绍了包括卡尔曼滤波、双边滤波和 Guided 滤波等在内的 10 种滤波算法；第 4 篇系统介绍了包括 DCT、小波变换和 Hough 变换等在内的 7 种变换算法；第 5 篇为本书重点，系统介绍了包括 OTSU、PCNN、Mean Shift、SVM、2D PCA 和 2D LDA 等在内的 29 种特征分析方法。

《视频分析算法 60 讲》由国防科技大学电子科学与工程学院谢剑斌教授与新加坡国立大学 (National University of Singapore) 电子与计算机工程学院的颜水成 (YAN Shuicheng) 教授合作完成。近几年，我和颜教授交流颇多，知道他年轻有为，在视频分析与检索领域建树多。该书是他与国防科大谢教授团队合作的结果。该书介绍的理论、算法、成果，对于本领域的研究生、科研和工程技术人员提高研究能力和素养，具有参考价值，值得一读。



中国工程院院士

2014 年 11 月于北京

前 言

本书是作者二十多年研究视频分析算法的心血，可作为信息、计算机、自动化、电子与通信等学科专业高年级本科生和研究生的教材，也可作为从事视频分析与智能识别领域的技术人员实现经典算法的参考资料。

全书分为 5 篇。第 1 篇介绍 Moravec、Forstner、Harris、SUSAN、CSS、FAST、DoG、LoG 等常用算子。第 2 篇介绍 Hu 矩、Legendre 矩、傅里叶、HOG、LBP、Haar、SIFT、SURF 等常用描述子。第 3 篇介绍 Butterworth 滤波、Chebyshev 滤波、椭圆滤波、递归中值滤波、最小二乘滤波、维纳滤波、卡尔曼滤波、同态滤波、双边滤波、Guided 滤波等常用滤波方法。第 4 篇介绍 K-L 变换、DCT、Gabor 变换、小波变换、Haar 变换、LPT、Hough 变换等常用变换方法。第 5 篇介绍相似性度量方法、直方图双峰法、分水岭方法、区域分裂合并方法、OTSU 方法、最大二维熵方法、二维交叉熵方法、PCNN 方法、侧抑制网络、背景减除法、时间差分法、数学形态学、光流法、Mean Shift 方法、Camshift 方法、梯度下降法、牛顿迭代法、共轭梯度法、禁忌搜索方法、罚函数方法、模拟退火方法、贝叶斯方法、K 均值聚类方法、AdaBoost 方法、SVM 方法、PCA 方法、2D PCA 方法、LDA 方法、2D LDA 方法等常用方法。

本书由国防科学技术大学电子科学与工程学院数字视频课题组和新加坡国立大学机器学习课题组联合编著。谢剑斌教授负责中文版，Yan Shuicheng 教授负责英文版。闫玮、刘通、李沛秦、穆春迪等参与实验仿真。在编著过程中得到国防科学技术大学庄钊文教授、唐朝京教授的大力支持，刘双亚、李润华等为本书编著做了大量工作，国家自然科学基金项目(61303188)对本书的相关研究工作进行了资助，在此一并致谢！由于时间有限，可能没有列全参考文献，请读者或相关作者联系作者，在本书再版和提供配套资源的网站上加入并致谢。

配套资源的网站为：www.kedachang.com。

作 者

2013 年 11 月于新加坡

目 录

序 前言

第 1 篇 算 子

第 1 讲	Moravec 算子	3
第 2 讲	Forstner 算子	5
第 3 讲	Harris 算子	8
第 4 讲	SUSAN 算子	12
第 5 讲	CSS 算子	16
第 6 讲	FAST 算子	21
第 7 讲	DoG 算子	23
第 8 讲	LoG 算子	26

第 2 篇 描 述 子

第 9 讲	Hu 矩描述子	31
第 10 讲	Legendre 矩描述子	34
第 11 讲	傅里叶描述子	37
第 12 讲	HOG 描述子	40
第 13 讲	LBP 描述子	43
第 14 讲	Haar 描述子	48
第 15 讲	SIFT 描述子	53
第 16 讲	SURF 描述子	61

第 3 篇 滤 波

第 17 讲	Butterworth 滤波	69
第 18 讲	Chebyshev 滤波	71
第 19 讲	椭圆滤波	74
第 20 讲	递归中值滤波	78
第 21 讲	最小二乘滤波	80
第 22 讲	维纳滤波	85

