

# 利用 FME 实现 GIS 与 CAD 的语义转换

曾巧玲<sup>1</sup> 张书亮<sup>1</sup> 姜永发<sup>1,2</sup> 阎国年<sup>1</sup>

<sup>1</sup>(南京师范大学地理信息科学江苏省重点实验室,南京 210097)

<sup>2</sup>(江苏省基础地理信息中心,南京 210013)

E-mail:janeyrez@163.com

**摘要** 论文分析了 GIS 和 CAD 的数据结构,介绍了如何利用 FME 实现 GIS 数据与 CAD 的语义转换及其优势所在,采用这种方法进行数据转换,可以在数据转换过程中重构数据实体,在一定程度上解决了以往数据转换中存在的问题。

**关键词** CAD GIS 语义转换 数据结构 SDE

文章编号 1002-8331-(2005)13-0214-04 文献标识码 A 中图分类号 TP311

## Semantic Translation of GIS and CAD Using FME

Zeng Qiaoling<sup>1</sup> Zhang Shuliang<sup>1</sup> Jiang Yongfa<sup>1,2</sup> Lü Guonian<sup>1</sup>

<sup>1</sup>(Key Lab of Jiangsu Geographic Information Science, Nanjing Normal University, Nanjing 210097)

<sup>2</sup>(Jiangsu Fundamental Geographic Information Center, Nanjing 210013)

**Abstract:** The paper analyses the data-structure of GIS and CAD, and introduces how to translate data between GIS and CAD. Using that method, data-entity could be rebuild during translation, which solves some problems existing in data-translation a certain extent.

**Keywords:** CAD (Computer Aided Drafting), GIS (Geographic Information System), semantic translation, data-structure, SDE (Spatial Database Engine)

CAD (Computer Aided Drafting: 计算机辅助设计) 技术,以其操作简单、易于编辑、出图美观、可二次开发等优点广泛应用于诸多行业,如建筑、市政以及电力等行业,这些行业部门都有大量的 CAD 数据档案。GIS (Geographic Information System: 地理信息系统) 技术具有便捷的地图显示、地理信息查询和强大的空间分析功能<sup>[1]</sup>,在数字产品的管理与应用方面明显优于 CAD 技术<sup>[2]</sup>。GIS 与 CAD 相结合可以取长补短,寻求二者的结合,即解决 GIS 与 CAD 数据共享与互操作已成为当前 GIS 的研究热点之一。由于 GIS 与 CAD 的图形数据格式各异,没有统一支持的格式标准,这给信息共享带来了极大的不便。实现 GIS 数据和 CAD 数据的相互转换是解决数据共享的一个主要途径,但是用传统的方式转换数据,往往造成数据信息的丢失和冗余,或者在使用 GIS 或 CAD 系统前,必须对转换过来的数据进行二次加工处理,严重影响了数据质量。论文主要探讨了基于语义转换的理念,利用加拿大 Safe Software 公司的软件 FME (Feature Manipulate Engine: 要素操作引擎) 实现 GIS 数据和 CAD 数据的转换,并且在南京市排水 GIS 系统中得到了应用。

## 1 语义转换

### 1.1 语义转换

语义转换是由 OpenGIS 协会 (1996 年) 提出的一个术语,允许用户在数据转换过程中重新构造数据,用户可以根据其特殊的需求,提取相同数据源的不同层面和内容,而不是以单一格式输入数据。

语义转换是基于语义层次上的空间数据转换,它除了数据

结构的转换外,更重要的是对语义数据模型的转换和操作。语义转换相比传统数据转换有着很大的不同。如图 1 所示。

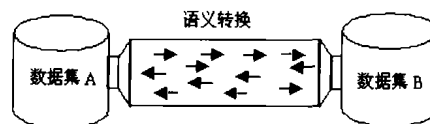


图 1 语义转换概念图

如图所示,语义转换就像一个宽宽的管子把两个数据集连接起来,数据转换双方可以自由地进行数据的转换与共享。在转换器中不仅有与数据相关的几何/空间信息还有语义信息。语义转换具有以下特点:

