

# 数字摄影测量与计算机视觉

张祖勋<sup>1</sup>

(1 武汉大学遥感信息工程学院, 武汉市珞喻路 129 号, 430079)

**摘要:**从摄影测量的角度出发,在回顾摄影测量的发展历程及其最新进展的基础上,系统分析了数字摄影测量与计算机视觉的共同点及其本质差异,指出数字摄影测量的下一步发展必须与计算机视觉的理论和最新发展相结合。

**关键词:**数字摄影测量;计算机视觉;差异;发展

**中图法分类号:**P231.5

摄影测量是测绘学科的一个分支,它是对由摄影机摄取的影像(二维)进行量测,测定物体在三维空间的位置、形状、大小乃至物体的运动。摄影测量在近百年的历史中经历了模拟、解析与数字摄影测量三个阶段,其典型仪器如图 1 所示。模拟摄影测量仪器完全依赖于精密的光学-机械仪器,它常用两根金属导杆代替投影光线,将两个同名像点的投影光线“交会”得到它们的空间点位。20 世纪 30 年代,德国著名的摄影测量学者 Gruber 指出:“摄影测量是一门可以避免技术的技术”<sup>[1]</sup>。应该指出的是,摄影测量的模拟测图仪本身就是一台“模拟计算机”。



图 1 摄影测量三个发展阶段的三种典型仪器

Fig. 1 Three Typical Instruments in Accordance with the Development of Photogrammetry

事实上,摄影测量的发展史实质上是对计算机应用的历史。到解析摄影测量时代,计算机不仅应用于大规模的空中三角测量、区域网平差、数字测图等,还实时计算共线方程,代替模拟测图仪

的金属导杆(称为数字导杆),控制像片盘的实时运动,交会空间点位,构成如图 1(b)所示的解析测图仪。而进入计算机化的数字摄影测量时代后,代表性的数字摄影测量工作站(digital photogrammetry workstation, DPW)几乎无需光学、机械部件,DPW 就是一台计算机+各种功能的摄影测量软件。因此,美国曾将数字摄影测量称为 Softcopy Photogrammetry(软拷贝摄影测量)。

DPW 处理的原始数据就是数字影像,如图 2 所示,它利用数字图像处理、模式识别等技术,开始实现摄影测量的自动化。如利用“框标”模板对数字影像进行“模板识别”,实现摄影测量内定向自动化;利用影像兴趣算子(interesting operator)在一张影像上(如左影像)提取特征点,然后以它为模板,在另一张影像(如右影像)上确定同名点,实现摄影测量相对定向的自动化、空中三角测量的自动化、DEM(数字地面模型)生成的自动化、正射影像生成的自动化等。



图 2 数字摄影测量工作站

Fig. 2 Digital Photogrammetry Workstation

1992 年,国际摄影测量与遥感学会首次在华

盛顿大会上推出商用的 DPW,至今仅 12 年。1994 年,具有我国自主知识产权的基于 SGI 工作站的 DPW-VirtuoZo 首次在澳大利亚黄金海岸推出,至今也已经 10 年。特别是 21 世纪以来,我国数字摄影测量以世人难以想像的速度发展,DPW 在中国的摄影测量生产中获得了普遍的应用与推广。摄影测量的教学也由过去只有少数院校才能进行的“贵族”式的教学,得到了极大的普及,全国至少有 40 多所大专院校的测绘工程专业开设摄影测量课程,极大地拓展了摄影测量所需人才的培养渠道。

摄影测量进入数字摄影测量时代,它开始并将不断地利用计算机代替作业员的眼睛,毫无疑问,它已经与计算机视觉紧密联系在一起。美国俄亥俄大学的 Schenk 教授在其著作《数字摄影测量学》的序言中指出:“数字摄影测量是一门相对年轻的、并且迅速发展的学科。它的许多基本概念与方法来自影像处理与计算机视觉。但是不管它们对她(它)的影响有多强烈,数字摄影测量还是一门有自己特色的学科<sup>[2]</sup>。”

摄影测量与计算机视觉的关系是什么?究竟具有什么样的特色?本文将从以下两个方面进一步讨论数字摄影测量与计算机视觉之间的关系。

必须破除传统摄影测量来自“双眼视觉”的束缚,从计算机视觉的角度思考数字摄影测量的进一步发展;较详细地讨论了摄影测量与计算机视觉的差别,以说明摄影测量是一门有自己特色的学科,从而说明学科交叉,基于量测的计算机视觉——摄影测量的计算机视觉(photogrammetric computer vision)的必要性。

