

从摄影测量到计算机视觉

龚健雅^{1,2} 季顺平¹

1 武汉大学遥感信息工程学院,湖北 武汉,430079
2 武汉大学测绘遥感信息工程国家重点实验室,湖北 武汉,430079

摘 要:首先回顾了摄影测量的历史,从透视几何、成像设备、摄影平台、测量法和测量工具等 4 个方面较系统地总结了前人的贡献。其次,简要介绍了计算机视觉的起源,并从几何角度分析了计算机视觉与摄影测量之间的紧密联系,探讨了两者在实用上的一些区别。再次,从语义方面,分析了遥感学科的发展,与机器学习和计算机视觉之间的关系,以及目前深度学习和连接主义的盛行。最后,展望了摄影测量的未来,指出与计算机视觉、人工智能等学科的进一步交叉融合是摄影测量发展的必然之路。

关键词:摄影测量;计算机视觉;遥感;机器学习

中图法分类号:P237.9; P208 **文献标志码:**A

1 摄影测量溯源

摄影测量学是一门利用光学像片研究被摄物体的形状、位置、大小、特性及相互位置关系的学科。一般的共识是,法国测量学家 Fourcade 首先发现了用立体照片可重建立体视觉,从而促成了摄影测量学的诞生。若从技术拼图的局部出发,仍可继续溯流而上。光学成像的物理法则是小孔成像,对应的几何原理是透视变换,这也是摄影测量学的理论基础,拼图的第 1 个板块。早在文艺复兴初期,意大利建筑师布鲁内莱斯基就已经建立起几何意义上的透视法;同时代的著名画家马萨乔,第一个运用透视法作画。从此开始,西方画作能够如照片般以人眼对景深的真实感知来表现;晚清传教士带来透视法后,中国美术家也开始透视图作图。1838 年,物理学家惠斯顿发明了实体镜,第一次发现和定义了立体视觉,也为 Fourcade 的工作铺平了道路。

摄影测量的第 2 个板块是成像设备。区别于大地测量,相机和照片帮助人们将野外测量工作搬运至室内。“内业”工作成为主体,照片替代了三脚架、经纬仪和标尺,成了主要的研究对象。19

世纪早期,德国教授舒尔兹发现银的混合物在日光下会变黑。1839 年,法国画家达盖尔发明了银版摄影法,并制作了世界上第一台真正的照相机。直到上世纪 80 年代,银的混合物都是地球上最好的感光材料。

不过,竞争对手早在上世纪初就已出现。1887 年,物理学家赫兹发现电子可以捕获光。1900 年,普朗克提出能量量子化假说。1905 年,爱因斯坦利用量子化假说成功诠释了赫兹发现的光电效应。64 年后,贝尔实验室的博伊尔和史密斯发明了一种称作为电荷耦合器件(charge coupled device, CCD)的元件^[1],这是一种高感光度的半导体单晶材料。从此,经过不断的技术改进,基于量子物理的感光材料逐渐赶超基于化学原理的银感光材料,并在本世纪初淘汰了后者。基于光电效应的成像方法称作数字成像,对应的相机称作数码相机。这不但是促成模拟摄影测量转变为数字摄影测量的关键,同时也极大促进了计算机视觉的发展。

摄影测量的第 3 个板块是载体。狭义上的摄影测量一般指航空摄影测量,成熟的飞行平台是重要的组成部分。利用比重原理,制造比空气轻的物体是实现飞行的古老思路。早在三国时代,传说诸葛亮发明了孔明灯,背后原理即热空气比

常温空气更轻。真正的载人平台要等到 18 世纪,西方的孟格菲兄弟重新发明热气球,并于 1783 年第一次载人航行。一年后,法国的罗伯特兄弟乘坐氢气球飞上天空。1858 年,法国摄影师纳达尔乘坐气球拍摄了世界上第一张航空影像。1892 年,著名的卡尔文勋爵论断:比空气重的飞行器是飞不起来的。11 年后,莱特兄弟驾驶自制的世界上第一架飞机上了天。随后,两次世界大战迫使

科技突飞猛进。1957 年,第一颗卫星被发射到外太空,同时开启了卫星摄影测量时代。2000 年前后,各国陆续开始深空探测项目,比如中国的嫦娥登月 and 美国的火星探测。虽然没有“深空摄影测量”的明确说法,摄影测量学者对基于光学的深空探测项目也贡献良多。图 1 中列出了摄影测量丰富多彩的载体平台。