第 23 讲	卡尔曼滤波	90
第 24 讲	同态滤波	95
第 25 讲	双边滤波	99
第 26 讲	Guided 滤波	103

第 4 篇 变 换

第 27 讲	K-L 变换	107
第 28 讲	DCT	110
第 29 讲	Gabor 变换	116
第 30 讲	小波变换	119
第 31 讲	Haar 变换	125
第 32 讲	LPT	130
第 33 讲	Hough 变换	134

第 5 篇 方 法

第 34 讲	相似性度量方法	143
第 35 讲	直方图双峰法	147
第 36 讲	分水岭方法	150
第 37 讲	区域分裂合并方法	153
第 38 讲	OTSU 方法	155
第 39 讲	最大二维熵方法	158
第 40 讲	二维交叉熵方法	164
第 41 讲	PCNN 方法	171
第 42 讲	侧抑制网络	176
第 43 讲	背景减除法	181
第 44 讲	时间差分法	187
第 45 讲	数学形态学	192
第 46 讲	光流法	199
第 47 讲	Mean Shift 方法	205
第 48 讲	CamShift 方法	209
第 49 讲	梯度下降法	211
第 50 讲	牛顿迭代法	214
第 51 讲	共轭梯度法	218
第 52 讲	禁忌搜索方法	221
第 53 讲	罚函数方法	226

第 54 讲	模拟退火方法	229
第 55 讲	贝叶斯方法	233
第 56 讲	K 均值聚类方法	236
第 57 讲	AdaBoost 方法	239
第 58 讲	SVM 方法	245
第 59 讲	PCA 方法	251
第 60 讲	2D PCA 方法	255
第 61 讲	LDA 方法	258
第 62 讲	2D LDA 方法	261

第 1 篇 算 子

本篇重点介绍 Moravec、Forstner、Harris、SUSAN、CSS、FAST、DoG、LoG 等常用算子。

Moravec 算子是最早提出的角点检测算子，计算速度快；对噪声干扰非常敏感；兴趣值的计算方向偏少。

Forstner 算子是摄影测量中著名的点定位算子，计算速度快、精度高；受图像灰度、对比度变化的影响较大。

Harris 算子是比较稳定的点特征提取算子，对图像旋转、灰度变化、噪声和视点变换不敏感，不具有尺度不变性。

SUSAN 算子可以检测角点和边缘，精度好，具有很好的稳定性；存在采用固定阈值和定位不够精确的问题。

CSS 算子在曲线尺度空间采用高斯平滑法，滤掉噪声和不重要的微弱结构，角点检测效果好；难以确定复杂视频图像的尺度。

FAST 算子具有平移和旋转不变性、可靠性高、计算量小的特点，阈值设定依赖于人的干涉，抗噪性能较差。

DoG 算子可以很好地近似视网膜神经节细胞的视野，增加边缘和细节的可见性，实现简单；在调整图像对比度时信息量会减少。

LoG 算子结合高斯平滑滤波和拉普拉斯锐化滤波，先平滑掉噪声，再检测边缘，定位精度高；在边缘定位精度和消除噪声级间存在矛盾。

第 1 讲 Moravec 算子

Moravec 算子由美国斯坦福大学 (Stanford University) 的 Moravec^[1-2]于 1977 年提出, 通过定义兴趣值 (interest value) 进行闭值处理和非最大值抑制, 最终确定角点。

一、基本原理

点特征指图像中的明显点 (如角点、圆点等), 是图像匹配和定位中的常用特征。用于点特征提取的算子称为兴趣算子, 如图 1-1 所示, 自 20 世纪 70 年代以来出现多种各有特色的兴趣算子。

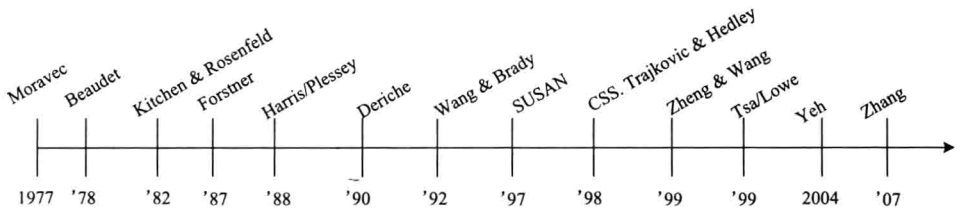


图 1-1 兴趣算子提出时间表

Moravec 算子计算待处理图像每一个像素四个主要方向 (水平、垂直、两对对角线, 即 0° 、 45° 、 90° 、 135°) 上的灰度方差, 并选择灰度方差符合最大-最小条件的像素点作为待处理图像的特征点。首先以像素四个主要方向上的最小灰度方差表示该像素与邻近像素的灰度变化情况, 即像素的兴趣值; 然后在图像的局部选择具有最大兴趣值的点作为特征点, 即灰度变化明显的点^[3]。

