

几何形态测量法在生物形态学研究中的应用

白 明^{1, 2*} 杨星科^{1**}

(1. 中国科学院动物研究所 北京 100080; 2 中国科学院研究生院 北京 100039)

Application of geometric morphometrics in biological researches. BAI Ming^{1, 2*}, YANG Xing-Ke^{1**} (1. *Institute of Zoology, Chinese Academy of Sciences, Beijing 100080, China*; 2 *Graduate School, Chinese Academy of Sciences, Beijing 100039, China*)

Abstract The analysis of shape is a fundamental part of many biological researches. As the field of statistics developed, so have the sophistication of the analysis of these types of data. In the 1980s, a fundamental change began in the nature of the data gathered and analyzed. This change focused on the coordinates of landmarks and the geometric information about their relative positions. This new approach, called “geometric morphometrics”, had benefits that lead Rohlf and Marcus (1993) to proclaim a “revolution” in morphometrics. In this paper, we briefly introduce Geometric morphometrics.

Key words morphology, measurement, geometry, outline, landmark.

摘 要 形态分析是生物系统学及其多样性研究中很重要的一部分。随着统计学的发展, 我们可以对非常复杂的数据进行分析, 这在客观上导致了多变量形态测量的出现。在 20 世纪 80 年代, 在数据收集和分析上产生了重要突破——标点和标点相对位置的几何信息的匹配, 从而可以将多变量分析的标点集叠加到生物原始图上, 它不仅仅是生成散点图, 而是试图客观反映生物的形态性状。这项研究被称为几何形态测量法(geometric morphometrics), Rohlf 和 Marcus(1993)称其为形态测量方法上的一次革命。文章简要介绍了该方法。

关键词 形态, 测量, 几何, 标点, 轮廓

几个世纪以来, 生物解剖性状的比较已经是生物研究的核心问题, 因为生物的系统分类和生物多样性的研究, 主要基于形态描述。在 20 世纪早期, 生物学开始从描述和定性学科逐渐向量化过渡, 形态学分析也如此。形态学研究出现了可测量性状的测量及对不同性状采用平均值进行类群间的比较等。统计方法的发展提供了更先进的严格的定量分析, 截至 20 世纪中期, 形态性状的数量描述综合了类群内和类群间的形态变化的统计分析, 这也是现代形态测量方法的开始。

形态测量法主要研究形态变量及其协同变量与其他变量的关系。传统的形态测量法是多变量统计分析在形态变量数值化中的应用, 如长度、宽度和高度。在 20 世纪 80 年代后期和 90 年代前期, 在形态结构数值化和数据分析的

方法上有了重要突破。该方法最初是强调捕捉感兴趣的几何形态结构, 通过一定规则测量其结构, 将其转换成数字信息, 并分析保留这些信息。Rohlf 和 Marcus 称其为几何形态测量法(geometric morphometrics)^[1]。经过 10 余年的发展, 更多的新方法产生出来, 几何形态测量法进入一个更加成熟的阶段。本文简单概述目前几何形态测量法的研究进展, 努力去说明这个现代形态分析中的重要方法, 试图探讨几个将来可能的发展方向。

1 传统的形态测量法

在 20 世纪 60 ~ 70 年代, 生物统计学家开

* E-mail: baim@ioz.ac.cn

** 通讯作者, E-mail: yangxk@ioz.ac.cn

始应用多变量统计工具描述类群内和类群间的形态变异,这被称为传统的形态测量法,或称为多变量形态测量法,包括应用多变量统计分析方法对形态变量的分析,如线性距离的测量,数量、比例或者角度大小。由于线性距离测量通常与大小非常相关,因此可以对不定形结构的形态变量进行分析和对形态变异的模式进行研究。对于异速生长的研究也很多。

