 DOI: https://doi.org/10.12349/iser.v7i6.10802

A Plotting Scoring Algorithm Based on Large-Sample Fuzzy Matching

Tianyu Pan Bing Guo Wenlong Dun Qingqing Miao

The 28th Research Institute of China Electronics Technology Group Corporation, Nanjing, Jiangsu, 210000, China

Abstract

This paper designs a plotting scoring algorithm based on large-sample fuzzy matching. It transforms the traditional method of manually judging and scoring map plotting assignments one by one. By deconstructing the generation process and principles of graphical plotting, the algorithm performs text parsing and relationship mapping on the generated SML files. It assigns different weights to various evaluation types and calculates a weighted sum. A positive feedback mechanism is introduced, allowing instructors to rate the output results (high or low), thereby enabling the algorithm to continuously optimize evaluation factors and weight coefficients. Finally, the algorithm is validated through a software implementation. The results indicate that with the addition of a sufficient number of evaluation factors, the algorithm's results closely approximate the actual scores of the assignments.

Keywords

plotting scoring; SML parsing; fuzzy matching; evaluation factors; self-learning algorithm

一种基于大样本模糊匹配的标绘评分算法

潘天宇 郭冰 顿文龙 苗清清

中国电子科技集团第二十八研究所，中国 · 江苏 南京 210000

摘 要

本文设计了一种基于大样本模糊匹配的标绘评分算法，改变了传统图上标绘作业需要人工进行逐个判断评分的方式，通过拆解图形标绘的生成过程和原理，对标图生成的SML文件进行文本解析和关系映射，对不同的评价类型赋不同的权重进行加权求和。引入正反馈机制，让教员对输出的结果进行高低评价，让算法不断修正评价因子和权重系数的方式实现算法优化。最后通过该算法实现的软件程序进行验证，结果表明，在添加足够多的评价因子后该算法的结果已接近作业的真实得分。

关键词

标绘评分；SML解析；模糊匹配；评价因子；自学习算法

1 引言

当前地理信息系统（GIS）、现代标绘技术以及大模型技术的不断发展，图上标绘的能力已跃升为军事领域、突发应急救援及城市规划布局等关键领域的核心作业环节。标绘作业的精准与否直接决定了态势研判的可靠性以及指挥决策的科学化水平。然而，针对标绘成果的考评工作主要倚赖资深教员或专家进行人工目进行批阅，这种传统的人工模式在实际应用中会遇到很多问题：首先，面对大规模演习或考核产生的海量标绘数据，人工逐一比对不仅耗时费力，且处理效率极其低下；其次，评分标准往往受限于评审者的个人经验与主观认知，这种“人治”模式导致不同评审员之间，甚至同一评审员在不同时段的评判尺度存在显著差异，严重削弱了评分结果的客观公正性与可复现性；此外，人工肉眼难以对微小的几何偏差（例如细微的位置漂移、角度偏转等）也难以做到客观的判断。

鉴于此，如何将标绘评分标准进行量化特征提取与科学的数学建模，通过不同评价维度赋予合理的权重系数，利用计算机技术自动识别作业与标准答案间的差异，并通过构建一套可解释的自定义评分算法来测算误差范围，最终输出客观、高效且合理的评分结果，已成为标绘评价的重要研究方向。基于此，本文致力于设计并实现一种基于大样本模糊匹配机制的智能化标绘评分算法，旨在为相关领域的教学训练与效能评估提供坚实的技术支撑。

2 国内外研究现状

标绘作业的自动化评估是一门多科学交叉的研究内容，主要涉及到文件解析、目标识别、差值拟合等多个工程领域。目前，国内外的相关研究主要集中在以下几个方向：

基于规则的严格匹配模式。将图形化的表征转化为数值的计算。例如，Kong 等人 [2] 提出了一种将矢量地图转化

【作者简介】潘天宇（1993–），男，中国江苏南京人，硕士，工程师，从事工程信息化研究。

为 XML 配置的方法，通过预定义的结构化的数学定义解释图形的信息。这类方法直观易懂、结构严谨，在实际的发展过程中中也进行的大规模的应用，但对于多批次标号的统一评价问题未进行过深入的实践。

