

# 基于图像识别技术的固井质量评价方法研究

邹卓峰, 张宝权\*, 李辉, 王建华, 王海涛, 管震

(昆仑数智科技有限责任公司, 北京 102206)

**摘要:**关于固井实施方法和固井质量工程探测方法研究颇多,而关于固井质量的图像解析却鲜有研究。本文就使用评价图进行固井质量评估问题,提出了用于聚焦图像解析范围的文本坐标位置参照算法,以及基于直方图相似度和梯度距离的图像解析算法。以图像解析算法和业务机理相结合的方式,从评价图中抽取可评估固井质量的量化数据。所提出的图像解析算法能准确识别图片中的固井质量信息,两种算法的解析结果与实际结果误差率均在10%以内。本文所提出的方法可有效解决人工手动分析效率低且分析误差大的缺点,并对固井质量电子数据缺失仅留存固井质量评价图的情况提供了切实可行的解决方案,所提出的算法对于固井质量图片的深入挖掘具有较强的探索性意义。

**关键词:**固井质量;直方图;梯度距离;评价图;图像解析

**中图分类号:**TE256;P634 **文献标识码:**A **文章编号:**2096-9686(2024)S1-0104-08

## Cementing quality evaluation method research by image recognition technology

ZOU Zhuofeng, ZHANG Baoquan\*, LI Hui, WANG Jianhua, WANG Haitao, GUAN Zhen

(Kunlun Digital Technology Limited Liability Company, Beijing 102206, China)

**Abstract:** There are many researches on cementing implementation methods and engineering detection methods of cementing quality, but the analysis of picture results on cementing quality is relatively rare. In this paper, a text coordinate position reference algorithm for focusing image analysis range and an image analysis algorithm based on histogram similarity and gradient distance are proposed to evaluate cement quality using evaluation image. A combination of image analysis algorithm and business mechanism is used to extract quantitative data from the evaluation image to evaluate cementing quality. Through comparison of model experiments, the proposed image analysis algorithm can accurately identify the cement quality information in the picture, and the error rate between the analytic results and the actual results is less than 10%. The method proposed in this paper can effectively solve the shortcomings of low efficiency and large error of manual analysis, and provides a practical solution to the situation that electronic data of cementing quality is missing and only the cementing quality evaluation image is kept. The proposed algorithm has strong exploratory significance for deep mining of cementing quality pictures.

**Key words:** cementing quality; histogram; gradient distance; evaluation image; image analysis

## 0 引言

固井工作是石油钻井过程中不可或缺的一个环节<sup>[1]</sup>,固井作业旨在增强井壁稳定性,确保持续安全钻进,同时封隔井下油、气、水层<sup>[2]</sup>。固井工作通

常是一次性完成的,因此其质量至关重要。通常采用声幅变密度法对固井质量进行检测,并根据固井质量检测数据生成固井质量评价图。然而,评价图目前仅提供定性分析结果,未能呈现详细的数据化

收稿日期:2024-03-22;修回日期:2024-06-06 DOI:10.12143/j.ztgc.2024.S1.015

基金项目:中国石油天然气集团有限公司科研项目“钻完井及井下作业智能优化系统研发”(编号:2021DJ7401)

第一作者:邹卓峰,男,汉族,1983年生,工程师,勘查技术与工程专业,从事智能钻完井技术与信息系统建设工作,北京市石景山区京原路7号,zouzhuofeng@cnpc.com.cn。

通信作者:张宝权,男,汉族,1988年生,工程师,计算科学与技术专业,硕士,从事智能钻完井算法研发工作,北京市石景山区京原路7号,zhangbaoquan01@cnpc.com.cn。

引用格式:邹卓峰,张宝权,李辉,等.基于图像识别技术的固井质量评价方法研究[J].钻探工程,2024,51(S1):104-111.

ZOU Zhuofeng, ZHANG Baoquan, LI Hui, et al. Cementing quality evaluation method research by image recognition technology[J]. Drilling Engineering, 2024, 51(S1):104-111.

分析结论。因此,对固井质量评价图进行图像识别,分析特定范围内声波幅度、变密度以及第一界面和第二界面优、中、差图例的比例,实现从图像到数字化展示转换,具有重要意义。

本文基于固井工程声幅及声幅变密度测井的评价结果,开展了基于图像识别技术的固井质量评价方法研究。建立了表头文本信息的解析方法、图像聚焦与解析技术,并提出了创新的图像解析算法,该算法结合了梯度距离和直方图相似度,用于准确划分第一、二界面的图例类别。通过这些方法,成功实现了对固井质量评价结果的量化分析。结果显示,测试集中 90% 的测试样本的固井合格率与原数据合格率的误差控制在 5% 以内。