虽然传统形态测量法结合了多变量统计和形态性状的数值化,但仍然存在一些问题。例如:(1)线性距离的同源性很难评估,因为很多距离(如最大宽度)不能被作为同源性状。(2)从不同的形态可以获得相同的距离数据,因为测量点并未包括在数据中。如测量卵形和泪珠形的最大长度和最大宽度可能会获得相同的数据,但他们是完全不同的形状。因此,统计学在辨别形状时的能力被高估了。(3)通常不可能生成图形来代表线性距离,因为变量间的几何关系并未获得保存(一系列线性距离通常很难捕获原始物体的几何形态)。因此,物体的有些方面性状不可避免的会丢失。这些缺点的存在,给应用带来很大困难。

2 几何形态测量方法的革命

由于以上的问题,研究人员试图通过不断努力,来克服它的缺陷,所以大家不断探索出其他量化形态性状和分析方法。他们的研究兴趣主要集中在几何形态结构的数据(包括轮廓数据和标点数据)获得和这些数据的分析方法上。与这些研究同时出现的是 David Kendall 和其他统计学家提出了一个更加严谨的形态分析理论,该理论可以综合多变量统计和将生物形态直观化^[2]。Bookstein 称其为形态测量的综合方法^[3]。Rohlf 和 Marcus 对它们的评价是几何形态测量方法的革命,为什么呢,我们将对其进行详述。

2.1 轮廓线法(outline methods)

轮廓法是首先应用的几何形态测量方法,当某结构或某区域的边缘具有同源性时,在边缘曲线上选取的样点并不要求完全一致。这个

方法通常被用在将轮廓上的样点数字化上,这些样点符合一定数学函数(通常是 Fourier 分析),然后将形态变量作为函数系数的多变量进行分析就可以比较曲线的差别。多变量参数区间(如 Fourier 系数)中的点可以转化为生物体物理上的距离区间,并可以直观化为轮廓。早期的方法是应用从边缘到中心的半径长度,中心可以是标点,也可以是物体的质心。因为这个方法仅限于简单物体的边缘,所以其他方法如参考轮廓线上每点的切线的角度变化,对沿曲线距离函数的 Δx 和 Δy 值进行分析,或者将沿曲线上的点相关性视为一系列复杂的数字。虽然这些方法有一定的解决能力,但问题是基于不同方法得出的统计结果并不一致,而且对研究人员如何选择最好的理论仍然没有一致意见。

2.2 标点法(landmark methods)

标点是在生物结构中选取的样点,而且在不同标本中结构的变化可以通过标点反映出来。标点法最初是基于生物可定义的标点的 2 维或 3 维相关性上。然而标本上标点的位置、取向和比例的变化会引起直接分析的相关性的变化。因此,必须利用数学去除非形态变异对分析的干扰。一旦非形态变异的干扰能够去除,形态变异的变量可作为比较不同样本间统计工具,图形化的比较也成为可能。

要去除非形态变异的干扰,虽然有一些方法获得应用,但这些方法使用的理论和最优化原则稍有不同。目前常用的是叠印法。图 1^[4]示例了叠印法的 3 个主要步骤:(1)未加工数据的量化;(2)去除非形态变异的影响;(3)统计分析和结果的图形化。2 点配准是一种简单的叠印方法,它是 Bookstein 在 20 世纪 80 年代晚期提出的形态分析理论的基础。该方法通过基于最优化原则叠加标点消除了非形态变异的影响。另外,还有 2 种方法,一种是 GPA,另一种是 GRF。GPA (generalized procrustes analysis; 早期指 GLS, generalized least squares)是使用最小正方形估计平移和旋转参数,从而将标点布局进行叠印。在这个过程中,要重复计算形状的

平均值,但无法对叠印进行提前估计。当很多形态变量仅因为很少的标点而受到限制时,GRF(generalized resistant fit)可以用于这种变量模型。GRF 可将叠印参数作为中值,而不是最小正方形估计。旋转角度和比例作为标点子集的中值,平移是简单的坐标中值。在 GPA 中,这个步骤可以反复的允许样本叠印。与 GPA 相比,GRF 不能进行进一步的统计分析。