基于图像处理的相似性度量方法。此类技术将标绘成果视作常规光栅图像，依托计算机视觉算法实施比对分析。例如，Zhang 等学者 [3] 采用结构相似性（SSIM）指标来量化两幅地图影像的视觉一致性。另有研究利用卷积神经网络（CNN）提取地图图像的深层特征向量，进而借助余弦相似度算法完成匹配任务。尽管该类方法的优势在于无需解析底层数据架构，但其局限性亦十分突出：一方面，矢量图形所包含的关键语义信息（诸如符号类别、属性参数及拓扑关联）在栅格化过程中流失，导致系统难以精准定位错误性质（比如无法有效甄别“符号误标”与“要素漏标”）；另一方面，该类算法对图像的分辨率缩放、色彩模式变更及背景噪声干扰表现出较高的敏感度。

基于矢量数据的几何与拓扑解析方法 [4-5]。鉴于标绘数据的本质属性为矢量格式，部分学者主张直接在矢量层面开展分析工作。Chen 等研究人员提出了一种融合 Hausdorff 距离与 Fréchet 距离的混合度量模型，旨在评估轨迹线或多边形要素的相似程度。该方法在应对线性或面状要素的几何形态偏差方面表现优异。然而，通用的几何度量标准往往难以兼顾军事或特定专业标绘中符号所承载的专用语义及作业规范，导致在特定场景下的适用性受限。

3 标绘和 SML 解析

GIS 标绘上图主要分为 WebGIS 和桌面 GIS 两种展现方式。其中 WebGIS 因其独特的 B/S 架构，具备轻量化部署和易访问特征而更受用户青睐 [6]。

在设计标绘自动评分算法前，要理解在 GIS 上实现标绘和显示的原理。通过预置的各类标准的图形标号库，用户快速选择想要的标号，通过对标号的再编辑和组合，构建标绘图的基本元素。其核心的图形标号本质为基于 JS 的前端控件，将矢量图进行三维绘制，并转化为一段描述其几何形状的 JS 代码如图。将该图形标号以数字代号加入原有的图形标号库中，通过该代号快速查询到这个图形标号。对于不同类型的标号则需要对该图形标号的前端代码有更多的描述。

SML 标绘文件是一个标准的可扩展标记语言文件（XML 文件），它包含了每一个图形标号上图显示需要的所有特征信息，以系统能够识别的配置和用户自定义的数据结构形式，如表 1 是一个典型的 SML 片段。

表 1 SML 文件标号信息

```
<GOBJ JBID="0_0_76_ 点 (160-27)">
<SYMATTR>
<CODE>27</CODE>
<LIBID>160</LIBID>
<USERID>76</USERID>
</SYMATTR>
<GEOATTR>
<LEFT>121.609291</LEFT>
<TOP>31.306522</TOP>
<RIGHT>121.609291</RIGHT>
<BOTTOM>31.306522</BOTTOM>
<PX>121.609291</PX>
<PY>31.306522</PY>
</GEOATTR>
<DRAWATTR>
<COLOR>255</COLOR>
<LINEW>2</LINEW>
<ANGLE>0</ANGLE>
<BASESCALE>10946.995117</BASESCALE>
</DRAWATTR>
</GOBJ>
```

通过解析 SML 文件，我们可以精确获取标号的代号、库号、边界经纬度、中心点坐标、颜色、线宽、旋转角度以及基础缩放比例等关键参数 [7]。这些参数构成了算法进行自动比对的数据基石。

4 算法构建

本文设计的评分算法核心逻辑是将复杂的标绘作业分解为可量化的评价因子，并通过加权求和的方式得出最终分数。评价主要关注“程度”和“精度”问题，如位置偏差、大小缩放、旋转角度等。这类评价需要建立连续的度量函数。

