

基于张量投票的昆虫翅脉提取算法

段大高 龚乐 万月亮 韩忠明

(北京工商大学计算机与信息工程学院 北京 100048)

摘要 昆虫翅脉提取对于昆虫自动分类意义重大。鉴于传统翅脉提取算法中存在断裂、边缘不整齐等缺点,提出一种基于张量投票的昆虫翅脉提取算法。首先对翅脉图像进行去噪、二值化、形态学等预处理,得到稀疏二值点图,然后计算每点的张量值,结合 Gestalt 定律的邻近性、相似性规则,对邻域内点进行张量投票,并设定投票阈值,最终获取昆虫翅脉轮廓。实验结果表明,由于引入张量和 Gestalt 规则,文中所提算法可以提取出更加符合感知规则的翅脉结构,得到较为完整且平滑的翅脉轮廓,同时对于出现少许断裂的翅脉图像,依然可以获得较为完整的翅脉边缘,这为后期的昆虫自动分类奠定基础。

关键词 张量投票,昆虫翅脉,边缘轮廓,视觉感知

中图分类号 TP391.4 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.12.047

Venation Extraction Algorithm for Insect Vein Based on Tensor Voting

DUAN Da-gao GONG Le WAN Yue-liang HAN Zhong-ming

(College of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China)

Abstract Insects venation extraction for its automatic classification is of great significance. The paper discussed some problems existing in the traditional algorithms of vein extraction and proposed a new method based on tensor voting. First, a sparse binary point figure is obtained by preprocessing such as denoising, binarization, morphology and so on, and the tensor of each pixel is calculated. After performing tensor voting between pixels in neighborhood by setting up a threshold, the final vein edge is obtained. The tensor voting process combines with Gestalt law of proximity, similarity, and etc. The experiment results show that by mean of Gestalt and tensor voting, the structure corresponding to visual perception can be extracted and a more smooth and complete edge contour image can be got. What's more, even if the vein image has a little fault, the proposed algorithm still can obtain a complete outline. The result vein contour image can greatly improve the performance of the later automatic classification of insects.

Keywords Tensor voting, Insect vein, Edge contour, Visual perception

1 引言

昆虫是地球上一个主导的生命群体,有关昆虫识别的研究一直是环境污染监控、物种多样性保护、海关进出口检验检疫等领域的重点之一。早期的昆虫识别方法,多数是基于形态学及分子生物学知识,通常只有特定的昆虫学家才能掌握,并且操作起来复杂、耗时。随着计算机技术的发展,出现了一批基于形状、颜色等图像特征的昆虫自动鉴定系统。如中国农业大学的沈佐锐教授等人融合昆虫形态学与数学形态学的特点,从图像分割的角度开发出 BugVisux 自动鉴定系统^[1];王江宁^[2]等人利用人工神经网络和支持向量机技术,开发出一个基于子层次的昆虫自动鉴定系统;波兰的 Tofilski 利用昆虫翅脉交叉点等形态特征开发的 DrawWing 系统^[3],为蜂类自动鉴别提供了有力工具。然而,由于缺乏合适的模型,昆

虫自动识别在计算机视觉处理中依然是一个难题。

翅膀目昆虫是昆虫纲中的一个大的分类。翅膀上面纵横交错的脉络有着特定的形状、数目等分布特点,称为“脉序”。不同种属的昆虫,脉序特征差异万千,因此脉序在昆虫识别中占据重要地位。利用边缘检测技术提取出脉序轮廓走向、整体翅膀形状,结合翅膀的颜色分布等特征,是昆虫图像自动识别系统中的常用方法。传统提取边缘的方法通常是基于微分算子,如基于一阶微分的 Canny 算子、Roberts 算子、Prewitt 算子、Sobel 算子等,基于二阶微分的 Laplacian 算子、LOG 算子等。利用相邻像素点间的变化检测到的边缘效果通常模糊不清,且不符合视觉结构,比如效果较好的 Canny 算子针对不同的图像往往因阈值的选取问题,在图像细节和整体的取舍上难以两全,对于昆虫翅脉图像来说,小阈值保留了太多的纹理细节而呈现出管状,不利于识别过程的形状匹配,而大阈