叠印后,形态差异可以通过物体间标点的相关一致性来体现。这种差异也可以通过形态变异中多变量比较数据体现,如薄板样条函数

(thin-plate spline)可以被用于绘制不同物体间形态上的变异。在该方法中形态差异是通过 D'Arcy Thompson 的栅格转化理论经严格数学计算实现的。该理论首先将一个物体通过变形或反卷为另外一个物体,接着物体的差异可被描述为栅格变形的差异。描述这些变形的参数(部分反卷数,partial warp scores)可以作为形态变异总体内和总体间的统计比较时的变量。在 20 世纪 90 年代早期该类研究被称为相对反卷分析(relative warp analysis),并获得很大关注。

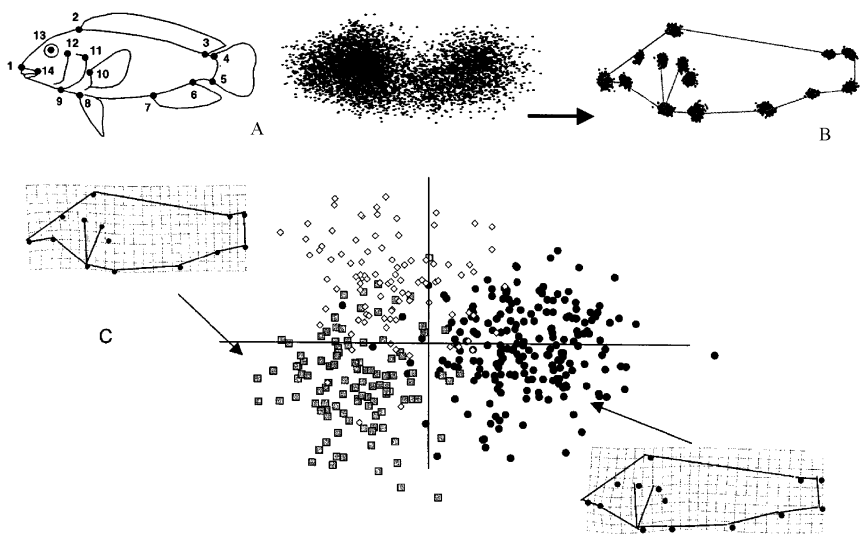


图 1 叠印法标点数据分析的步骤(引自 Rüber and Adams 2001)^[4]

- A. 未加工数据的量化(数据取自非鲫(丽鲷科)) B. 去除非形态变异的影响(从 412 号标本获取数据)
- C. 统计分析和结果的图形化(对左侧 *Spathodus erythron* 和右侧 *Eretmodus cyanostictus* 栅格化变形)

3 几何形态测量法的成熟

形态测量法革命性发展的标志是形态学研究从线性距离分析到标点和轮廓的分析。在 20 世纪 90 年代早期,几何形态测量法获得广泛认可,生物学家在不断熟练使用的基础上将标点法和轮廓法逐渐应用到众多领域。另外,20 世纪 80 年代时,Kendall^[5, 9]及其他专家在标点法方面所做的工作,使标点法获得了更加深厚的数学基础的支持,20 世纪 90 年代的工作

是对方法的深入理解,而不是基础理论的突破。

随着该方法的广泛应用,近些年在理论方面也获得了重要突破。例如,20 世纪 90 年代早期,很多不同的标点方法同时出现,但对如何选择这些方法却没有一致意见。后来本领域的学者通过辛勤的工作,对存在的缺陷进行了弥补,因此出现了对标点分析的一套方法,这些方法基于普氏距离(procrustes distances)或他们的切线距离近似值。当今标点分析几乎全用标点数据分析法(analysis of landmark data)。一些最

近的研究同时使用分析和模拟的方法作为比较标点的工具, 经证明这些方法是形态统计比较中比较好的方法。另外, 在轮廓数据分析中也涌现出很多新方法。滑动半标点法 (sliding semilandmarks) 将轮廓法和标点法结合起来分析数据, 可以对形状进行更详细的描述。最后, 因为几何形态测量数据被用于更多的生物学领域, 从而需要更多新的理论方法。