### 1 相关工作

目前国内外关于固井质量的研究主要集中在不同的固井方法和质量探测技术上。潜晓晨<sup>[3]</sup>发现泥浆配方、地层岩性、套管型号、测井时间及仪器偏心等因素对固井质量评价具有显著影响;吴华等<sup>[4]</sup>以页岩油水平井为例,通过提高水泥强度、降低杨氏模量等关键技术来保障固井质量;邵一臻等<sup>[5]</sup>提出了将高温固井水泥浆应用于山西干热岩井,并构建了干热岩的固井技术体系;王明辉等<sup>[6]</sup>提出了基于超声 Lamb 波的固井质量评价方法,提高了低密度轻质水泥固井质量评价的准确性。针对特殊条件下固井质量评价结果与实际情况差异较大的问题,陈向军等<sup>[7]</sup>提出了结合地层信息和 VDL 波形来判断和评价界面胶结质量的方法。

随着人工智能技术的兴起,机器学习技术和图像字符识别(OCR)技术也逐渐被应用于固井领域。肖红等<sup>[8]</sup>利用改进的 DenseNet 卷积神经网络解决了固井质量评价效率和准确率不高的问题;Jing 等<sup>[9]</sup>、郑双进等<sup>[10]</sup>则将经典的支持向量机(SVM)和支持向量回归(SVR)算法应用于固井评价;李忠庆等<sup>[11]</sup>基于气液多相流和达西渗流理论,建立了气体渗流数学模型,以指导气井固井作业。OCR 技术是计算机视觉领域的一个分支,它通常作为辅助技术手段,应用于标书、带有盖章的文书等具体的业务需求中<sup>[12-13]</sup>;钟琳<sup>[14]</sup>利用 OCR 技术识别设备状态和生产数据,以提高海上石油生产安全。与本文类似将图像解析技术应用于固井工作,杨玉卿等<sup>[15]</sup>、彭越等<sup>[16]</sup>将电成像测井图像识别、图像自动识别技术

应用于地层裂缝的研究中。

## 2 固井质量评价提升效率难点分析

### 2.1 工程技术评价难点

目前,国内外油气井固井作业完成后,通常采用声幅及声幅变密度方法进行测井,然而,在评价工作开展过程中,针对特定位置处固井质量的评价,往往依赖于人工定性的分析,效率较低。此外,在固井质量评价图中,虽然根据深度信息可以准确地反映特定井段的水泥胶结状况,磁定位、声幅、全波列、变密度等为不同测井技术所获得的固井评价信息,在分析的过程中可能会因多解性而导致固井质量评价结果的差异。

与此同时,在固井质量评价图分析的过程中,其中第一界面与第二界面使用表头(图 1)中的优、中、差图例来标示了如图 2 所示的不同深度的胶结质量情况。通过这个图例能够直观地区分不同深度范围内的水泥胶结质量。虽然第一、二界面的图例结果形成基于声幅曲线等数据得出,声幅曲线、全波列、变密度等参数列虽然可以更形像的展示不同深度的水泥固井情况,但在计算机图像识别过程中,声幅曲线中的标尺网络线会严重干扰声幅值的解析,全波列和变密度以波动密集程度反映不同深度的差异,但它们不提供具体的度量值,因此,在综合对比评价图中的各类图版后,解析以不同等级标示的第一、二界面结果更具可行性。

### 2.2 图像技术评价难点

图像识别技术目前已经开始在地层产状分析、

固井质量评价图																			
地区	溢吉桑	钻头程序	445.00 mm X 276.50 m 311.00 mm X 2268.00 m 118.00 mm X 2875.00 m																
坐标	X:4772716.98 Y:16310564.82	套管程序	339.70 mm X 276.08 m 244.48 mm X 2268.00 m 95.25 mm X 2375.00 m																
海拔高度	518.76 m	水泥设计返高	2168.00 m																
水泥浆密度	1.87 g/cm <sup>3</sup>	水泥实际返高	2168.00 m																
人工井底	2838.65 m	阻流环位置	2860.21 m																
测量井段	2168.00-2860.21 m	仪器型号	EILog																
测井日期	2019.1.14	解释程序	LEAD																
固井日期	2019.1.13	操作员	李仲彪																
比例尺	1:500	审核																	
测井小队	C1699 小队																		
解释绘图	刘正军																		
图例说明	<table border="0"> <tr> <td>自由套管</td> <td>油气储层图例</td> <td>油层</td> <td>干层</td> </tr> <tr> <td>胶结良好</td> <td>水层</td> <td>气水同层</td> <td></td> </tr> <tr> <td>胶结中等</td> <td>产油层</td> <td>低产油层</td> <td></td> </tr> <tr> <td>胶结差</td> <td>水淹层</td> <td>堵水层</td> <td></td> </tr> </table>			自由套管	油气储层图例	油层	干层	胶结良好	水层	气水同层		胶结中等	产油层	低产油层		胶结差	水淹层	堵水层	
自由套管	油气储层图例	油层	干层																
胶结良好	水层	气水同层																	
胶结中等	产油层	低产油层																	
胶结差	水淹层	堵水层																	
技术说明	本井深度依据2019年1月14日所测三样零长深度决定。																		