## 1 基于计算机视觉考虑数字摄影测量的进一步发展

就计算机视觉(特别是计算机立体视觉)的研究内容而言,它与摄影测量十分相近。计算机视觉的研究目标是使计算机具有通过二维图像认知三维环境信息的能力,这种能力将不仅使机器感知三维环境中物体的几何信息,包括其形状、位置、姿态、运动等,而且能对它们进行描述、存储、识别与理解<sup>[3]</sup>。事实上,计算机界也在利用立体视觉的原理进行从影像测绘地形图的研究。最近,人们利用立体视觉的原理,正在探索综合运用同一地区的多张不同角度拍摄所得的照片来恢复出地面的高度信息,以获得真正的三维地形图,并已取得了相当的进展<sup>[4]</sup>。在一个立体像对内进行

量测的过程中,两者也基本相似。一个完整的立体视觉系统通常可分为图像获取、摄像机定标、特征提取、立体匹配、深度确定及内插等 6 个部分<sup>[5]</sup>。

### 1.1 基于“广义点”的摄影测量

由于摄影测量渊源于测量学中测点的前方交会与后方交会,因此,共线方程(即物点、像点与摄影中心位于一条直线上)是整个摄影测量的核心。但是在建筑物的提取、建筑摄影测量、工业零件测量中,大量存在的是直线,由此,基于直线的摄影测量得到了深入的研究与应用<sup>[6]</sup>。但是在现实世界中,还存在大量的曲线,如地面上的道路、河流、湖泊等,建筑测量、工业测量中的圆、圆弧、曲线等。

目前,在摄影测量实际生产(特别是利用影像进行地图修测)与遥感图像的几何精纠正中,一般要在地形图与影像之间确定对应“点”作为控制点,这是一项比较困难的任务。因此,利用地图与影像上存在的大量的“线”作为控制信息进行配准,具有重大的理论与现实意义。

为此,笔者提出了“广义点”理论。传统摄影测量中的“点”指的是“物理”意义上(或称为“可视”)的点,而广义点则是“数学”意义上的点,因为任何一条“线”都是由“点”组成。由“广义点”理论,曲线(或直线)上任意一个点都可以被用为“控制点”,而且可以直接应用“共线方程”,但是只能在两个 $(x, y)$ 共线方程中选取一个。因此,很容易将点、直线、圆、圆弧、任意曲线归纳为一个数学模型——共线方程,进行统一平差。

### 1.2 多基线立体(multi-base stereo)

到目前为止,人都是由一条眼基线的“双眼”感受三维世界,而摄影测量(无论是模拟、解析,还是数字摄影测量)大多是沿用“由一条基线、两张影像构成的一个立体像对”进行测量。但是数字摄影测量利用计算机匹配替代人眼测定影像同名点时,由于存在大量的误匹配,即测量中所说的粗差,使自动匹配的结果很不可靠。为了提高自动匹配的可靠性,要求减小立体像对的“交会角”,但是由此产生的交会误差变大,精度变低。为了降低双目匹配的难度,计算机界很早就开始研究三目立体视觉系统<sup>[5]</sup>、三目机器人视觉系统<sup>[7]</sup>、多目立体匹配<sup>[8]</sup>。为解决周期性重复特征所引起的误匹配,计算机界提出了对倒距离的 SSD(sum of square difference)求和<sup>[9]</sup>。

多目视觉对于现在的航空数码相机显得尤为重要。当今的面阵数码相机多为“矩形”,特别是航线方向的相机视场角较小(Z/I 的 DMC 视场角

为 43°,奥地利 Vexcel 的 UltraCamD 视场角为 37°,按常规 60%航向重叠度构成的立体像对,交会角仅分别为 17.5°与 15°。而交会角太小,导致高程(计算机视觉中的深度)精度较低。但是 UltraCamD 航空相机的重叠度大,每个地面点至少出现在 5 张影像上 (each object in at least 5 images),即所谓多光线立体 (multi-ray-stereo),如图 3 所示。“多目立体匹配”可以利用摄影测量的空中三角测量原理,对多度重叠点进行“多方向的前方交会”,既能较有效地解决随机的误匹配问题,同时又能增加交会角,提高高程测量的精度。多基线立体对于数字近景摄影测量同样至关重要,它的重叠度可以完全由摄影者控制。多基线立体的出现将有力地解决双目立体匹配的歧义性,并完全有别于传统的摄影测量与传统的近景摄影测量。