- (1) 定义了数据集间的相互映射 (mapping);
- (2) 使用要素 (feature) 操作语言;
- (3) 允许转换方案的重定义;
- (4) 提供查找、复杂计算的函数;
- (5) 提供从原始数据生成过程数据的方法。

语义转换如同一个引擎,通过要素操作语言,不仅可以对输入数据而且也可以对输出数据进行数据的重定义。因为在语义转换的背后存在一个丰富的数据转换模型,它具有内部一致性和外部可扩展性。

如图 2 所示,数据模型 A 和数据模型 B 可以自由映射到数据转换模型中,而这种映射不是基于最低数据转换标准的映射机,即不是基于公共要素的映射。

在这种映射中,针对输入、输出数据,数据转换模型提供了一系列的方法来实现数据模型之间的定义和转换。这种功能使

得数据的转入方和数据的转出方之间可以自由地转换,并且可以继续地使用各自独立的系统和数据格式。在数据转换模型中的映射不仅能够实现高度的定制,而且这种映射是双向的。因此潜在信息丢失的问题被如何更好地扩展信息内容、如何更好地满足用户的需求所替代。这样,用户或者开发者只需要了解输入数据的实际描述和输出数据希望达到的要求就可以完成转换过程。

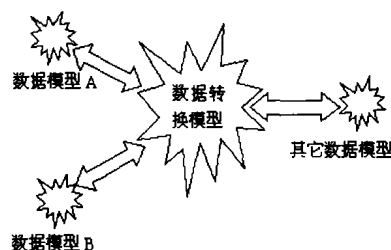


图2 基于语义数据转换模型示意图

## 1.2 FME 简介

数据转换的方法有很多,论文主要利用长期致力于无缝数据转换解决方案研究的加拿大 Safe Software 公司推出的空间数据转换处理系统 FME Suite,实现 GIS 数据与 CAD 的语义转换。该系统基于 OpenGIS 组织提出的新的数据转换理念“语义转换”,通过提供在转换过程中重新构造数据的功能,实现了超过 100 种不同空间数据格式/模型之间的相互转换。语义转换允许用户在转换过程中重新构造数据实体,以解决以往同构转换软件中存在的问题<sup>[4]</sup>。

FME 是对语义数据转换方法的实现,它不再将数据转换问题看作是从一种格式到另一种格式的变换,而是力求将 GIS 要素同构化并向用户提供组件以使用户能够将数据处理为所需的表达方式。FME 已被业界广泛承认和采用,实现不同格式或系统间的数据转换。FME 所执行的整个数据转换过程都通过语义映射文件来控制,该文件指定执行数据转换的方式。

FME 由三个主要模块组成:语义数据转换引擎:这是 FME 的核心模块,负责实际的转换工作。语义映射文件注册表:该注册表提供了用于管理众多不同语义映射文件的机制。该注册表提供了简单的过滤操作,该操作用于在数据转换过程中选择所需的映射文件。注册表提供了分类机制,用于在不同用户间开发和共享数据转换规范。自动语义映射文件生成器:可使 FME 以完全自动的方式执行转换。FME 数据转换体系结构如图 3 所示。

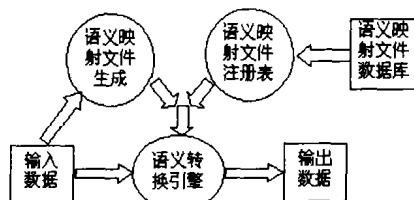


图3 FME 数据转换体系结构

## 2 数据结构分析

### 2.1 GIS 的数据结构

GIS 数据结构基本上可分为两大类:矢量结构和栅格结构,也可以称为矢量模型和栅格模型。两类结构都可用来描述地理实体的点、线、面三种基本类型<sup>[9]</sup>。矢量结构即通过记录坐

标的方式尽可能精确地表示点、线、多边形等地理实体,它的一个显著特点是能够完全表达点、线、面的空间位置和它们之间的所有空间关系,建立拓扑结构,从而提高空间网络分析的能力。栅格结构是将地球表面划分为大小均匀紧密相邻的网络阵列,每个网格作为一个像元或像素由行、列定义,并包含一个代码表示该像素的属性类型或量值,或仅仅包括指向其属性纪录的指针。栅格数据结构中,地理实体使用网格单元的行和列作为位置标记符,这些离散的数据网格具有叠加方便、迅速的特点,所以常被用来作为空间物体的背景图。