图 1 固井评价图表头示例

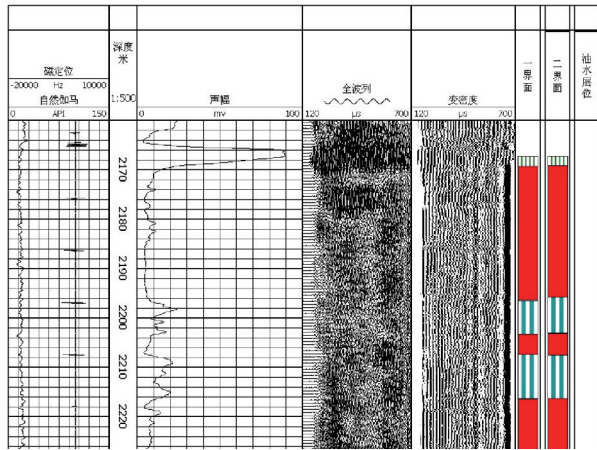


图2 固井评价图详情示例

储层裂缝识别等工程场景中得到应用,并且证明了较好的效果。其本质在于应用深度学习和神经网络处理固井质量图像中的模式和特征,并在数据集上进行训练之后,使用这些经过训练的模型能够准确识别新的固井质量评价图。在分析固井质量评价图的过程中,以每一种固井示例图作为标签,将全幅固井质量评价图上的每一部分区域与其对应的标签进行匹配,从而高效完成固井质量评价的量化分析。但是在图像技术应用的过程中,也存在两个方面的难点需要解决:

(1)固井质量评价图表头的解析精准度有待提高。在固井质量评价图中,通常表头部分以表格的形式呈现井深、水泥返高、固井水泥浆密度等信息,

表头是后续开展数据分析的依据。但是在表头中往往同时包含数字、以及横向或者纵向的文字,实现对其精准的识别是图像技术应用的关键。

(2)固井质量评价的量化存在困难。固井质量评价图中的优、中、差图例是分析结果的基本参照,依据这些图例,需要在全井固井质量评价图的胶结界面中找到具有相同特点的界面区间,同时,还需要在详情部分划定需要特定分析的范围以分析相关信息列,在细节和局部特征的识别和量化上面较大的难度。

### 3 考虑固井质量评价图的图像识别方法

本文采用了以神经网络为基础的开源 Paddle OCR 图像识别技术,该技术可以单独针对文本检测、文本方向分类、文字识别 3 个不同的模型进行预测。

在固井质量评价图识别过程中,首先会对评价图的表头信息进行解析,明确水泥返高等关键数据,为分析固定井段提供明确的限定范围;其次,根据水泥返高信息以及深度栏在各个深度的标签信息,划定分析评价图子图范围;最后,在特定分析井段上,聚焦于第一、二界面的子图,依据表头给出优、中、差图例,按类别统计优、中、差等级的占比,从而实现高效量化分析固井质量评价图的目的。具体流程如图 3 所示。

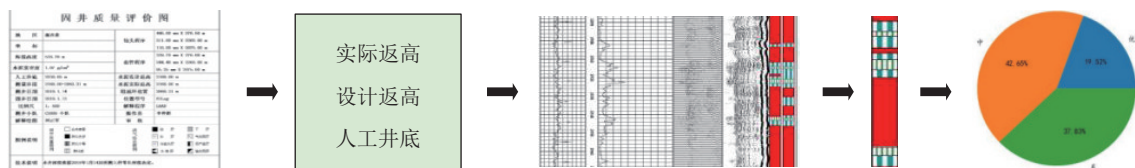


图3 固井评价图分析流程

#### 3.1 表头文本信息解析

通过 PaddleOCR 模型算法,可以对于评价图中各文本块的位置坐标及预测概率进行量化,结果以二元组形式输出。输出形式如图 4,其中  $X_1 \sim X_4$  均为包含横、纵轴的坐标点,0.95 为文本的预测概率。

