

青藏高原黑海浅层地震剖面野外实测与 能量比法解译数据集

孙运坤^{1*}, 冯辛韦², 王振红¹, 邓 华¹

1. 浙江华东建设工程有限公司, 杭州 310000; 2. 浙江大学求是高等研究院, 杭州 310000

摘 要: 浅地层地震剖面自动解译法以其能快速高效的解译地震剖面且结果更为客观的优势在地震资料处理方面有着非常好的应用前景。本文利用 GeoPulse3.5 kHz 浅地层剖面系统分析青藏高原北部的黑海采集的浅地层地震剖面数据, 首先以优化变长滑动时窗能量比法计算出剖面数据的能量比值曲线, 其次采用均值曲线法提取地层边界点, 最后进行边界点的连续性检验, 通过 Matlab 编程实现浅地层地震剖面的自动解译, 并获得野外实测与能量比法解译数据集。经过一系列实验数据与实测数据对比分析得到, 当变长滑动时窗的取值范围在 2–20 之间且均值曲线峰值提取窗口取值在 70–80 之间时, 黑海剖面数据的自动解译与人工解译得到的地层界线吻合度大于 0.65, 提取的边界点在界线附近的密集度大于 0.8。青藏高原黑海浅层地震剖面野外实测与能量比法解译数据集包括青藏高原黑海地理位置地理信息系统数据、该点位浅地层地震剖面的 2011 年实测数据和能量比法解译数据。其中, 野外实测包括 3 sn、4 ns 和 4 we 三个剖面的数据。这三个剖面的数据分别包含 3001 道、1918 道和 8001 道数据, 每道数据由 1,664 个样点的数据组成 (*.txt)。能量比法解译数据包括: 3 sn 剖面地层界线数据; 单道剖面数据和 19 列能量比值数据; 参数化方案 ($(L_{\min}-L_{\max})/L_{\text{win}}$) 及提取的边界点数量; ($L_{\min}-L_{\max}$) 取值方案下不同 L_{win} 的提取效果; 黑海单道地震剖面数据与西洋单道地震剖面数据。数据集存储为 .kmz、.shp、.txt 和 .xlsx 格式, 包含 5 个数据文件, 数据量为 73.9 MB, (压缩为 1 个文件, 27.7 MB)。

关键词: 青藏高原; 黑海; 浅层地震; 实测数据; 能量比法

DOI: 10.3974/geodp.2018.02.11

1 前言

浅地层剖面探测技术已广泛应用于海洋、湖泊等区域的地质调查中, 剖面数据中包含的地层信息能为工程建设、科学研究等提供基础数据^[1–2]。在进行剖面数据解译时, 传统的数据解译工作主要以耗时耗力的人工目视解译大量的繁杂的剖面数据为主, 且解译成果受人为主观因素影响较大。因此通过数据转换提取地震波属性中包含的地层信息, 利用计算机批量处理大量剖面数据而实现数据解译工作的自动化, 不仅能提高解译效率, 而且获得的成果更为客观和准确。

收稿日期: 2018-06-08; 修订日期: 2018-06-20; 出版日期: 2018-06-25

基金项目: 国家科学技术部 (2013CB956500); 江苏省自然科学基金重点项目 (BK20130056)

*通讯作者: 孙运坤 A-4112-2017, 浙江华东建设工程有限公司, sun_yk@ecidi.com

数据引用方式: [1] 孙运坤, 冯辛韦, 王振红等. 青藏高原黑海浅层地震剖面野外实测与能量比法解译数据集[J]. 全球变化数据学报, 2018, 2(2): 190–197. DOI: 10.3974/geodp.2018.02.11.

[2] 孙运坤, 张瑞琦, 周连福等. 青藏高原黑海浅层地震剖面野外实测与能量比法解译数据集 [DB/OL]. 全球变化科学研究数据出版系统, 2016. DOI: 10.3974/geodb.2016.09.10.V1.