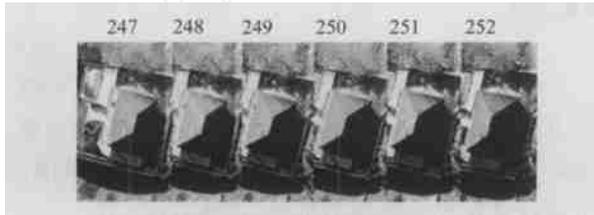


图 3 多光线立体

Fig. 3 Multi-Ray Stereo

### 1.3 单像建模(建筑物) ——灭点理论

灭点是空间一组平行线的无穷远点在影像上的构像(它不一定在影像的范围内),即该组空间平行线在影像上的交点,可以认为该空间的无穷远点与对应的灭点是一对“对应点”,它们满足共线方程。如航空影像,除“底点”(空间一组铅垂线的灭点)位于影像内(但它一般不是明显点),其他的“灭点”多数位于影像外。当航空影像接近水平时,水平方向的灭点几乎在“无穷远”处。因此,在模拟、解析摄影测量中,灭点不可能有任何实用意义,除了其定义外,对其理论也没有多少研究,即使在目前的 DPW 中也没有应用灭点。但在计算机图形学与数字摄影测量的下一步发展中,对空间平行线的自动分类、灭点的提取与应用是一个重要的研究方向,它不仅仅是利用单张非量测相机所获得的影像进行建筑物三维重建的基础<sup>[10, 11]</sup>,而且将成为利用城市大比例尺地形图进行空中三角测量的控制信息,并被用于城市大比例尺地形图的数据更新。

## 2 数字摄影测量与计算机视觉的差异

### 2.1 出发点不同导致基本参数物理意义的差异

摄影测量是测绘地形图的重要生产手段,它以空间(地面)坐标系为基准,如在一个地区进行航空摄影测量,所有摄影机的空间位置( $X_S, Y_S, Z_S$ )与影像的坐标系都相对于该空间坐标系。而计算机视觉是研究怎样用计算机模拟人的眼睛,实现机器人的视觉,它是以眼睛(摄影机)坐标系为准,它定义的平移量( $t_x, t_y, t_z$ )是空间坐标系相对于摄影机坐标系的平移量,二者的关系如图 4 所示,其数学表达式为  $t = RX_S$ 。可见,计算机视觉的平移量还与影像坐标系旋转矩阵  $R$  有关。

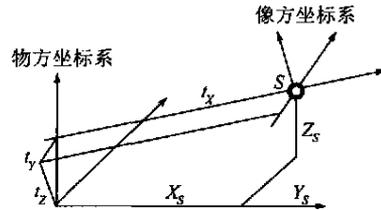


图 4 两个平移量的关系

Fig. 4 Relation Between Two Offsets

因此,在摄影测量中的“外定向”是确定影像在空间相对于物体的位置与方位;而计算机视觉通常从另一个方向描述这个问题:搜索物体相对于影像的位置与方位<sup>[12]</sup>。

### 2.2 出发点不同导致基本公式的差异

描述影像坐标( $x, y$ )与物方坐标( $X, Y, Z$ )的摄影测量基本关系式为:

$$\begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} X - X_S \\ Y - Y_S \\ Z - Z_S \end{bmatrix} \quad (1)$$

由此可得到摄影测量中最基本的共线方程:

$$\begin{aligned} x - x_0 &= \\ &- f \frac{a_1(X - X_S) + b_1(Y - Y_S) + c_1(Z - Z_S)}{a_3(X - X_S) + b_3(Y - Y_S) + c_3(Z - Z_S)} \\ y - y_0 &= \\ &- f \frac{a_2(X - X_S) + b_2(Y - Y_S) + c_2(Z - Z_S)}{a_3(X - X_S) + b_3(Y - Y_S) + c_3(Z - Z_S)} \end{aligned} \quad (2)$$