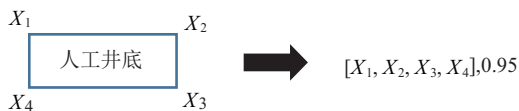


图4 PaddleOCR 输出形式

以图 1 为例, PaddleOCR 识别到的部分实际数据信息如表 1 所示,以识别人工井底深度‘2838.65m’为例,首先定位到相同行的‘人工井底’和‘水泥设计返高’两项的坐标位置,具体数值信息的水平位置则在两项文本信息之间,垂直位置和二者近似,因此根据坐标限制,可以查找到所关注的的数据。

基于文本特点,在解析文本内容时使用两项解析规则。对于非数字型文本,参考坐标位置,将非常邻近的字符合并为词组,如表 1 中的‘地’和‘区’



表 1 图 1 的 PaddleOCR 识别结果

文本内容	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	预测概率
固井质量评价图	[202.0, 14.0]	[544.0, 4.0]	[544.0, 4.0]	[202.0, 4.0]	0.91
地	[56.0, 285.0]	[75.0, 85.0]	[75.0, 06.0]	[56.0, 06.0]	0.99
区	[110.0, 84.0]	[128.0, 284.0]	[128.0, 306.0]	[110.0, 306.0]	0.99
温吉桑	[148.0, 284.0]	[200.0, 284.0]	[200.0, 306.0]	[148.0, 306.0]	0.96
人工井底	[56.0, 448.0]	[127.0, 448.0]	[127.0, 466.0]	[56.0, 466.0]	0.95
2838.65 m	[149.0, 448.0]	[225.0, 448.0]	[225.0, 465.0]	[149.0, 465.0]	0.97
水泥设计返高	[383.0, 448.0]	[489.0, 447.0]	[489.0, 465.0]	[383.0, 466.0]	0.94
2168.00 m	[504.0, 448.0]	[580.0, 448.0]	[580.0, 465.0]	[504.0, 465.0]	0.93

为邻近非数字文本;对于数字型文本,采用预置限定字符集合查询方法,对于单个字符为 0 至 9 或小数点,则记作单个有效字符,其它为无效字符。采用上述方法测试多张评价图,文本条目均能准确预测,预测准确率在 90% 以上。

### 3.2 图像聚焦

精准聚集于相应的固井质量面板是准确分析固井质量的前提。图像聚焦主要是为找到特定井深范围的固井质量面板,主要分为两步:首先根据需求划分出特定井深的分析范围,其次确定固井质量分析面板在评价图中的范围。

在对于特定井深范围划定过程中,可以使用坐标位置参照方法。图 5 为图 2 井深标示放大、旋转后的局部区域示例,通过 PaddleOCR 可以获取图 5 所示中‘深度’在固井评价图中水平、垂直的坐标位置,深度‘2170’、‘2180’等井深标示全部在‘深度’正下方位置,固井评价图中井深和‘深度’等标示在距离评价图左边界 10%~40% 位置。通过此种方式可以逐步缩小解析深度信息在评价图中的范围,进而达到提高解析准确率的目的。

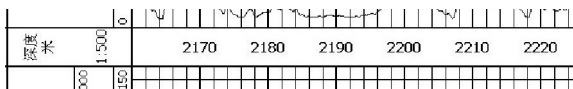


图 5 井深标识

通过深度标示数据的坐标可以精确定位井深在评价图中的位置。以图 5 为例,其中‘2170’坐标信息为 [383.0, 60.0], [448.0, 60.0], [448.0, 92.0], [383.0, 92.0], 即‘2170’字符最左端、最右端至该子图的左边界距离分别为 383、448 像素,字符本身占据 65(448~383)个像素,因此字符‘2170’的位置可精确定位在 383 和 448 中间的 416 像素位置。同理

可获得井深字符‘2180’至最左端的像素距离,按坐标间距离与井深距离等比例计算可以获得‘2171’、‘2172’等未标示的深度像素位置。相比于深度下方的‘1:500’,此种方式获取井深与标示的比例关系更为准确。

评价图中包含声幅、全波列、变密度等多项反映固井质量的面板,研究开展过程以评价图中不同图例区分固井质量等级的第一、二界面的面板为例进行分析。搜索范围限定在水平坐标相对比例位置 70%~100% 的区域。评价图纸张为白色且第一、二界面的左、右两边均有黑色边界线,灰度图像像素值范围是 0~255,0 为黑色,255 为白色,基于此特点,固定评价图第一、二界面范围,从右至左搜索整列像素值均小于 10 的列作为黑色线条,按黑色线条数量和白色边界可以确定第一、二界面的双侧范围。如图 6 中灰色方框线所示即为第一、二界面附近区域的黑色边界线像素值。