### 2.2 CAD 图形的数据结构

CAD 的图形元素按照一定的数据格式与数据结构经过特殊的压缩处理记录在 DWG 文件中<sup>[10]</sup>。DWG 文件格式可读性差,通常用户都是从 CAD 的数据交换文件(后缀为.DXF)中来读取数据。

图形数据具有量大、结构复杂等特点,且包括几何信息与非几何信息。下面分析某一 CAD 文件片段中所包含的几何信息中涉及的实体类型与图形数据结构。

假设在 line 层中画一条从(20.56,32.56)到(33.24,54.11)的直线,然后用 AutoLISP 读出该直线的图形数据为:

```
((24.<ENTITYNAME:60000018>)
```

```
(0."LINE")
```

```
(8."LINE")
```

```
(6."CONTINUOUS")
```

```
(62.3)
```

```
(10 20.56 32.56 0.0)
```

```
(11 33.24 54.11 0.0))
```

该表中的大部分数据与图形交换文件(后缀为.DXF)格式是一致的。在上面包含直线图形数据的表中,第一个子表的组码为 21,表示实体名,其组值表示实体数据表所表示的实体。第二个子表的组码为 0,表示开始一个实体,此实体的类型由此后的组值的字符串的值给定,它可以是“LINE”(直线),“CIRCLE”(圆),“ARC”(圆弧),“POLYLINE”(多义线)等,且该字符串的所有字符都用大写字母表示。第三个子表的组码为 8,其值为一个表示该实体所在图层名称的字符串(其中的字符全用大写表示)。这三个子表在每个实体的数据中是不可缺省的。第四个子表的组码为 6,其组值表示线型名,如“CONTINUOUS”或“DASHED”等。第五个子表的组码为 62,其组值表示颜色号,当线型名和颜色号不为“BYLAYER”时,则此子表出现,否则缺省。

## 3 数据转换前数据组织与设计

### 3.1 合理组织数据

合理组织数据是保证数据转换质量的一项重要保证。各类数据应根据数据类型、专题要素进行合理分层。图层的划分宜细不宜粗,因为将不同层上的实体合并成一层比较容易,而将一层上的内容再分成多层则要困难得多。在图层划分时,应遵循以下原则:各层要素类型单一,不同专题要素分层存放,图层命名规范,即一层中只有点要素,或者只有线、面要素,并且一层中只有一种专题要素。

由于 GIS 与 CAD 中支持的线型、符号类型不同,要使转换后的数据与原数据尽量一致,在数据转换前在 CAD 中建立与 GIS 相对应的符号,并建立符号库表,使 GIS 与 CAD 中的符号

类型,符号名称一一对应。

### 3.2 设计 GIS 数据结构

在合理组织数据的基础上,按照 CAD 图形的数据结构设计思想设计 GIS 的数据结构,也就是在 GIS 中设置与 CAD 图形数据相对应的属性字段(如:颜色,线宽,符号样式等属性)。其中,注记在 GIS 中的表现方式与 CAD 不同,在 CAD 中有专门的文本层来表示注记,而在 GIS 中只有点、线、面层,没有文本图层,通常以间接的方式来表现注记,把注记的内容及样式等记录在图层的属性里。GIS 的数据结构如表 1 所示。

表 1 GIS 数据结构

| 对象    | 数据结构内容             |
|-------|--------------------|
| 点、线、面 | ID、颜色、大小、符号样式      |
| 文本    | ID、高度、颜色、旋转角度、字体内容 |

### 4 应用案例

排水管网系统是重要的市政基础设施,是一个城市的地下“动脉”,它与城市建设发展的速度和城区的环境效益息息相关。随着经济快速发展,城市建设突飞猛进,排水管网系统变得越来越复杂、越来越庞大,具有大量的文档和图形资料,传统的方式已无法适应城市管线现代化管理的需要,用计算机信息系统进行管理迫在眉睫。南京市排水 GIS 系统正是基于目前排水管网管理的这种需求开发的,它实现了对城市排水设备设施的定量化管理,并辅助排水管网规划决策。