它将摄影机的内、外方位元素全部分开表达。而计算机视觉最基本的公式用齐次坐标的投影方程表达为:

$$Z_C \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & x_0 & 0 \\ 0 & f & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3)$$

$M_1 M_2 X = MX$

式中,  $M$  为  $3 \times 4$  阶的投影矩阵, 它将摄影机的内、外方位元素全部隐含在其中, 它与摄影测量的直接线性变换公式<sup>[13]</sup>一致:

$$x + \frac{L_1 X + L_2 Y + L_3 Z + L_4}{L_9 X + L_{10} Y + L_{11} Z + 1} = 0$$

$$y + \frac{L_5 X + L_6 Y + L_7 Z + L_8}{L_9 X + L_{10} Y + L_{11} Z + 1} = 0 \quad (4)$$

一般摄影测量多使用量测相机(即预先对相机标定), 因而测定每张影像的 6 个外方位元素是摄影测量的主要任务。最新发展的 POS 系统可以在航空摄影过程中直接测定 6 个外方位元素。而在计算机视觉的许多应用场合中, 求得  $M$  矩阵后, 不必再分解相机的内外方位元素, 直接将  $M$  矩阵作为摄像机参数, 但是它们没有具体的物理意义<sup>[3]</sup>。

### 2.3 处理的对象不同, 处理流程也不同

航空摄影测量需要对一个区域测绘地形图, 因此影像按航线的条带排列, 影像在航带内具有 60% 的重叠, 而在航线间具有 20% 的重叠, 构成区域如图 5 所示。

因此, 航带内相邻影像的相对定向(构成一个立体模型)、模型间的连接(构成一条航线的立体模型)、航线间的连接(构成区域)、空中三角测量、区域网平差, 从而确定每张影像的外方位元素, 是摄影测量最重要的流程之一。但是计算机视觉处理的范围一般较小, 通常为一个立体像对(立体视觉), 上述的很多处理方法(如空中三角测量和区域网平差)在计算机视觉中并不需要。

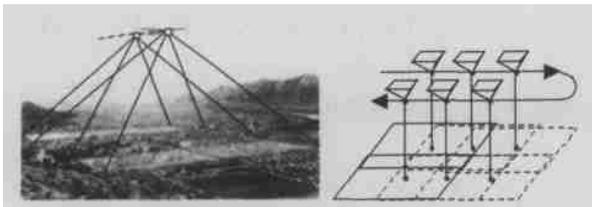
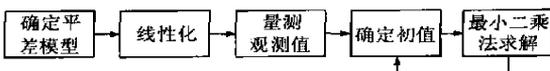


图 5 区域摄影测量

Fig. 5 Area Photogrammetry

### 2.4 算法的差异

由于摄影测量源于测绘学科, 测量数据处理方法——最小二乘法迭代求解几何贯穿于数字摄影测量的全过程为:



现以两者都关注的摄影机标定问题为例说明它们的差异。Hartley 提出, 保持摄影机中心不变、改变摄影机的角度对任意空间物体摄取多张影像, 进行相机标定<sup>[14]</sup>。由于相机不动, 移位置  $t = 0$ , 设局部空间坐标系原点位于摄影中心。因此, 空间点  $x = (x, y, z, 1)^T$  摄影于影像  $j$  上的像点坐标为  $u_j = K(R_j(0)(x, y, z, 1)^T) = KR_j(x, y, z)^T$ , 其中,  $K$  为由影像内方位元素所构成的  $3 \times 3$  矩阵;  $R_j$  为由 3 个旋转角构成的  $3 \times 3$  正交矩阵。

若以影像 0 为基准, 即它的旋角为 0, 则  $u_0 = K(x, y, z)^T$ , 由此可得  $u_j = KR_j x = KR_j K^{-1} u_0$ 。因此, 由影像 0 到  $j$  的投影为:

$$u_j = P_j u_0 \quad (5)$$

根据影像与一组同名点就可获得投影矩阵  $P_j$ , 由  $P_j = KR_j K^{-1}$  可得正交矩阵  $R_j = K^{-1} P_j K$ 。由于  $R = R^{-T}$ , 有  $(KK^T) P_j^{-T} = P_j (KK^T)$ , 令  $C = KK^T$ , 得:

$$CP_j^{-T} = P_j C \quad (6)$$

由  $n$  张影像建立上述线性方程组, 共  $9n$  个线性方程组, 用最小二乘法解(线性系统), 求得对称矩阵  $C$ , 再用 Choleski 分解求得相机参数矩阵  $K$ 。这是计算机视觉典型的推导、计算方法。