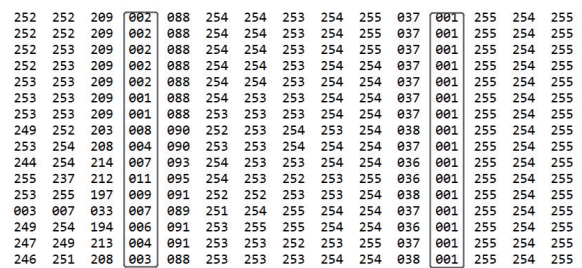


图 6 第一、二界面的黑色边界线像素值

如上所述,使用坐标位置参照及按图形比例限定搜索范围的方法可准确定位井深标示数据;根据井深的坐标距离分析可精细到每米井深的具体位置;参考第一、二界面像素值接近 0 的黑色边界线可以聚焦到第一、二界面。

### 3.3 图像解析

对比国内多家油田公司固井质量评价图可知,评价图的第一、二界面图例多用简单的线条或颜色区分不同固井质量。这种类型图像特征的识别与常规的整张图片分类或目标检测算法有所不同,因为它们更注重细节和局部特征的识别。针对此,提出了采用梯度距离和直方图相似度图像解析算法来完成对第一、二界面的图例类别划分的解决方案。其中:梯度算法首先通过计算图像中不同区域或特征之间的梯度距离,来识别和区分不同的图例,梯度距离能够有效地衡量图像中相邻像素间的差异,帮助算法识别出线条或颜色的变化;直方图相似度算法被用来分析这些特征的分布和模式,通过比较不同图像之间的直方图,可以量化它们的相似度,从而确定它们所属的类别。

#### 3.3.1 梯度距离

分析线条差异明显的第一、二界面可知不同等级的图像水平梯度变化非常大。图像梯度是指图像某像素在水平和垂直两个方向上的变化率,它反映了图像相邻像素颜色过度是否平滑,过度越平滑则图像梯度越小,反之梯度越大。图像梯度是包含水平  $X$  轴和垂直  $Y$  轴的二维向量,计算方式如式(1),其中  $g_x$  是一个像素在点水平方向上与其相邻像素点的差值,  $g_y$  是垂直方向上相邻点的像素值差值,  $x, y$  为像素点在水平、垂直方向上的位置索引值,以该差值向量  $\nabla f(x, y)$  为对应像素点的梯度,以此反映该点像素与周围像素颜色的变化程度。

$$\nabla f(x, y) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} f(x+1, y) - f(x-1, y) \\ f(x, y+1) - f(x, y-1) \end{bmatrix} \quad (1)$$

评价固井质量的第一、二界面在整口井的深度范围上呈现为竖条形,因此区分不同固井质量等级仅需关注每行像素在水平方向上的梯度  $g_x$ , 即一个像素点左、右像素值差值,再进一步分析第一界面彩色图像灰度化后图像。可发现:非0梯度值之间距离具有明显区分性特征,为增强距离特征,将图形阈值化处理,图像阈值化处理方式,计算方式如式(2):

$$gray(x, y) = \begin{cases} 0 & gray(x, y) \leq T \\ 255 & gray(x, y) > T \end{cases} \quad (2)$$

式中:  $gray(x, y)$ ——在行、列索引值为  $x, y$  处像素阈值化后的灰度值;  $T$ ——设定的阈值。

在计算机图像中每个像素的范围为  $0 \sim 255$ , 数值越小亮度越暗,越接近黑色;数值越大颜色越亮,越接近白色。阈值化目的是为将不同灰度范围的像素归类纯黑或纯白,通过这种方式区分不同类型图例梯度距离更为准确。第一界面部分区域灰度化、阈值化后的展示效果示例如图7所示。



图7 第一界面灰度化、阈值化示例

对图7中阈值化后的图像,应用式(1)可以得到第1、2界面每行像素在水平方向上的梯度向量  $g_x$ , 在水平方向上,由像素值为0的纯黑色过渡到像素值为255的纯白色梯度值为255,反之由纯白色过渡到纯黑色,梯度值则为-255,而其它纯黑、纯白域梯度值为0,由于梯度为255或-255均表示黑白颜色的交替变化,因此对出现负值的梯度作绝对值处理。最后根据式(3)计算第1、2界面每行像素中相邻梯度值不为0的像素点之间的平均距离。

$$distance = \frac{1}{n} \sum_{i=1}^n |i - j| s.t. \begin{cases} grad_i \neq 0 \\ grad_j \neq 0 \end{cases} \quad i, j = 1 \dots n \quad (3)$$

式(3)中  $grad_i$  和  $grad_j$  为第1、2界面某行像素列位置索引在  $i$  和  $j$  水平方向上的梯度值,其中  $n$  为梯度交替出现的次数。以图7为例,纯红色类型梯度距离为0,而绿色较细线条类型梯度距离相较于青色粗线条梯度距离较大,再通过设定经验阈值比较,可完成不同等级固井质量类型划分。

