doi:10. 3969/j.issn. 1672-4623.2023.02.026

1∶10 000 比例尺多源数据的修测更新入库

甘迎娟[1]
（1. 中煤航测遥感集团，陕西 西安 710199）

摘　要：对已有1∶10 000比例尺多源数据（格式为GDB、DEM、DGN）分析，通过编写FME模板完成原始数据库（GDB）-修测矢量（DLG）-更新成果数据库（MDB）三者之间的无缝转换，以及成果数据库MDB的质量检查，实现多源数据融合修测更新入库。整个实现过程高效率、高质量，具有良好的经济效益，解决了常规数据库修测更新过程中，必须配备专门的数据平台、操作交互性差及效率低的问题。

关键词：多源数据；修测更新；FME；无缝转换；质量检查
中图分类号：P284　　**文献标志码：**B　　**文章编号：**1672-4623（2023）02-0125-05

1∶10 000 Scale Multi-source Data Revision and Update into the Database

GAN Yingjuan[1]
(1. Aerial Photogrammetry and Remote Sensing Group Co., Ltd., Xi'an Shaanxi,710199, China)

Abstract: After analyzing the 1 ∶ 10 000 scale multi-source data(the format is GDB, DEM, DGN), through writing FME template, we completed the seamless transformation among the original database GDB, revision vector DLG and update database MDB, and the quality check of MDB, which realized multi-source data revision, update into the database. The whole process is of high efficiency, high quality and good economic benefit, which solves the problems of special data platform, poor operation interaction and low efficiency in the process of routine database revision and update.

Key words: multi-source data, revision and update, FME, seamless conversion, quality inspection

FME是加拿大safe Software公司推出的用于不同数据格式之间转换的一款系统软件。该系统基于OpenGIS组织提出的新的数据转换理念“语义转换”，提供转换过程中重新构造数据功能，实现了超过500多种数据格式之间的转换，可为规划技术服务提供快速处理和实现的解决方案。FME的优点是向用户提供了比同类GIS软件更加强大的操作灵活性与自由度，为用户的技术处理思路提供了无限可能[1]。

1　数据准备

以原始数据库（GDB）-修测矢量（DLG）-更新成果数据库（MDB）的生产流程为主线，叙述了FME在数据准备、数据入库及质量检查过程中的应用。通过运用上百个/次Transformers函数，在不借助第三方软件平台的基础上实现上述流程并取得了较高的成果质量和生产效率[2]。多源数据修测更新入库技术流程如图1所示。

1.1　数据源简介

多源性一方面表现为数据获取手段的多样性，另一方面表现为不同采集方法和管理系统中数据具有的不同存储、交换格式[3]。

多源数据的特点是多元异构，多元指数据来源多样化。多源数据融合是指将各种不同的数据信息进行综合，吸取不同数据源的特点，然后从中提取出统一的，比单一数据更好、更丰富的信息。项目已有多源数据（格式为GDB、DEM、DGN），其中DLG来源于最新时相的航空影像采集，GDB来源于地理国情普查成果，DEM来源于农村承包经营权确权成果。异构指数据的类型、结构、特征等的不一致性。项目中，DLG的格式为*.DGN，是一种矢量文件；GDB的格式为*.GDB，是一种空间数据库集；DEM的格式为*.GRID，是一种数字高程模型的存储方式。

使用航空摄影测量方法，借助FME数据重构功能，全面修测更新已有的基础地形地貌数据库。同时，扩充一些必要的属性数据，建立空间与非空间数据之间的联系，为各行业提供更准确、全面的数据服务。