地震数据中的振幅属性中包含地层信息, 为了利用计算机自动提取剖面数据中振幅包含的地层信息, 首先需要将特殊格式的剖面数据转换为可直接操作格式的数据^[3], 再根据振幅特征与地层层序之间的对应关系提取地层信息, 通过计算机批量处理剖面数据, 可快速解译剖面数据。为了实现该目的, 刘秀娟等根据地震波振幅在海底表面的反射特征, 采用最大振幅法成功提取了浅地层剖面数据中的海底地形数据^[4]。徐钰等通过改进传统的能量比法, 实现了浅层地震数据中初至波的自动拾取^[5]。振幅分析法的成功运用为实现计算机批量自动解译浅地层剖面数据中的地层信息奠定了基础。

本工作以青藏高原黑海的浅地层地震剖面数据为研究对象。在地质构造上, 黑海湖北部为两组东西走向的大断裂, 即东昆仑中断裂和南沟下断裂; 东部为昆仑河-野牛沟大断裂, 呈东西走向; 南部为东西走向的西大滩断裂和东昆仑南断裂, 构造活动丰富^[6]。在湖中进行数据采集时, 风浪较弱, 所使用的船体较小, 因此周围环境噪声干扰较小, 剖面数据质量较好。进行剖面数据自动解译时, 首先将 SEG-Y 格式的剖面数据转换为 txt 格式的可直接操作的数据, 通过参数化方案选取合适的自动提取参数集, 利用选取的参数集经过优化的能量比法计算每道地震记录的能量比值, 采用均值曲线法提取边界点, 再通过边界点的连续性检验, 剔除孤立点, 最后获得自动解译地层图。

2 数据集元数据简介

青藏高原黑海浅层地震剖面野外实测与能量比法解译数据集^[7]的名称、作者、地理区域、数据年代、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

3 数据研发方法

本文选用的数据是 2011 年在青藏高原东北部的湖泊——黑海所采集的浅地层地震剖面数据。黑海位于青藏高原北部东昆仑山的北麓昆仑河-野牛沟断裂谷地内(图 1)。平均海拔 4 477 m, 沿谷地 NE-SW 向延伸, 面积 38.7 km², 湖中心最大水深 21.9 m。黑海处在高原地区, 人类活动少, 总体悬浮物浓度低, 且湖中周围环境中的噪声影响较弱, 利于数据采集。

野外数据采集使用英国 GeoAcoustic 公司制造的 GeoPulse 3.5 kHz 浅地层剖面系统, 理想状态下, 有效探测深度可达 70 m。系统由 4 个 3.5 kHz 的换能器与拖体、发射机、接收机和 GeoPro II 工作站组成, 数据采集时设定的主要参数值为: 脉冲发射周期与记录长度为 100 ms, 信号采集率为 60 μ s, 高低通滤波 300 Hz-7 kHz, 增益为 6 dB 或 9 dB, 数据存储格式为 SEG-Y 格式。

3.1 算法原理

(1) SEG-Y 数据格式转换

研究工作中, 对 SEG-Y 文件的处理方式是: (1) 在 Visual C++ 平台下先将文件中的二进制数据转换为十六进制的数据, 找出 SEG-Y 文件的数据存储方式。(2) 将二进制数据转换为十进制的 txt 文本文件。最终得到的 txt 文本文件格式为: 每一道数据前存入道编号、样点数两个固定的信息, 之后为地震剖面数据。

表 1 青藏高原黑海浅层地震剖面野外实测与能量比法解译数据集元数据简表

条 目	描 述
数据集名称	青藏高原黑海浅层地震剖面野外实测与能量比法解译数据集
数据集短名	SubSeismicProfileHeihaiTibet
作者信息	孙运坤 A-4112-2017, 浙江华东建设工程有限公司, sun_yk@ecidi.com 张瑞琦 A-5164-2017, 复旦大学计算机科学与技术学院, 734722018@qq.com 周连福 南京大学地理与海洋科学学院, 812975031@qq.com 戴晨 A-4153-2017, 南京大学地理与海洋科学学院, 812975031@qq.com 张永战 A-3143-2017, 南京大学地理与海洋科学学院, zhangyzh@nju.edu.cn
地理区域	青藏高原东北部的湖泊——黑海 (35°59'13.84"N, 93°15'51.87"E)
数据年代	2011 年
研究区域	8.98 km×4.41 km
数据格式	.xlsx、.txt、.shp、.kmz
数据量	27.7 MB (压缩后)
数据集组成	青藏高原黑海的地理位置: HeihaiLake.kmz; 现场采集得到的 3 条浅剖数据: SeismicProfileData_4ns.txt、SeismicProfileData_3sns.txt、SeismicProfileData_4we.txt; 经过解译处理得到的解译数据: ProssedSeismicProfileData.xlsx
基金项目	国家科学技术部 (2013CB956500); 江苏省自然科学基金重点项目 (BK20130056)
出版与共享服务平台	全球变化科学研究数据出版系统 http://www.geodoi.ac.cn
地址	北京市朝阳区大屯路甲 11 号 100101, 中国科学院地理科学与资源研究所
数据共享政策	全球变化科学研究数据出版系统的“数据”包括元数据 (中英文)、实体数据 (中英文) 和通过《全球变化数据学报》(中英文) 发表的数据论文。其共享政策如下: (1) “数据”以最便利的方式通过互联网系统免费向全社会开放, 用户免费浏览、免费下载; (2) 最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源; (3) 增值服务用户或以任何形式散发和传播 (包括通过计算机服务器) “数据”的用户需要与《全球变化数据学报》(中英文) 编辑部签署书面协议, 获得许可; (4) 摘取“数据”中的部分记录创作新数据的作者需要遵循 10% 引用原则, 即从本数据集中摘取的数据记录少于新数据集总记录量的 10%, 同时需要对摘取的数据记录标注数据来源 ^[8]