#### 3.3.2 直方图相似度

在计算机中,图像可以看作是红、绿、蓝3种不同颜色通道的二维数组互相叠加在一起形成的,而每种颜色通道中二维数组数值范围是固定的从  $0 \sim 255$  离散的整数值,代表了该颜色在像素点上的强度。图像直方图是反映一个图像像素分布的统计表,其中横坐标表示图像各个像素值,可以是图像整体灰度化后的像素值,也可以是红、绿、蓝某个单

通道的像素值。纵坐标代表了每一种颜色值在图像中的像素总数或者占有所有像素个数的百分比。

图8形象地描述了图像直方图计算过程,4×4个16个不同的灰度级别小方块表示16个不同的像素值,右侧数轴中的柱形图表示对应像素值的数量,该柱形图形成的像素分布通常被称之为图像直方图。图像直方图数据存储形式为1维向量,向量的索引下标为对应的像素值,而向量值为像素的统计个数或百分比。图像的直方图向量表示易于完成多个图像的相似度比较,欧基里德距离和余弦相似度是度量一维向量相似度的有效方法,根据固井质量评价图特点,本文采用了余弦相似度方法,第一、二界面的图像直方图对比分类过程如下:

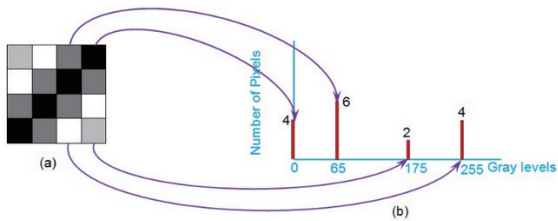


图8 图像直方计算示例

(1)将像素值0~255的整体范围等分成 $n$ 个区间( $n \in [10 \sim 20]$ ),此目的为缩小直方向图向量维度。

(2)根据表头中不同固井质量等级的图例,按步骤(1)计算不同固井质量等级图例相邻 $k$ 行像素的直方图向量 $V_a$ ,以该组向量作为评估第1、2界面的对比基向量

(3)分别遍历第一、二界面每 $k$ 行像素按步骤

(1)计算直方图向量 $V_b$ 。

(4)按公式(4)将直方图向量 $V_a$ 、 $V_b$ 中各个元素归一化为0至1的浮点型向量 $a$ 、 $b$ ,以易于区分直方向图向量内部各个区间的差异,公式(4)中 $x_i$ 为直方向量内像素数量。

$$vector = \frac{x_i}{\sum x_i}, i = 1 \dots n \quad (4)$$

(5)按余弦相似度公式(5),计算第一、二界的直方向向量 $b$ 与不同等级基向量 $a$ 之间余弦距离,并将其归类于与之距离最小的基向量等级,公式(5)中 $\|a\|$ 和 $\|b\|$ 分别表示向量 $a$ 、 $b$ 的模。

$$\cos \theta = \frac{a \cdot b}{\|a\| \cdot \|b\|} \quad (5)$$

### 3.4 质量评估

在固井评价图上的多步聚焦能准确定位到第1、2界面,如上所述图像解析算法可完成界面的单行像素质量等级划分。图9为聚焦和等级类型划分的整体示例,其中图9(a)为使用文本坐标参照方法得到的特定的深度范围,图9(b)为根据第一、二界面相对位置按评价图宽度比例宽泛划定的搜索范围,图9(c)为使用黑色边界线找到的第一界面固井质量分析面板,图9(d)为使用梯度距离分类方法所完成的质量等级划分,包含蓝色区域的列为第一界面中绿色线条的类别划分,其中蓝色区域表示其被划分为绿色线条类型等级,而灰色区域表示它未被划分到此类等级。同理,含绿色区域的列为第一界面中青色线条的类别划分,最右侧含红色区域的列为第一界面中红色线条的类别划分。从图9(d)中可以看出梯度距离的度量方式可以有效判断第一界面中的图例类别。