到稿日期:2013-06-25 返修日期:2013-08-16 本文受国家自然科学基金面上项目(61170112),2013 年度教育部人文社会科学研究青年基金项目(13YJC860006),北京市属高等学校科学技术与研究生教育创新工程建设项目(PXM2013_014213_000030_00042300)资助。

段大高(1976—),男,博士,副教授,CCF 会员,主要研究方向为多媒体信息处理与数据挖掘,E-mail:duandg@th. btbu. edu. cn;龚乐(1988—),女,硕士生,CCF 会员,主要研究方向为多媒体信息处理;万月亮(1973—),男,博士,副教授,CCF 会员,主要研究方向为互联网智能信息处理;韩忠明(1972—),男,博士后,副教授,CCF 会员,主要研究方向为数据挖掘。

值则又会丢失一些边缘信息,从而导致识别困难。

同时,昆虫翅膀非常薄且脆弱,有些昆虫物种极其稀有,标本不可能重复制作,所能得到的昆虫翅膀图像往往有限,标本制作及保存过程中易发生断裂,有斑点的翅脉图像也会导致最终得到的翅脉边缘出现不连续。在这些情况下,基于上述种种翅脉提取方法得到的边缘信息就不全面。若能在翅脉出现少许断裂的情况下,从人类视觉感知结构出发,依然得到较好的边缘输出,将有助于翅膀目昆虫自动识别的后期处理。

格式塔(Gestalt)心理学家认为,人类之所以能从复杂的背景中快速准确地分辨出目标物体,是因为人类对这些物体的感知符合一定的规律,这些规律被高度总结和概括为格式塔规律^[4]。针对目前昆虫翅脉边缘提取问题,本文引入张量和格式塔规律,提出一种基于张量投票的昆虫翅脉提取算法,算法通过图像预处理和张量投票处理,最终得到较为完整且平滑的翅脉边缘,以便应用在后期昆虫自动识别与分类系统中。

2 算法原理

2.1 张量简介

张量代数是线性代数的推广,在数学上表示为一个 $n \times n$ 的 n 维矩阵,更一般的定义表示为改变向量空间坐标基变换的有序数集,它是由一个向量与该向量的转置之积生成的,所以是一个对称的正定矩阵。任一正定对称矩阵都可分解成特征值与其对应特征向量乘积的形式。二维空间中的张量是一个 2×2 的对称矩阵。假设一张量 T 的两个特征值分别为 λ_1 和 λ_2 ($\lambda_1 \geq \lambda_2 \geq 0$, 正定矩阵特征值不小于零), 对应的特征向量分别为 \vec{e}_1 和 \vec{e}_2 , 那么 T 可用式(1)表示。张量可以表示成几何空间上的椭圆或椭球。对二阶张量来说,张量就是位于二维空间上的一个椭圆 λ_1 和 λ_2 分别对应该椭圆的长半轴长和短半轴长, \vec{e}_1 和 \vec{e}_2 分别为平行与椭圆长半轴及短半轴的向量。

$$T = \lambda_1 \vec{e}_1 \vec{e}_1^T + \lambda_2 \vec{e}_2 \vec{e}_2^T \quad (1)$$

$$\begin{cases} T = (\lambda_1 - \lambda_2) \vec{e}_1 \vec{e}_1^T + \lambda_2 (\vec{e}_1 \vec{e}_1^T + \vec{e}_2 \vec{e}_2^T) \\ T = T_{Sick} + T_{Ball} \end{cases} \quad (2)$$

式(1)又可分解为式(2)的形式。其中前者称为线型张量,后者称为球型张量。二维空间中的张量有两种结构:曲线结构和点结构。张量的目的在于根据线型张量及球型张量的置信度来提取特征结构,对于二维空间来说,即曲线结构或点结构。