在建立南京市排水 GIS 系统之前,南京市排水管网的资料大多是些文档以及 CAD 格式的规划图。充分利用这些已有数据可以大大缩减 GIS 系统的建设时间,避免重复劳动;同时,在利用 GIS 管理排水管网的同时,充分发挥 CAD 的制图、出图等的优势,对于城市排水管网的管理将会更加有效。为了在排水管网的管理中实现二者的结合,实现 GIS 数据与 CAD 数据的相互转换是必要的。在南京市排水 GIS 系统的建设中,作者利用 FME 实现了 GIS 数据与 CAD 的语义转换,而且达到了较高的数据质量。

下面,将以排水管网数据从 GIS 向 CAD 为例,详细阐述如何利用 FME 实现 GIS 与 CAD 数据的语义转换。基于同样的方法,可以实现数据从 CAD 向 GIS 的转换,限于篇幅,不再一一赘述。

#### 4.1 分析源数据

南京市排水 GIS 系统的空间数据,使用 ESRI 公司的空间数据库管理软件 SDE(Spatial Database Engine:空间数据库引擎)来管理,各不同数据类型、专题内容的数据分层存储,包括管线(雨水管、合流管、污水管等),管点(雨水井、出水口、泵站等)等排水基本设备图层,注记、管线流向箭头等装饰图层,以及道路、水系等基础地理背景图层。在该系统中,主要的图层如表 2 所示。

表 2 源数据分析

| 图层名称                 | 几何类型 | 专题内容 |
|----------------------|------|------|
| sde_SDE_RoadBoundary | 线    | 雨水管  |
| sde_SDE_PSGH_YSG     | 线    | 雨水管  |
| sde_SDE_PSGH_JCJ     | 点    | 检查井  |
| sde_SDE_PSGH_ARROW   | 面    | 箭头   |
| sde_SDE_PSGH_anno    | 点    | 注记   |
| .....                |      |      |

如表 2 所示,源数据包括点层(sde\_SDE\_PSGH\_JCJ 等)、

线层(sde\_SDE\_PSGH\_YSG 等)、面层(sde\_SDE\_PSGH\_ARROW 等);其中,在 SDE 中文本层(注记)事实上是点层,点的 X、Y 位置值表示注记在图上的位置,把注记的内容存储在‘ZJNR’字段里,表示注记字体的高度、颜色、旋转角度等也用相应的属性字段来表示。

#### 4.2 新建工作空间

现在利用 FME Workbench 功能模块,定制数据转换完成复杂的数据转换任务,而且达到较高的数据质量。首先创建一个工作空间(workspace),选择源数据(Esri Geodatabase(Arcsde))格式和目标数据(AUTOCAD DWG/DXF)格式,以及源数据文件。FME 自动建立了源数据和目标数据对应关系,即自动生成了语义映射文件,如图 4 所示。