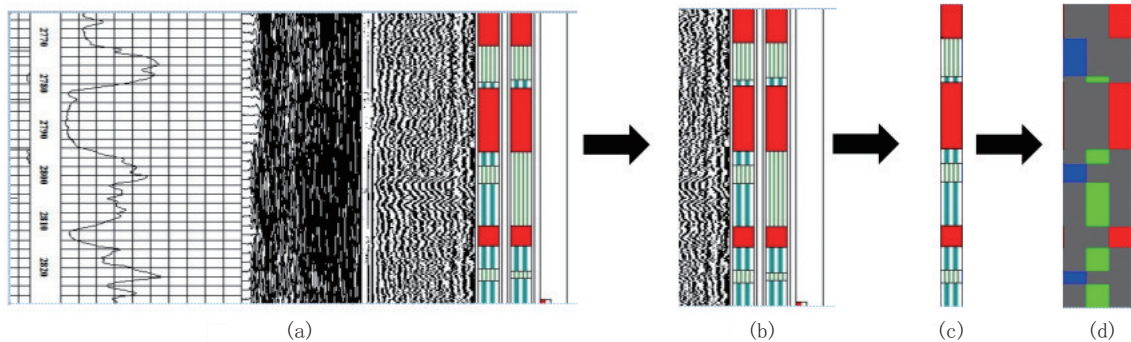


图9 聚焦和类型划分整体示例



对图9(d)所示的第一、二界面统计占比,可获得相应井深范围的固井质量优、中、差各部分的比例。以评价图表头中的水泥预计返高和人工井底范围作为整口井的分析范围,以图中的第一界面分析为例,评价图描述性分析结果如下:温侧X井,人工井底深度2838.65 m,水泥设计返高2168 m,实际返高2168 m,实际井深2875 m,实际识别固井质量长度670.65 m,水泥胶结优质长度580.11 m,占比85.6%,水泥胶结中等长度60.35 m,占比9%,水泥胶结差长度36.21 m,占比5.4%。

#### 4 效果评估

为验证和调优模型参数,收集了A油田50口钻井的固井质量评价图,以A油田公司所提供的评价图原始数据实际合格率为标准,其中合格率是指固

井质量为优等和中等两个等级的占比,计算方式如公式(6),其中P是指百分比。为更准确评估模型的泛化能力,本文随机挑选20张固井质量评价图为训练样本,30张图为测试样本以调优和评估模型效果。

$$pass\ rate = \frac{P_{优} + P_{中}}{P_{优} + P_{中} + P_{差}} \quad (6)$$

在最优参数的基础上,进一步评估模型对合格率的识别情况。有研究开展过程中主要对梯度距离和图像直方图相似度两类图像解析算法进行了调优测试,图像直方图算法测试了第一、二界面不同颜色通道对合格率的影响,表2为10口井第一界面红色通道、绿色通道、蓝色通道、灰度图、红绿蓝通和灰度图相拼接4种不同类型直方图向量对合格率的分析结果。

表2 不同类型直方图算法合格率

井名	第一界面合格率/%					
	实际值	红通道	绿通道	蓝通道	灰度图	红绿蓝通道+灰度图
葡北X井	92.34	74.36	20.54	30.47	85.21	87.38
红台X井	89.47	70.31	37.26	45.95	93.7	87.64
玉北X井	67.45	55.55	45.25	52.9	62.3	72.31
玉东X井	78.56	62.64	45.39	50.22	82.53	81.21
丘东X井	88.78	71.4	42.67	54.2	74.72	80.9
雁X井	95.67	82.33	26.83	40.51	85.46	89.42
神X井	99.45	99.5	10.44	15.4	81.78	92.41
葡X井	94.35	80.5	21.43	36.83	78.5	87.65
牛东X井	89.85	65.87	32.64	40.56	95.85	93.2
马56X井	91.69	71.43	29.73	35.3	82.0	84.87

对比不同方式计算直方图向量对合格率的影响,可以看出红、绿、蓝3种颜色通道中红色通道方式效果最好,而绿色和蓝色通道效果最差,导致这种情况的原因在于:第1、2界面中优等质量等级使用红色,并且固井优等占通常较高。比较灰度图、红绿蓝+灰度图两种方式,综合更多特征效果为更好。

梯度距离算法准确性主要在于梯度距离阈值的设定,使用网格搜索的方式可找到在测试集上最优的阈值。表3为最优梯度距离和直方图相似度算法的效果对比,从表3可以看出梯度距离算法更优。直方图相似度效果相对较差的原因主要在于直方图向量容易受噪点影响而造成误判。此外,分析结

果表明固井质量合格率越高,识别结果误差越小。

整个井筒的固井实际合格率与预测值对比可以反映所提出算法的整体效果,但是算法在局部井深段的效果难以呈现。表4给出了葡北X井以10为间隔的分段部分深度的对比情况。从表4中可以看出序号为1、2、6、8行整段均为单一类别可以准确识别,而对于序号为4、5两种或3种类似交替出现的片断会出现一定的误差,但每类识别的各类比例接近于真实情况。

30张测试样本中,其中有27张图片所识别的固井合格率与原数据合格率的误差在5%以内,3张图片误差范围在5%~10%之间,实现对于固井质量评价图自动分析的同时,误差范围满足了工程现场的需要。