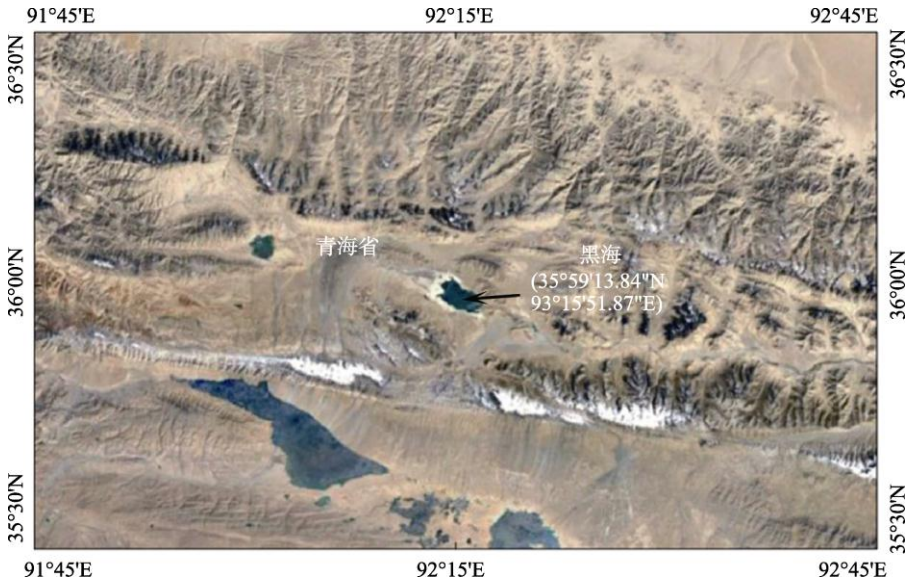


图 1 黑海地理位置图 (引用于 Google Earth)

(2) 能量比法

完整的一条地震剖面是由若干道地震记录组成, 单道数据存储为离散的采样点的地震波振幅值。时窗定义在单道记录的采样点上, 其大小以采样点的数量为单位。能量比值的计算公式为^[9]:

$$R = \frac{\sqrt{\sum_{t=T_0}^{T_2} x^2(t) + \alpha E}}{\sqrt{\sum_{t=T_1}^{T_0} x^2(t) + \alpha E}} \quad (1)$$

$$E = \sqrt{\sum_{t=0}^N x^2(t)} \quad (2)$$

式中, R 为时窗内后半段 ($T_0 - T_2$) 与前半段 ($T_1 - T_0$) 的能量比, $x(t)$ 为地震记录振幅值, T_1 为时窗起点, T_0 为时窗中点, T_2 为时窗终点, E 为单道地震道的相对能量, N 为单道地震道的样点总数, α 为稳定系数。

(3) 能量比值均值曲线的计算

根据能量比值曲线提取地层边界点之前, 首先需要将计算得到的能量比值曲线进行平均化处理, 得到能量比值均值曲线。经过平均化得到的能量比值均值曲线比能量比值曲线能更加清晰的突出峰值所在的位置。