法线与切线垂直,所以张量可由该点的法向量生成,亦可由它的切向量生成,一般选用法线生成。 $(\lambda_1 - \lambda_2)$ 为线型张量的置信度,它表示此张量属于曲线结构的可能性,如果用法线方向生成张量,那么 $(\lambda_1 - \lambda_2)$ 的大小就表示生成该张量的点位于法向量平行与椭圆长半轴方向的曲线上的可能性大小; λ_2 为球型张量的置信度,它表示此张量属于点结构的可能性,即生成该张量的点位于多条曲线的交点上或该点是一孤立点的可能性。经计算得出,曲线结构张量的特征值为 $\lambda_1 = 1, \lambda_2 = 0$, 点结构的特征值为 $\lambda_1 = \lambda_2 = 1$ 。

2.2 算法原理

张量应用非常广泛,Medioni 首次将张量引入图像轮廓提取和分组中,提出张量投票(Tensor Voting)方法,其通过

稀疏投票和稠密投票两个步骤取得了良好的效果^[5,6]。Fischer 等研究了稀疏投票过程的迭代及曲率的调整对投票效果的改善^[7]。Loss 等提出的基于迭代、多尺度张量投票的方法,可以提高边缘检测的效果^[8]。

张量算法基于两个核心步骤,1)将像素点用张量来表示;2)张量之间的信息传递。本文选用像素点的梯度方向,即法向量,生成像素点的二阶张量,张量信息传递通过投票方式来计算。投票过程的输入是已经二值化的稀疏点图,将每一个点表示成一个二阶的对称张量,每一个点生成的张量用一个 token 表示,这些 token 的集合称为张量场。在投票过程中,每一个 token 都将自己的方向及置信度信息通过张量场的形式传递下去。本文采取两种投票方式:稀疏投票和稠密投票,稀疏投票指从一个 token 到另一个 token 间的投票,即像素点间的相互投票;稠密投票指每个 token 都向图像中所有点投票。首先进行基于球型张量的稀疏投票,重新计算张量场后,再进行基于线型张量的稠密投票。

如图1所示,设 P 点和 Q 点为张量场中的两点,它们可能是曲线结构或是点结构, O 为通过 P 点和 Q 点的密切圆的圆心, R 为此密切圆的半径, θ 为直线 PQ 与 \vec{e}_2 (短半轴方向)之间的夹角。若 P 为投票点, Q 为获票点,即 P 给 Q 投票, P 点所投出的票值大小表示 P 与 Q 同属一个结构元素的权值大小。

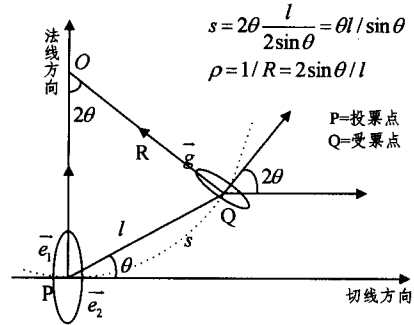


图1 投票示意图

根据 Gestalt 规则,空间上相邻、结构上相似的组织更易被感知为一个知觉结构。所以在构造 Q 点受 P 点的衰减函数(Decay Function, DF)时就要考虑到 PQ 间距离及通过 P 点和 Q 点的曲线的平滑性。 P 点投给 Q 点的票值应随着 PQ 间的距离的增大而减小,且经过 P 点和 Q 点的所有曲线中,从感知规则出发,小曲率的更符合人眼视觉的选择,即当直线和曲线共存时,倾向于选择直线形式^[9]。综合考虑后,DF 的定义如式(3)所示。 s 为 P 点与 Q 点间密切圆的弧长, C 为控制曲率衰减的因子,通常取固定值, ρ 为密切圆的曲率大小, σ 为投票尺度, s 及 ρ 的大小如图1所示,通过 DF 的定义可知 σ 是 DF 的唯一参变量。

$$DF = \exp(-(s^2 + c\rho^2)/\sigma^2) \quad (3)$$

$$\vec{g} = \begin{bmatrix} -\sin 2\theta \\ \cos 2\theta \end{bmatrix} \quad (4)$$