1.2　数据源预处理

1.2.1　DEM提取曲线

农村承包经营权确权DEM成果在项目中有2个应

收稿日期：2021-06-19。

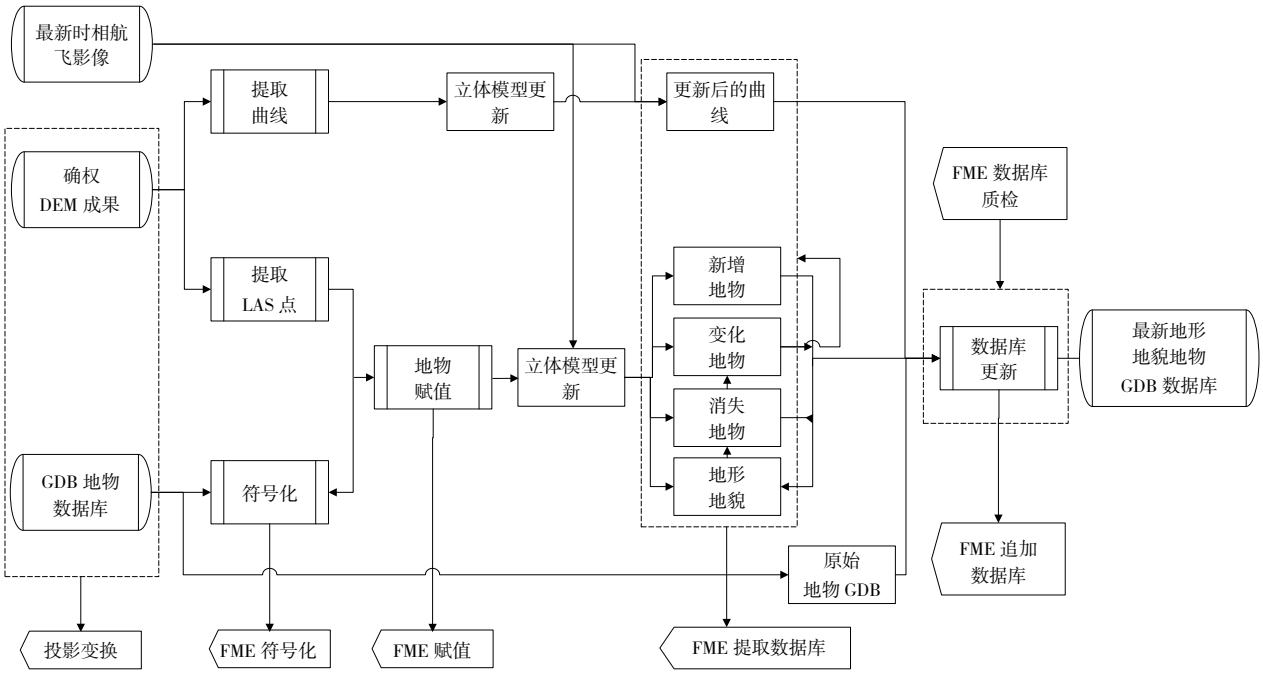


图1 多源数据修测更新入库技术流程

用，DEM生成LAS点文件用于地理国情普查成果GDB数据库的高程赋值。DEM同时生成等高线用于立体模型修测。

1.2.2 GDB符号化

通过FME的模板中的“Joiner”函数，以Excel表中的GB对应项为识别信息码，挂接FH.xlsx属性映射对照表，属性映射对照表由GB列、符号名列、线型名列组成，并在转换过程中挂接seed.dgn的种子文件，通过转换将数据库中的点状数据符号化，线状数据赋于线型信息。完成从GDB到DLG的数据格式转换，GDB数据符号化示意如图2所示。

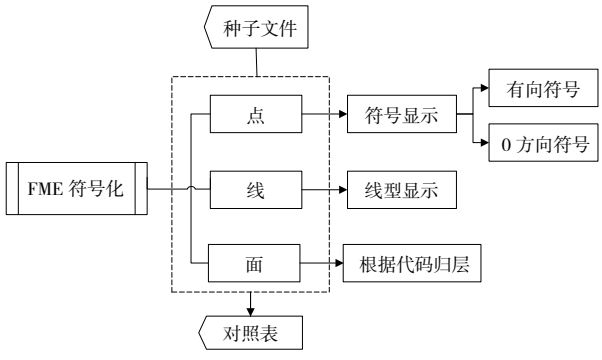


图2 GDB数据符号化示意图

1.2.3 GDB赋Z值

通过符号化作为修测的GDB文件，必须要赋三维值后才能在立体模型中正确显示和使用。在格式转换过程中，同时进行GDB与DME生成的LAS点叠加分析，将LAS点的三维值（Z值）赋给GDB文件。赋值过程中，使用“SurfaceDraper”，“CoordinateExtractor”，“FeatureMerger”，“3DForcer”等关键函数，最终输出3D的矢量DLG成果。符号化成果赋Z值示意如图3所示。