图 1 多样化的摄影平台(上排:手持仪器架、地面移动测图系统、无人机;下排:无人飞艇、国产运 12 航摄飞机、测绘卫星;中排:嫦娥月球探测车)

Fig. 1 Multiple Photogrammetry Platform (up: Handheld Device, Mobile Measurement System and Drone; Below: Unmanned Airship, Domestic Yun-12 Plane for Aerial Photogrammetry and Mapping Satellite; Middle: Change Lunar Rover)

摄影测量的第 4 个板块是测量法和测量工具。1795 年,18 岁的高斯发明了最小二乘法。这是大地测量(包括摄影测量)最实用、流传至今的误差处理方法。可惜的是,那个时代还没有航空摄影测量,用大地测量方法绘制汉诺威公国的地图,花去了高斯 30 年的时光。1959 年,德国的 Schmid 教授提出光束法区域网平差,这是小孔成像的物理原理与最小二乘的完美结合。光束法平差不但在相当长的时间内代言了摄影测量,也是计算机视觉几何部分的重心。

至于测量工具的演进,分别开启了摄影测量的 3 个时代,也反映了机械制造、计算机和数学运算的发展。1900~1960 年,一般称作模拟摄影测量时代,即用光学和机械的模拟仪器来重现光线的直线传播和交会。这些仪器包括立体坐标量测仪和立体测图仪两大类(见图 2)。前者用于像片的定位,后者用于测图。1960~1990 年,一般称为解析摄影测量时代,推动者是迅猛发展的计算机以及专业的解析测图仪。光线的重现和交会不

再依赖于模拟仪器,而是在计算机中以虚拟的形式实现并解算,但像点的坐标量测和最终的测图,依然要靠机械和人工作业完成。从 20 世纪 80 年代末期开始,数码相机逐渐取代光学相机,计算机更加强大,基于机器学习、模式识别等先进技术,像点自动量测和自动测图成为可能,由此开启数字摄影测量时代。

2 摄影测量与计算机视觉在几何上的联系

计算机视觉可简单概括为“用计算机代替人眼,从图片中重建和解译世界”,在定义上与摄影测量就有很高的重叠度。1940 年左右,贝尔实验室的 Julesz 发现任意视差图都可恢复立体感,而无需事先识别单幅图像的含义,而此前心理和神经学家认为人需要先感知才能产生立体感。由此 Marr 认识到复杂的神经和心理过程可用直接的数学计算表达,并由此提出 3D 重建的计算视觉

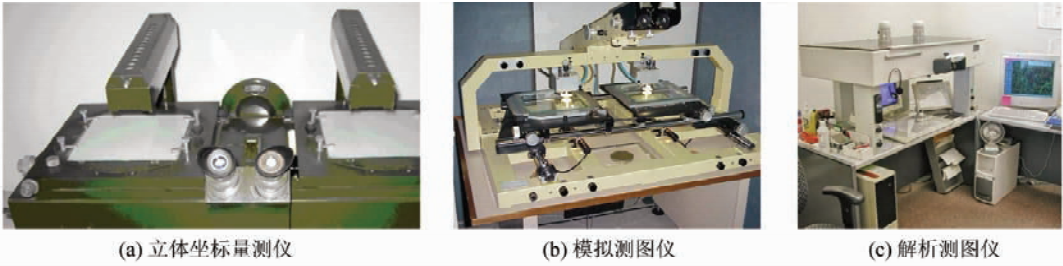


图 2 立体坐标量测仪、模拟测图仪和解析测图仪
Fig. 2 Stereocomparator, Analog Plotter and Analytical Plotter

理论。他在 1982 年发表的《视觉:从计算的视角研究人的视觉信息表达与处理》^[2]详细分析了二维图像的表达、立体图像的对应和重建、算法以及硬件的实现,是计算机视觉的开山之作。20 世纪 80 年代也是计算机视觉的第一段黄金时期。许多经典的算法和算子,如 Canny 边缘检测^[3]、Shape from Shading^[4]、Hough 变换^[5]、LoG (Laplace of Gaussian)^[6]等都起源于七八十年代。除了图像处理,早期的计算机视觉同样关注几何问题。不过,在计算机和数码相机成熟之前,摄影测量与计算机视觉当时的状况距离 Marr 所提倡的 3D 重建理论依然相去甚远。