二、仿真实验

求取 Moravec 算子的具体流程如下。

(1) 计算图像中各像素的兴趣值。

如图 1-2 所示, 计算像素 (c, r) 的兴趣值。在以像素 (c, r) 为中心的 $n \times n$ 的图像窗口中 (如 5×5), 计算四个主要方向 (水平、垂直、两个对角线, 即图中的 a 、 b 、 c 、 d 四条线) 相邻像素灰度 (g) 差的平方和, 其表达式为

$$\begin{aligned} V_1 &= \sum_{i=-k}^{k-1} (g_{r+i,c} - g_{r+i+1,c})^2, & V_2 &= \sum_{i=-k}^{k-1} (g_{r+i,c+i} - g_{r+i+1,c+i+1})^2 \\ V_3 &= \sum_{i=-k}^{k-1} (g_{r,c+i} - g_{r,c+i+1})^2, & V_4 &= \sum_{i=-k}^{k-1} (g_{r+i,c-i} - g_{r+i+1,c-i-1})^2 \end{aligned} \quad (1-1)$$

式中, $k = \text{INT}(n/2)$ 。

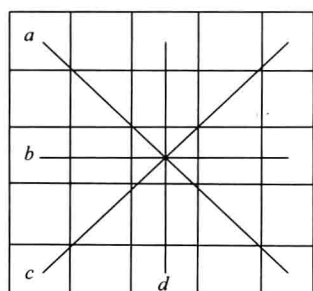


图 1-2 Moravec 算子的四个方向

取式 (1-1) 中的最小的一个值作为像素 (c, r) 的兴趣值, 即

$$V = \min \{V_1, V_2, V_3, V_4\}$$

(2) 给定一个经验阈值, 将兴趣值大于该阈值的像素作为候选点。阈值的选择应以候选点中包含所需要的主要特征点而又不包含过多的非特征点为原则。

(3) 在一定大小的窗口内, 将候选点中兴趣值最大者作为该窗口区域的特征点。该步骤称为“抑制局部非最大”, 是一种应用广泛的思想^[4]。

(4) 如果两个特征点之间的距离过短, 则去掉其中一个特征点。

三、算法特点

Moravec 算子是最早提出的角点检测算子, 简单直观, 计算速度快。

Moravec 算子没有对图像进行降噪处理, 对噪声干扰非常敏感; 对图像的边缘响应很敏感; 在计算像素点的兴趣值时考虑的不够全面。

参 考 文 献

- [1] Moravec H P. Towards automatic visual obstacle avoidance// Proceedings 5th International Joint Conference on Artificial Intelligence, 1977: 584-600.
- [2] Moravec H P. Obstacle avoidance and navigation in the real world by a seeing robot rover. Tech Report CMU-RI-TR-80-03, 1980.
- [3] 吴萌, 龚可. 关于 Moravec 算子的一些讨论. 中国科技成果, 2011, 16: 67-69.
- [4] 陈淑养. 数字图像特征点提取及匹配的研究[硕士学位论文]. 西安: 西安科技大学, 2009.

第 2 讲 Forstner 算子

Forstner 算子是德国斯图加特大学 (University of Stuttgart) 的 Förstner 等^[1]于 1987 年的 ISPRS (International Society for Photogrammetry and Remote Sensing) 研讨会上提出的, 是一种从视频图像中提取角点、圆点等特征的有效算子。

一、基本原理

Forstner 算子以待处理图像中各像素的 Robert 梯度和灰度协方差矩阵为兴趣值, 通过抑制局部极小值准则衡量各像素点的兴趣值, 提取待处理图像中的特征点^[2]。

Forstner 算子的求取过程分为以下两步。

1. 最佳窗口

如图 2-1 所示, 对于以每个像素为中心、大小为 5×5 的局部窗口, 计算其对应的兴趣值 q 和 w 。将所有兴趣值大于给定阈值 (经验值) 的窗口作为候选最佳窗口, 进而通过抑制局部非最大候选最佳窗口, 得到最佳窗口。

获取所有像素点的 q 、 w 值的计算量很大, 减少计算量的可行方法之一是先计算每一个像素点在 x 正反方向和 y 正反方向上共计 4 个 Robert 梯度值的绝对值, 然后在这 4 个值均大于某个给定阈值时才进行 q 、 w 值的计算。