3.1 三维数据的获得、分析及其问题

早期使用几何形态测量法所获得和分析的数据均为二维数据, 研究对象也是二维物体或者三维物体的二维照片。而具有可比性的照片必须通过被研究物体的固定轴向拍照。三维数据的获得则需要特殊且较昂贵的设备的支持。如映射显微镜 (reflex microscope), 3D 激光扫描仪 (3D laser scanners) 等设备。数据分析方法与二维数据分析相似。三维形态测量法难以普及, 其主要原因除了数据获得需要昂贵设备之外, 还有一个原因就是成果发表问题。目前大多数的学术刊物仍然是纸质的二维媒介, 对于三维的图像无法充分表达。但随着越来越多的刊物网络版的发行, 三维图像可以通过网络在计算机上完美的展示。目前获得三维数据的设备价格已便宜了很多, 如果将来大多数刊物都有网络版的话, 那么三维数据的图像发表和方法普及都将不再成为问题。

3.2 几何形态测量法与系统发育分析

几何形态测量法使用的数据对形态结构的描述非常深入和完全, 所以这些数据应该可以用于支序系统发育分析。但将这两者结合, 却有非常大的困难。首先是几何形态测量法要捕捉全部可能的形态变异, 其变量是连续的, 但支序分析中的性状是非连续的。其次支序分析中要获得的是最简的系统发育树, 但对几何形态测量法来说很难做到这点。

两者的结合虽然很难, 但还是有一些相关研究进展。基于一定的理论标准, 将原始的形态数据编码为非连续的性状状态, 从而可以作为支序分析的输入数据^[7-9]。但是在编码过程中, 有些形态的信息会丢失。

4 前景和展望

如今几何形态测量法通过与其他研究手段结合, 开始向其他研究领域渗透。如非对称性结构的研究, 内部解剖对称结构的任意横向距离的研究, 不定型非对称性结构的研究。尤为重要的是利用该方法在定量遗传学中的应用。如对小鼠上颌遗传和表型的协方差研究; 通过子代形态性状对亲代性状的回归, 从而对形态的遗传可能性做出评估。

几何形态测量法目前在医学、考古学、人类学和脊椎动物学等领域已有广泛应用, 但限于昆虫的个体尺度较小, 该方法在昆虫学领域的应用仍处于起步阶段。Pretorius 和 Scholtz 利用几何形态测量法基于后胸叉骨 (metendosternite) 对金龟总科高级阶元系统发育关系进行了推断, 他们还积极倡导将几何形态测量法应用到昆虫进化研究中^[10]。我们认为, 如果能够利用 3D 扫描仪采集昆虫模式标本的三维数据, 经过软件处理成为三维图像后发布到互联网上的数据库, 广大的研究人员将足不出户就可以“检视”到模式标本。因为三维图像是模式标本三维结构的还原, 该图像是三维且可以任意旋转, 理论上讲, 观察该图像与从任意角度观察模式标本几乎没有区别。也可以通过与 3D 打印机、金属雕刻机或者机床联用, 将三维图像“打印”出来, 从而可以获得模式标本的复制品, 甚至可以获得高精度 (体毛、浅沟纹等的完美再现) 和按比例放大的模式标本复制品。更为重要的是模式标本三维数据的共享, 将为进一步的研究提供极大的便利。

虽然经过 10 余年的发展和成熟, 但当前几何形态测量法仍然不是占据主流地位的研究方法。其中原因有三: 首先, 该方法仍然在不断发展, 日益涌现的新方法和新理论会令研究人员无所适从; 其次, 该方法需要比较多的数学知识和计算机软件应用能力; 第三, 三维数据的获得需要昂贵的设备支持。我们相信随着研究人员对该方法的熟悉和应用以及相关设备造价的降低, 该方法一定会有更大的发展, 也会应用到更

多的研究领域。

致谢 以下学者为本文完成提供部分资料, 并且提出很好的建议, 在此一并感谢: F. James Rohlf 教授(美国纽约大学), Dennis E. Slice 博士(奥地利维也纳大学)。