20 世纪 90 年代后,摄影测量与计算机视觉都得到物理和技术领域的强大推动。仅几何而言,两门学科具有同样的理论基础,即小孔成像和双目视觉原理,因此两个领域的学者们都在处理相似的问题,但也有细微的区别。如相机检校,摄影测量一般布设高精度三维检校场,而计算机视觉常采用二维平面棋盘。如空中三角测量,摄影测量一般用严密的光束法区域网平差,而在计算机视觉中一般称为从运动恢复结构(structure from motion, SfM)^[7],除了用全局的光束法平差,也采用一些非全局解法,如增量式的局部平差、滤波^[8]等。这些差异来源于摄影测量需要更高的测量精度,即所谓的摄影测量精度。此外,还有一些近似等价的概念,由于学科的发展而具有不同的称呼。如平差中的粗差探测,摄影测量中称为选权迭代的方法,在计算机视觉中称为权衰减法;如为了处理法方程系数阵(即 Hessian 矩阵)接近奇异的问题,计算机视觉中常用 L-M 算法^[9],而卫星摄影测量中常用岭估计法^[10]。一般而言,计算机视觉中理论的严密性要高于摄影测量,这可从两门学科在几何上的代表作,《摄影测量原理》^[11]与《Multiple View Geometry in Computer Vision》^[12]中可以看出明显的差别;当然,摄影测量在工程和实用性方面可能更占优势。

摄影测量与计算机视觉的最大区别并非在几何理论上,而是在实践中。传统摄影测量集中于航空和航天平台,处理航空图像和卫星遥感图像,所用的相机通常为专业量测相机,最初的目的是制作地形图和各类专题图,为测绘行业服务。计算机视觉以普通相机、大众数据为主,相对而言行业面更加广泛。然而,在进入 21 世纪后,两者的融合速度又得到进一步提升。这很大程度上得益于两者共同关注的无人机和车载移动平台。无人机航摄不仅是摄影测量中的一个方便快捷的成像技术,也是计算机视觉所关注的未来焦点,是各类信息获取的重要渠道。与此类似,摄影测量中以地面移动测量系统(mobile mapping system, MMS)采集道路和街景;而计算机视觉同样关注道路信息(以及室内场景)的提取与重建,并应用于机器人、城市地图、智能交通和自动驾驶汽车中,并由此产生了一个称为同时定位与地图构建(simultaneous localization and mapping, SLAM)^[13]的研究支流(见图 3)。早期的 SLAM 以激光扫描仪为主,这也是名词中 mapping 的由来;后来基于视觉的 SLAM,即 Visual SLAM 成为主流,并与摄影测量特别是实时摄影测量在各个研究点上(匹配、平差、定位和重建)都有共通之处。SLAM 与空三的明显区别在于,SLAM 定位的同时生成了半密集点云,这些点云可以通过激光扫描获得,也可通过图像匹配获得。

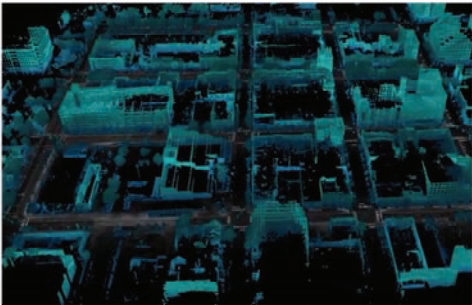


图 3 同时定位与地图构建(SLAM)
Fig. 3 Simultaneous Localization and Mapping(SLAM)

3 遥感、机器学习与未来

§ 2 讨论了摄影测量与计算机视觉在几何上的关系。事实上,在语义方面,两者的联系同样紧密,同时语义也是未来共同的主流研究方向。摄影测量的延伸是遥感。1972 年,第一颗 Landsat 卫星升空,标志着遥感学科的建立。在摄影测量已经解决大部分几何问题的前提下,遥感的工作重点就集中在“解译”上。解译是回答“是什么”和“为什么”的问题,与语义方向的计算机视觉、模式识别、机器

学习等异曲同工。在遥感中,相对于摄影测量中的可见光,光谱段被大大扩充(至多光谱)并细化(至高光谱)。这些辐射信息有利于解决农业、测绘、环境、地质和地理所关心的宏观问题,如土地利用覆盖分类、农作物趋势分析、大气的长期变化监测、泥石流和洪水等自然灾害的评估与预测(见图 4)。因为数据源的特殊性,遥感领域显然也发展了一些特有的算法。但是,大部分机器学习方法依然在这些相关领域通用,如支持向量机^[14]、随机决策树^[15]、条件随机场^[16]等。