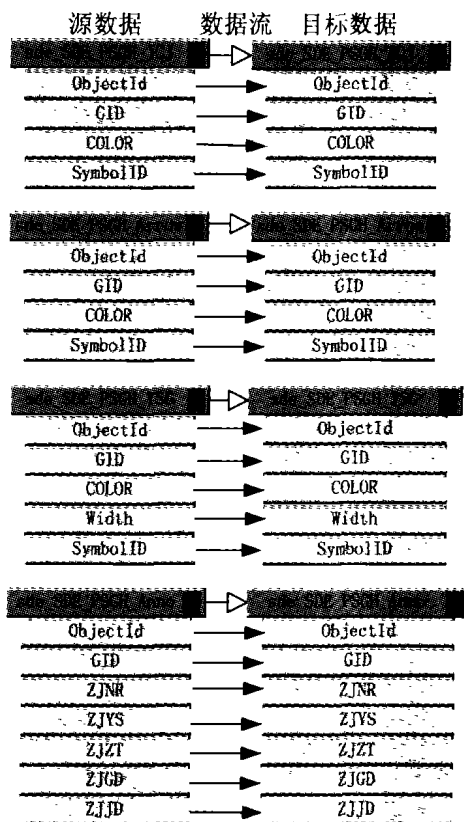


图 4 新建工作空间关联图

如图所示,用可视化的方式(几何图形)来表示源数据和目标数据,在几何图形里,用长方形的条块来表示这种空间数据中的属性数据结构,用箭头来表示这个字段是否被导入或被导出,用带箭头的线段来表示源数据和目标数据的对应属性关系。在此过程中,在空间坐标系注册器中存在源数据的空间坐标系,空间坐标系自动从源数据转换到目标数据中。

#### 4.3 自定义源数据和目标数据的对应关系

上一步已经创建了工作空间,并且建立了源数据和目标数据的默认关联关系,生成了自动语义映射文件,为了使转换的数据达到预期要求,在这里需要重新建立源数据和目标数据的对应关系,即属性和图形的对应关系,其实质就是重新整理数据,按要求生成语义映射文件,以用户希望的方式表达数据。重新关联过的源数据和目标数据,如图 5 所示。

同样,在可视化环境中,定制数据转换形式,对源数据和目标数据作相应处理,并且可以加入一些附加信息。

具体到南京市排水 GIS 系统,其中,sde\_SDE\_PSGH\_JCJ 图层是点层,把目标数据的 CAD 层的 autocad\_entity 属性赋值

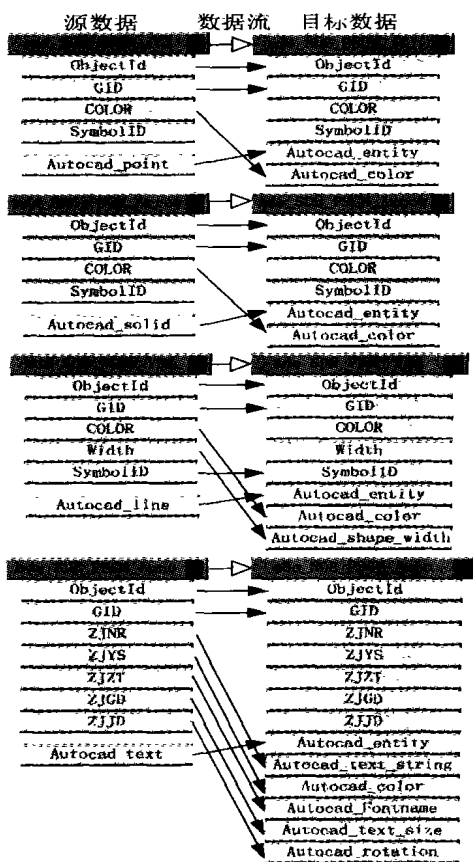


图5 重新定义源数据和目标数据的对应关系

为 autocad\_point, 此处 autocad\_point 表示该层的实体类型是 point; 以此类推, 线层 sde\_SDE\_PSGH\_YSG 在 CAD 中的实体类型是 line; 面层 sde\_SDE\_PSGH\_ARROW 在 CAD 中的实体类型是 solid; 同时, 可以把各层符号的颜色 COLOR 对应为 CAD 中的 AutoCAD\_color 属性, 对线层的宽度字段 Width 对应为 CAD 中的 AutoCAD\_shape\_width 属性。图层 sde\_SDE\_PSGH\_anno 比较特殊, 在 GIS 中该层主要是表示文本注记信息, 转换到 CAD 中该层的实体类型是 text, 即文本类型, 同时在 GIS 中的 ZJNR, ZJYS, ZJZT, ZJGD, ZJJD 字段分别表示注记的内容, 颜色, 字体, 高度, 角度, 对应到 CAD 中, 它们也可以表示 CAD 中相应对象的图形属性。

(上接 162 页)

表3 测试结果(单元格中数据是: 正确召回的文档数 / 召回的总文档数)

| 次数 | 文娱          | 财经        | 体育          |
|----|-------------|-----------|-------------|
| 1  | (1239/1507) | (846/914) | (975/977)   |
| 2  | (1188/1264) | (796/866) | (1074/1120) |
| 3  | (1171/1229) | (844/893) | (1114/1189) |
| 4  | (1255/1413) | (824/861) | (1100/1105) |