参 考 文 献

1 Rohlf F. J., Marcus L. F. *Trends Ecol. Evol.*, 1993 **8**(4): 129~132

2 Adams D. C., Rohlf F. J., Slice D. E. *Ital. J. Zool.*, 2004, **71**(1): 5~16

3 Bookstein F. L. *Bull. Math. Biol.*, 1996, **58**(2): 313~365

4 Rüber L., Adams D. C. *J. Evol. Biol.*, 2001, **14**(2): 325

~ 332

5 Kendall D. G. *Adv. Appl. Prob.*, 1985, **17**(2): 308~329.

6 Kendall D. G. *Bull. London Math. Soc.*, 1984, **16**(1): 81~121

7 Zelditch M. L., Swiderski D. L., Fink W. L. In: Weins J. J. (eds.), *Phylogenetic Analysis of Morphological Data.*, Smithsonian Institution Press: Washington, 2000. 37~83.

8 Fink W. L., Zelditch M. L. *Syst. Biol.*, 1995, **44**(3): 343~360

9 Swiderski D. L., Zelditch M. L., Fink W. L. In: MacLeod N., Forey P. L. (eds.), *Morphology, Shape and Phylogeny*, Syst. Ass. Spec. Taylor and Francis: London. 2000. 67~99.

10 Pretorius E., Scholtz C. H. *Biol. J. Linn. Soc.*, 2001, **74**(1): 35~50.

秉志与昆虫学的渊源

瞿启慧

(中国科学院动物研究所 北京 100080)

秉志这个名字, 对于当年轻一代的昆虫学工作者来说, 肯定是陌生的。因此, 有必要先介绍一下秉志其人其事。秉志是我国著名动物学家, 我国生物学界的开创者, 近代动物学的主要奠基人。早在 90 多年前, 他就大力倡导我国的科学事业, 与留美同学共同发起组织我国最早的民办自然科学学术团体——中国科学社, 并集资创刊发行我国最早的学术刊物《科学》。从 20 世纪 20 年代起, 他积极从事我国生物学的组织领导工作和教学、研究工作。他是我国第一个生物学系和第一个生物研究机构创办人, 中国动物学会的创始人。他培养出一批不同分支领域的早期动物学英才, 成为 20 世纪我国教育界和科技界的重要骨干。他在脊椎动物形态学、神经生理学、动物区系分类学、古动物学等领域进行了大量开拓性的研究。解放后, 他全面系统地研究了鲤鱼实验形态学, 充实和提高了鱼类生物学的理论基础。他毕生为开创和发展我国的生物学事业做出了卓越的历史性贡献。

秉志与昆虫学有什么关系呢? 他于 1909 年参加当时清政府游美学务处第一期庚款留学生考试。录取后赴美留学, 进入美国康奈尔大学农学院昆虫系学习, 于 1913 年获学士学位。然后他继续在该校研究院深造, 师从著名昆虫学家 J.G.Needham 研究昆虫学, 于 1918 年获博士学位。他于 1915 年在《Pomona J. Entomol. & Zoo》上发表首篇学术论文“加拿大金杆草上虫瘿内的昆虫”。加拿大金杆草是一种菊科植物, 其茎部有许多虫瘿, 是由一种蝇类 *Eurosta solidaginis* Fitch 的幼虫所形成。他在仔细观察了 3 300 个虫瘿的基础上详尽地描述了该种蝇类的生活史, 其幼虫形成虫瘿之过程, 以及虫瘿内的其他昆虫种类, 包括鞘翅目、膜翅目、双翅目、鳞翅目等多种寄食昆虫和寄生昆虫的生活史, 进入虫瘿的途径, 等等。该文是中国人在国外正式发表的最早的昆虫学研究论文。1917 年又发表了关于摇蚊的习性的论文。

秉志又是第一位在美国研究昆虫学获得博士学位的中国人, 他的博士学位论文“一种咸水