摄影测量的推导、计算方法则不同。对于两个平面(影像  $j$  与影像 0), 存在平面的直接线性变换:

$$x_j = \frac{m_1 x + m_2 y + m_3}{m_7 x + m_8 y + m_9}$$

$$y_j = \frac{m_4 x + m_5 y + m_6}{m_7 x + m_8 y + m_9} \quad (7)$$

直接线性变换的 9(实际是 8)个参数, 它实际是上述的投影矩阵  $P$ 。摄影测量的计算过程为:

6 首先由同名点的最小二乘法求得参数  $m_1$ 、 $m_2$ 、 $m_3$ 、...

6 再由下式求得  $x$ 、 $y$  方向的灭点:

$$x_{x-} = \frac{m_1}{m_7}, y_{x-} = \frac{m_4}{m_7}$$

$$x_{y-} = \frac{m_1}{m_7}, y_{y-} = \frac{m_4}{m_7} \quad (8)$$

6 假定相机主点(初值)位于数字影像中心, 其坐标为  $x_0$ 、 $y_0$ , 从而获得灭点相对于主点的坐标  $x_x = x_{x-} - x_0$ ,  $y_x = y_{x-} - y_0$ , 进而可得相机焦距、倾角的初值<sup>[10]</sup>;

6 最后利用自检校空中三角测量, 利用上面获得的初值对  $n$  张影像进行区域网平差, 解得相机的内方位元素。

由上可以看出,对同一问题两者解决的方法是不同的。特别是在数字摄影测量中引入最小二乘匹配<sup>[15]</sup>之后,它使最小二乘法同样也贯穿于数字摄影测量的全过程。

计算机立体视觉与数字摄影测量的根本问题之一是利用计算机代替人眼识别影像同名点。最小二乘匹配是将影像的灰度  $g$  用一个函数表示  $g = f(x, y)$ ,如图 6 所示。设左影像的灰度为  $g_1$ ,右影像的灰度为  $g_2$ ,判断同名点的最简单的准则就是同名点窗口内所有点的灰度应该相等,或灰度差的绝对值和为最小:

$$|g_1(x, y) - g_2(x + i, y + j)| = \min \quad (9)$$

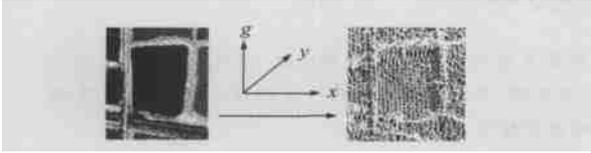


图 6 图像的三维表示  $g = f(x, y)$

Fig. 6 3D Representation of Image

在此以机械零件(图 7 所示)的方孔测量说明最小二乘法已经贯穿了数字摄影测量的全过程。直线提取需要最小二乘法匹配;圆弧提取需要最小二乘法匹配;最后需要引入“几何约束条件”,利用带条件的间接观测平差按最小二乘法获得最优解,它既能获得稳定的解算结果,又能达到高精度。

除研究怎样获得高精度的观测值外,怎样确定平差的物理模型,怎样获得“初值”,在平差过程中怎样剔除观测值中的粗差,都是数字摄影测量重要的研究内容。

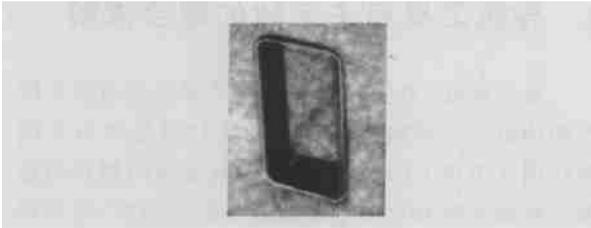


图 7 方孔的测量

Fig. 7 Measurement of Round Rectangle

数字摄影测量所解算的多数是非线性系统,解算过程中都是将它线性化,进行迭代求解。这与计算机视觉强调矩阵分解、线性方程求解有很大差异,计算机视觉总是设法将非线性问题转化为线性问题,尽可能避免求解非线性方程。