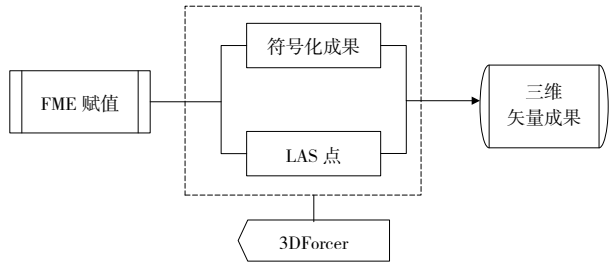


图3 符号化成果赋Z值示意图

1.2.4 立体矢量修测

3D的矢量DLG成果在最新时相的航空影像立体模型下，根据相应的UCM表进行修测更新，修测内容包括变化的、新增的、删除的[4-5]。

2 数据入库

2.1 数据库的原则

统一性：要素内容、表达方式、采集指标、数据模型、组织结构等内容统一，保证数据生产与建库过程中数据成果的准确性和一致性。

兼容性：满足该省1∶10 000DLG数据生产、建库和应用的需要，统筹兼顾数据库快速出图的要求，以及其他行业的需求。

可扩展性：预留数据层、要素代码和要素属性的扩充空间，要素代码和属性项应与规范相衔接。

2.2　数据分层要求

数据分层的命名采用5个字符，第一个字符为数据比例尺代码（G为1∶10 000，H为1∶5 000），第二个至第四个字符为数据内容的缩写，第五个字符代表几何类型（A:面，L:线，P:点）数据分层见表1所示。

表1　1∶10 000DLG数据分层表

要素分类	数据内容	数据分层	数据体几何特征
水系（H）	水系要素	GHYDA	面
		GHYDL	线
		GHYDP	点
	水系附属设施	GHFCA	面
		GHFCL	线
		GHFCP	点
居民地及设施（R）	居民地	GRESA	面
		GRESL	线
		GRESP	点
…	…	…	…

2.3　数据入库的思想

使用FME Workbench入库非常便捷，通过前期的数据分析、修测整理、编写映射文件及模板，实现使用FME完成一键入库。流程简洁、操作简便、交互性好。FME修测入库流程如图4所示。

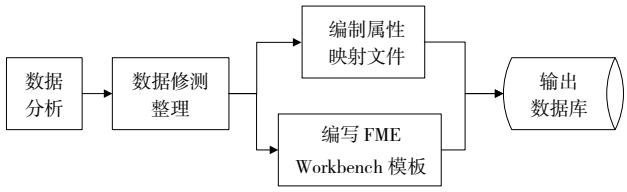


图4　FME修测入库流程图

属性映射文件内容包含修测矢量文件的层、色、粗度及属性库的层、GB码等信息，通过GB码的唯一性提取数据库并填写相应的属性信息。属性映射文件见表2所示。

2.4　Workbench模板编写

数据入库的基本思路为在Bently公司的microstationv8软件下，进行矢量数据修测并整理为以点、线、面为表达格式的矢量文件。在FME软件下，通过编写数据转换模板，运用“joiner”，“ExpressionEvaluator”，“AttributeCreator”，“CoordinateExtractor”，等69个/次Transformers函数完成数据库的提取。

表2　属性映射文件

ID	Name	Color	WT	Layer	描述	GB
30	Level 39	4	0	GHFCL	单坡堤	7602001
31	Level 39	0	0	GLFCL	路堤路堑	7602000
32	Level 54	11	0	GLFCL	沟堑	760200
1	Level 49	1	2	GTERL	计曲线	710102
2	Level 49	0	0	GTERL	首曲线	710101
36	Level 50	5	2	GTERL	草绘计曲线	710202
37	Level 50	4	0	GTERL	草绘首曲线	710201
5	Level 51	4	0	GTERP	高程点	720100
13	Level 53	43	0	GTFCA	露岩地	750702
…	…	…	…	…	…	…

数据库最终成果需添加投影信息。一般情况是在ArcGIS使用ArcTOOLBOX下的工具就能实现，但对于大数据量来说，FME的批量操作更高效。FME中Esri-Reprojector转换器的思想为使用ESRI重投影库，把要素坐标从一个坐标系统重投影到另一个坐标系统，使用Batch Deploy命令批量完成投影变换。

3　质量检查

3.1　质量要求

3.1.1　数据处理原则

规范性（要素分类、数据分层、属性项定义、要素属性等需与国标一致），一致性（DLG数据编码标准、拓扑关系、格式概念须一致），完整性（数据层、覆盖区域以及内部文件必须满足完整性要求），正确性（数据在满足图形拓扑关系的同时，各属性项顺序、名称、长度、类型、分类、属性等相关内容必须正确）。

3.1.2　图形要求

所有面、线要素中均不允许存在圆弧、B样条曲线和Multiline的线串，源数据中如果存在该类型的线串，则需处理为与源形状基本一致的多段线类型。

3.1.3　拓扑要求

各类点、线、面状要素都不能有自重叠。点、

线、面要素之间也不能有不合理的相互重叠。线、面要素不能存在自相交、折返线。构面要素与原始构面边线应保持线面的节点套合，最大误差不超过0.01 m。

3.1.4 属性要求

必填属性不得为空，要素图形表示与“分类代码”属性保持一致；图形的名称、类型等属性与图面注记一致。

对矢量数据进行质量检查是获取精确性较高的矢量数据的必要手段[6]。根据上述质量检查内容及质量要求结合项目本身的特点，在数据质量检查设计过程中，以要素类型为研究对象，从点、线、面3个方面，重点关注以下几个方面的内容。数据质检内容见图5所示。

