

Integration and Updating of Multi-source Heterogeneous Geospatial Data: A Case Study of “Tianditu·Shanghai”

Qiyaoyao Liu

Shanghai Fandu Geographic Information Technology Co., Ltd., Shanghai, 200333, China

Abstract

As the National Geographic Information Public Service Platform (Tianditu) transitions to a distributed service architecture, efficiently integrating multi-source and heterogeneous data from national, provincial, and municipal nodes has become a critical challenge. Taking the annual integration and update practice of “Tianditu·Shanghai” as the research subject, this paper systematically discusses the technical challenges and solutions encountered in fusing geospatial data across multiple scales, time periods, and standards. First, it analyzes the heterogeneity between national main node and Shanghai municipal node data in terms of geometric accuracy, attribute structure, timeliness, and mathematical foundations. Then, it proposes a three-tier fusion principle prioritizing timeliness, high precision, and consistency, establishing a complete technical workflow encompassing coordinate system conversion, feature layer extraction, processing for geometric and attribute consistency, and quality control. The paper focuses on addressing key technical issues such as road network integration in complex urban environments, intelligent deduplication of place names, addresses, and points of interest (POIs), and color balancing of multi-source imagery. Finally, quantitative evaluation verifies that the fused data achieve significant improvements in completeness, consistency, timeliness, and accuracy. The paper also outlines future directions, including intelligent fusion based on geographic knowledge graphs and crowdsourced collaborative updates.

Keywords

Tianditu; multi-source data fusion; spatial data updating; quality control; geographic information public service

多源异构地理空间数据的融合与更新实践——以“天地图·上海”为例

刘其瑶

上海凡度地理信息技术有限公司，中国·上海 200333

摘要

随着国家地理信息公共服务平台（天地图）向分布式服务架构转型升级，如何高效整合国家级与省市级节点的多源异构数据成为关键挑战。本文以“天地图·上海”的年度融合更新实践为研究对象，系统阐述了在多尺度、多时相、多标准的地理空间数据融合过程中面临的技术难题与解决方案。文章首先对国家级主节点以及上海市级节点数据在几何精度、属性结构、现势性还有数学基础等方面所存在的异构特性进行了分析；进而提出了“现势性优先、高精度优先、统一性优先”的三级融合原则，构建了从坐标系转换、要素分层提取、几何与属性一致性的处理工作到质量控制的完整技术流程。重点探讨复杂城市环境之下道路网络融合、地名地址与兴趣点即POI智能去重以及多源影像色彩均衡等关键技术难题的应对策略。最后，量化评估工作验证了融合后的数据在完整性、一致性、现势性以及精度这些方面得以显著提升，同时还对基于地理知识图谱的智能融合、众源协同更新等未来发展方向进行了展望。

关键词

天地图；多源数据融合；空间数据更新；质量控制；地理信息公共服务

1 引言

“天地图·上海”自2015年启动数据融合工作以来，已形成年度更新机制。在实际工作中，面临着数据源异构性突出、城市要素变化快速、融合质量控制复杂等多重挑战。传统的周期性、半人工融合模式已难以满足高现势性服务

需求。

本文以上海市为例，旨在系统总结国家级与省级节点数据融合的技术框架与实践经验。首先分析多源数据的特征与融合难点；然后详细阐述融合原则、技术流程与关键方法；接着通过量化指标评估融合效果；最后探讨未来智能化融合的发展方向。研究期望为其他地区乃至国家层面的地理信息数据融合工作提供可借鉴的技术路线和解决方案。

【作者简介】刘其瑶（1997-），女，中国河南南阳人，本科，助理工程师，从事地理信息研究。

2 研究区概况与多源数据特征分析

2.1 研究区概况

上海作为国际化大都市，期城市空间结构是比较复杂，且地理要素更新快。截止到2024年底，全市道路的总里程超过了2.1万公里，其中城市快速路以及立交桥等复杂交通设施的占比要高于全国平均水平。同时，随着城市更新以及基础设施建设的加速开展，年均新增的建筑面积超过了3000万平方米，地理空间数据的变化频率以及复杂度给数据融合工作提出了较高的要求。

2.2 多源数据特征与差异分析

“天地图·上海”数据融合涉及到了三类数据源。

2.2.1 国家级主节点数据特征

国家主节点的数据是由国家基础地理信息中心来进行统一制作以及分发的，拥有标准化的数据模型以及严格的质量控制，其优势在于：

1. 数学基础统一：运用CGCS2000国家大地坐标系，来保障全国空间的一致性。
2. 要素表达规范：严格按照《地理信息公共服务平台电子地图数据规范》来执行，其属性结构完整。
3. 质量稳定且可靠：经过国家级的质检，拓扑错误率低于0.1%。