表 3 梯度距离和直方图相似度算法效果对比

井名	第一界面合格率/%			第二界面合格率/%		
	实际值	梯度距离	直方图相似度	实际值	梯度分距离	直方图相似度
葡北 X 井	92.34	90.62	87.38	92.23	91.45	94.37
红台 X 井	89.47	90.53	87.64	88.43	87.52	85.83
玉北 X 井	67.45	65.84	72.31	60.45	55.31	53.22
玉东 X 井	78.56	76.64	81.21	75.36	71.21	69.87
丘东 X 井	88.78	84	80.9	84.38	80.25	80.1
雁 X 井	95.67	95.5	89.42	94.45	92.15	89.57
神 X 井	99.45	99.2	92.41	99.32	97.64	98.0
葡 X 井	94.35	93.46	87.65	92.39	91.67	94.1
牛东 X 井	89.85	88.95	93.2	85.96	82.25	81.34
马 56X 井	91.69	88.9	84.87	89.77	86.63	83.59

表 4 葡北 X 井深分段细分对比

序号	井深/m		实际值占比/%			梯度距离预测值占比/%		
	开始	结束	优	中	差	优	中	差
1	2000	2010	100	0	0	100	0	0
2	2110	2120	100	0	0	100	0	0
3	2350	2360	58.2	41.8	0	53.7	46.3	0
4	2570	2580	63.1	28.4	8.5	61.2	33.4	5.4
5	2770	2780	0	100	0	0	100	0
6	2780	2790	0	74.3	35.7	0	72.1	37.9
7	2840	2850	0	0	100	0	0	100

5 结论

(1)建立了基于图像识别技术开展固井质量评价的具体方法,提出了采用梯度距离和直方图相似度图像解析算法来完成对第一、二界面的图例类别划分的创新解决方案。

(2)建立了坐标位置参照算法和两种评价图局部特征域解析算法,使用评价图文本坐标位置参照算法,可以准确限定待分析信息在评价图中的位置范围。

(3)基于所建立的基于图像识别技术开展固井质量评价方法,开展了 50 口井固井质量评价,其中 30 张测试样本中,有 27 张图片所识别的固井合格率与原数据合格率的误差在 5% 以内,3 张图片误差范围在 5%~10% 之间,满足了工程现场的技术需要。

参考文献:

[1] 吴天乾,宋文宇,谭凌方,等.超低密度水泥固井质量评价方法[J].石油钻探技术,2022,50(1):65-70.

[2] 郭洪志,雷达.双管柱封隔固井质量评价技术方法及应用[J].世界石油工业,2022,29(3):54-60.

[3] 潜晓晨.固井质量评价测井影响因素分析[J].西部探矿工程,2023,35(11):92-94.

[4] 吴华,刘连恺,王磊,等.GY5-1-4H 页岩油水平井提高固井质量关键技术研究与实践[J].钻探工程,2023,50(4):135-141.

[5] 邵一臻,赵福金,荆京,等.山西干热岩 GR1 井高温固井技术研究与实践[J].钻探工程,2022,49(6):42-47.

[6] 王明辉,李进,韩耀图,等.基于超声 Lamb 波的固井质量评价新技术与应用[J].钻井液与完井液,2023,40(4):519-526.

[7] 陈向军,李维彦,杨顺辉.固井质量评价面临的新问题及思路.测井技术,2015,39(1):1-5.

[8] 肖红,钱祎鸣.基于改进 DenseNet 的固井质量评价新方法[J].计算机技术与发展,2024,34(1):193-199.

[9] Jing Cheng L , Hong Tu W , Shun Peng Z , et al. Intelligent Evaluation Model for Cementing Quality Based on PSO-SVM and Application [C]//International Conference on Green Building, Materials and Civil Engineering, 2011.

[10] 郑双进,程霖,龙震宇,等.基于 GA-SVR 算法的顺北区块固井质量预测[J].石油钻采工艺,2021,43(4):467-473.

[11] 李忠庆,李传武,朱达江.利用环空压力测试数据评价高温高压气井固井水泥环的密封完整性[J].钻探工程,2021,48(6):15-21.

[12] 芦琦,刘洋,秦辉等.基于 OCR 的标书文件信息获取技术应用研究[J].信息与电脑(理论版),2023,35(9):166-169.

[13] 曹菁,陈康,齐宁等.基于 OCR 和图像检测的盖章文书图像自动审核方法[J].应用科学学报,2023,41(6):1058-1067.

[14] 钟琳.基于 RPA+OCR 技术的海上油田生产过程监测研究[J].中国高新科技,2023(18):33-35.

[15] 杨玉卿,崔维平,张翔.基于电成像测井图像识别求取地层产状的新方法[J].石油物探,2017,56(2):302-308

[16] 彭越,张满郎,李明秋,等.基于图像识别技术的裂缝发育程度定量评价新方法——以安岳气田须二气藏为例[J].非常规油气,2024,11(1):12-21.

(编辑 荐华)