(4) 边界点提取和连续性检验

1) 边界点的提取

根据地层对地震波的反射特征可知, 边界点在能量比值曲线中表现为峰值, 能量比值曲线中的峰值即为边界点。本文采用均值曲线法提取能量比值曲线的峰值, 为了突显每个点的能量比值特征, 先计算方差, 剔除与均值相差较大的若干个值, 直到剩余值的方差在设定范围以内或者剔除值的个数达到设定的最大个数为止, 最后再计算剩余数据的平均值, 得到均值曲线。在均值曲线法中, 同样采用变长滑动时窗法来判别峰值点。

2) 边界点连续性检验

在实际地层中, 地层在一定范围内的分布是连续的, 因此当某一道记录中有地层边界点出现时, 与其相邻的若干道记录中应该有与之连续的边界点。因此, 针对识别出的地层边界点, 需要根据地层连续性原则, 将相对离散的孤立点剔除, 即可获得连续的地层界线。边界点连续性检验的基本原理为: 遍历剖面中所有的边界点, 判断各个点所在指定深度范围内, 指定若干道相邻记录范围内是否存在一定数量的其它点, 若满足条件则认为此点属于边界点, 将其保留; 若不满足条件则认为此点是孤立点, 将其剔除^[10]。

3.2 技术路线

在进行剖面数据自动解译时, 首先将 SEG-Y 格式的剖面数据转化为常用的.txt 格式的数据文件, 其次计算每道剖面数据的能量比值曲线和能量比均值曲线, 最后进行边界点提取和连续性检验, 进而获得目标界线。具体技术路线如图 2 所示。

4 数据结果与验证

4.1 数据结果组成

本数据集包括:

(1).kmz 格式的青藏高原黑海的地理位置数据: HeihaiLake.kmz。

(2).txt 格式的 2011 年于黑海现场采集的原始数据经过转换后得到的 3 条 .txt 格式的浅剖数据: SeismicProfileData_4ns.txt, SeismicProfileData_3sn.txt, SeismicProfileData_4 we.txt。

(3).xlsx 格式的经过解译处理得到的解译数据: ProssedSeismic Profile Data.xlsx。

1) 黑海 3 sn 剖面地层线数据

人工解译边界点的数目是根据每道地震记录中穿过的地层界线的条数确定的, 统计时记录边界点的垂直样点序号, 最后得到一条剖面中的所有边界点及相应样点序号 (图 3)。

2) 单道地震剖面数据和 19 列能量比值数据

本文采用均值曲线法提取能量比值曲线的峰值, 为了更能突显每个点的能量比值特征, 在计算能量比值均值之前, 先计算方差, 剔除与均值相差较大的若干个值, 直到剩余值的方差在设定范围以内或者剔除值的个数达到设定的最大个数为止, 最后再计算剩余数据的平均值, 得到均值曲线。在均值曲线法中, 同样采用变长滑动时窗法来判别峰值点。

3) 参数化方案 ($(L_{min} - L_{max}) / L_{win}$) 及边界点提取效果

自动解译剖面时, 解译精度由多个参数共同决定, 不同参数设定得到的解译效果亦不相同。在进行剖面解译时, 滑动时窗的长度范围 ($L_{min} - L_{max}$)、均值曲线法峰值提取窗口长度 (L_{win}) 是影响解译精度的主要参数 (图 4), 参数化方案便以这两组参数为基础进行设计, 得到提取效果最好的参数化方案 (图 5)。提取效果以边界点密集度和边界线相似度来衡量。

4) 4 ns 剖面和 4 we 剖面解译数据

根据剖面解译的技术路线及解译参数 ($(L_{min} - L_{max})$ 取 2-20, L_{win} 取 76, 参数化方案列于表 2), 选取黑海两条地层分层较为明显的剖面 (4 ns、4 we) 进行整条剖面的解译, 得到整条