从测试结果来看, 经过修正型增量式反馈后, 过滤效果虽然在局部范围有所动荡, 但整体上明显改善。

## 6 结束语

作者针对海量信息处理需求提出了精细过滤模型, 在精细过滤模型的基于向量空间模型的精细过滤阶段, 利用修正型增量式反馈和扩充再还原修改用户兴趣向量的方法不断逼近用户的真实需求。从测试效果来看, 这个方法是有效的。但同时也发现有许多值得再改进的地方, 比如在给扩展的关键词确定初

最后语义转换引擎根据语义映射文件, 按要求转换数据, 结果可以使转换的 CAD 数据基本上符合用户的需求, 并且达到与 GIS 中同样的图形效果, 极大地减少了数据转换后的图形修改的工作量。以上是 GIS 数据向 CAD 数据的转换, 同理, 只要定义规范的语义映射文件即可以实现 CAD 数据向 GIS 数据的平滑转换。

## 5 结语

语义数据转换比传统的数据转换方法具有更强大的功能和灵活性, 使数据在众多格式间更加容易转换, 而且转换的数据具有较高的数据质量, 转换后的数据经过少量处理或不用处理就可以达到预期的效果。当然, 目前利用 FME 的语义数据转换也还具有一些缺陷。属性数据问题: 有的 GIS 软件的数据格式中属性数据是以文件的方式和图形数据存放在一起, 有的属性数据则存放在专业数据库中(如 MAPINFO 的 Tab 格式), 由于文件和数据库, 以及各数据库之间所支持的数据类型不一致, 所以属性数据也会有所丢失; 拓扑关系: 当从支持拓扑关系的数据格式向不支持拓扑关系的数据格式转入数据时, 必须打散拓扑关系, 拓扑关系丢失, 当从不支持拓扑的数据格式向支持拓扑的数据格式转入数据时, 不能重新建立拓扑关系; 符号: 由于各个 GIS 系统, 及 CAD 对符号的表示方法各异, 在数据转换过程中, 符号的转换需要做事先处理。完善语义数据转换, 解决数据转换过程中的瓶颈问题, 还需要进一步的研究。

(收稿日期: 2004 年 8 月)

## 参考文献

1. 潘建群, 周方, 唐礼勇等. 综合地理信息系统及态势图处理系统在 C3I 系统中的应用[J]. 系统工程与电子技术, 1998; (1): 4~8
2. 周勇前. GIS 与 CAD 数据交换[J]. 测绘信息与工程, 1996; (3): 16~19
3. U Visser\*, H Stuckenschmidt, G Schuster et al. Ontologies for geographic information processing Computers & Geosciences, 2002; 28: 103~117
4. FME 技术白皮书
5. 郭伦, 刘瑜等. 地理信息系统——原理、方法和应用[M]. 科学出版社, 2001
6. 郭朝勇等. AUTOCAD 2002 定制与开发[M]. 北京: 清华大学出版社, 2002

始权重时, 是不是应该有更科学的方法; 在伪相关反馈时, 表征类的核的文档、用户反馈的文档以及典型不相关文档之间的分布是不是有更好的模型。(收稿日期: 2004 年 12 月)

## 参考文献

1. James Allan. Incremental Relevance Feedback for Information Filtering. <http://citeseer.nj.nec.com/allan96incremental.html>
2. Bin Wang, Hongbo Xu, Zhifeng Yang et al. TREC-10 Experiments at CAS-ICT: Filtering, Web and QA. <http://trec.nist.gov/pubs/trec10/papers/CASICT.pdf>
3. Kjersti Aas. A Survey on Personalised Information Filtering Systems for the World Wide Web[R]. Norwegian Computing Center, Report No. 922, 1997
4. G Salton, M J McGill. An Introduction to Modern Information Retrieval[M]. McGraw-Hill, 1983
5. Ricardo Baeza-Yates, Berthier Ribeiro-Neto. Modern Information Retrieval[M]. ACM PRESS