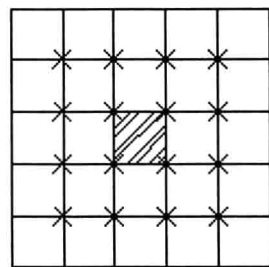


图 2-1 Forstner 算子的窗口

Robert 梯度的计算公式为

$$g(x, y) = \left[\left(\sqrt{f(x, y)} - \sqrt{f(x+1, y+1)} \right)^2 + \left(\sqrt{f(x+1, y)} - \sqrt{f(x, y+1)} \right)^2 \right]^{\frac{1}{2}} \quad (2-1)$$

式中, $f(x, y)$ 为像素灰度值。

2. 角点定位

在最佳窗口内, 通过衡量经过每个像素点的梯度直线的加权中心化结果, 可以实现圆状点的检测; 通过衡量经过每个像素点的边缘直线 (垂直于梯度方向) 的加权中心化结果, 可以实现角点的检测。

如图 2-2 所示, 设最佳窗口的左上角像素为坐标原点, (r, c) 为最佳窗口内任意像素点, 则经过 (r, c) 的边缘直线 l 可用如下方程表示, 即

$$\rho = r \cos \theta + c \sin \theta$$

式中, ρ 为坐标原点与直线 l 的垂直距离; θ 为对应的梯度角。

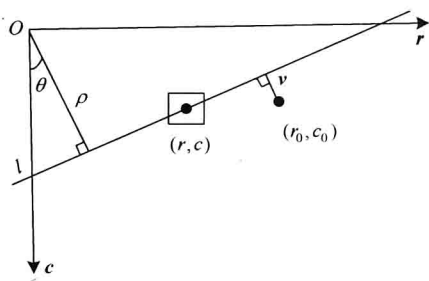


图 2-2 角点定位示意图

设角点坐标为 (r_0, c_0) , v 是角点到直线 l 的垂直距离, 则

$$\tan \theta = g_c / g_r \quad (2-2a)$$

$$\rho + v = r_0 \cos \theta + c_0 \sin \theta \quad (2-2b)$$

$$\omega(r, c) = |\nabla_g|^2 = g_r^2 + g_c^2 \quad (2-2c)$$

式中, g_c 、 g_r 为点 (r, c) 的 Robert 梯度; 权 $\omega(r, c)$ 实质上是一个边缘尺度^[3]。

对式 (2-2c) 法化, 得到法方程为

$$\begin{bmatrix} \sum g_r^2 & \sum g_r g_c \\ \sum g_r g_c & \sum g_c^2 \end{bmatrix} \begin{bmatrix} r_0 \\ c_0 \end{bmatrix} = \begin{bmatrix} r_0 \sum g_r^2 + c_0 \sum g_r g_c \\ r_0 \sum g_r g_c + c_0 \sum g_c^2 \end{bmatrix} \quad (2-3)$$

式 (2-3) 的解 (r_0, c_0) 即为所求的角点坐标。

二、仿真实验

求取 Forstner 算子流程如下。

- (1) 计算各像素的 Robert 梯度。
- (2) 计算 $n \times n$ 窗口中灰度协方差矩阵。
- (3) 计算兴趣值 q 与 w 。
- (4) 确定待选点。
- (5) 选取极值点。

三、算法特点

Forstner 算子是摄影测量中著名的点定位算子, 借助于加权中心化这一操作, 可以在最佳窗口内将定位精度提高到亚像素, 且计算速度快^[4]。

Forstner 算子需要确定阈值, 受图像灰度、对比度变化的影响较大^[5]。

参 考 文 献

- [1] Förstner W, Gulch E. A fast operator for detection and precise location of distinct points, corners and centres of circular features// ISPRS, Interlaken, 1987.
- [2] 朱庆, 吴波, 万能, 等. 具有良好重复率与信息量的立体影像点特征提取方法. 电子学报, 2006, 2: 205-210.
- [3] 王广学, 黄晓涛, 周智敏. SAR 图像尺度不变特征提取方法研究. 中国图象图形学报, 2011, 12: 2199-2205.
- [4] 刘二洋. 摄影测量中图像匹配的探讨和研究[硕士学位论文]. 西安: 西安科技大学, 2010.
- [5] 韩斌, 周增雨, 王士同. 改进的亚像素级快速角点检测算法. 江苏科技大学学报(自然科学版), 2009, 23(2): 146-149.