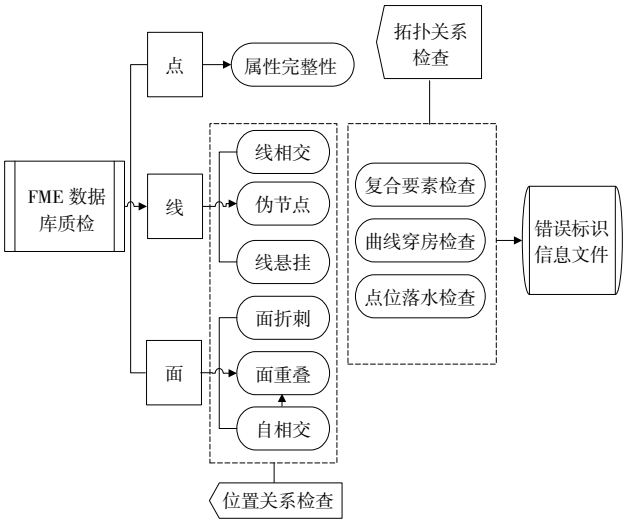


图5 数据质检内容

3.2 质量检查实现

DLG数据转换为MDB格式后，需要对数据库进行面折刺检查，空间关系合理性检查，复合要素检查等及投影变换等。质量检查可以通过ArcGIS实现，也可通过FME编写模板实现，使用FME编写检查模板有3个方面的优点：第一，交互性好且检查效率高，输出结果仅需数秒且可直接导入矢量数据修改。第二，可进行一对一的定制，对于拓扑关系检查，在模板中写定拓扑规则，操作时无需选择层及规则。第三，对于有的错误项检查过程即修改过程，运行完检查模板后自动修复错误项。加快制图更新速度、推进生产单位整体工作效率，不断提高数据成果质量[7]。

3.2.1 位置关系检查

面折刺在ArcGIS检查的思想为通过ArcTOOLBOX完成要素折点转点和要素转线，然后在Arccatalog下创建属性字段，通过python编写脚本计算出各个折次点的角度。根据属性表中的角度字段筛选出最小锐角进行排查折次。实现过程非常繁琐，且无法自动修改。

FME下“SpikeRemover”转换器的思想为每一对由3个连续不同的点组成的线段。如果两线段角（度）小于或等于指定的最大角度，那么中间点（为尖锐）被删除。函数原理如图6所示。使用FME检查模板仅数秒就完成了面折刺的检查和修改。

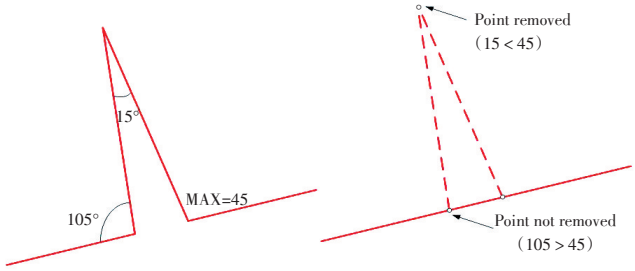


图6 SpikeRemover函数原理

自相交使用“SelfIntersector”，“Intersector”，“Tester”等函数筛选。自相交检查模板如图7所示。

伪节点使用“chopper”，“PointOnPointOverlayer”，“DuplicateRemover”，“SpatialFilter”，“Tester”等函数筛选。伪节点模板如图8所示。