以 P 为坐标原点, Q 点的单位法向量 \vec{g} 的大小可求得,为式(4),其方向平行于 Q 到 O 点的直线,那么 Q 点的张量就表示为 T_a ,如式(5)所示,实验发现当 $\theta > 45^\circ$ 时,投票值 $DF \ll 0.01$,可忽略不记。定义 P 对 Q 点的线型投票函数 $S(l, \theta)$ 及

球型投票函数 $B(q)$ 如式(6)所示,球型投票可理解为是对线型投票在 0 到 2π 上的积分。

$$T_q = \vec{g} \vec{g}^T = \begin{bmatrix} -\sin 2\theta \\ \cos 2\theta \end{bmatrix} \begin{bmatrix} -\sin 2\theta & \cos 2\theta \end{bmatrix} = \begin{bmatrix} \sin^2 2\theta & -\sin 2\theta \cos 2\theta \\ -\sin 2\theta \cos 2\theta & \cos^2 2\theta \end{bmatrix} \quad (5)$$

$$\begin{cases} S(l, \theta) = DF * T_q \\ B(q) = \int_0^{2\pi} R_q^{-1} S(l_q, \theta) R_q^T d\theta \end{cases} \quad (6)$$

二阶张量的线型和球型投票域的形状及颜色强度分别如图2所示。结合图1,由图2中左侧的线型投票域可知离投票中心点较近的点所获得的票值大,颜色强度深,且当 $\theta > 45^\circ$ 时,认为是零舍去。而球型投票(图2右侧)显得没有方向性,颜色强度也呈发散状。这样每个点都对属于它投票域中的点进行投票,选择它所认为的方向,某一点的受票值为所有对它投票的点的票值的线性叠加,最后根据各结构显著性判断它所属的结构是曲线结构还是点结构。

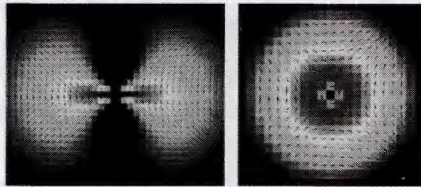


图2 线型、球型投票域

3 算法实现

本文提出的翅脉边缘提取算法分为两个步骤:图像的预处理和张量投票特征提取,如图3所示。

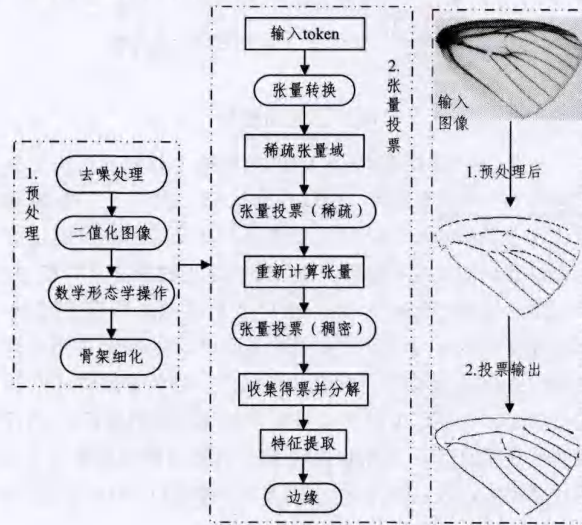


图3 算法流程

3.1 图像预处理

首先对翅脉图像进行预处理,包括采用偏微分方程(Partial Differential Equation, PDE)去噪、二值化和数学形态学操作(去除图像中孤立的亮点,断开图像中的H型连接,对图像进行闭操作等)。运用数学形态学的优点,清除掉一些孤立的点及翅脉之间的细小横脉,然后细化骨架。

通过预处理过程,得到稀疏二值图像,如图3右侧所示,原图是一幅翅脉脉络有少许断裂的图像,经预处理后的图像大致上呈断断续续的点状二值图。

3.2 张量投票

将预处理后的稀疏二值图作为张量投票的输入数据,求每个稀疏点的张量,并在投票作用域对点张量进行投票,计算图像的边缘特征。具体步骤如下:

1)求张量场:将原图从标量场转换为向量场,然后将每一点的向量乘以它的转置,即将向量场转换成了张量场。

2)基于张量场的稀疏投票:边缘像素点与点间的投票。由于点结构不具有方向性,因此对每个 token 先进行稀疏投票,即球型投票,将每个边缘点的特征值取 $\lambda_1 = \lambda_2 = 1$,并累计每个点的获票数(通过张量累加实现),重新计算张量场,这样原始的球型张量就具备了一定的方向性。

3)基于张量场的稠密投票:空间所有位置间的投票。稠密投票是每个 token 向空间中所有位于投票域中的所有点投票(如图2左侧所示,线型投票),即每个点仅投票给位于它投票邻域内的点。此时球型张量已经没有作用,仅对 λ_1 对应特征向量生成的张量作稠密投票。对每个点进行线型投票后,重新计算张量场。

4)特征提取阶段:将每个张量分解成式(2)的形式。 $\lambda_1 - \lambda_2$ 为线型张量的置信度, $\vec{e}_1 \vec{e}_1^T$ 为线型张量显著性结构, λ_2 为球型张量的置信度, $(\vec{e}_1 \vec{e}_1^T + \vec{e}_2 \vec{e}_2^T)$ 为球型张量的显著性结构(见式(2))。通过设定阈值,在各显著性图中将小于阈值的部分去掉。

5)张量场到标量场的转换:张量场到向量场,向量场到标量场的转换,即复原出原图。

4 实验结果

实验数据为完整昆虫翅膀标本图像、有少许断裂的标本图像以及理想翅脉图像。实验将本文算法与标准 Canny 算法进行对比,同时采用文献[10]中的评价方法,对本文采用的算法进行客观评价。假设 E_{DO} 和 B_{DO} 分别表示理想输出图像的边缘和背景的像素集, E_D 和 B_D 分别表示各种检测算法的输出轮廓及背景的像素集。那么真实的轮廓集合就可以表示为 $E = E_D \cap E_{DO}$, 遗漏的轮廓集合就可以表示为 $E_{FN} = B_D \cap E_{DO}$, 虚假的轮廓就可以表示为 $E_{FP} = E_D \cap B_{DO}$ 。则评价函数 P 可表示为式(7):

$$P = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})} \quad (7)$$

其中, $\text{card}(X)$ 代表求集合 X 中的元素的个数。式(7)也即检测到的真实边缘与一个交集的比值,这个交集是理想边缘与检测到的边缘的交。同时从式(7)可知, P 是 $[0, 1]$ 之间的一个数值,如果算法能将理想的轮廓完全提取出来并且没有将背景误认为轮廓,那么 $P=1$, P 越接近 1,则算法性能越好;反之越差。使用大量昆虫标本图像进行实验,表1列出其中4种类型的昆虫翅脉实验结果,可以看出本文算法结果明显优于 Canny 算法,这是因为 Canny 算法得到的翅脉轮廓显现管状,存在大量的虚假边缘,而本文算法能很好地保留翅脉轮廓边缘,去除了孤立噪声细节。本文算法的 P 值远大于 Canny 的结果,为后续分类与比对提供了稳定的图像轮廓。

表 1 实验结果对比

原始图像	Canny 阈值=0.2	Canny 阈值=0.4	Canny 阈值=0.6	本文算法
舞毒蛾 1	0.0804	0.0765	0.0130	0.1653
舞毒蛾 2	0.0572	0.0689	0.0460	0.2733
折带黄毒蛾	0.0437	0.0302	0.0111	0.2260
角棕毒蛾	0.0439	0.0446	0.0264	0.1178

图 4 使用统计学中的盒图展示了本实验中用到的 5 幅图的对比效果,它揭示了数据的离散分布情况,其中每一个盒图中最上方是实验结果的最大值,也即是实验效果最好的;相对应,最下方是实验数值结果的最小值。这里 Canny 的阈值分别取为 0.2、0.3、0.4、0.5、0.6, T 表示本文算法, C 表示 Canny 算法。从图 4 可以看出, Canny 算法的评价值较小,大致取值在 0.05 左右,本文算法的评价值大都在 0.2 左右,是 Canny 的 4 倍,远远大于 Canny 的评价值,本文算法在稳定性和一致性上都优于 Canny 算法。本文算法由于引入了张量计算和投票过程,因此计算复杂度有所提高,复杂度与 token 的数量成正比。