2.2.2 上海市级节点数据特征

上海市级数据来源于本地基础地理信息数据库，具有显著的本地化优势：

1. 现势性突出：依托季度更新机制，重要要素更新滞后不超过3个月。
2. 几何精度高：基于上海市连续运行参考站系统(CORS)，平面精度优于2米。
3. 要素内容丰富：包含了大量地方特色要素，如里弄建筑、历史保护街区等。

2.2.3 影像数据特征

融合采用的影像数据包括国家主节点2米分辨率影像和上海市0.5米高分辨率影像。主要矛盾体现在：

1. 时相不一致：国家影像现势性为2023年，而上海影像为2025年。
2. 色彩差异显著：不同传感器、不同时相导致色调、对比度存在明显差异。
3. 几何配准误差：即使在正射纠正后，同名地物在两种影像上仍可能存在1-3个像素的位置偏差。

3 数据融合总体框架与原则

天地图数据融合核心目标：提供多源数据优势互补的特性，生成现势性、精度、完整性、一致性方面优于单一数据源，并且标准统一的服务数据集。这个框架包含了4个层次：

1. 数据预处理层：坐标系转换、把格式进行统一并且进行质量初检。

2. 核心融合层：对于要素类型开展差异化融合策略。

3. 质量控制层：贯穿始终的质量检查与修正

4. 成果输出层：开展标准化服务数据集的生成工作以及切分工作

4 多源数据融合关键技术流程

4.1 矢量数据融合流程与关键技术

4.1.1 坐标系转换与几何配准

将上海市节点数据从上海2000坐标系转换至CGCS2000国家大地坐标系，运用七参数布尔莎模型：

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + (1+k)R(\epsilon_x, \epsilon_y, \epsilon_z) \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} \quad (1)$$

当中的转换参数是由上海市测绘院提供的，平面转换的精度要优于0.05米，能够契合1:5000比例尺的制图要求。

几何配准运用基于特征点的二次多项式校正方法，在城乡结合部等变化显著的区域中，布设156个均匀分布的检查点，从而保证配准后的中误差能够控制在1.5像素以内。

4.1.2 道路网络融合的难点与解决方案

上海城市道路网络所具有的复杂性给融合工作带来了特殊挑战，具体表现为：

1. 立体交通冲突：高架和地面道路的二维投影出现了重叠情况。
2. 立交系统比较复杂：大型互通式立交拥有多层次以及多方向的连接关系。
3. 内部道路模糊：居住区以及园区当中的道路和绿地、建筑边界难以进行区分。

解决方案：

1. 开展立体要素的分离：创建“高程属性”字段，将高架、地面、地下道路分层存储，在可视化时通过符号化来进行区分，融合规则：在进行立体交叉时，把上海市级数据的高程信息当作标准来使用，也就是其经过激光雷达扫描得到精确高程的高程信息。

2. 立交系统的规则化表达：针对苜蓿叶式以及涡轮式这类典型立交，制定标准化的表达模板，从而保证道路连接关系的层级是正确的，如图1所示，

3. 立交系统规则化表：内部道路处理的策略为：基于“继承优先”的原则，保留国家主节点所表达的居住区内部道路；添加新的区域时把上海数据当作参考依据，需要结合0.5米影像来进行人工核实真实性。

4.1.3 水系数据融合策略

针对国家主节点水系存在的简化严重这一问题，采用“几何取上海、属性补国家”的融合策略：

1. 几何融合：运用上海市节点高精度水系数据，含大量支流和小型水体的数据。
2. 属性融合：把国家主节点标准化属性，像河流编码、

管理等级这些，匹配到上海水系要素当中。

3. 拓扑检查：要确保水系以及桥梁、水闸等构筑物的空间关系正的，并人工去核查穿堤涵洞、地下河道等特殊要素。

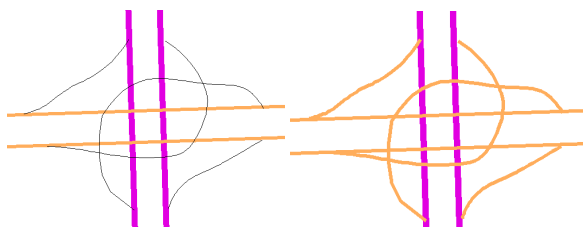


图1 匝道表示

4.1.4 建筑物融合的几何一致性处理

建筑物融合的主要缘由是轮廓差异以及属性缺失方面的挑战，采用以下流程：

1. 空间位置比对：把高精度影像当作基准，比较两套数据的建筑物边界。
2. 轮廓优化：若位置偏差大于1米，采用上海数据；若偏差小于或等于1米，取两者的平均位置。
3. 属性继承：把国家数据当中的“建筑结构类型”“竣工年代”等属性补充到上海数据。
4. 拓扑重建：重建建筑物、道路以及围墙之间拓扑关系，能确保建筑不会压占道路红线。