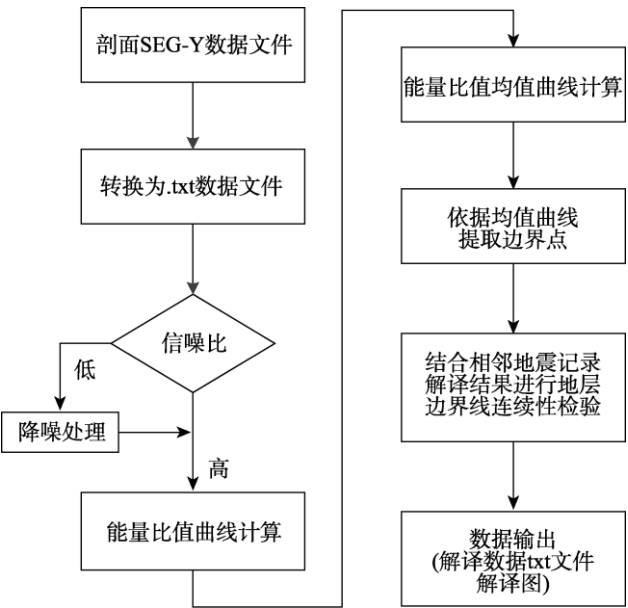


图2 数据集研发技术路线

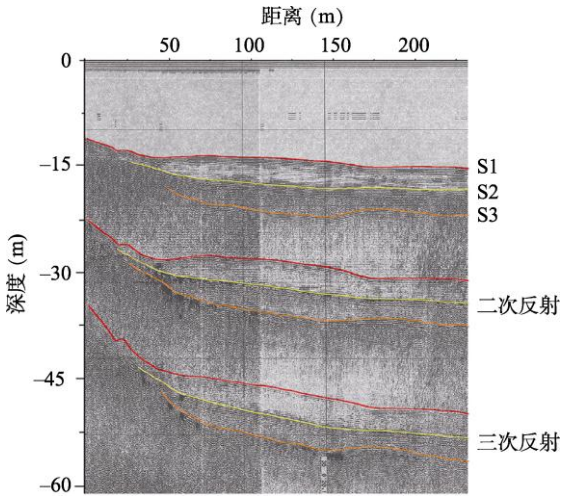


图3 黑海 3sn 剖面人工目视解译结果

剖面的解译数据（图 6）。

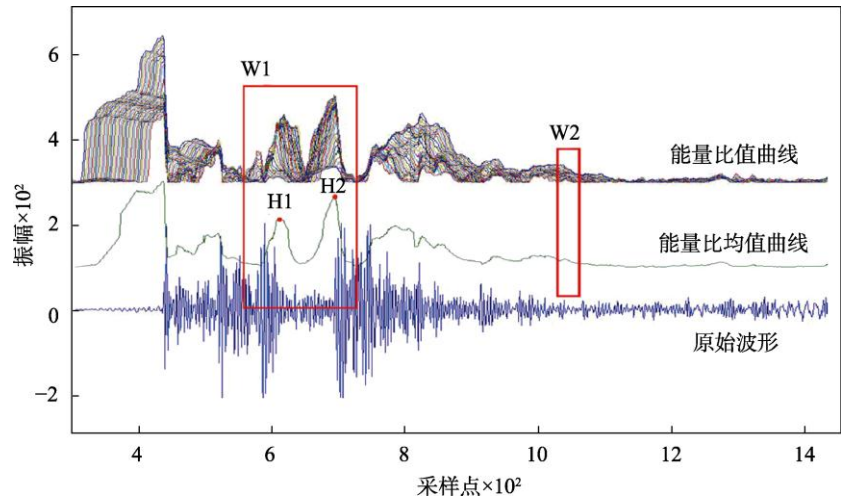


图 4 单道地震记录的能量比值曲线、均值曲线、原始波形

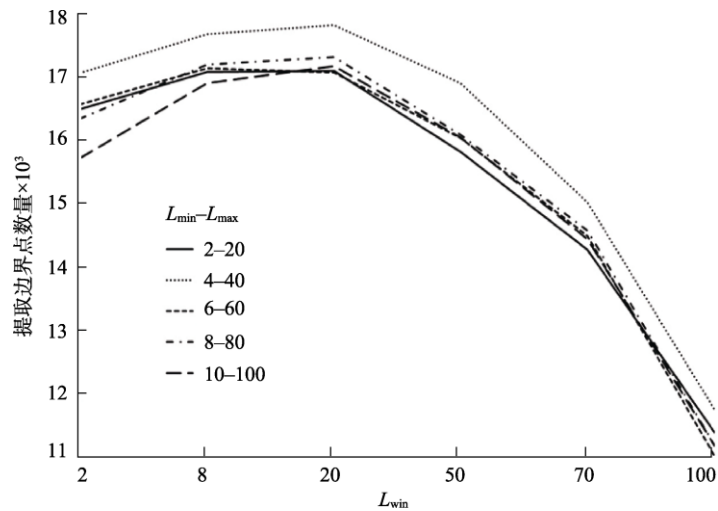


图 5 不同 L_{win} 和 $(L_{min}-L_{max})$ 取值的边界点提取效果^[10]