### 3 结 语

数字摄影测量事实上已经进入计算机视觉的领域,数字摄影测量的进一步发展必须破除传统摄影测量的束缚,从计算机视觉的观点出发。同时,尽管数字摄影测量与计算机视觉有差异,但是随着量测(特别是高精度量测)型的计算机视觉的需求以及数字近景摄影测量的发展,两者的学科交叉必将愈来愈重要,似乎这种交叉将会在计算机视觉中形成一个新的分支——摄影测量的计算机视觉,但是它不应仅仅局限于地学信息(Geo-information)<sup>[16]</sup>。

### 参 考 文 献

- 1 Wang Z. From Photogrammetry to Geomatics—a Commemoration of Accomplishment That is VirtuZo. VirtuZo Manuscript, Brisbane, Australia, 1995
- 2 Schenk T. Digital Photogrammetry. USA: TerraScience, 1999
- 3 马颂德,张正友. 计算机视觉——计算理论与算法基础. 北京:科学出版社,1998
- 4 田 峰. 模式识别技术及其应用. 中国科学报,1995-09-27
- 5 游素亚,徐光耀. 立体视觉研究的现状与进展. 中国图像图形学报,1997,12(1)
- 6 Heuvel F A. A Line-photogrammetric Mathematical Model for the Reconstruction of Polyhedral Objects. SPIE, 1999, 3641:60~70
- 7 Ayache N L. Trinocular Stereo Vision for Robotics. IEEE Trans. on PAMI, 1991, 13(1): 73~85
- 8 章毓晋. 图像工程——图像理解与计算机视觉. 北京:清华大学出版社,2000
- 9 Okutomi M, Kanade T. A Multiple-baseline Stereo. IEEE PAMI-15, 1993. 353~363
- 10 张祖勋,张剑清. 单影像房屋三维重建. 2001年两岸自动化数字工程测量研究会,台北,2001
- 11 Cipolla R, Robertson D, Boyer E. Photobuilder-3D Models of Architectural Scenes from Uncalibrated Images. IEEE International Conference on Multimedia Computing and Systems, Firenze, 1999
- 12 Heuvel F A. Exterior Orientation Using Coplanar Parallel Lines. The 10th Scandinavian Conference on Image Analysis, Lappeenranta, 1997
- 13 王之卓. 摄影测量原理. 北京:测绘出版社,1996. 321~324

(下转第 1105 页)

**Abstract:** Three-dimensional datum transformation is widely used in geodesy, photogrammetry, engineering survey and other subjects. The Bursa model, molodensky model etc. are usually used. General speaking, they are mainly suitable in small angle's three-dimensional datum transformation. But in actual, a lot of large angles' three dimensional datum transformation need to be performed. This paper puts forward a simplified model of three-dimensional datum transformation that can suit any angle.

**Key words:** orthogonal matrix; three-dimensional datum transformation; big rotation angle

**About the first author:** CHEN Yi, associate professor, concentrated on the research in space geodesy and satellite geodesy.

E-mail: chenyi@mail.tongji.edu.cn

(责任编辑: 宏远)

(上接第 1039 页)

14 Hartley R I. Self-calibration from Multiple Views with a Rotating Camera. The 3rd European Conference on Computer Vision, Stockholm, 1994

Photogrammetric Computer Vision'. Highlights of ISPRS, 2001,6(2)

15 Ackermann F. Digital Image Correlation: Performance and Potential Application in Photogrammetry. The Photogrammetric Record, 1984, 11 (64)

作者简介:张祖勋,教授,博士生导师,中国工程院院士。现主要从事数字摄影测量与计算机视觉等领域的研究。

E-mail: zxzhang@supresoft.com.cn

16 Leberl F W. ISPRS Commission III ' Towards

## Digital Photogrammetry and Computer Vision

ZHANG Zuxun<sup>1</sup>

(1 School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

**Abstract:** This paper gives an overview of the history and recent developments of photogrammetry. Then the common ground and differences between digital photogrammetry and computer vision are analyzed systematically. It is evident that the advantages of computer vision must be integrated into the further progress of digital photogrammetry.

**Key words:** digital photogrammetry; computer vision; difference; development

**About the author:** ZHANG Zuxun, professor, Ph. D supervisor, academician of Chinese Academy of Engineering. He is mainly engaged in the research on digital photogrammetry and computer vision.

E-mail: zxzhang@supresoft.com.cn

(责任编辑: 宏光)