4.2 地名地址与兴趣点 (POI) 数据融合

4.2.1 多源 POI 智能去重算法 多特征相似度计算方法

$$S = \alpha S_{name} + \beta S_{address} + \gamma S_{location} + \delta S_{category} \quad (2)$$

其中， S_{name} 为名称相似度， $S_{address}$ 为地址相似度， $S_{location}$ 为空间相似度， $S_{category}$ 为分类相似度，依据天地图 POI 分类树计算层级距离。决策规则：对于相似度大于 0.6 的数据，利用影像、地图街景数据进行核实，去除错误数据，如正确但存在重复，则保留公开版地图库 POI 点位，主节点数据进行补充，对相似度 0.6 以下数据，人工对统计结果进行检查，如发现重复数据，则进行修改。

4.2.2 POI 与基础要素的空间关系协同更新

当道路、水系等基础要素在融合后发生几何变化时，POI 的相对位置必须同步调整。基于拓扑关系的自动调整算法：

1. 关系类型判断：判断 POI 与基础要素的关系类型（临路、临河、院内等）。
2. 相对距离保持：计算原始数据中 POI 到参考要素的垂直距离或沿线距离。
3. 位置重建：将相同的相对关系应用到融合后的基础要素上，重建 POI 坐标。
4. 人工核查：对调整距离 > 10 米的 POI 进行人工核实。当道路中心线因精度提升而偏移后，沿街商铺 POI 自动进行相应位移，保持与道路的临接关系不变。

5 质量控制体系与融合效果评估

通过开展生产过程的质量控制，及时发现影响成果质

量的普遍性，倾向性问题，督促和指导作业队伍纠正技术偏差、调整作业方式，确保生产过程质量符合技术设计要求，主要有如下内容：

成果质量要求。在数据生产过程中应当对数据各方面进行质量控制，包括数据的数学基础正确、数据完整性、数据逻辑一致性、数据拓扑正确、数据接边正确、数据现势性与数据源现势性一致。

强化全过程质量控制。必须严格执行生产作业技术标准，加强各工序质量控制，落实人员质量责任、确保生产过程符合设计要求。

落实全过程质量检查。根据生产作业实际情况，组织专职质量检查人员对承担的生产任务开展全过程质量检查，对发现的问题及时归纳分析，并督促整改落实。

数据安全。在作业过程中应作好防护、保密等措施，以防范所涉及的数据资料泄密。地图上表示的内容严格遵守国家有关保密法律、法规的规定。所有对外发布的地图切片都应经过数据变形处理。

6 结论

本文基于“天地图·上海”数据融合实践，系统地进行国家级以及省级节点多源地理空间数据融合方法体系的总结，得出主要结论：

1. 提出适宜于多级节点数据融合的三级决策原则，“现势性优先、几何精度优先、要素一致性优先”的原则体系，为处理多源数据冲突提供清晰思路。
2. 构建全流程技术解决方案：鉴于矢量、POI、影像等不同类型数据的特性，设计差异化的融合流程，特别是在道路网络立体表达方面 POI 智能去重、影像色彩均衡等关键技术得以实现突破。
3. 建立可量化质量控制体系：借助全过程质量控制以及量化评估指标，实现融合成果在现势性、精度、完整性以及一致性方面的显著提升。

参考文献

- [1] 上海市测绘院. 上海市基础地理信息要素数据字典[R]. [出版地不详]: 上海市测绘院, 2023.
- [2] 国家测绘地理信息局. 国家地理信息公共服务平台涉密版建设指南[S]. 北京: [出版者不详], 2017.
- [3] YANG Y, GAO W, ZHANG X, et al. Unified algorithm for transformation between cgcs2000 and local coordinate systems[J]. Journal of Geodesy, 2020, 94(1):1-15.
- [4] 陈军, 刘建军, 赵勇. 多源矢量空间数据融合与更新中的几何匹配方法[J]. 测绘学报, 2018, 47(8):1121-1131.
- [5] 罗国玮, 叶嘉媛, 王金凤. 基于多特征相似性的多源 POI 匹配方法[J]. 测绘通报, 2022(004):000.
- [6] ZHANG Y, ZHANG J, CHEN J, et al. A color consistency adjustment method for multi-temporal remote sensing images[J]. Remote Sensing, 2019, 11(18):2130.