图 4 性能评价

直观上,本文算法也优于 Canny 算法,如图 5 所示,图像分为 3 列(a,b,c)。从前两列可以发现,Canny 算法提取到的轮廓与理想轮廓相差较远,第三列 c 图在第二列 b 图的基础上图像样本缺少少许部分(见方框标示),进行 Canny 算法与本文算法的边缘轮廓提取对比。从实验结果可以看出,当翅膀部分出现少许断裂时,由 Canny 算法得到的结果中也有明显断裂,不能直接用于识别,而本文算法结果在有少许断裂时与原完整图像的边缘没有区别。本文算法提取到的轮廓不仅与理想输出较为接近,且在在有少许断裂的情况下,输出没有明显差别,同时轮廓边缘也较为平滑。

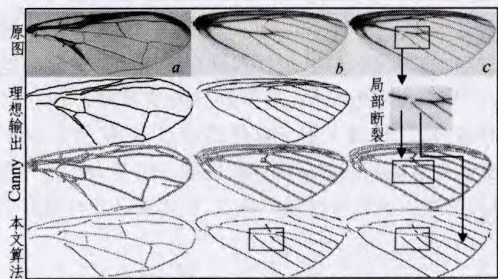


图 5 实验结果对比

结束语 昆虫翅膀在昆虫鉴别工作中起重要作用,而传统的轮廓提取方法得到的翅脉轮廓存在图像细节和整体难以取舍的问题。针对昆虫翅脉特点,本文引入张量和 Gestalt 规则,提出基于张量投票的翅脉提取算法。该算法能提取出符合感知结构的轮廓边缘,且在翅脉出现少许断裂的情况下,依然可以得到较完整的边缘信息;细化了翅脉边缘图像,在保留了局部细节特征的同时,轮廓边缘也更为圆滑,为分类与比对提供稳定的图像轮廓,为后期的昆虫自动分类奠定了基础。

参考文献

- [1] 赵汗青,沈佐锐,于新文. 数学形态特征应用于昆虫自动鉴别的研究[J]. 中国农业大学学报,2002,7(3):38-42
- [2] 王江宁,林聪田,纪力强,等. 一个基于昆虫目级阶元的新型自动鉴定系统[J]. 知识工程,2012,33:102-110
- [3] Tofilski A. DrawWing, a program for Numerical Description of Insect Wings[J]. Journal of Insect Science,2004,4(17):1-5
- [4] Kubovy M, Wagemans J. Grouping by proximity and multistability in dot lattices: A quantitative Gestalt theory[J]. Psychological Science,1995,6(4):225-234
- [5] Medioni G, Kang S B. Emerging Topics in Computer Vision [M]. Prentice Hall PTR Upper Saddle River, USA, 2004
- [6] Leo R L, Gerard M, Eduardo B C. Registration on 2D points using geometric algebra and tensor voting[J]. Journal of Mathematical Imaging and Vision,2010,37(3):249-266
- [7] Fischer S, Bayerl P, Neumann H, et al. Iterated tensor voting and curvature improvement [J]. Signal Processing, 2007, 87(11):2503-2515
- [8] Loss L, Bebis G, Nicolescu M, et al. An iterative multi-scale tensor voting scheme for perceptual grouping of natural shapes in cluttered backgrounds[J]. Computer Vision and Image Understanding, 2009, 113(1):126-149
- [9] Grigorescu C, Petkov N, Westenberg M A. Contour detection based on nonclassical receptive field inhibition[J]. IEEE Transactions on Image Processing, 2003, 12(7):729-739
- [10] Gioi R G, Delon J, Morel J M. The collaboration of grouping laws in vision[J]. Journal of Physiology-Paris, 2012, 106(5/6): 266-283