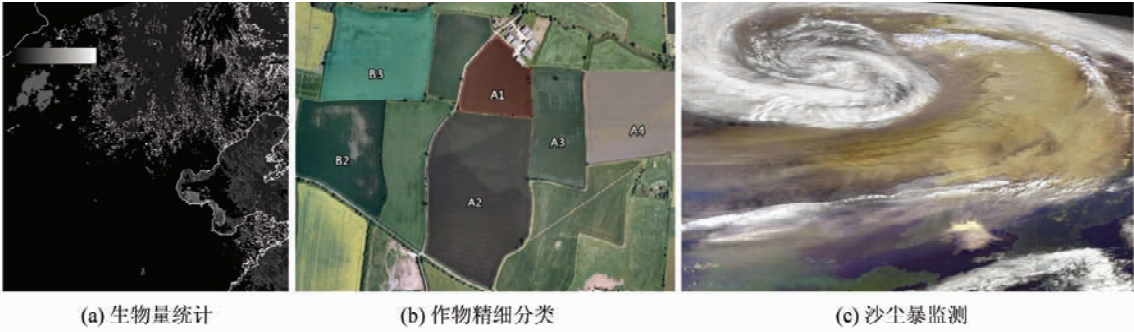


图 4 遥感关注的宏观问题

Fig. 4 Those Remote Sensing Focused on Macroscopic Problems

抛开数据源的区别,从理论源头上看,符号主义流派在 1956 年提出“人工智能”的概念,并统治了该领域将近半个世纪,基于统计学习的思想被广泛应用于摄影测量与遥感、计算机视觉。而连接主义者所倡导的神经网络模型和感知机在经过八九十年代的寒冬后,在 2000 年后卷土重来,更名为“深度学习”,并大有一统天下的趋势。深度学习利用无需人工设计的过量参数去拟合一个可能非常复杂的非线性目标函数,显示了强大的泛化应用能力。2012 年,在 ImageNet 挑战赛中,深度学习的方法夺得第一,并一举超过传统方法 10 个百分点^[17];而第 2 至第 4 名相差不超过 1%,显示了传统方法的天花板与艰难处境。随后的大量实验表明,在图像分类、物体识别^[18]、语音识别^[19]、遥感应用^[20]等关于学习和语义的研究领域,深度学习都占据上风。目前,许多摄影测量与遥感中的实际应用,如道路网的提取、作物的精细分类、车牌和交通标志的识别等,都逐渐被深度学习占领,传统方法仅能保住深度学习无法涉及之处,即几何领域。这也是大势所趋。然而,连接主义令人不满之处在于,虽然理论上多层网络确实可能学习出最优的函数模型,但它无法解释该模型如何构建以及模型背后的含义,并导致传统、优雅的理论研究工作被缺乏激情、简单的“调参”所取代。这种缺失在未来会

将机器学习和人工智能带往何方,依然是一个需要长期思考的问题。

4 结 语

本文简要回顾了摄影测量的历史,并从几何和语义两个方面探讨了其与计算机视觉的联系与区别,以及未来的发展与展望。作为一门古老的应用型学科,摄影测量的许多研究内容,特别是几何部分已经成熟。然而,受到传感器技术、计算机技术和相关数学理论的推动,它也一直处于持续发展中。特别是平台技术,除了现有的航空、航天、无人机和车载平台,摄影测量也逐渐向着深空、水下和地下平台发展。此外,与计算机视觉、机器学习、人工智能等专业的进一步交叉融合,是摄影测量与遥感学科向着更深的理论基础、更广泛的应用前景和更实际的自动化解题能力前进的必然之路。

参 考 文 献

[1] Boyle W S, Smith G E. Charge Coupled Semiconductor Devices[J]. *Bell Syst. Tech*, 1970, 49 (4): 587-593
[2] Marr D. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information[M]. Cambridge, MA: MIT Press, 1982