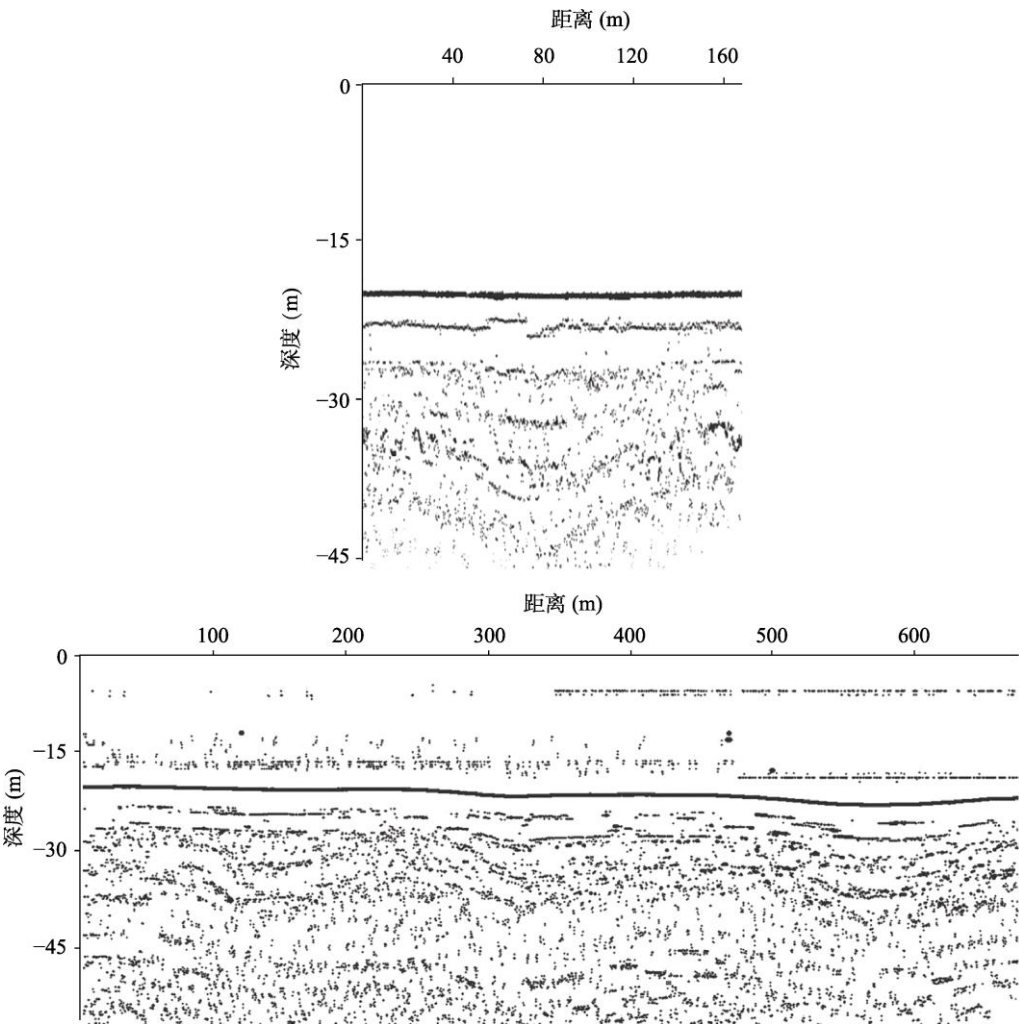
表 2 参数化方案 $((L_{min}-L_{max}) / L_{win})$ 及边界点提取效果^[9]

序号	数据名称	数据说明	数据量（个）
1	参数化方案 $((L_{min}-L_{max}) / L_{win})$ 及提取边界点数量	L_{win} 取值 2、8、20、50、70、100， $(L_{min}-L_{max})$ 取值 2-20、4-40、6-60、8-80、10-100 时，提取 的边界点数量	30
2	$(L_{min}-L_{max})$ 取 2-20 时 不同 L_{win} 的提取效果		30
3	$(L_{min}-L_{max})$ 取 6-60 时 不同 L_{win} 的提取效果	$(L_{min}-L_{max})$ 分别取值 2-20、6-60、8-80 时， L_{win} 取 2、8、20、50、70、100 时的解译情况， 即边界点密集度和边界线相似度	30
4	$(L_{min}-L_{max})$ 取 8-80 时 不同 L_{win} 的提取效果		30