对于面状图形，计算其最小外切矩形面积；对于线状或不规则图形，计算特征点构成的多边形面积。计算学员作业与标准答案的面积差值，以及特征点坐标的欧氏距离误差。为了处理人为操作的误差，算法引入了模糊匹配机制。对于定量评价，我们设定一个误差容忍阈值时，视为无误差，得满分；当时，采用指数衰减函数进行扣分，以模拟人类评分的宽容度。

5 软件验证

为了验证算法的有效性，我们开发了自动评分软件系统，并进行了大规模的实验。实验环境为 Windows 10 系统，Intel i7 处理器，16GB 内存。实验数据选取了某次军事演习中的 50 份标准答案及对应的 200 份学员作业。我们将标准答案和学员作业导出为 SML 文件后，导入到自动评分系统中进行计算，同时邀请 3 位资深教员进行人工评分，取平均值作为基准，结果如表 2。实验结果显示，该算法能够有效识别各类常见错误。例如，当学员标号与正确标号中心点距离超过阈值时，系统自动判定位置错误并扣分；当特征点缺失时，系统能精准定位。

表 2 软件自动评分与人工评分对比

作业编号	错误类型描述	自动评分	人工评分	误差
A01	直升机位置偏差 2.4km，大小不符	82.5	84.0	-1.5
A02	缺少 1 个特征点，漏标 1 个标号	75.0	73.5	+1.5
A03	数量不对（多标 1 个）	88.0	89.0	-1.0
A04	旋转角度偏差过大	80.5	79.0	+1.5

从上表可知，该算法的评分结果与人工评分结果高度相似，最大误差控制在 ±2 分以内，平均相关系数达到 0.92。通过软件验证，该算法能够实现对标绘作业的客观评价，尤其对于作业量大、标号较多的场景可有效提升教员的批改效率。通过枚举尽可能多的评价因子，设置不同评价因子之间的相互影响域，规定不同的偏差下的扣分区间从而构建正确的评价算法。

同时教员在分析结果后，可不断修正权重因子数值，从而使算法的输出结果更加的合理。在后期工作中，可引入大模型方法，通过建立作业案例库和评分结果知识库，引导算法根据不同的场景找到合适的权重。引入正反馈机制，让教员对输出的结果进行高低评价，通过大模型自学习能力，让算法不断修正因子和权重方式实现算法优化。

6 结论

本文创新提出了一种基于大样本模糊匹配的标绘评分算法。通过深度解析 SML 矢量文件，将复杂的标绘评估问题分解为可量化的评价因子，并通过加权求和模型得出综合评分。实践证明，该方法不仅显著提高了评分效率，同时还保证了评分的客观性和一致性。通过引入案例库、知识库和正反馈学习机制，算法具备了自我优化的能力，不断逼近专家级的评判水平。后期将引入大模型的推理能力，通过优秀的教师评分案例引导系统的自我学习。

参考文献

[1] 史云松. 基于ArcGIS Server的网络地图服务系统研究与实现［D］. 南京：南京林业大学，2012

[2] Kong，X.，Zhang，L.，& Li，D. (2018). A rule-based approach for consistency checking of vector map data using XML schema. ISPRS International Journal of Geo-Information，7(5)，189.

[3] Zhang，Y.，Wang，H.，& Liu，Y. (2019). Map similarity assessment based on structural similarity index.Cartography and Geographic Information Science，46(4)，321-330.

[4] 朱卫东.基于 GIS 的通用军标标绘软件的设计与实现[D].南京:东南大学，2015.

[5] 袁华.海军战术仿真想定空间辅助生成研究[M].北京:海潮出版社,2010.

[6] 潘辰. 基于 WebGIS 的轻量级态势显示系统设计与实现［D］. 北京: 北京邮电大学，2018.

[7] 牛红光,李汉荣,陈超等.海图标绘系统的开发和实现[J].第二十届海洋测绘综合性学术 研讨会论文集,2008(9):236~239.