- [3] Canny A. A Computational Approach to Edge Detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, 8(6): 679-698
- [4] Horn B. Robot Vision[M]. Cambridge, MA: MIT Press, 1986
- [5] Duda R, Hart P. Use of the Hough Transformation to Detect Lines and Curves in Pictures[J]. *Comm. ACM*, 1975, 15(1): 11-15
- [6] Marr D, Hildreth E. Theory of Edge Detection[J]. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 1980, 207 (1 167): 187-217
- [7] Forsyth D A, Ponce J. Computer Vision: A Modern Approach[M]. N J: Prentice Hall Professional Technical Reference, 2002:133-149
- [8] Thrun S, Fox D, Burgard W, et al. Robust Monte Carlo Localization for Mobile Robots[J]. *Artif. Intell.*, 2001, 128: 99-141
- [9] Marquardt D W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters[J]. *Journal of the Society for Industrial & Applied Mathematics*, 2006, 11(2):431-441
- [10] Hansen P C. Analysis of Discrete Ill-Posed Problems by Means of the L-Curve[J]. *Siam Review*, 2006, 34 (4):561-580
- [11] Wang Zhizhuo. Photogrammetry[M]. Wuhan: Wuhan University Press, 2007(王之卓. 摄影测量原理 [M]. 武汉:武汉大学出版社, 2007)
- [12] Hartley R, Zisserman A. Multiple View Geometry in Computer Vision[M]. Cambridge: Cambridge University Press, 2003
- [13] Cummins M, Newman P. FAB-MAP: Probabilistic Localization and Mapping in the Space of Appearance [J]. *International Journal of Robotics Research*, 2008, 27(6):647-665
- [14] Cortes C, Vapnik V. Support Vector Network[J]. *Machine Learning*, 1995, 20(3):273-297
- [15] Ho T K. Random Decision Forests[C]. International Conference on Document Analysis and Recognition, Montreal, Canada, 1995
- [16] Lafferty J D, McCallum A, Pereira F C N. Conditional Random Fields: Probabilistic Models for Segmenting And Labeling Sequence Data[C]. The Eighteenth International Conference on Machine Learning, Maryland, 2001
- [17] Krizhevsky A, Sutskever I, Hinton G E. Image Net Classification with Deep Convolutional Neural Networks [C]. International Conference on Neural Information Processing Systems, Nevada, 2012
- [18] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition [C]. *IEEE Transactions on Computer Vision and Pattern Recognition*, 2015(1): 770-778
- [19] Hinton G, Deng L, Yu D, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups[J]. *IEEE Transactions on Signal Processing Magazine*, 2012, 29(6):82-97
- [20] Chen Y, Lin Z, Zhao X, et al. Deep Learning-Based Classification of Hyperspectral Data[J]. *IEEE Transactions on Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2014, 7(6): 2 094-2 107

From Photogrammetry to Computer Vision

GONG Jianya^{1,2} JI Shunping¹

1 School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

2 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Abstract: We outline the history of photogrammetry from the aspects of, perspective geometry, camera, platform, measure methods and measure instruments, and summarize previous contributions to photogrammetry. A brief review of computer vision history is given. The tight connections between computer vision and photogrammetry are discussed in terms of geometric principles, and some differences in applications are also considered. From the aspect of semantics, we analyze the development of remote sensing and its relations to machine learning and computer vision, including their common approaches and different applications. The prevailing deep learning raised from connectionism is also reviewed and its successful applications in photogrammetry are analyzed. At last, we expect that the fu

(下转第 1615 页)

risk as optimization objectives. This solution is based on ant colony algorithm. Experimental results verify the effectiveness of the scheduling method.

Key words: emergency surveying and mapping; UAV scheduling; mission allocation; ant colony algorithm

First author: ZHU Qing, PhD, professor, Distinguished Professor of Changjiang Scholars Program of China, specializes in digital photogrammetry, multi-dimensional and dynamic GIS, virtual geographic environment. E-mail: zhuq66@263.net

Corresponding author: YU Jie, PhD. E-mail: yujie_gsis@whu.edu.cn

Foundation support: The National Natural Science Foundation of China, Nos. 41471320, 41571390,41501463; the Science and Technology Support Project of Sichuan Surveying and Mapping Geographic Information Bureau ,No. J2014ZC11.

.....
(上接第 1522 页)

ture development of photogrammetry will be more tightly cross-integrated with computer vision, machine learning and artificial intelligence.

Key words: photogrammetry; computer vision; remote sensing; machine learning

First author: GONG Jianya, PhD, professor, Academician of the Chinese Academy of Sciences, specializes in geo-informatics and photogrammetry. E-mail: gongjy@whu.edu.cn

Corresponding author: JI Shunping, PhD, professor. E-mail: jishunping@whu.edu.cn

Foundation support: The National Natural Science Foundation of China, Nos. 41471288, 61403285.