4.2 数据结果验证

为验证参数集的可靠性，选取黑海两条剖面（4 ns、4 we）进行解译。

图 6 为 2 条剖面的自动提取的地层界线图。剖面 4 ns 中自动提取的地层界线与人工识别的地层界线^[10]吻合较好，25 m 以浅的地层中，地层界线 S2、S3 吻合度非常好；25–45 m 深度范围内，S4、S5 的走向一致；45 m 深度以下，地震反射信号揭示的沉积物分层已不明显，难以识别明显的地层界线。剖面 4 we 的一段连续起伏的埋藏谷地，自动解译能够提取出不整合面 S2，30–45 m 深度范围内 S3 的趋势也基本吻合。



（上图为剖面 4 ns；下图为剖面 4 we）

图 6 自动提取的地层界线图

5 讨论和总结

基于青藏高原黑海数据集，以变长滑动时窗优化的能量比法为基础，通过参数化方案选取合适的参数集，实现了对浅地层剖面数据的自动解译。本数据集能够批量解译同一地

区获得的浅地层剖面数据, 得到剖面地层图, 可以帮助建立研究区的基本地层结构框架, 为进一步的人工判读和详细分析提供重要基础。另外, 本数据集可根据不同的研究区域设计参数化方案, 确定适合于研究区域的浅地层剖面数据自动解译的关键参数, 提高自动解译精度, 获得最优的剖面数据解译成果。其不足之处是本数据集对次一级的地层单元中的地层界线提取精度还不够, 需要寻找其他对提取精度影响较大的参数进行参数化方案, 或采用合适的方法对剖面数据进行预处理, 以进一步提高自动解译精度。对数据集的初步分析表明, 本数据集能够成功提取出黑海湖的浅地层剖面数据中多数的地层界线, 且在精度上, 两条剖面中主要地层界线的相似度达到 60% 以上, 能够确保数据的使用精度。

参考文献

- [1] 滕吉文. 地震探测中三个科学难题的涉足与创新征程[J]. 地球物理学进展, 2014, 29(3): 0971-0992, DOI: 10.6038/pg20140301.
- [2] Lafferty, B., Quinn, R., Breen, C. A side-scan sonar and high-resolution chirp sub-bottom profile study of the natural and anthropogenic sedimentary record of Lower Lough Erne, northwestern Ireland [J]. *Journal of Archaeological Science*, 2006, 33(6): 756-766. DOI: 10.1016/j.jas.2005.10.007.
- [3] 张瑞琦. SEG-Y 格式地震波数据的提取与边界点识别研究[D]. 南京: 南京大学, 2012.
- [4] 刘秀娟, 高抒, 杨旸. 浅地层剖面仪 GeoChirp II 图像中风浪变形效应的校正方法初探[J]. 海洋科学进展, 2007, 25(2): 160-167. DOI: 10.3969/j.issn.1671-6647.2007.02.005.
- [5] 徐钰, 段卫星, 徐维秀等. 高精度初至自动拾取综合方法研究[J]. 物探与化探, 2010, 34(5): 595-599.
- [6] 周连福. 晚更新世以来藏北高原黑海湖泊沉积环境演化过程研究[D]. 南京: 南京大学, 2014.
- [7] 孙运坤, 张瑞琦, 周连福等. 青藏高原黑海浅层地震剖面野外实测与能量比法解译数据集[DB/OL]. 全球变化科学研究数据出版系统, 2016. DOI: 10.3974/geodb.2016.09.10.V1.
- [8] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策[OL]. DOI: 10.3974/dp.policy.2014.05 (2017 年更新).
- [9] 刘喜武, 张宁, 勾永峰等. 地震勘探信号时频分析方法对比与应用分析[J]. 地球物理学进展, 2008, 23(3): 743-753.
- [10] 孙运坤, 张瑞琦, 周连福等. 基于能量比法的浅地层地震剖面自动解译[J]. 地球物理学进展, 2015, 30(3): 1141-1150. DOI: 10.6038/pg20150318.