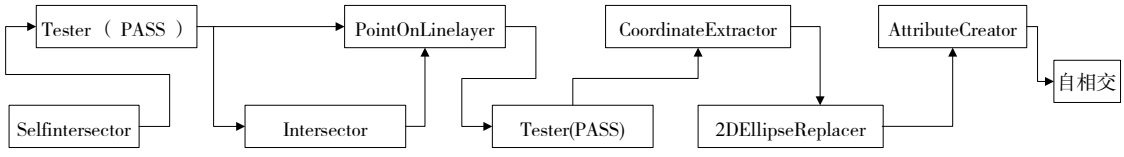


图7 自相交检查模板

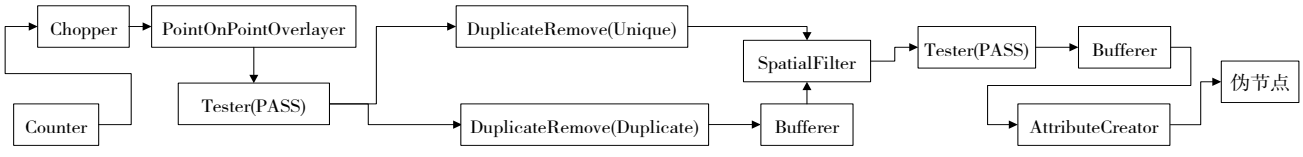


图8 伪节点检查模板

其余描述线悬挂、线相交、小面检查、面重叠检查等位置关系错误模板不再一一叙述。

3.2.2　拓扑关系检查

拓扑关系检查功能通过使用ArcGIS软件中与拓扑处理相关的组件实现，过程为设置容差、添加投影信息、添加拓扑规则等一系列。该检查方式操作繁琐，且只能检查SHP、MDB等库体数据，交互性和操作性差，严重影响作业效率。

以曲线穿过房屋、比高点及高程点落水2种情况描述FME中实现拓扑关系检查的过程：使用“LineOnAreaOverlayer”,“PointOnAreaOverlayer”,“Tester”3个转换器，通过点、线、面不同类型GB码的空间分析叠加，以及叠加结果的关系判断完成曲线是否穿过房屋面检查和比高点及高程点是否落入至水系面检查，并根据测试的结果以DWG格式的矢量文件输出错误要素。同样，模板运行时间仅数秒就完成该项拓扑关系的检查。

复合要素检查。ArcGIS下处理复合要素非常容易，但是检查出复合要素需要借助python编写脚本才能实现，这对没有编程基础的测绘工作者实现起来有些难度。

FME使用“AggregateFilter”转换器，该转换器的思想为根据聚合体的几何图形的类型、或者根据输出单个几何对象，过滤出不同的要素。将原始数据库输入至FME检查模板中，使用该转换函数筛选出非法线聚合体和非法面聚合体，同样数秒即可输出检查结果。

3.2.3　属性完整性检查

点状符号数据完整性检查主要是在数据库的提取过程中，对必填项进行强制赋值。以高程点为例，高程点属性结构如表3所示。

表3　高程点属性结构表

数据分层	属性项	中文说明	字段类型	长度	必填	填写示例
GTERP	GB	国标分类码	TEXT	6	YES	720300
	ELEV	高程值	DOUBLE	–	YES	153.7
	DATE	现势性	TEXT	8	NO	201106
	STACOD	更新状态标识	TEXT	4	NO	增加/删除/修改
	FEAID	数据库标识	LONG	–	NO	23422
	VERS	版本标识	TEXT	10	NO	2012

上表中的必填项GB、ELEV，其中GB为唯一标识码。在数据库提取中通过函数“AttributeRenamer”重命名获得高程值属性字段。现势性通过函数“AttributeCreator”派生创建更新状态、数据库、版本标识属性字段。

4　结　语

使用FME模板极大地提高了数据处理效率和数据质量，节约了人力资源，是多源数据库修测更新的强大技术支撑。同时，项目所取得的数据成果现势性好、质量优良，对于满足国家经济建设和社会发展的需要，提高测绘地理信息的服务保障能力具有十分重要的作用和价值[8]。

参考文献

[1]　陈琦，黄伟涛. 基于FME的城市更新基础数据自动化入库[J]. 地理空间信息，2019，17(7):57-59

[2]　唐辉，韩维喆，蒲德祥，等. FME在大比例尺基础测绘数据建库中的应用[J]. 地理空间信息，2020，18(11):91-93

[3]　高光军，孙灵杰，李自力，等. 基于“多规多源”的不动产权籍数据库存储及更新机制研究[J]. 现代测绘，2017，40(2):10-12

[4]　李路英. 1∶500基础地理信息数据库更新方法探讨[J]. 测绘通报，2017(5):132-135

[5]　郭兰勤，吴燕平，王海恒，等. 基于增量式更新的1∶50 000地形数据库融合方法探讨[J]. 测绘标准化，2019，35(3):36-38

[6]　张标，陈楠，高瑛，等. 矢量数据质检工具的设计与实现[J]. 北京测绘，2020，34(10):1 447-1 450

[7]　黄波，谢三五，王辉，等. 基于图库一体化的制图更新技术探究[J]. 地理空间信息，2020，18(2):45-49

[8]　陈军，王东华，商瑶玲，等. 国家1∶50 000基础地理信息数据库更新工程及实施[J]. 地理信息世界，2012(1):8-12

作者简介：甘迎娟（1982—），硕士，从事航空摄影测量技术质量工作。